UNIVERSIDAD DE OVIEDO
Departamento de Estadística, I.O y D.M

Medidas de variación para elementos aleatorios imprecisos

Tesis Doctoral

M. Asunción Lubiano Gómez

A mi familia

Agradecimientos

En primer lugar, deseo expresar mi más sincero reconocimiento a los directores de esta memoria, María Ángeles Gil Álvarez, Miguel López Díaz y María Teresa López García, quienes durante la realización de la misma me han hecho sentir apoyada en todos los aspectos.

Quiero también dar las gracias a todos mis compañeros de área, que de una u otra forma me han ayudado a llegar hasta aquí y, en especial, a nuestro director de Departamento, Profesor Pedro Gil.

Han sido muchas las personas que me han asesorado en la elaboración de este trabajo. Entre ellas, es obligado destacar a los Profesores Carlo Bertoluzza de la Universidad de Pavía (Italia) y María del Carmen Alonso, Teófilo Brezmes, Hortensia López y Antonia Salas de la Universidad de Oviedo por sus contribuciones en algunos puntos de la memoria. Deseo manifestar así mismo mi agradecimiento a mis compañeros Luis José Rodríguez y Manuel Montenegro por su meticulosa revisión de este trabajo.

Mención aparte debo hacer al Profesor Norberto Corral, ya que varias de las ideas que han inspirado el trabajo se han debido a él.

También deseo subrayar mi profunda gratitud al Profesor Wolfgang Näther por su invitación a visitar en Marzo de 1998 su grupo de investigación en la Universidad de Freiberg (Alemania). Esta oportunidad ha resultado realmente valiosa, por su reconocida experiencia en el tema objeto de esta memoria, por las fructíferas discusiones científicas mantenidas con él y el Doctor Ralf Körner, así como por las sugerencias aportadas por ambos y por el Profesor Dietrich Stoyan durante mi estancia en Freiberg.

Los datos reales que han servido para ilustrar algunos de los estudios desarrollados en la memoria han sido proporcionados por un lado por Yolanda Colubi, Orencio Hernández y Teresa Sánchez del Servicio de Conservación de la Naturaleza, Dirección Regional de Montes y Medio Ambiente en la Consejería de Agricultura del Principado de Asturias y, por otro, por Aurora Fonseca, Jefe de Sección del Servicio de Nefrología, y Miguel de Zárraga, Director Médico, del Hospital Valle del Nalón de Langreo en Asturias. Quiero agradecerles sinceramente su amable colaboración y apoyo.

A Tomás Aranda y Ana Colubi he de darles las gracias por su amistad, y por su paciencia a nivel personal y su inestimable apoyo a nivel científico durante todo este tiempo.

Deseo hacer constar que a lo largo de estos años he sido miembro del equipo de investigación del Proyecto de Investigación N. 37/PB-TIC97-02 de la FICYT, cuya ayuda económica debe agradecerse a la Consejería de Cultura del Principado de Asturias.

Para terminar, no quiero dejar pasar esta ocasión para agradecer de forma muy especial a mis amigos, por serlo, y a mis padres y hermanos por su cariño y la confianza que siempre han depositado en mí.

Prólogo

Uno de los objetivos fundamentales en Estadística es la descripción de un conjunto de observaciones en términos de unas pocas medidas que resumen ese conjunto. Entre las principales medidas resumen cabe destacar las de localización (y de forma especial el valor esperado), y las de variación (o dispersión).

Cuando no existe variación entre las observaciones disponibles, la metodología estadística pierde prácticamente todo su interés. Además, la cuantificación de la variación cobra mayor sentido cuando se desean comparar poblaciones, muestras, variables, estimadores, etc.

La variación correspondiente a la magnitud aleatoria a partir de la que se obtienen las observaciones puede cuantificarse mediante medidas absolutas o mediante medidas relativas. Las primeras suelen medir la variación en las mismas unidades que la magnitud considerada (o en las unidades al cuadrado), y vienen a dar una idea de hasta qué punto la medida de localización (o, en general, cierto valor de referencia) representa los valores de esa magnitud. Las medidas de variación relativas (más concretamente, las medidas de desigualdad) suelen ser índices carentes de unidades, y vienen a dar una idea de hasta qué punto la medida de localización (o valor de referencia) queda por encima o por debajo de los valores de la magnitud que se considere.

Las medidas más usuales de variación absoluta se definen en términos de desviaciones o distancias entre ciertos valores (habitualmente, entre los valores de la magnitud y el valor de referencia), de forma que resultan invariantes por traslaciones de los valores de esa magnitud. En consecuencia, se adaptan de forma especialmente idónea a las variables aleatorias reales cuyas observaciones estén realizadas en una escala de intervalo.

Las medidas más usuales de variación relativa se definen en términos de cocientes entre ciertos valores, de forma que resultan invariantes por escala. En consecuencia, se adaptan de manera especial al tratamiento de las variables aleatorias reales cuyas observaciones estén realizadas en una escala de razón.

En esta memoria va a presentarse la extensión de algunas de las medidas de variación tradicionales para elementos aleatorios con valores imprecisos de cierto tipo: valores difusos y de intervalo.

El modelo adoptado para caracterizar los elementos aleatorios que toman valores difusos es el dado por las variables aleatorias difusas (también llamados conjuntos difusos aleatorios), según han sido definidas por Puri & Ralescu (1986), mientras que el modelo para los elementos aleatorios que toman valores de intervalo es el de los conjuntos aleatorios compactos y convexos (ver Matheron 1975).

Las medidas de variación (absoluta y relativa) que se introducen en la memoria toman valores reales. No obstante, algunos autores han señalado que la variación (y, en particular, la desigualdad) es una característica imprecisa en sí misma, incluso cuando afecta a magnitudes con valores reales. Sin embargo, como a lo largo de la memoria se supone que la misión fundamental de las medidas de variación va a ser la de servir de base para la comparación entre poblaciones, magnitudes, etc., se ha optado por la definición de índices cuyos valores sean comparables directamente.

En la memoria aparece un capítulo de introducción en el que se recogen los conceptos y resultados de apoyo para los estudios desarrollados.

El núcleo del trabajo se presenta desglosado en tres capítulos:

El Capítulo 1 está dedicado a la introducción de una medida generalizada de la variación absoluta (desviación) de una variable aleatoria difusa. A continuación, se estudian propiedades de esta medida, y se llevan a cabo diversas inferencias sobre su valor en una población a partir de muestras extraídas de la misma. Posteriormente, se analizan dos aplicaciones estadísticas de la medida introducida: la cuantificación del error de muestreo en la estimación del valor esperado de una variable aleatoria difusa en el muestreo de poblaciones finitas y la cuantificación del error asociado a una relación lineal o funcional general entre dos variables aleatorias difusas para la determinación de las relaciones óptimas.

En el Capítulo 2 se introduce una medida generalizada de la variación relativa (desigualdad) de una variable aleatoria difusa. Se examinan sus propiedades y se llevan a cabo dos tipos de inferencias estadísticas: la estimación de una medida particular en muestreos aleatorios de poblaciones finitas, y el estudio de la distribución asintótica en poblaciones finitas (y bajo condiciones muy generales) de la medida generalizada asociada a una muestra aleatoria a partir de una variable aleatoria difusa.

En el Capítulo 3 se particularizan los conceptos y resultados de los Capítulos 1 y 2 al caso especial en que la variable aleatoria difusa se reduce a un conjunto aleatorio compacto y convexo.

Cada uno de los tres capítulos se concluye con una valoración final sobre las contribuciones del mismo y algunos problemas abiertos relacionados, y contiene varios ejemplos ilustrativos.

En el Epílogo de la memoria se comentan sucintamente algunos de los aspectos generales de la misma, y se incluye una breve discusión sobre la metodología seguida.

Por último, en sendos apéndices aparecen los resultados relativos a la estructura métrica del espacio que sirve de base para la generalización de las medidas de variación absoluta, y algunos resultados de caracterización de las funciones que dan lugar a la generalización de las medidas de variación relativa.

Contenido

Índice de Figuras

Índice de Tablas

Conceptos previos y fundamentos

En este capítulo se recogen conceptos y resultados básicos de la Teoría de Conjuntos Difusos y la Teoría de Conjuntos Aleatorios, junto con otros conceptos y resultados de apoyo sobre la cuantificación de la desigualdad asociada a una variable aleatoria.

0.1 Conceptos y resultados sobre conjuntos difusos

A lo largo de todo el trabajo, los datos experimentales involucrados se suponen imprecisos. Los modelos que van a tratar esa imprecisión serán por un lado los intervalos reales y, por otro, ciertos conjuntos difusos del espacio $\mathbb{R}$ de los números reales.

Los subconjuntos difusos de $\mathbb{R}$ con los que se va a trabajar en la memoria satisfacen ciertas condiciones que se recogen en la definición siguiente:

Definición 0.1.1. *El conjunto $\mathcal{F}_c(\mathbb{R})$ es la clase de los subconjuntos difusos de $\mathbb{R}$, $\widetilde{V} : \mathbb{R} \to [0,1]$, que verifican las siguientes condiciones:*

i) *el α-corte de $\widetilde{V}$, $\widetilde{V}_\alpha = \{x \in \mathbb{R} \,|\, \widetilde{V}(x) \geq \alpha\}$, es un conjunto compacto para todo $\alpha \in (0,1]$,*

ii) $\widetilde{V}_1 = \{x \in \mathbb{R} \,|\, \widetilde{V}(x) = 1\} \neq \emptyset$ *(es decir, $\widetilde{V}$ es normal),* *iii)* $\widetilde{V}$ *es un conjunto convexo, es decir, los conjuntos* $\widetilde{V}_\alpha$ *son conjuntos convexos clásicos para todo* $\alpha \in (0,1]$,

iv) *la envolvente convexa cerrada del soporte de* $\widetilde{V}$, *denotada por* $\widetilde{V}_0$ *(con* $\text{sop}\,\widetilde{V} = \{x \in \mathbb{R} \,|\, \widetilde{V}(x) > 0\}$*), que en este caso coincide con su clausura (es decir,* $\widetilde{V}_0 = \text{cl}\,[\text{sop}\,\widetilde{V}]$*), es compacta.*

Obviamente, $\mathcal{F}_c(\mathbb{R})$ puede describirse brevemente como la clase de los subconjuntos difusos $\widetilde{V}$ de $\mathbb{R}$ tales que $\widetilde{V}_\alpha \in \mathcal{K}_c(\mathbb{R})$ para todo $\alpha \in [0,1]$, supuesto que $\mathcal{K}_c(\mathbb{R})$ es la clase de los intervalos compactos no vacíos de $\mathbb{R}$.

Se hará referencia de aquí en adelante a los elementos de $\mathcal{F}_c(\mathbb{R})$ con la denominación genérica de *números difusos.*

0.1.1 Operaciones entre números difusos

El manejo estadístico de datos imprecisos con valores difusos o de intervalo requiere usualmente la consideración de operaciones elementales (especialmente la suma y el producto por valores reales) entre esos datos. Las operaciones entre intervalos son muy conocidas y su resultado se define como la imagen del producto cartesiano de ambos mediante las aplicaciones que representan las operaciones correspondientes.

Las operaciones entre subconjuntos difusos (y, en particular, entre números difusos) se definen a partir del *Principio de extensión* (Zadeh 1975) como sigue:

Principio de extensión Sea $f : \mathbb{R}^k \to \mathbb{R}$, y sean $\widetilde{V}_i$ subconjuntos difusos de $\mathbb{R}$, para $i \in \{1, \ldots, k\}$. Se define la *imagen (difusa) de* $\left(\widetilde{V}_1, \ldots, \widetilde{V}_k\right)$ *inducida* por f, $\widetilde{W} = f\left(\widetilde{V}_1, \ldots, \widetilde{V}_k\right)$ como un subconjunto difuso de $\mathbb{R}$ dado por:

$$\widetilde{W}(y) = f\left(\widetilde{V}_1, \ldots, \widetilde{V}_k\right)(y)$$

$$= \begin{cases} \displaystyle\sup_{\substack{(x_1,\ldots,x_k)| \\ f(x_1,\ldots,x_k)=y}} \text{mín}\{\widetilde{V}_1(x_1), \ldots, \widetilde{V}_k(x_k)\} & \text{si } f^{-1}(\{y\}) \neq \emptyset \\ 0 & \text{si } f^{-1}(\{y\}) = \emptyset. \end{cases}$$

Observación 0.1.1. Puede demostrarse a partir de los resultados de Nguyen (1978), que si f es una función inyectiva o continua se verifica para todo $\alpha \in (0,1]$ que:

$$\left(f\left(\widetilde{V}_1, \ldots, \widetilde{V}_r\right)\right)_\alpha = f\left(\left(\widetilde{V_1}\right)_\alpha, \ldots, \left(\widetilde{V_r}\right)_\alpha\right),$$

(para una demostración detallada de este resultado ver, por ejemplo, Sánchez de Posada Martínez 1998).

La aplicación del Principio de extensión al establecimiento de las *operaciones algebraicas entre subconjuntos difusos* (ver, por ejemplo, Dubois & Prade 1978, 1987, Kaufmann & Gupta 1991, Mareš 1994), conduce a las definiciones de suma $\oplus$ (o, alternativamente, $\widetilde{\sum}$), substracción $\ominus$, producto por escalar $\odot$, producto $\otimes$ y cociente $\oslash$ difusos, que para $\widetilde{V}$ y $\widetilde{W} \in \mathcal{F}_c(\mathbb{R})$ (las dos últimas supuestas definidas para elementos de $\mathcal{F}_c((0,+\infty))$ -es decir, $\widetilde{V}$, $\widetilde{W} \in \mathcal{F}_c(\mathbb{R})$ con $\widetilde{V}_0$, $\widetilde{W}_0 \in \mathcal{P}((0,+\infty))$-) son las siguientes:

$$\left(\widetilde{V} \oplus \widetilde{W}\right)(z) = \sup_{(x,y)|x+y=z} \min\{\widetilde{V}(x), \widetilde{W}(y)\},$$

$$\left(\widetilde{V} \ominus \widetilde{W}\right)(z) = \sup_{(x,y)|x-y=z} \min\{\widetilde{V}(x), \widetilde{W}(y)\},$$

$$\left(\lambda \odot \widetilde{V}\right)(z) = \begin{cases} \widetilde{V}(z/\lambda) & \text{si } \lambda \neq 0, \\ \mathbf{1}_{\{0\}}(z) & \text{si } \lambda = 0, \end{cases}$$

$$\left(\lambda \odot \widetilde{V}\right)(z) = \widetilde{V}(z/\lambda), \ \text{ si } \lambda \neq 0,$$

$$\left(\widetilde{V} \otimes \widetilde{W}\right)(z) = \sup_{(x,y)|x\times y=z} \min\{\widetilde{V}(x), \widetilde{W}(y)\},$$

$$\left(\widetilde{V} \oslash \widetilde{W}\right)(z) = \sup_{(x,y)|x/y=z} \min\{\widetilde{V}(x), \widetilde{W}(y)\}.$$

Observación 0.1.2. En virtud de las conclusiones derivadas a partir de Nguyen (1978) la aritmética entre elementos de $\mathcal{F}_c(\mathbb{R})$ puede reducirse a la aritmética de intervalos, en el sentido de que para $\alpha \in [0,1]$ (el caso $\alpha = 0$

puede encontrarse demostrado en López García 1997) los α-cortes de los resultados de las operaciones en $\mathcal{F}_c(\mathbb{R})$ pueden expresarse en términos de operaciones de elementos de $\mathcal{K}_c(\mathbb{R})$, más concretamente de los α-cortes de los elementos de $\mathcal{F}_c(\mathbb{R})$ considerados. De este modo, si $\widetilde{V}$ y $\widetilde{W} \in \mathcal{F}_c(\mathbb{R})$ se tiene para todo $\alpha \in [0,1]$ que $\left(\widetilde{V} \oplus \widetilde{W}\right)_\alpha = \widetilde{V}_\alpha + \widetilde{W}_\alpha$ (con $+ =$ suma de Minkowski), es decir:

$$\left(\widetilde{V} \oplus \widetilde{W}\right)_\alpha = \left[\inf \widetilde{V}_\alpha + \inf \widetilde{W}_\alpha, \sup \widetilde{V}_\alpha + \sup \widetilde{W}_\alpha\right],$$

que $\left(\widetilde{V} \ominus \widetilde{W}\right)_\alpha = \widetilde{V}_\alpha - \widetilde{W}_\alpha$ (con $- =$ substracción de Minkowski), es decir:

$$\left(\widetilde{V} \ominus \widetilde{W}\right)_\alpha = \left[\inf \widetilde{V}_\alpha - \sup \widetilde{W}_\alpha, \sup \widetilde{V}_\alpha - \inf \widetilde{W}_\alpha\right],$$

y si $\lambda \in \mathbb{R} \setminus \{0\}$, se cumple que $\left(\lambda \odot \widetilde{V}\right)_\alpha = \lambda \widetilde{V}_\alpha$, es decir:

$$\left(\lambda \odot \widetilde{W}\right)_\alpha = \left[\lambda \cdot \inf \widetilde{W}_\alpha, \lambda \cdot \sup \widetilde{W}_\alpha\right], \quad \text{si } \lambda > 0,$$

$$\left(\lambda \odot \widetilde{W}\right)_\alpha = \left[\lambda \cdot \sup \widetilde{W}_\alpha, \lambda \cdot \inf \widetilde{W}_\alpha\right], \quad \text{si } \lambda < 0.$$

Si $\widetilde{V}$ y $\widetilde{W} \in \mathcal{F}_c((0,+\infty))$ se cumple para todo $\alpha \in [0,1]$ que $\left(\widetilde{V} \otimes \widetilde{W}\right)_\alpha = \widetilde{V}_\alpha \cdot \widetilde{W}_\alpha$, es decir:

$$\left(\widetilde{V} \otimes \widetilde{W}\right)_\alpha = \left[\inf \widetilde{V}_\alpha \cdot \inf \widetilde{W}_\alpha, \sup \widetilde{V}_\alpha \cdot \sup \widetilde{W}_\alpha\right],$$

y que $\left(\widetilde{V} \oslash \widetilde{W}\right)_\alpha = \widetilde{V}_\alpha / \widetilde{W}_\alpha$, es decir:

$$\left(\widetilde{V} \oslash \widetilde{W}\right)_\alpha = \left[\inf \widetilde{V}_\alpha / \sup \widetilde{W}_\alpha, \sup \widetilde{V}_\alpha / \inf \widetilde{W}_\alpha\right].$$

A partir de las operaciones difusas suma y producto por escalar pueden establecerse de forma análoga al caso real las correspondientes operaciones matriciales (ver Dubois & Prade 1980).

Sobre las operaciones entre números difusos pueden verse las conclusiones de Kaufmann & Gupta (1991), Mareš (1994) y la recopilación reciente de Sánchez de Posada Martínez (1998).

0.1.2 Ordenación de números difusos e intervalos reales compactos

Cuando se lleva a cabo un estudio estadístico con datos imprecisos, especialmente cuando uno de los objetivos fundamentales es la comparación entre poblaciones, variables, etc. debe recurrirse a menudo a establecer comparaciones entre valores imprecisos. Existen muchos criterios de ordenación de números difusos, en ocasiones basados en ordenaciones de intervalos (ver, por ejemplo, Adamo 1980, Yager 1981, Bortolan & Degani 1985, Ramík & Římánek 1985, Kołodziejczyk 1986, Nakamura 1986, Delgado *et al.* 1988, Campos & González 1989, Tseng & Klein 1989, González & Vila 1992 y también Sánchez de Posada Martínez 1998 para una recopilación reciente).

Algunos de estos criterios están basados en relaciones nítidas, es decir, determinan un orden (o preorden) categórico entre los números difusos comparables, y otros están basados en relaciones difusas, de manera que indican el grado de verdad con el que un número difuso se considera al menos igual a otro. Dentro de los primeros cabe distinguir otros dos grupos: los que conducen a un orden (o preorden) total, y los que llevan a un orden (o preorden) parcial.

En esta memoria se hará uso de dos criterios diferentes, a los que se recurrirá a lo largo del Capítulo 2.

Criterio de Ramík & Římánek

Es un criterio que da lugar a un orden parcial sobre $\mathcal{F}_c(\mathbb{R})$ y que podría catalogarse de universalmente aceptable. Se trata del criterio basado en la relación siguiente (Ramík & Římánek 1985):

Definición 0.1.2. *Si* $\widetilde{V}, \widetilde{W} \in \mathcal{F}_c(\mathbb{R})$, *entonces* $\widetilde{V} \succeq_S \widetilde{W}$ *si, y sólo si:*

$$\inf \widetilde{V}_\alpha \geq \inf \widetilde{W}_\alpha \ y \ \sup \widetilde{V}_\alpha \geq \sup \widetilde{W}_\alpha$$

para todo $\alpha \in [0,1]$.

La relación $\succeq_S$ puede verse como la generada por las operaciones reticulares $\widetilde{\text{máx}}$ y $\widetilde{\text{mín}}$ (basadas ambas en el Principio de extensión de Zadeh) sobre $\mathcal{F}_c(\mathbb{R})$, de modo que $\widetilde{V}\succeq_S\widetilde{W}$ si, y sólo si, $\widetilde{\text{máx}}\{\widetilde{V},\widetilde{W}\} = \widetilde{V}$ y $\widetilde{\text{mín}}\{\widetilde{V},\widetilde{W}\} = \widetilde{W}$. La relación $\succeq_S$ puede verse también como la extensión del criterio de dominancia fuerte de González & Vila (1992) para un sistema de ordenación denso en $[0,1]$ (ver López García 1997).

El inconveniente práctico de este criterio es que no siempre es aplicable, y existen muchos pares de números difusos no comparables a través de este criterio, aunque el cumplimiento de $\widetilde{V}\succeq_S\widetilde{W}$ y $\widetilde{W}\succeq_S\widetilde{V}$ (que se denotaría por $\widetilde{V}\sim_S\widetilde{W}$) equivale a que $\widetilde{V}=\widetilde{W}$.

El criterio siguiente evita la falta de comparabilidad que surge a veces con el criterio precedente, y es acorde con éste cuando es aplicable.

Criterio de Yager

Este criterio está basado en una de las funciones de ordenación definidas por Yager (1981) y que se establece como sigue:

Definición 0.1.3. *Si $\widetilde{V}$, $\widetilde{W} \in \mathcal{F}_c(\mathbb{R})$, se dice que $\widetilde{V}$ es preferible o indiferente a $\widetilde{W}$ de acuerdo con el criterio de ordenación de Yager, y se denota por $\widetilde{V} \geq_Y \widetilde{W}$, si, y sólo si, se verifica que $F(\widetilde{V}) \geq F(\widetilde{W})$, donde la función de ordenación F asigna a cada valor difuso $\widetilde{V}$ de $\mathcal{F}_c(\mathbb{R})$ su valor $0,5$-promedio, dado por:*

$$F(\widetilde{V}) = \frac{1}{2}\int_{[0,1]} \left[\sup \widetilde{V}_\alpha + \inf \widetilde{V}_\alpha\right] d\alpha.$$

La función de ordenación F tiene una interpretación muy interesante (ver González 1990) en términos del valor medio del número difuso $\widetilde{V}$ definido por Dubois & Prade (1987) como el intervalo $E(\widetilde{V}) = \left[E_*(\widetilde{V}), E^*(\widetilde{V})\right]$, con $E_*(\widetilde{V}) = \int_{[0,1]} \inf \widetilde{V}_\alpha \, d\alpha$ y $E^*(\widetilde{V}) = \int_{[0,1]} \sup \widetilde{V}_\alpha \, d\alpha$, que se entienden como el inferior y el superior de los valores esperados de $\widetilde{V}$ respectivamente.

Entonces, para cada $\widetilde{V} \in \mathcal{F}_c(\mathbb{R})$, se tiene que:

$$F(\widetilde{V}) = \frac{E^*(\widetilde{V}) + E_*(\widetilde{V})}{2}.$$

El criterio de Yager determina un preorden total sobre $\mathcal{F}_c(\mathbb{R})$ y es un caso especial del criterio paramétrico introducido por Campos & González (1989).

0.1.3 Distancias entre números difusos

Existen diversas definiciones de posibles métricas sobre el espacio $\mathcal{F}_c(\mathbb{R})$. En Diamond & Kloeden (1994) puede verse una amplia recopilación de varias de estas métricas.

En este apartado sólo van a recordarse las dos que van a servir de apoyo a los conceptos básicos de la memoria.

Distancia d_∞

La primera de estas métricas, en ocasiones denominada "distancia de Hausdorff generalizada" fue introducida por Puri & Ralescu (1981,1983) y en ella está fundamentada la noción de medibilidad para los elementos aleatorios con valores en $\mathcal{F}_c(\mathbb{R})$.

Definición 0.1.4. *Si $\widetilde{V}$, $\widetilde{W} \in \mathcal{F}_c(\mathbb{R})$, la distancia "generalizada" de Hausdorff entre $\widetilde{V}$ y $\widetilde{W}$, $d_\infty(\widetilde{V}, \widetilde{W})$, viene dada por el valor:*

$$d_\infty(\widetilde{V}, \widetilde{W}) = \sup_{\alpha \in [0,1]} d_H(\widetilde{V}_\alpha, \widetilde{W}_\alpha),$$

donde d_H representa la distancia de Hausdorff en $\mathcal{K}_c(\mathbb{R})$, que para K, $K' \in \mathcal{K}_c(\mathbb{R})$, corresponde a:

$$d_H(K, K') = \max \left\{ \sup_{a \in K} \inf_{b \in K'} |a - b|, \sup_{b \in K'} \inf_{a \in K} |a - b| \right\},$$

y al ser K y K' convexos puede definirse alternativamente por:

$$d_H(K, K') = \max \{ |\inf K - \inf K'|, |\sup K - \sup K'| \}.$$

En Klement *et al.* (1986) se prueba que $(\mathcal{F}_c(\mathbb{R}), d_\infty)$ es un espacio métrico no separable.

Aunque, como se verá más tarde, el empleo de la métrica d_∞ para formalizar la medibilidad de los elementos aleatorios $\mathcal{F}_c(\mathbb{R})$-valorados dota a estos elementos de propiedades de gran interés, d_∞ no es siempre la métrica más apropiada y operativa en $\mathcal{F}_c(\mathbb{R})$ cuando se emplea para medir el error en la "estimación" de ciertas características.

Distancia D_S

Bertoluzza *et al.* (1995a) han sugerido con este último propósito la siguiente distancia basada en la selección previa de una medida S asociada al espacio medible $([0,1], \mathcal{B}_{[0,1]})$ (con $\mathcal{B}_{[0,1]}$ la σ-álgebra de Borel sobre $[0,1]$), que se supone es una medida normalizada ponderada, que puede expresarse como la suma de una medida absolutamente continua respecto a la medida de Lebesgue m en $[0,1]$ y una medida discretizada en un conjunto finito de puntos $\lambda_1, \ldots, \lambda_T$, es decir:

$$dS = g\,dm \ \text{ (o, simplemente, } dS(\lambda) = g(\lambda)\,d\,\lambda)$$

con:

$$g(\lambda) = \overline{g}(\lambda) + \sum_{t=1}^{T} k_t\,\delta(\lambda - \lambda_t),$$

donde $\overline{g}$ representa una función medible Lebesgue, y δ corresponde a la distribución de Dirac (es decir, $\delta(\lambda - \lambda_t) = 1$ si $\lambda = \lambda_t$, $\delta(\lambda - \lambda_t) = 0$ en el resto), y la función g verifica las siguientes propiedades:

$$g(\lambda) \geq 0,$$

$$\int_{[0,1]} g(\lambda)\,d\lambda = 1,$$

$$g(0) > 0, \quad g(1) > 0,$$

$$\lambda_1 = 0, \quad \lambda_T = 1 \quad \text{si } T > 1.$$

Entonces:

Definición 0.1.5. *Si S es una medida sobre $([0,1], \mathcal{B}_{[0,1]})$ que satisface las condiciones anteriores, se define la S-distancia entre dos elementos de $\mathcal{F}_c(\mathbb{R})$ como la aplicación $D_S : \mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \to [0, +\infty)$ tal que para $\widetilde{V}$, $\widetilde{W} \in \mathcal{F}_c(\mathbb{R})$:*

$$D_S(\widetilde{V}, \widetilde{W}) = \sqrt{\int_{[0,1]} \left[d_S(\widetilde{V}_\alpha, \widetilde{W}_\alpha) \right]^2 d\,\alpha},$$

donde:

$$\begin{aligned} d_S(\widetilde{V}_\alpha, \widetilde{W}_\alpha) &= \sqrt{\int_{[0,1]} \left[f_{\widetilde{V}}(\alpha, \lambda) - f_{\widetilde{W}}(\alpha, \lambda) \right]^2 d\,S(\lambda)} \\ &= \sqrt{\int_{[0,1]} \left[f_{\widetilde{V}_\alpha}(\lambda) - f_{\widetilde{W}_\alpha}(\lambda) \right]^2 d\,S(\lambda)}, \end{aligned}$$

con:

$$f_{\widetilde{V}_\alpha}(\lambda) = \lambda \sup \widetilde{V}_\alpha + (1 - \lambda) \inf \widetilde{V}_\alpha = f_{\widetilde{V}}(\alpha, \lambda).$$

Se supone que S verifica las condiciones consideradas antes de la Definición 0.1.5 para garantizar que D_S sea una métrica. De otra forma, $D_S(\widetilde{V}, \widetilde{W}) = 0$ podría no implicar necesariamente que $\widetilde{V} = \widetilde{W}$. Así, si se supone que $g(1) = 1$ y $g(\lambda) = 0$ para todo $\lambda \in [0, 1)$, y $\widetilde{V}$ es el número difuso triangular $\mathrm{Tri}(0, 1, 2)$ y $\widetilde{W}$ es el número difuso trapezoidal $\mathrm{Tra}(0, 0, 1, 2)$, entonces $\widetilde{V}$ y $\widetilde{W}$ no coinciden pero $D_S(\widetilde{V}, \widetilde{W}) = 0$.

A diferencia de lo que ocurría con d_∞, el espacio $(\mathcal{F}_c(\mathbb{R}), D_S)$ es un espacio métrico separable (ver Apéndice A).

La métrica D_S puede particularizarse al caso en el que $S = S_{\vec{\lambda}}$ (con $\vec{\lambda} = (k_1, k_2, k_3)$, $k_1 + k_2 + k_3 = 1$ y $k_1, k_2, k_3 \in [0, 1))$, donde $S_{\vec{\lambda}}$ es la medida ponderada parametrizada tal que la función asociada g satisface que $g(0) = k_1$, $g(0,5) = k_2$, $g(1) = k_3$ y $g(\lambda) = 0$ en cualquier otro caso, llegando a la distancia introducida por Salas (1991) y considerada por Lubiano *et al.* (1999a).

Si se denota por S_2 la medida $S_{\vec{\lambda}}$ correspondiente a $\vec{\lambda} = (0,5,\ 0,\ 0,5)$, se obtiene que D_{S_2} coincide en el caso $\mathcal{F}_c(\mathbb{R})$ con la distancia δ_2 considerada por Näther (1997) y Körner (1997ab).

La distancia δ_2 es un caso especial de la distancia δ_p que está definida (ver Diamond & Kloeden 1994) para $\widetilde{V}$, $\widetilde{W} \in \mathcal{F}_c(\mathbb{R}^d)$ (clase de los subconjuntos difusos de $\mathbb{R}^d$, es decir, aplicaciones de $\mathbb{R}^d$ en $[0,1]$, que satisfacen las condiciones *(i)*-*(iv)* de la Definición 0.1.1) y $p \in [1,+\infty)$ como:

$$\delta_p(\widetilde{V},\widetilde{W}) = \left(\int_{[0,1]}\int_{S^{d-1}} \left|s_{\widetilde{V}}(u,\alpha) - s_{\widetilde{W}}(u,\alpha)\right|^p \,\nu(d\,u)\,d\alpha\right)^{\frac{1}{p}},$$

donde ν es la medida de Lebesgue normalizada en la esfera unidad S^{d-1}, $s_{\widetilde{V}}$ representa la función soporte de $\widetilde{V}$ (Puri & Ralescu 1985), es decir:

$$s_{\widetilde{V}}(u,\alpha) = \sup\{\langle u,v\rangle \,|\, v \in \widetilde{V}_\alpha\} \text{ para todo } u \in S^{d-1} \text{ y } \alpha \in [0,1],$$

donde $\langle\cdot,\cdot\rangle$ es el producto interno en $\mathbb{R}^d$.

Los argumentos que apoyan la consideración de la métrica generalizada D_S, y en particular la prioridad concedida a ciertas elecciones de S, se fundamentan en la breve discusión que se realiza a continuación y que se ilustra con un ejemplo.

De este modo, Bertoluzza *et al.* (1995a) han apuntado algunas críticas en el caso $S = S_{\vec{\lambda}}$ a la elección de S_2, para motivar la consideración de una elección más general. Las elecciones más comunes en ese caso son $k_2 \leq k_1 = k_3$, y de forma especial, $\vec{\lambda} = (0,375,\ 0,25,\ 0,375)$ o $\vec{\lambda} = (1/3, 1/3, 1/3)$.

Un ejemplo muy simple ilustra las ventajas de la elección de $\vec{\lambda} = (0,375,\ 0,25, 0,375)$ o $S =$ la medida de Lebesgue m en $[0,1]$, frente a $S_{\vec{\lambda}} = S_2$ cuando D_S se emplea para cuantificar un error de estimación. Se consideran los siguientes cuatro pares de números difusos triangulares, $(\widetilde{S}_1,\widetilde{S}_2)$, $(\widetilde{U}_1,\widetilde{U}_2)$, $(\widetilde{V}_1,\widetilde{V}_2)$ y $(\widetilde{W}_1,\widetilde{W}_2)$, donde $\widetilde{S}_1 = \text{Tri}(-2,0,2)$ y $\widetilde{S}_2 = \text{Tri}(-1,0,1)$ representan dos descripciones (más o menos "afinadas") del valor (o propiedad) 'ALREDEDOR DE 0', $\widetilde{U}_1 = \text{Tri}(-\sqrt{2},0,\sqrt{2})$ que es otra descripción del valor 'ALREDEDOR DE 0', $\widetilde{U}_2 = \text{Tri}(0,0,\sqrt{2})$ que es una descripción del valor 'ALREDEDOR DE 0 y positiva', $\widetilde{V}_1 = \text{Tri}(-1,0,0)$ que es una descripción de 'ALREDEDOR DE 0 y negativo', $\widetilde{V}_2 = \text{Tri}(0,0,1)$ que es una descripción de 'ALREDEDOR DE 0 y positivo', $\widetilde{W}_1 = \text{Tri}(-2,0,1)$ que es una descripción de 'ALREDEDOR DE 0 y con TENDENCIA negativa', y $\widetilde{W}_2 = \text{Tri}(-1, 0, 2)$ que es una descripción de 'ALREDEDOR DE 0 y con TENDENCIA positiva' (ver Figura 0.1).

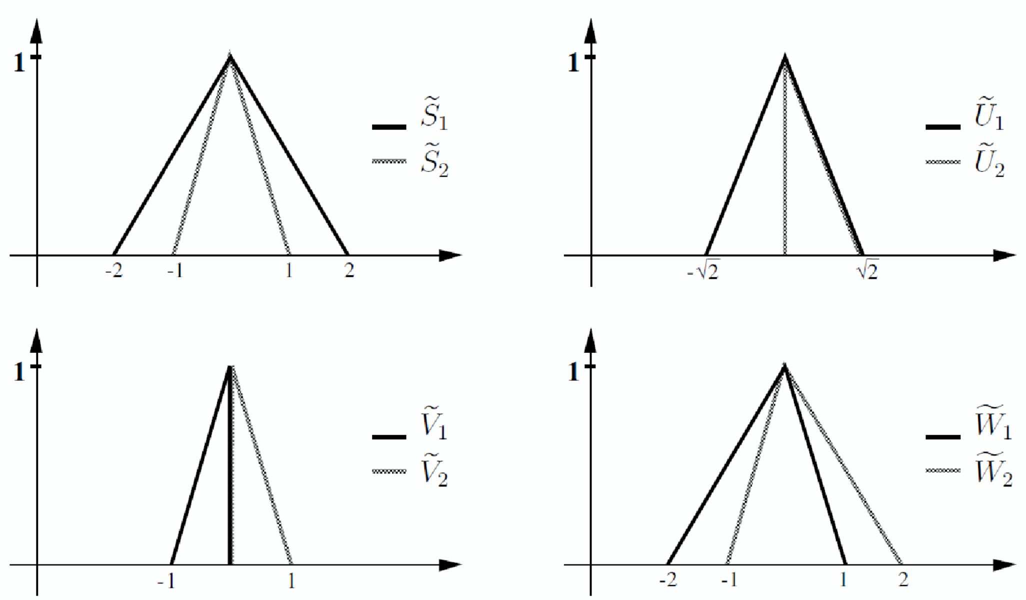


Fig. 0.1: Representaciones gráficas de $(\widetilde{S}_1,\widetilde{S}_2)$, $(\widetilde{U}_1,\widetilde{U}_2)$, $(\widetilde{V}_1,\widetilde{V}_2)$ y $(\widetilde{W}_1,\widetilde{W}_2)$.

Entonces, se tiene que $D^2_{S_2}(\widetilde{S}_1, \widetilde{S}_2) = D^2_{S_2}(\widetilde{U}_1, \widetilde{U}_2) = D^2_{S_2}(\widetilde{V}_1, \widetilde{V}_2) = D^2_{S_2}(\widetilde{W}_1, \widetilde{W}_2) = 0,333$, mientras que $D^2_{(0{,}375,\, 0{,}25,\, 0{,}375)}(\widetilde{S}_1,\ \widetilde{S}_2) = 0,25$, $D^2_{(0{,}375,\, 0{,}25,\, 0{,}375)}(\widetilde{U}_1,\ \widetilde{U}_2) = 0,292$, $D^2_{(0{,}375,\, 0{,}25,\, 0{,}375)}(\widetilde{V}_1,\ \widetilde{V}_2) = D^2_{(0{,}375,\, 0{,}25,\, 0{,}375)}(\widetilde{W}_1,\ \widetilde{W}_2) = 0,333$ y $D^2_m(\widetilde{S}_1,\ \widetilde{S}_2) = 0,111$, $D^2_m(\widetilde{U}_1,\ \widetilde{U}_2) = 0,222$, $D^2_m(\widetilde{V}_1,\ \widetilde{V}_2) = D^2_m(\widetilde{W}_1, \widetilde{W}_2) = 0,333$. Así, cuando se intenta estimar el número difuso en la primera componente de cada par anterior por medio del que aparece en la segunda componente del mismo par, el error que se comete en la estimación está mucho mejor cuantificado en términos de $D^2_{(0{,}375,\, 0{,}25,\, 0{,}375)}$ o D^2_m, que en términos de $D^2_{S_2}$.

Por otro lado, ese mismo ejemplo da idea de la conveniencia de utilizar D_S frente a d_∞ en el problema considerado. Es evidente que $D_S\left(\widetilde{V}, \widetilde{W}\right) \leq d_\infty\left(\widetilde{V}, \widetilde{W}\right)$, para todo $\widetilde{V}$, $\widetilde{W} \in \mathcal{F}_c(\mathbb{R})$ y S, por lo que la mayoría de los resultados sobre convergencias o acotaciones basadas en d_∞ serán más fuertes que los que se establecen para D_S, si bien la separabilidad de $(\mathcal{F}_c(\mathbb{R}), D_S)$ frente a la no separabilidad de $(\mathcal{F}_c(\mathbb{R}), d_\infty)$, hace que algunos resultados para el primero no sean válidos para el segundo. Sin embargo, en el ejemplo anterior, se tiene que $d_\infty(\widetilde{S}_1, \widetilde{S}_2) = d_\infty(\widetilde{V}_1, \widetilde{V}_2) = d_\infty(\widetilde{W}_1, \widetilde{W}_2) = 1$ y $d_\infty(\widetilde{U}_1, \widetilde{U}_2) = \sqrt{2}$, con lo que tampoco d_∞ parece la más adecuada si se emplea para cuantificar el error asociado con la estimación de la primera componente en cada par por medio de la segunda. La razón de este comportamiento reside en el uso que del supremo se hace en d_∞ en contraste con el del "promedio" que se hace en D_S.

En esta memoria se van a considerar experimentos aleatorios que en lugar de involucrar un proceso de cuantificación numérica, asimilable con una variable aleatoria real (es decir, una función medible Borel con valores en $\mathbb{R}$), se supone que consideran procesos de cuantificación imprecisos que se formalizarán como valorados en intervalo o valorados difusos (o, más concretamente, con valores en $\mathcal{K}_c(\mathbb{R})$ o en $\mathcal{F}_c(\mathbb{R})$).

Conviene subrayar que la imprecisión se asume como intrínseca a los procesos, y no derivada de una imprecisión en la percepción o en la transmisión de los valores de un proceso en escala numérica.

Los conceptos y resultados que se presentan en las dos secciones siguientes, corresponden a procesos de cuantificación que convierten los resultados de un experimento aleatorio en elementos de $\mathcal{K}_c(\mathbb{R})$ o $\mathcal{F}_c(\mathbb{R})$. A lo largo de estas dos secciones se partirá de la suposición de que el modelo matemático para el experimento aleatorio inicial viene dado por el espacio probabilístico $(\Omega, \mathcal{A}, P)$. Una variable aleatoria asociada a este espacio (o función $(\mathcal{A}, \mathcal{B}_{\mathbb{R}})$-medible de Ω en $\mathbb{R}$, con $\mathcal{B}_{\mathbb{R}}$ la σ-álgebra de Borel sobre $\mathbb{R}$) representaría el nivel de precisión más elevado, y el nivel de generalización inferior desde un punto de vista teórico.

0.2 Conceptos y resultados sobre conjuntos aleatorios

El nivel de precisión inferior, aunque se trate de un nivel de generalización matemática intermedio, lo ocupa el concepto de conjunto aleatorio. En esta sección va a considerarse el caso especial de *conjunto aleatorio compacto y convexo* con valores en $\mathcal{K}_c(\mathbb{R})$ (si bien puede establecerse sin dificultades una definición más general -ver, por ejemplo, Kendall 1974, Matheron 1975, Molchanov 1993-).

Si $(\Omega, \mathcal{A}, P)$ es un espacio probabilístico, entonces:

Definición 0.2.1. *Una aplicación $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ se dice que es un conjunto aleatorio compacto y convexo asociado a $(\Omega, \mathcal{A})$ si es $(\mathcal{A}, \mathcal{B}_{d_H})$-medible, con $\mathcal{B}_{d_H}$ la σ-álgebra engendrada por la topología asociada a d_H sobre $\mathcal{K}_c(\mathbb{R})$.*

El *valor esperado* de un conjunto aleatorio compacto y convexo se define en términos del concepto de integral con respecto a una medida de probabilidad de una función con valores de conjunto medible, introducido por Aumann (1965) como sigue:

Definición 0.2.2. *Si $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ es un conjunto aleatorio compacto y convexo, el valor esperado de X es la integral de Aumann de X sobre Ω respecto a P, es decir, el valor:*

$$E(X) = \left\{ \int_\Omega f(\omega)\, dP(\omega) \,\middle|\, f : \Omega \to \mathbb{R}, \right.$$
$$\left. \text{con } f \in L^1(\Omega, \mathcal{A}, P) \ \ y \ f \in X \ c.s. \, [P] \right\}.$$

Cuando se necesite especificar la medida de probabilidad P, $E(X)$ se denotará alternativamente por $E(X|P)$.

La noción siguiente establece una condición suficiente para asegurar que el conjunto $E(X)$ no es vacío y es un elemento de $\mathcal{K}_c(\mathbb{R})$.

Definición 0.2.3. *Se dice que un conjunto aleatorio compacto y convexo $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ es integrablemente acotado si existe una función $h : \Omega \to$*

$\mathbb{R}$, con $h \in L^1(\Omega, \mathcal{A}, P)$, de modo que $|X(\omega)| \leq h(\omega)$ para todo $\omega \in \Omega$, donde $|X(\omega)| = \sup_{x \in X(\omega)} |x| = \text{máx}\{|\inf X(\omega)|, |\sup X(\omega)|\}$.

La suposición de que X toma valores en $\mathcal{K}_c(\mathbb{R})$, y más concretamente la convexidad de esos valores, permite expresar $E(X)$ en forma muy simple como sigue (ver López-Díaz & Gil 1998a):

$$\begin{aligned} E(X) &= [E(\inf X), E(\sup X)] \\ &= \left[\int_\Omega \inf X(\omega) d P(\omega), \int_\Omega \sup X(\omega) d P(\omega)\right], \end{aligned}$$

ya que $\inf X$ y $\sup X$ son en este caso variables aleatorias reales.

Algunos casos especiales de conjuntos aleatorios a los que se hará referencia en la memoria son los siguientes:

Definición 0.2.4. *Se dice que un conjunto aleatorio compacto y convexo $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ es degenerado, si existe un intervalo $K \in \mathcal{K}_c(\mathbb{R})$ tal que $X = K$ casi seguro $[P]$. En particular, cuando K se reduce a un conjunto unitario de $\mathbb{R}$, se dice que X es un conjunto aleatorio compacto y convexo degenerado en un valor real.*

Definición 0.2.5. *Un conjunto aleatorio compacto y convexo $X : \Omega \to \mathcal{K}_c((0, +\infty))$ (es decir, $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ con $X(\Omega) \subset \mathcal{P}((0, +\infty))$), se dice que es un conjunto aleatorio compacto y convexo positivo.*

Observación 0.2.1. En el caso de conjuntos aleatorios compactos y convexos positivos, es inmediato ver que la condición de integrabilidad acotada de X es equivalente a $E(\sup X) < +\infty$.

Por último, el concepto de *independencia* para conjuntos aleatorios se formaliza como sigue:

Definición 0.2.6. *Se dice que dos conjuntos aleatorios compactos y convexos $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ e $Y : \Omega \to \mathcal{K}_c(\mathbb{R})$ son independientes si para cualesquiera elementos $K, K' \in \mathcal{B}_{d_H}$ se cumple que $P(X \in K, Y \in K') = P(X \in K) \cdot P(Y \in K')$.*

0.3 Conceptos y resultados sobre variables aleatorias difusas

Un nivel de precisión intermedio, pero correspondiente al nivel de generalización matemática superior, lo ocupa el concepto de variable aleatoria difusa. En esta sección va a considerarse el caso especial de variables aleatorias difusas con valores en $\mathcal{F}_c(\mathbb{R})$ (si bien Puri & Ralescu -1986- establecieron una noción más general que hacía referencia a un espacio euclídeo de dimensión finita y no obligaba a la convexidad de los valores).

Si $(\Omega, \mathcal{A}, P)$ es un espacio probabilístico, entonces:

Definición 0.3.1. *Una aplicación $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ se dice que es una variable aleatoria difusa (también llamado conjunto difuso aleatorio) si es $(\mathcal{A}, \mathcal{B}_{d_\infty})$-medible, con $\mathcal{B}_{d_\infty}$ la σ-álgebra que define la topología asociada a d_∞ sobre $\mathcal{F}_c(\mathbb{R})$.*

Observación 0.3.1. Es conveniente reseñar la conexión existente entre las nociones de variable aleatoria difusa y conjunto aleatorio. Concretamente, se puede comprobar que si $\mathcal{X}$ es una variable aleatoria difusa, entonces para todo $\alpha \in [0,1]$ la función α-corte $\mathcal{X}_\alpha$, donde $\mathcal{X}_\alpha : \Omega \to \mathcal{K}_c(\mathbb{R})$ está definida como $\mathcal{X}_\alpha(\omega) = (\mathcal{X}(\omega))_\alpha$ para todo $\omega \in \Omega$, es un conjunto aleatorio compacto y convexo (ver Puri & Ralescu 1986, Klement *et al.* 1986).

El *valor esperado* de una variable aleatoria difusa fue introducido por Puri & Ralescu (1986) como sigue:

Definición 0.3.2. *Si $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ es una variable aleatoria difusa, el valor esperado de $\mathcal{X}$ es el único subconjunto difuso de $\mathbb{R}$ (si existe), $\widetilde{E}(\mathcal{X})$, para el que se cumple para todo $\alpha \in (0,1]$ que $\left(\widetilde{E}(\mathcal{X})\right)_\alpha = E(\mathcal{X}_\alpha)$, es decir, el α-corte del valor esperado de $\mathcal{X}$ es la integral de Aumann del conjunto aleatorio compacto y convexo $\mathcal{X}_\alpha$.*

Cuando sea necesario especificar la medida de probabilidad P, $\widetilde{E}(\mathcal{X})$ se denotará alternativamente por $\widetilde{E}(\mathcal{X}|P)$.

La noción siguiente (Puri & Ralescu 1986) establece una condición suficiente para asegurar que $\widetilde{E}(\mathcal{X})$ está bien definido y es un elemento de $\mathcal{F}_c(\mathbb{R})$.

Definición 0.3.3. *Se dice que una variable aleatoria difusa $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ es integrablemente acotada si $\mathcal{X}_0$ es un conjunto aleatorio compacto y convexo integrablemente acotado.*

Observación 0.3.2. Stojaković (1994) demuestra que $\left(\widetilde{E}(\mathcal{X})\right)_0 = E\left(\mathcal{X}_0\right)$.

Observación 0.3.3. Se puede comprobar que, al tomar $\mathcal{X}$ valores en $\mathcal{F}_c(\mathbb{R})$, se cumple que $\left(\widetilde{E}(\mathcal{X})\right)_\alpha = \left[E\left(\inf \mathcal{X}_\alpha\right),\, E\left(\sup \mathcal{X}_\alpha\right)\right]$ para todo $\alpha \in [0,1]$.

Es conveniente subrayar que el criterio de ordenación de Yager en la Definición 0.1.3 resulta especialmente operativo cuando se combina con el valor esperado (difuso) de una variable aleatoria difusa. Concretamente, y sobre la base de los resultados de López-Díaz & Gil (1998a) (establecidos para el criterio más general de Campos & González 1989) se puede concluir que si $\mathcal{X}$ es una variable aleatoria difusa integrablemente acotada con valor esperado $\widetilde{E}(\mathcal{X})$, se cumple que:

$$F\left(\widetilde{E}(\mathcal{X})\right) = E(F \circ \mathcal{X}),$$

de forma que el valor de la función de ordenación F para el valor esperado (difuso) de la variable aleatoria difusa $\mathcal{X}$ se reduce al valor esperado de la variable aleatoria con valores reales $F \circ \mathcal{X}$.

Algunos casos especiales de variables aleatorias difusas a las que se hará referencia en la memoria son las siguientes:

Definición 0.3.4. *Se dice que una variable aleatoria difusa $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ es degenerada si existe un número difuso $\widetilde{V} \in \mathcal{F}_c(\mathbb{R})$ tal que $\mathcal{X} = \widetilde{V}$*

casi seguro $[P]$*. En particular, cuando* $\widetilde{V}$ *se reduce a la función indicador de un conjunto de* $\mathcal{K}_c(\mathbb{R})$*, se dice que* $\mathcal{X}$ *es una variable aleatoria difusa degenerada en un intervalo compacto real, y cuando ese conjunto es unitario se dice que* $\mathcal{X}$ *es una variable aleatoria difusa degenerada en un valor real.*

Definición 0.3.5. *Una variable aleatoria difusa* $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ *(es decir,* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *con* $\mathcal{X}_0(\Omega) \subset \mathcal{P}((0,+\infty))$*), se dice que es una variable aleatoria difusa positiva.*

Observación 0.3.4. En el caso de una variable aleatoria difusa positiva, es inmediato ver que la condición de integrabilidad acotada de $\mathcal{X}$ es equivalente a $E(\sup \mathcal{X}_0) < +\infty$.

El concepto de *independencia* de variables aleatorias difusas se formaliza como sigue:

Definición 0.3.6. *Sea* $(\Omega, \mathcal{A}, P)$ *un espacio de probabilidad. Se dice que dos variables aleatorias difusas* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *e* $\mathcal{Y} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *son independientes si para cualesquiera elementos* A*,* $B \in \mathcal{B}_{d_\infty}$ *se cumple que* $P(\mathcal{X} \in A, \mathcal{Y} \in B) = P(\mathcal{X} \in A)P(\mathcal{Y} \in B)$*.*

Por último, un resultado que va a servir de apoyo para algunos resultados de la memoria es la extensión del *Teorema de la doble esperanza* que aparece probado en López-Díaz & Gil (1998b):

Teorema 0.3.1. *Sea* $(\Omega, \mathcal{A}, P)$ *un espacio probabilístico y sean* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *e* $\mathcal{Y} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *dos variables aleatorias difusas integrablemente acotadas. Sean* $\sigma_{\mathcal{X}}$ *y* $\sigma_{\mathcal{Y}}$ *las* σ*-álgebras en* $\mathcal{F}_c(\mathbb{R})$ *inducidas a partir de* $\mathcal{A}$ *por* $\mathcal{X}$ *e* $\mathcal{Y}$*, respectivamente (es decir,* $\sigma_{\mathcal{X}} = \{B \subset \mathcal{F}_c(\mathbb{R}) \,|\, \mathcal{X}^{-1}(B) \in \mathcal{A}\}$*,* $\sigma_{\mathcal{Y}} = \{B \subset \mathcal{F}_c(\mathbb{R}) \,|\, \mathcal{Y}^{-1}(B) \in \mathcal{A}\}$*). Sean* $P_{\mathcal{X}}$ *y* $P_{\mathcal{Y}}$ *las medidas de probabilidad inducidas a partir de* P *por* $\mathcal{X}$ *e* $\mathcal{Y}$*, respectivamente.*

Se considera el espacio de probabilidad producto $(\mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{X}} \otimes \sigma_{\mathcal{Y}}, P_{\mathcal{X}} \otimes P_{\mathcal{Y}})$*, y la variable aleatoria difusa integrablemente acotada* $\mathcal{Y}^* :$

$\mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \to \mathcal{F}_c(\mathbb{R})$ *definida de forma que* $\mathcal{Y}^*(\tilde{x}, \tilde{y}) = \tilde{y}$ *para cualesquiera* $\tilde{x}, \tilde{y} \in \mathcal{F}_c(\mathbb{R})$. *Se admite que cuando* $\mathcal{X} = \tilde{x}$ *la distribución de probabilidad condicionada inducida por* $\mathcal{Y}$ *corresponde a una distribución condicionada regular en* $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{Y}})$ *denotada por* $P_{\tilde{x}}$, *es decir:*

- $P_{\tilde{x}}$ *es una medida de probabilidad sobre* $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{Y}})$ *para cada* $\tilde{x} \in \mathcal{X}(\Omega)$, *y*

- *para cada* $B \in \sigma_{\mathcal{Y}}$ *la aplicación* $h_B : \mathcal{X}(\Omega) \to [0,1]$ *tal que* $h_B(\tilde{x}) = P_{\tilde{x}}(B)$ *es una variable aleatoria real asociada al espacio medible* $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{X}})$ *que satisface que para todo* $A \in \sigma_{\mathcal{X}}$

$$P(\mathcal{X} \in A, \mathcal{Y} \in B) = \int_A P_{\tilde{x}}(B)\, d\, P_{\mathcal{X}}.$$

Si se define la aplicación $\varphi : \mathcal{X}(\Omega) \to \mathcal{F}_c(\mathbb{R})$ *tal que* $\varphi(\tilde{x}) = \widetilde{E}\,(\mathcal{Y} \,|\, \mathcal{X} = \tilde{x})$ *donde se denota por* $\widetilde{E}\,(\mathcal{Y} \,|\, \mathcal{X} = \tilde{x}) = \widetilde{E}\,(\mathcal{Y}^*_{\tilde{x}} \,|\, P_{\tilde{x}})$ *para cada* $\tilde{x} \in \mathcal{X}(\Omega)$, *entonces:*

a) φ *es una variable aleatoria difusa integrablemente acotada asociada a* $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{X}}, P_{\mathcal{X}})$;

b) $\widetilde{E}\,(\mathcal{Y} \,|\, P) = \widetilde{E}\left(\varphi \,\middle|\, P_{\mathcal{X}}\right)$, *es decir,* $\widetilde{E}\,(\mathcal{Y} \,|\, P) = \widetilde{E}\left(\widetilde{E}\,(\mathcal{Y}^*_{\tilde{x}} \,|\, P_{\tilde{x}}) \,\middle|\, P_{\mathcal{X}}\right)$.

0.4 Medidas de desigualdad asociadas a una variable aleatoria

En esta sección se recogen las nociones básicas relativas a la desigualdad de una población asociada a una variable aleatoria con valores reales.

La *desigualdad* de una población con respecto a un aspecto cuantitativo que puede formalizarse matemáticamente por medio de una variable aleatoria (habitualmente supuesta positiva), es una característica que evalúa la variación relativa del aspecto en la población. El análisis de la desigualdad, su medición numérica y las propiedades de las medidas propuestas para esta característica forman parte de un tema con múltiples aplicaciones en campos como la Economía (desigualdad de rentas, beneficios, etc.) y la Industria (concentración industrial, etc.).

En la bibliografía sobre la cuantificación de la desigualdad de una población respecto de cierta variable aleatoria, se han propuesto un gran número de índices. Muchos de éstos han sido aceptados mayoritariamente por la comunidad científica internacional, tanto matemática como de los campos de aplicación, y en los últimos años algunos de los que más se utilizan son aquellos que coinciden con (o son funciones crecientes de) los índices de desigualdad aditivamente descomponibles de orden α.

Los estudios sobre la desigualdad suponen habitualmente que los valores del aspecto considerado pertenecen al intervalo real $(0, +\infty)$, puesto que el aspecto es habitualmente monetario (en la desigualdad de rentas, beneficios, etc.) o corresponde al tamaño de una subpoblación (en la concentración industrial). Además, las variables para las que interesa medir la desigualdad suelen ser de escala de razón.

Si X es una variable aleatoria positiva, los *índices aditivamente descomponibles* de orden α se definen (ver, por ejemplo, Bourguignon 1979, Cowell 1980, Shorrocks 1980, Cowell & Kuga 1981, Eichhorn & Gehrig 1982, Zagier 1983, Gil *et al.* 1989b), si existen, como sigue:

$$\mathrm{I}^{\alpha}(X) = [\alpha(\alpha-1)]^{-1}\left[E\left(\left(\frac{X}{E(X)}\right)^{\alpha}\right) - 1\right] \text{ si } \alpha \neq 0, 1,$$

$$\mathrm{I}^{0}(X) = -E\left(\log\left(\frac{X}{E(X)}\right)\right),$$

$$\mathrm{I}^{1}(X) = E\left(\frac{X}{E(X)}\log\left(\frac{X}{E(X)}\right)\right),$$

si bien, en ocasiones, se prescinde del factor $[\alpha(\alpha-1)]^{-1}$, pero se tiene en cuenta su signo.

El papel que juega α es el de ponderación del "grado de aversión a la desigualdad", o grado de sensibilidad relativa a las transferencias entre las distintas clases.

Los índices I^{0} e I^{1} satisfacen que $\mathrm{I}^{0}(X) = \lim_{\alpha\to 0} \mathrm{I}^{\alpha}(X)$ e $\mathrm{I}^{1}(X) = \lim_{\alpha\to 1} \mathrm{I}^{\alpha}(X)$, cualquiera que sea la variable X a la que vayan asociados. I^{1} es el conocido *índice de Theil* (1967). El índice I^{0} (o *índice de tipo Shannon*) ha recibido atención especial en algunos estudios (ver Bourguignon 1979, Gil 1979, 1981, 1982) y el índice I^{-1} (o *índice hiperbólico*), y su interés práctico, se ha examinado con detalle en otros trabajos anteriores (ver Gil & Gil 1989, Gil *et al.* 1989ab, Martínez 1991).

Una familia de índices más general, que incluye todos los índices aditivamente descomponibles salvo el de Theil y el de tipo Shannon, es la familia introducida por Gastwirth (1975) (ver también Gastwirth *et al.* 1986) o familia de los *ϕ-índices* de desigualdad que se define de forma genérica para una variable aleatoria X con valores en $(0, +\infty)$ por:

$$\mathrm{I}_{\phi}(X) = E\left(\frac{\phi(X)}{\phi\left(E(X)\right)} - 1\right),$$

donde $\phi : (0, +\infty) \to \mathbb{R}$, de clase C^1 y *cóncava hacia arriba* (habitualmente supuesta esta concavidad de forma *estricta*, es decir, para todo $\lambda \in (0, 1)$ y x, $y \in (0, +\infty)$ con $x \neq y$ se cumple que $f(\lambda x+(1-\lambda)y) < \lambda f(x)+(1-\lambda)f(y)$).

Recientemente (ver Alonso *et al.* 1998), se ha definido una familia generalizada de índices de desigualdad, que incluye a todos los índices aditivamente descomponibles (también los de tipo Shannon y de Theil). Estos índices están basados en la familia generalizada de medidas de divergencia dirigida de Csiszár (1967).

Se ha señalado en trabajos previos (ver, por ejemplo, Gil *et al.* 1989b) que existe una conexión formal entre la cuantificación de la desigualdad de una población respecto a una variable aleatoria real y la Teoría de la Información. De este modo, las medidas más conocidas pueden obtenerse como versiones particulares o a partir de medidas de divergencia dirigida entre dos distribuciones de probabilidad.

En concreto, cada uno de los índices de desigualdad aditivamente descomponibles de orden α coincide, salvo por una constante, con una de las medidas de divergencia dirigida no-aditiva de orden α (Rathie 1971) para $\alpha \neq 0, 1$, o con la medida de divergencia dirigida de Kullback-Leibler (1951) para $\alpha = 0$ ó 1, entre ciertas distribuciones particulares.

Este hecho, junto con el interés incuestionable de adoptar familias de medidas muy amplias que permitan abordar convenientemente cada problema con alguna medida de la familia, nos sugiere la posibilidad de construir *medidas de desigualdad generalizadas*.

Las medidas de divergencia entre dos distribuciones, se definen como expresiones funcionales de tales distribuciones que cuantifican el "grado de discrepancia" entre ellas (entendiendo que dos distribuciones discrepan más, cuanto más fácil resulta discriminar entre ellas).

En 1967 Csiszár introdujo una familia de medidas que extiende la medida de Kullback-Leibler, que a su vez incluye como casos especiales las medidas de divergencia dirigida no-aditivas (Rathie 1971). La definición de esta familia se estableció como sigue: dadas las medidas de probabilidad P y Q sobre un espacio medible $(\Omega, \mathcal{A})$, con funciones de densidad p y q, respectivamente, respecto a una medida σ-finita ν asociada a $(\Omega, \mathcal{A})$, la *f-divergencia de Csiszár* entre P y Q respecto a P, se define a través del valor (si existe) dado por:

$$\mathrm{D}_f(P; Q) = \int_\Omega p(\omega)\, f\left(\frac{q(\omega)}{p(\omega)}\right)\, d\,\nu(\omega),$$

donde $f : [0, +\infty) \to \mathbb{R}$ es una función cóncava hacia arriba arbitraria que satisface las condiciones siguientes:

$$f(0) = \lim_{u \downarrow 0} f(u), \quad 0 \cdot f\left(\frac{0}{0}\right) = 0, \quad 0 \cdot f\left(\frac{a}{0}\right) = \lim_{\varepsilon \downarrow 0} \varepsilon \cdot f\left(\frac{a}{\varepsilon}\right) = a \cdot \lim_{u \uparrow \infty} \frac{f(u)}{u}$$

(con $a \geq 0$).

A partir de la f-divergencia de Csiszár es posible construir una nueva familia generalizada de índices de desigualdad que incluye el índice hiperbólico, la varianza normalizada, el índice de tipo Shannon y el de Theil (lo que no ocurre con la familia de Gastwirth 1975).

El índice hiperbólico ha resultado muy útil en el tratamiento de poblaciones finitas, permitiendo la determinación de estimadores insesgados en los muestreos probabilísticos y proporcionando (según diversos estudios de simulación) empíricamente menores errores relativos de muestreo que otros índices más "populares" (ver, por ejemplo, Martínez 1991, Gil *et al* 1989b, Gil & Gil 1989).

El índice de tipo Shannon es el que dispone de una propiedad de descomponibilidad aditiva más operativa y con interpretación muy intuitiva.

Aunque el índice de Theil está incluido en esta familia generalizada, cuando se extiendan los índices a ambiente difuso no se considerará, porque se exigirá monotonía a la función f para poder encontrar expresiones operativas de dichos índices, y el índice de Theil procede de una función no monótona.

Así, si X es una variable aleatoria asociada a $(\Omega, \mathcal{A})$ y se consideran las medidas de probabilidad P y Q relacionadas de forma que si p y q son sus densidades respecto a ν, $q(\omega) = p(\omega)X(\omega)/E(X)$ para todo $\omega \in \Omega$, y se supone que $f(x) = x^{-1} - 1$ para todo $x \in (0, +\infty) = \mathrm{Dom} f$, se obtiene el *índice hiperbólico*:

$$\mathrm{I}_H(X) = E\left(\frac{E(X)}{X} - 1\right).$$

Para las mismas medidas de probabilidad y la función $f(x) = -\log x$ para todo $x \in (0, +\infty) = \mathrm{Dom} f$, se obtiene el *índice de tipo Shannon*:

$$\mathrm{I}_{Sh}(X) = -E\left(\log\left(\frac{X}{E(X)}\right)\right).$$

Por último, para conseguir el índice de Theil, sólo hace falta considerar la función $f(x) = x\log(x)$ para todo $x \in (0, +\infty) = \mathrm{Dom} f$:

$$\mathrm{I}_T(X) = E\left(\frac{X}{E(X)}\log\left(\frac{X}{E(X)}\right)\right).$$

Sobre la base de las medidas anteriores se establece la siguiente:

Definición 0.4.1. *Sea* $(\Omega, \mathcal{A}, P)$ *un espacio de probabilidad y sea* $X : \Omega \to (0, +\infty)$ *una variable aleatoria positiva asociada a* $(\Omega, \mathcal{A})$*. Sea* $f : (0, +\infty) \to \mathbb{R}$ *una función estrictamente cóncava hacia arriba, continua y tal que* $f(1) = 0$*. El* f*-índice de desigualdad asociado a* X *viene dado por el valor (si existe):*

$$\mathrm{I}_f(X) = E\left(f\left(\frac{X}{E(X)}\right)\right).$$

Las propiedades siguientes, que son deseables en un índice que cuantifique la desigualdad, se cumplen para los índices que acaban de ser definidos. Así, si $(\Omega, \mathcal{A}, P)$ es un espacio de probabilidad, y $X : \Omega \to (0, +\infty)$ es una variable aleatoria integrable y positiva, se verifica que:

- **(Independencia de la media)** *Para todo* $k \in (0, +\infty)$ *se tiene que* $\mathrm{I}_f(k \cdot X) = \mathrm{I}_f(X)$.

- **(No negatividad)** $\mathrm{I}_f(X) \geq 0$.

- **(Minimalidad)** $\mathrm{I}_f(X) = 0$ *si, y sólo si,* X *es una variable aleatoria degenerada.*

Si además X está definida sobre una población $\Omega = \{\omega_1, \ldots, \omega_N\}$ y $\mathcal{A} = \mathcal{P}(\Omega)$, se verifican las siguientes propiedades:

- **(Simetría)** *Si* σ *es una permutación sobre* Ω*, entonces se verifica que* $\mathrm{I}_f(X \circ \sigma) = \mathrm{I}_f(X)$.

- **(Homogeneidad de la población)** *Si se denota por* X^{*r} *la variable aleatoria positiva definida como la extensión de* X *a la población de* $N \times r$ *individuos, obtenida a partir de la original mediante la repetición de cada individuo* r *veces, entonces* $\mathrm{I}_f(X^{*r}) = \mathrm{I}_f(X)$.

- **(Continuidad)** *Si se define una nueva variable sobre* Ω*,* $X_{h,\varepsilon}$ *con* $X_{h,\varepsilon}(\omega_h) = X(\omega_h) + \varepsilon$ *y* $X_{h,\varepsilon}(\omega_j) = X(\omega_j)$ *para todo* $j \in \{1, \ldots, N\} \setminus \{h\}$ *para un* $\varepsilon \in \mathbb{R}$ *tal que* $X(\omega_h) + \varepsilon > 0$*, entonces:*

$$\lim_{\varepsilon \to 0} \mathrm{I}_f(X_{h,\varepsilon}) = \mathrm{I}_f(X).$$

- **(Schur-convexidad)** *Si* (μ_{jl}) *es una matriz* $N \times N$ *doblemente estocástica y se define sobre* Ω *una variable aleatoria positiva* X' *que toma los valores dados por:*

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_N \end{pmatrix} = \begin{pmatrix} \mu_{11} & \cdots & \mu_{1N} \\ \vdots & \ddots & \vdots \\ \mu_{N1} & \cdots & \mu_{NN} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$$

entonces, se tiene que $\mathrm{I}_f(X) \geq \mathrm{I}_f(X')$ *con igualdad si, y sólo si,* (μ_{jl}) *equivale a una permutación sobre* Ω*.*

- **(Compatibilidad con el criterio de Lorenz)** *Si se consideran dos variables aleatorias positivas* $X : \Omega \to (0, +\infty)$ *y* $X' : \Omega \to (0, +\infty)$*, tales que* $X(\omega_N) \geq \ldots \geq X(\omega_1), X'(\omega_N) \geq \ldots \geq X'(\omega_1)$*,* $E(X) = E(X')$ *y* $X(\omega_1) + \cdots + X(\omega_k) \geq X'(\omega_1) + \cdots + X'(\omega_k)$ *para todo* $k \in \{1, \ldots, N\}$ *con desigualdad estricta para al menos un* k*, entonces se cumple que* $\mathrm{I}_f(X) < \mathrm{I}_f(X')$*.*

- **(Principio progresivo de transferencias)** *Si se supone que existen* $h, l \in \{1, \ldots, N\}$*, con* $X(\omega_h) \geq X(\omega_l) > 0$*, y se considera un valor real* $\varepsilon \in [0, X(\omega_h) - X(\omega_l)]$ *y se define* X' *a partir de* X *como sigue:*

$$X'(\omega_h) = X(\omega_h) - \varepsilon\,,\; X'(\omega_l) = X(\omega_l) + \varepsilon,$$

$$X'(\omega_j) = X(\omega_j),\ \textit{para}\ j \in \{1, \ldots, N\} \setminus \{h, l\},$$

entonces, se tiene que $\mathrm{I}_f(X) \geq \mathrm{I}_f(X')$*, con igualdad si, y sólo si,* $X' = X$ *ó* $X' = X \circ \sigma_{hl}$ *con* σ_{hl} *la permutación sobre* Ω *que intercambiaría los individuos h-ésimo y l-ésimo (es decir,* $\varepsilon = X(\omega_h) - X(\omega_l)$*).*

- **(Principio regresivo de transferencias)** *Si se supone que existen* $h, l \in \{1, \ldots, N\}$*, con* $X(\omega_h) \geq X(\omega_l) > 0$*, y se considera un valor real* $\varepsilon \in [0, X(\omega_l)]$*, y se define* X' *a partir de* X *como sigue:*

$$X'(\omega_h) = X(\omega_h) + \varepsilon\,,\; X'(\omega_l) = X(\omega_l) - \varepsilon,$$

$$X'(\omega_j) = X(\omega_j),\ \textit{para}\ j \in \{1, \ldots, N\} \setminus \{h, l\},$$

  entonces, se tiene que $\mathrm{I}_f(X') \geq \mathrm{I}_f(X)$*, con igualdad si, y sólo si,* $X' = X$ *(es decir,* $\varepsilon = 0$*).*

- **(Efectos de la agrupación)** *Si se clasifica la población* Ω *de acuerdo con una partición que la divide en* M *grupos de* $N_1, \ldots, N_M$ *individuos (*$N = N_1 + \cdots + N_M$*), de forma que el m-ésimo grupo* $\Omega_m = \{\omega_{m1}, \ldots, \omega_{mN_m}\}$*,* $m = 1, \ldots, M$*, y se supone que* Ω *está dotada con la distribución uniforme* P *y que* $\mathcal{P} = \{\Omega_m\}_{m=1}^M$ *denota la partición anterior. Si* $X : \Omega \to (0, +\infty)$ *es una variable aleatoria positiva asociada a* $(\Omega, \mathcal{P}(\Omega), P)$*, y* $X_{\mathcal{P}} : \mathcal{P} \to (0, +\infty)$ *es la variable aleatoria tal que* $X_{\mathcal{P}}(\Omega_m) =$ *el valor esperado de* X *sobre* Ω_m *(*$m = 1, \ldots, M$*), y* X_{Ω_m} *denota la restricción de* X *de* Ω *a* Ω_m *(*$m = 1, \ldots, M$*), entonces se cumple que:*

$$\mathrm{I}_f(X) \geq \mathrm{I}_f(X_{\mathcal{P}}).$$

  Por otra parte, $\mathrm{I}_f(X)$ *es igual a* $\mathrm{I}_f(X_{\mathcal{P}})$ *si, y sólo si,* X_{Ω_m} *es una variable degenerada para cada* $m \in \{1, \ldots, M\}$*.*

Para el caso particular de los índices aditivamente descomponibles se verifica, como refuerzo del *efecto de la agrupación*, la siguiente propiedad:

- **(Descomponibilidad aditiva)** *Si se clasifica la población de acuerdo con la partición* $\mathcal{P}$ *de la propiedad anterior, entonces se tiene que:*

$$\mathrm{I}^{\alpha}(X) = \mathrm{I}^{\alpha}(X_{\mathcal{P}}) + \sum_{m=1}^{M} \left(\frac{N_m}{N}\right)^{\alpha} \left(\frac{N_m X_{\mathcal{P}}(\Omega_m)}{N E(X)}\right)^{1-\alpha} \mathrm{I}^{\alpha}(X_{\Omega_m}).$$

  Además, $\mathrm{I}^{\alpha}(X)$ *es igual a* $\mathrm{I}^{\alpha}(X_{\mathcal{P}})$ *si, y sólo si, en la m-ésima subpoblación* X_{Ω_m} *es una variable degenerada para cada* $m \in \{1, \ldots, M\}$*.*

Como se ha indicado, al extender los f-índices de desigualdad al caso de variables aleatorias difusas y conjuntos aleatorios compactos y convexos se impondrán a la función f algunas condiciones adicionales para poder conseguir índices sencillos de calcular y que satisfagan propiedades de interés. En contrapartida a esta exigencia únicamente habrá de prescindirse de la extensión de algún índice, como el de Theil.

La familia de funciones f que verifican las condiciones exigidas en los Capítulos 2 y 3 para definir la f-medida de desigualdad generalizada para conjuntos aleatorios compactos y convexos y variables aleatorias difusas va a ser muy amplia. No obstante, y bajo condiciones bastante generales, se puede hacer cierta caracterización de la misma, así como de ciertas subfamilias de interés a las que se hace referencia en los Capítulos 2 y 3 para extender la propiedad de minimalidad del caso real (ver Apéndice B).

Capítulo 1

Medida generalizada de la variación absoluta (desviación) de una variable aleatoria difusa

En el estudio de las variables aleatorias difusas resulta importante describir el comportamiento de las mismas por medio de ciertas medidas que resumen algunas de sus características más relevantes. De este modo, el valor esperado de una variable aleatoria difusa se ha introducido (Puri & Ralescu 1986) como una medida resumen (con valores difusos) de la "tendencia central" de los valores de la variable.

Otra característica de interés en el caso de variables aleatorias reales es la "desviación" o variación absoluta. Esta característica permite la comparación entre diferentes variables o poblaciones, habitualmente a través de la medición de la desviación de los valores de una variable aleatoria con respecto a un valor determinado.

Ya que la desviación es un rasgo que se define principalmente con el objeto de establecer comparaciones, sería útil que (incluso cuando se trabaja con variables aleatorias difusas) la medida asociada tomase valores reales, ya que la desviación asociada a una variable aleatoria difusa respecto a un valor difuso predeterminado (en particular, a su valor esperado) reduciría en este caso las comparaciones entre variables, poblaciones o valores, a la ordenación usual entre números reales.

Una medida de la desviación media de una variable aleatoria difusa respecto a un valor difuso, debe expresar el "error" con el que se espera que este último sea una descripción de los valores de la variable. Este error puede calcularse de forma natural, en términos de una función creciente de una distancia adecuada entre números difusos.

En el caso de una variable aleatoria real, el error se expresa usualmente en términos del cuadrado de la distancia euclídea en el espacio $\mathbb{R}$. Las razones que justifican el uso del valor esperado del cuadrado de la distancia euclídea para medir la desviación media en el caso clásico son, entre otras, las siguientes: todos los valores de la variable están incluidos en su determinación; puede estimarse fácilmente en los muestreos aleatorios probabilísticos en poblaciones finitas o no; permite operar fácilmente con combinaciones lineales de variables aleatorias (especialmente, en el caso de independencia e identidad de distribución); se adapta apropiadamente al tratamiento inferencial estadístico.

En este capítulo se introduce una medida de la variación absoluta o desviación, que es una extensión del concepto de momento de segundo orden (y, en particular, del concepto de varianza) del caso clásico a una variable aleatoria difusa respecto a un valor difuso. Esa medida se establece sobre la base de la distancia generalizada de Bertoluzza *et al.* (1995a) que se ha presentado en la Definición 0.1.5.

A continuación, se examinan condiciones para la existencia de esa extensión, y se analizan propiedades de la medida de variación definida, así como la extensión de algunos resultados notables de interés posterior.

Se desarrolla también la estimación insesgada de la medida de variación introducida, en los muestreos aleatorios simple y con reposición de poblaciones finitas, cuantificándose de forma exacta el error asociado al muestreo en dicha estimación. Este estudio se complementa con el de la distribución asintótica de la medida de variación muestral, que permite la construcción de técnicas inferenciales aproximadas sobre la medida poblacional.

Por último, se presentan dos aplicaciones de la medida de variación: la cuantificación del error de muestreo asociado a la estimación difusa del parámetro difuso valor esperado de una variable aleatoria difusa en los muestreos aleatorio simple y con reposición; la cuantificación del error asociado a una aproximación funcional, en particular lineal, de la relación entre dos variables aleatorias difusas.

1.1 La S-dispersión cuadrática media asociada a una variable aleatoria difusa

En esta sección se introduce en primer lugar una medida de la variación absoluta asociada a una variable aleatoria difusa respecto a un elemento de $\mathcal{F}_c(\mathbb{R})$, y en particular respecto al valor esperado de esa variable. Se examina también una condición suficiente para su existencia.

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y sea $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ una variable aleatoria difusa integrablemente acotada asociada a $(\Omega, \mathcal{A}, P)$.

Definición 1.1.1. *La S-dispersión cuadrática media asociada a $\mathcal{X}$ con respecto a un número difuso $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$, viene dada por el valor real (si existe):*

$$DCM_S\left(\mathcal{X}, \widetilde{A}\right) = E\left(\left[D_S\left(\mathcal{X}(\,\cdot\,), \widetilde{A}\right)\right]^2\right) = \int_\Omega \left[D_S\left(\mathcal{X}(\omega), \widetilde{A}\right)\right]^2 dP(\omega).$$

La S-dispersión cuadrática media (abreviadamente S-DCM) no necesariamente existe para una variable aleatoria difusa. Van a estudiarse algunas condiciones sobre la variable $\mathcal{X}$ para garantizar que $DCM_S\left(\mathcal{X}, \widetilde{A}\right) \in \mathbb{R}$ cualquiera que sea $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$.

En primer lugar, se probará que la medibilidad de $\mathcal{X}$ implica la $(\mathcal{A}, \mathcal{B}_{\mathbb{R}})$-medibilidad de $\left[D_S\left(\mathcal{X}(\,\cdot\,), \widetilde{A}\right)\right]^2$.

Teorema 1.1.1. *Sea $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ una variable aleatoria difusa asociada a un espacio de probabilidad $(\Omega, \mathcal{A}, P)$, y sea $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$. Entonces, la función $\left[D_S\left(\mathcal{X}(\,\cdot\,), \widetilde{A}\right)\right]^2 : \Omega \to \mathbb{R}$ es $(\mathcal{A}, \mathcal{B}_{\mathbb{R}})$-medible.*

Demostración:

Sea $\left([0,1], \mathcal{M}_{[0,1]}, m\right)$ el espacio de medida de Lebesgue sobre $[0,1]$ y sea $([0,1],\ \mathcal{B}_{[0,1]}, S)$ un espacio de medida sobre $[0,1]$, en el que la medida S se supone que satisface las condiciones de la Definición 0.1.5. Se considera la aplicación $h_{\mathcal{X}} : \Omega \times [0,1] \times [0,1] \to \mathbb{R}$ definida como:

$$h_{\mathcal{X}}(\omega, \alpha, \lambda) = f_{\mathcal{X}(\omega)}(\alpha, \lambda) = \lambda \sup \mathcal{X}_\alpha(\omega) + (1-\lambda) \inf \mathcal{X}_\alpha(\omega).$$

Como $h_1 : \Omega \times [0,1] \to \mathbb{R}$ y $h_2 : \Omega \times [0,1] \to \mathbb{R}$ tales que $h_1(\omega, \alpha) = \inf \mathcal{X}_\alpha(\omega)$ y $h_2(\omega, \alpha) = \sup \mathcal{X}_\alpha(\omega)$ son aplicaciones $\mathcal{A} \otimes \mathcal{M}_{[0,1]}$-medibles (ver López-Díaz & Gil 1997b), entonces $h_{\mathcal{X}}$ es $\mathcal{A} \otimes \mathcal{M}_{[0,1]} \otimes \mathcal{B}_{[0,1]}$-medible.

Por otra parte, las aplicaciones $h_1^* : \Omega \times [0,1] \to \mathbb{R}$ y $h_2^* : \Omega \times [0,1] \to \mathbb{R}$ definidas como $h_1^*(\omega, \alpha) = \inf \widetilde{A}_\alpha$ y $h_2^*(\omega, \alpha) = \sup \widetilde{A}_\alpha$, son continuas por la izquierda con respecto a α en $[0,1]$ y no dependen de ω, y por tanto son $\mathcal{A} \otimes \mathcal{M}_{[0,1]}$-medibles. En consecuencia, la aplicación $f_{\widetilde{A}} : [0,1] \times [0,1] \to \mathbb{R}$ definida de forma que $f_{\widetilde{A}}(\alpha, \lambda) = \lambda \sup \widetilde{A}_\alpha + (1-\lambda) \inf \widetilde{A}_\alpha$ es $\mathcal{A} \otimes \mathcal{M}_{[0,1]} \otimes \mathcal{B}_{[0,1]}$-medible, de donde $\left[f_{\mathcal{X}(\cdot)}(\,\cdot\,,\,\cdot\,) - f_{\widetilde{A}}(\,\cdot\,,\,\cdot\,)\right]^2$ es $\mathcal{A} \otimes \mathcal{M}_{[0,1]} \otimes \mathcal{B}_{[0,1]}$-medible.

Como las medidas S y m son finitas, el Teorema de Fubini permite concluir que la aplicación $\left[D_S\left(\mathcal{X}(\,\cdot\,), \widetilde{A}\right)\right]^2$ es $(\mathcal{A}, \mathcal{B}_{\mathbb{R}})$-medible. □

A continuación, se establece una condición adicional con respecto a la integrabilidad acotada, para asegurar la existencia de la S-dispersión cuadrática media asociada a $\mathcal{X}$ con respecto a $\widetilde{A}$.

Teorema 1.1.2. *Sea $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ una variable aleatoria difusa asociada a un espacio de probabilidad $(\Omega, \mathcal{A}, P)$, y sea $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$. Si $|\mathcal{X}_0| \in L^2(\Omega, \mathcal{A}, P)$, la función $\left[D_S\left(\mathcal{X}(\,\cdot\,), \widetilde{A}\right)\right]^2 : \Omega \to \mathbb{R}$ es de clase $L^1(\Omega, \mathcal{A}, P)$.*

Demostración:

Como para todo $\omega \in \Omega$ y $\alpha \in [0,1]$ se tiene que:

$$|\sup \mathcal{X}_\alpha(\omega)| \leq |\mathcal{X}_0|(\omega),\ |\inf \mathcal{X}_\alpha(\omega)| \leq |\mathcal{X}_0|(\omega),$$

y $|\mathcal{X}_0| \in L^2(\Omega, \mathcal{A}, P)$, se cumple que las aplicaciones h_1 y h_2 de la demostración del Teorema 1.1.1 pertenecen a $L^2(\Omega\times[0,1], \mathcal{A}\otimes\mathcal{M}_{[0,1]}, P\otimes m)$.

Por otro lado, como $\widetilde{A}_0$ no depende ni de ω ni de α, y la función $\max\{|\sup \widetilde{A}_0|, |\inf \widetilde{A}_0|\} \in L^2(\Omega \times [0,1], \mathcal{A} \otimes \mathcal{M}_{[0,1]}, P \otimes m)$ y domina a $|\sup \widetilde{A}_\alpha|$ e $|\inf \widetilde{A}_\alpha|$, queda probado que h_1^* e h_2^* pertenecen a $L^2(\Omega\times[0,1], \mathcal{A}\otimes \mathcal{M}_{[0,1]}, P\otimes m)$. Por lo tanto, $\left[f_{\mathcal{X}(\cdot)}(\cdot\,,\cdot) - f_{\widetilde{A}}(\cdot\,,\cdot)\right]^2 \in L^1(P\otimes m\otimes S)$, por lo que, en virtud del Teorema de Fubini puede concluirse que la aplicación $\left[D_S\left(\mathcal{X}(\cdot), \widetilde{A}\right)\right]^2 \in L^1(\Omega \times [0,1] \times [0,1], \mathcal{A} \otimes \mathcal{M}_{[0,1]} \otimes \mathcal{B}_{[0,1]}, P \otimes m \otimes S)$.□

La S-dispersión cuadrática media asociada a una variable aleatoria difusa, puede definirse con respecto al valor esperado difuso de esta variable. Esta medida extiende la noción de varianza de una variable aleatoria real y se formaliza como sigue:

Definición 1.1.2. *La S-dispersión cuadrática media central asociada a $\mathcal{X}$ se define como el valor real (si existe):*

$$\Delta_S^2(\mathcal{X}) = DCM_S\left(\mathcal{X}, \widetilde{E}(\mathcal{X})\right) = \int_\Omega \left[D_S\left(\mathcal{X}(\omega), \widetilde{E}(\mathcal{X})\right)\right]^2 dP(\omega),$$

donde $\widetilde{E}(\mathcal{X})$ representa el valor esperado (difuso) de $\mathcal{X}$ con respecto a P.

Ocasionalmente, y cuando sea necesario especificar la medida de probabilidad P en el espacio probabilístico, se denotará $\Delta_S^2(\mathcal{X})$ alternativamente por $\Delta_S^2(\mathcal{X}\,|\,P)$.

De forma inmediata, si se cumplen las condiciones del Teorema 1.1.2, se puede asegurar la existencia de $\Delta_S^2(\mathcal{X})$.

1.2 Propiedades de la S-dispersión cuadrática media asociada a una variable aleatoria difusa

En esta sección se examinan algunas de las propiedades más importantes de la S-DCM. La mayoría de ellas son extensión de las que verifican los momentos de segundo orden y la varianza de las variables aleatorias reales.

A lo largo de esta sección, se supondrá que se satisfacen condiciones que garanticen la existencia de la S-DCM de la variable aleatoria difusa con respecto al valor difuso que se considere.

En primer lugar, va a verificarse que *la S-DCM extiende* al caso difuso *el momento de segundo orden* de una variable aleatoria real con respecto a un valor real.

Teorema 1.2.1 (Extensión del caso real). *Si $\mathcal{X}$ es una variable aleatoria real (es decir, $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ es una variable aleatoria difusa con $\mathcal{X}(\Omega) \subset \{\mathbf{1}_{\{x\}} \,|\, x \in \mathbb{R}\}$) y $a \in \mathbb{R}$, se cumple que $DCM_S(\mathcal{X}, \mathbf{1}_{\{a\}})$ corresponde al momento de segundo orden de $\mathcal{X}$ con respecto a a. En particular, $\Delta_S^2(\mathcal{X}) = \mathrm{Var}(\mathcal{X})$, donde* Var *denota la varianza de una variable aleatoria real.*

Demostración:

En efecto, en las condiciones supuestas se tiene que $\mathcal{X}_\alpha(\omega) = [\mathcal{X}(\omega), \mathcal{X}(\omega)]$ y $(\mathbf{1}_{\{a\}})_\alpha = \{a\}$ para todo $\alpha \in [0,1]$, de forma que $\left[f_{\mathcal{X}(\omega)}(\alpha,\lambda) - f_{\mathbf{1}_{\{a\}}}(\alpha,\lambda)\right]^2$ $= [\mathcal{X}(\omega) - a]^2$ para todo $\omega \in \Omega$. En consecuencia:

$$DCM_S(\mathcal{X}, \mathbf{1}_{\{a\}}) = \mathrm{E}\left[(\mathcal{X} - a)^2\right] = \int_\Omega (\mathcal{X}(\omega) - a)^2 \, dP(\omega),$$

que coincide con el momento de segundo orden de $\mathcal{X}$ respecto a a. □

El resultado siguiente establece la *no negatividad* de la S-dispersión cuadrática media:

Teorema 1.2.2 (No negatividad). *La S-DCM asociada a $\mathcal{X}$ con respecto a $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ es no negativa, cualesquiera que sean $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ y S que satisfaga las condiciones de la Definición* 0.1.5*.*

La propiedad siguiente *caracteriza a las variables aleatorias difusas degeneradas* en términos de la anulación de la S-DCM.

Teorema 1.2.3 (Minimalidad). *Se cumple que $DCM_S\left(\mathcal{X}, \widetilde{A}\right) = 0$ si, y sólo si, $\mathcal{X}$ es una variable aleatoria difusa degenerada en $\widetilde{A}$, es decir, $\mathcal{X} = \widetilde{A}$ c.s. $[P]$, cualesquiera que sean $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ y S que satisfaga las condiciones de la Definición* 0.1.5*.*

Demostración:

En efecto, la no negatividad de $\left[D_S\left(\mathcal{X}(\,\cdot\,), \widetilde{A}\right)\right]^2$ sobre Ω asegura que $DCM_S\left(\mathcal{X}, \widetilde{A}\right) = 0$ si, y sólo si, $\left[D_S\left(\mathcal{X}(\omega), \widetilde{A}\right)\right]^2 = 0$ c.s. $[P]$, y al ser D_S una distancia en $\mathcal{F}_c(\mathbb{R})$ lo anterior ocurre si, y sólo si, $\mathcal{X} = \widetilde{A}$ c.s. $[P]$. □

El resultado siguiente presenta un *procedimiento para calcular la S-DCM* (y, por lo tanto, *la S-DCM central*) asociada a una variable aleatoria difusa $\mathcal{X}$, a partir de los momentos de segundo orden (y varianzas) de ciertas variables aleatorias reales (las variables $f_{\mathcal{X}}(\alpha, \lambda)$).

Teorema 1.2.4 (Simplificación del Cálculo). *Cualesquiera que sean $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ y S que satisfaga las condiciones de la Definición* 0.1.5*, se tiene que:*

$$DCM_S\left(\mathcal{X}, \widetilde{A}\right) = \int_{[0,1]} \int_{[0,1]} \left(\int_{\Omega} \left[f_{\mathcal{X}(\omega)}(\alpha, \lambda) - f_{\widetilde{A}}(\alpha, \lambda)\right]^2 \, d\,P(\omega)\right) d\,S(\lambda)\, d\,\alpha.$$

En particular:

$$\Delta_S^2(\mathcal{X}) = \int_{[0,1]} \int_{[0,1]} \mathrm{Var}\left[f_{\mathcal{X}}(\alpha, \lambda)\right]\, d\,S(\lambda)\, d\,\alpha.$$

Demostración:

Puesto que en virtud del Teorema 1.1.2 se satisface que $\left[f_{\mathcal{X}(\cdot)}(\cdot\,,\,\cdot)\right.$ $\left.-f_{\widetilde{A}}(\cdot\,,\,\cdot)\right]^2 \in L^1(\Omega \times [0,1] \times [0,1], \mathcal{A} \otimes \mathcal{M}_{[0,1]} \otimes \mathcal{B}_{[0,1]}, P \otimes m \otimes S)$, la aplicación del Teorema de Fubini prueba la primera igualdad.

La particularización a $\Delta_S^2(\mathcal{X})$ se lleva a cabo de forma inmediata, teniendo en cuenta que para una variable aleatoria difusa $\mathcal{X}$ con valores en $\mathcal{F}_c(\mathbb{R})$ se cumple para todo $\alpha \in [0,1]$ que $\sup \widetilde{E}(\mathcal{X})_\alpha = E\left(\sup \mathcal{X}_\alpha\right)$, e $\inf \widetilde{E}(\mathcal{X})_\alpha = E\left(\inf \mathcal{X}_\alpha\right)$ (ver López-Díaz & Gil 1998a). □

En el siguiente lema se presenta un resultado de apoyo, al que se recurrirá en varias ocasiones a lo largo de la memoria:

Lema 1.2.5. *Sea $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ una variable aleatoria difusa integrablemente acotada asociada a un espacio de probabilidad $(\Omega, \mathcal{A}, P)$. Entonces, se tiene que:*

$$E\left(f_{\mathcal{X}}(\alpha, \lambda)\right) = f_{\widetilde{E}(\mathcal{X})}(\alpha, \lambda).$$

Demostración:

Puesto que $\mathcal{X}$ toma valores en $\mathcal{F}_c(\mathbb{R})$, se cumple que $\sup \widetilde{E}(\mathcal{X})_\alpha = E\left(\sup \mathcal{X}_\alpha\right)$ e $\inf \widetilde{E}(\mathcal{X})_\alpha = E\left(\inf \mathcal{X}_\alpha\right)$, y por lo tanto:

$$E\left(f_{\mathcal{X}}(\alpha, \lambda)\right) = \int_\Omega f_{\mathcal{X}(\omega)}(\alpha, \lambda)\, d\,P(\omega)$$
$$= \lambda E\left(\sup \mathcal{X}_\alpha\right) + (1-\lambda) E\left(\inf \mathcal{X}_\alpha\right) = f_{\widetilde{E}(\mathcal{X})}(\alpha, \lambda).$$

□

La propiedad siguiente establece una *conexión entre la S-dispersión cuadrática media* asociada a una variable aleatoria difusa con respecto a un valor difuso *y la S-dispersión cuadrática media central.*

Teorema 1.2.6 (Conexión entre la S-DCM y la S-DCM central). *Cualesquiera que sean $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ y S que satisfaga las condiciones de la Definición* 0.1.5*, se cumple que:*

$$DCM_S\left(\mathcal{X}, \widetilde{A}\right) = \Delta_S^2(\mathcal{X}) + \left[D_S\left(\widetilde{E}(\mathcal{X}), \widetilde{A}\right)\right]^2.$$

Demostración:

Fijados α y λ, $f_{\mathcal{X}}(\alpha, \lambda)$ es una variable aleatoria real y $f_{\widetilde{A}}(\alpha, \lambda)$ es un número real, por lo que al ser $\int_{\Omega} \left[f_{\mathcal{X}(\omega)}(\alpha, \lambda) - f_{\widetilde{A}}(\alpha, \lambda)\right]^2 d\,P(\omega)$ el error cuadrático medio asociado a $f_{\mathcal{X}}(\alpha, \lambda)$ en $f_{\widetilde{A}}(\alpha, \lambda)$, se cumple que esa integral coincide con $\mathrm{Var}\left[f_{\mathcal{X}}(\alpha, \lambda)\right] + \left[E\left(f_{\mathcal{X}}(\alpha, \lambda)\right) - f_{\widetilde{A}}(\alpha, \lambda)\right]^2$.

El lema anterior, y la aplicación del Teorema 1.2.4 y del Teorema de Fubini prueba el resultado. □

El teorema siguiente es una consecuencia inmediata de la anterior, y formaliza el hecho de que la $DCM_S\left(\mathcal{X}, \widetilde{A}\right)$ alcanza su valor mínimo en $\mathcal{F}_c(\mathbb{R})$ para el valor esperado difuso de $\mathcal{X}$. De este modo, la extensión que se propone en esta memoria junto con la definición de valor esperado de Puri y Ralescu son *acordes con la aproximación de Fréchet*.

Teorema 1.2.7 (Concordancia con la aproximación de Fréchet). *Cualquiera que sea S que satisfaga las condiciones de la Definición* 0.1.5, *la función $\mathcal{G}_{\mathcal{X},S} : \mathcal{F}_c(\mathbb{R}) \to [0, +\infty)$ definida de forma que $\mathcal{G}_{\mathcal{X},S}\left(\widetilde{A}\right) = DCM_S\left(\mathcal{X}, \widetilde{A}\right)$ para todo $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$, alcanza su valor mínimo cuando $\widetilde{A} = \widetilde{E}(\mathcal{X})$.*

Demostración:

En efecto, por el teorema anterior se deduce que $\mathcal{G}_{\mathcal{X},S}$ alcanza su valor mínimo cuando $\left[D_S\left(\widetilde{E}(\mathcal{X}), \widetilde{A}\right)\right]^2$ es mínima, y esto ocurre cuando $\widetilde{A} = \widetilde{E}(\mathcal{X})$ por ser D_S una métrica en $\mathcal{F}_c(\mathbb{R})$. □

A continuación se presentan varias propiedades que conciernen a la S-DCM central.

A partir del Teorema 1.2.6, se deriva que esta medida puede *calcularse alternativamente* usando la relación siguiente:

Teorema 1.2.8 (Expresión alternativa). *Cualquiera que sea S que satisfaga las condiciones de la Definición* 0.1.5*, se tiene que:*

$$\Delta_S^2(\mathcal{X}) = DCM_S\left(\mathcal{X}, \mathbf{1}_{\{0\}}\right) - \left[D_S\left(\widetilde{E}(\mathcal{X}), \mathbf{1}_{\{0\}}\right)\right]^2.$$

En los problemas estadísticos en los que se consideran las variables aleatorias difusas como modelo de los procesos que suministran datos experimentales difusos, las operaciones más comunes con elementos de $\mathcal{F}_c(\mathbb{R})$ serán la suma $\oplus$ y el producto $\odot$ por un elemento de $\mathbb{R}$. Las propiedades que se exponen a continuación examinan el comportamiento de la S-DCM central en conexión con estas dos operaciones.

El resultado siguiente establece que la S-DCM central es *invariante por traslación*, es decir, cuando se suma un valor difuso constante a cada valor de una variable aleatoria difusa, la S-DCM central no cambia. Así:

Teorema 1.2.9 (Invarianza por traslación). *Cualesquiera que sean* $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ *y* S *que satisfaga las condiciones de la Definición* 0.1.5*, se cumple que:*

$$\Delta_S^2(\mathcal{X} \oplus \widetilde{A}) = \Delta_S^2(\mathcal{X}).$$

Demostración:

Es inmediato, ya que para cada $\alpha \in [0,1]$ y cada $\omega \in \Omega$ se tiene que $\left(\mathcal{X}(\omega) \oplus \widetilde{A}\right)_\alpha = \mathcal{X}_\alpha(\omega) + \widetilde{A}_\alpha$, y por las propiedades de la varianza de las variables aleatorias reales y la definición de $f_{\mathcal{X}}(\alpha, \lambda)$, se deduce que:

$$\text{Var}\left[f_{\mathcal{X}\oplus\widetilde{A}}(\alpha,\lambda)\right] = \text{Var}\left[f_{\mathcal{X}}(\alpha,\lambda) + f_{\widetilde{A}}(\alpha,\lambda)\right] = \text{Var}\left[f_{\mathcal{X}}(\alpha,\lambda)\right]$$

cualesquiera que sean $\alpha \in [0,1]$ y $\lambda \in [0,1]$.

Por lo tanto, sobre la base del Teorema 1.2.4 se obtiene el resultado de esta proposición. □

El teorema siguiente estudia el efecto de multiplicar una variable aleatoria difusa por una constante no negativa (es decir, de realizar *un cambio de escala* en los valores de la variable); en ese caso, la S-DCM central queda multiplicada por el cuadrado de esta constante. De este modo:

Teorema 1.2.10 (Efectos del cambio de escala). *Cualquiera que sea* S *que satisfaga las condiciones de la Definición* 0.1.5 *y* $a \in [0, +\infty)$*, se tiene que:*

$$\Delta_S^2(a \odot \mathcal{X}) = a^2 \Delta_S^2(\mathcal{X}).$$

Demostración:

En efecto, como para cada $\alpha \in (0,1]$ se tiene que $(a \odot \mathcal{X})_\alpha = a\,\mathcal{X}_\alpha$ y por las propiedades de la varianza de las variables aleatorias reales, se deduce que:

$$\operatorname{Var}\left[f_{a \odot \mathcal{X}}\left(\alpha,\lambda\right)\right] = \operatorname{Var}\left[a\,\cdot f_{\mathcal{X}}\left(\alpha,\lambda\right)\right] = a^2 \operatorname{Var}\left[f_{\mathcal{X}}\left(\alpha,\lambda\right)\right],$$

en virtud del Teorema 1.2.4 se obtiene el resultado de esta proposición. □

Observación 1.2.1. Para $a \in (-\infty, 0)$, y una medida S que satisfaga las condiciones de la Definición 0.1.5 y corresponda a una distribución simétrica respecto a $\lambda = 0,5$, también se deriva la misma conclusión del último teorema, pero a diferencia del caso clásico, el resultado de la misma no es válido para cualquier $a \in \mathbb{R}$ cuando g no es simétrica respecto a $\lambda = 0,5$.

A continuación va a establecerse la relación entre *la S-DCM central de la suma de dos variables aleatorias difusas* definidas sobre un mismo espacio probabilístico y la suma de sus S-DCM centrales. Así:

Teorema 1.2.11 (S-DCM central de la suma de variables aleatorias difusas). *Cualquiera que sea S que satisfaga las condiciones de la Definición* 0.1.5*, se satisface que:*

$$\Delta_S^2(\mathcal{X} \oplus \mathcal{Y}) = \Delta_S^2(\mathcal{X}) + \Delta_S^2(\mathcal{Y}) + 2\Delta_S(\mathcal{X},\mathcal{Y}),$$

donde:

$$\Delta_S(\mathcal{X},\mathcal{Y}) = \int_{[0,1]}\int_{[0,1]} \operatorname{Cov}\left[f_{\mathcal{X}}\left(\alpha,\lambda\right), f_{\mathcal{Y}}\left(\alpha,\lambda\right)\right]\, d\,S(\lambda)\, d\,\alpha,$$

donde Cov *denota la covarianza entre las variables aleatorias reales correspondientes.*

Demostración:

En efecto, en virtud del Teorema 1.2.4, se tiene que:

$$\Delta_S^2(\mathcal{X} \oplus \mathcal{Y}) = \int_{[0,1]}\int_{[0,1]} \operatorname{Var}\left[f_{\mathcal{X}\oplus\mathcal{Y}}\left(\alpha,\lambda\right)\right]\, d\,S(\lambda)\, d\,\alpha,$$

y por las propiedades de la varianza de la suma de variables aleatorias reales, se cumple que:

$$\mathrm{Var}\left[f_{\mathcal{X}\oplus\mathcal{Y}}\left(\alpha,\lambda\right)\right]=\mathrm{Var}\left[\lambda\left(\sup\mathcal{X}_\alpha+\sup\mathcal{Y}_\alpha\right)+\left(1-\lambda\right)\left(\inf\mathcal{X}_\alpha+\inf\mathcal{Y}_\alpha\right)\right]$$

$$=\mathrm{Var}\left[\lambda\sup\mathcal{X}_\alpha+(1-\lambda)\inf\mathcal{X}_\alpha\right]+\mathrm{Var}\left[\lambda\sup\mathcal{Y}_\alpha+(1-\lambda)+\inf\mathcal{Y}_\alpha\right]$$

$$+2\,\mathrm{Cov}\left[\lambda\sup\mathcal{X}_\alpha+(1-\lambda)\inf\mathcal{X}_\alpha,\lambda\sup\mathcal{Y}_\alpha+(1-\lambda)+\inf\mathcal{Y}_\alpha\right]$$

$$=\mathrm{Var}\left[f_{\mathcal{X}}\left(\alpha,\lambda\right)\right]+\mathrm{Var}\left[f_{\mathcal{Y}}\left(\alpha,\lambda\right)\right]+2\,\mathrm{Cov}\left[f_{\mathcal{X}}\left(\alpha,\lambda\right),f_{\mathcal{Y}}\left(\alpha,\lambda\right)\right].$$

La linealidad de la integral de Lebesgue con respecto a una medida finita en [0,1] conduce a la conclusión de esta proposición. □

En particular, la *S-DCM central de la suma de variables aleatorias difusas independientes* coincide con la suma de las S-DCM centrales, es decir:

Teorema 1.2.12 (S-DCM central de la suma de variables aleatorias difusas independientes). *Si $\mathcal{X}$ e $\mathcal{Y}$ son variables aleatorias difusas independientes, para cualquier S que satisfaga las condiciones de la Definición* 0.1.5*, se verifica que:*

$$\Delta_S^2(\mathcal{X}\oplus\mathcal{Y})=\Delta_S^2(\mathcal{X})+\Delta_S^2(\mathcal{Y}).$$

Demostración:

Si $\mathcal{X}$ e $\mathcal{Y}$ son variables aleatorias difusas independientes, entonces para cada $\alpha\in[0,1]$, $\mathcal{X}_\alpha$ y $\mathcal{Y}_\alpha$ son conjuntos aleatorios compactos y convexos independientes, por lo que $\sup\mathcal{X}_\alpha$ y $\sup\mathcal{Y}_\alpha$ son variables aleatorias reales independientes, y también lo son $\inf\mathcal{X}_\alpha$ e $\inf\mathcal{Y}_\alpha$, lo que garantiza la independencia de $f_{\mathcal{X}}\left(\alpha,\lambda\right)$ y $f_{\mathcal{Y}}\left(\alpha,\lambda\right)$, con lo que, sobre la base del resultado anterior, se concluye el presente. □

El resultado siguiente es *una extensión de la Desigualdad de Tchebychev* para variables aleatorias difusas.

Teorema 1.2.13 (Extensión de la Desigualdad de Tchebychev). *Cualesquiera que sean* $\varepsilon > 0$, $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ *y* S *que satisfaga las condiciones de la Definición* 0.1.5*, se cumple que:*

$$P\left(\left\{\omega \in \Omega \,\middle|\, D_S\left(\mathcal{X}(\omega), \widetilde{A}\right) \le \varepsilon\right\}\right) \ge 1 - \frac{DCM_S(\mathcal{X}, \widetilde{A})}{\varepsilon^2}.$$

Demostración:

En el Teorema 1.1.1 se demostró que $D_S\left(\mathcal{X}(\cdot), \widetilde{A}\right)$ es $(\mathcal{A}, \mathcal{B}_{\mathbb{R}})$-medible, es decir, es una variable aleatoria real. Aplicando la Desigualdad de Tchebychev a esta variable se obtiene el resultado del enunciado. □

Por último, el resultado siguiente presenta un método que permite obtener a partir de una variable aleatoria difusa, otra que tenga el mismo valor esperado y menor S-DCM.

Este método *extiende las ideas del Teorema de Rao-Blackwell* (ver, por ejemplo, Dudewicz & Mishra 1988, Casella & Berger 1990) en su versión más simple (es decir, sin recurrir al concepto de suficiencia), y su aplicación a la estimación difusa de un parámetro difuso es inmediata.

Teorema 1.2.14 (Extensión del Teorema de Rao-Blackwell). *Sea* $(\Omega, \mathcal{A}, P)$ *un espacio probabilístico y sean* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *e* $\mathcal{Y} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *dos variables aleatorias difusas integrablemente acotadas. Sean* $\sigma_{\mathcal{X}}$ *y* $\sigma_{\mathcal{Y}}$ *las* σ*-álgebras en* $\mathcal{F}_c(\mathbb{R})$ *inducidas a partir de* $\mathcal{A}$ *por* $\mathcal{X}$ *e* $\mathcal{Y}$*, respectivamente (es decir,* $\sigma_{\mathcal{X}} = \{B \subset \mathcal{F}_c(\mathbb{R}) \,|\, \mathcal{X}^{-1}(B) \in \mathcal{A}\}$, $\sigma_{\mathcal{Y}} = \{B \subset \mathcal{F}_c(\mathbb{R}) \,|\, \mathcal{Y}^{-1}(B) \in \mathcal{A}\}$*). Sean* $P_{\mathcal{X}}$ *y* $P_{\mathcal{Y}}$ *las medidas de probabilidad inducidas a partir de* P *por* $\mathcal{X}$ *e* $\mathcal{Y}$*, respectivamente.*

Se considera el espacio de probabilidad producto $(\mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{X}} \otimes \sigma_{\mathcal{Y}}, P_{\mathcal{X}} \otimes P_{\mathcal{Y}})$*, y la variable aleatoria difusa integrablemente acotada* $\mathcal{Y}^* : \mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \to \mathcal{F}_c(\mathbb{R})$ *definida de forma que* $\mathcal{Y}^*(\tilde{x}, \tilde{y}) = \tilde{y}$ *para cualesquiera* $\tilde{x}, \tilde{y} \in \mathcal{F}_c(\mathbb{R})$*. Se admite que cuando* $\mathcal{X} = \tilde{x}$ *la distribución de probabilidad condicionada inducida por* $\mathcal{Y}$ *corresponde a una distribución condicionada regular en* $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{Y}})$ *denotada por* $P_{\tilde{x}}$*, es decir:*

- $P_{\tilde{x}}$ *es una medida de probabilidad sobre* $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{Y}})$ *para cada* $\tilde{x} \in \mathcal{X}(\Omega)$, *y*

- *para cada* $B \in \sigma_{\mathcal{Y}}$ *la aplicación* $g_B : \mathcal{X}(\Omega) \to [0,1]$ *tal que* $g_B(\tilde{x}) = P_{\tilde{x}}(B)$ *es una variable aleatoria real asociada al espacio medible* $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{X}})$ *que satisface que para todo* $A \in \sigma_{\mathcal{X}}$

$$P(\mathcal{X} \in A, \mathcal{Y} \in B) = \int_A P_{\tilde{x}}(B)\, d\, P_{\mathcal{X}}.$$

Si $\widetilde{V} \in \mathcal{F}_c(\mathbb{R})$ *cumple que* $\widetilde{E}(\mathcal{Y} \,|\, P) = \widetilde{V}$ *con* $\Delta_S^2(\mathcal{Y} \,|\, P) < \infty$, *y* $\varphi : \mathcal{X}(\Omega) \to \mathcal{F}_c(\mathbb{R})$ *se define de modo que* $\varphi(\tilde{x}) = \widetilde{E}\,(\mathcal{Y} \,|\, \mathcal{X} = \tilde{x}) = \widetilde{E}\,(\mathcal{Y}_{\tilde{x}}^* \,|\, P_{\tilde{x}})$ *para todo* $\tilde{x} \in \mathcal{X}(\Omega)$, *entonces se satisface que* $\widetilde{E}(\varphi(\mathcal{X}) \,|\, P) = \widetilde{V}$, *y* $\Delta_S^2(\varphi(\mathcal{X}) \,|\, P) \leq \Delta_S^2(\mathcal{Y} \,|\, P)$, *alcanzándose la igualdad si, y sólo si,* $f_{\mathcal{Y}(\omega)}(\alpha, \lambda) = f_{\varphi(\mathcal{X}(\omega))}(\alpha, \lambda)$ *c.s.* $[m \otimes S \otimes P]$.

Demostración:

En efecto, ya que de acuerdo con los resultados del Teorema 0.3.1 en relación con el cálculo de esperanzas iteradas de variables aleatorias difusas respecto al orden de integración, se tiene que:

$$\widetilde{E}(\mathcal{Y} \,|\, P) = \widetilde{E}\left(\widetilde{E}\,(\mathcal{Y}_{\tilde{x}}^* \,|\, P_{\tilde{x}}) \,\Big|\, P_{\mathcal{X}}\right).$$

Si $\varphi(\tilde{x}) = \widetilde{E}\,(\mathcal{Y}_{\tilde{x}}^* \,|\, P_{\tilde{x}})$, entonces:

$$\widetilde{V} = \widetilde{E}(\mathcal{Y} \,|\, P) = \widetilde{E}\left(\widetilde{E}\,(\mathcal{Y}_{\tilde{x}}^* \,|\, P_{\tilde{x}}) \,\Big|\, P_{\mathcal{X}}\right) = \widetilde{E}(\varphi(\mathcal{X}) \,|\, P).$$

Por otro lado:

$$\Delta_S^2(\mathcal{Y} \,|\, P) = \int_\Omega \int_{[0,1]} \int_{[0,1]} \left[f_{\mathcal{Y}(\omega)}(\alpha, \lambda) - f_{\widetilde{V}}(\alpha, \lambda)\right]^2 d\, S(\lambda)\, d\, \alpha\, d\, P(\omega).$$

Por el Teorema de Fubini , se cumple que:

$$\Delta_S^2(\mathcal{Y} \,|\, P) = \int_{[0,1]} \int_{[0,1]} \left(\int_\Omega \left[f_{\mathcal{Y}(\omega)}(\alpha, \lambda) - f_{\widetilde{V}}(\alpha, \lambda)\right]^2 d\, P(\omega)\right) d\, S(\lambda)\, d\, \alpha$$

$$= \int_{[0,1]} \int_{[0,1]} \left(\int_\Omega \left[f_{\mathcal{Y}(\omega)}(\alpha, \lambda) - f_{\varphi(\tilde{x})}(\alpha, \lambda)\right]^2 d\, P(\omega)\right) d\, S(\lambda)\, d\, \alpha$$

$$+\int_{[0,1]}\int_{[0,1]}\left(\int_{\Omega}\left[f_{\varphi(\tilde{x})}(\alpha,\lambda)-f_{\widetilde{V}}(\alpha,\lambda)\right]^{2}\,d\,P(\omega)\right)\,d\,S(\lambda)\,d\,\alpha$$

$$+\,2\int_{[0,1]}\int_{[0,1]}\left(\int_{\Omega}\left[f_{\mathcal{Y}(\omega)}(\alpha,\lambda)-f_{\varphi(\tilde{x})}(\alpha,\lambda)\right]\right.$$

$$\left.\cdot\left[f_{\varphi(\tilde{x})}(\alpha,\lambda)-f_{\widetilde{V}}(\alpha,\lambda)\right]\,d\,P(\omega)\right)\,d\,S(\lambda)\,d\,\alpha.$$

Como $\varphi(\tilde{x})=\widetilde{E}\left(\mathcal{Y}^{*}_{\tilde{x}}\,|\,P_{\tilde{x}}\right)$, se tiene entonces que para todo $\alpha\in[0,1]$ y $\lambda\in[0,1]$:

$$\int_{\Omega}\left[f_{\mathcal{Y}(\omega)}(\alpha,\lambda)-f_{\varphi(\tilde{x})}(\alpha,\lambda)\right]\left[f_{\varphi(\tilde{x})}(\alpha,\lambda)-f_{\widetilde{V}}(\alpha,\lambda)\right]\,d\,P(\omega)$$

$$=\int_{\mathcal{X}(\Omega)}\left[f_{\varphi(\tilde{x})}(\alpha,\lambda)-f_{\widetilde{V}}(\alpha,\lambda)\right]$$

$$\left(\int_{\mathcal{Y}(\Omega)}\left[f_{\tilde{y}}(\alpha,\lambda)-f_{\varphi(\tilde{x})}(\alpha,\lambda)\right]\,d\,P_{\tilde{x}}(\tilde{y})\right)\,d\,P_{\mathcal{X}}(\tilde{x}),$$

con lo que para todo $\alpha\in[0,1]$ y $\lambda\in[0,1]$, se obtiene que:

$$\int_{\mathcal{Y}(\Omega)}\left[f_{\tilde{y}}(\alpha,\lambda)-f_{\varphi(\tilde{x})}(\alpha,\lambda)\right]\,d\,P_{\tilde{x}}(\tilde{y})$$

$$=\int_{\mathcal{Y}(\Omega)}f_{\tilde{y}}(\alpha,\lambda)\,d\,P_{\tilde{x}}(\tilde{y})-f_{\varphi(\tilde{x})}(\alpha,\lambda)$$

$$=\int_{\mathcal{Y}(\Omega)}\left[\lambda\sup\tilde{y}_{\alpha}+(1-\lambda)\inf\tilde{y}_{\alpha}\right]\,d\,P_{\tilde{x}}(\tilde{y})$$

$$-\left[\lambda\sup\left(\widetilde{E}\left(\mathcal{Y}^{*}_{\tilde{x}}\,|\,P_{\tilde{x}}\right)\right)_{\alpha}+(1-\lambda)\inf\left(\widetilde{E}\left(\mathcal{Y}^{*}_{\tilde{x}}\,|\,P_{\tilde{x}}\right)\right)_{\alpha}\right].$$

Como:

$$\left(\widetilde{E}\left(\mathcal{Y}^{*}_{\tilde{x}}\,|\,P_{\tilde{x}}\right)\right)_{\alpha}=\left[\int_{\mathcal{Y}(\Omega)}\inf\tilde{y}_{\alpha}\,d\,P_{\tilde{x}}(\tilde{y}),\int_{\mathcal{Y}(\Omega)}\sup\tilde{y}_{\alpha}\,d\,P_{\tilde{x}}(\tilde{y})\right],$$

entonces para todo $\alpha\in[0,1]$ y $\lambda\in[0,1]$ se concluye que:

$$\int_{\mathcal{Y}(\Omega)}\left[f_{\tilde{y}}(\alpha,\lambda)-f_{\varphi(\tilde{x})}(\alpha,\lambda)\right]\,d\,P_{\tilde{x}}(\tilde{y})=0$$

cualquiera que sea $\tilde{x}\in\boldsymbol{\mathcal{X}}(\Omega)$.

Por lo tanto:

$$\Delta_S^2(\mathcal{Y}\,|\,P) = \int_{[0,1]}\int_{[0,1]}\left(\int_\Omega \left[f_{\mathcal{Y}(\omega)}(\alpha,\lambda) - f_{\varphi(\tilde{x})}(\alpha,\lambda)\right]^2 \, d\,P(\omega)\right) d\,S(\lambda)\, d\,\alpha$$

$$+\int_{[0,1]}\int_{[0,1]}\left(\int_\Omega \left[f_{\varphi(\tilde{x})}(\alpha,\lambda) - f_{\widetilde{V}}(\alpha,\lambda)\right]^2 \, d\,P(\omega)\right) d\,S(\lambda)\, d\,\alpha.$$

Además:

$$\Delta_S^2(\varphi(\mathcal{X})\,|\,P) = \int_{[0,1]}\int_{[0,1]}\left(\int_\Omega \left[f_{\varphi(\tilde{x})}(\alpha,\lambda) - f_{\widetilde{V}}(\alpha,\lambda)\right]^2 \, d\,P(\omega)\right) d\,S(\lambda)\, d\,\alpha,$$

por lo que:

$$\Delta_S^2(\mathcal{Y}\,|\,P) \geq \Delta_S^2(\varphi(\mathcal{X})\,|\,P).$$

Por otro lado, $\Delta_S^2(\mathcal{Y}\,|\,P) = \Delta_S^2(\varphi(\mathcal{X})\,|\,P)$ si, y sólo si, $f_{\mathcal{Y}(\omega)}(\alpha,\lambda) = f_{\varphi(\mathcal{X}(\omega))}(\alpha,\lambda)$ c.s. $[m \otimes S \otimes P]$. □

En los ejemplos siguientes se ilustra el cálculo de la S-DCM y la S-DCM central, así como su empleo para la comparación de ciertas poblaciones.

Ejemplo 1.2.1 . Cuando se quiere calificar un día de acuerdo con su temperatura, es habitual recurrir a adjetivos como FRÍO, FRESCO, NORMAL, CÁLIDO o CALUROSO. Según esa calificación, el tipo de día podría considerarse como una variable aleatoria difusa, $\mathcal{X}$, cuyos valores serían los cinco adjetivos precedentes, y una posible aproximación de los mismos a lo largo del verano podría llevarse a cabo por medio de números difusos con soportes contenidos en $[8,40]$ (en $^\circ$C) basados en S- y Π-curvas (ver, por ejemplo, Cox 1994), como los siguientes: FRÍO $= S(14,19)$, FRESCO $= \Pi(14,19,24)$, NORMAL $= \Pi(19,24,29)$, CÁLIDO $= \Pi(24,29,34)$ y CALUROSO $= 1 - S(29,34)$ (ver Figura 1.1), donde:

$$S(a,b)(t) = \begin{cases} 0 & \text{si } t \leq a \\ 2\left(\dfrac{t-a}{b-a}\right)^2 & \text{si } t \in \left[a, \dfrac{a+b}{2}\right] \\ 1 - 2\left(\dfrac{t-b}{b-a}\right)^2 & \text{si } t \in \left[\dfrac{a+b}{2}, b\right] \\ 1 & \text{en el resto,} \end{cases}$$

$$\Pi(a,(a+b)/2,b)(t) = \begin{cases} S(a,(a+b)/2) & \text{si } t \leq (a+b)/2 \\ 1 - S((a+b)/2,a) & \text{en el resto,} \end{cases}$$

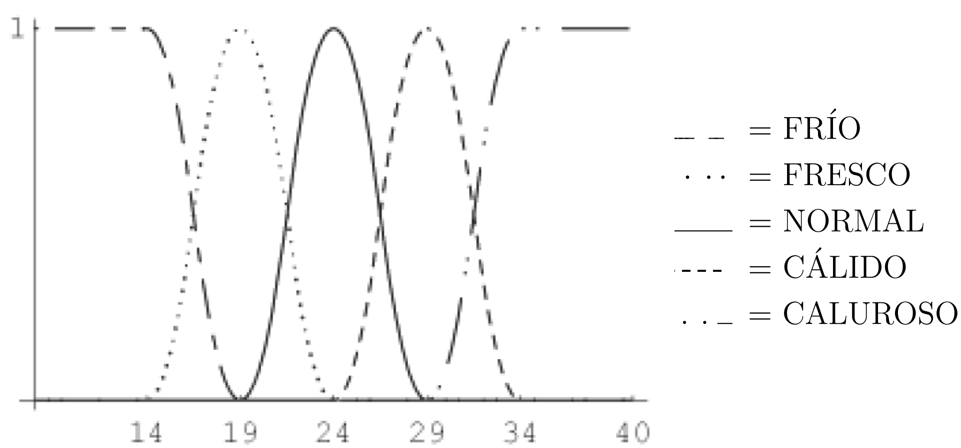


Fig. 1.1: Valores difusos del tipo de día

En este primer ejemplo se analiza la S-dispersión cuadrática media asociada a $\mathcal{X}$ para la medida m de Lebesgue en $[0,1]$, de forma que $L = 0$, y $\overline{g}(\lambda) = 1$ si $\lambda \in [0,1]$. La población considerada es la de los 31 días del mes de Julio de cierto año y en cierta zona (Ω_1), supuesto que a lo largo de ese mes 2 días han sido FRÍOS, 4 han sido FRESCOS, 7 NORMALES, 17 CÁLIDOS y 1 CALUROSO. Se desea estudiar la m-dispersión cuadrática media asociada a $\mathcal{X}$ con respecto a $\widetilde{A}_1 = \widetilde{E}(\mathcal{X})$, $\widetilde{A}_2 = 24$ y $\widetilde{A}_3 = \text{NORMAL}$.

Como se tiene que $\left[D_m\left(\text{FRÍO}, \widetilde{A}_1\right)\right]^2 = 67{,}084$, $\left[D_m\left(\text{FRESCO}, \widetilde{A}_1\right)\right]^2 = 47{,}549$, $\left[D_m\left(\text{NORMAL}, \widetilde{A}_1\right)\right]^2 = 3{,}598$, $\left[D_m\left(\text{CÁLIDO}, \widetilde{A}_1\right)\right]^2 = 9{,}646$, y $\left[D_m\left(\text{CALUROSO}, \widetilde{A}_1\right)\right]^2 = 19{,}738$. Entonces:

$$\Delta_m^2(\mathcal{X}) = DCM_m\left(\mathcal{X}, \widetilde{A}_1\right) = 17,202.$$

De forma análoga pueden hallarse las otras dos dispersiones:

$$DCM_m\left(\mathcal{X}, \widetilde{A}_2\right) = 22,995,\ DCM_m\left(\mathcal{X}, \widetilde{A}_3\right) = 20,8.$$

Ejemplo 1.2.2. En este segundo ejemplo se examina la comparación, a través de la m-dispersión cuadrática media central asociada a $\mathcal{X}$ de la distribución de los tipos de días en el mes de Julio del año y en la zona considerados en el Ejemplo 1.2.1 con una distribución hipotética "uniforme" en un mes de 30 días (población que denotaremos por Ω_2), de forma que 6 días de ese mes habrían sido FRÍOS, 6 FRESCOS, 6 NORMALES, 6 CÁLIDOS y 6 CALUROSOS.

En este caso, procediendo como en el Ejemplo 1.2.1, se obtendría que en Ω_2:

$$\Delta_m^2(\mathcal{X}) = 25,875,$$

de modo que $\mathcal{X}$ resultaría sensiblemente menos desviada respecto a su valor esperado en Ω_1 que en Ω_2.

Ejemplo 1.2.3. En este tercer ejemplo se estudia la comparación, a través de la m-dispersión cuadrática media central asociada a $\mathcal{X}$ de la distribución de los tipos de días en el mes de Julio del año y en la zona considerados en el Ejemplo 1.2.1 con la distribución de los tipos de días en el mes de Agosto del año y en la zona del Ejemplo 1.2.1 (población Ω_3), supuesto que a lo largo de ese Agosto ningún día ha sido FRÍO, 3 han sido FRESCOS, 6 NORMALES, 12 CÁLIDOS y 10 CALUROSOS.

Siguiendo un desarrollo similar al del Ejemplo 1.2.1, se concluiría que en Ω_3:

$$\Delta_m^2(\mathcal{X}) = 12,862,$$

de modo que $\mathcal{X}$ resultaría ligeramente más desviada respecto a su valor esperado en Julio que en Agosto.

1.3 Estimación insesgada de la medida de variación absoluta central en muestreos aleatorios de poblaciones finitas

En esta sección va a estudiarse el problema de la estimación de la medida generalizada de variación absoluta introducida en la Sección 1.1, en el muestreo aleatorio de poblaciones finitas.

Para ello, va a comprobarse que es posible construir un estimador insesgado de la medida generalizada en los muestreos simple y con reposición, y va a determinarse la precisión correspondiente a tal estimador en ambos muestreos.

Supongamos que se considera una población finita Ω de N unidades, $\omega_1, \ldots, \omega_N$, y una variable aleatoria difusa $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ asociada a un espacio medible definido sobre Ω y que se supone dotado con la distribución uniforme.

Si se elige una muestra de tamaño n al azar y sin reposición a partir de Ω, υ representa una muestra aleatoria simple genérica de tamaño n, y $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ son las unidades en υ, entonces la *S-dispersión cuadrática media central muestral* de $\mathcal{X}$ en υ viene dada por:

$$\Delta_S^2\left(\mathcal{X}[\upsilon]\right) = \frac{1}{n}\sum_{i=1}^{n}\left[D_S\left(\mathcal{X}(\omega_{\upsilon i}), \overline{\mathcal{X}}_n[\upsilon]\right)\right]^2,$$

donde $\overline{\mathcal{X}}_n[\upsilon] = \dfrac{1}{n} \odot [\mathcal{X}(\omega_{\upsilon 1}) \oplus \cdots \oplus \mathcal{X}(\omega_{\upsilon n})]$ es el valor esperado muestral de $\mathcal{X}$ en υ.

$\Delta_S^2\left(\mathcal{X}[\,\cdot\,]\right)$ es una variable aleatoria real asociada al espacio probabilístico $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ (siendo Υ_n el espacio de las $C_{N,n} = \binom{N}{n}$ posibles muestras aleatorias distintas sin reposición de tamaño n de la población dada, $\mathcal{P}(\Upsilon_n)$ el conjunto partes de Υ_n, y $p[\upsilon] = 1/C_{N,n}$ para todo $\upsilon \in \Upsilon_n$), y por lo tanto define un estimador de la *S-dispersión cuadrática media central poblacional*, que corresponde al valor:

$$\Delta_S^2(\mathcal{X} \mid P) = \frac{1}{N} \sum_{j=1}^{N} \left[D_S \left(\mathcal{X}(\omega_j), \overline{\mathcal{X}} \right) \right]^2 ,$$

con $\overline{\mathcal{X}} = \dfrac{1}{N} \odot [\mathcal{X}(\omega_1) \oplus \cdots \oplus \mathcal{X}(\omega_N)]$.

Para obtener a partir de la medida muestral un estimador insesgado de la poblacional, la primera debe corregirse de forma inmediata, teniendo en cuenta los resultados para variables aleatorias reales y la propiedad del Teorema 1.2.4. De este modo:

Teorema 1.3.1. *En el muestreo aleatorio simple de tamaño n a partir de Ω, se cumple que el estimador $\widehat{\Delta_S^2}(\mathcal{X}[\,\cdot\,])$ que a una muestra υ asocia el valor:*

$$\widehat{\Delta_S^2}(\mathcal{X}[\upsilon]) = \frac{(N-1)n}{N(n-1)} \Delta_S^2 \left(\mathcal{X}[\upsilon] \right),$$

es un estimador insesgado de $\Delta_S^2(\mathcal{X} \mid P)$.

Demostración:

En efecto, como $f_{\mathcal{X}}(\alpha, \lambda)$ es una variable aleatoria real para cada $\alpha \in [0,1]$ y $\lambda \in [0,1]$, entonces $\mathrm{Var}\left[f_{\mathcal{X}}(\alpha, \lambda) \right]$ puede estimarse insesgadamente en el muestreo aleatorio simple mediante el estadístico $\dfrac{N-1}{N} s_n^2 \left(f_{\mathcal{X}}(\alpha, \lambda) [\,\cdot\,] \right)$ que a la muestra υ le asocia el valor:

$$\frac{N-1}{N} s_n^2 \left(f_{\mathcal{X}}(\alpha, \lambda)[\upsilon] \right) = \frac{N-1}{N(n-1)} \sum_{i=1}^{n} \left[f_{\mathcal{X}(\omega_{\upsilon i})}(\alpha, \lambda) - \overline{f_{\mathcal{X}}(\alpha, \lambda)[\upsilon]} \right]^2 ,$$

donde:

$$\overline{f_{\mathcal{X}}(\alpha, \lambda)[\upsilon]} = \frac{1}{n} \sum_{i=1}^{n} f_{\mathcal{X}(\omega_{\upsilon i})}(\alpha, \lambda) = f_{\overline{\mathcal{X}}_n[\upsilon]}(\alpha, \lambda) .$$

En consecuencia, y en virtud del Teorema 1.2.4, $\Delta_S^2(\mathcal{X} \mid P)$ podrá estimarse insesgadamente en ese muestreo mediante el estimador que a la muestra υ le asocia el valor:

$$\widehat{\Delta_S^2}\left(\mathcal{X}[\upsilon] \right) = \int_{[0,1]} \int_{[0,1]} \frac{N-1}{N} s_n^2 \left(f_{\mathcal{X}}(\alpha, \lambda)[\upsilon] \right) \, d\,S(\lambda) \, d\,\alpha$$

$$= \frac{(N-1)n}{N(n-1)} \frac{1}{n} \sum_{i=1}^{n} \int_{[0,1]} \int_{[0,1]} \left[f_{\mathcal{X}(\omega_{vi})}(\alpha,\lambda) - f_{\overline{\mathcal{X}}_{n}[v]}(\alpha,\lambda) \right]^2 \, d\,S(\lambda)\, d\,\alpha$$

$$= \frac{(N-1)n}{N(n-1)} \Delta_S^2 \left(\mathcal{X}[v] \right).$$

□

Para establecer la precisión del estimador $\widehat{\Delta_S^2}(\mathcal{X}[\,\cdot\,])$ anterior, se determina a continuación el error cuadrático medio del mismo, que coincide con su varianza.

Teorema 1.3.2. *En el muestreo aleatorio simple de tamaño n a partir de Ω, se cumple que (si $f = n/N$ representa la fracción muestral):*

$$\mathrm{Var}\left(\widehat{\Delta_S^2}(\mathcal{X}[\,\cdot\,])\right) = \frac{(1-f)}{n(n-1)N^2(N-2)(N-3)}$$

$$\cdot \Bigg\{ 2(N-n-1)(N-1) \sum_{j=1}^{N} \sum_{l=1}^{N} \Bigg(\int_{[0,1]} \int_{[0,1]} \left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda) \right]$$

$$\cdot \left[f_{\mathcal{X}(\omega_l)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda) \right] \, d\,S(\lambda)\, d\,\alpha \Bigg)^2$$

$$-[4(n+2)N-(6n-1)]N \, \left(\Delta_S^2(\mathcal{X} \mid P) \right)^2$$

$$+[(n-1)N^3 - 2N^2 - 3(n-3)N + (6n-8)]N \; \mathrm{Var}\left(\left[D_S(\mathcal{X}, \overline{\mathcal{X}}) \right]^2 \Big| \, P \right) \Bigg\}.$$

Demostración:

En efecto:

$$\mathrm{Var}\left(\widehat{\Delta_S^2}(\mathcal{X}[\,\cdot\,])\right) = \frac{(N-1)^2 n^2}{N^2(n-1)^2}$$

$$\cdot \,\mathrm{Var}\left(\frac{1}{n} \sum_{j=1}^{N} \int_{[0,1]} \int_{[0,1]} \left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda) - f_{\overline{\mathcal{X}_n}}(\alpha,\lambda) \right]^2 a_j \, d\,S(\lambda)\, d\,\alpha \right)$$

$$= \frac{(N-1)^2}{N^2(n-1)^2} \mathrm{Var} \Bigg(\sum_{j=1}^{N} \int_{[0,1]} \int_{[0,1]} \Big\{ \left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda) \right]^2 a_j$$

$$+ \left[f_{\overline{\mathcal{X}}}(\alpha,\lambda) - f_{\overline{\mathcal{X}}_n}(\alpha,\lambda) \right]^2 a_j - 2 \left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda) \right]$$

$$\cdot \left[f_{\overline{\mathcal{X}}}\,(\alpha,\lambda) - f_{\overline{\mathcal{X}}_n}\,(\alpha,\lambda)\right]\, a_j \Big\}\, d\,S(\lambda)\, d\,\alpha\Big)$$

$$= \frac{(N-1)^2}{N^2(n-1)^2}\,\mathrm{Var}\left(\sum_{j=1}^{N}\left[D_S\left(\mathcal{X}(\omega_j),\overline{\mathcal{X}}\right)\right]^2\, a_j - n\left[D_S\left(\overline{\mathcal{X}}_n,\overline{\mathcal{X}}\right)\right]^2\right),$$

donde a_j es la variable aleatoria de Bernoulli asociada a la presencia de la unidad ω_j en la muestra y definida sobre el espacio probabilístico (Υ_n, $\mathcal{P}(\Upsilon_n)$, p).

Si se emplea la notación:

$$\mathrm{Var}\left(\widehat{\Delta^2_S}(\mathcal{X}[\,\cdot\,])\right) = \frac{(N-1)^2}{N^2(n-1)^2}\,\mathrm{Var}\left(\sum_{j=1}^{N}\mathrm{W}_{jj}\, a_j - \frac{1}{n}\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jl}\, a_j a_l\right),$$

donde:

$$\mathrm{W}_{jl} = \mathrm{W}_{lj} = \int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{X}(\omega_j)}\,(\alpha,\lambda) - f_{\overline{\mathcal{X}}}\,(\alpha,\lambda)\right]$$

$$\cdot\left[f_{\mathcal{X}(\omega_l)}\,(\alpha,\lambda) - f_{\overline{\mathcal{X}}}\,(\alpha,\lambda)\right]\, d\,S(\lambda)\, d\,\alpha,$$

para $j, l \in \{1, \ldots, N\}$, se tiene que:

$$\mathrm{Var}\left(\widehat{\Delta^2_S}(\mathcal{X}[\,\cdot\,])\right) = \frac{(N-1)^2}{N^2(n-1)^2}\left[\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jj}\mathrm{W}_{ll}\ \mathrm{Cov}(a_j\, a_l)\right.$$

$$+\frac{1}{n^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N}\sum_{l'=1}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll'}\ \mathrm{Cov}(a_j a_{j'},\, a_l a_{l'})$$

$$\left.-\frac{2}{n}\sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll}\,\mathrm{Cov}(a_l, a_j\, a_{j'})\right]$$

$$= \frac{(N-1)^2}{N^2(n-1)^2}\left[\sum_{j=1}^{N}\mathrm{W}^2_{jj}\,\mathrm{Var}(a_j) + \sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jj}\mathrm{W}_{ll}\,\mathrm{Cov}(a_j, a_l)\right.$$

$$+\frac{1}{n^2}\left\{\sum_{j=1}^{N}\mathrm{W}^2_{jj}\ \mathrm{Var}(a_j^2) + 4\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jl}\mathrm{W}_{jj}\ \mathrm{Cov}(a_j\, a_l, a_j^2)\right.$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{W}_{jj}\mathrm{W}_{ll}\ \mathrm{Cov}(a_j^2,\ a_l^2)+\ 2\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{W}_{jl}^2\ \mathrm{Var}(a_j\,a_l)$$

$$+\,2\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll}\ \mathrm{Cov}(a_j\,a_{j'},a_l^2)$$

$$+\,4\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{jl}\ \mathrm{Cov}(a_j\,a_{j'},a_j\,a_l)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\sum_{\substack{l'=1\\l'\neq l,j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll'}\ \mathrm{Cov}(a_j\,a_{j'},a_l\,a_{l'})\Bigg\}$$

$$-\frac{2}{n}\Bigg\{\sum_{j=1}^{N}\mathrm{W}_{jj}^2\ \mathrm{Cov}(a_j,a_j^2)+2\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{W}_{jl}\mathrm{W}_{jj}\ \mathrm{Cov}(a_j,a_j\,a_l)$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{W}_{jj}\mathrm{W}_{ll}\ \mathrm{Cov}(a_l,a_j^2)+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll}\ \mathrm{Cov}(a_l,a_j\,a_{j'})\Bigg\}\Bigg].$$

Como se cumple para la distribución p que:

$$\mathrm{Var}(a_j)=\mathrm{Var}(a_j^2)=\mathrm{Cov}(a_j,a_j^2)=(1-f)\frac{n}{N}\ \text{ para }\ j\in\{1,\,\ldots,\,N\},$$

$$\mathrm{Var}(a_j\,a_l)=\left(1-\frac{n(n-1)}{N(N-1)}\right)\frac{n(n-1)}{N(N-1)},$$

$$\mathrm{Cov}(a_j,a_ja_l)=\mathrm{Cov}(a_j\,a_l,a_j^2)=(1-f)\frac{n(n-1)}{N(N-1)},$$

$$\mathrm{Cov}(a_l,a_j^2)=\mathrm{Cov}(a_j^2,a_l^2)=\mathrm{Cov}(a_j,a_l)=-(1-f)\frac{n}{N(N-1)}$$

$$\text{para }\ j\in\{1,\,\ldots,\,N\},\,l\in\{1,\,\ldots,\,N\}\setminus\{j\},$$

$$\mathrm{Cov}(a_j\,a_{j'},a_j\,a_l)=\frac{(1-f)n(n-1)}{N(N-1)^2(N-2)}[(n-1)(N-2)-N],$$

$$\mathrm{Cov}(a_l,a_j\,a_{j'})=\mathrm{Cov}(a_j\,a_{j'},a_l^2)=\frac{-2(1-f)n(n-1)}{N(N-1)(N-2)}$$

para $j \in \{1, \ldots, N\}$, $j' \in \{1, \ldots, N\} \setminus \{j\}$, $l \in \{1, \ldots, N\} \setminus \{j, j'\}$, y

$$\mathrm{Cov}(a_j\, a_{j'}, a_l\, a_{l'}) = \frac{(1-f)n(n-1)}{N(N-1)^2(N-2)(N-3)}[N(6-4n)+6(n-1)]$$

para $j \in \{1, \ldots, N\}$, $j' \in \{1, \ldots, N\} \setminus \{j\}$, $l \in \{1, \ldots, N\} \setminus \{j, j'\}$, $l' \in \{1, \ldots, N\} \setminus \{j, j', l\}$, se tiene que:

$$\mathrm{Var}\left(\widehat{\Delta^2_S}(\mathcal{X}[\,\cdot\,])\right) = \frac{(1-f)(N-1)^2}{nN^3}\left\{\sum_{j=1}^N \mathrm{W}_{jj}^2\right.$$

$$-\frac{1}{N-1}\sum_{j=1}^N\sum_{\substack{l=1\\l\neq j}}^N \mathrm{W}_{jj}\mathrm{W}_{ll} - \frac{4}{N-1}\sum_{j=1}^N\sum_{\substack{l=1\\l\neq j}}^N \mathrm{W}_{jl}\mathrm{W}_{jj}$$

$$+\frac{2(N+n-1)}{(n-1)(N-1)^2}\sum_{j=1}^N\sum_{\substack{l=1\\l\neq j}}^N \mathrm{W}_{jl}^2 + \frac{4}{(N-1)(N-2)}\sum_{j=1}^N\sum_{\substack{j'=1\\j'\neq j}}^N\sum_{\substack{l=1\\l\neq j,j'}}^N \mathrm{W}_{jj'}\mathrm{W}_{ll}$$

$$+\frac{4}{(n-1)(N-1)^2(N-2)}\sum_{j=1}^N\sum_{\substack{j'=1\\j'\neq j}}^N\sum_{\substack{l=1\\l\neq j,j'}}^N \mathrm{W}_{jj'}\mathrm{W}_{jl}$$

$$\left.+\frac{[N(6-4n)+6(n-1)]}{(n-1)(N-1)^2(N-2)(N-3)}\sum_{j=1}^N\sum_{\substack{j'=1\\j'\neq j}}^N\sum_{\substack{l=1\\l\neq j,j'}}^N\sum_{\substack{l'=1\\l'\neq l,j,j'}}^N \mathrm{W}_{jj'}\mathrm{W}_{ll'}\right\}$$

$$= \frac{(1-f)}{n(n-1)N^2(N-2)(N-3)}\left\{\frac{[N(6-4n)+6(n-1)]}{N}\left(\sum_{j=1}^N\sum_{l=1}^N \mathrm{W}_{jl}\right)^2\right.$$

$$+4(n-2)(N-1)\sum_{j=1}^N\left(\sum_{l=1}^N \mathrm{W}_{jl}\right)^2 + 4[(n-1)(N-1)-n]\sum_{j=1}^N\sum_{j'=1}^N\sum_{l=1}^N \mathrm{W}_{jj'}\mathrm{W}_{ll}$$

$$+2(N-n-1)(N-1)\sum_{j=1}^N\sum_{l=1}^N \mathrm{W}_{jl}^2 - (n-1)(N-1)^2\left(\sum_{j=1}^N \mathrm{W}_{jj}\right)^2$$

$$-4(N-1)[(n-1)(N-1)-n]\sum_{j=1}^N\sum_{l=1}^N \mathrm{W}_{jl}\mathrm{W}_{jj}$$

$$+[n(N-2)(N^2-3)-(N-1)(N^2+N-8)]\sum_{j=1}^{N}\mathrm{W}_{jj}^2\Big\},$$

y como $\sum_{l=1}^{N}\mathrm{W}_{jl} = 0$ para todo $j \in \{1,\ldots,N\}$, $\sum_{j=1}^{N}\mathrm{W}_{jj} = N\,\Delta_S^2(\mathcal{X}\,|\,P)$, y $\sum_{j=1}^{N}\mathrm{W}_{jj}^2 = N\,E\left(\left[D_S(\mathcal{X},\overline{\mathcal{X}})\right]^4 \,\Big|\, P\right)$, se obtiene entonces el resultado del presente teorema. □

Si en vez de adoptar una selección aleatoria sin reposición, se considera una selección aleatoria con reposición de n unidades de la población $\Omega = \{\omega_1,\ldots,\omega_N\}$, y υ representa una muestra aleatoria con reposición genérica de tamaño n, y $\omega_{\upsilon 1},\ldots,\omega_{\upsilon n}$ son las unidades en υ, la S-*dispersión cuadrática media central muestral* de $\mathcal{X}$ en υ viene dada ahora por:

$$\Delta_S^2\left(\mathcal{X}[\upsilon]\right) = \frac{1}{n}\sum_{i=1}^{n}\left[D_S\left(\mathcal{X}(\omega_{\upsilon i}),\overline{\mathcal{X}}_n[\upsilon]\right)\right]^2$$

$$= \frac{1}{n}\sum_{j=1}^{N}\left[D_S\left(\mathcal{X}(\omega_j),\overline{\mathcal{X}}_n[\upsilon]\right)\right]^2 t_j[\upsilon] = \frac{1}{n}\Bigg\{\sum_{j=1}^{N}\left[D_S\left(\mathcal{X}(\omega_j),\overline{\mathcal{X}}\right)\right]^2 t_j[\upsilon]$$

$$-\frac{1}{n}\sum_{j=1}^{N}\sum_{l=1}^{N}\Bigg(\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda)\right]$$

$$\cdot\left[f_{\mathcal{X}(\omega_l)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda)\right]\,d\,S(\lambda)\,d\,\alpha\Bigg)t_j[\upsilon]t_l[\upsilon]\Bigg\},$$

donde t_j es la variable aleatoria definida sobre el espacio probabilístico (Υ_n^w, $\mathcal{P}(\Upsilon_n^w), p^w$) (siendo Υ_n^w el espacio de las $CR_{N,n} = \binom{N+n-1}{n}$ posibles muestras aleatorias con reposición distintas de tamaño n de la población dada, y $p^w[\upsilon]$ la probabilidad de elegir la muestra $\upsilon \in \Upsilon_n^w$, que no determina una distribución uniforme en Υ_n^w para el muestreo considerado), de manera que $t_j[\upsilon]$ representa el "número de veces que ω_j aparece en la muestra υ".

$\Delta_S^2\left(\mathcal{X}[\,\cdot\,]\right)$ es una variable aleatoria real asociada al espacio probabilístico $(\Upsilon_n^w,\mathcal{P}(\Upsilon_n^w),p^w)$, y por lo tanto define un estimador de la S-DCM central poblacional.

Al igual que en el muestreo aleatorio simple, para construir un estimador insesgado de $\Delta_S^2(\mathcal{X} \mid P)$ a partir del muestral, basta con corregir este último como sigue:

Teorema 1.3.3. *En el muestreo aleatorio con reposición de tamaño n a partir de Ω, se cumple que el estimador $\widehat{\Delta_S^2}^w(\mathcal{X}[\,\cdot\,])$ que a una muestra υ asocia el valor:*

$$\widehat{\Delta_S^2}^w(\mathcal{X}[\upsilon]) = \frac{n}{(n-1)}\Delta_S^2(\mathcal{X}[\upsilon]),$$

es un estimador insesgado de $\Delta_S^2(\mathcal{X} \mid P)$.

Demostración:

En efecto, razonando como en el Teorema 1.3.1, $\Delta_S^2(\mathcal{X})$ puede estimarse insesgadamente en el muestreo aleatorio con reposición mediante el estadístico:

$$\widehat{\Delta_S^2}^w(\mathcal{X}[\upsilon]) = \int_{[0,1]}\int_{[0,1]} s_n^2\left(f_{\mathcal{X}}(\alpha,\lambda)[\upsilon]\right)\, d\,S(\lambda)\, d\,\alpha = \frac{n}{n-1}\Delta_S^2(\mathcal{X}[\upsilon]).$$

□

El error cuadrático medio asociado de $\widehat{\Delta_S^2}^w(\mathcal{X}[\,\cdot\,])$ vendría dado en este muestreo por:

Teorema 1.3.4. *En el muestreo aleatorio con reposición de tamaño n a partir de Ω, se cumple que:*

$$\operatorname{Var}\left(\widehat{\Delta_S^2}^w(\mathcal{X}[\,\cdot\,])\right) = \frac{1}{n(n-1)^2N^3}$$

$$\cdot\Bigg\{2(n-1)(N-2)\sum_{j=1}^{N}\sum_{l=1}^{N}\Bigg(\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda)\right]$$

$$\cdot\left[f_{\mathcal{X}(\omega_l)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda)\right]\, d\,S(\lambda)\, d\,\alpha\Bigg)^2$$

$$+2(nN+n^2-2)N^2\left(\Delta_S^2(\mathcal{X} \mid P)\right)^2$$

$$+[(n-1)^2N - 2(n^2-5n+5)]N^2\ \operatorname{Var}\left(\left[D_S(\mathcal{X},\overline{\mathcal{X}})\right]^2 \,\middle|\, P\right)\Bigg\}.$$

Demostración:

En efecto, utilizando la notación del Teorema 1.3.2:

$$\operatorname{Var}\left(\widehat{\Delta_S^2}^{w}(\mathcal{X}[\,\cdot\,])\right) = \frac{1}{(n-1)^2}\operatorname{Var}\left(\sum_{j=1}^{N} \mathrm{W}_{jj}t_j - \frac{1}{n}\sum_{j=1}^{N}\sum_{l=1}^{N} \mathrm{W}_{jl}t_jt_l\right)$$

$$= \frac{1}{(n-1)^2}\Bigg[\sum_{j=1}^{N} \mathrm{W}_{jj}^2 \operatorname{Var}(t_j) + \sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{W}_{jj}\mathrm{W}_{ll}\operatorname{Cov}(t_j,t_l)$$

$$+\frac{1}{n^2}\Bigg\{\sum_{j=1}^{N} \mathrm{W}_{jj}^2 \operatorname{Var}(t_j^2) + 4\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{W}_{jl}\mathrm{W}_{jj}\operatorname{Cov}(t_j\,t_l, t_j^2)$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{W}_{jj}\mathrm{W}_{ll}\operatorname{Cov}(t_j^2,\, t_l^2) + 2\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{W}_{jl}^2\operatorname{Var}(t_j\,t_l)$$

$$+2\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N} \mathrm{W}_{jj'}\mathrm{W}_{ll}\operatorname{Cov}(t_j\,t_{j'}, t_l^2)$$

$$+4\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N} \mathrm{W}_{jj'}\mathrm{W}_{jl}\operatorname{Cov}(t_j\,t_{j'}, t_j\,t_l)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\sum_{\substack{l'=1\\ l'\neq l,j,j'}}^{N} \mathrm{W}_{jj'}\mathrm{W}_{ll'}\operatorname{Cov}(t_j\,t_{j'}, t_l\,t_{l'})\Bigg\}$$

$$-\frac{2}{n}\Bigg\{\sum_{j=1}^{N} \mathrm{W}_{jj}^2\operatorname{Cov}(t_j,\, t_j^2) + 2\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{W}_{jl}\mathrm{W}_{jj}\operatorname{Cov}(t_j, t_j\,t_l)$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{W}_{jj}\mathrm{W}_{ll}\operatorname{Cov}(t_l, t_j^2) + \sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N} \mathrm{W}_{jj'}\mathrm{W}_{ll}\operatorname{Cov}(t_j, t_j\,t_{j'})\Bigg\}\Bigg].$$

Como se cumple para la distribución p^w que:

$$\operatorname{Var}(t_j) = \frac{n(N-1)}{N^2},$$

$$\mathrm{Var}(t_j^2) = \frac{n}{N^4}\{N^2(N-1) + (n-1)[2N(3N+2n-6) + (6-4n)]\}$$

$$\mathrm{Cov}(t_j, t_j^2) = \frac{n(N-1)(N+2(n-1))}{N^3}$$

para $j \in \{1, \ldots, N\}$,

$$\mathrm{Var}(t_j\, t_l) = \frac{n(n-1)}{N^4}[N(N+2n-4) + (6-4n)],$$

$$\mathrm{Cov}(t_j, t_j\, t_l) = \frac{n(n-1)(N-2)}{N^3},$$

$$\mathrm{Cov}(t_j\, t_l, t_j^2) = \frac{n(n-1)}{N^4}[N(N+2n-6) + (6-4n)],$$

$$\mathrm{Cov}(t_j^2, t_l^2) = \frac{n[(n-1)(6-4n-4N) - N^2]}{N^4},$$

$$\mathrm{Cov}(t_l, t_j^2) = -\frac{2n(n-1)}{N^3},\ \mathrm{Cov}(t_j,\, t_l) = -\frac{n}{N^2}$$

para $j \in \{1, \ldots, N\},\ l \in \{1, \ldots, N\} \setminus \{j\}$,

$$\mathrm{Cov}(t_j\, t_{j'}, t_j\, t_l) = \frac{n(n-1)[N(n-2) + (6-4n)]}{N^4},$$

$$\mathrm{Cov}(t_l, t_j\, t_{j'}) = \frac{-2(n-1)}{N^3},$$

$$\mathrm{Cov}(t_j\, t_{j'}, t_l^2) = \frac{n(n-1)[(6-4n) - 2N]}{N^4}$$

para $j \in \{1, \ldots, N\},\ j' \in \{1, \ldots, N\} \setminus \{j\},\ l \in \{1, \ldots, N\} \setminus \{j, j'\}$, y

$$\mathrm{Cov}(t_j\, t_{j'}, t_l\, t_{l'}) = \frac{n(n-1)[(6-4n) - 2N]}{N^4}$$

para $j \in \{1, \ldots, N\},\ j' \in \{1, \ldots, N\} \setminus \{j\},\ l \in \{1, \ldots, N\} \setminus \{j, j'\}$, $l' \in \{1, \ldots, N\} \setminus \{j, j', l\}$, se tiene que:

$$\mathrm{Var}\left(\widehat{\Delta_S^2}^{w}(\mathcal{X}[\,\cdot\,])\right) = \frac{1}{n(n-1)^2N^4}\Bigg[(n-1)(N-1)[(n-1)N^2$$

$$+(6-4n)(N-1)]\sum_{j=1}^{N} \mathrm{W}_{jj}^2$$

$$+ [(n-1)[2N-(6-4n)(N-1)] - N^2(n^2+1)]\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{W}_{jj}\mathrm{W}_{ll}$$

$$-4(N-1)[N^2(n-1)+(6-4n)(N-1)]\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{W}_{jl}\mathrm{W}_{jj}$$

$$+2(n-1)[N(N+2n-4)+(6-4n)]\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{W}_{jl}^2$$

$$+2(n-1)(6-4n)\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll}$$

$$+4(n-1)[N(n-2)+(6-4n)]\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{jl}$$

$$+(n-1)[(6-4n)-2N]\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\sum_{\substack{l'=1\\l'\neq l,j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll'}\Bigg]$$

$$= \frac{1}{n(n-1)N^3}\Bigg\{\frac{(6-4n)-2N}{N}\left(\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jl}\right)^2$$

$$+4n\sum_{j=1}^{N}\left(\sum_{l=1}^{N}\mathrm{W}_{jl}\right)^2 + 4\sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll}$$

$$+2(N-2)\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jl}^2 - \frac{(6-4n)(n-1)+N(n^2+1)}{(n-1)}\left(\sum_{j=1}^{N}\mathrm{W}_{jj}\right)^2$$

$$-4[(N-2)(n+1)+6]\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jl}\mathrm{W}_{jj}$$

$$+\frac{N[N(n-1)^2-2(n^2-5n+5)]}{n-1}\sum_{j=1}^{N}\mathrm{W}_{jj}^2\Bigg\},$$

y de nuevo, al ser $\sum_{l=1}^{N} \mathrm{W}_{jl} = 0$ para todo $j \in \{1, \ldots, N\}$, $\sum_{j=1}^{N} \mathrm{W}_{jj} = N\, \Delta_S^2(\mathcal{X} \mid P)$, y $\sum_{j=1}^{N} \mathrm{W}_{jj}^2 = N\, E\left(\left[D_S(\mathcal{X}, \overline{\mathcal{X}})\right]^4 \,\middle|\, P\right)$, se obtiene el resultado del teorema. □

Los resultados que acaban de establecerse pueden emplearse con el fin de comparar la precisión en la estimación de la S-DCM central poblacional asociada a distintas variables, construir intervalos de confianza y procedimientos de contraste de hipótesis para la misma, y determinar tamaños muestrales adecuados en la estimación.

Sin embargo, el desconocimiento de la distribución exacta de los estimadores $\widehat{\Delta_S^2}$ y $\widehat{\Delta_S^2}^w$ obligaría a construir técnicas basadas en la desigualdad de Tchebychev o aproximaciones similares, lo que daría lugar a procedimientos excesivamente conservadores en la mayoría de las situaciones. Por otro lado, estos métodos involucrarían valores poblacionales desconocidos. Aunque en esta memoria no va a abordarse tal estudio, podría analizarse el efecto, sobre las inferencias derivadas de la aplicación de esos métodos, del reemplazo de esos valores poblacionales por estimaciones insesgadas de los mismos. Esta estimación se definiría fácilmente a partir de los valores de

$$E(a_j) = \frac{n}{N} \text{ para } j \in \{1, \ldots, N\},$$

$$E(a_j\, a_l) = \frac{n(n-1)}{N(N-1)} \text{ para } j \in \{1, \ldots, N\}, l \in \{1, \ldots, N\} \setminus \{j\}.$$

$$E(t_j) = \frac{n}{N} \text{ para } j \in \{1, \ldots, N\},$$

$$E(t_j\, t_l) = \frac{n(n-1)}{N^2} \text{ para } j \in \{1, \ldots, N\},\ l \in \{1, \ldots, N\} \setminus \{j\}.$$

Algunos de los resultados principales de la Sección 1.3 van a servir de apoyo a otros que se presentan en la Sección 1.5, en la que se ilustrará su aplicación.

1.4 Distribución asintótica de la S-dispersión cuadrática media central muestral en poblaciones finitas. Aplicación al desarrollo de procedimientos inferenciales aproximados para el valor poblacional

En la sección anterior se ha verificado que la S-DCM central poblacional admite un estimador insesgado en el muestreo de poblaciones finitas, aunque el desarrollo de otros procedimientos inferenciales exactos sobre ese valor poblacional no es viable y, como se ha señalado, las técnicas que podrían construirse a partir de los resultados de la Sección 1.3 y otros complementarios serían muy conservadoras. Sin embargo, el estudio que se va a llevar a cabo en esta sección hará posible la construcción de procedimientos inferenciales aproximados que proporcionarán conclusiones muy útiles.

Como implicaciones de ese estudio y de los procedimientos inferenciales que se derivan del mismo, cabe destacar su aplicación cuando se consideran grandes muestras o bien extraídas al azar y con reemplazamiento de una población cualquiera, o extraídas al azar y sin reemplazamiento de una población de tamaño sustancialmente más grande que el tamaño muestral. Bajo tales condiciones, las suposiciones de independencia e identidad de distribución para las variables aleatorias difusas que conforman la muestra aleatoria de gran tamaño son válidas exacta o aproximadamente, respectivamente.

Sea $\Omega = \{\omega_1, \ldots, \omega_N\}$ una población finita y sea $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ una variable aleatoria difusa. Sobre Ω se puede definir el espacio probabilístico $(\Omega, \mathcal{P}(\Omega), P)$, donde P representa la medida de probabilidad correspondiente a la distribución uniforme sobre Ω (es decir, a la selección aleatoria de los individuos de Ω).

Si sobre la población Ω la variable $\mathcal{X}$ toma r valores distintos, $\tilde{x}_1^*, \ldots, \tilde{x}_r^*$, y para $l \in \{1, \ldots, r\}$ se denota por p_l la probabilidad $P(\{\omega \in \Omega \mid \mathcal{X}(\omega) = \tilde{x}_l^*\})$, y S es una medida en $[0,1]$ que satisface las condiciones en la Definición 0.1.5, la *S-DCM central poblacional* asociada a $\mathcal{X}$ en Ω puede expresarse como:

$$\Delta_S^2(\mathcal{X} \mid \mathbf{p}) = \sum_{l=1}^{r} p_l \left[D_S\left(\tilde{x}_l^*, \widetilde{E}(\mathcal{X} \mid \mathbf{p})\right)\right]^2,$$

con $\widetilde{E}(\mathcal{X} \mid \mathbf{p}) = \bigodot\!\!\!\!\!\!\sum_{l=1}^{r} p_l \odot \tilde{x}_l^*$ y $\mathbf{p} = (p_1, \ldots, p_{r-1})$ $(p_r = 1 - \sum_{l=1}^{r-1} p_l)$.

Si se selecciona al azar una muestra de tamaño n de la población y f_{nl} representa la frecuencia relativa del valor $\tilde{x}_l^*$ de $\mathcal{X}$ en esa muestra, la *S-DCM central muestral* corresponde al valor:

$$\Delta_S^2(\mathcal{X} \mid \mathbf{f}_n) = \sum_{l=1}^{r} f_{nl} \left[D_S\left(\tilde{x}_l^*, \overline{\mathcal{X}}_n\right)\right]^2,$$

con $\overline{\mathcal{X}}_n = \bigodot\!\!\!\!\!\!\sum_{l=1}^{r} f_{nl} \odot \tilde{x}_l^*$, y $\mathbf{f}_n = (f_{n1}, \ldots, f_{n(r-1)})$ $(f_{nr} = 1 - \sum_{l=1}^{r-1} f_{nl})$.

El siguiente teorema recoge la distribución asintótica de la S-DCM central muestral, y en él se establece la normalidad asintótica y consistencia del mismo (de hecho, puede concluirse que la S-DCM central muestral es el mejor estimador asintóticamente normal -ver, por ejemplo, Zacks 1971, págs. 248-249-) en la estimación de $\Delta_S^2(\mathcal{X})$.

Teorema 1.4.1. *Para cada $n \in \mathbb{N}$, se consideran n variables aleatorias difusas independientes e idénticamente distribuidas que la variable aleatoria difusa $\mathcal{X}$ (es decir, una muestra aleatoria simple de tamaño n a partir de $\mathcal{X}$), definida sobre la población finita $\Omega = \{\omega_1, \ldots, \omega_N\}$ de forma que $P(\{\omega \in \Omega \mid \mathcal{X}(\omega) = \tilde{x}_l^*\}) = p_l$ con $p_l \in (0,1)$, $l = 1, \ldots, r$, $\sum_{l=1}^{r} p_l = 1$. Sea S una medida normalizada sobre* $[0,1]$, *que satisface las condiciones en la Definición* 0.1.5. *Se cumple entonces que:*

i) *Si* $\mathbf{f}_n = (f_{n1}, \ldots, f_{n(r-1)}) \in [0,1]^{r-1}$, *con* f_{nl} = *frecuencia relativa de* $\tilde{x}_l^*$ *en la correspondiente realización de la muestra aleatoria simple de tamaño* n $(l = 1, \ldots, r-1,\ f_{nr} = 1 - \sum_{l=1}^{r-1} f_{nl})$, *y* $\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n)$ *es la S-DCM central muestral correspondiente, entonces* $\{\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n)\}_n$ *es una sucesión de estimadores de* $\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})$, *que es fuertemente consistente, es decir, cuando* $n \to \infty$ *se tiene que:*

$$\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) \overset{c.s.}{\to} \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}),$$

cualquiera que sea $\mathbf{p} = (p_1, \ldots, p_{r-1})$ *(con* $p_1, \ldots, p_{r-1} \in (0,1)$ *y* $\sum_{l=1}^{r-1} p_l < 1$*).*

ii) $\{\sqrt{n}\,(\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}))\}_n$ *es una sucesión de variables aleatorias que converge en ley hacia una normal unidimensional* $N(0, \sigma^2(\mathbf{p}))$, *con:*

$$\sigma^2(\mathbf{p}) = \mathrm{Var}\left(\left[D_S\left(\mathcal{X}, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2\right),$$

siempre que $\sigma^2(\mathbf{p}) > 0$.

iii) *Si* $\sigma^2(\mathbf{p}) = 0$, *y para algún par* (i,j) *con* $i, j \in \{1, \ldots, r-1\}$ *se cumple que:*

$$h_{ij} = \frac{\partial^2 \Delta_S^2(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i \partial p_j}$$
$$= \frac{\partial}{\partial p_i}\left(\left[D_S\left(\tilde{x}_j^*, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2 - \left[D_S\left(\tilde{x}_r^*, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2\right) > 0,$$

entonces $\{2n\,(\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}))\}_n$ *es una sucesión de variables aleatorias que converge en ley hacia una combinación lineal de, a lo sumo,* $r-1$ *variables chi-cuadrado* χ_1^2 *independientes.*

Demostración:

Sobre la población finita considerada, $\mathcal{X}$ se "distribuye" dependiendo de un parámetro $\mathbf{p} = (p_1, \ldots, p_{r-1})$, con $p_1, \ldots, p_{r-1} \in (0,1)$ y $p_r = 1 - \sum_{l=1}^{r-1} p_l \in (0,1)$.

Se satisfacen, por tanto, las condiciones siguientes:

- El espacio paramétrico correspondiente es $\mathbb{P} = [0,1]^{r-1}$, y el verdadero valor de $\mathbf{p}$ pertenece al interior de $\mathbb{P}$, que es el espacio $(0,1)^{r-1}$. Además, como se cumple que $\mathbf{p} \in (0,1)^{r-1}$, para valores distintos de $\mathbf{p}$ las medidas de probabilidad correspondientes deben diferir.

- El conjunto de valores distintos de $\mathcal{X}$ en la población, $\{\tilde{x}_1^*, \ldots, \tilde{x}_r^*\}$, no depende de $\mathbf{p}$.

- Para cada valor $\mathbf{p} \in (0,1)^{r-1}$, se cumple que si $A_l = \{\omega \in \Omega \mid \mathcal{X}(\omega) = \tilde{x}_l^*\}$ y $P(A_l \,|\, \mathbf{p}) = p_l$ para cada $l \in \{1, \ldots, r\}$, entonces $\log P(A_l \,|\, \mathbf{p})$ es derivable parcialmente hasta el tercer orden con respecto a las componentes p_i de $\mathbf{p}$ y en un entorno de $\mathbf{p}$ se satisface además que
$$\sum_{l=1}^{r} p_l \left| \frac{\partial^3}{\partial p_i \partial p_j \partial p_k} \log P(A_l \,|\, \mathbf{p}) \right| < \infty$$
para cualesquiera $i, j, k \in \{1, \ldots, r-1\}$.

- La matriz de información de Fisher asociada a la familia $\wp$ (familia de las medidas de probabilidad correspondientes a los distintos valores de $\mathbf{p} \in \mathbb{P}$) para el valor $\mathbf{p}$ viene dada por:
$$\mathrm{I}_{\mathcal{X}}^F(\mathbf{p}) = \left[\mathrm{I}_{ij}^F(\mathbf{p})\right] = \left[\sum_{l=1}^{r} p_l \, \frac{\partial \log p_l}{\partial p_i} \cdot \frac{\partial \log p_l}{\partial p_j}\right]$$
$$= \left[\frac{\delta_{ij}}{p_i} + \frac{1}{p_r}\right] \quad \text{con } \delta_{ij} = \text{delta de Kronecker}$$
$(i, j \in \{1, \ldots, r-1\})$. $\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})$ está bien definida (ya que $\mathrm{I}_{ij}^F(\mathbf{p})$ es finito para cualquier par (i,j)) y es una matriz definida positiva (es decir, la forma cuadrática asociada es definida positiva) para cada $\mathbf{p} \in (0,1)^{r-1}$, ya que para cualquier $i \in \{1, \ldots, r-1\}$ se cumple que:
$$\begin{vmatrix} \mathrm{I}_{11}^F(\mathbf{p}) & \cdots & \mathrm{I}_{1i}^F(\mathbf{p}) \\ \vdots & \ddots & \vdots \\ \mathrm{I}_{i1}^F(\mathbf{p}) & \cdots & \mathrm{I}_{ii}^F(\mathbf{p}) \end{vmatrix} = \frac{p_1 + \cdots + p_i + p_r}{p_1 \cdots p_i \cdot p_r} > 0$$

(lo que garantiza -ver, por ejemplo, Rao 1973, pág. 36- que la matriz de información de Fisher es definida positiva).

Por otro lado, la sucesión $\{\mathbf{f}_n\}_n$ es una sucesión de estimadores de $\mathbf{p}$ para esa muestra aleatoria, que son solución del sistema de ecuaciones de verosimilitud, puesto que son los estimadores máximo-verosímiles de $\mathbf{p}$. La sucesión $\{\mathbf{f}_n\}_n$ es fuertemente consistente y tiene distribución asintótica normal (r-1)-dimensional $N\left(\mathbf{p}, \frac{\left[\mathrm{I}_{ij}^{F}(\mathbf{p})\right]^{-1}}{n}\right)$ cuando $n \to \infty$, con:

$$\left[\mathrm{I}_{ij}^{F}(\mathbf{p})\right]^{-1} = [p_i(\delta_{ij} - p_j)],$$

ya que $\left|\left[\mathrm{I}_{ij}^{F}(\mathbf{p})\right]\right| = \left(\prod_{l=1}^{r} p_l\right)^{-1}$ y, para $i, j \in \{1, \dots, r-1\}$, el adjunto de $\mathrm{I}_{ij}^{F}(\mathbf{p})$ es ${}^{\alpha}\mathrm{I}_{ij}^{F}(\mathbf{p}) = p_i(\delta_{ij} - p_j)\left(\prod_{l=1}^{r} p_l\right)^{-1}$.

Además, $n(\mathbf{f}_n - \mathbf{p})\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})(\mathbf{f}_n - \mathbf{p})^t$ converge en distribución a una chi-cuadrado χ^2_{r-1} cuando $n \to \infty$ (ver, por ejemplo, Rao 1973, págs. 359-363, Serfling 1980, págs. 144-153, Lehmann 1983, págs. 403-436).

A partir de estas condiciones van a probarse los resultados del presente teorema.

i) En las condiciones que se acaban de señalar se garantiza que $\mathbf{f}_n \overset{c.s.}{\to} \mathbf{p}$, por lo que de acuerdo con los resultados de la Teoría Asintótica en Inferencia Paramétrica (ver, por ejemplo, Serfling 1980, pág. 24) y al ser $\Delta_S^2(\mathcal{X} \mid \mathbf{p})$ continua en un entorno de $\mathbf{p}$, se concluye que:

$$\Delta_S^2(\mathcal{X} \mid \mathbf{f}_n) \overset{c.s.}{\to} \Delta_S^2(\mathcal{X} \mid \mathbf{p})$$

cuando $n \to \infty$, es decir, $\{\Delta_S^2(\mathcal{X} \mid \mathbf{f}_n)\}_n$ es una sucesión estimadora de $\Delta_S^2(\mathcal{X} \mid \mathbf{p})$ fuertemente consistente.

ii) Por las condiciones supuestas, para n suficientemente grande es posible desarrollar $\Delta_S^2(\mathcal{X} \mid \mathbf{f}_n)$ en un entorno de $\mathbf{p}$. De este modo, el desarrollo de Taylor de primer orden de $\Delta_S^2(\mathcal{X} \mid \mathbf{f}_n)$ será:

$$\Delta^2_S(\mathcal{X}\,|\,\mathbf{f}_n) = \Delta^2_S(\mathcal{X}\,|\,\mathbf{p}) + \nabla\Delta^2_S(\mathbf{p})(\mathbf{f}_n - \mathbf{p})^t$$
$$+\frac{1}{2}(\mathbf{f}_n - \mathbf{p})H\Big(\Delta^2_S(\mathbf{p}^*_n)\Big)\,(\mathbf{f}_n - \mathbf{p})^t,$$

con $\nabla\Delta^2_S(\mathbf{p})$ = vector gradiente de $\Delta^2_S(\mathcal{X}\,|\,\cdot)$ en $\mathbf{p}$, y $H(\Delta^2_S(\mathbf{p}))$ = matriz hessiana $(r-1)\times(r-1)$ dada por:

$$H\Big(\Delta^2_S(\mathbf{p}^*_n)\Big) = [h_{ij}] = \left[\left(\frac{\partial^2\Delta^2_S(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i \partial p_j}\right)_{\mathbf{p}=\mathbf{p}^*_n}\right],$$

y con $\mathbf{p}^*_n \in \mathbb{P}$ tal que $\|\mathbf{p}^*_n - \mathbf{p}\| \le \|\mathbf{f}_n - \mathbf{p}\|$.

Por lo tanto:

$$\sqrt{n}\left[\Delta^2_S(\mathcal{X}\,|\,\mathbf{f}_n) - \Delta^2_S(\mathcal{X}\,|\,\mathbf{p})\right]$$
$$= \nabla\Delta^2_S(\mathbf{p})\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t + \frac{1}{2}(\mathbf{f}_n - \mathbf{p})H\Big(\Delta^2_S(\mathbf{p}^*_n)\Big)\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t.$$

El vector gradiente vendrá dado por la matriz fila:

$$\nabla\Delta^2_S(\mathbf{p}) = \left(\frac{\partial}{\partial p_1}\Delta^2_S(\mathcal{X}\,|\,\mathbf{p}) \ \cdots \ \frac{\partial}{\partial p_{r-1}}\Delta^2_S(\mathcal{X}\,|\,\mathbf{p})\right),$$

y, para $i \in \{1, \ldots, r-1\}$ se cumple que:

$$\frac{\partial}{\partial p_i}\Delta^2_S(\mathcal{X}\,|\,\mathbf{p}) = \frac{\partial}{\partial p_i}\left[\sum_{l=1}^{r-1} p_l \left\{\left[D_S\left(\tilde{x}^*_l, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2\right.\right.$$
$$\left.\left.-\left[D_S\left(\tilde{x}^*_r, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2\right\} + \left[D_S\left(\tilde{x}^*_r, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2\right]$$
$$= \left[D_S\left(\tilde{x}^*_i, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2 - \left[D_S\left(\tilde{x}^*_r, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2$$
$$+\sum_{l=1}^{r-1} p_l \left\{\frac{\partial}{\partial p_i}\left[D_S\left(\tilde{x}^*_l, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2\right.$$
$$\left.-\frac{\partial}{\partial p_i}\left[D_S\left(\tilde{x}^*_r, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2\right\} + \frac{\partial}{\partial p_i}\left[D_S\left(\tilde{x}^*_r, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2.$$

Si $l \in \{1, \ldots, r\}$, se tiene que:

$$\frac{\partial}{\partial p_i} \left[D_S \left(\tilde{x}_l^*, \widetilde{E}(\mathcal{X} \,|\, \mathbf{p}) \right) \right]^2$$

$$= -2 \int_{[0,1]} \int_{[0,1]} \left[f_{\tilde{x}_l^*}(\alpha, \lambda) - f_{\widetilde{E}(\mathcal{X} \,|\, \mathbf{p})}(\alpha, \lambda) \right]$$

$$\cdot \left(\frac{\partial}{\partial p_i} f_{\widetilde{E}(\mathcal{X} \,|\, \mathbf{p})}(\alpha, \lambda) \right) d\, S(\lambda)\, d\, \alpha$$

$$= -2 \int_{[0,1]} \int_{[0,1]} \left[f_{\tilde{x}_l^*}(\alpha, \lambda) - f_{\widetilde{E}(\mathcal{X} \,|\, \mathbf{p})}(\alpha, \lambda) \right]$$

$$\cdot \left[f_{\tilde{x}_i^*}(\alpha, \lambda) - f_{\tilde{x}_r^*}(\alpha, \lambda) \right] d\, S(\lambda)\, d\, \alpha.$$

Como $\sqrt{n}(\mathbf{f}_n - \mathbf{p})$ se distribuye asintóticamente como una normal $(r-1)$-dimensional $N\Big(\mathbf{0},\, \left[\mathrm{I}\,{}^F_{\mathcal{X}}(\mathbf{p}) \right]^{-1} \Big)$, según las propiedades de la convergencia en ley (ver, por ejemplo, Serfling 1980, pág. 26), se tiene que cuando $n \to \infty$:

$$\nabla \Delta_S^2(\mathbf{p}) \left(\sqrt{n}(\mathbf{f}_n - \mathbf{p}) \right)^t \stackrel{\mathcal{L}}{\to} N \left(0,\, \nabla \Delta_S^2(\mathbf{p}) \left[\mathrm{I}\,{}^F_{\mathcal{X}}(\mathbf{p}) \right]^{-1} \left(\nabla \Delta_S^2(\mathbf{p}) \right)^t \right),$$

siempre que $\nabla \Delta_S^2(\mathbf{p}) \left[\mathrm{I}\,{}^F_{\mathcal{X}}(\mathbf{p}) \right]^{-1} \left(\nabla \Delta_S^2(\mathbf{p}) \right)^t > 0$.

En virtud de las propiedades relativas a las convergencias en ley y en probabilidad (ver, por ejemplo, Serfling 1980, págs. 19, 24, 26), al ser $\|\mathbf{p}_n^* - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$ se tiene que cuando $n \to \infty$:

$$\left(\frac{\partial^2 \Delta_S^2(\mathcal{X} \,|\, \mathbf{p})}{\partial p_i \partial p_j} \right)_{\mathbf{p} = \mathbf{p}_n^*} \stackrel{p}{\to} \frac{\partial^2 \Delta_S^2(\mathcal{X} \,|\, \mathbf{p})}{\partial p_i \partial p_j},$$

para cualesquiera $i, j \in \{1, \ldots, r-1\}$. Por otro lado, $\mathbf{f}_n \stackrel{c.s.}{\to} \mathbf{p}$, y por lo tanto $\mathbf{f}_n \stackrel{p}{\to} \mathbf{p}$, cuando $n \to \infty$, de donde:

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p}) H\Big(\Delta_S^2(\mathbf{p}_n^*) \Big) \stackrel{p}{\to} \mathbf{0},$$

y, como $\sqrt{n}(\mathbf{f}_n - \mathbf{p}) \xrightarrow{\mathcal{L}} N\left(\mathbf{0},\, \left[\mathrm{I}\,^F_{\mathcal{X}}(\mathbf{p})\right]^{-1}\right)$, se tiene que:

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p})H\Big(\Delta_S^2(\mathbf{p}_n^*)\Big)\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \xrightarrow{\mathcal{L}} 0,$$

y, en consecuencia:

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p})H\Big(\Delta_S^2(\mathbf{p}_n^*)\Big)\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \xrightarrow{p} 0.$$

Los resultados anteriores garantizan que:

$$\sqrt{n}\left[\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \Delta_S^2(\mathcal{X}\,|\,\mathbf{p})\right]$$

$$\xrightarrow{\mathcal{L}}\; N\left(0,\, \nabla\Delta_S^2(\mathbf{p})\left[\mathrm{I}\,^F_{\mathcal{X}}(\mathbf{p})\right]^{-1}\left(\nabla\Delta_S^2(\mathbf{p})\right)^t\right)$$

cuando $n \to \infty$, por lo tanto se verifica *ii)* siempre que $\sigma^2(\mathbf{p}) = \nabla\Delta_S^2(\mathbf{p})\left[\mathrm{I}\,^F_{\mathcal{X}}(\mathbf{p})\right]^{-1}(\nabla\Delta_S^2(\mathbf{p}))^t = \mathrm{W}^*\Sigma\mathrm{W}^{*t} > 0$, con $\mathrm{W}^* =$ matriz fila $(\mathrm{W}_1^*, \ldots, \mathrm{W}_r^*)$, donde $\mathrm{W}_l^* = \left[D_S\left(\tilde{x}_l^*, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2$ y:

$$\Sigma = \begin{pmatrix} p_1(1-p_1) & \cdots & -p_1p_r \\ \vdots & \ddots & \vdots \\ -p_rp_1 & \cdots & p_r(1-p_r) \end{pmatrix}$$

ya que:

$$\sigma^2(\mathbf{p}) = \sum_{i=1}^{r-1}\sum_{j=1}^{r-1} p_i(\delta_{ij} - p_j)\left(\frac{\partial}{\partial p_i}\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})\right)\left(\frac{\partial}{\partial p_j}\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})\right)$$

$$= \mathrm{W}^*\Sigma\mathrm{W}^{*t} = \mathrm{Var}\left(\left[D_S\left(\mathcal{X}, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2\right).$$

iii) La matriz $\mathrm{I}\,^F_{\mathcal{X}}(\mathbf{p})$ es definida positiva, por lo que la forma cuadrática asociada a $\mathrm{I}\,^F_{\mathcal{X}}(\mathbf{p})$ y a $\left[\mathrm{I}\,^F_{\mathcal{X}}(\mathbf{p})\right]^{-1}$ es definida positiva. En consecuencia, si $\sigma^2(\mathbf{p}) = \nabla\Delta_S^2(\mathbf{p})\left[\mathrm{I}\,^F_{\mathcal{X}}(\mathbf{p})\right]^{-1}(\nabla\Delta_S^2(\mathbf{p}))^t = 0$ se debe cumplir que $\nabla\Delta_S^2(\mathbf{p}) = 0$.

Si se considera ahora el desarrollo de Taylor de segundo orden de $\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n)$, puede asegurarse que para n suficientemente grande:

$$\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) = \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}) + \frac{1}{2}(\mathbf{f}_n - \mathbf{p})H\Big(\Delta_S^2(\mathbf{p})\Big)\,(\mathbf{f}_n - \mathbf{p})^t$$
$$+\frac{1}{6}\sum_{i=1}^{r-1}\sum_{j=1}^{r-1}\sum_{k=1}^{r-1}\left(\frac{\partial^3\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i\partial p_j\partial p_k}\right)_{\mathbf{p}=\mathbf{p}_n^{**}}(f_{ni}-p_i)(f_{nj}-p_j)(f_{nk}-p_k),$$

con $\mathbf{p}_n^{**} \in \mathbb{P}$ tal que $\|\mathbf{p}_n^{**} - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$. Por lo tanto:

$$2n\left[\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \Delta_S^2(\mathcal{X}\,|\,\mathbf{p})\right]$$
$$= (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))H\Big(\Delta_S^2(\mathbf{p})\Big)\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t$$
$$+\frac{1}{3}\sum_{i=1}^{r-1}\sum_{j=1}^{r-1}\sum_{k=1}^{r-1}\left(\frac{\partial^3\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i\partial p_j\partial p_k}\right)_{\mathbf{p}=\mathbf{p}_n^{**}}$$
$$\cdot\,(f_{ni}-p_i)\left(\sqrt{n}(f_{nj}-p_j)\right)\left(\sqrt{n}(f_{nk}-p_k)\right).$$

Razonando como en el apartado *ii)*, puede concluirse que:

$$(f_{ni}-p_i)\left(\sqrt{n}(f_{nj}-p_j)\right)\left(\sqrt{n}(f_{nk}-p_k)\right) \xrightarrow{p} 0.$$

Además (ver, por ejemplo, Serfling 1980, pág. 25) se tiene que:

$$\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)H\Big(\Delta_S^2(\mathbf{p})\Big)\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \xrightarrow{\mathcal{L}} YH\Big(\Delta_S^2(\mathbf{p})\Big)\,Y^t,$$

donde Y es un vector aleatorio con distribución normal $(r-1)$-dimensional $N\Big(\mathbf{0},\Big[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\Big]^{-1}\Big)$.

Como se supone que alguna de las derivadas segundas $h_{ij} = \dfrac{\partial^2\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i\partial p_j}$ es superior a cero (y, por lo tanto, que $H(\Delta_S^2(\mathbf{p}))$ no se reduce a la matriz nula de orden $(r-1)\times(r-1)$), y ya que el rango de $\Big[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\Big]^{-1}$ es $r-1$, debe cumplirse que el vector Y puede expresarse como $Y = ZB$ con Z vector aleatorio $(r-1)$-dimensional cuyas componentes son $r-1$ variables aleatorias independientes e idénticamente distribuidas, con distribución $N(0,1)$, y B es una matriz $(r-1)\times(r-1)$ tal que $BB^t = \left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}$. Además, existe una transformación $Z = UC$ en la que C es una matriz ortogonal, de forma que:

$$\begin{aligned} Y H\Big(\Delta_S^2(\mathbf{p})\Big)\, Y^t &= Z B H\Big(\Delta_S^2(\mathbf{p})\Big)\, B^t Z^t \\ &= U C B H\Big(\Delta_S^2(\mathbf{p})\Big)\, B^t C^t U^t = \lambda_1 U_1^2 + \cdots + \lambda_q U_q^2, \end{aligned}$$

con $\lambda_1, \ldots, \lambda_q$ $(q \le r-1)$ los autovalores no nulos de la matriz $(r-1)\times(r-1)$ dada por $BH(\Delta_S^2(\mathbf{p}))\, B^t$, y donde $U_1, \ldots, U_q$ son variables aleatorias independientes e idénticamente distribuidas con distribución $N(0,1)$.

En consecuencia, el estudio de la forma cuadrática $YH(\Delta_S^2(\mathbf{p}))\, Y^t$ se reduce en primer lugar al de una forma cuadrática de un vector normal $(r-1)$-dimensional con componentes independientes y $N(0,1)$, y este último se reduce a su vez al de una combinación lineal de variables aleatorias independientes e idénticamente distribuidas, con distribución $N(0,1)$, es decir, una combinación lineal de cuadrados de variables aleatorias independientes e idénticamente distribuidas, con distribución chi-cuadrado χ_1^2 (ver, por ejemplo, Rao 1973, págs. 186-188 y Serfling 1980, págs. 25, 128-130, para una revisión de los resultados básicos de las afirmaciones precedentes). $\square$

Observación 1.4.1 . Conviene señalar que (ver, por ejemplo, Rao 1973, págs. 186-188 y Serfling 1980, págs. 25, 128-130), al ser $\left|\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}\right| = p_1 \cdots p_r \neq 0$, la variable $YH(\Delta_S^2(\mathbf{p}))\, Y^t$ tendrá distribución chi-cuadrado si, y sólo si:

$$H\Big(\Delta_S^2(\mathbf{p})\Big)\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1} H\Big(\Delta_S^2(\mathbf{p})\Big) = H\Big(\Delta_S^2(\mathbf{p})\Big),$$

en cuyo caso el número de grados de libertad correspondiente coincidiría con el rango de $H(\Delta_S^2(\mathbf{p}))\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}$, que por la regularidad de $\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}$ coincide con el rango de $H(\Delta_S^2(\mathbf{p}))$.

Observación 1.4.2. Los resultados anteriores y los resultados y procedimientos que se desarrollan en el resto de la sección son también aplicables directamente sobre poblaciones infinitas, siempre que el número de valores distintos de la variable en esas poblaciones sea finito, que es la situación más común en la práctica.

Observación 1.4.3. Se puede verificar fácilmente a partir del Teorema 1.3.4 que:

$$\sigma^2(\mathbf{p}) = \lim_{N\to\infty} \lim_{n\to\infty} \mathrm{Var}\left(\sqrt{n}\, \widehat{\Delta_S^2}^{w}(\mathcal{X}[\,\cdot\,]) \,\Big|\, p^w\right).$$

El apartado *ii)* del teorema anterior puede ser modificado para poder calcular la varianza asintótica del estimador $\Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n)$ en la práctica, y poder desarrollar (aunque sea de forma aproximada) inferencias como la estimación por intervalo y algunos contrastes de hipótesis. Más concretamente, cuando $\sigma^2(\mathbf{p})$ se reemplaza por su estimación analógica, $\sigma^2(\mathbf{f}_n)$, se obtiene la conclusión siguiente:

Teorema 1.4.2. *En las condiciones del Teorema* 1.4.1*, se cumple que:*

$$\left\{ \frac{\sqrt{n}\,(\Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n) - \Delta_S^2(\mathcal{X} \,|\, \mathbf{p}))}{\sqrt{\sigma^2(\mathbf{f}_n)}} \right\}_n$$

converge en ley hacia una distribución normal $N(0,1)$ *cuando* $n \to \infty$, *siempre que* $\sigma^2(\mathbf{p}) > 0$ *y* $\sigma^2(\mathbf{f}_n) > 0$.

Demostración:

Como las componentes de $\nabla\Delta_S^2(\mathbf{p})$ y de $\left[\mathrm{I}\,{}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}$ son continuas en un entorno de $\mathbf{p}$, se cumple que cuando $n \to \infty$:

$$\nabla\Delta_S^2(\mathbf{f}_n) \xrightarrow{p} \nabla\Delta_S^2(\mathbf{p}),$$

y que:

$$\left[\mathrm{I}\,{}_{\mathcal{X}}^{F}(\mathbf{f}_n)\right]^{-1} \xrightarrow{p} \left[\mathrm{I}\,{}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}.$$

En consecuencia:

$$\sqrt{\sigma^2(\mathbf{f}_n)} \xrightarrow{p} \sqrt{\sigma^2(\mathbf{p})}$$

cuando $n \to \infty$, y al cumplirse que:

$$\sqrt{n}\left(\Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n) - \Delta_S^2(\mathcal{X} \,|\, \mathbf{p})\right) \xrightarrow{\mathcal{L}} N(0, \sigma^2(\mathbf{p})),$$

se satisface (ver, por ejemplo, Serfling 1980, pág. 19) que:

$$\frac{\sqrt{n}\,(\Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n) - \Delta_S^2(\mathcal{X} \,|\, \mathbf{p}))}{\sqrt{\sigma^2(\mathbf{f}_n)}} \xrightarrow{\mathcal{L}} N(0, 1).$$

□

El estudio desarrollado hasta el momento en esta sección, permite realizar inferencias adicionales sobre la medida de variación absoluta poblacional.

Por lo que se refiere a la *estimación por intervalo*, el procedimiento que se expone a continuación nos suministrará un rango de valores posibles para la S-DCM central poblacional que cubrirá el verdadero valor de esta medida con (en este caso aproximadamente) una probabilidad prefijada.

El Teorema 1.4.2 nos permite establecer de forma aproximada los límites del rango de valores anterior para muestras grandes, como sigue:

Teorema 1.4.3. *En las condiciones de los Teoremas* 1.4.1 *y* 1.4.2*, el intervalo aleatorio:*

$$\left[\Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n) - z_\alpha\sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}},\ \Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n) + z_\alpha\sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}}\,\right],$$

con $z_\alpha =$ *cuantil de orden* $1 - \alpha/2$ *de la distribución* $N(0,1)$*, proporciona para cada muestra de* n *observaciones independientes a partir de* $\mathcal{X}$ *intervalos de confianza de* $\Delta_S^2(\mathcal{X} \,|\, \mathbf{p})$ *con coeficiente aproximadamente igual a* $1 - \alpha$ *(con* $\alpha \in [0,1]$*).*

Por último, y de modo inmediato, se derivan los *contrastes de hipótesis* siguientes para grandes muestras:

Teorema 1.4.4. *En las condiciones de los Teoremas* 1.4.1 *y* 1.4.2*:*

(i) Para contrastar al nivel de significación $\alpha \in [0,1]$ *la hipótesis nula:*

$$H_0 : \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}) = \delta_0$$

frente a la alternativa:

$$H_1 : \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}) \neq \delta_0,$$

si $|\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \delta_0| > z_\alpha \sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$ *debe rechazarse* H_0*. El p-valor del contraste vendrá dado, aproximadamente, por:*

$$p = 2\left[1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\left|\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \delta_0\right|\right)\right]$$

($\Phi =$ *función de distribución de* $N(0,\,1)$).

(ii) Para contrastar al nivel de significación α *la hipótesis nula:*

$$H_0 : \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}) \geq \delta_0$$

frente a la alternativa:

$$H_1 : \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}) < \delta_0,$$

si $\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \delta_0 < z_{2\alpha} \sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$ *debe rechazarse* H_0*. El p-valor del contraste vendrá dado, aproximadamente, por:*

$$p = \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\left(\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \delta_0\right)\right).$$

(iii) Para contrastar al nivel de significación α *la hipótesis nula:*

$$H_0 : \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}) \leq \delta_0$$

frente a la alternativa:

$$H_1 : \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}) > \delta_0,$$

si $\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \delta_0 > z_{2\alpha}\sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$ *debe rechazarse* H_0. *El p-valor del contraste vendrá dado, aproximadamente, por:*

$$p = 1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\left(\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \delta_0\right)\right).$$

En las dos secciones siguientes de este capítulo, van a examinarse dos aplicaciones estadísticas de la medida generalizada de variación absoluta de una variable aleatoria difusa, introducida y analizada en la secciones precedentes. En ambas aplicaciones, la S-DCM jugará un papel análogo al del error cuadrático medio en el caso de variables aleatorias reales. En la Sección 1.5 se estudiará el problema de la estimación de un parámetro difuso (el valor esperado) asociado a una variable aleatoria difusa en los muestreos aleatorios simple y con reposición de poblaciones finitas. En la Sección 1.6 se analizará el problema de regresión lineal y funcional entre variables aleatorias difusas, obteniéndose la solución bajo condiciones bastante generales.

En la Sección 1.5 se ilustrará el empleo de alguno de los procedimientos recogidos en el Teorema 1.4.4, combinado con los resultados de las Secciones 1.3 y 1.5.

1.5 Aplicación de la medida de variación absoluta a la estimación del valor esperado de variables aleatorias difusas en muestreos aleatorios de poblaciones finitas

En esta sección se considera en primer lugar el problema de estimar el valor esperado de una variable aleatoria difusa en una población finita, cuando se dispone de la información proporcionada por una muestra aleatoria simple o con reposición de esta población. Se complementa tal estudio con el análisis de la precisión del proceso de estimación.

En la teoría de la Estadística Matemática, se supone a menudo que las variables aleatorias reales involucradas siguen una distribución perteneciente a una clase paramétrica conocida y se trabaja con poblaciones cualesquiera, mientras que en la Teoría de Muestras suele admitirse que se consideran poblaciones finitas y no se dispone del conocimiento sobre la clase a la que pertenece la distribución poblacional.

En la literatura sobre variables aleatorias difusas, el estudio de modelos de "distribuciones" para variables aleatorias difusas se reduce prácticamente al correspondiente a las variables difusas normales (ver Puri & Ralescu 1985, Ralescu 1995c), de modo que en el marco de las variables aleatorias difusas cobra más sentido el desarrollo de estudios análogos a los de la Teoría de Muestras.

Por lo que se refiere a la estimación del valor esperado poblacional de una variable aleatoria difusa en los muestreos aleatorios con y sin reposición, va a comprobarse que el valor esperado muestral de esa variable es un estimador insesgado (en el sentido del valor esperado definido por Puri & Ralescu 1986).

En este estudio la S-DCM servirá de apoyo para medir la precisión del "estimador difuso" valor esperado muestral del "parámetro difuso" valor esperado poblacional.

Si se considera una población finita de N unidades, $\Omega = \{\omega_1, \ldots, \omega_N\}$, para las que la variable aleatoria difusa $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ asociada al espacio de probabilidad $(\Omega, \mathcal{P}(\Omega), P)$ (con P la distribución uniforme sobre Ω) toma los valores $\mathcal{X}(\omega_1), \ldots, \mathcal{X}(\omega_N)$, entonces, de acuerdo con las propiedades del valor esperado para variables aleatorias difusas con valores convexos, el *valor esperado de $\mathcal{X}$ en esta población* viene dado por:

$$\overline{\mathcal{X}} = \widetilde{E}(\mathcal{X} \mid P) = \frac{1}{N} \odot \bigoplus_{j=1}^{N} \mathcal{X}(\omega_j),$$

es decir, para cada $\alpha \in [0,1]$:

$$\overline{\mathcal{X}}_\alpha = \left[\frac{1}{N} \sum_{j=1}^{N} \inf \mathcal{X}_\alpha(\omega_j), \frac{1}{N} \sum_{j=1}^{N} \sup \mathcal{X}_\alpha(\omega_j) \right].$$

Supongamos que se selecciona una muestra de tamaño n al azar y sin reposición a partir de la población Ω, que υ representa una muestra aleatoria simple genérica de tamaño n, y que $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ denotan las unidades en la muestra υ. Entonces, el *valor esperado muestral* de $\mathcal{X}$ en υ viene dado por:

$$\overline{\mathcal{X}}_n[\upsilon] = \frac{1}{n} \odot \sum_{i=1}^{n} \mathcal{X}(\omega_{\upsilon i}),$$

y $\overline{\mathcal{X}}_n$ define una variable aleatoria difusa asociada con el espacio de probabilidad $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ (ver Sección 1.3).

La variable aleatoria difusa $\overline{\mathcal{X}}_n$ constituye por tanto un *estimador difuso* de $\overline{\mathcal{X}}$ y, su valor esperado sobre el espacio $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ coincide con $\overline{\mathcal{X}}$, lo que permite concluir que $\overline{\mathcal{X}}_n$ es un *estimador insesgado difuso de* $\overline{\mathcal{X}}$ *en el muestreo aleatorio simple* (la falta de sesgo se supone entendida en el sentido del valor esperado definido por Puri & Ralescu 1986). De esta forma:

Teorema 1.5.1. *En el muestreo aleatorio simple de tamaño n de la población Ω con N unidades, el estimador $\mathcal{F}_c(\mathbb{R})$-valorado $\overline{\mathcal{X}}_n$ es insesgado para estimar $\overline{\mathcal{X}}$, es decir, $\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p) = \overline{\mathcal{X}}$.*

Demostración:

En efecto:

$$\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p) = \sum_{\upsilon \in \Upsilon_n} p[\upsilon] \odot \overline{\mathcal{X}}_n[\upsilon].$$

Por otro lado:

$$\overline{\mathcal{X}}_n[\upsilon] = \frac{1}{n} \odot \sum_{j=1}^{N} a_j[\upsilon] \odot \mathcal{X}(\omega_j),$$

donde a_j es la variable Bernoulli definida en $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ que está asociada con la presencia de ω_j en cada muestra.

En virtud de las propiedades del valor esperado de una variable aleatoria difusa con valores en $\mathcal{F}_c(\mathbb{R})$ (ver, por ejemplo, Puri & Ralescu 1986), se tiene que:

$$\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p) = \frac{1}{n} \odot \sum_{j=1}^{N} E(a_j) \odot \mathcal{X}(\omega_j),$$

donde $E(a_j) = n/N$ y, por lo tanto, $\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p) = \overline{\mathcal{X}}$. □

Para *cuantificar la precisión de un estimador difuso* de un parámetro difuso va a considerarse la S-dispersión cuadrática media del estimador respecto al parámetro. El interés de esta cuantificación reside en la posibilidad de comparar (sin utilizar técnicas de ordenación de números difusos) diferentes procedimientos de muestreo o distintos estimadores, y en la de formalizar el problema de elección del tamaño de muestra para asegurar una precisión prefijada.

Con respecto a la estimación de $\overline{\mathcal{X}}$ a través de $\overline{\mathcal{X}}_n$, puede establecerse que:

Teorema 1.5.2. *En el muestreo aleatorio simple de tamaño n de una población Ω con N unidades, si $f = n/N$ la S-dispersión cuadrática media de $\overline{\mathcal{X}}_n$ respecto a $\overline{\mathcal{X}}$ viene dada por:*

$$DCM_S\left(\overline{\mathcal{X}}_n, \overline{\mathcal{X}}\right) = \Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p\right) = \frac{(1-f)N}{n(N-1)}\Delta_S^2\left(\mathcal{X} \,|\, P\right).$$

Demostración:

En efecto, de acuerdo con las propiedades de la S-DCM se tiene que:

$$\Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p\right) = \Delta_S^2\left(\frac{1}{n} \odot \bigoplus_{j=1}^{N} a_j \odot \mathcal{X}(\omega_j)\right)$$

$$= \frac{1}{n^2}\sum_{j=1}^{N}\Delta_S^2\left(a_j \odot \mathcal{X}(\omega_j)\right) + \frac{2}{n^2}\sum_{j=1}^{N-1}\sum_{l=j+1}^{N}\Delta_S\left(a_j \odot \mathcal{X}(\omega_j), a_l \odot \mathcal{X}(\omega_l)\right).$$

Para cada $j \in \{1, \ldots, N\}$ se cumple (ver Teorema 1.2.4) que:

$$\Delta_S^2\left(a_j \odot \mathcal{X}(\omega_j)\right) = \int_{[0,1]}\int_{[0,1]} \operatorname{Var}\left[f_{a_j \odot \mathcal{X}(\omega_j)}(\alpha, \lambda)\right] d\,S(\lambda)\, d\,\alpha,$$

donde para todo $\alpha \in [0,1]$ y $\lambda \in [0,1]$:

$$f_{a_j \odot \mathcal{X}(\omega_j)}(\alpha, \lambda) = \lambda \sup\left(a_j \odot \mathcal{X}(\omega_j)\right)_\alpha + (1-\lambda)\inf\left(a_j \odot \mathcal{X}(\omega_j)\right)_\alpha$$

$$= \lambda a_j \sup \mathcal{X}(\omega_j)_\alpha + (1-\lambda) a_j \inf \mathcal{X}(\omega_j)_\alpha = a_j f_{\mathcal{X}(\omega_j)}(\alpha, \lambda),$$

y, por lo tanto:

$$\mathrm{Var}\left[f_{a_j \odot \mathcal{X}(\omega_j)}(\alpha,\lambda)\right] = \mathrm{Var}(a_j)\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda)\right]^2,$$

de donde:

$$\Delta_S^2\left(a_j \odot \mathcal{X}(\omega_j)\right) = \mathrm{Var}(a_j)\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda)\right]^2 d\,S(\lambda)\,d\,\alpha.$$

Por otro lado, para cada $j, l \in \{1, \ldots, N\}$ con $l \neq j$:

$$\Delta_S\left(a_j \odot \mathcal{X}(\omega_j), a_l \odot \mathcal{X}(\omega_l)\right)$$

$$= \int_{[0,1]}\int_{[0,1]} \mathrm{Cov}\left[f_{a_j \odot \mathcal{X}(\omega_j)}(\alpha,\lambda), f_{a_l \odot \mathcal{X}(\omega_l)}(\alpha,\lambda)\right] d\,S(\lambda)\,d\,\alpha,$$

donde:

$$\mathrm{Cov}\left[f_{a_j \odot \mathcal{X}(\omega_j)}(\alpha,\lambda), f_{a_l \odot \mathcal{X}(\omega_l)}(\alpha,\lambda)\right]$$

$$= \mathrm{Cov}\left[a_j f_{\mathcal{X}(\omega_j)}(\alpha,\lambda), a_l f_{\mathcal{X}(\omega_l)}(\alpha,\lambda)\right]$$

$$= \left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda)\right]\left[f_{\mathcal{X}(\omega_l)}(\alpha,\lambda)\right]\mathrm{Cov}\left(a_j, a_l\right),$$

por lo que:

$$\Delta_S\left(a_j \odot \mathcal{X}(\omega_j), a_l \odot \mathcal{X}(\omega_l)\right)$$

$$= \mathrm{Cov}\left(a_j, a_l\right)\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda)\right]\left[f_{\mathcal{X}(\omega_l)}(\alpha,\lambda)\right] d\,S(\lambda)\,d\,\alpha.$$

Por las expresiones que adoptan $\mathrm{Var}(a_j)$ y $\mathrm{Cov}(a_j, a_l)$ (ver Teorema 1.3.2), se tiene que:

$$\Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p\right) = \frac{1-f}{n}\int_{[0,1]}\int_{[0,1]}\left[\frac{1}{N}\sum_{j=1}^{N}\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda)\right]^2\right.$$

$$\left. - \frac{2}{N(N-1)}\sum_{j=1}^{N-1}\sum_{l=j+1}^{N}\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda)\right]\left[f_{\mathcal{X}(\omega_l)}(\alpha,\lambda)\right]\right] d\,S(\lambda)\,d\,\alpha,$$

$$= \frac{(1-f)N}{n(N-1)}\int_{[0,1]}\int_{[0,1]}\left[\frac{1}{N}\sum_{j=1}^{N}\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda)\right]^2\right.$$

$$-\frac{1}{N^2}\sum_{j=1}^{N}\sum_{l=1}^{N}\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda)\right]\left[f_{\mathcal{X}(\omega_l)}(\alpha,\lambda)\right]\Bigg]\,d\,S(\lambda)\,d\,\alpha$$

$$=\frac{(1-f)N}{n(N-1)}\int_{[0,1]}\int_{[0,1]}\mathrm{Var}\left[f_{\mathcal{X}(\omega)}(\alpha,\lambda)\right]\,d\,S(\lambda)\,d\,\alpha$$

$$=\frac{(1-f)N}{n(N-1)}\Delta_S^2\left(\mathcal{X}\,|\,P\right).$$

□

El resultado del Teorema 1.5.2 no sirve habitualmente para calcular el error muestral en la estimación de $\overline{\mathcal{X}}$ por medio de $\overline{\mathcal{X}}_n$, ya que $\Delta_S^2\left(\mathcal{X}\,|\,P\right)$ suele ser desconocido. Como consecuencia, en la mayoría de los objetivos para los que se calcula la medida de error anterior (aproximación de la magnitud de la S-dispersión cuadrática media central de los valores de $\mathcal{X}$ en la población, o determinación del tamaño de muestra necesario para asegurar la precisión deseada), $\Delta_S^2\left(\mathcal{X}\,|\,P\right)$ debe ser estimada a partir de los datos muestrales difusos.

El Teorema 1.3.1 nos permite establecer tal estimación como sigue:

Teorema 1.5.3. *En el muestreo aleatorio simple de tamaño n de la población Ω con N unidades, el estimador que a la muestra υ le asocia el valor:*

$$\widehat{\Delta_S^2}\left(\overline{\mathcal{X}}_n[\upsilon]\right)=\frac{(N-1)n}{N(n-1)}\Delta_S^2\left(\overline{\mathcal{X}}_n[\upsilon]\right)$$

$$=\frac{(N-1)n}{N(n-1)}\int_{[0,1]}\int_{[0,1]}\mathrm{Var}\left(f_{\overline{\mathcal{X}}_n}(\alpha,\lambda)\right)d\,S(\lambda)d\,\alpha,$$

es insesgado para estimar $\Delta_S^2\left(\overline{\mathcal{X}}_n\,|\,p\right)$.

Obviamente, un incremento en el tamaño muestral supone un incremento en la precisión de $\overline{\mathcal{X}}_n$ como estimador de $\overline{\mathcal{X}}$, aunque el coste del muestreo también aumentará. Supóngase que, teniendo en cuenta ambos aspectos (precisión y costes), se desea estimar $\overline{\mathcal{X}}$ mediante el valor esperado de $\mathcal{X}$ en una muestra aleatoria simple de tamaño n a partir de Ω, tal que con un *nivel de confianza* o *riesgo* $\alpha\in[0,1]$, se garantiza que la probabilidad de que el error de estimar $\overline{\mathcal{X}}$ por medio de $\overline{\mathcal{X}}_n$ supere una tolerancia d no sea inferior a α. De este modo, el objetivo es obtener el *mínimo tamaño muestral* $n \in \mathbb{N}$ con el que pueda asegurarse que:

$$\mathrm{P}\Big(D_S(\overline{\mathcal{X}}_n, \overline{\mathcal{X}}) > d\Big) \leq \alpha.$$

Para este propósito se va a recurrir a la extensión de la Desigualdad de Tchebychev recogida en el Teorema 1.2.13, de acuerdo con la cual:

Teorema 1.5.4. *En el muestreo aleatorio simple de la población* Ω *con* N *unidades muestrales, el tamaño muestral*

$$n = \left.\frac{N\Delta_S^2(\mathcal{X}\,|\,P)}{(N-1)d^2\alpha + \Delta_S^2(\mathcal{X}\,|\,P)}\right]$$

satisface que $\mathrm{P}\Big(D_S(\overline{\mathcal{X}}_n, \overline{\mathcal{X}}) > d\Big) \leq \alpha$, *donde* $a]$ = *menor entero mayor o igual que* $a \in \mathbb{R}$.

Demostración:

En efecto, sobre la base de la extensión de la Desigualdad de Tchebychev, una condición suficiente para que $\mathrm{P}\Big(D_S(\overline{\mathcal{X}}_n, \overline{\mathcal{X}}) > d\Big) \leq \alpha$ es que:

$$\frac{\Delta_S^2(\overline{\mathcal{X}}_n\,|\,p)}{d^2} \leq \alpha,$$

es decir:

$$\frac{1}{n} - \frac{1}{N} \leq \frac{(N-1)d^2\alpha}{N\Delta_S^2(\mathcal{X}\,|\,P)},$$

y el valor mínimo en $\mathbb{N}$ que cumple esta condición es:

$$n = \left.\frac{N\Delta_S^2(\mathcal{X}\,|\,P)}{(N-1)d^2\alpha + \Delta_S^2(\mathcal{X}\,|\,P)}\right].$$

□

El tamaño muestral propuesto en el Teorema 1.5.4 no es realmente el mínimo tamaño muestral que se ajusta al objetivo perseguido puesto que se ha basado en una condición suficiente. De hecho, y como en el caso de variables aleatorias reales, la extensión de la Desigualdad de Tchebychev determina un procedimiento habitualmente muy conservador para elegir el tamaño muestral. De cualquier forma, el tamaño propuesto en el Teorema 1.5.4 dependerá de $\Delta_S^2(\mathcal{X} \,|\, P)$, que suele ser desconocido. Un modo adecuado (siempre que sea factible) para aproximar este valor desconocido, es considerar una muestra aleatoria simple previa de un tamaño moderado n_1, y emplear esta muestra para estimar $\Delta_S^2(\mathcal{X} \,|\, P)$ por medio del estimador con valores reales $\widehat{\Delta_S^2}\left(\overline{\mathcal{X}}_{n_1}[\,\cdot\,]\right)$. Si:

$$n = \left.\frac{N\widehat{\Delta_S^2}\left(\overline{\mathcal{X}}_{n_1}[\,\cdot\,]\right)}{(N-1)d^2\alpha + \widehat{\Delta_S^2}\left(\overline{\mathcal{X}}_{n_1}[\,\cdot\,]\right)}\right\rceil > n_1,$$

entonces se completará la muestra previa eligiendo una muestra aleatoria simple adicional de tamaño $n - n_1$.

Supongamos que se selecciona una muestra de tamaño n al azar y con reposición a partir de toda la población Ω. El *valor esperado muestral* de $\mathcal{X}$, $\overline{\mathcal{X}}_n$, define ahora una variable aleatoria difusa asociada con el espacio de probabilidad $(\Upsilon_n^w, \mathcal{P}(\Upsilon_n^w), p^w)$ (ver Sección 1.3).

Como en el muestreo aleatorio simple, $\overline{\mathcal{X}}_n$ es un *estimador insesgado difuso de* $\overline{\mathcal{X}}$ *en el muestreo aleatorio con reposición.* De este modo:

Teorema 1.5.5. *En el muestreo aleatorio con reposición de tamaño n de la población Ω con N unidades muestrales, $\omega_1, \ldots, \omega_N$, el estimador $\mathcal{F}_c(\mathbb{R})$-valorado $\overline{\mathcal{X}}_n$ es insesgado para estimar el valor esperado poblacional $\overline{\mathcal{X}}$, es decir, $\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p^w) = \overline{\mathcal{X}}$.*

Demostración:

En efecto:

$$\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p^w) = \sum_{\upsilon \in \Upsilon_n^w} p^w[\upsilon] \odot \overline{\mathcal{X}}_n[\upsilon],$$

y

$$\overline{\mathcal{X}}_n[\upsilon] = \frac{1}{n} \odot \sum_{j=1}^{N} t_j[\upsilon] \odot \mathcal{X}(\omega_j),$$

donde $t_j[v]$ es el número de veces que ω_j aparece en $v \in \Upsilon_n^w$, $j = 1, \ldots, N$ (ver Teorema 1.3.4).

Por las propiedades del valor esperado de una variable aleatoria difusa, se tiene que:

$$\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p^w) = \frac{1}{n} \odot \sum_{j=1}^{N} E(t_j) \odot \mathcal{X}(\omega_j),$$

donde $E(t_j) = n/N$, y por tanto $\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p^w) = \overline{\mathcal{X}}$. □

El estudio de la cuantificación de la *precisión* de $\overline{\mathcal{X}}_n$ en la estimación de $\overline{\mathcal{X}}$, permite concluir para este procedimiento de muestreo que:

Teorema 1.5.6. *En el muestreo aleatorio con reposición de tamaño n de la población Ω con N unidades muestrales, la S-dispersión cuadrática media de $\overline{\mathcal{X}}_n$ respecto a $\overline{\mathcal{X}}$ viene dada por:*

$$DCM_S\left(\overline{\mathcal{X}}_n, \overline{\mathcal{X}}\right) = \Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p^w\right) = \frac{\Delta_S^2\left(\mathcal{X} \,|\, P\right)}{n}.$$

Demostración:

En efecto, siguiendo argumentos similares a los del Teorema 1.5.2, se tiene que:

$$\Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p^w\right) = \Delta_S^2\left(\frac{1}{n} \odot \sum_{j=1}^{N} t_j \odot \mathcal{X}(\omega_j)\right),$$

y, por las expresiones que adoptan $\mathrm{Var}(t_j)$ y $\mathrm{Cov}(t_j, t_l)$ (ver Teorema 1.3.4) se tiene que:

$$\Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p^w\right) = \frac{1}{n} \int_{[0,1]} \int_{[0,1]} \left(\frac{N-1}{N^2} \sum_{j=1}^{N} \left[f_{\mathcal{X}(\omega_j)}(\alpha, \lambda) \right]^2 \right.$$

$$\left. - \frac{2}{N^2} \sum_{j=1}^{N-1} \sum_{l=j+1}^{N} \left[f_{\mathcal{X}(\omega_j)}(\alpha, \lambda) \right] \left[f_{\mathcal{X}(\omega_l)}(\alpha, \lambda) \right] \right) d\,S(\lambda)\, d\,\alpha$$

$$= \frac{\Delta_S^2\left(\mathcal{X} \,|\, P\right)}{n}.$$

□

Como consecuencia inmediata de los Teoremas 1.5.2 y 1.5.7 se deduce que el muestreo aleatorio simple es más preciso que el muestreo aleatorio con reposición en la estimación de $\overline{\mathcal{X}}$.

El siguiente resultado que establece la estimación de $\Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p^w\right)$ en el muestreo aleatorio con reposición se basa en el Teorema 1.3.3:

Teorema 1.5.7. *En el muestreo aleatorio con reposición de tamaño n de la población Ω con N unidades muestrales, el estimador que a la muestra υ le asocia el valor:*

$$\widehat{\Delta_S^2}^{w}\left(\overline{\mathcal{X}}_n[\upsilon]\right) = \frac{n}{n-1}\Delta_S^2\left(\overline{\mathcal{X}}_n[\upsilon]\right),$$

es insesgado para estimar $\Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p^w\right)$.

El problema de *elección del tamaño de muestra más apropiado* se puede resolver ahora como sigue:

Teorema 1.5.8. *En el muestreo aleatorio con reposición de tamaño n de la población Ω con N unidades muestrales, el tamaño muestral*

$$n = \left.\frac{\Delta_S^2(\mathcal{X} \,|\, P)}{d^2\alpha}\right\rceil$$

satisface que $\mathrm{P}\Big(D_S(\overline{\mathcal{X}}_n, \overline{\mathcal{X}}) > d\Big) \leq \alpha$.

Antes de concluir esta sección, va a examinarse un resultado relativo a la estimación del valor esperado poblacional $\overline{\mathcal{X}}$ en el muestreo aleatorio con reposición de poblaciones finitas, mediante el cual se ilustra en parte la aplicación de la extensión del Teorema de Rao-Blackwell, que se estudió en el Teorema 1.2.14 (ver también Lubiano *et al.* 1999b), a la estimación difusa de un parámetro difuso.

Conviene observar que la variable $\mathcal{Y}$ del Teorema 1.2.14 se corresponde en este caso con una variable aleatoria con valores reales.

Supongamos que se elige una muestra de tamaño n al azar y con reposición de la población $\Omega = \{\omega_1, \ldots, \omega_N\}$. Según los resultados obtenidos en esta sección, $\overline{\mathcal{X}}_n$ es un estimador insesgado de $\overline{\mathcal{X}}$ en ese muestreo, y $\Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p^w\right) = \Delta_S^2\left(\mathcal{X} \,|\, P\right)/n$.

El empleo de la extensión del Teorema de Rao-Blackwell desarrollada en la Sección 1.2 de este capítulo, permite construir un estimador insesgado con menor S-DCM central.

Concretando más, supongamos ordenado en forma arbitraria el conjunto $\bigcup_{k=1}^{n} \Upsilon_k$ de todas las posibles muestras aleatorias simples de tamaños inferiores o iguales a n a partir de Ω. Sea $M = \mathrm{card}(\bigcup_{k=1}^{n} \Upsilon_k) = \sum_{i=1}^{n} \binom{N}{i}$ y sea $\mathbb{M} = \{1, \ldots, M\}$.

Asociada al espacio probabilístico $(\Upsilon_n^w, \mathcal{P}(\Upsilon_n^w), p^w)$ sobre el que se define $\overline{\mathcal{X}}_n$, puede establecerse a su vez una variable aleatoria real $Y : \Upsilon_n^w \to \mathbb{M}$, que a cada muestra aleatoria con reposición de tamaño n le hace corresponder la posición de la muestra aleatoria simple que determinan las unidades distintas de la primera en la ordenación adoptada.

Si $m \in Y(\Upsilon_n^w)$ y corresponde a una muestra aleatoria simple y_m constituida por k unidades distintas, $\omega_1^*(y_m), \ldots, \omega_k^*(y_m)$, la probabilidad condicionada por m con la que $\omega_i^*(y_m)$ aparece en una muestra de Υ_n^w para la que Y toma el valor m es igual a $1/k$, $i = 1, \ldots, k$, de modo que:

$$\widetilde{E}(\overline{\mathcal{X}}_n \,|\, m) = \frac{1}{k} \odot \sum_{i=1}^{k} \mathcal{X}(\omega_i^*(y_m)).$$

En consecuencia, $\widetilde{E}(\overline{\mathcal{X}}_n \,|\, m)$ equivale al valor esperado condicionado por m del estimador $\overline{\mathcal{X}}_\nu$ que a cada muestra aleatoria con reposición de tamaño n le asocia la media muestral de $\mathcal{X}$ en las unidades distintas de esa muestra (que es una variable aleatoria difusa definida sobre el espacio $(\Upsilon_n^w, \mathcal{P}(\Upsilon_n^w), p^w)$ y cuya distribución depende del *tamaño muestral efectivo* ν, o número de unidades distintas en la muestra considerada -ver, por ejemplo, Thompson 1992-), es decir, se cumple que:

$$\widetilde{E}(\overline{\mathcal{X}}_\nu \,|\, m) = \widetilde{E}(\overline{\mathcal{X}}_n \,|\, m).$$

Por lo tanto, si se considera $\varphi(Y)$ tal que $\varphi(m) = \widetilde{E}(\overline{\mathcal{X}}_n \,|\, m) = \widetilde{E}(\overline{\mathcal{X}}_\nu \,|\, m)$, se satisface que:

$$\widetilde{E}(\varphi(Y) \,|\, p^w) = \bigodot_{k=1}^{n} \left[\bigodot_{m \in Y(\Upsilon_n^w) \,|\, \# y_m = k} \varphi(m) \odot P(Y = m \,|\, \nu = k) \right] \odot P_\nu(k)$$

$$= \bigodot_{k=1}^{n} \widetilde{E}(\overline{\mathcal{X}}_\nu \,|\, \nu = k) \odot P_\nu(k),$$

(donde por $\# y_m = k$ se entiende que y_m tiene k unidades distintas) que, de acuerdo con López-Díaz & Gil (1998b), coincide con $\widetilde{E}(\overline{\mathcal{X}}_\nu \,|\, p^w)$.

Como, en virtud del Teorema 1.5.1, se verifica que $\widetilde{E}(\overline{\mathcal{X}}_\nu \,|\, \nu = k) = \overline{\mathcal{X}}$ para $k = 1, \cdots, n$, entonces:

$$\widetilde{E}(\varphi(Y) \,|\, p^w) = \overline{\mathcal{X}}.$$

Además, el Teorema 1.2.14 garantiza que:

$$\Delta_S^2(\varphi(Y) \,|\, p^w) \leq \Delta_S^2 \left(\overline{\mathcal{X}}_\nu \,|\, p^w \right),$$

y:

$$\Delta_S^2 \left(\overline{\mathcal{X}}_\nu \,|\, p^w \right) = \int_{[0,1]} \int_{[0,1]} \mathrm{Var} \left[f_{\overline{\mathcal{X}}_\nu}(\alpha, \lambda) \right] \, d\,S(\lambda) \, d\,\alpha$$

$$= \int_{[0,1]} \int_{[0,1]} \left[\sum_{k=1}^{n} \mathrm{Var} \left[f_{\overline{\mathcal{X}}_\nu}(\alpha, \lambda) \,|\, \nu = k \right] P_\nu(k) \right.$$

$$\left. + \mathrm{Var} \left(E \left[f_{\overline{\mathcal{X}}_\nu}(\alpha, \lambda) \,|\, \nu \right] \right) \right] d\,S(\lambda) \, d\,\alpha$$

$$= \int_{[0,1]} \int_{[0,1]} \Big[\sum_{k=1}^{n} \mathrm{Var} \left(\overline{(f_{\mathcal{X}}(\alpha, \lambda))}_\nu \,|\, \nu = k \right) P_\nu(k)$$

$$+ \mathrm{Var} \left(E \left(\overline{(f_{\mathcal{X}}(\alpha, \lambda))}_\nu \,|\, \nu \right) \right) \Big] \, d\,S(\lambda) \, d\,\alpha,$$

donde $\overline{(f_{\mathcal{X}}(\alpha, \lambda))}_\nu$ representa la media muestral de $f_{\mathcal{X}}(\alpha, \lambda)$ para las unidades distintas en cada muestra, que de acuerdo con la Teoría de Muestras para variables aleatorias reales (ver, por ejemplo, Raj & Khamis 1958, Thompson 1992, págs. 20, 90) es un estimador insesgado de $\overline{f_{\mathcal{X}}(\alpha, \lambda)} = f_{\overline{\mathcal{X}}}(\alpha, \lambda)$ para cualquier valor de ν, de forma que $\mathrm{Var} \left(E \left(\overline{(f_{\mathcal{X}}(\alpha, \lambda))}_\nu \,|\, \nu \right) \right) = 0$.

Por otro lado, y también según resultados para variables aleatorias reales, se satisface que:

$$\mathrm{Var}\left(\overline{(f_{\mathcal{X}}(\alpha,\lambda))}_\nu \,|\, \nu = k\right) = \left(\frac{1}{k} - \frac{1}{N}\right) \mathrm{Var}\left(f_{\mathcal{X}}(\alpha,\lambda)\right),$$

de donde:

$$\Delta_S^2\left(\overline{\mathcal{X}}_\nu \,|\, p^w\right) = \int_{[0,1]}\int_{[0,1]} \left[\sum_{k=1}^{n}\left(\frac{1}{k} - \frac{1}{N}\right)\mathrm{Var}\left(f_{\mathcal{X}}(\alpha,\lambda)\right) P_\nu(k)\right] d\,S(\lambda)\, d\,\alpha$$

$$= \left[E\left(\frac{1}{\nu}\right) - \frac{1}{N}\right]\frac{N\Delta_S^2(\mathcal{X})}{N-1}.$$

Como, de acuerdo con Raj & Kharmis (1958), se cumple que $E\left(\frac{1}{\nu}\right) \leq \frac{1}{N} + \frac{N-1}{nN}$, con igualdad si, y sólo si, $n = 2$, se cumple que:

$$\Delta_S^2(\varphi(Y) \,|\, p^w) \leq \Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p^w\right),$$

con igualdad si, y sólo si, $n = 2$.

Observación 1.5.1. Los estudios realizados en esta sección sobre la estimación difusa del parámetro difuso valor esperado $\overline{\mathcal{X}} = \widetilde{E}(\mathcal{X} \,|\, P)$ mediante el valor esperado muestral $\overline{\mathcal{X}}_n$, puede completarse con la particularización a poblaciones finitas del resultado establecido por Colubi *et al.* (1999), según el cual $\{\overline{\mathcal{X}}_n\}_n$ sería además una sucesión de estimadores difusos "fuertemente consistentes" (esta consistencia fuerte entendida en el sentido de la convergencia en la métrica d_∞).

Varios de los estudios realizados en las tres últimas secciones van a ilustrarse a continuación mediante un ejemplo.

Ejemplo 1.5.1. Un psicólogo quiere hacer un sondeo sobre las preferencias de edad en los 25.000 habitantes de cierta ciudad con el fin de estimar la media de la edad preferida.

Con este propósito, el psicólogo selecciona una muestra aleatoria simple previa υ_1 de $n_1 = 100$ habitantes de esa ciudad y les pregunta sobre la etapa de su vida que consideran ha sido/es/será la mejor (en el supuesto de que no padezcan ninguna enfermedad durante ese periodo).

Supongamos que las respuestas de las personas encuestadas fueron las siguientes:

15 'cuando se es JOVEN' ($\tilde{x}_1$),
2 'cuando se es MUY JOVEN' ($\tilde{x}_2$),
2 'cuando se es EXTREMADAMENTE JOVEN' ($\tilde{x}_3$),
3 'cuando se es BASTANTE JOVEN' ($\tilde{x}_4$),
10 'cuando se tiene MEDIANA EDAD' ($\tilde{x}_5$),
8 'cuando SE SOBREPASA la MEDIANA EDAD' ($\tilde{x}_6$),
15 'ANTES DE LLEGAR A MEDIANA EDAD' ($\tilde{x}_7$),
10 'cuando se está EN TORNO A la MEDIANA EDAD' ($\tilde{x}_8$),
4 'cuando se tiene APROXIMADAMENTE entre 40 y 50 años' ($\tilde{x}_9$),
7 'cuando se es MAYOR' ($\tilde{x}_{10}$),
2 'cuando se es MUY MAYOR' ($\tilde{x}_{11}$),
1 'cuando se es EXTREMADAMENTE MAYOR' ($\tilde{x}_{12}$),
15 'cuando NO se es MAYOR' ($\tilde{x}_{13}$),
6 'cuando se es BASTANTE MAYOR' ($\tilde{x}_{14}$).

Los valores anteriores para la variable $\mathcal{X}$ = etapa de la vida, sobre la población Ω de los 25.000 habitantes de la ciudad considerada, están claramente mal definidos y no se han establecido universalmente cotas exactas para los mismos. Supongamos que para expresar las respuestas anteriores se usan los valores difusos con soportes en $[0, 100]$ basados en S-curvas (ver, por ejemplo, Cox 1994), así como los modificadores lingüísticos (MUY, ANTES DE LLEGAR A, EN TORNO A, etc.) que se han definido usando las ideas de Bandemer & Gottwald (1995) (de esta forma, los modificadores combinan una "traslación" y un cambio funcional en la forma del valor difuso). En particular, hemos supuesto que $\tilde{x}_i$ se define como sigue:

$$\begin{aligned}
&\tilde{x}_1 = 1 - S(25,40),\\
&\tilde{x}_2(t) = (\tilde{x}_1(t+5))^2,\\
&\tilde{x}_3(t) = (\tilde{x}_1(t+10))^3,\\
&\tilde{x}_4(t) = \sqrt{\tilde{x}_1(t-5)},
\end{aligned}$$

$$\tilde{x}_5 = \begin{cases} S(25,35) & \text{en } [0,35] \\ 1 & \text{en } [35,55] \\ 1 - S(55,65) & \text{en otro caso} \end{cases}$$

$$\begin{aligned}
&\tilde{x}_6 = S(55,65),\\
&\tilde{x}_7 = 1 - S(25,35),
\end{aligned}$$

$$\tilde{x}_8(t) = \begin{cases} (\tilde{x}_5(t-5))^{0,7} & \text{si } t \geq 45 \\ (\tilde{x}_5(t+5))^{0,7} & \text{en otro caso} \end{cases}$$

$$\tilde{x}_9(t) = \begin{cases} (\tilde{x}_5(t+5))^{2} & \text{si } t \geq 45 \\ (\tilde{x}_5(t-5))^{2} & \text{en otro caso} \end{cases}$$

$$\begin{aligned}
&\tilde{x}_{10} = S(60,85),\\
&\tilde{x}_{11}(t) = (\tilde{x}_{10}(t-5))^2,\\
&\tilde{x}_{12}(t) = (\tilde{x}_{10}(t-10))^3,\\
&\tilde{x}_{13}(t) = 1 - \tilde{x}_{10}(t),\\
&\tilde{x}_{14}(t) = \sqrt{\tilde{x}_{10}(t+5)},
\end{aligned}$$

en $[0,100]$, y $\tilde{x}_i$ es nulo en $\mathbb{R}\setminus[0,100]$ (ver Figuras 1.2, 1.3 y 1.4).

Con la información muestral disponible, se puede estimar $\Delta^2_S(\mathcal{X})$ para la medida S_0 discretizada de forma que $L = 3$, $\lambda_1 = 0$, $\lambda_2 = 0,5$, $\lambda_3 = 1$, $k_1 = k_2 = k_3 = 1/3$, y $\overline{g}(\lambda) = 0$ si $\lambda \in (0,1) \setminus \{0,5\}$, en virtud del Teorema 1.3.1, mediante el valor:

$$\widehat{\Delta^2_{S_0}}(\mathcal{X}[\upsilon_1]) = 371,075.$$

De acuerdo con el Teorema 1.5.4, esta estimación preliminar de $\Delta^2_{S_0}(\mathcal{X})$ nos permite determinar el tamaño de muestra adecuado para estimar la

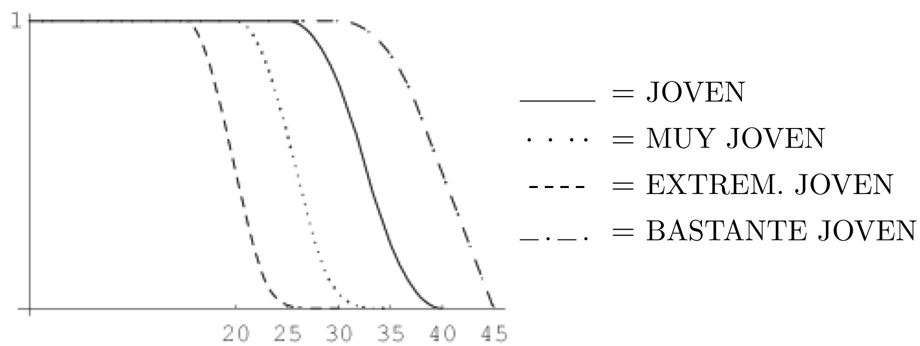


Fig. 1.2: Valor difuso JOVEN y valores relacionados modificados

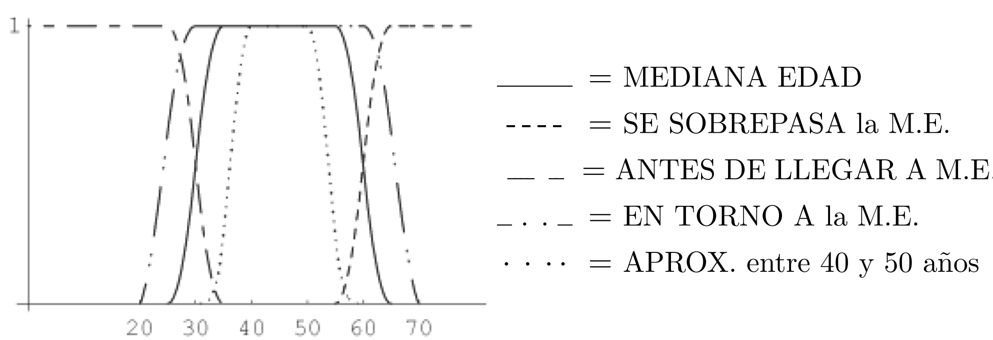


Fig. 1.3: Valor difuso MEDIANA EDAD y valores relacionados modificados

media poblacional de la edad preferida, a un nivel de aceptación $\alpha = 0,05$ y una tolerancia $d =$ '5 años', como el valor:

$$n = \left. \frac{25.000 \cdot 371,075}{24.999 \cdot 25 \cdot 0,05 + 371,075} \right] = 294.$$

Supongamos que se extrae de la población una nueva muestra aleatoria simple de tamaño 194, que se añade a la muestra anterior, y que las respuestas a la pregunta sobre la edad preferida para la muestra global, υ, son: 64 'cuando se es JOVEN', 10 'cuando se es MUY JOVEN', 4 'cuando se es EXTREMADAMENTE JOVEN', 39 'cuando se es BASTANTE JOVEN', 30 'cuando se tiene MEDIANA EDAD', 15 'cuando SE SOBREPASA la MEDIANA EDAD', 17 'ANTES DE LLEGAR A MEDIANA EDAD', 34

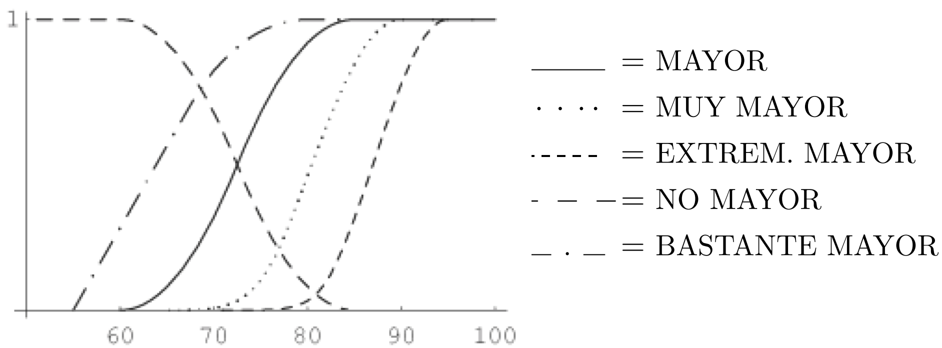


Fig. 1.4: Valor difuso MAYOR y valores relacionados modificados

'cuando se está EN TORNO A la MEDIANA EDAD', 19 'cuando se tiene APROXIMADAMENTE entre 40 y 50 años', 18 'cuando se es MAYOR', 4 'cuando se es MUY MAYOR', 3 'cuando se es EXTREMADAMENTE MAYOR', 26 'cuando NO se es MAYOR', y 11 'cuando se es BASTANTE MAYOR'.

La edad preferida esperada en la población se estima, según el Teorema 1.5.1, por el valor esperado en la muestra global de $n = 294$ habitantes, que se presenta en la Figura 1.5.

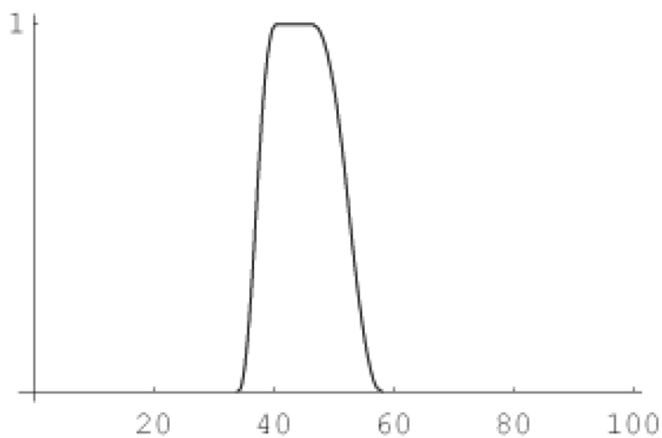


Fig. 1.5: Media muestral de la "edad preferida"

Además, "la variabilidad de las preferencias" se puede aproximar por la estimación final de $\Delta^2_{S_0}(\mathcal{X})$ en la muestra aleatoria simple global, que de nuevo según el Teorema 1.3.1 viene dado por:

$$\widehat{\Delta^2_{S_0}}(\mathcal{X}[\upsilon]) = 315,317.$$

Con el fin de completar el estudio de la estimación de la variación absoluta poblacional de las preferencias, puede utilizarse el resultado del Teorema 1.4.3 para obtener un intervalo de confianza con coeficiente $1-\alpha$, que vendrá dado para υ por:

$$\left[\Delta^2_{S_0}(\mathcal{X}[\upsilon]) - z_\alpha\sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}},\ \Delta^2_{S_0}(\mathcal{X}[\upsilon]) + z_\alpha\sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}}\right].$$

Si se considera, por ejemplo, $\alpha = 0,5$, el intervalo de confianza con coeficiente aproximadamente igual a $= 0,95$ correspondería a:

$$\left[315,317 - 1,96\sqrt{\frac{115978,04585}{294}},\ 315,317 + 1,96\sqrt{\frac{115978,04585}{294}}\right]$$

$$= [276,388,\ 354,245].$$

1.6 Aplicación de la medida de variación absoluta a la regresión lineal y general con variables aleatorias difusas

Otro problema estadístico en el que se puede aplicar una medida inspirada en la S-dispersión cuadrática media, es el de la *predicción* del valor de una variable aleatoria difusa dado el valor de otra mediante alguna transformación.

Este problema es el objetivo del llamado Análisis de Regresión con datos difusos, que es un tema que se ha estudiado exhaustivamente. La literatura sobre el mismo es muy variada, y en ella se distinguen diferentes aproximaciones, dependiendo de los elementos del problema que están afectados por la borrosidad, del modelo adoptado para tal problema y del tratamiento seguido para su resolución.

Las dos aproximaciones que han recibido mayor atención investigadora son la que trata el problema de regresión lineal difusa como un problema de Programación Lineal, y la que se basa en una extensión del criterio clásico de Mínimos Cuadrados.

En relación con la primera aproximación, cabe destacar los trabajos de Tanaka *et al.* (1980, 1982, entre otros), en los que las variables aleatorias involucradas se suponen reales o vectoriales y los parámetros de la relación funcional (de tipo lineal) se consideran difusos. La selección de los valores de estos parámetros difusos, y de las funciones implicadas en el problema de Programación Lineal que formaliza los objetivos del problema de regresión, son cruciales en la resolución del último. Moskowitz & Kim (1993), y Redden & Woodall (1994, 1996) analizan el efecto de tal selección en la solución del problema de Programación Lineal correspondiente (ver, por ejemplo, Corral *et al.* 1998, Kruse *et al.* 1999, para un breve resumen sobre este y otros estudios).

La aproximación consistente en extender el método de Mínimos Cuadrados, depende principalmente a su vez de la extensión que se considere para la distancia entre los valores observados y esperados por la relación del modelo.

Diamond (1987, 1988, 1990) ha desarrollado el método de Mínimos Cuadrados que recurre a la métrica δ_2 cuando las variables de predicción y respuesta toman como valores números difusos triangulares, y los parámetros del ajuste son números reales. La proyección sobre un cono o subespacio lineal cerrado, permite utilizar las técnicas del método de Mínimos Cuadrados con variables aleatorias reales.

Näther (1997) y Körner (1997ab) han abordado el problema de regresión lineal cuando se considera que la variable de predicción es una variable aleatoria real (o un vector aleatorio clásico) y la variable respuesta es una variable aleatoria difusa, con la suposición de que el término independiente de la relación lineal es también una variable aleatoria difusa, mientras que el parámetro que multiplica a la variable de predicción es real o vectorial.

Salas (1991) y Bertoluzza *et al.* (1995ab) han introducido la métrica D_S y desarrollado la extensión del criterio de Mínimos Cuadrados cuando se considera un número finito de pares de datos difusos y para cierta selección de la función S.

Guo & Chen (1992) han tratado el problema de la regresión múltiple difusa con vectores de números difusos triangulares y el mejor ajuste se determina minimizando los cuadrados de las diferencias de las funciones que representan los datos difusos correspondientes.

Para una revisión más profunda y detallada de los modelos y métodos existentes sobre el Análisis de Regresión con elementos difusos, pueden verse la monografía de Kacprzyk & Fedrizzi (1992) y la recopilación recientemente elaborada por Diamond & Tanaka (1998).

En esta sección se desarrolla un estudio que generaliza el de Salas (1991) y Bertoluzza *et al.* (1995ab), en el que la borrosidad se supone que aparece en la variable de predicción y en la variable respuesta y tal borrosidad se va a representar mediante la consideración del modelo establecido por las variables aleatorias difusas para ambas, y o bien una relación lineal con parámetros reales o una relación funcional general. En este estudio, se presenta una extensión del método de Mínimos Cuadrados, midiéndose el error entre los valores observados y estimados difusos mediante el valor medio de la S-dispersión cuadrática media asociado a la variable aleatoria difusa respuesta respecto a cada uno de los valores de la de predicción.

Con el fin de simplificar las notaciones, se emplearán las funciones $V_S : \mathcal{F}_c(\mathbb{R}) \to \mathbb{R}$, $W_S : \mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \to \mathbb{R}$ y $W'_S : \mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \to \mathbb{R}$ definidas de forma que:

$$V_S\left(\widetilde{A}\right) = \int_{[0,1]} \int_{[0,1]} f_{\widetilde{A}}(\alpha, \lambda)\, d\,S(\lambda)\, d\,\alpha,$$

$$W_S\left(\widetilde{A}, \widetilde{B}\right) = \int_{[0,1]} \int_{[0,1]} f_{\widetilde{A}}(\alpha, \lambda)\, f_{\widetilde{B}}(\alpha, \lambda)\, d\,S(\lambda)\, d\alpha,$$

y

$$W'_S\left(\widetilde{A},\widetilde{B}\right)=\int_{[0,1]}\int_{[0,1]} f_{\widetilde{A}}(\alpha,1-\lambda)\, f_{\widetilde{B}}(\alpha,\lambda)\; d\,S(\lambda)\, d\alpha.$$

Sean $\mathcal{X}$ e $\mathcal{Y}$ dos variables aleatorias difusas integrablemente acotadas asociadas al espacio de probabilidad $(\Omega,\mathcal{A},P)$ y consideremos en primer lugar una *relación lineal* de la variable aleatoria difusa $\mathcal{Y}$ respecto a $\mathcal{X}$ del tipo:

$$\mathcal{Y}=(a\odot\mathcal{X})\oplus b,$$

donde a y b son números reales.

Para explicar el significado de una relación lineal como la anterior se puede considerar el ejemplo recogido en las Figuras 1.6 y 1.7.

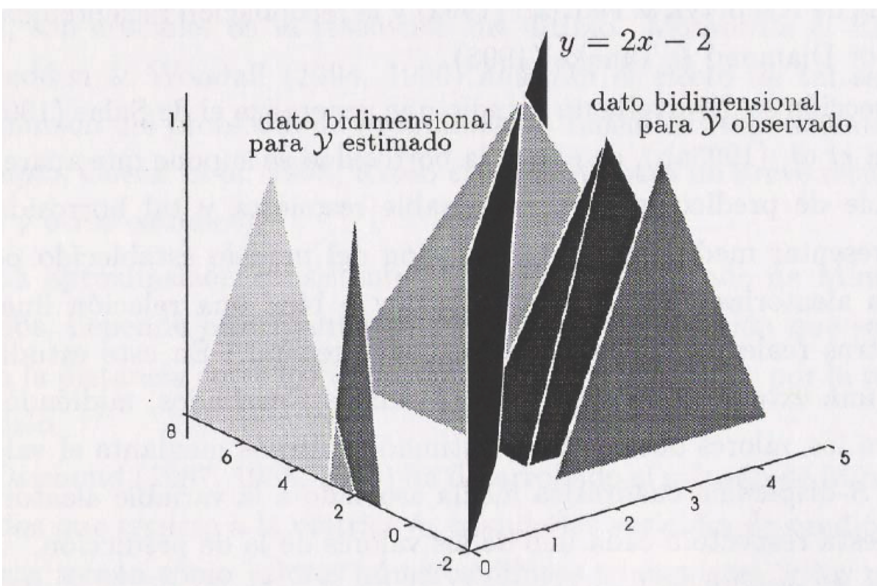


Fig. 1.6: Datos difusos bidimensionales para los valores de $\mathcal{Y}$ servado y estimado (por la relación lineal $(2\odot\mathcal{X})\oplus(-2)$)

En la Figura 1.6 se muestra el caso en el que se considera por un lado el dato difuso bidimensional correspondiente al producto cartesiano del valor (observado) de $\mathcal{X}$ (dado por el número difuso triangular Tri(2,3,5,5)) y el valor observado de $\mathcal{Y}$ (Tri(1,2,3)) y, por otro, el dato difuso bidimensional correspondiente al producto cartesiano del mismo valor de $\mathcal{X}$ y el valor estimado de $\mathcal{Y}$ por la relación lineal $\mathcal{Y} = (2 \odot \mathcal{X}) \oplus (-2)$ (que será el número difuso Tri(2,5,8)).

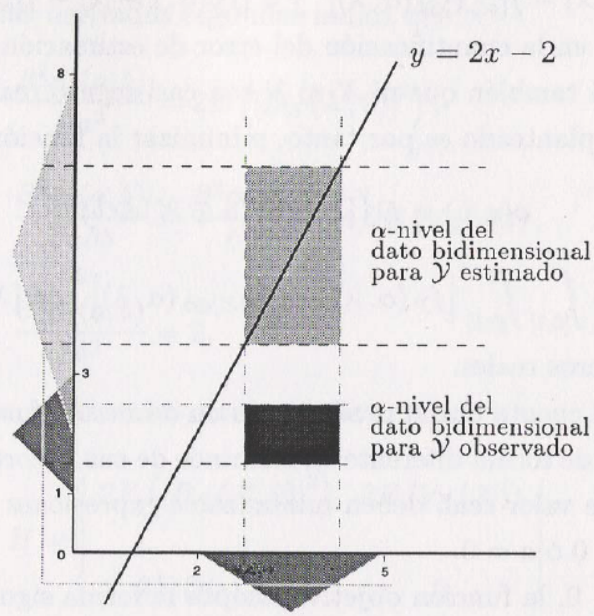


Fig. 1.7: α-nivel del dato difuso bidimensional observado y estimado

La Figura 1.7 examina la situación de la figura anterior al nivel $\alpha = 0,5$.

El objetivo del *Análisis de Regresión Lineal* para esta situación consiste en determinar los valores de los parámetros a y b que minimicen el error asociado a la relación lineal $\mathcal{Y} = (a \odot \mathcal{X}) \oplus b$. Ese error va a cuantificarse en la presente sección mediante la S-DCM o, más concretamente, una medida definida en forma análoga, ya que la aleatoriedad aparece ahora involucrada en las dos variables que intervienen en ese error, y se supondrá habitualmente que la medida S es simétrica en el sentido de que la función de densidad asociada g sea una función simétrica respecto al valor $\lambda = 0,5$, es decir, $g(\lambda) = g(1 - \lambda)$.

La suposición precedente no entraña una pérdida de generalidad sustancial, ya que el error asociado a la relación lineal $\mathcal{Y} = (a \odot \mathcal{X}) \oplus b$ se corresponde en el problema de regresión con una medida de la "distancia" entre los valores observados para $\mathcal{Y}$ y los esperados por la relación lineal $(a\odot\mathcal{X})\oplus b$. En este sentido, no parece haber razones que justifiquen que los pesos asignados a $[f_{\mathcal{Y}}(\alpha,\lambda)-f_{(a\odot\mathcal{X})\oplus b}(\alpha,\lambda)]^2$ y a $[f_{\mathcal{Y}}(\alpha,1-\lambda)-f_{(a\odot\mathcal{X})\oplus b}(\alpha,1-\lambda)]^2$ sean diferentes en la cuantificación del error de estimación.

Se supondrá también que ni $\mathcal{X}$ ni $\mathcal{Y}$ son casi seguro reales.

El objetivo planteado es por tanto, minimizar la función:

$$\phi(a,b)=E\left([D_S(\mathcal{Y},(a\odot\mathcal{X})\oplus b)]^2\right)$$

$$=E\bigg(\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{Y}}(\alpha,\lambda)-f_{(a\odot\mathcal{X})\oplus b}(\alpha,\lambda)\right]^2\,d\,S(\lambda)\,d\,\alpha\bigg),$$

con a y b números reales.

Teniendo en cuenta que el producto de un número difuso por un número real se expresa de forma diferente en términos de sus α-cortes, dependiendo del signo de ese valor real, deben minimizarse expresiones diferentes según que $a>0$, $a<0$ ó $a=0$.

Cuando $a\geq 0$, la función objetivo adopta la forma siguiente:

$$\phi_1(a,b)=E\bigg(\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{Y}}(\alpha,\lambda)-f_{(a\odot\mathcal{X})\oplus b}(\alpha,\lambda)\right]^2\,d\,S(\lambda)\,d\,\alpha\bigg)$$

$$=E\bigg(\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{Y}}(\alpha,\lambda)-a\,f_{\mathcal{X}}(\alpha,\lambda)\ -\ b\right]^2\,d\,S(\lambda)\,d\,\alpha\bigg)$$

$$=E\left([D_S(\mathcal{Y},0)]^2\right)+a^2E\left([D_S(\mathcal{X},0)]^2\right)$$

$$+\,b^2+2abE\left(V_S(\mathcal{X})\right)-2bE\left(V_S(\mathcal{Y})\right)-2aE\left(W_S(\mathcal{X},\mathcal{Y})\right).$$

La minimización de ϕ_1 cuando $a>0$ se lleva a cabo como sigue:

$$\frac{\partial\phi_1(a,b)}{\partial a}=2aE\left([D_S(\mathcal{X},0)]^2\right)+2bE\left(V_S(\mathcal{X})\right)-2E\left(W_S(\mathcal{X},\mathcal{Y})\right)=0,$$

$$\frac{\partial\phi_1(a,b)}{\partial b}=2b+2aE\left(V_S(\mathcal{X})\right)-2E\left(V_S(\mathcal{Y})\right)=0,$$

que conduce al sistema de ecuaciones cuya solución viene dada por:

$$a_1=\frac{E\left(W_S(\mathcal{X},\,\mathcal{Y})\right)-E\left(V_S(\mathcal{X})\right)E\left(V_S(\mathcal{Y})\right)}{E\left([D_S(\mathcal{X},0)]^2\right)-\left[E\left(V_S(\mathcal{X})\right)\right]^2},$$

$$b_1 = E\left(V_S\left(\mathcal{Y}\right)\right) - a_1 E\left(V_S\left(\mathcal{X}\right)\right),$$

y será válida si $a_1 > 0$.

El cálculo de las derivadas segundas indica que:

$$\frac{\partial^2 \phi_1(a,b)}{\partial a^2} = 2E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right),$$

$$\frac{\partial^2 \phi_1(a,b)}{\partial a \partial b} = \frac{\partial^2 \phi_1(a,b)}{\partial b \partial a} = 2E\left(V_S\left(\mathcal{X}\right)\right),$$

$$\frac{\partial^2 \phi_1(a,b)}{\partial b^2} = 2,$$

y, en consecuencia, el hessiano correspondiente será:

$$H = \begin{vmatrix} 2E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right) & 2E\left(V_S\left(\mathcal{X}\right)\right) \\ & \\ 2E\left(V_S\left(\mathcal{X}\right)\right) & 2 \end{vmatrix}$$

$$= 4\left\{E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right) - \left[E\left(V_S\left(\mathcal{X}\right)\right)\right]^2\right\}.$$

Para que la solución (a_1, b_1) sea un mínimo absoluto, el hessiano H tiene que ser positivo. La concavidad hacia arriba de $h(x) = x^2$ y la aplicación de la Desigualdad de Jensen permite concluir que:

$$\left[E\left(V_S\left(\mathcal{X}\right)\right)\right]^2 = \left[E\left(\int_{[0,1]}\int_{[0,1]} f_{\mathcal{X}}\left(\alpha,\lambda\right)\, d\,S(\lambda)\, d\,\alpha\right)\right]^2$$

$$\leq E\left[\left(\int_{[0,1]}\int_{[0,1]} f_{\mathcal{X}}\left(\alpha,\lambda\right)\, d\,S(\lambda)\, d\,\alpha\right)^2\right]$$

$$\leq E\left(\int_{[0,1]}\left[\int_{[0,1]} f_{\mathcal{X}}\left(\alpha,\lambda\right)\, d\,S(\lambda)\right]^2 d\,\alpha\right)$$

$$\leq E\left(\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{X}}\left(\alpha,\lambda\right)\right]^2 d\,S(\lambda)\, d\,\alpha\right) = E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right).$$

Para que H sea nulo, todas las desigualdades anteriores deben reducirse a igualdades, y en virtud de la Desigualdad de Jensen estas igualdades se alcanzan si, y sólo si, $f_{\mathcal{X}(\cdot)}(\cdot,\cdot)$ es constante casi seguro $[P \otimes m \otimes S]$, lo que equivale a que $\mathcal{X}$ sea una variable aleatoria difusa degenerada en un valor real.

La solución para el caso $a = 0$ será $(0, E(V_S(\mathcal{Y}))$.

Análogamente, si $a < 0$ la función objetivo será:

$$\phi_2(a,b) = E\Big(\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{Y}}(\alpha,\lambda) - f_{(a\odot\mathcal{X})\oplus b}(\alpha,\lambda)\right]^2 d\,S(\lambda)\,d\,\alpha\Big)$$

$$= E\Big(\int_{[0,1]}\int_{[0,1]}[f_{\mathcal{Y}}(\alpha,\lambda) - a\,f_{\mathcal{X}}(\alpha,1-\lambda)\,-\,b]^2\,d\,S(\lambda)\,d\,\alpha\Big).$$

Por la simetría de la medida S se tiene que:

$$E(V_S(\mathcal{X})) = E\Big(\int_{[0,1]}\int_{[0,1]} f_{\mathcal{X}}(\alpha,\lambda)\;d\,S(\lambda)\,d\,\alpha\Big)$$

$$= E\Big(\int_{[0,1]}\int_{[0,1]} f_{\mathcal{X}}(\alpha,1-\lambda)\;d\,S(\lambda)\,d\,\alpha\Big),$$

$$E\left([D_S(\mathcal{X},0)]^2\right) = E\Big(\int_{[0,1]}\int_{[0,1]}[f_{\mathcal{X}}(\alpha,\lambda)]^2\;d\,S(\lambda)\,d\,\alpha\Big)$$

$$= E\Big(\int_{[0,1]}\int_{[0,1]}[f_{\mathcal{X}}(\alpha,1-\lambda)]^2\;d\,S(\lambda)\,d\,\alpha\Big),$$

de donde:

$$\phi_2(a,b) = E\left([D_S(\mathcal{Y},0)]^2\right) + a^2 E\left([D_S(\mathcal{X},0)]^2\right)$$

$$+ b^2 + 2abE(V_S(\mathcal{X})) - 2bE(V_S(\mathcal{Y})) - 2aE(W'_S(\mathcal{X},\mathcal{Y})).$$

Las derivadas parciales de primer orden de ϕ_2 vienen dadas por:

$$\frac{\partial\phi_2(a,b)}{\partial a} = 2aE\left([D_S(\mathcal{X},0)]^2\right) + 2bE(V_S(\mathcal{X})) - 2E(W'_S(\mathcal{X},\mathcal{Y})),$$

$$\frac{\partial\phi_2(a,b)}{\partial b} = 2b + 2aE(V_s(\mathcal{X})) - 2E(V_S(\mathcal{Y})),$$

y la solución del sistema formado por ambas viene dada por:

$$a_2 = \frac{E\left(W'_S\left(\mathcal{X},\mathcal{Y}\right)\right) - E\left(V_S\left(\mathcal{X}\right)\right)E\left(V_S\left(\mathcal{Y}\right)\right)}{E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right) - \left[E\left(V_S\left(\mathcal{X}\right)\right)\right]^2},$$

$$b_2 = E\left(V_S\left(\mathcal{Y}\right)\right) - a_2 E\left(V_S\left(\mathcal{X}\right)\right),$$

y será válida si $a_2 < 0$.

Las segundas derivadas vienen dadas por:

$$\frac{\partial^2 \phi_2(a,b)}{\partial a^2} = 2E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right),$$

$$\frac{\partial^2 \phi_2(a,b)}{\partial a \partial b} = \frac{\partial^2 \phi_2(a,b)}{\partial b \partial a} = 2E\left(V_S\left(\mathcal{X}\right)\right),$$

$$\frac{\partial^2 \phi_2(a,b)}{\partial b^2} = 2,$$

con lo que se obtiene el hessiano:

$$H = 4\left\{E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right) - \left[E\left(V_S\left(\mathcal{X}\right)\right)\right]^2\right\},$$

que resulta positivo siempre que $\mathcal{X}$ no sea una variable aleatoria difusa degenerada en un valor real.

Si g es simétrica respecto a $\lambda = 0,5$ se tiene que:

$$E\left(W_S\left(\mathcal{X},\mathcal{Y}\right)\right) - E\left(W'_S\left(\mathcal{X},\mathcal{Y}\right)\right)$$

$$= E\left(\int_{[0,1]}\int_{[0,1]} g(\lambda) f_{\mathcal{Y}}\left(\alpha,\lambda\right)\left[f_{\mathcal{X}}\left(\alpha,\lambda\right) - f_{\mathcal{X}}\left(\alpha,1-\lambda\right)\right]\, d\lambda\, d\alpha\right)$$

$$= E\left(\int_{[0,1]}\int_{[0,\,0,5]} g(\lambda) f_{\mathcal{Y}}\left(\alpha,\lambda\right)\left[f_{\mathcal{X}}\left(\alpha,\lambda\right) - f_{\mathcal{X}}\left(\alpha,1-\lambda\right)\right]\, d\lambda\, d\alpha\right.$$

$$\left.+\int_{[0,1]}\int_{[0,5,\,1]} g(1-\lambda) f_{\mathcal{Y}}\left(\alpha,\lambda\right)\left[f_{\mathcal{X}}\left(\alpha,\lambda\right) - f_{\mathcal{X}}\left(\alpha,1-\lambda\right)\right]\, d\lambda\, d\alpha\right)$$

$$= E\Bigg(\int_{[0,1]}\int_{[0,\,0,5]} g(\lambda)\Big[f_{\mathcal{Y}}\left(\alpha,\lambda\right) f_{\mathcal{X}}\left(\alpha,\lambda\right) + f_{\mathcal{Y}}\left(\alpha,1-\lambda\right) f_{\mathcal{X}}\left(\alpha,1-\lambda\right)$$

$$-f_{\mathcal{Y}}(\alpha,\lambda)\, f_{\mathcal{X}}(\alpha,1-\lambda) - f_{\mathcal{Y}}(\alpha,1-\lambda)\, f_{\mathcal{X}}(\alpha,\lambda)\Big]\, d\lambda\, d\alpha\Big)$$

$$= K(S)\int_{[0,1]} E\left((\sup \mathcal{X}_\alpha - \inf \mathcal{X}_\alpha)(\sup \mathcal{Y}_\alpha - \inf \mathcal{Y}_\alpha)\right)\, d\,\alpha,$$

con $K(S) = \dfrac{1}{2}\displaystyle\int_{[0,1]} (1-2\lambda)^2\, d\,S(\lambda) > 0$.

Por lo tanto, como $(\sup \mathcal{X}_\alpha - \inf \mathcal{X}_\alpha)(\sup \mathcal{Y}_\alpha - \inf \mathcal{Y}_\alpha) \geq 0$ para todo $\alpha \in [0,1]$ y $\omega \in \Omega$, se tiene que:

$$E\left(W_S\left(\mathcal{X},\mathcal{Y}\right)\right) \geq E\left(W'_S\left(\mathcal{X},\mathcal{Y}\right)\right),$$

con igualdad si, y sólo si, $\sup \mathcal{X}_{(\,\cdot\,)}(\,\cdot\,) = \inf \mathcal{X}_{(\,\cdot\,)}(\,\cdot\,)$ y $\sup \mathcal{Y}_{(\,\cdot\,)}(\,\cdot\,) = \inf \mathcal{Y}_{(\,\cdot\,)}(\,\cdot\,)$ c.s. $[m \otimes P]$.

En virtud de la continuidad por la izquierda de los extremos de los α-cortes de los elementos de $\mathcal{F}_c(\mathbb{R})$ respecto a α, para cualquier $\omega \in \Omega$ tal que $\sup \mathcal{X}_{(\,\cdot\,)}(\omega) = \inf \mathcal{X}_{(\,\cdot\,)}(\omega)$ c.s. $[m]$ se cumple que $\sup \mathcal{X}_\alpha(\omega) = \inf \mathcal{X}_\alpha(\omega)$ para todo $\alpha \in [0,1]$ y, análogamente ocurre para $\mathcal{Y}$. En consecuencia, $E\left(W_S\left(\mathcal{X},\mathcal{Y}\right)\right) = E\left(W'_S\left(\mathcal{X},\mathcal{Y}\right)\right)$ si, y sólo si, $\mathcal{X}$ e $\mathcal{Y}$ toman valores en $\mathbb{R}$ c.s. $[P]$.

La última afirmación indica que, salvo para variables aleatorias difusas casi seguro reales, se cumple que el valor a_1 es superior al valor a_2, y por lo tanto y debido a la continuidad de la función $\phi(a,b)$ se puede concluir que el mínimo de $\phi(a,b)$ no puede alcanzarse en un punto (a^*,b^*) con $a^* = 0$, ya que $a_1 \leq 0$ obligaría a $a_2 < 0$ y $a_2 \geq 0$ obligaría a $a_1 > 0$, con lo que el mínimo de $\phi(a,b)$ se alcanzaría en el primer caso para $a^* = a_2$ y en el segundo para $a^* = a_1$.

Las conclusiones precedentes se resumen en el resultado siguiente:

Teorema 1.6.1. *Si $\mathcal{X}$ e $\mathcal{Y}$ son dos variables aleatorias difusas integrablemente acotadas que no son casi seguro reales, y S es una medida cuya densidad asociada g respecto a la medida de Lebesgue es simétrica en $\lambda = 0,5$, entonces la función $\phi(a,b) = E\left(\left[D_S\left(\mathcal{Y},(a \odot \mathcal{X}) \oplus b\right)\right]^2\right)$ alcanza su valor mínimo en el punto (a^*,b^*) con $a^* \neq 0$, tal que:*

$$\phi(a^*,b^*) = \min\{\phi(a_1,b_1), \phi(a_2,b_2)\},$$

donde:

$$a^* = \frac{E\left(W_S\left(\mathcal{X},\, \mathcal{Y}\right)\right) - E\left(V_S\left(\mathcal{X}\right)\right) E\left(V_S\left(\mathcal{Y}\right)\right)}{E\left(\left[D_S\left(\mathcal{X}, 0\right)\right]^2\right) - \left[E\left(V_S\left(\mathcal{X}\right)\right)\right]^2},$$

$$b^* = E\left(V_S\left(\mathcal{Y}\right)\right) - a^* E\left(V_S\left(\mathcal{X}\right)\right),$$

si o bien se cumple que $E\left(W'_S\left(\mathcal{X},\, \mathcal{Y}\right)\right) \geq E\left(V_S\left(\mathcal{X}\right)\right)\ E\left(V_S\left(\mathcal{Y}\right)\right),$ *o bien* $E\left(W_S\left(\mathcal{X},\, \mathcal{Y}\right)\right) > E\left(V_S\left(\mathcal{X}\right)\right) E\left(V_S\left(\mathcal{Y}\right)\right) > E\left(W'_S\left(\mathcal{X},\, \mathcal{Y}\right)\right),$ *y:*

$$a^* = \frac{E\left(W'_S\left(\mathcal{X}, \mathcal{Y}\right)\right) - E\left(V_S\left(\mathcal{X}\right)\right) E\left(V_S\left(\mathcal{Y}\right)\right)}{E\left(\left[D_S\left(\mathcal{X}, 0\right)\right]^2\right) - \left[E\left(V_S\left(\mathcal{X}\right)\right)\right]^2}$$

$$b^* = E\left(V_S\left(\mathcal{Y}\right)\right) - a^* E\left(V_S\left(\mathcal{X}\right)\right),$$

si o bien se cumple que $E\left(V_S\left(\mathcal{X}\right)\right)\ E\left(V_S\left(\mathcal{Y}\right)\right) \geq E\left(W_S\left(\mathcal{X},\, \mathcal{Y}\right)\right),$ *o bien* $E\left(W_S\left(\mathcal{X},\, \mathcal{Y}\right)\right) > E\left(V_S\left(\mathcal{X}\right)\right)\ E\left(V_S\left(\mathcal{Y}\right)\right) > E\left(W'_S\left(\mathcal{X},\, \mathcal{Y}\right)\right).$

Supongamos ahora que $\mathcal{X}$ e $\mathcal{Y}$ son dos variables aleatorias difusas integrablemente acotadas asociadas al espacio de probabilidad $(\Omega, \mathcal{A}, P)$ y que se considera una relación funcional general entre $\mathcal{X}$ e $\mathcal{Y}$, del tipo:

$$\mathcal{Y} = h(\mathcal{X})$$

con $h:\ \mathcal{F}_c(\mathbb{R}) \ \rightarrow\ \mathcal{F}_c(\mathbb{R})$ tal que h sea $(\mathcal{B}_{d_\infty}, \mathcal{B}_{d_\infty})$-medible, de modo que $h(\mathcal{X})$ sea en consecuencia variable aleatoria difusa.

El objetivo del *Análisis de Regresión Funcional general* para esta situación consiste en determinar la función h^* que minimice respecto a h el error asociado a la relación $\mathcal{Y} = h(\mathcal{X})$, admitiéndose que el error se mide a través del valor:

$$F(h) = E\left(\left[D_S\left(\mathcal{Y}, h(\mathcal{X})\right)\right]^2 \,\middle|\, P\right)$$

$$= E\left(\int_{[0,1]} \int_{[0,1]} \left[f_{\mathcal{Y}}(\alpha, \lambda) - f_{h(\mathcal{X})}(\alpha, \lambda)\right]^2 \, d\,S(\lambda)\, d\,\alpha \,\middle|\, P\right).$$

Si se aplica el Teorema 0.3.1 entonces para todo $\alpha \in [0,1]$ y $\lambda \in [0,1]$, se tiene que:

$$E\left(\left[f_{\mathcal{Y}}(\alpha,\lambda) - f_{h(\mathcal{X})}(\alpha,\lambda)\right]^2 \,\middle|\, P\right)$$

$$= E\left(E\left(\left[f_{\mathcal{Y}}(\alpha,\lambda) - f_{h(\tilde{x})}(\alpha,\lambda)\right]^2 \,\middle|\, P_{\tilde{x}}\right) \,\middle|\, P_{\mathcal{X}}\right),$$

y, por lo tanto:

$$F(h) = E\left(DCM_S(\mathcal{Y}, h(\tilde{x})) \,\middle|\, P_{\mathcal{X}}\right),$$

con $DCM_S(\mathcal{Y}, h(\tilde{x}))$ la S-dispersión cuadrática media asociada a $\mathcal{Y}$ respecto a $h(\tilde{x})$ sobre el espacio probabilístico condicionado por $\tilde{x}$.

Fijado $\tilde{x} \in \mathcal{X}(\Omega)$ se tiene que, en virtud del Teorema 1.2.7, $DCM_S(\mathcal{Y}, h(\tilde{x}))$ es mínima para $h^*(\tilde{x}) = \widetilde{E}(\mathcal{Y} \,|\, \mathcal{X} = \tilde{x})$, de donde se concluye que:

Teorema 1.6.2. *Si $\mathcal{X}$ e $\mathcal{Y}$ son dos variables aleatorias difusas integrablemente acotadas y S es una medida que satisface las condiciones de la Definición* 0.1.5, *entonces la función F anterior, definida sobre el conjunto $\mathcal{H} = \{h : \mathcal{F}_c(\mathbb{R}) \to \mathcal{F}_c(\mathbb{R}) \,|\, h(\mathcal{X})$ variable aleatoria difusa$\}$, alcanza su valor mínimo para la función $h^* : \mathcal{F}_c(\mathbb{R}) \to \mathcal{F}_c(\mathbb{R})$ tal que:*

$$h^*(\tilde{x}) = \widetilde{E}(\mathcal{Y} \,|\, \mathcal{X} = \tilde{x}),$$

para cada $\tilde{x} \in \mathcal{X}(\Omega)$.

Obviamente, las funciones $h_1 : \mathcal{F}_c(\mathbb{R}) \to \mathcal{F}_c(\mathbb{R})$ tales que $h_1(\widetilde{V}) = (a \odot \widetilde{V}) \oplus b$ para todo $\widetilde{V} \in \mathcal{F}_c(\mathbb{R})$, con a y b valores reales arbitrarios, cumplen que $h_1 \in \mathcal{H}$, de modo que el óptimo establecido en el Teorema 1.6.2 será al menos "tan bueno" como el del Teorema 1.6.1.

Observación 1.6.1. La relación general óptima no depende de la elección de la medida S.

Observación 1.6.2. En principio el estudio último sólo permite hacer predicciones de valores de $\mathcal{Y}$ para valores observados de $\mathcal{X}$ que hayan aparecido en la población (o muestra) considerada. En la Sección 1.7 se hacen algunos comentarios sobre posibles formas de solventar este inconveniente.

Los resultados de los Teoremas 1.6.1 y 1.6.2, así como su interés práctico, se ilustran a continuación mediante un ejemplo real, cuyos datos fueron proporcionados por miembros del Servicio de Conservación de la Naturaleza, Dirección Regional de Montes y Medio Ambiente de la Consejería de Agricultura del Principado de Asturias.

Ejemplo 1.6.1. Se considera la población Ω de 50 días de un determinado año, en la cual se ha observado las variables NUBOSIDAD ($\mathcal{X}$) y VISIBILIDAD ($\mathcal{Y}$).

La variable $\mathcal{X}$ toma los valores SOLEADO ($\tilde{x}_1$), DESPEJADO ($\tilde{x}_2$), NUBES y CLAROS ($\tilde{x}_3$), NUBOSO ($\tilde{x}_4$) y TOTALMENTE CUBIERTO ($\tilde{x}_5$), y la variable $\mathcal{Y}$ toma los valores PERFECTA ($\tilde{y}_1$), BUENA ($\tilde{y}_2$), REGULAR ($\tilde{y}_3$), MALA ($\tilde{y}_4$) y MUY MALA ($\tilde{y}_5$). Expertos en la medición de estos valores los han descrito en términos de los siguientes números difusos con soportes en $[0, 100]$ (entendidos como porcentajes) basados en S- y Pi-curvas, y en números difusos triangulares y trapezoidales, con soporte contenido estrictamente en $[0, 100]$ siguientes (ver Figuras 1.8 y 1.9):

$$\tilde{x}_1 = S(0, 20),$$

$$\tilde{x}_2 = \Pi(10, 20, 40),$$

$$\tilde{x}_3 = \begin{cases} S(20, 40) & \text{en } [20, 40] \\ 1 & \text{en } [40, 50] \\ 1 - S(50, 70) & \text{en } [50, 70] \\ 0 & \text{en otro caso} \end{cases}$$

$$\tilde{x}_4 = \begin{cases} S(60, 70) & \text{en } [60, 70] \\ 1 & \text{en } [70, 80] \\ 1 - S(80, 90) & \text{en } [80, 90] \\ 0 & \text{en otro caso} \end{cases}$$

$$\tilde{x}_5 = 1 - S(80, 100),$$

$$\tilde{y}_1 = \mathrm{Tra}(90, 95, 100, 100),$$

$$\tilde{y}_2 = \mathrm{Tri}(70, 90, 100),$$

$$\tilde{y}_3 = \begin{cases} S(40,50) & \text{en } [40,50] \\ 1 & \text{en } [50,70] \\ 1 - S(70,80) & \text{en } [70,80] \\ 0 & \text{en otro caso} \end{cases}$$

$$\tilde{y}_4 = \begin{cases} S(20,30) & \text{en } [20,30] \\ 1 & \text{en } [30,40] \\ 1 - S(40,50) & \text{en } [40,50] \\ 0 & \text{en otro caso} \end{cases}$$

$\tilde{y}_5 = S(0,20)$.

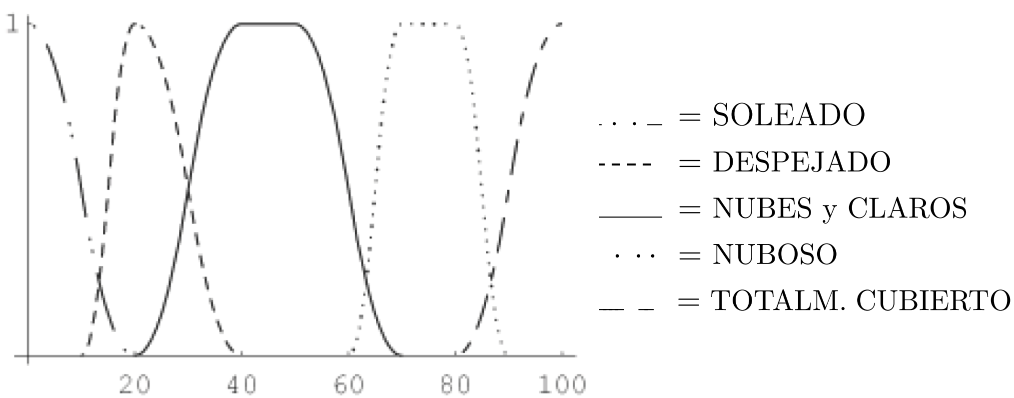


Fig. 1.8: Valores de la variable NUBOSIDAD

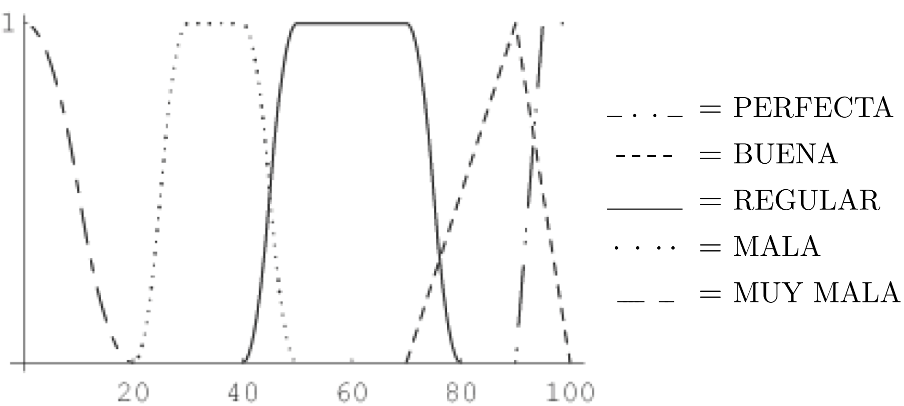


Fig. 1.9: Valores de la variable VISIBILIDAD

Para la población considerada los datos difusos bidimensionales observados se recogen en la Tabla 1.1.

$\mathcal{X} \setminus \mathcal{Y}$	$\tilde{y}_1$	$\tilde{y}_2$	$\tilde{y}_3$	$\tilde{y}_4$	$\tilde{y}_5$
$\tilde{x}_1$	3	7			
$\tilde{x}_2$	2	7	1		
$\tilde{x}_3$		5	4	1	
$\tilde{x}_4$		2	5	3	
$\tilde{x}_5$			2	5	3

Tabla 1.1: Tabla de contingencia para las variables NUBOSIDAD y VISIBILIDAD en Ω

Cuando se busca la mejor relación lineal de la VISIBILIDAD respecto a la NUBOSIDAD, en el sentido de expresar la primera como una función lineal difusa de la segunda, y usando el resultado del Teorema 1.6.1 para una medida S discretizada de forma que $L = 3,\ \lambda_1 = 0,\ \lambda_2 = 0,5,\ \lambda_3 = 1$, $k_1 = k_3 =, 375,\ k_2 = 0,25$ y $\overline{g}(\lambda) = 0$ si $\lambda \in (0,1) \setminus \{0,5\}$, entonces la función $\phi(a,b)$ adopta la forma de la Figura 1.10 y la relación óptima es aquella en la que la pendiente corresponde al valor $a^* = -0,95$ y el término independiente a $b^* = 111,77$, es decir, $\mathcal{Y} = (-0,95 \odot \mathcal{X}) \oplus 111,77$.

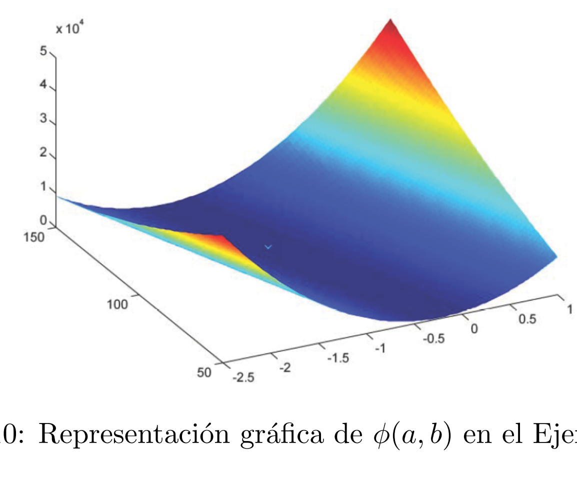


Fig. 1.10: Representación gráfica de $\phi(a,b)$ en el Ejemplo 1.6.1

Teniendo en cuenta esta relación, podemos predecir la VISIBILIDAD para un día que se ha clasificado con respecto a la NUBOSIDAD como un día con NUBES Y CLAROS, mediante el valor:

$$\tilde{y} = (-0,95 \odot \tilde{x}_3) \oplus 111,77 = \begin{cases} S(45,27,64,27) & \text{si } x \in [45,27,64,27] \\ 1 & \text{si } x \in [64,27,73,77] \\ 1 - S(73,77,92,77) & \text{si } x \in [73,77,92,77] \\ 0 & \text{en otro caso} \end{cases}$$

que correspondería a una VISIBILIDAD que (más o menos) podría catalogarse como MAS BIEN BUENA (Figura 1.11).

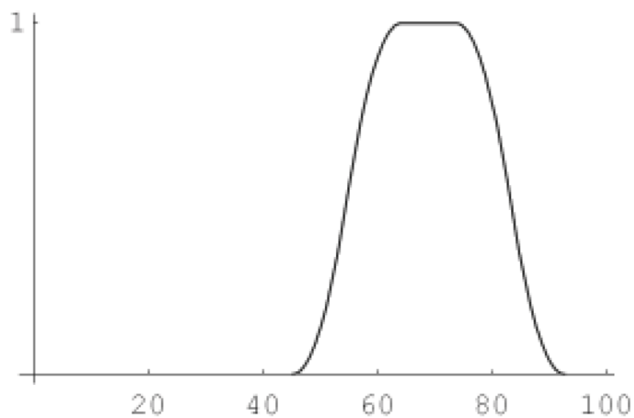
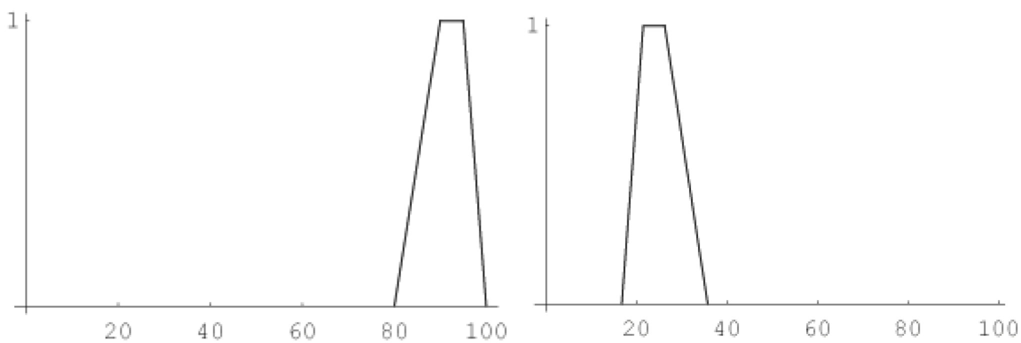


Fig. 1.11: Predicción de la VISIBILIDAD para un día con NUBES Y CLAROS según la regresión lineal óptima

También puede predecirse la VISIBILIDAD para un día cuyo valor queda fuera de los que $\mathcal{X}$ toma en Ω y clasificado con un "valor" que no aparece en los datos en los que se ha basado la obtención de la relación óptima. Así, la VISIBILIDAD para un día que se ha clasificado como MUY NUBOSO para la variable NUBOSIDAD y descrito por los expertos mediante el número trapezoidal $\tilde{x} = \text{Tra}(80, 90, 95, 100)$. El valor estimado para la VISIBILIDAD será:

$$\tilde{y} = (-0,95 \odot \tilde{x}) \oplus 111,77 = \text{Tra}(16,77,\ 21,52,\ 26,27,\ 35,77),$$

que podría interpretarse (más o menos) como BASTANTE MALA(Figura 1.12).



(a) (b)

Fig. 1.12: (a) Representación gráfica de un día MUY NUBOSO
(b) Predicción de la VISIBILIDAD para un día MUY NUBOSO
según la regresión lineal óptima

Cuando se busca la mejor relación funcional entre la VISIBILIDAD y la NUBOSIDAD, la solución viene dada por el Teorema 1.6.2, de modo que (cualquiera que fuera S), la predicción de la VISIBILIDAD para un día MUY NUBOSO no podría llevarse a cabo directamente. Sólo tendrían sentido directo las predicciones de valores de $\mathcal{Y}$ para los valores de $\mathcal{X}$ que aparecen en la población o muestra que nos ha proporcionado los datos que han sido la base de la relación óptima. Así, por ejemplo, la predicción de la VISIBILIDAD para un día de NUBES Y CLAROS vendría dada por:

$$\tilde{y} = E(\mathcal{Y} \,|\, \mathcal{X} = \tilde{x}_3) = \sum_{i=1}^{5} P(\tilde{y}_i \,|\, \mathcal{X} = \tilde{x}_3) \odot \tilde{y}_i$$

$$= \left(\frac{5}{10} \odot \tilde{y}_2\right) \oplus \left(\frac{4}{10} \odot \tilde{y}_3\right) \oplus \left(\frac{1}{10} \odot \tilde{y}_4\right),$$

valor que aparece representado en la Figura 1.13 y que es una versión más “afinada” de la predicción obtenida por la relación lineal.

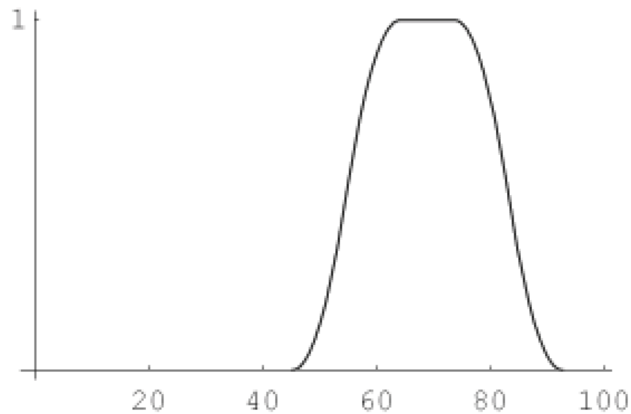


Fig. 1.13: Predicción de la VISIBILIDAD para un día con NUBES Y CLAROS según la regresión general óptima

1.7 Valoración final y problemas abiertos

Como valoración final del Capítulo 1 se concluye que la extensión del momento de segundo orden para variables aleatorias difusas basada en el empleo de la métrica D_S de Bertoluzza *et al.* (1995a) conserva todas la propiedades de interés para las variables aleatorias reales y se adapta en forma análoga a las aplicaciones estadísticas de dichos momentos que se recogen en este capítulo. No obstante, otros estudios basados en la extensión llevada a cabo en este capítulo no van a presentar un paralelismo similar y algunos de estos problemas se señalan dentro de esta sección.

Desde el punto de vista de la operatividad, las dificultades de la extensión se reducen básicamente a la descripción de los valores de la variables y el cálculo de una serie de integrales que se resuelven con el tratamiento informático (basado en buena parte en el desarrollado por López García 1997, y en las aproximaciones usuales de esas integrales).

En relación con la investigación que se presenta en este capítulo quedan muchos problemas abiertos que están siendo examinados o lo serán en un futuro próximo.

- Las propiedades de la Sección 1.2 podrían completarse de forma prácticamente automática con el desarrollo de algunas Leyes de los Grandes

Números. Debido a la existencia de una ley fuerte para variables difusas independientes dos a dos e igualmente distribuidas basada en la convergencia en el sentido de la métrica d_∞ (ver Colubi *et al.* 1999), que implicaría una ley análoga para la métrica D_S, sería interesante establecer leyes que no exigieran las condiciones anteriores para las variables aleatorias difusas de la sucesión (en este sentido, pueden servir de referencia los resultados de Körner 1997b).

- El llamado Teorema de Rao-Blackwell "mejorado" podría tratar de extenderse para variables aleatorias difusas, si bien su establecimiento precisaría de la formalización del concepto de suficiencia para un estimador difuso en relación con un parámetro difuso, lo que no es en absoluto inmediato.

- Los resultados de las Secciones 1.3-1.5 podrían extenderse a muestreos más complejos; en este sentido, reviste especial interés el muestreo estratificado, que se correspondería en el caso asintótico con la consideración de mixturas finitas de distribuciones. También sería interesante examinar la idoneidad de posibles criterios de estratificación a través de la definición de alguna medida de la ganancia de precisión relativa debida a la estratificación (siguiendo las ideas de Alonso & Gil 1995). En este punto un inconveniente que surge, en comparación con los resultados del caso clásico, es el de la determinación de una buena aproximación del tamaño muestral adecuado para estimar la S-DCM poblacional sobre la base de la distribución asintótica, puesto que no sería posible con los resultados disponibles hasta el momento desarrollar una técnica análoga a la que en el caso clásico nos proporciona la aplicación del Teorema de Stone. Será, por tanto, conveniente buscar procedimientos que mejoren los tamaños de los Teoremas 1.5.4 y 1.5.8.

- Los estudios asintóticos de la Sección 1.4 pueden llevarse a cabo cuando se consideran dos poblaciones y se analiza la diferencia entre sus

S-DCM, con el objeto de comparar las mismas.

- También será necesario discutir los tamaños muestrales para los que los resultados asintóticos de la Sección 1.4 proporcionan buenas aproximaciones. Esta discusión podría realizarse mediante la simulación a partir de variables aleatorias difusas, lo que nos lleva a un doble problema abierto: el de la consideración de modelos operativos de "distribuciones" para variables aleatorias difusas, y el de la simulación a partir de los mismos.

- El estudio de regresión lineal desarrollado en la Sección 1.6 considera que ni $\mathcal{X}$ ni $\mathcal{Y}$ son variables que puedan tomar casi seguro valores reales. Si una de las dos variables toma casi seguro valores reales, y en particular si se trata de una variable aleatoria real, el estudio llevado a cabo en este capítulo no sería aplicable. Sería útil desarrollar el estudio correspondiente a esta situación particular discutiendo en qué casos se produce la anulación del parámetro que representa la pendiente de la relación lineal. No obstante, este problema cobra mayor sentido cuando se admite que los parámetros de la relación lineal pueden tomar valores difusos, lo que en el caso en que $\mathcal{Y}$ fuera la variable difusa y $\mathcal{X}$ la real nos llevaría al problema tratado por Näther (1997) y Körner (1997ab).

- Uno de los problemas abiertos de este capítulo que parecen más atractivos, es el de la Correlación Lineal (y General) entre variables aleatorias difusas, especialmente por cuanto de diferente va a resultar este estudio con respecto al caso clásico. En este respecto conviene observar que cuando se consideran dos datos bidimensionales difusos, no necesariamente existirá una relación lineal que describa con exactitud la existente entre esos datos. También sería útil estudiar las relaciones entre los dos posibles "planos de regresión lineal" que representan las relaciones lineales óptimas de $\mathcal{Y}$ respecto a $\mathcal{X}$ y de $\mathcal{X}$ respecto a $\mathcal{Y}$.

- Otra cuestión asociada de interés es la búsqueda de métodos para predecir valores de la variable de predicción $\mathcal{Y}$ a partir de los de la observada $\mathcal{X}$, en el Análisis de Regresión General, cuando se consideran valores de $\mathcal{Y}$ que no aparecen en la muestra o población considerada para determinar la relación óptima. Con este fin, podrían tratar de adaptarse métodos adecuados de interpolación de datos difusos (ver, por ejemplo, Lowen 1990, Kaleva 1994, Gal 1996). Además, algunas de las funciones que aparecen involucradas en los estudios de interpolación servirían de inspiración para la búsqueda de modelos de regresión con relaciones no lineales, pero que satisficieran ciertas condiciones de "continuidad" y "regularidad".

- Por último, los estudios llevados a cabo en este capítulo requieren un análisis sobre la robustez desde distintas perspectivas: por un lado, la sensibilidad de las valoraciones y comparaciones basadas en la S-DCM en función de la elección de la medida S; por otro lado, la sensibilidad de las valoraciones, comparaciones y resultados descriptivos e inferenciales en función de la forma de los conjuntos difusos elegidos como modelo de los valores de la variable considerada.

Capítulo 2

Medida generalizada de la variación relativa (desigualdad) de una variable aleatoria difusa

Al igual que la desviación, la "desigualdad" o variación relativa de una variable aleatoria permite la comparación entre variables o entre poblaciones. Sin embargo, mientras que la desviación suele cuantificarse en términos de las distancias entre los valores de la variable y cierto valor de referencia (habitualmente el valor esperado), la desigualdad suele cuantificarse en términos de las razones de esos valores con respecto al de referencia. Además, en la desigualdad interesa considerar "cuántas veces mayor es cada valor de la variable que el valor de referencia" y se distingue por tanto entre el caso en que un valor sea superior y el caso en que sea inferior al de referencia (lo que no hace la desviación para valores equidistantes del de referencia).

Otro rasgo diferencial de la desigualdad frente a la desviación es que la "contribución" de cada valor de la variable a la desigualdad asociada a esa variable, no siempre es no negativa, sino que es positiva para valores por debajo del de referencia, negativa para los valores por encima y nula para valores iguales al de referencia.

En el capítulo de introducción se ha señalado que las variables para las que se definen los índices de desigualdad se suponen habitualmente positivas, y en la mayoría de las situaciones (aunque no se indique de forma expresa) se admite que sus valores vienen dados en una escala de razón.

En este capítulo, se parte del hecho práctico de que existen diversos rasgos intrínsecos a su aspecto cuantitativo (repercusión social/política, publicidad/perjuicio añadida/o, imprecisión en la información o en las valoraciones, etc.), que pueden dar lugar a que la escala usual para el aspecto no refleje en cada caso el verdadero valor. Ejemplos de esta naturaleza motivaron la introducción del concepto de variable aleatoria difusa y, en particular, el de utilidad difusa en los problemas de decisión (ver, por ejemplo, Gil & López-Díaz 1996, Gil *et al.* 1998b).

Atkinson (1983, pág. 53) señala que las medidas de desigualdad "...se emplean ampliamente con dos fines:

(i) comparar distribuciones (de rentas, beneficios, etc.);

(ii) dar alguna medida del grado de desigualdad, o alguna idea de si la desigualdad es 'grande' o 'pequeña'..."

Para lograr el segundo propósito se han desarrollado algunos estudios (Basu 1987, Ok 1994, 1995, 1996, López García 1997, Colubi *et al.* 1997, Colubi Cervero 1997, Gil *et al.* 1998a, López-García & Corral 1998) en los que se han establecido medidas con valores difusos del grado de desigualdad de una población con respecto a una variable aleatoria (real o difusa).

Sin embargo, si quiere darse respuesta al primer propósito, debe recurrirse o bien a una medida de la desigualdad que tome valores reales, o a aplicar algún procedimiento posterior de ordenación de números difusos sobre las medidas con valores difusos. En la familia de medidas que se presenta a continuación se recogen ambas opciones: toma valores reales, y proviene de la composición de una función de ordenación de números difusos y una medida de desigualdad difusa generalizada.

De este modo, en este capítulo se define y estudia el comportamiento de una familia generalizada de índices de desigualdad con valores reales respecto a una variable aleatoria difusa. Esta familia extiende la que aparece en la Definición 0.4.1 e incluye las extensiones a variables aleatorias difusas de todos los índices aditivamente descomponibles, salvo el de Theil. Concretamente, el índice hiperbólico y el de tipo Shannon aparecen extendidos en la familia definida.

A continuación, se examinan condiciones para la existencia de esa extensión, y se analizan propiedades de la medida de variación relativa definida.

En el caso del índice hiperbólico extendido, se desarrolla también la estimación insesgada en los muestreos aleatorios simple y con reposición de poblaciones finitas, cuantificándose de forma exacta el error asociado al muestreo en dicha estimación. Este estudio se complementa con el de la distribución asintótica para la mayoría de los índices de la familia definida, que permite la construcción de técnicas inferenciales aproximadas sobre la medida poblacional.

2.1 Los f-índices de desigualdad para variables aleatorias difusas

Supongamos que se considera una población Ω cualquiera, el espacio de probabilidad asociado a ella, $(\Omega, \mathcal{A}, P)$, y una variable aleatoria difusa integrablemente acotada y positiva asociada a $(\Omega, \mathcal{A}, P)$ (es decir, $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ variable aleatoria difusa y tal que $\sup \mathcal{X}_0 \in L^1(\Omega, \mathcal{A}, P)$).

Sea $f : (0, +\infty) \to \mathbb{R}$ una función monótona y estrictamente cóncava hacia arriba que verifica que $f(u) + f(1/u) \geq 0$ para todo $u \in (0, +\infty)$ y $f(1) = 0$, y F la función de ordenación introducida por Yager (1981) (expuesta en la Subsección 0.1).

Para medir la desigualdad asociada a $\mathcal{X}$ en Ω mediante una cuantificación con valores reales, se proponen los índices siguientes:

Definición 2.1.1 . *El f-índice de desigualdad (con valores reales) asociado a $\mathcal{X}$ en la población viene dado por el valor (si existe):*

$$\mathrm{I}_f(\mathcal{X}) = F\left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right),$$

donde:

$$\widetilde{\mathrm{I}}_f(\mathcal{X}) = \widetilde{E}\left[f\left(\mathcal{X} \oslash \widetilde{E}(\mathcal{X})\right)\right],$$

*es el f-índice de desigualdad difuso asociado con $\mathcal{X}$ en la población (*Colubi *et al.* 1997, Colubi Cervero 1997*), donde se denota por $f\left(\mathcal{X} \oslash \widetilde{E}(\mathcal{X})\right)$ la imagen de $\mathcal{X} \oslash \widetilde{E}(\mathcal{X})$ mediante la extensión de f por medio del principio de Zadeh (es decir, $\left(f\left(\mathcal{X} \oslash \widetilde{E}(\mathcal{X})\right)\right)_\alpha = \left[\text{mín}\left\{ f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right), f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right\}, \text{máx}\left\{ f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right), f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right\}\right]$ para todo $\alpha \in [0,1]$).*

Ocasionalmente, y cuando sea necesario especificar la medida de probabilidad P en el espacio probabilístico, se denotará $\mathrm{I}_f(\mathcal{X})$ alternativamente por $\mathrm{I}_f(\mathcal{X} \mid P)$.

Las condiciones exigidas sobre f las cumplen tanto las funciones asociadas a los índices aditivamente descomponibles para $\alpha \neq 0, 1$, como la función $f(x) = -\log x$ asociada al índice de tipo Shannon (ver Apéndice B). En cambio, la función correspondiente al índice de Theil ($f(x) = x\log(x)$) no es monótona, con lo que queda excluida del estudio.

En general, la existencia de $\widetilde{\mathrm{I}}_f$ no puede asegurarse para cualquier variable o población, y las condiciones para la existencia varían dependiendo de la función f.

Las condiciones siguientes garantizan la existencia de $\widetilde{\mathrm{I}}_f(\mathcal{X})$ (ver Colubi Cervero 1997) y, en consecuencia, la de $\mathrm{I}_f(\mathcal{X})$:

Se considera un espacio de probabilidad $(\Omega, \mathcal{A}, P)$, y se supone que $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ es una variable aleatoria difusa integrablemente acotada y positiva asociada a ese espacio y que $f : (0, +\infty) \to \mathbb{R}$ es una función monótona y estrictamente cóncava hacia arriba y de clase C^1, que satisface que $f(u) + f(1/u) \geq 0$ para todo $u \in (0, +\infty)$ y $f(1) = 0$.

Entonces, el f-índice de desigualdad difuso $\widetilde{\mathrm{I}}_f(\mathcal{X})$ está bien definido y pertenece a $\mathcal{F}_c(\mathbb{R})$ si, y sólo si, f verifica que:

(1) $f(\inf \mathcal{X}_0 / E(\sup \mathcal{X}_0)) \in L^1(\Omega, \mathcal{A}, P)$, en el caso de f no creciente;

(2) $f(\sup \mathcal{X}_0 / E(\inf \mathcal{X}_0)) \in L^1(\Omega, \mathcal{A}, P)$, en el caso de f no decreciente.

Por otra parte, el f-índice de desigualdad difuso asociado a $\mathcal{X}$ en la población (cuando existe en las condiciones supuestas para f), es el número difuso en $\mathcal{F}_c(\mathbb{R})$ cuyos α-cortes, para todo $\alpha \in [0,1]$, vienen dados por:

$$\left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha = \left[\inf \left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha, \sup \left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha\right],$$

donde:

$$\inf \left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha = E\left(\text{mín}\left\{f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right), f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right\}\right),$$

$$\sup \left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha = E\left(\text{máx}\left\{f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right), f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right\}\right).$$

Sobre la base de estas afirmaciones, y teniendo en cuenta las propiedades de la función de ordenación F, se tiene que:

Teorema 2.1.1. *Sea* $(\Omega, \mathcal{A}, P)$ *un espacio de probabilidad y* $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ *una variable aleatoria difusa integrablemente acotada. Se considera una aplicación* $f : (0, +\infty) \to \mathbb{R}$ *monótona y estrictamente cóncava hacia arriba y de clase* C^1*, que satisface que* $f(u) + f(1/u) \geq 0$ *para todo* $u \in (0, +\infty)$ *y* $f(1) = 0$.

Entonces, si se cumple o bien (1) o (2), se tiene que:

$$\mathrm{I}_f(\mathcal{X}) = \frac{1}{2} \int_{[0,1]} E\left[f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right) + f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right] d\alpha.$$

Obviamente, la eliminación de las condiciones de convexidad para los valores de $\mathcal{X}$ y de monotonía para la función f no permitiría una caracterización como la que se acaba de dar, y esto supondría un inconveniente, tanto práctico como teórico, para los estudios desarrollados posteriormente.

En la sección siguiente se examinarán varias propiedades deseables y convenientes en la medición de la desigualdad de una población respecto a una variable aleatoria difusa.

2.2 Propiedades de los f-índices de desigualdad

De aquí en adelante, sin hacer mención expresa en cada momento, se considerará el espacio de probabilidad $(\Omega, \mathcal{A}, P)$, y una aplicación $f : (0, +\infty) \to \mathbb{R}$ monótona y cóncava hacia arriba en sentido estricto, de clase C^1 y que satisface que $f(u)+f(1/u) \geq 0$ para todo $u \in (0, +\infty)$ y $f(1) = 0$. Se supondrá también que se cumple o bien la condición (1) o la (2) de la Sección 2.1. Si se reduce la concavidad estricta a la concavidad, las condiciones para la igualdad en varias de las propiedades que siguen no podrían establecerse.

Los índices introducidos no experimentan cambios al considerar una variación real equiproporcional en los valores de la variable aleatoria difusa. En otras palabras, de acuerdo con la terminología de Kölm (1976ab) estos índices son "medidas derechistas", y siguiendo a Blackorby & Donaldson (1978) se trata de medidas de "desigualdad relativa". De este modo, el resultado siguiente extiende a variables aleatorias difusas la propiedad de *independencia de la media* de la mayoría de los índices de desigualdad clásicos (también llamada *invariancia por escala,* u *homogeneidad de grado 0*).

Teorema 2.2.1 (Independencia de la media). *Si* $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ *es una variable aleatoria difusa integrablemente acotada, entonces para todo* $k \in (0, +\infty)$ *se tiene que* $\mathrm{I}_f(k \odot \mathcal{X}) = \mathrm{I}_f(\mathcal{X})$.

Demostración:

Como $k \odot \mathcal{X}$ es a su vez una variable aleatoria difusa integrablemente acotada y positiva, con $k\odot\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ definida por $(k \odot \mathcal{X})(\omega) =$

$k \odot \mathcal{X}(\omega)$ para todo $\omega \in \Omega$, entonces para todo $\alpha \in [0,1]$ y $k \in (0,+\infty)$ se tiene que:

$$\frac{\inf(k \odot \mathcal{X})_\alpha}{E\left(\sup(k \odot \mathcal{X})_\alpha\right)} = \frac{k \inf \mathcal{X}_\alpha}{kE(\sup \mathcal{X}_\alpha)} = \frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)},$$

y:

$$\frac{\sup(k \odot \mathcal{X})_\alpha}{E\left(\inf(k \odot \mathcal{X})_\alpha\right)} = \frac{k \sup \mathcal{X}_\alpha}{kE(\inf \mathcal{X}_\alpha)} = \frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)},$$

de donde:

$$\mathrm{I}_f(k \odot \mathcal{X}) = \mathrm{I}_f(\mathcal{X}).$$

□

La *conservación del signo* (o *condición de no negatividad*) va a cumplirse para todos los índices introducidos en la sección anterior. Así:

Teorema 2.2.2 (No negatividad). *Sea* $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ *una variable aleatoria difusa integrablemente acotada. Entonces, se cumple que* $\mathrm{I}_f(\mathcal{X}) \geq 0$*.*

Demostración:

En efecto, como la función f es cóncava hacia arriba, sin más que aplicar la Desigualdad de Jensen a la expresión anterior, y por las condiciones impuestas sobre f, se obtiene que:

$$E\left(f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right)\right) + E\left(f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right)$$

$$\geq f\left(E\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right)\right) + f\left(E\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right)$$

$$= f\left(\frac{E(\inf \mathcal{X}_\alpha)}{E(\sup \mathcal{X}_\alpha)}\right) + f\left(\frac{E(\sup \mathcal{X}_\alpha)}{E(\inf \mathcal{X}_\alpha)}\right) \geq 0.$$

□

La *positividad* (o *sensibilidad en caso de no igualdad*) se formaliza en la propiedad siguiente:

Teorema 2.2.3 (Sensibilidad en caso de no igualdad). *Sea* $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ *una variable aleatoria difusa integrablemente acotada. Si* $\mathrm{I}_f(\mathcal{X}) = 0$, *entonces* $\mathcal{X}$ *debe ser una variable aleatoria difusa degenerada.*

Demostración:

En virtud del Teorema 2.1.1, y de la continuidad por la izquierda de $\inf \left(f\left(\mathcal{X} \oslash \widetilde{E}(\mathcal{X})\right)\right)_\alpha$ y $\sup \left(f\left(\mathcal{X} \oslash \widetilde{E}(\mathcal{X})\right)\right)_\alpha$ respecto a α en $[0,1]$, $\mathrm{I}_f(\mathcal{X}) = 0$ si, y sólo si, para cada $\alpha \in [0,1]$:

$$\mathrm{E}\left[f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right)\right] + \mathrm{E}\left[f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right] = 0.$$

Para ello, sería necesario que la relación de desigualdad obtenida por aplicación de la Desigualdad de Jensen en el Teorema 2.2.2 fuera igualdad, lo cual ocurre si, y sólo si, para cada $\alpha \in [0,1]$ se satisface que $\inf \mathcal{X}_\alpha$ y $\sup \mathcal{X}_\alpha$ son variables aleatorias reales degeneradas, y por lo tanto todas las funciones α-corte de $\mathcal{X}$ deben ser conjuntos aleatorios compactos y convexos degenerados. En consecuencia (ver López García 1997, pág. 191), $\mathcal{X}$ debe ser una variable aleatoria difusa degenerada. □

La *insensibilidad* o *anulación* del f-índice de desigualdad no puede garantizarse sin embargo para una variable aleatoria difusa degenerada, cualquiera que sea f. El siguiente resultado, establece que cuando la variable aleatoria difusa es degenerada en un valor real, sí se produce tal insensibilidad.

Teorema 2.2.4 (Insensibilidad). *Sea* $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ *una variable aleatoria difusa integrablemente acotada. Si* $\mathcal{X}$ *es degenerada en un valor real positivo, entonces* $\mathrm{I}_f(\mathcal{X}) = 0$.

Demostración:

En efecto, si $\mathcal{X}$ es una variable aleatoria difusa degenerada en un valor real positivo, entonces para cada $\alpha \in [0,1]$ se cumple que $\inf \mathcal{X}_\alpha = \sup \mathcal{X}_\alpha$ c.s. $[P]$, y, en consecuencia, $E(\inf \mathcal{X}_\alpha) = E(\sup \mathcal{X}_\alpha)$, de modo que $\inf \left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha = \sup \left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha = 0$, lo que prueba que $\mathrm{I}_f(\mathcal{X}) = 0$. □

A continuación se exponen dos propiedades de *minimalidad* diferentes para los f-índices según ciertas condiciones satisfechas por la función f.

Teorema 2.2.5 (Minimalidad I). *Sea* $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ *una variable aleatoria difusa integrablemente acotada. Si* f *verifica que* $f(u) + f(1/u) = 0$ *si, y sólo si,* $u = 1$*, entonces* $\mathrm{I}_f(\mathcal{X}) = 0$ *si, y sólo si,* $\mathcal{X}$ *es una variable aleatoria difusa degenerada en un valor real positivo.*

Demostración:

En virtud de los Teoremas 2.2.2 y 2.2.3, $\mathrm{I}_f(\mathcal{X}) = 0$ si, y sólo si, $\inf \mathcal{X}_\alpha$ y $\sup \mathcal{X}_\alpha$ son variables aleatorias reales degeneradas y (por la condición adicional satisfecha por f) $E(\inf \mathcal{X}_\alpha) = E(\sup \mathcal{X}_\alpha)$, es decir, si, y sólo si, $\inf \mathcal{X}_\alpha$ y $\sup \mathcal{X}_\alpha$ son variables aleatorias reales degeneradas en el mismo valor, lo que obliga a que $\mathcal{X}$ sea una variable aleatoria difusa degenerada en ese valor. □

Como se señala en el Apéndice B de la memoria, la condición adicional $f(u) + f(1/u) = 0$ si, y sólo si, $u = 1$, la cumplen muchas funciones f (en particular, las que sirven de base a la extensión de los índices aditivamente descomponibles para $\alpha \neq 0, 1$), aunque no la verifican otras como la función $f(x) = -\log x$ (base del índice de tipo Shannon) que cumple que $f(u) + f(1/u) = 0$ para todo $u \in (0, +\infty)$ (ver también Apéndice B). En este caso, la condición necesaria y suficiente para la anulación de $\mathrm{I}_f(\mathcal{X})$ aparece recogida en el resultado siguiente:

Teorema 2.2.6 (Minimalidad II). *Si* $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ *es una variable aleatoria difusa integrablemente acotada, y* $f(u) + f(1/u) = 0$ *para todo* $u \in (0, +\infty)$*, entonces* $\mathrm{I}_f(\mathcal{X}) = 0$ *si, y sólo si,* $\mathcal{X}$ *es una variable aleatoria difusa degenerada en un elemento de* $\mathcal{F}_c((0, +\infty))$*.*

Demostración:

La condición necesaria es válida por el Teorema 2.2.3.

Recíprocamente, si $\mathcal{X}$ es una variable aleatoria difusa degenerada, $\inf \mathcal{X}_\alpha$ y $\sup \mathcal{X}_\alpha$ son variables aleatorias reales degeneradas, y por lo tanto:

$$\frac{1}{2}\left[f\left(\frac{E(\inf \mathcal{X}_\alpha)}{E(\sup \mathcal{X}_\alpha)}\right)+f\left(\frac{E(\sup \mathcal{X}_\alpha)}{E(\inf \mathcal{X}_\alpha)}\right)\right]=0,$$

obedeciendo esta anulación a la condición adicional supuesta para f. □

Observación 2.2.1. Según los Teoremas 2.2.5 y 2.2.6, cuando $\mathcal{X}$ es una variable aleatoria difusa degenerada en un valor de $\mathcal{F}_c((0,+\infty))$ no necesariamente se va a anular el valor del f-índice de desigualdad. Así, por ejemplo, si $\mathcal{X}$ toma casi seguro el valor $\tilde{x} = \mathrm{Tri}(1,2,3)$ sobre Ω y se considera $f(x) = x^{-1} - 1$ para todo $x \in (0,1)$, se obtiene que $\mathrm{I}_f(\mathcal{X}) = 0,099$.

El motivo por el que los f-índices no se anulan siempre para variables aleatorias difusas degeneradas, obedece al hecho de que varios de esos índices (en particular, todos los asociados a funciones f tales que $f(u)+f(1/u)=0$ si, y sólo si, $u=1$) tienen en cuenta además de la desigualdad entre los valores de $\mathcal{X}$ la desigualdad "dentro" de cada valor. En este sentido, y como caso especial, puede probarse la *descomposición aditiva* siguiente para el índice hiperbólico (es decir, $f(x) = x^{-1} - 1$), según la cual se obtiene que si $\Omega = \{\omega_1, \ldots, \omega_N\}$, $\mathcal{X}(\Omega) = \{\tilde{x}_1^*, \ldots, \tilde{x}_r^*\}$ y $p_l = P(\{\omega \in \Omega \,|\, \mathcal{X}(\omega) = \tilde{x}_l^*\})$, $l = 1, \ldots, r$:

$$\mathrm{I}_H(\mathcal{X}) = \sum_{l=1}^{r} p_l^2\, \mathrm{I}_H(\{\tilde{x}_l^*\}) + \mathrm{I}_H^{ev}(\mathcal{X})$$

($\mathrm{I}_H^{ev}(\mathcal{X})$ representa una especie de índice *inter*valores, e $\mathrm{I}_H(\{\tilde{x}_l^*\})$ representa el índice *intra*valor para $\tilde{x}_l^*$) con $\mathrm{I}_H^{ev}(\mathcal{X}) = 0$ si, y sólo si, $\mathcal{X}$ es una variable aleatoria difusa degenerada, aunque esta descomposición no es válida para cualquier f.

Cuando se trabaja con poblaciones finitas, existen propiedades del f-índice con clara aplicación y significado. Por ello, a partir de ahora se considerará un espacio de probabilidad $(\Omega, \mathcal{P}(\Omega), P)$ donde $\Omega = \{\omega_1, \ldots, \omega_N\}$ es la población y P es la distribución uniforme sobre Ω.

La particularización del Teorema 2.1.1 para poblaciones finitas y funciones que satisfagan las condiciones señaladas al comienzo de esta sección, se presenta a continuación:

Teorema 2.2.7 (Expresión sobre poblaciones finitas). *Sea $\mathcal{X}$ una variable aleatoria difusa positiva definida sobre una población de N individuos $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$. Entonces, se cumple que:*

$$\mathrm{I}_f(\mathcal{X}) = \frac{1}{2}\int_{[0,1]} \left[\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right) + \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right)\right] d\alpha.$$

La *simetría* de los f-índices formaliza el hecho de que la identidad o numeración de los individuos es irrelevante en la cuantificación de la desigualdad, es decir, los f-índices de desigualdad son objetivos. De este modo:

Teorema 2.2.8 (Simetría). *Si $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ es una variable aleatoria difusa, entonces $\mathrm{I}_f(\mathcal{X} \circ \sigma) = \mathrm{I}_f(\mathcal{X})$ para cualquier permutación σ sobre Ω.*

Demostración:

Se verifica, obviamente, que $\widetilde{E}(\mathcal{X} \circ \sigma) = \widetilde{E}(\mathcal{X})$. Además, para todo $\alpha \in [0,1]$, se sigue que:

$$E\left(f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{\sigma(j)})}{\sup E((\mathcal{X}\circ\sigma)_\alpha)}\right)\right) = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{\sigma(j)})}{\sup E((\mathcal{X}\circ\sigma)_\alpha)}\right)$$

$$= \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right),$$

$$E\left(f\left(\frac{\sup \mathcal{X}_\alpha(\omega_{\sigma(j)})}{\inf E((\mathcal{X}\circ\sigma)_\alpha)}\right)\right) = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_{\sigma(j)})}{\inf E((\mathcal{X}\circ\sigma)_\alpha)}\right)$$

$$= \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right),$$

de donde se deduce que $\mathrm{I}_f(\mathcal{X} \circ \sigma) = \mathrm{I}_f(\mathcal{X})$. □

El *principio poblacional* (u *homogeneidad poblacional*) resalta el hecho de que los f-índices de desigualdad de una población dada Ω con respecto a una variable aleatoria difusa positiva coinciden con los que corresponden a una población $\Omega^{(r)} = \{\omega_{11}, \ldots, \omega_{1N}, \ldots, \omega_{r1}, \ldots, \omega_{rN}\}$ que se haya obtenido a partir de la primera mediante su réplica un número finito y arbitrario r de veces (es decir, $\omega_{ij} = \omega_j$ para todo i, j), con respecto a la extensión inmediata de esa variable. De forma más concisa, la desigualdad depende sólo de las proporciones en la población de cada valor difuso del aspecto $\mathcal{X}$ (es decir, de la estructura de la población) y no del tamaño de la misma.

Teorema 2.2.9 (Homogeneidad de la población). *Si* $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ *es una variable aleatoria difusa y* $\mathcal{X}^{(r)} : \Omega^{(r)} \to \mathcal{F}_c((0, +\infty))$ *es la variable aleatoria difusa que extiende* $\mathcal{X}$ *a* $\Omega^{(r)}$*, entonces* $\mathrm{I}_f(\mathcal{X}^{(r)}) = \mathrm{I}_f(\mathcal{X})$.

Demostración:

Teniendo en cuenta que la variable aleatoria difusa $\mathcal{X}^{(r)}$ está también definida sobre una población finita, para todo $\alpha \in [0, 1]$ se tiene que:

$$\frac{1}{N\,r}\sum_{i=1}^{r}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha^{(r)}(\omega_{ij})}{\dfrac{1}{N\,r}\displaystyle\sum_{i=1}^{r}\sum_{j=1}^{N}\sup \mathcal{X}_\alpha^{(r)}(\omega_{ij})}\right)$$

$$= \frac{1}{N\,r}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\dfrac{1}{N\,r}\displaystyle\sum_{j=1}^{N} r \sup \mathcal{X}_\alpha(\omega_j)}\right)\cdot r = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\dfrac{1}{N}\displaystyle\sum_{j=1}^{N}\sup \mathcal{X}_\alpha(\omega_j)}\right),$$

$$\frac{1}{N\,r}\sum_{i=1}^{r}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha^{(r)}(\omega_{ij})}{\dfrac{1}{N\,r}\displaystyle\sum_{i=1}^{r}\sum_{j=1}^{N}\inf \mathcal{X}_\alpha^{(r)}(\omega_{ij})}\right)$$

$$= \frac{1}{N\,r}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\dfrac{1}{N\,r}\displaystyle\sum_{j=1}^{N} r \inf \mathcal{X}_\alpha(\omega_j)}\right)\cdot r = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\dfrac{1}{N}\displaystyle\sum_{j=1}^{N}\inf \mathcal{X}_\alpha(\omega_j)}\right),$$

lo que asegura la igualdad del f-índice de desigualdad para ambas poblaciones. □

Otra propiedad importante de la familia de los f-índices es la *continuidad,* según la cual "pequeños cambios" (difusos) en los valores de la variable difusa provocan también "pequeñas variaciones" (reales) en los f-índices de desigualdad. Para formalizar la continuidad usaremos la métrica en la que se basa la medibilidad de las variables aleatorias difusas. Así:

Teorema 2.2.10 (Continuidad). *Sea $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ una variable aleatoria difusa, y sea $\mathcal{X}_{l,\widetilde{\mathcal{E}}}$ la variable aleatoria difusa definida en Ω tal que $\mathcal{X}_{l,\widetilde{\mathcal{E}}}(\omega_j) = \mathcal{X}(\omega_j)$ para todo $j \in \{1, \ldots, N\} \setminus \{l\}$ y $\mathcal{X}_{l,\widetilde{\mathcal{E}}}(\omega_l) = \mathcal{X}(\omega_l) \oplus \widetilde{\mathcal{E}}$ con $\widetilde{\mathcal{E}} \in \mathcal{F}_c\big(\mathbb{R}\big)$ tal que $\mathcal{X}(\omega_l) \oplus \widetilde{\mathcal{E}} \in \mathcal{F}_c((0, +\infty))$. Entonces:*

$$\lim_{d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}}) \to 0} \mathrm{I}_f(\mathcal{X}_{l,\widetilde{\mathcal{E}}}) = \mathrm{I}_f(\mathcal{X}).$$

Demostración:

En efecto:

$$d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}}) = \sup_{\alpha \in [0,1]} \max\{|\sup \widetilde{\mathcal{E}}_\alpha|, |\inf \widetilde{\mathcal{E}}_\alpha|\},$$

y, además:

$$|\mathrm{I}_f(\mathcal{X}_{l,\widetilde{\mathcal{E}}}) - \mathrm{I}_f(\mathcal{X})| = \left| \frac{1}{N} \sum_{j=1}^{N} \int_{[0,1]} \frac{1}{2} \left[f\left(\frac{\inf(\mathcal{X}_{l,\widetilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\widetilde{\mathcal{E}}})_\alpha)} \right) - f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)} \right) \right] \right.$$

$$\left. + \frac{1}{2} \left[f\left(\frac{\sup(\mathcal{X}_{l,\widetilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\widetilde{\mathcal{E}}})_\alpha)} \right) - f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)} \right) \right] d\,\alpha \right|.$$

Es inmediato que para todo $\alpha \in [0, 1]$ y $j \in \{1, \ldots, N\}$ se cumple que:

$$\left| \inf(\mathcal{X}_{l,\widetilde{\mathcal{E}}}(\omega_j))_\alpha - \inf \mathcal{X}_\alpha(\omega_j) \right| \leq |\inf \widetilde{\mathcal{E}}_\alpha| \leq d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}}),$$

$$\left| \sup(\mathcal{X}_{l,\widetilde{\mathcal{E}}}(\omega_j))_\alpha - \sup \mathcal{X}_\alpha(\omega_j) \right| \leq |\sup \widetilde{\mathcal{E}}_\alpha| \leq d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}}),$$

de donde $d_\infty(\mathcal{X}_{l,\widetilde{\mathcal{E}}}(\omega_j), \mathcal{X}(\omega_j)) \leq d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}})$, y:

$$\left|\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha) - \inf E(\mathcal{X}_\alpha)\right| \leq |\inf \widetilde{\mathcal{E}}_\alpha|/N \leq d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}})/N,$$

$$\left|\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha) - \sup E(\mathcal{X}_\alpha)\right| \leq |\sup \widetilde{\mathcal{E}}_\alpha|/N \leq d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}})/N,$$

y, por lo tanto, $d_\infty(\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}}), \widetilde{E}(\mathcal{X})) \leq d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}})/N$.

En consecuencia, se tiene que para todo $j \in \{1, \ldots, N\}$:

$$\lim_{d_\infty(\tilde{\mathcal{E}}, \mathbf{1}_{\{0\}}) \to 0} d_\infty\left(\frac{\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j)}{\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}})}, \frac{\mathcal{X}(\omega_j)}{\widetilde{E}(\mathcal{X})}\right) = 0.$$

Además, por el Teorema del Valor Medio del Cálculo Diferencial, al ser f una función monótona, cóncava hacia arriba en sentido estricto y de clase C^1, se cumple que:

$$\left|f\left(\frac{\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}\right) - f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right)\right|$$

$$= |f'(c_{1j}(\alpha))| \left|\frac{\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)} - \frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right|,$$

y, análogamente:

$$\left|f\left(\frac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}\right) - f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right)\right|$$

$$= |f'(c_{2j}(\alpha))| \left|\frac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)} - \frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right|,$$

donde c_{1j} y c_{2j} son valores entre $\dfrac{\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}$ y $\dfrac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}$, y entre $\dfrac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}$ y $\dfrac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}$, respectivamente.

Como la función f es de clase C^1, si $[a_j, b_j]$ es un elemento de $\mathcal{K}_c((0, +\infty))$ que contiene a $\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_0 / \inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_0)$ y $\sup \mathcal{X}_0(\omega_j)/\inf E(\mathcal{X}_0)$ si k_j es el máximo de la función derivada f' en $[a_j, b_j]$ se satisface que:

$$\left|f\left(\frac{\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}\right) - f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right)\right|$$

$$\leq k_j \left| \frac{\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)} - \frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)} \right|,$$

y:

$$\left| f\left(\frac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)} \right) - f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)} \right) \right|$$

$$\leq k_j \left| \frac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)} - \frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)} \right|.$$

Por lo tanto, para todo $j \in \{1, \ldots, N\}$:

$$d_H \left(f\left(\frac{(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{(\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}}))_\alpha} \right), f\left(\frac{\mathcal{X}_\alpha(\omega_j)}{E(\mathcal{X}_\alpha)} \right) \right)$$

$$\leq d_\infty \left(f\left(\frac{\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j)}{\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}})} \right), f\left(\frac{\mathcal{X}(\omega_j)}{\widetilde{E}(\mathcal{X})} \right) \right)$$

$$\leq k_j d_\infty \left(\frac{\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j)}{\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}})}, \frac{\mathcal{X}(\omega_j)}{\widetilde{E}(\mathcal{X})} \right),$$

de donde, para todo $j \in \{1, \ldots, N\}$ se cumple que:

$$\lim_{d_\infty(\tilde{\mathcal{E}}, \mathbf{1}_{\{0\}}) \to 0} d_\infty \left(f\left(\frac{\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j)}{\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}})} \right), f\left(\frac{\mathcal{X}(\omega_j)}{\widetilde{E}(\mathcal{X})} \right) \right) = 0.$$

En consecuencia, como para cualesquiera $\alpha \in [0, 1]$ y $j \in \{1, \ldots, N\}$, se satisface que:

$$\frac{1}{2} \left[f\left(\frac{\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)} \right) - f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)} \right) \right]$$

$$+ \frac{1}{2} \left[f\left(\frac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)} \right) - f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)} \right) \right]$$

$$\leq d_\infty \left(f\left(\frac{\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j)}{\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}})} \right), f\left(\frac{\mathcal{X}(\omega_j)}{\widetilde{E}(\mathcal{X})} \right) \right),$$

de donde se concluye que:

$$\lim_{d_\infty(\tilde{\mathcal{E}},\mathbf{1}_{\{0\}})\to 0} \mathrm{I}_f(\mathcal{X}_{l,\tilde{\mathcal{E}}}) = \mathrm{I}_f(\mathcal{X}).$$

□

Algunas de las propiedades principales y deseables en los índices de desigualdad son aquellas relativas a la reacción de los mismos frente a la "redistribución" de los valores del atributo considerado.

La *Schur-convexidad estricta* de I_f, formaliza el hecho de que la sustitución de valores del aspecto difuso por combinaciones lineales convexas de los mismos, no puede aumentar el f-índice de desigualdad en la población.

Teorema 2.2.11 (Schur-convexidad estricta). *Si $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ es una variable aleatoria difusa, (μ_{jl}) una matriz $N \times N$ doblemente estocástica, y $\mathcal{X}' : \Omega \to \mathcal{F}_c((0,+\infty))$ una variable aleatoria difusa definida de forma que:*

$$\begin{pmatrix} \mathcal{X}'(\omega_1) \\ \mathcal{X}'(\omega_2) \\ \vdots \\ \mathcal{X}'(\omega_N) \end{pmatrix} = \begin{pmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1N} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{N1} & \mu_{N2} & \cdots & \mu_{NN} \end{pmatrix} \odot \begin{pmatrix} \mathcal{X}(\omega_1) \\ \mathcal{X}(\omega_2) \\ \vdots \\ \mathcal{X}(\omega_N) \end{pmatrix},$$

entonces, se tiene que $\mathrm{I}_f(\mathcal{X}) \geq \mathrm{I}_f(\mathcal{X}')$ con igualdad si, y sólo si, $\mathcal{X}' = \mathcal{X} \circ \sigma$ para cierta permutación sobre Ω.

Demostración:

En primer lugar se va a comprobar que bajo las condiciones supuestas, se verifica que $\widetilde{E}(\mathcal{X}') = \widetilde{E}(\mathcal{X})$. Para ello es suficiente tener en cuenta que, por la definición de $\mathcal{X}'$, para todo $j \in \{1,\ldots,N\}$ y $\alpha \in [0,1]$ se tiene que:

$$\inf \mathcal{X}'_\alpha(\omega_j) = \sum_{l=1}^{N} \mu_{jl} \inf \mathcal{X}_\alpha(\omega_j),$$

$$\sup \mathcal{X}'_\alpha(\omega_j) = \sum_{l=1}^{N} \mu_{jl} \sup \mathcal{X}_\alpha(\omega_j).$$

En consecuencia, para todo $\alpha \in [0,1]$ se cumple que:

$$E\left(\inf \mathcal{X}'_\alpha\right) = \frac{1}{N}\sum_{j=1}^{N} \inf \mathcal{X}'_\alpha(\omega_j) = \frac{1}{N}\sum_{j=1}^{N}\sum_{l=1}^{N} \mu_{jl} \inf \mathcal{X}_\alpha(\omega_l)$$

$$= \frac{1}{N}\sum_{l=1}^{N}\left(\inf \mathcal{X}_\alpha(\omega_l)\sum_{j=1}^{N}\mu_{jl}\right) = \frac{1}{N}\sum_{l=1}^{N} \inf \mathcal{X}_\alpha(\omega_l) = E\left(\inf \mathcal{X}_\alpha\right),$$

y, de forma análoga, para todo $\alpha \in [0,1]$, se satisface que:

$$\sup E\left(\mathcal{X}'_\alpha\right) = \sup E\left(\mathcal{X}_\alpha\right),$$

de modo que $\widetilde{E}(\mathcal{X}') = \widetilde{E}(\mathcal{X})$.

Además, para todo $\alpha \in [0,1]$ la aplicación de la Desigualdad de Jensen y la doble estocasticidad de la matriz (μ_{jl}), implica que:

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_j)}{\sup E(\mathcal{X}'_\alpha)}\right) = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\displaystyle\sum_{l=1}^{N}\mu_{jl}\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)}\right)$$

$$\leq \frac{1}{N}\sum_{j=1}^{N}\sum_{l=1}^{N}\mu_{jl}\, f\left(\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)}\right) = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right),$$

y, de la misma manera:

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}'_\alpha(\omega_j)}{\inf E(\mathcal{X}'_\alpha)}\right) \leq \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right).$$

De lo anterior se concluye que $\mathrm{I}_f(\mathcal{X}) \geq \mathrm{I}_f(\mathcal{X}')$.

Por otro lado, la igualdad se alcanza si, y sólo si, después de permutar apropiadamente los elementos de Ω (lo cual, por la simetría establecida en el Teorema 2.2.8, no conlleva una modificación en el valor del f-índice) y las columnas correspondientes en la matriz (μ_{jl}), tenemos que o bien esta es la matriz identidad, o los $\mu_{ll} \in (0,1)$ corresponden a los individuos $\omega_l \in \Omega$ para las que $\mathcal{X}(\omega_l)$ coincide con las $\mathcal{X}(\omega_j)$ para los $j \neq l$ con $\mu_{jl} \in (0,1)$

o $\mu_{lj} \in (0,1)$, por lo que $\mathcal{X}$ y $\mathcal{X}'$ tomarán los mismos valores en estos individuos (tal vez permutados de la población original). $\square$

El conocido criterio de Lorenz para el caso real, puede extenderse para la ordenación de los vectores de los valores de una variable aleatoria difusa en una población finita mediante la consideración de la relación $\succ_S$ de Ramík & Římánek (1985). Así, si se admite la notación $V \succ_S W$ si, y sólo si, $V \succeq_S W$ pero no $V \preceq_S W$, la *compatibilidad con el criterio de Lorenz* de la f-familia de índices de desigualdad se puede establecer de la siguiente forma:

Teorema 2.2.12 (Compatibilidad con el criterio de Lorenz). *Se consideran $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ y $\mathcal{X}' : \Omega \to \mathcal{F}_c((0,+\infty))$ dos variables aleatorias difusas tales que $\mathcal{X}(\omega_N) \succeq_S \ldots \succeq_S \mathcal{X}(\omega_1), \mathcal{X}'(\omega_N) \succeq_S \ldots \succeq_S \mathcal{X}'(\omega_1)$, $\widetilde{E}(\mathcal{X}) = \widetilde{E}(\mathcal{X}')$ y $\mathcal{X}\,{}_L\,\mathcal{X}'$ (por lo que se entiende que $\mathcal{X}(\omega_1) \oplus \cdots \oplus \mathcal{X}(\omega_k) \succeq_S \mathcal{X}'(\omega_1) \oplus \cdots \oplus \mathcal{X}'(\omega_k)$ para todo $k \in \{1,\ldots,N\}$ con $\succ_S$ para al menos un k). Entonces, se cumple que $\mathrm{I}_f(\mathcal{X}) < \mathrm{I}_f(\mathcal{X}')$.*

Demostración:

Si $\mathcal{X}\,{}_L\,\mathcal{X}'$, se cumple que las variables aleatorias reales $\sup \mathcal{X}_\alpha$ y $\sup \mathcal{X}'_\alpha$ conservan la ordenación (con el criterio de Lorenz clásico) de las variables aleatorias difusas de las que provienen, es decir, $(\sup \mathcal{X}_\alpha)\,{}_L\,(\sup \mathcal{X}'_\alpha)$ y, análogamente, que $(\inf \mathcal{X}_\alpha)\,{}_L\,(\inf \mathcal{X}'_\alpha)$ para todo $\alpha \in [0,1]$. En virtud de un conocido resultado (ver, por ejemplo, Dasgupta *et al.* 1973, Marshall & Olkin 1979 y Eichhorn & Gehrig 1982), se cumple entonces, para todo $\alpha \in [0,1]$, que existen dos matrices doblemente estocásticas $(\mu_{jl}^{\sup}(\alpha))$ y $(\mu_{jl}^{\inf}(\alpha))$ tales que:

$$\begin{pmatrix} \sup \mathcal{X}_\alpha(\omega_1) \\ \vdots \\ \sup \mathcal{X}_\alpha(\omega_N) \end{pmatrix} = \begin{pmatrix} \mu_{11}^{\sup}(\alpha) & \cdots & \mu_{1N}^{\sup}(\alpha) \\ \vdots & \ddots & \vdots \\ \mu_{N1}^{\sup}(\alpha) & \cdots & \mu_{NN}^{\sup}(\alpha) \end{pmatrix} \odot \begin{pmatrix} \sup \mathcal{X}'_\alpha(\omega_1) \\ \vdots \\ \sup \mathcal{X}'_\alpha(\omega_N) \end{pmatrix},$$

y:

$$\begin{pmatrix} \inf \mathcal{X}_\alpha(\omega_1) \\ \vdots \\ \inf \mathcal{X}_\alpha(\omega_N) \end{pmatrix} = \begin{pmatrix} \mu_{11}^{\inf}(\alpha) & \cdots & \mu_{1N}^{\inf}(\alpha) \\ \vdots & \ddots & \vdots \\ \mu_{N1}^{\inf}(\alpha) & \cdots & \mu_{NN}^{\inf}(\alpha) \end{pmatrix} \odot \begin{pmatrix} \inf \mathcal{X}'_\alpha(\omega_1) \\ \vdots \\ \inf \mathcal{X}'_\alpha(\omega_N) \end{pmatrix}.$$

En consecuencia, y aplicando la Desigualdad de Jensen, se tiene que:

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right) = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\displaystyle\sum_{l=1}^{N} \mu_{jl}^{\inf}(\alpha) \inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)}\right)$$

$$\leq \frac{1}{N}\sum_{j=1}^{N}\sum_{l=1}^{N} \mu_{jl}^{\inf}(\alpha) f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)}\right) = \frac{1}{N}\sum_{l=1}^{N} f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)}\right),$$

con igualdad si, y sólo si, o bien $\left(\mu_{jl}^{\inf}(\alpha)\right)$ es la matriz identidad o los $\mu_{ll}^{\inf}(\alpha) \in (0,1)$ corresponden a los individuos ω_l para los que $\mathcal{X}(\omega_l)$ coincide con $\mathcal{X}(\omega_j)$ para todo $j \neq l$ tal que $\mu_{jl}^{\inf}(\alpha) \in (0,1)$ o $\mu_{lj}^{\inf}(\alpha) \in (0,1)$, de modo que los valores de $\inf \mathcal{X}_\alpha$ resultarían una permutación de los de $\inf \mathcal{X}'_\alpha$, lo cual es imposible por $(\inf \mathcal{X}_\alpha) \, L \, (\inf \mathcal{X}'_\alpha)$.

Por lo tanto, para todo $\alpha \in [0,1]$ se cumple que:

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right) < \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_j)}{\sup E(\mathcal{X}'_\alpha)}\right).$$

Análogamente, se satisface que:

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right) < \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}'_\alpha(\omega_j)}{\inf E(\mathcal{X}'_\alpha)}\right),$$

de modo que $\mathrm{I}_f(\mathcal{X}) < \mathrm{I}_f(\mathcal{X}')$. □

Los *principios de transferencia progresivo* y *regresivo* formalizan la idea del efecto que produce la redistribución llevada a cabo por medio de la transferencia de un valor a otro inferior, manteniendo el valor esperado del aspecto difuso en el valor del f-índice de desigualdad de la población. Precisando más, el principio de transferencia progresivo indica que:

Teorema 2.2.13 (Principio progresivo de transferencia). *Si* $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ *es una variable aleatoria difusa, y* $\mathcal{X}' : \Omega \to \mathcal{F}_c((0,+\infty))$ *es otra variable aleatoria difusa tal que* $\mathcal{X}(\omega_j) = \mathcal{X}'(\omega_j)$ *para todo* $j \in \{1,\dots,N\}\backslash\{l,l'\}$, $\mathcal{X}(\omega_l) \succeq_S \mathcal{X}'(\omega_l) \succeq_S \mathcal{X}(\omega_{l'})$, $\mathcal{X}(\omega_l) \succeq_S \mathcal{X}'(\omega_{l'}) \succeq_S \mathcal{X}(\omega_{l'})$ *y* $\widetilde{E}(\mathcal{X}) = \widetilde{E}(\mathcal{X}')$. *Entonces, se cumple que* $\mathrm{I}_f(\mathcal{X}) \geq \mathrm{I}_f(\mathcal{X}')$.

Además, $\mathrm{I}_f(\mathcal{X}) = \mathrm{I}_f(\mathcal{X}')$ *si, y sólo si, o bien* $\mathcal{X} = \mathcal{X}'$ *en* Ω, *o* $\mathcal{X}' = \mathcal{X} \circ \sigma_{ll'}$ *para la permutación* $\sigma_{ll'}$ *sobre* Ω *que intercambia* ω_l *y* $\omega_{l'}$, *o para todo* $\alpha \in [0,1]$ *se cumple que o bien* $\inf \mathcal{X}_\alpha(\omega_l) = \inf \mathcal{X}'_\alpha(\omega_{l'})$, $\inf \mathcal{X}_\alpha(\omega_{l'}) = \inf \mathcal{X}'_\alpha(\omega_l)$, $\sup \mathcal{X}_\alpha(\omega_l) = \sup \mathcal{X}'_\alpha(\omega_l)$ *y* $\sup \mathcal{X}_\alpha(\omega_{l'}) = \sup \mathcal{X}'_\alpha(\omega_{l'})$, *o* $\inf \mathcal{X}_\alpha(\omega_l) = \inf \mathcal{X}'_\alpha(\omega_l)$, $\inf \mathcal{X}_\alpha(\omega_{l'}) = \inf \mathcal{X}'_\alpha(\omega_{l'})$, $\sup \mathcal{X}_\alpha(\omega_l) = \sup \mathcal{X}'_\alpha(\omega_{l'})$ *y* $\sup \mathcal{X}_\alpha(\omega_{l'}) = \sup \mathcal{X}'_\alpha(\omega_l)$.

Demostración:

Por las hipótesis del teorema, se verifica que $\sup E(\mathcal{X}'_\alpha) = \sup E(\mathcal{X}_\alpha)$, y:

$$\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)} \geq \frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)} \geq \frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)},$$

$$\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)} \geq \frac{\inf \mathcal{X}'_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)} \geq \frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)},$$

$$\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)} + \frac{\inf \mathcal{X}'_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)} = \frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)} + \frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)},$$

por lo que se tiene que debe existir $\lambda_1(\alpha) \in [0,1]$ tal que:

$$\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)} = \lambda_1(\alpha)\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)} + (1-\lambda_1(\alpha))\frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)},$$

$$\frac{\inf \mathcal{X}'_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)} = (1-\lambda_1(\alpha))\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)} + \lambda_1(\alpha)\frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)}.$$

En virtud de la Desigualdad de Jensen, se concluye que:

$$f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)}\right) \leq \lambda_1(\alpha) f\left(\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)}\right) + (1-\lambda_1(\alpha)) f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)}\right),$$

$$f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)}\right) \leq (1-\lambda_1(\alpha)) f\left(\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)}\right) + \lambda_1(\alpha) f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)}\right).$$

Sumando las desigualdades anteriores, se obtiene que:

$$f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)}\right) + f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)}\right) \leq f\left(\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)}\right) + f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)}\right),$$

y por las desigualdades precedentes, se satisface que:

$$\begin{aligned}
\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_j)}{\sup E(\mathcal{X}'_\alpha)}\right) &= \frac{1}{N}\left[\left(\sum_{\substack{j=1\\ j\neq l,l'}}^{N} f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_j)}{\sup E(\mathcal{X}'_\alpha)}\right)\right)\right. \\
&\qquad \left. + f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)}\right) + f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)}\right)\right] \\
&\leq \frac{1}{N}\left[\left(\sum_{\substack{j=1\\ j\neq l,l'}}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right)\right) + f\left(\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)}\right) + f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)}\right)\right] \\
&= \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right),
\end{aligned}$$

y razonando en forma similar (a través de un valor $\lambda_2(\alpha) \in [0,1]$):

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}'_\alpha(\omega_j)}{\inf E(\mathcal{X}'_\alpha)}\right) \leq \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right),$$

de donde se deduce que $\mathrm{I}_f(\mathcal{X}) \geq \mathrm{I}_f(\mathcal{X}')$.

La condición suficiente para la igualdad de $\mathrm{I}_f(\mathcal{X})$ e $\mathrm{I}_f(\mathcal{X}')$ es obvia. Que tal condición es también necesaria se deduce de forma inmediata mediante un razonamiento por el absurdo. De este modo, si $\mathcal{X}$ y $\mathcal{X}'$ no se ajustaran a las condiciones indicadas, entonces (por la continuidad por la izquierda de $\lambda_1(\alpha)$ y $\lambda_2(\alpha)$ en $[0,1]$) debe existir algún $\alpha \in [0,1]$ tal que $\lambda_1(\alpha) \in (0,1)$ o $\lambda_2(\alpha) \in (0,1)$, en cuyo caso, y siguiendo la demostración de la primera parte de este teorema, resultaría $\mathrm{I}_f(\mathcal{X}) > \mathrm{I}_f(\mathcal{X}')$. $\square$

Por otro lado, el principio de transferencia regresivo tiene un enunciado que se deriva inmediatamente del anterior, sin más que intercambiar los papeles de $\mathcal{X}$ e $\mathcal{X}'$ en el mismo.

Teorema 2.2.14 (Principio regresivo de transferencia). *Si* $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ *es una variable aleatoria difusa, y* $\mathcal{X}' : \Omega \to \mathcal{F}_c((0,+\infty))$, *es otra variable aleatoria difusa tal que* $\mathcal{X}(\omega_j) = \mathcal{X}'(\omega_j)$ *para todo* $j \in \{1,\ldots,N\} \setminus \{l,l'\}$, $\mathcal{X}'(\omega_l) \succeq_S \mathcal{X}(\omega_l) \succeq_S \mathcal{X}'(\omega_{l'})$, $\mathcal{X}'(\omega_l) \succeq_S \mathcal{X}(\omega_{l'}) \succeq_S \mathcal{X}'(\omega_{l'})$ *y* $\widetilde{E}(\mathcal{X}) = \widetilde{E}(\mathcal{X}')$*. Entonces, se tiene que* $\mathrm{I}_f(\mathcal{X}') \geq \mathrm{I}_f(\mathcal{X})$*.*

Además, $\mathrm{I}_f(\mathcal{X}) = \mathrm{I}_f(\mathcal{X}')$ *si, y sólo si, o bien* $\mathcal{X} = \mathcal{X}'$ *en* Ω*, o bien* $\mathcal{X}' = \mathcal{X} \circ \sigma_{ll'}$*, o bien para todo* $\alpha \in [0,1]$ *se cumple que* $\inf \mathcal{X}'_\alpha(\omega_l) = \inf \mathcal{X}_\alpha(\omega_{l'})$, $\inf \mathcal{X}'_\alpha(\omega_{l'}) = \inf \mathcal{X}_\alpha(\omega_l)$, $\sup \mathcal{X}'_\alpha(\omega_l) = \sup \mathcal{X}_\alpha(\omega_l)$ *y* $\sup \mathcal{X}'_\alpha(\omega_{l'}) = \sup \mathcal{X}_\alpha(\omega_{l'})$*, o* $\inf \mathcal{X}'_\alpha(\omega_l) = \inf \mathcal{X}_\alpha(\omega_l)$, $\inf \mathcal{X}'_\alpha(\omega_{l'}) = \inf \mathcal{X}_\alpha(\omega_{l'})$, $\sup \mathcal{X}'_\alpha(\omega_l) = \sup \mathcal{X}_\alpha(\omega_{l'})$ *y* $\sup \mathcal{X}'_\alpha(\omega_{l'}) = \sup \mathcal{X}_\alpha(\omega_l)$*.*

El siguiente resultado corresponde a la propiedad que formaliza el *efecto* de la *"agrupación" de datos difusos* en la cuantificación del f-índice de desigualdad. Concretando más, esta propiedad expresa la "relación de orden" que existe entre la desigualdad de la población, y la desigualdad entre los grupos de una partición (clásica) dada de la población, cuando cada uno de los grupos está representado mediante su valor difuso esperado. De este resultado se puede concluir que la "agrupación" no aumenta el f-índice de desigualdad.

Teorema 2.2.15 (Efectos de la agrupación). *Se considera una población finita* $\Omega = \{\omega_{11},\ldots,\omega_{1N_1},\ldots\ldots,\omega_{M1},\ldots,\omega_{MN_M}\}$ *(con* $N = N_1 + \cdots + N_M$*) que se divide en* M *subpoblaciones* $\Omega_m = \{\omega_{m1},\ldots,\omega_{mN_m}\}$, $m = 1,\ldots,M$*, y supongamos que* $(\Omega, \mathcal{P}(\Omega))$ *está dotado con la distribución uniforme* P *y que* $\mathcal{P} = \{\Omega_m\}_{m=1}^M$ *denota la partición anterior. Si* $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ *es una variable aleatoria difusa asociada a* $(\Omega, \mathcal{P}(\Omega), P)$*, y* $\mathcal{X}_\mathcal{P} : \mathcal{P} \to \mathcal{F}_c((0,+\infty))$ *es la variable aleatoria difusa tal que* $\mathcal{X}_\mathcal{P}(\Omega_m) =$ *el valor esperado de* $\mathcal{X}$ *sobre* Ω_m *(*$m = 1,\ldots,M$*), y* $\mathcal{X}_{\Omega_m}$ *denota la restricción de* $\mathcal{X}$ *de* Ω *a* Ω_m *(*$m = 1,\ldots,M$*), entonces se cumple que:*

$$\mathrm{I}_f(\mathcal{X}) \geq \mathrm{I}_f(\mathcal{X}_\mathcal{P}).$$

Por otra parte, $\mathrm{I}_f(\mathcal{X})$ *es igual a* $\mathrm{I}_f(\mathcal{X}_{\mathcal{P}})$ *si, y sólo si, en la m-ésima subpoblación* $\mathcal{X}_{\Omega_m}$ *es una variable aleatoria difusa degenerada para cada* $m \in \{1, \ldots, M\}$.

Demostración:

Se cumple obviamente que $\widetilde{E}(\mathcal{X}) = \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)$. Por otro lado, aplicando la Desigualdad de Jensen, se obtiene que:

$$\sum_{m=1}^{M} \frac{N_m}{N} f\left(\frac{\inf \mathcal{X}_{\mathcal{P}}(\Omega_m)_\alpha}{\sup \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)_\alpha}\right) = \sum_{m=1}^{M} \frac{N_m}{N} f\left(\frac{\frac{1}{N_m}\sum_{i=1}^{N_m} \inf \mathcal{X}_\alpha(\omega_{mi})}{\sup E(\mathcal{X}_\alpha)}\right)$$

$$\leq \frac{1}{N}\sum_{m=1}^{M}\sum_{i=1}^{N_m} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{mi})}{\sup E(\mathcal{X}_\alpha)}\right) = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right).$$

Análogamente, se tiene que:

$$\sum_{m=1}^{M} \frac{N_m}{N} f\left(\frac{\sup \mathcal{X}_{\mathcal{P}}(\Omega_m)_\alpha}{\inf \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)_\alpha}\right) \leq \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right),$$

de donde se concluye, por lo tanto, que:

$$\mathrm{I}_f(\mathcal{X}) = \frac{1}{2}\int_{[0,1]} \left[\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right) + \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right)\right] d\,\alpha$$

$$\geq \frac{1}{2}\int_{[0,1]} \left[\sum_{m=1}^{M} \frac{N_m}{N} f\left(\frac{\sup \mathcal{X}_{\mathcal{P}}(\Omega_m)_\alpha}{\inf \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)_\alpha}\right)\right.$$

$$\left. + \sum_{m=1}^{M} \frac{N_m}{N} f\left(\frac{\inf \mathcal{X}_{\mathcal{P}}(\Omega_m)_\alpha}{\sup \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)_\alpha}\right)\right] d\,\alpha = \mathrm{I}_f(\mathcal{X}_{\mathcal{P}}).$$

La igualdad se alcanza si, y sólo si:

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right) + \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right)$$

$$= \sum_{m=1}^{M} \frac{N_m}{N} f\left(\frac{\sup \mathcal{X}_{\mathcal{P}}(\Omega_m)_\alpha}{\inf \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)_\alpha}\right) + \sum_{m=1}^{M} \frac{N_m}{N} f\left(\frac{\inf \mathcal{X}_{\mathcal{P}}(\Omega_m)_\alpha}{\sup \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)_\alpha}\right),$$

para todo $\alpha \in [0,1]$, lo cual ocurre si, y sólo si, se tiene la igualdad entre los respectivos extremos, lo que ocurre cuando en todas las expresiones a las que se aplica la Desigualdad de Jensen resultan igualdades. Esto ocurre si, y sólo si, cualquiera que sea la subpoblación Ω_m, $m \in \{1, \ldots, M\}$ y para cada $\alpha \in [0,1]$, los valores $\inf \mathcal{X}_\alpha(\omega_{mi})$ son iguales para $i = 1, \ldots, N_m$ y los valores $\sup \mathcal{X}_\alpha(\omega_{mi})$ son iguales para $i = 1, \ldots, N_m$, es decir, si los valores de cada una de las aplicaciones $\mathcal{X}_{\Omega_m}$ sobre la subpoblación m-ésima no dependen del individuo ω_{mi}. Por consiguiente, la variable $\mathcal{X}$ es una variable aleatoria difusa degenerada en cada subpoblación. □

Observación 2.2.2. La *descomponibilidad aditiva* de los índices definidos en la Página 19 de la introducción de esta memoria, se pierde al pasar al caso difuso, salvo para el índice de tipo Shannon. De este modo, si $f(x) = -\log x$ para todo $x \in (0, +\infty)$, se satisface que

$$\mathrm{I}_{Sh}(\mathcal{X}) = \mathrm{I}_{Sh}(\mathcal{X}_{\mathcal{P}}) + \sum_{m=1}^{M} \frac{N_m \mathrm{I}_{Sh}(\mathcal{X}_{\Omega_m})}{N},$$

es decir, que la desigualdad global coincide con la suma de la desigualdad entre los grupos (más específicamente, entre los valores esperados de $\mathcal{X}$ para los distintos grupos) y el promedio de la desigualdad dentro de los grupos.

En los ejemplos siguientes se ilustra el cálculo de ciertos f-índices de desigualdad y su empleo en la comparación de poblaciones.

Ejemplo 2.2.1. Se realiza un sondeo telefónico sobre la población Ω de los 105 socios varones de un centro deportivo a los que se les pide que se clasifiquen en uno de los cuatro grupos siguientes: BAJO, NO ALTO, ALTO y MUY ALTO. Supóngase que se obtienen como respuestas que 12 se han considerado BAJOS, 23 NO ALTOS, 57 ALTOS y 13 MUY ALTOS. Este tipo de clasificación puede identificarse con una variable aleatoria difusa cuyos valores son los cuatro grupos precedentes. Supongamos que para describir cada uno de esos valores se recurre a la caracterización dada por

Norwich & Turksen (1984), basada en un procedimiento de valoración directa media para ciertos puntos de referencia en $[54, 88]$ (en el supuesto de que las unidades son pulgadas) y en una interpolación lineal para los restantes puntos, que da lugar a las funciones poligonales de la Figura 2.1.

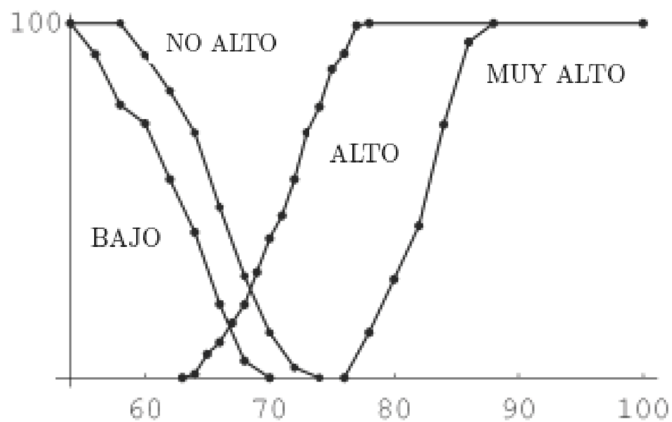


Fig. 2.1: Representación gráfica de ALTO y valores relacionados

Si quiere medirse la desigualdad de alturas de esa población sobre la base de ese sondeo, puede considerarse, por ejemplo, el f-índice que extiende la varianza normalizada del caso real (correspondiente a $f(x) = x^2 - 1$) vendría dado por:

$$\mathrm{I}_{\mathrm{NVar}}(\mathcal{X}_\Omega) = 0,04.$$

Ejemplo 2.2.2. Se considera la variable RENTA ANUAL, $\mathcal{X}$, según la clasificación que se adopta en cierto sistema para la concesión de ciertos créditos en algunas entidades. De acuerdo con Cox (1994) esa variable puede contemplarse como una variable cuyos valores (difusos) son $\tilde{x}_1 =$ POCO ALTA, $\tilde{x}_2 =$ MEDIANAMENTE ALTA, $\tilde{x}_3 =$ ALTA y $\tilde{x}_4 =$ MUY ALTA, donde $\tilde{x}_1$, $\tilde{x}_2$, $\tilde{x}_3$ y $\tilde{x}_4$ se describen (en dólares USA) mediante las S- y Π-curvas siguientes:

$$\tilde{x}_1 = 1 - S(100, 125),$$
$$\tilde{x}_2 = \Pi(100, 125, 150),$$
$$\tilde{x}_3 = \Pi(125, 147,5, 170),$$
$$\tilde{x}_4 = S(147,5, 170),$$

y se supone que sop $\tilde{x}_i \subset [90, 180]$, $i = 1, 2, 3, 4$, para los clientes que solicitan el préstamo según ese sistema (ver Figura 2.2).

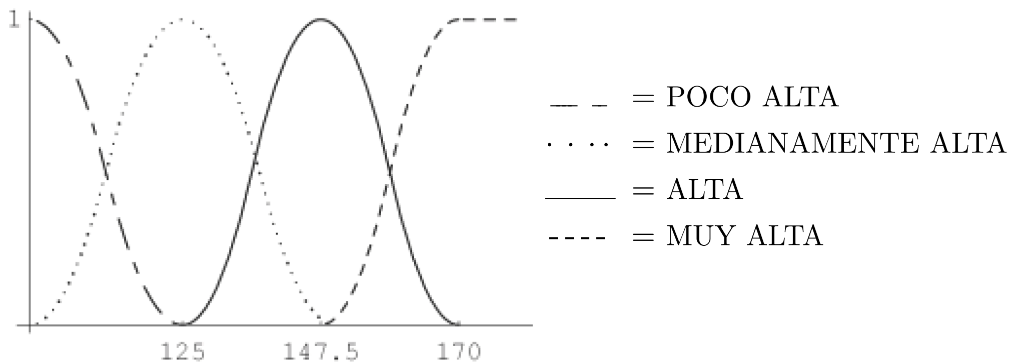


Fig. 2.2: Valores difusos de la variable RENTA ANUAL

Supongamos que una entidad que adopta el sistema anterior quiere comparar dos ciudades distintas a través de su desigualdad de la renta anual, para lo que observa los valores de la variable $\mathcal{X}$ en las oficinas centrales de esa entidad en cada una de esas dos ciudades.

Si el número de clientes que han solicitado crédito en una de sus oficinas (Ω_1) durante cierto periodo de tiempo es de 125, de los cuales 28 tienen una renta anual POCO ALTA, 43 MEDIANAMENTE ALTA, 31 ALTA y 23 MUY ALTA, mientras que el número de clientes que lo ha solicitado en la otra oficina (Ω_2) en el mismo periodo de tiempo es de 178, de los cuales 63 tienen una renta anual POCO ALTA, 79 MEDIANAMENTE ALTA, 27 ALTA y 9 MUY ALTA, y se emplea el f-índice de desigualdad con $f(x) = -\log x$, se obtiene que:

$$\mathrm{I}_{Sh}(\mathcal{X}_{\Omega_1}) = 4,892,$$

$$\mathrm{I}_{Sh}(\mathcal{X}_{\Omega_2}) = 4,815,$$

de modo que podemos concluir que las dos ciudades tienen un comportamiento muy similar en cuanto a la desigualdad de la renta anual.

2.3 Estimación insesgada del índice de desigualdad hiperbólico en muestreos aleatorios de poblaciones finitas

En esta sección se considera el problema de la estimación del índice de desigualdad poblacional asociado a una variable aleatoria difusa en una población finita, cuando sólo se dispone de la información proveniente de una muestra aleatoria extraída de tal población.

Para ello, va a comprobarse que es posible construir un estimador insesgado del índice hiperbólico (asociado a la función $f(x) = x^{-1} - 1$), para muestras de tamaño cualquiera en los muestreos aleatorios simple y con reposición. Sin embargo, para los demás índices de desigualdad introducidos en el presente capítulo, la construcción de estimadores insesgados resulta muy compleja y a menudo es inviable.

Supongamos que se considera una población finita Ω de N unidades, $\omega_1, \ldots, \omega_N$, y una variable aleatoria difusa $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ asociada a un espacio probabilístico definido sobre Ω que se considera dotada con la distribución uniforme.

Si se elige una muestra de tamaño n al azar y sin reposición a partir de Ω, υ representa una muestra aleatoria simple genérica de tamaño n, y $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ son las unidades en υ, entonces el *índice de desigualdad hiperbólico muestral* de $\mathcal{X}$ en la muestra υ viene dado por:

$$\mathrm{I}_H(\mathcal{X}[\upsilon]) = F\left(\frac{1}{n^2} \odot \bigoplus_{i=1}^{n} \bigoplus_{i'=1}^{n} \left(\mathcal{X}(\omega_{\upsilon i}) \oslash \mathcal{X}(\omega_{\upsilon i'})\right) \ominus 1\right)$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{i'=1}^{n} F\left(\mathcal{X}(\omega_{\upsilon i}) \oslash \mathcal{X}(\omega_{\upsilon i'})\right) - 1.$$

$\mathrm{I}_H(\mathcal{X}[\cdot])$ es una variable aleatoria real asociada al espacio probabilístico $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ (ver Capítulo 1, pág. 45), y por lo tanto define un estimador del *índice de desigualdad hiperbólico poblacional*, que corresponde al valor:

$$
\begin{aligned}
\mathrm{I}_H(\mathcal{X} \mid P) &= F\left(\frac{1}{N^2} \odot \sum_{j=1}^{N}\sum_{j'=1}^{N} \left(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_{j'})\right) \ominus 1\right) \\
&= \frac{1}{N^2} \sum_{j=1}^{N}\sum_{j'=1}^{N} F\left(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_{j'})\right) - 1.
\end{aligned}
$$

Para obtener a partir del índice muestral un estimador insesgado del poblacional, el primero debe corregirse. La corrección que va a aplicarse está basada en el resultado siguiente sobre el valor esperado del índice muestral:

Teorema 2.3.1 . *En el muestreo aleatorio simple de tamaño n a partir de la población Ω, si $f = n/N$ se cumple que:*

$$
E\left(\mathrm{I}_H\left(\mathcal{X}[\,\cdot\,]\right)\right) = \frac{(n-1)N}{n(N-1)}\mathrm{I}_H(\mathcal{X} \mid P) + \frac{1-f}{n(N-1)} \sum_{j=1}^{N} \mathrm{I}_H(\{\mathcal{X}(\omega_j)\}),
$$

donde $\mathrm{I}_H(\{\mathcal{X}(\omega_j)\})$ representa la (intra)desigualdad hiperbólica de la variable aleatoria difusa degenerada en el valor $\mathcal{X}(\omega_j) \in \mathcal{F}_c((0,+\infty))$.

Demostración:

En efecto:

$$
E\left(\mathrm{I}_H\left(\mathcal{X}[\,\cdot\,]\right)\right) = \sum_{\upsilon \in \Upsilon_n} p[\upsilon]\mathrm{I}_H\left(\mathcal{X}[\upsilon]\right)
$$

$$
= \sum_{\upsilon \in \Upsilon_n} p[\upsilon]\frac{1}{n^2} \sum_{j=1}^{N}\sum_{j'=1}^{N} F\left(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_{j'})\right) \cdot a_j[\upsilon]a_{j'}[\upsilon] - 1,
$$

con a_j las variables de Bernoulli introducidas en el Capítulo 1, pág. 48.

En consecuencia:

$$
E\left(\mathrm{I}_H\left(\mathcal{X}[\,\cdot\,]\right)\right) = \frac{1}{n^2} \sum_{j=1}^{N}\sum_{j'=1}^{N} F\left(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_{j'})\right) E\left(a_j\, a_{j'}\right) - 1
$$

$$
= \frac{1}{n^2} \sum_{j=1}^{N} F\left(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_j)\right) \frac{n}{N}
$$

$$+\frac{1}{n^2}\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N} F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_{j'})\right)\frac{n(n-1)}{N(N-1)}-1$$

$$=\frac{(n-1)N}{n(N-1)}\mathrm{I}_H(\mathcal{X})+\frac{1-f}{n(N-1)}\Big[\sum_{j=1}^{N} F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_j)\right)-1\Big]$$

$$=\frac{(n-1)N}{n(N-1)}\mathrm{I}_H(\mathcal{X}\,|\,P)+\frac{1-f}{n(N-1)}\mathrm{I}_H(\mathcal{X}(\omega_j)).$$

□

A partir del resultado que acaba de probarse, puede concluirse que:

Teorema 2.3.2. *En el muestreo aleatorio simple de tamaño n a partir de Ω, si $f=n/N$ se cumple que el estimador $\widehat{\mathrm{I}_H}(\mathcal{X}[\,\cdot\,])$ tal que:*

$$\widehat{\mathrm{I}_H}(\mathcal{X}[v])=\frac{n(N-1)}{N(n-1)}\mathrm{I}_H\left(\mathcal{X}[v]\right)-\frac{1-f}{n}\left\{\frac{1}{n-1}\sum_{i=1}^{n}\mathrm{I}_H\left(\{\mathcal{X}(\omega_{vi})\}\right)\right\},$$

es un estimador insesgado de $\mathrm{I}_H(\mathcal{X}\,|\,P)$.

Demostración:

En efecto, como la media muestral de los índices hiperbólicos *intra*valores cumple que es un estimador insesgado de la correspondiente media poblacional, es decir:

$$E\left(\frac{1}{n}\sum_{i=1}^{n}\mathrm{I}_H\left(\{\mathcal{X}(\omega_{\cdot i})\}\right)\right)=\frac{1}{N}\sum_{j=1}^{N}\mathrm{I}_H\left(\{\mathcal{X}(\omega_j)\}\right),$$

se tiene que:

$$E\left(\frac{1-f}{n}\left(\frac{1}{n}\sum_{i=1}^{n}\mathrm{I}_H\left(\{\mathcal{X}(\omega_{\cdot i})\}\right)\right)\right)=\frac{1-f}{N(n-1)}\sum_{j=1}^{N}\mathrm{I}_H\left(\{\mathcal{X}(\omega_j)\}\right),$$

y por el resultado del Teorema 2.3.1 se satisface que:

$$E\left(\widehat{\mathrm{I}_H}(\mathcal{X}[\,\cdot\,])\right)=\mathrm{I}_H(\mathcal{X}\,|\,P).$$

□

Para establecer la precisión del estimador de $\mathrm{I}_H(\mathcal{X}\,|\,P)$ anterior, determinamos el error cuadrático medio asociado, en este caso coincidente con la varianza $\mathrm{Var}\left(\widehat{\mathrm{I}_H}(\mathcal{X}[\,\cdot\,])\right)$, que se recoge a continuación:

Teorema 2.3.3. *En el muestreo aleatorio simple de tamaño n a partir de* Ω, *si* $f = n/N$ *se cumple que:*

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}(\mathcal{X}[\cdot])\right) = \frac{(1-f)}{n(n-1)N^2(N-1)(N-2)(N-3)}$$

$$\cdot\Bigg\{[N(6-4n)+6(n-1)]N^3(N-1)\left(\mathrm{I}_H(\mathcal{X}\,|\,P)\right)^2$$

$$+[4N^2(3-2n)+13N(n-1)+3(n-3)]N^2(N-1)\,\mathrm{I}_H(\mathcal{X}\,|\,P)$$

$$+[N^2(7-3n)+N(5n-7)-4(n-2)]\left(\sum_{j=1}^{N}\mathrm{I}_H(\{\mathcal{X}(\omega_j)\})\right)^2$$

$$+[N^2(n-5)+N(5n+1)-10(n-2)]\sum_{j=1}^{N}\left(\mathrm{I}_H(\{\mathcal{X}(\omega_j)\})\right)^2$$

$$+\,2[N^3(3-n)+N^2(3n-8)+N(n+9)-10(n-2)]\sum_{j=1}^{N}\mathrm{I}_H(\{\mathcal{X}(\omega_j)\})$$

$$+\,4(n-2)N^2(N-1)\mathrm{I}_H(\mathcal{X}\,|\,P)\sum_{j=1}^{N}\mathrm{I}_H(\{\mathcal{X}(\omega_j)\})$$

$$+(n-2)(N-1)(N-2)\sum_{j=1}^{N}\left(\sum_{l=1}^{N}[F(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_l))+F(\mathcal{X}(\omega_l)\oslash\mathcal{X}(\omega_j))]\right)^2$$

$$+(N-n+1)(N-1)(N-3)\sum_{j=1}^{N}\sum_{l=1}^{N}F(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_l))$$

$$\cdot\,[F(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_l))+F(\mathcal{X}(\omega_l)\oslash\mathcal{X}(\omega_j))]$$

$$-[N(3n-7)-3(n-3)](N-1)\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{I}_H(\{\mathcal{X}(\omega_l)\})F(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_l))$$

$$+[N^5(6-4n)+8N^4n+N^3(n-9)-6N^2(n-2)+N(5n+1)-10(n-2)]\Bigg\}.$$

Demostración:

Para simplificar las fórmulas, utilizaremos para todo $j, l \in \{1, \dots, N\}$ la siguiente notación:

$$F\left(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_l)\right) = \mathrm{V}_{jl}.$$

En consecuencia:

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}(\mathcal{X}[\,\cdot\,])\right) = \frac{1}{n^2(n-1)^2}\Bigg\{ \left(\frac{N-1}{N}\right)^2$$

$$\cdot \sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N}\sum_{l'=1}^{N} \mathrm{V}_{jj'}\mathrm{V}_{ll'} \ \mathrm{Cov}(a_j\, a_{j'},\, a_l\, a_{l'})$$

$$+ \left(\frac{N-n}{N}\right)^2 \sum_{j=1}^{N}\sum_{l=1}^{N} \mathrm{V}_{jj}\mathrm{V}_{ll} \ \mathrm{Cov}(a_j,\, a_l)$$

$$- 2\frac{(N-1)(N-n)}{N^2} \sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N} \mathrm{V}_{jj'}\mathrm{V}_{ll}\, \mathrm{Cov}(a_j\, a_{j'},\, a_l)\Bigg\}$$

$$= \frac{N-1}{nN(n-1)}\Bigg\{ \sum_{j=1}^{N} \mathrm{V}_{jj}^2 \,\mathrm{Var}(a_j^2) \ + \ 2\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{V}_{lj}\mathrm{V}_{jj} \ \mathrm{Cov}(a_l\, a_j,\, a_j^2)$$

$$+ 2\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{V}_{jl}\mathrm{V}_{jj} \ \mathrm{Cov}(a_j\, a_l,\, a_j^2) \ + \ \sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{V}_{jj}\mathrm{V}_{ll} \ \mathrm{Cov}(a_j^2,\, a_l^2)$$

$$+ \sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{V}_{jl}\mathrm{V}_{lj} \ \mathrm{Var}(a_j\, a_l) + \sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{V}_{jl}^2 \ \mathrm{Var}(a_j\, a_l)$$

$$+ 2\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,\, j'}}^{N} \mathrm{V}_{jj'}\mathrm{V}_{ll} \ \mathrm{Cov}(a_j\, a_{j'},\, a_l^2)$$

$$+ 2\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,\, j'}}^{N} \mathrm{V}_{jj'}\mathrm{V}_{lj} \ \mathrm{Cov}(a_j\, a_{j'},\, a_l\, a_j)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,\,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{lj'}\ \mathrm{Cov}(a_j\,a_{j'},\,a_l\,a_{j'})$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,\,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{jl}\ \mathrm{Cov}(a_j\,a_{j'},\,a_j\,a_l)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,\,j'}}^{N}\sum_{\substack{l'=1\\l'\neq l,\,j,\,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll'}\ \mathrm{Cov}(a_j\,a_{j'},\,a_l\,a_{l'})\Bigg\}$$

$$+\left(\frac{N-n}{nN(n-1)}\right)^2\left\{\sum_{j=1}^{N}\mathrm{V}_{jj}^2\ \mathrm{Var}(a_j)\ +\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(a_j,\,a_l)\right\}$$

$$-\,2\frac{(N-1)(N-n)}{[nN(n-1)]^2}\Bigg\{\sum_{j=1}^{N}\mathrm{V}_{jj}^2\ \mathrm{Cov}(a_j^2,\,a_j)\ +\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jl}\mathrm{V}_{jj}\ \mathrm{Cov}(a_j\,a_l,\,a_j)$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jl}\mathrm{V}_{ll}\ \mathrm{Cov}(a_j\,a_l,\,a_l)\ +\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(a_j^2,\,a_l)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,\,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll}\ \mathrm{Cov}(a_j\,a_{j'},\,a_l)\Bigg\},$$

y, en virtud de los momentos de las variables a_j (ver Cap. 1, pág. 48), se tiene que:

$$\mathrm{Var}\left(\widehat{\mathrm{I}}_H(\mathcal{X}[\,\cdot\,])\right)=\frac{1-f}{nN^3}\Bigg\{\sum_{j=1}^{N}\mathrm{V}_{jj}^2\ +\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\Bigg[2\,(\mathrm{V}_{jl}\mathrm{V}_{ll}+\mathrm{V}_{jl}\mathrm{V}_{jj})$$

$$-\frac{1}{N-1}\mathrm{V}_{jj}\mathrm{V}_{ll}+\frac{N+n-1}{n-1}\left(\mathrm{V}_{jl}\mathrm{V}_{lj}+\mathrm{V}_{jl}^2\right)\Bigg]+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,\,j'}}^{N}\Bigg[\frac{-4}{(N-2)}\mathrm{V}_{jj'}\mathrm{V}_{ll}$$

$$+\frac{[N(n-2)-2(n-1)]}{(n-1)(N-2)}\,(\mathrm{V}_{jj'}\mathrm{V}_{lj'}+\mathrm{V}_{jj'}\mathrm{V}_{jl}+2\mathrm{V}_{jj'}\mathrm{V}_{lj})\Bigg]$$

$$+\frac{[N(6-4n)+6(n-1)]}{(n-1)(N-2)(N-3)}\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\sum_{\substack{l'=1\\ l'\neq l,j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll'}\Bigg\}$$

$$=\frac{(1-f)}{n(n-1)N^2}\Bigg\{\frac{[N(6-4n)+6(n-1)]}{N(N-2)(N-3)}\left(\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jl}\right)^2$$

$$+\frac{n-2}{N-3}\left[\sum_{j=1}^{N}\left(\sum_{l=1}^{N}(\mathrm{V}_{jl}+\mathrm{V}_{lj})\right)^2\right]$$

$$+\frac{N-n+1}{N-2}\sum_{j=1}^{N}\sum_{l=1}^{N}\left(\mathrm{V}_{jl}^2+\mathrm{V}_{jl}\mathrm{V}_{lj}\right)+\frac{4(n-2)}{(N-2)(N-3)}\left(\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jl}\right)\left(\sum_{j=1}^{N}\mathrm{V}_{jj}\right)$$

$$-\frac{N(3n-7)-3(n-3)}{(N-2)(N-3)}\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jl}\mathrm{V}_{ll}$$

$$+\frac{N^2(7-3n)+N(5n-7)-4(n-2)}{(N-1)(N-2)(N-3)}\left(\sum_{j=1}^{N}\mathrm{V}_{jj}\right)^2$$

$$+\frac{N^2(n-5)+N(5n+1)-10(n-2)}{(N-1)(N-2)(N-3)}\sum_{j=1}^{N}\mathrm{V}_{jj}^2\Bigg\}.$$

Como $\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jl} = N^2\left(\mathrm{I}_H(\mathcal{X}\,|\,P)+1\right)$ y $\mathrm{V}_{jj} = \mathrm{I}_H(\{\mathcal{X}(\omega_j)\})$ para todo $j\in\{1,\ldots,N\}$, puede concluirse el resultado del presente teorema. □

Si en vez de adoptar una selección aleatoria sin reposición, se considera una selección aleatoria con reposición de n unidades de la población $\Omega = \{\omega_1,\ldots,\omega_N\}$, y υ representa una muestra aleatoria con reposición genérica de tamaño n, y $\omega_{\upsilon 1},\ldots,\omega_{\upsilon n}$ son las unidades en υ, entonces el *índice de desigualdad hiperbólico muestral* de $\mathcal{X}$ en la muestra υ viene dado ahora por:

$$\mathrm{I}_H(\mathcal{X}[\upsilon]) = \frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_{j'})\right)\cdot t_j[\upsilon]t_{j'}[\upsilon]-1,$$

con t_j las variables aleatorias introducidas en el Capítulo 1, pág. 51.

$\mathrm{I}_H(\mathcal{X}[\,\cdot\,])$ es una variable aleatoria real asociada al espacio probabilístico $(\Upsilon_n^w, \mathcal{P}(\Upsilon_n^w), p^w)$ (ver Capítulo 1, pág. 51), y por lo tanto define un estimador del índice hiperbólico poblacional.

Al igual que en el muestreo aleatorio simple, para construir un estimador insesgado del índice poblacional a partir del muestral, se examina en primer término el valor esperado del índice hiperbólico muestral en el muestreo con reposición.

Teorema 2.3.4. *En el muestreo aleatorio con reposición de tamaño n a partir de la población Ω, se cumple que:*

$$E\left(\mathrm{I}_H(\mathcal{X}[\,\cdot\,])\right) = \frac{(n-1)}{n}\mathrm{I}_H(\mathcal{X}\,|\,P) + \frac{1}{nN}\sum_{j=1}^{N}\mathrm{I}_H\left(\{\mathcal{X}(\omega_j)\}\right).$$

Demostración:

En efecto:

$$E\left(\mathrm{I}_H(\mathcal{X}[\,\cdot\,])\right) = \sum_{\upsilon\in\Upsilon_n^w} p^w[\upsilon]\,\mathrm{I}_H(\mathcal{X}[\upsilon])$$

$$= \sum_{\upsilon\in\Upsilon_n^w} p^w[\upsilon]\frac{1}{n^2}\sum_{j=1}^{N}\sum_{j'=1}^{N} F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_{j'})\right)\cdot t_j[\upsilon]t_{j'}[\upsilon] - 1.$$

En consecuencia:

$$E\left(\mathrm{I}_H(\mathcal{X}[\,\cdot\,])\right) = \frac{1}{n^2}\sum_{j=1}^{N}\sum_{j'=1}^{N} F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_{j'})\right)\,E\left(t_j\,t_{j'}\right) - 1$$

$$= \frac{1}{n^2}\sum_{j=1}^{N} F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_j)\right)\,\frac{n(N+n-1)}{N^2}$$

$$+\frac{1}{n^2}\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N} F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_{j'})\right)\,\frac{n(n-1)}{N^2} - 1$$

$$= \frac{(n-1)}{n}\mathrm{I}_H(\mathcal{X}\,|\,P) + \frac{1}{nN}\sum_{j=1}^{N}\mathrm{I}_H\left(\{\mathcal{X}(\omega_j)\}\right).$$

□

A partir del resultado que acaba de probarse, puede concluirse que:

Teorema 2.3.5. *En el muestreo aleatorio con reposición de tamaño n a partir de* Ω, *se cumple que el estimador* $\widehat{\mathrm{I}_H}^w(\mathcal{X}[\cdot])$ *tal que:*

$$\widehat{\mathrm{I}_H}^w(\mathcal{X}[v]) = \frac{n}{(n-1)}\mathrm{I}_H(\mathcal{X}[v]) - \frac{1}{n}\left\{\frac{1}{n-1}\sum_{i=1}^{n}\mathrm{I}_H(\{\mathcal{X}(\omega_{vi})\})\right\},$$

es un estimador insesgado de $\mathrm{I}_H(\mathcal{X}\,|\,P)$.

Demostración:

En efecto, como en el muestreo aleatorio simple, la media muestral de los índices hiperbólicos *intra*valores cumple que es un estimador insesgado de la correspondiente media poblacional, se tiene que:

$$E\left(\frac{1}{n}\left\{\frac{1}{n-1}\sum_{i=1}^{n}\mathrm{I}_H(\{\mathcal{X}(\omega_{\cdot i})\})\right\}\right) = \frac{1}{N(n-1)}\sum_{j=1}^{N}\mathrm{I}_H(\{\mathcal{X}(\omega_j)\}),$$

y por el resultado del Teorema 2.3.4 se satisface que:

$$E\left(\widehat{\mathrm{I}_H}^w(\mathcal{X}[\cdot])\right) = \mathrm{I}_H(\mathcal{X}\,|\,P).$$

□

El error cuadrático medio asociado al estimador del Teorema 2.3.5, coincidente con la varianza $\mathrm{Var}\left(\widehat{\mathrm{I}_H}^w(\mathcal{X}[\cdot])\right)$, y vendrá dado por:

Teorema 2.3.6. *En el muestreo aleatorio con reposición de tamaño n a partir de* Ω, *se cumple que:*

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}^w(\mathcal{X}[\cdot])\right) = \frac{1}{n(n-1)N^3}$$

$$\cdot\left\{(6-4n)N^3\left(\mathrm{I}_H(\mathcal{X}\,|\,P)\right)^2 + 2(6-4n)N^3\mathrm{I}_H(\mathcal{X}\,|\,P)\right.$$

$$-[N-4(n-1)]\left(\sum_{j=1}^{N}\mathrm{I}_H(\{\mathcal{X}(\omega_j)\})\right)^2 + [N-4(n-1)]\sum_{j=1}^{N}\left(\mathrm{I}_H(\{\mathcal{X}(\omega_j)\})\right)^2$$

$$-2[N-4(n-1)](N-1)\sum_{j=1}^{N}\mathrm{I}_H(\{\mathcal{X}(\omega_j)\})$$

$$+(n-2)\sum_{j=1}^{N}\left(\sum_{l=1}^{N}\left[F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_l)\right)+F\left(\mathcal{X}(\omega_l)\oslash\mathcal{X}(\omega_j)\right)\right]\right)^2$$

$$+N\sum_{j=1}^{N}\sum_{l=1}^{N}F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_l)\right)\left[F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_l)\right)+F\left(\mathcal{X}(\omega_l)\oslash\mathcal{X}(\omega_j)\right)\right]$$

$$\left.+[N^2(5-4n)+N(4n-3)-4(n-1)]N\right\}.$$

Demostración:

Para simplificar las fórmulas, utilizaremos la notación V_{jl} del Teorema 2.3.3.

$$\mathrm{Var}\left(\widehat{\mathrm{I}}_H^{\,w}(\mathcal{X}[\,\cdot\,])\right)=\frac{1}{n^2(n-1)^2}\left\{\sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N}\sum_{l'=1}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll'}\ \mathrm{Cov}(t_j\,t_{j'},\,t_l\,t_{l'})\right.$$

$$\left.+\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j,\,t_l)\ -\ 2\sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll}\,\mathrm{Cov}(t_j\,t_{j'},\,t_l)\right\}$$

$$=\frac{1}{n^2(n-1)^2}\left\{\sum_{j=1}^{N}\mathrm{V}_{jj}^2\,\mathrm{Var}(t_j^2)\ +\ 2\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{lj}\mathrm{V}_{jj}\ \mathrm{Cov}(t_l\,t_j,\,t_j^2)\right.$$

$$+\,2\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jl}\mathrm{V}_{jj}\ \mathrm{Cov}(t_j\,t_l,\,t_j^2)\ +\ \sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j^2,\,t_l^2)$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jl}\mathrm{V}_{lj}\ \mathrm{Var}(t_j\,t_l)+\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jl}^2\ \mathrm{Var}(t_j\,t_l)$$

$$+\,2\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,\,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j\,t_{j'},\,t_l^2)+2\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,\,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{lj}\ \mathrm{Cov}(t_j\,t_{j'},\,t_l\,t_j)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,\,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{lj'}\ \mathrm{Cov}(t_j\,t_{j'},\,t_l\,t_{j'})+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,\,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{jl}\ \mathrm{Cov}(t_j\,t_{j'},\,t_j\,t_l)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\sum_{\substack{l'=1\\ l'\neq l,j,j'}}^{N} \mathrm{V}_{jj'}\mathrm{V}_{ll'}\ \mathrm{Cov}(t_j\, t_{j'},\, t_l\, t_{l'})$$

$$+\sum_{j=1}^{N}\mathrm{V}_{jj}^2\ \mathrm{Var}(t_j)\ +\ \sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j,\, t_l)$$

$$-\,2\sum_{j=1}^{N}\mathrm{V}_{jj}^2\ \mathrm{Cov}(t_j^2,\, t_j)\ -\,2\ \sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{V}_{jl}\mathrm{V}_{jj}\ \mathrm{Cov}(t_j\, t_l,\, t_j)$$

$$-2\ \sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{V}_{jl}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j\, t_l,\, t_l)\ -\,2\ \sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j^2,\, t_l)$$

$$\left.-2\ \sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j\, t_{j'},\, t_l)\right\}.$$

En virtud de las expresiones de los momentos de las variables t_j (ver Cap. 1, pág. 51), se tiene que:

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}^{w}(\mathcal{X}[\,\cdot\,])\right) = \frac{1-f}{n^2(n-1)^2}\left\{\frac{n(n-1)(N-1)[2N+(4n-6)]}{N^4}\sum_{j=1}^{N}\mathrm{V}_{jj}^2\right.$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\left[\frac{2n(n-1)}{N^4}[2N(n-2)+(6-4n)]\,(\mathrm{V}_{jl}\mathrm{V}_{ll}+\mathrm{V}_{jl}\mathrm{V}_{jj})\right.$$

$$+\frac{n}{N^4}[-N^2(n+1)+4N(n-1)^2+(6-4n)(n-1)]\mathrm{V}_{jj}\mathrm{V}_{ll}$$

$$\left.+\frac{n(n-1)}{N^4}[N(N+2n-4)+(6-4n)]\left(\mathrm{V}_{jl}\mathrm{V}_{lj}+\mathrm{V}_{jl}^2\right)\right]$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\left[\frac{2n(n-1)(6-4n)}{(N^4)}\mathrm{V}_{jj'}\mathrm{V}_{ll}\right.$$

$$\left.+\frac{n(n-1)}{N^4}[N(n-2)+(6-4n)]\,(\mathrm{V}_{jj'}\mathrm{V}_{lj'}+\mathrm{V}_{jj'}\mathrm{V}_{jl}+2\mathrm{V}_{jj'}\mathrm{V}_{lj})\right]$$

$$+\frac{n(n-1)(6-4n)}{N^4}\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\sum_{\substack{l'=1\\ l'\neq l,j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll'}\Bigg\}$$

$$=\frac{1}{n(n-1)N^2}\Bigg\{\frac{6-4n}{N^2}\left(\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jl}\right)^2$$

$$+\frac{n-2}{N}\sum_{j=1}^{N}\left(\sum_{l=1}^{N}(\mathrm{V}_{jl}+\mathrm{V}_{lj})\right)^2+\sum_{j=1}^{N}\sum_{l=1}^{N}\left(\mathrm{V}_{jl}^2+\mathrm{V}_{jl}\mathrm{V}_{lj}\right)$$

$$-\frac{N(n-1)-4(n-1)^2}{N(n-1)}\left(\left(\sum_{j=1}^{N}\mathrm{V}_{jj}\right)^2-\sum_{j=1}^{N}\mathrm{V}_{jj}^2\right)\Bigg\},$$

y razonando como en el Teorema 2.3.3 se obtiene el resultado del presente teorema. □

Al igual que se señaló en la Sección 1.3, los resultados que acaban de establecerse pueden emplearse con el fin de comparar la precisión en la estimación del índice de desigualdad hiperbólico poblacional asociado a distintas variables, construir intervalos de confianza y procedimientos de contraste de hipótesis para el índice anterior y determinar tamaños muestrales adecuados, si bien los procedimientos serían conservadores y requerirían la estimación de los elementos poblacionales en la varianza de los estimadores.

2.4 Distribución asintótica de los índices de desigualdad muestrales en poblaciones finitas. Aplicación al desarrollo de procedimientos inferenciales aproximados para los índices poblacionales

En la sección anterior se ha verificado que el índice de desigualdad hiperbólico admite un estimador insesgado en el muestreo de poblaciones finitas. Esta afirmación no puede extenderse inmediatamente a los restantes índices, bien porque no resulta viable su estimación insesgada, bien porque tal estimación entraña una complejidad elevada. Conclusiones análogas pueden obtenerse en relación con la determinación de la precisión de las estimaciones y con las restantes inferencias realizadas.

Sin embargo, el estudio de la distribución asintótica de los índices de desigualdad muestrales va a permitir bajo ciertas condiciones la estimación asintóticamente insesgada, la determinación de la precisión asintótica de las estimaciones, y el desarrollo de aproximaciones de otras inferencias, para la mayoría de los índices poblacionales.

Como implicaciones de este estudio y de los procedimientos inferenciales aproximados que se derivan del mismo, cabe destacar su aplicación cuando se consideran grandes muestras extraídas al azar y con reemplazamiento de una población cualquiera, o extraídas al azar y sin reemplazamiento de una población de tamaño suficientemente más grande que el tamaño muestral.

Supongamos que se considera una población finita $\Omega = \{\omega_1, \ldots, \omega_N\}$ de tamaño N, y que $\mathcal{X}$ es una variable aleatoria difusa positiva. Si sobre la población Ω se admite que la variable $\mathcal{X}$ toma r valores distintos, $\tilde{x}_1^*, \ldots, \tilde{x}_r^* \in \mathcal{F}_c((0, +\infty))$, y para $l \in \{1, \ldots, r\}$ se denota por p_l la probabilidad $P(\{\omega \in \Omega \mid \mathcal{X}(\omega) = \tilde{x}_l^*\})$, donde P es la medida correspondiente a la distribución uniforme sobre Ω. Si $f : (0, +\infty) \to \mathbb{R}$ es una función monótona, cóncava hacia arriba en el sentido estricto y de clase C^1, que satisface que $f(u) + f(1/u) \geq 0$ para todo $u \in (0, +\infty)$ y $f(1) = 0$, el *f-índice de desigualdad poblacional* de $\mathcal{X}$ viene dado por el valor:

$$\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) = F\left(\sum_{l'=1}^{r} p_{l'} \odot f\left(\tilde{x}_{l'}^* \oslash \left(\sum_{l=1}^{r} p_l \odot \tilde{x}_l^*\right)\right)\right)$$
$$= \sum_{l'=1}^{r} p_{l'} F\left(f\left(\tilde{x}_{l'}^* \oslash \left(\sum_{l=1}^{r} p_l \odot \tilde{x}_l^*\right)\right)\right).$$

Si se selecciona al azar una muestra de tamaño n de la población y f_{nl} representa la frecuencia relativa del valor $\tilde{x}_l^*$ de $\mathcal{X}$ en esa muestra, el *f-índice de desigualdad muestral* de $\mathcal{X}$ corresponde al valor:

$$\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) = F\left(\sum_{l'=1}^{r} f_{nl'} \odot f\left(\tilde{x}^*_{l'} \oslash \left(\sum_{l=1}^{r} f_{nl} \odot \tilde{x}^*_l\right)\right)\right)$$
$$= \sum_{l'=1}^{r} f_{nl'} F\left(f\left(\tilde{x}^*_{l'} \oslash \left(\sum_{l=1}^{r} f_{nl} \odot \tilde{x}^*_l\right)\right)\right).$$

El siguiente teorema recoge la distribución asintótica del índice muestral.

Teorema 2.4.1. *Para cada $n \in \mathbb{N}$, se consideran n variables aleatorias difusas independientes e idénticamente distribuidas que la variable aleatoria difusa $\mathcal{X}$ (es decir, una muestra aleatoria simple de tamaño n a partir de $\mathcal{X}$), definida sobre la población finita $\Omega = \{\omega_1, \ldots, \omega_N\}$ de forma que $P(\{\omega \in \Omega \mid \mathcal{X}(\omega) = \tilde{x}^*_l\}) = p_l$ con $p_l \in (0,1)$, $l = 1, \ldots, r$, $\sum_{l=1}^{r} p_l = 1$. Sea $f : (0,+\infty) \to \mathbb{R}$ una función monótona y cóncava hacia arriba en sentido estricto, que satisface que $f(u) + f(1/u) \geq 0$ para todo $u \in (0,+\infty)$ y que $f(1) = 0$, y f admite derivadas de tercer orden finitas. Se cumple entonces que:*

i) Si $\mathbf{f}_n = (f_{n1}, \ldots, f_{n(r-1)}) \in [0,1]^{r-1}$, con f_{nl} = frecuencia relativa de $\tilde{x}^*_l$ en la correspondiente realización de la muestra aleatoria simple de tamaño n ($l = 1, \ldots, r-1$, $f_{nr} = 1 - \sum_{l=1}^{r-1} f_{nl}$), e $\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)$ es el índice de desigualdad muestral asociado, entonces $\{\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)\}_n$ es una sucesión de estimadores de $\mathrm{I}_f(\mathcal{X}) = \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})$, que es fuertemente consistente, es decir, cuando $n \to \infty$:

$$\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) \xrightarrow{c.s.} \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}),$$

cualquiera que sea $\mathbf{p} = (p_1, \ldots, p_{r-1})$ (con $p_1, \ldots, p_{r-1} \in (0,1)$ y $\sum_{l=1}^{r-1} p_l < 1$).

ii) $\{\sqrt{n}\,(\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}))\}_n$ es una sucesión de variables aleatorias que converge en ley hacia una normal unidimensional $N(0, \sigma^2(\mathbf{p}))$, *con:*

$$\sigma^2(\mathbf{p}) = \sum_{l'=1}^{r} p_{l'} (V_{l'}^*)^2 - \left(\sum_{l'=1}^{r} p_{l'} V_{l'}^* \right)^2 ,$$

donde:

$$V_{l'}^* = F\left(f\left(\tilde{x}_{l'}^* \oslash E(\mathcal{X} \,|\, \mathbf{p})\right)\right) - \sum_{l=1}^{r} p_l \left[\frac{1}{2} \int_{[0,1]} \left\{ \inf(\tilde{x}_{l'}^*)_\alpha \right.\right.$$

$$\cdot f'\left(\frac{\sup(\tilde{x}_l^*)_\alpha}{E\left(\inf \mathcal{X}_\alpha \,|\, \mathbf{p}\right)} \right) \cdot \left(\frac{\sup(\tilde{x}_l^*)_\alpha}{\left[E\left(\inf \mathcal{X}_\alpha \,|\, \mathbf{p}\right)\right]^2} \right)$$

$$\left.\left. + \sup(\tilde{x}_{l'}^*)_\alpha \cdot f'\left(\frac{\inf(\tilde{x}_l^*)_\alpha}{E\left(\sup \mathcal{X}_\alpha \,|\, \mathbf{p}\right)} \right) \cdot \left(\frac{\inf(\tilde{x}_l^*)_\alpha}{\left[E\left(\sup \mathcal{X}_\alpha \,|\, \mathbf{p}\right)\right]^2} \right) \right\} d\,\alpha \right],$$

siempre que $\sigma^2(\mathbf{p}) > 0$.

iii) *Si* $\sigma^2(\mathbf{p}) = 0$, *y para algún par* (i, j) *con* $i, j \in \{1, \ldots, r-1\}$ *se cumple que:*

$$h_{ij} = \frac{\partial^2 \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})}{\partial p_i \partial p_j} = \frac{\partial}{\partial p_i} [\, V_j^* - V_r^*] > 0,$$

entonces $\{2n\,(\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) - \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}))\}_n$ *es una sucesión de variables aleatorias que converge en ley hacia una combinación lineal de, a lo sumo,* $r-1$ *variables chi-cuadrado* χ_1^2 *independientes.*

Demostración:

Si se revisa la introducción de la demostración del Teorema 1.4.1, se concluye que las condiciones satisfechas por $\mathbf{f}_n$, $\mathbf{p}$, el espacio paramétrico y el conjunto de los valores de la variable $\mathcal{X}$, garantizan que:

i) $\mathbf{f}_n \stackrel{c.s.}{\rightarrow} \mathbf{p}$ y al ser $\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})$ continua en un entorno de $\mathbf{p}$, se concluye que:

$$\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) \stackrel{c.s.}{\rightarrow} \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}),$$

cuando $n \to \infty$, es decir, $\{\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n)\}_n$ es una sucesión estimadora de $\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})$ fuertemente consistente.

ii) Por las condiciones supuestas, para n suficientemente grande es posible desarrollar $\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n)$ en un entorno de $\mathbf{p}$. De este modo, el desarrollo de Taylor de primer orden de $\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n)$ será:

$$\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) = \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) + \nabla \mathrm{I}_f(\mathbf{p})(\mathbf{f}_n - \mathbf{p})^t$$
$$+\frac{1}{2}(\mathbf{f}_n - \mathbf{p}) H(\mathrm{I}_f(\mathbf{p}_n^*))\,(\mathbf{f}_n - \mathbf{p})^t,$$

con $\nabla \mathrm{I}_f(\mathbf{p})$ = vector gradiente de $\mathrm{I}_f(\mathcal{X} \,|\, \cdot)$ en $\mathbf{p}$, y $H(\mathrm{I}_f(\mathbf{p}))$ = matriz hessiana $(r-1) \times (r-1)$ dada por:

$$H(\mathrm{I}_f(\mathbf{p}_n^*)) = \left[\left(\frac{\partial^2 \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})}{\partial p_i \partial p_j} \right)_{\mathbf{p}=\mathbf{p}_n^*} \right],$$

y con $\mathbf{p}_n^* \in \mathbb{P}$ tal que $\|\mathbf{p}_n^* - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$.

Por lo tanto:

$$\sqrt{n}\,[\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) - \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})]$$
$$= \nabla \mathrm{I}_f(\mathbf{p}) \left(\sqrt{n}(\mathbf{f}_n - \mathbf{p}) \right)^t + \frac{1}{2}(\mathbf{f}_n - \mathbf{p}) H(\mathrm{I}_f(\mathbf{p}_n^*)) \left(\sqrt{n}(\mathbf{f}_n - \mathbf{p}) \right)^t.$$

El vector gradiente vendrá dado por la matriz fila:

$$\nabla \mathrm{I}_f(\mathbf{p}) = \left(\frac{\partial}{\partial p_1} \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) \ \cdots \ \frac{\partial}{\partial p_{r-1}} \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) \right),$$

y, para $i \in \{1, \ldots, r-1\}$ se cumple que:

$$\frac{\partial}{\partial p_i} \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) = \frac{\partial}{\partial p_i} \left[\sum_{l=1}^{r-1} p_l \left\{ F\Big(f\left(\tilde{x}_l^* \oslash E(\mathcal{X} \,|\, \mathbf{p}) \right) \Big) \right. \right.$$
$$\left. \left. -F\Big(f\left(\tilde{x}_r^* \oslash E(\mathcal{X} \,|\, \mathbf{p}) \right) \Big) \right\} + F\Big(f\left(\tilde{x}_r^* \oslash E(\mathcal{X} \,|\, \mathbf{p}) \right) \Big) \right]$$
$$= F\Big(f\left(\tilde{x}_i^* \oslash E(\mathcal{X} \,|\, \mathbf{p}) \right) \Big) - F\Big(f\left(\tilde{x}_r^* \oslash E(\mathcal{X} \,|\, \mathbf{p}) \right) \Big)$$
$$+ \sum_{l=1}^{r-1} p_l \left\{ \frac{\partial}{\partial p_i} F\Big(f\left(\tilde{x}_l^* \oslash E(\mathcal{X} \,|\, \mathbf{p}) \right) \Big) \right.$$

$$-\frac{\partial}{\partial p_i} F\Big(f\,(\tilde{x}_r^* \oslash E(\mathcal{X}\,|\,\mathbf{p}))\Big)\Big\} + \frac{\partial}{\partial p_i} F\Big(f\,(\tilde{x}_r^* \oslash E(\mathcal{X}\,|\,\mathbf{p}))\Big).$$

Si $l \in \{1, \ldots, r\}$ se tiene que:

$$\frac{\partial}{\partial p_i} F\Big(f\,(\tilde{x}_l^* \oslash E(\mathcal{X}\,|\,\mathbf{p}))\Big)$$

$$= \frac{1}{2}\int_{[0,1]}\left\{\frac{\partial}{\partial p_i} f\left(\sup(\tilde{x}_l^*)_\alpha \Big/ \sum_{l'=1}^{r-1} p_{l'}\left[\inf(\tilde{x}_{l'}^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right]\right)\right.$$

$$\left. + \frac{\partial}{\partial p_i} f\left(\inf(\tilde{x}_l^*)_\alpha \Big/ \sum_{l'=1}^{r-1} p_{l'}\left[\sup(\tilde{x}_{l'}^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right]\right)\right\} d\,\alpha$$

$$= \frac{1}{2}\int_{[0,1]}\left\{\left[f'\left(\sup(\tilde{x}_l^*)_\alpha \Big/ \sum_{l'=1}^{r-1} p_{l'}\left[\inf(\tilde{x}_{l'}^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right]\right)\right.\right.$$

$$\cdot \left(-\sup(\tilde{x}_l^*)_\alpha \Big/ \left[\sum_{l'=1}^{r-1} p_{l'}\left(\inf(\tilde{x}_{l'}^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right)\right]^2\right)$$

$$\left.\cdot\,\left(\inf(\tilde{x}_i^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right)\right] + \left[f'\left(\inf(\tilde{x}_l^*)_\alpha \Big/ \sum_{l'=1}^{r-1} p_{l'}\left[\sup(\tilde{x}_{l'}^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right]\right)\right.$$

$$\cdot \left(-\inf(\tilde{x}_l^*)_\alpha \Big/ \left[\sum_{l'=1}^{r-1} p_{l'}\left(\sup(\tilde{x}_{l'}^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right)\right]^2\right)$$

$$\left.\left.\cdot\,\left(\sup(\tilde{x}_i^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right)\right]\right\} d\,\alpha.$$

En consecuencia, si $i \in \{1, \ldots, r-1\}$, se cumple que:

$$\frac{\partial}{\partial p_i} \mathrm{I}_{\,f}(\mathcal{X}\,|\,\mathbf{p}) = F\Big(f\,(\tilde{x}_i^* \oslash E(\mathcal{X}\,|\,\mathbf{p}))\Big) - F\Big(f\,(\tilde{x}_r^* \oslash E(\mathcal{X}\,|\,\mathbf{p}))\Big)$$

$$+\frac{1}{2}\sum_{l=1}^{r} p_l \cdot \int_{[0,1]}\left\{f'\left(\sup(\tilde{x}_l^*)_\alpha \Big/ \sum_{l'=1}^{r-1} p_{l'}\left[\inf(\tilde{x}_{l'}^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right]\right)\right.$$

$$\cdot\left(-\sup(\tilde{x}_l^*)_\alpha \Big/ \left[\sum_{l'=1}^{r-1} p_{l'}\left(\inf(\tilde{x}_{l'}^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right)\right]^2\right)\left(\inf(\tilde{x}_i^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right)$$

$$+f'\left(\inf(\tilde{x}_l^*)_\alpha \Big/ \sum_{l'=1}^{r-1} p_{l'}\left[\sup(\tilde{x}_{l'}^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right]\right)$$

$$\cdot\left(-\inf(\tilde{x}_l^*)_\alpha \Big/ \left[\sum_{l'=1}^{r-1} p_{l'}\left(\sup(\tilde{x}_{l'}^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right)\right]^2\right)$$

$$\cdot\left(\sup(\tilde{x}_i^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right)\Big\}d\,\alpha = \mathrm{V}_i^* - \mathrm{V}_r^*,$$

con:

$$\mathrm{V}_{l'}^* = F\left(f\left(\tilde{x}_{l'}^* \oslash E(\mathcal{X}\,|\,\mathbf{p})\right)\right)$$
$$-\sum_{l=1}^{r} p_l\left[\frac{1}{2}\int_{[0,1]}\left\{\inf(\tilde{x}_{l'}^*)_\alpha f'\left(\frac{\sup(\tilde{x}_l^*)_\alpha}{E\left(\inf\mathcal{X}_\alpha\,|\,\mathbf{p}\right)}\right)\cdot\left(\frac{\sup(\tilde{x}_l^*)_\alpha}{\left[E\left(\inf\mathcal{X}_\alpha\,|\,\mathbf{p}\right)\right]^2}\right)\right.\right.$$
$$\left.\left.+\sup(\tilde{x}_{l'}^*)_\alpha\cdot f'\left(\frac{\inf(\tilde{x}_l^*)_\alpha}{E\left(\sup\mathcal{X}_\alpha\,|\,\mathbf{p}\right)}\right)\cdot\left(\frac{\inf(\tilde{x}_l^*)_\alpha}{\left[E\left(\sup\mathcal{X}_\alpha\,|\,\mathbf{p}\right)\right]^2}\right)\right\}d\,\alpha\right]$$

para $l' \in \{1,\ldots,r\}$.

Como $\sqrt{n}(\mathbf{f}_n - \mathbf{p})$ es asintóticamente normal $(r-1)$-dimensional $N\Big(\mathbf{0}, \left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}\Big)$, según las propiedades de la convergencia en ley, se tiene que cuando $n\to\infty$:

$$\nabla\mathrm{I}_f(\mathbf{p})\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \xrightarrow{\mathcal{L}} N\left(0,\, \nabla\mathrm{I}_f(\mathbf{p})\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}\left(\nabla\mathrm{I}_f(\mathbf{p})\right)^t\right),$$

siempre que $\nabla\mathrm{I}_f(\mathbf{p})\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}\left(\nabla\mathrm{I}_f(\mathbf{p})\right)^t > 0$.

En virtud de las propiedades relativas a las convergencias en ley y en probabilidad, al ser $\|\mathbf{p}_n^* - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$ se tiene que cuando $n\to\infty$:

$$\left(\frac{\partial^2\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i\partial p_j}\right)_{\mathbf{p}=\mathbf{p}_n^*} \xrightarrow{p} \frac{\partial^2\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i\partial p_j}$$

($i,j \in \{1,\ldots,r-1\}$). Por otro lado, $\mathbf{f}_n \overset{c.s.}{\to} \mathbf{p}$, y por lo tanto $\mathbf{f}_n \overset{p}{\to} \mathbf{p}$, cuando $n \to \infty$, de donde:

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p})H(\mathrm{I}_f(\mathbf{p}_n^*)) \overset{p}{\to} \mathbf{0},$$

y, como $\sqrt{n}(\mathbf{f}_n - \mathbf{p}) \overset{\mathcal{L}}{\to} N\left(\mathbf{0},\ \left[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\right]^{-1}\right)$, se tiene que:

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p})H(\mathrm{I}_f(\mathbf{p}_n^*)) \left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \overset{\mathcal{L}}{\to} 0,$$

y, en consecuencia:

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p})H(\mathrm{I}_f(\mathbf{p}_n^*)) \left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \overset{p}{\to} 0.$$

Los resultados anteriores garantizan que:

$$\begin{aligned} &\sqrt{n}\left[\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})\right] \\ &\overset{\mathcal{L}}{\to}\ N\left(0, \nabla \mathrm{I}_f(\mathbf{p})\left[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\right]^{-1}(\nabla \mathrm{I}_f(\mathbf{p}))^t\right) \end{aligned}$$

cuando $n \to \infty$, siempre que $\sigma^2(\mathbf{p}) = \nabla \mathrm{I}_f(\mathbf{p})\left[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\right]^{-1}(\nabla \mathrm{I}_f(\mathbf{p}))^t$ $= \mathrm{V}^*\Sigma\mathrm{V}^{*t} > 0$, con $\mathrm{V}^* =$ matriz fila $(\mathrm{V}_1^*,\ldots,\mathrm{V}_r^*)$ y:

$$\Sigma = \begin{pmatrix} p_1(1-p_1) & \cdots & -p_1 p_r \\ \vdots & \ddots & \vdots \\ -p_r p_1 & \cdots & p_r(1-p_r) \end{pmatrix}$$

ya que:

$$\sigma^2(\mathbf{p}) = \sum_{i=1}^{r-1}\sum_{j=1}^{r-1} p_i(\delta_{ij} - p_j)\left(\frac{\partial}{\partial p_i}\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})\right)\left(\frac{\partial}{\partial p_j}\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})\right) = \mathrm{V}^*\Sigma\mathrm{V}^{*t}.$$

iii) La matriz $\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})$ es definida positiva, por lo que la forma cuadrática asociada a $\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})$ y a $\left[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\right]^{-1}$ es definida positiva. En consecuencia, si $\sigma^2(\mathbf{p}) = \nabla \mathrm{I}_f(\mathbf{p})\left[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\right]^{-1}(\nabla \mathrm{I}_f(\mathbf{p}))^t = 0$ se debe cumplir que $\nabla \mathrm{I}_{\mathcal{X}}^F(\mathbf{p}) = 0$.

Si se considera ahora el desarrollo de Taylor de segundo orden de $\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n)$, puede asegurarse que para n suficientemente grande:

$$\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) = \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) + \frac{1}{2}(\mathbf{f}_n - \mathbf{p}) H(\mathrm{I}_f(\mathbf{p}))\, (\mathbf{f}_n - \mathbf{p})^t$$
$$+ \frac{1}{6} \sum_{i=1}^{r-1} \sum_{j=1}^{r-1} \sum_{k=1}^{r-1} \left(\frac{\partial^3 \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})}{\partial p_i \partial p_j \partial p_k} \right)_{\mathbf{p}=\mathbf{p}_n^{**}} (f_{ni} - p_i)(f_{nj} - p_j)(f_{nk} - p_k),$$

con $\mathbf{p}_n^{**} \in \mathbb{P}$ tal que $\|\mathbf{p}_n^{**} - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$. Por lo tanto:

$$2n \left[\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) - \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) \right]$$
$$= \sqrt{n}(\mathbf{f}_n - \mathbf{p}) H(\mathrm{I}_f(\mathbf{p})) \left(\sqrt{n}(\mathbf{f}_n - \mathbf{p}) \right)^t$$
$$+ \frac{1}{3} \sum_{i=1}^{r-1} \sum_{j=1}^{r-1} \sum_{k=1}^{r-1} \left(\frac{\partial^3 \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})}{\partial p_i \partial p_j \partial p_k} \right)_{\mathbf{p}=\mathbf{p}_n^{**}}$$
$$\cdot (f_{ni} - p_i) \left(\sqrt{n}(f_{nj} - p_j) \right) \left(\sqrt{n}(f_{nk} - p_k) \right).$$

Razonando como en el apartado *ii)* puede concluirse que:

$$(f_{ni} - p_i) \left(\sqrt{n}(f_{nj} - p_j) \right) \left(\sqrt{n}(f_{nk} - p_k) \right) \xrightarrow{p} 0.$$

Además:

$$\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p}) \right) H(\mathrm{I}_f(\mathbf{p})) \left(\sqrt{n}(\mathbf{f}_n - \mathbf{p}) \right)^t \xrightarrow{\mathcal{L}} Y H(\mathrm{I}_f(\mathbf{p}))\, Y^t,$$

donde Y es un vector aleatorio con distribución normal $(r-1)$-dimensional $N\left(\mathbf{0}, \left[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\right]^{-1}\right)$.

Como se supone que $H(\mathrm{I}_f(\mathbf{p}))$ no se reduce a la matriz nula de orden $(r-1) \times (r-1)$, y ya que el rango de $\left[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\right]^{-1}$ es $r-1$, el vector Y puede expresarse como $Y = ZB^*$ con Z vector aleatorio $(r-1)$-dimensional cuyas componentes son $r-1$ variables aleatorias independientes e idénticamente distribuidas, con distribución $N(0,1)$, y B^* es una matriz $(r-1) \times (r-1)$ tal que $B^* B^{*t} = \left[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\right]^{-1}$.

Además, existe una transformación $Z = UC^*$ en la que C^* es una matriz ortogonal, de forma que:

$$Y H(\mathrm{I}_f(\mathbf{p}))\, Y^t = Z B^* H(\mathrm{I}_f(\mathbf{p}))\, B^{*t} Z^t$$

$$= U C^* B^* H(\mathrm{I}_f(\mathbf{p}))\, B^{*t} C^{*t} U^t = \lambda_1^* U_1^2 + \cdots + \lambda_q^* U_q^2,$$

donde $\lambda_1^*, \ldots, \lambda_q^*$ ($q \leq r-1$) son los autovalores no nulos de la matriz $(r-1) \times (r-1)$ dada por $B^* H(\mathrm{I}_f(\mathbf{p}))\, B^{*t}$, donde $U_1, \ldots, U_q$ son variables aleatorias independientes e idénticamente distribuidas, con distribución $N(0,1)$.

En consecuencia, el estudio de la forma cuadrática $Y H(\mathrm{I}_f(\mathbf{p}))\, Y^t$ se reduce en primer lugar al de una forma cuadrática de un vector normal $(r-1)$-dimensional con componentes independientes y $N(0,1)$, y este último se reduce a su vez al de una combinación lineal de cuadrados de variables aleatorias independientes e idénticamente distribuidas, con distribución $N(0,1)$, es decir, una combinación lineal de variables aleatorias independientes e idénticamente distribuidas, con distribución chi-cuadrado χ_1^2. □

Observación 2.4.1 . Cuando se particulariza la varianza asintótica $\sigma^2(\mathbf{p})$ del Teorema 2.4.1 al índice hiperbólico (es decir, se supone que $f(x) = x^{-1} - 1$ para todo $x \in (0, +\infty) = \mathrm{Dom} f$), se obtiene en términos de las notaciones de la demostración del Teorema 2.3.6 que $\sigma^2(\mathbf{p})$ equivale a la varianza de la variable aleatoria real que toma los valores

$$\mathrm{T}_j = \frac{1}{N} \sum_{l=1}^{N} (\mathrm{V}_{jl} + \mathrm{V}_{lj})$$

con probabilidades iguales a $1/N$ para $j \in \{1, \ldots, N\}$, de forma que:

$$\sigma^2(\mathbf{p}) = \frac{1}{N} \sum_{j=1}^{N} \mathrm{T}_j^2 - \left(\frac{1}{N} \sum_{j=1}^{N} \mathrm{T}_j \right)^2$$

$$= \frac{1}{N}\sum_{j=1}^{N}\left(\frac{1}{N}\sum_{l=1}^{N}[\mathrm{V}_{jl}+\mathrm{V}_{lj}]\right)^2 - 4\left(\frac{1}{N^2}\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jl}\right)^2$$

$$= \lim_{n\to\infty}\mathrm{Var}\left(\sqrt{n}\widehat{\mathrm{I}_H}^w(\mathcal{X}[\,\cdot\,])\right).$$

El apartado *ii)* del teorema anterior puede modificarse para poder calcular la varianza asintótica del estimador $\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)$ en la práctica, y poder desarrollar (aunque sea de forma aproximada) inferencias como la estimación por intervalo y algunos contrastes de hipótesis. Más concretamente, cuando $\sigma^2(\mathbf{p})$ se reemplaza por su estimación analógica, $\sigma^2(\mathbf{f}_n)$, se obtiene la conclusión siguiente:

Teorema 2.4.2. *En las condiciones del Teorema* 2.4.1*, se cumple que:*

$$\left\{\frac{\sqrt{n}\,(\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)-\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}))}{\sqrt{\sigma^2(\mathbf{f}_n)}}\right\}_n$$

converge en ley hacia una distribución normal $N(0,1)$ *cuando* $n\to\infty$*, siempre que* $\sigma^2(\mathbf{p})>0$ *y* $\sigma^2(\mathbf{f}_n)>0$.

Demostración:

Como las componentes de $\nabla\mathrm{I}_f(\mathbf{p})$ y de $\left[\mathrm{I}^F_{\mathcal{X}}(\mathbf{p})\right]^{-1}$ son continuas en un entorno de $\mathbf{p}$, se cumple que cuando $n\to\infty$:

$$\nabla\mathrm{I}_f(\mathbf{f}_n)\xrightarrow{p}\nabla\mathrm{I}_f(\mathbf{p}),$$

y que:

$$\left[\mathrm{I}^F_{\mathcal{X}}(\mathbf{f}_n)\right]^{-1}\xrightarrow{p}\left[\mathrm{I}^F_{\mathcal{X}}(\mathbf{p})\right]^{-1}.$$

En consecuencia:

$$\sqrt{\sigma^2(\mathbf{f}_n)}\xrightarrow{p}\sqrt{\sigma^2(\mathbf{p})}$$

cuando $n\to\infty$, y al cumplirse que:

$$\sqrt{n}\,(\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)-\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}))\xrightarrow{\mathcal{L}}N(0,\,\sigma^2(\mathbf{p})),$$

se satisface que:

$$\frac{\sqrt{n}\left(\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)-\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})\right)}{\sqrt{\sigma^2(\mathbf{f}_n)}}\xrightarrow{\mathcal{L}}N(0,\,1).$$

□

El estudio desarrollado hasta el momento en esta sección, permite realizar inferencias adicionales sobre la medida de variación relativa poblacional.

Por lo que se refiere a la *estimación por intervalo*, el procedimiento que se expone a continuación nos suministrará un rango de valores posibles para la medida de variación relativa poblacional que cubrirá el verdadero valor de esta medida con (en este caso aproximadamente) una probabilidad prefijada.

El Teorema 2.4.2 nos permite establecer de forma aproximada los límites del rango de valores anterior, como sigue:

Teorema 2.4.3. *En las condiciones de los Teoremas* 2.4.1 *y* 2.4.2*, el intervalo aleatorio*

$$\left[\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) - z_\alpha \sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}},\ \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) + z_\alpha \sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}} \right]$$

proporciona para cada muestra de n observaciones independientes a partir de $\mathcal{X}$ intervalos de confianza de $\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})$ con coeficiente aproximadamente $1-\alpha$ (para $\alpha \in [0,1]$).

Por último, y de modo inmediato, se derivan los *contrastes de hipótesis* siguientes:

Teorema 2.4.4. *En las condiciones de los Teoremas* 2.4.1 *y* 2.4.2*:*

(i) Para contrastar al nivel de significación $\alpha \in [0,1]$ la hipótesis nula:

$$H_0 : \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) = i_0$$

frente a la alternativa:

$$H_1 : \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) \neq i_0,$$

si $|\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) - i_0| > z_\alpha \sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$ debe rechazarse H_0. El p-valor del contraste vendrá dado, aproximadamente, por:

$$p = 2 \left[1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}} \, |\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) - i_0| \right) \right].$$

(ii) Para contrastar al nivel de significación α *la hipótesis nula:*

$$H_0 : \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}) \geq i_0$$

frente a la alternativa:

$$H_1 : \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}) < i_0,$$

si $\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - i_0 < z_{2\alpha}\sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$ *debe rechazarse* H_0. *El p-valor del contraste vendrá dado, aproximadamente, por:*

$$p = \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\,(\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - i_0)\right).$$

(iii) Para contrastar al nivel de significación α *la hipótesis nula:*

$$H_0 : \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}) \leq i_0$$

frente a la alternativa:

$$H_1 : \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}) > i_0,$$

si $\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - i_0 > z_{2\alpha}\sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$ *debe rechazarse* H_0. *El p-valor del contraste vendrá dado, aproximadamente, por:*

$$p = 1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\,(\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - i_0)\right).$$

Algunos de los estudios realizados en las dos últimas secciones van a ilustrarse a continuación mediante un ejemplo.

Ejemplo 2.4.1. En Klir & Yuan (1995) se señala que hay un amplio número de situaciones dentro de la Ingeniería de Construcción para las que la Teoría de los Conjuntos Difusos resulta especialmente útil, como son aquellas relativas a la valoración de diferentes tipos de construcciones. Entre éstas cabe considerar, la fatiga de una estructura metálica, la calidad del pavimento de una autopista, los daños ocasionados en un edificio tras un terremoto, etc.

Klir & Yuan cita como ejemplo en el estudio de las condiciones físicas de los puentes de autopista, la variable ESTADO ACTUAL DE LOS PILARES ($\mathcal{X}$) de un puente, cuyos valores son POBRE ($\tilde{x}_1$), REGULAR ($\tilde{x}_2$) y BUENO ($\tilde{x}_3$), y se consideran descritos (Klir & Yuan 1995) por los números difusos triangulares, $\tilde{x}_1 = \text{Tri}(1,2,3)$, $\tilde{x}_2 = \text{Tri}(2,3,5)$ y $\tilde{x}_3 = \text{Tri}(3,5,5)$ (ver Figura 2.3).

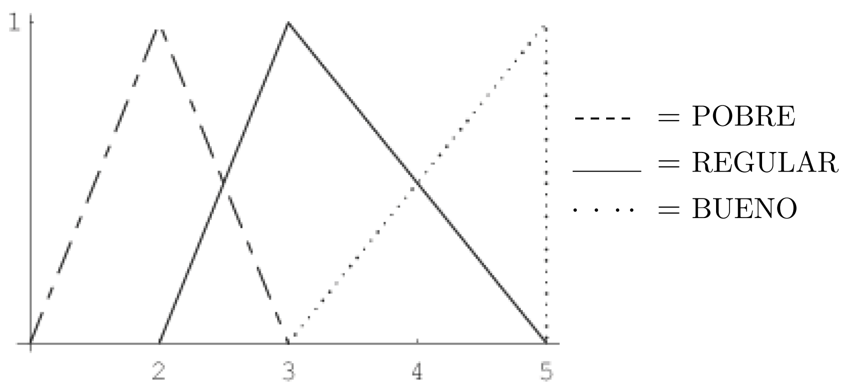


Fig. 2.3: Valores difusos del ESTADO ACTUAL DE LOS PILARES

En Estados Unidos existen unos 600.000 puentes de autopista, de los cuales aproximadamente la mitad se construyeron antes de 1940. Para estimar puntualmente, y contrastar hipótesis sobre el valor del f-índice de desigualdad de $\mathcal{X}$ en la población Ω de los 600.000 puentes, con $f = x^{-1} - 1$, se considera una muestra aleatoria simple v de $n = 400$ de esos puentes, obteniéndose los datos siguientes:

$\tilde{x}_l$	POBRE	REGULAR	BUENO
f_{nl}	15	56	329

En virtud del Teorema 2.3.2, $\text{I}_H(\mathcal{X})$ puede estimarse mediante el valor:

$$\widehat{\text{I}_H}\left(\mathcal{X}[v]\right) = \frac{400(599.999)}{600.000(399)} \text{I}_H\left(\mathcal{X}[v]\right)$$

$$-\frac{1-\dfrac{400}{600.000}}{400}\left\{\frac{1}{399}\sum_{l=1}^{3}\mathrm{I}_H\left(\tilde{x}_l\right)f_{nl}\right\}=0,099.$$

Por otro lado, si se quiere realizar el contraste de:

$$H_0 : \mathrm{I}_H(\mathcal{X}) \leq 0,05$$

frente a la alternativa:

$$H_1 : \mathrm{I}_H(\mathcal{X}) > 0,05,$$

el p-valor viene dado por:

$$p = 1 - \Phi\left(\sqrt{\frac{399}{\sigma^2(\mathbf{f}_n)}}\left(\mathrm{I}_H(\mathcal{X}[\upsilon]) - i_0\right)\right) = 0,19$$

con lo que se puede concluir que la hipótesis nula H_0 es bastante sostenible de acuerdo con la información muestral disponible.

2.5 Valoración final y problemas abiertos

Como valoración final del Capítulo 2 puede indicarse que para conservar algunas de las propiedades fundamentales sobre los índices de desigualdad del caso clásico, como son las relativas a los principios de transferencias y la compatibilidad con el criterio de Lorenz, se ha recurrido a la consideración del orden fuerte $\succeq_S$ (Ramík & Římánek 1985) para formalizar las suposiciones sobre la ordenación entre ciertos valores de la variable aleatoria difusa. Aunque se trata de una consideración bastante restrictiva en el caso difuso (ya que no establece un orden completo sino parcial en $\mathcal{F}_c(\mathrm{I\!R})$), permite llegar a afirmaciones tan fuertes como las del caso clásico en la comparación de los valores de cada índice de desigualdad.

También se pierde la descomponibilidad aditiva que tenían los índices I^α al tratar con variables aleatorias difusas, salvo para el índice de tipo Shannon.

A diferencia de los cálculos sugeridos por López García (1997) y Colubi Cervero (1997), los de los f-índices resultan mucho más simples y con el importante ahorro de tener que ofrecer un valor difuso (habitualmente su representación gráfica) como respuesta al problema. No obstante, buena parte del tratamiento informático presentado por López García y Colubi Cervero, por lo que se refiere a la descripción de los datos difusos y al cálculo de sus imágenes inducidas por las funciones f y los valores esperados de una variable aleatoria difusa, nos resulta válido en este caso como información complementaria, si bien el cálculo final de los f-índices únicamente requiere técnicas de aproximación bastante simples.

En relación con la investigación que se ha desarrollado en este capítulo, los problemas abiertos son similares a los del Capítulo 1:

- Los resultados de las Secciones 2.3 y 2.4 podrían extenderse a muestreos más complejos, como el estratificado.

- Las propiedades asintóticas y las inferencias derivadas de las mismas en la Sección 2.4, pueden desarrollarse en forma análoga para comparar la desigualdad de dos poblaciones sobre la base de sendas muestras aleatorias simples independientes. Tanto para este problema, como para el de una población considerado en esta memoria, es necesario analizar los tamaños muestrales que permiten garantizar la validez o buena aproximación de las aproximaciones asintóticas.

- Los estudios de este capítulo serían inmediatos de extender en su mayoría si se considerara un índice más general obtenido por composición con la función de ordenación λ-promedio introducida por Campos & González (1989), con $\lambda \in [0, 1]$, y cuya particularización al caso $\lambda = 0,5$ conduce a la función F de Yager (1981). El único (pero determinante) inconveniente de adoptar ese procedimiento más general es que para valores $\lambda \in [0,\ 0,5)$ podría perderse la propiedad de no negatividad, y las minimalidades no siempre pueden garantizarse (de hecho, el Teorema 2.2.5 sólo podría extenderse para el caso $\lambda \in [0,5, 1]$, y el Teorema 2.2.6 no puede enunciarse para valores de $\lambda \neq 0,5$). En consecuencia, la comparación de poblaciones/muestras y/o variables a través de los índices que pueden tomar valores negativos, llevaría a menudo a conclusiones claramente erróneas.

Capítulo 3

Medidas generalizadas de la variación absoluta y relativa de un conjunto aleatorio

En la cuantificación del orden de magnitud de la variación absoluta y relativa asociada a un atributo cuantitativo, se supone habitualmente que éste se formaliza en términos de una variable aleatoria con valores reales. No obstante, algunas variables (como, por ejemplo, rangos numéricos de ciertas características) no se ajustan al modelo de las variables aleatorias reales, sino más bien a una clase especial de variables aleatorias difusas. Se trata de aquellas variables que toman esencialmente valores de intervalo, y algunos de sus valores pueden solaparse.

Estos elementos aleatorios pueden identificarse formalmente con casos particulares de los llamados conjuntos aleatorios (más concretamente, de los conjuntos aleatorios compactos y convexos con valores en $\mathcal{K}_c(\mathbb{R})$).

Al ser estos conjuntos aleatorios una particularización de las variables aleatorias difusas que toman valores en $\mathcal{F}_c(\mathbb{R})$, los conceptos introducidos y las conclusiones obtenidas en los Capítulos 1 y 2 pueden aplicarse de forma inmediata a esos conjuntos aleatorios.

En el presente capítulo se va a llevar a cabo dicha aplicación, matizando algunos aspectos característicos de esta situación de especial interés, y prestando atención a las expresiones particularizadas de los estimadores, los parámetros de las distribuciones asintóticas, etc.

En primer lugar, se introduce una medida generalizada de la variación absoluta de un conjunto aleatorio compacto y convexo respecto a un intervalo real compacto. Se determina una estimación insesgada de la medida anterior y su correspondiente error asociado en los muestreos aleatorios simple y con reposición de poblaciones finitas, la distribución asintótica de la medida de variación absoluta muestral, y procedimientos inferenciales derivados de esa distribución. Se presentan las aplicaciones de la medida de variación absoluta en la cuantificación del error de muestreo asociado a la estimación de la integral de Aumann (poblacional) del conjunto aleatorio compacto y convexo en los muestreos aleatorio simple y con reposición de poblaciones finitas, y en el análisis de regresión entre dos conjuntos aleatorios compactos y convexos en poblaciones cualesquiera.

En segundo lugar, se introduce una medida generalizada de la desigualdad o variación relativa de un conjunto aleatorio compacto y convexo. Se determinan una estimación insesgada de una de las medidas de desigualdad (el índice hiperbólico) y la distribución asintótica de todas las que satisfacen ciertas condiciones de regularidad (de hecho, todas las medidas de desigualdad de mayor utilidad práctica) y algunos métodos inferenciales relacionados.

3.1 La S-dispersión cuadrática media asociada a un conjunto aleatorio. Resultados inferenciales en poblaciones finitas

En esta sección se presenta la particularización de la medida generalizada de la variación absoluta del Capítulo 1 al caso en que se considere un conjunto aleatorio compacto y convexo y un elemento de $\mathcal{K}_c(\mathbb{R})$.

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y sea $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ un conjunto aleatorio compacto y convexo integrablemente acotado asociado a $(\Omega, \mathcal{A}, P)$.

Definición 3.1.1. *La S-dispersión cuadrática media asociada a X con respecto al intervalo $K \in \mathcal{K}_c(\mathbb{R})$, viene dada por el valor real (si existe):*

$$DCM_S(X, K) = E\left([d_S(X(\cdot), K)]^2\right) = \int_\Omega [d_S(X(\omega), K)]^2 dP(\omega)$$

$$= \int_\Omega \int_{[0,1]} \left[f_{X(\omega)}(\lambda) - f_K(\lambda)\right]^2 dS(\lambda)\, dP(\omega),$$

con $f_K(\lambda) = \lambda \sup K + (1-\lambda) \inf K$ para todo $\lambda \in [0,1]$.

La S-dispersión cuadrática media central asociada a un conjunto aleatorio compacto y convexo se definirá como sigue:

Definición 3.1.2. *La S-dispersión cuadrática media central asociada a X se define como el valor real (si existe):*

$$\Delta_S^2(X) = DCM_S(X, E(X)) = \int_\Omega [d_S(X(\omega), E(X))]^2 dP(\omega).$$

Cuando se precise especificar la distribución P del espacio probabilístico se denotará $\Delta_S^2(X)$ por $\Delta_S^2(X \mid P)$.

Las condiciones de existencia y las propiedades de las Secciones 1.1 y 1.2 se particularizarían de forma inmediata para las medidas $DCM_S(X, K)$ y $\Delta_S^2(X)$.

El siguiente ejemplo ilustra el cálculo de la medida de variación absoluta de un conjunto aleatorio compacto y convexo en un ejemplo real, cuyos datos han sido proporcionados por el Servicio de Nefrología del Hospital Valle del Nalón de Langreo (Asturias).

Ejemplo 3.1.1. Los datos que aparecen en la Tabla 3.1 corresponden al "rango de la frecuencia cardíaca a lo largo de un día", X, observada en una

58-90	64-107	52-78	56-133	54-78	75-124
47-68	54-84	55-84	37-75	53-103	58-99
32-114	47-95	61-101	61-94	47-86	59-78
61-110	56-90	65-92	44-110	70-132	55-89
62-89	44-108	38-66	46-83	63-115	55-80
63-119	63-109	48-73	52-98	47-83	70-105
51-95	62-95	59-98	56-84	56-103	40-80
49-78	48-107	59-87	54-92	71-121	56-97
43-67	26-109	49-82	53-120	68-91	37-86
55-102	61-108	48-77	49-88	62-100	

Tabla 3.1: Datos sobre rangos de la frecuencia cardíaca

población Ω de 59 pacientes hospitalizados. Los valores de X se obtienen a partir de los distintos registros (habitualmente entre 60 y 70) de la frecuencia cardíaca realizados sobre cada paciente en distintos instantes de un mismo día.

Si se desea medir la variación del rango de la frecuencia cardíaca en un día en la población Ω y se considera la S-DCM central asociada a X para la medida S discretizada de forma que $L = 3$, $\lambda_1 = 0$, $\lambda_2 = 0,5$, $\lambda_3 = 1$, $k_1 = k_2 = k_3 = 1/3$, y $\overline{g}(\lambda) = 0$ si $\lambda \in (0,1) \setminus \{0,5\}$, se obtiene que

$$\Delta_S^2(X) = 146,833.$$

A continuación van a estudiarse las particularizaciones para conjuntos aleatorios de los resultados obtenidos en las Secciones 1.3 y 1.4 para el problema de la estimación de la medida generalizada de variación absoluta en el muestreo aleatorio de poblaciones finitas y para la determinación bajo ciertas condiciones de la distribución asintótica de la medida muestral en poblaciones finitas y de los procedimientos aproximados derivados de ella.

Supongamos que se considera una población finita Ω de N unidades, $\omega_1, \ldots, \omega_N$, y un conjunto aleatorio compacto y convexo $X : \Omega \to \mathcal{K}_c(\mathbb{R})$

asociado a un espacio medible definido sobre Ω y que se supone dotado con la distribución uniforme.

Si se elige una muestra de tamaño n al azar y sin reposición a partir de Ω, υ representa una muestra aleatoria simple genérica de tamaño n, y $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ son las unidades en υ, entonces la *S-dispersión cuadrática media central muestral* de X en υ viene dada por:

$$\Delta_S^2\left(X[\upsilon]\right) = \frac{1}{n}\sum_{i=1}^{n}\left[d_S\left(X(\omega_{\upsilon i}), \overline{X}_n[\upsilon]\right)\right]^2,$$

donde $\overline{X}_n[\upsilon] = \dfrac{1}{n}\left[X(\omega_{\upsilon 1}) + \cdots + X(\omega_{\upsilon n})\right]$ es el valor esperado muestral de X en υ.

$\Delta_S^2\left(X[\,\cdot\,]\right)$ es una variable aleatoria real asociada al espacio probabilístico $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ (ver Sección 1.3, pág. 45), y por lo tanto define un estimador de la *S-dispersión cuadrática media central poblacional* que viene dada ahora por:

$$\Delta_S^2(X \mid P) = \frac{1}{N}\sum_{j=1}^{N}\left[d_S\left(X(\omega_j), \overline{X}\right)\right]^2,$$

con $\overline{X} = \dfrac{1}{N}\left[X(\omega_1) + \cdots + X(\omega_N)\right].$

Las conclusiones que pueden obtenerse en cuanto a la estimación insesgada (en el sentido de la integral de Aumann sobre Υ_n) son las siguientes:

Teorema 3.1.1. *En el muestreo aleatorio simple de tamaño n a partir de Ω, se cumple que el estimador $\widehat{\Delta_S^2}(X[\,\cdot\,])$ que a una muestra υ asocia el valor:*

$$\widehat{\Delta_S^2}(X[\upsilon]) = \frac{(N-1)n}{N(n-1)}\Delta_S^2\left(X[\upsilon]\right),$$

es un estimador insesgado de $\Delta_S^2(X \mid P)$.

Teorema 3.1.2. *En el muestreo aleatorio simple de tamaño n a partir de Ω, si $f = n/N$ se cumple que:*

$$\operatorname{Var}\left(\widehat{\Delta_S^2}(X[\,\cdot\,])\right) = \frac{(1-f)}{n(n-1)N^2(N-2)(N-3)}$$

$$\cdot\Bigg\{2(N-n-1)(N-1)\sum_{j=1}^{N}\sum_{l=1}^{N}\Bigg(\int_{[0,1]}\left[f_{X(\omega_j)}(\lambda)-f_{\overline{X}}(\lambda)\right]$$

$$\cdot\left[f_{X(\omega_l)}(\lambda)-f_{\overline{X}}(\lambda)\right]d\,S(\lambda)\Bigg)^2+[(6-4N)n-8(N-1)]\left(\Delta_S^2(X\,|\,P)\right)^2$$

$$+[n(N-2)(N^2-3)-(N-1)(N^2+N-8)]N\ \mathrm{Var}\left(\left[d_S(X,\overline{X})\right]^2\,\Big|\,P\right)\Bigg\}.$$

Si se considera una selección aleatoria con reposición de n unidades de la población Ω, υ representa una muestra aleatoria con reposición genérica de tamaño n, y $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ son las unidades en υ, $\Delta_S^2(X[\cdot])$ es una variable aleatoria real asociada al espacio probabilístico $(\Upsilon_n^w, \mathcal{P}(\Upsilon_n^w), p^w)$ (ver Sección 1.3, pág. 51), y por lo tanto define un estimador de la S-DCM central poblacional. Las conclusiones son ahora las siguientes:

Teorema 3.1.3. *En el muestreo aleatorio con reposición de tamaño n a partir de Ω, se cumple que el estimador $\widehat{\Delta_S^2}^w(X[\cdot])$ que a una muestra υ asocia el valor:*

$$\widehat{\Delta_S^2}^w(X[\upsilon])=\frac{n}{(n-1)}\Delta_S^2(X[\upsilon]),$$

es un estimador insesgado de $\Delta_S^2(X\,|\,P)$.

Teorema 3.1.4. *En el muestreo aleatorio con reposición de tamaño n a partir de Ω, se cumple que:*

$$\mathrm{Var}\left(\widehat{\Delta_S^2}^w(X[\cdot])\right)=\frac{1}{n(n-1)N^3}$$

$$\cdot\Bigg\{2(n-1)(N-2)\sum_{j=1}^{N}\sum_{l=1}^{N}\Bigg(\int_{[0,1]}\left[f_{X(\omega_j)}(\lambda)-f_{\overline{X}}(\lambda)\right]$$

$$\cdot\left[f_{X(\omega_l)}(\lambda)-f_{\overline{X}}(\lambda)\right]d\,S(\lambda)\Bigg)^2+\ 2[n^2-Nn-2]N^2\ \left(\Delta_S^2(X\,|\,P)\right)^2$$

$$+[N(n-1)^2-2(n^2-5n+5)]N^2\ \mathrm{Var}\left(\left[d_S(X,\overline{X})\right]^2\,\Big|\,P\right)\Bigg\}.$$

Para desarrollar otros procedimientos inferenciales sobre la medida de variación absoluta poblacional de un conjunto aleatorio compacto y convexo, debe recurrirse o bien a muestras extraídas al azar y con reemplazamiento de una población cualquiera, o a muestras suficientemente grandes extraídas al azar y sin reemplazamiento de una población de tamaño substancialmente más grande que el tamaño muestral, y aplicar los resultados que van a establecerse ahora sobre la distribución asintótica de la S-DCM central muestral.

Sea $\Omega = \{\omega_1, \ldots, \omega_N\}$ una población finita y sea $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ un conjunto aleatorio compacto y convexo. Sobre Ω se puede definir el espacio probabilístico $(\Omega, \mathcal{P}(\Omega), P)$, donde P representa la medida de probabilidad correspondiente a la distribución uniforme sobre Ω.

Si sobre la población Ω la variable X toma r valores distintos, $x_1^*, \ldots, x_r^*$, y para $l \in \{1, \ldots, r\}$ se denota por p_l la probabilidad $P(\{\omega \in \Omega \mid X(\omega) = x_l^*\})$, y S es una medida en $[0,1]$ que satisface las condiciones en la Definición 0.1.5, la *S-DCM central poblacional* asociada a X en Ω puede expresarse como:

$$\Delta_S^2(X \mid \mathbf{p}) = \sum_{l=1}^{r} p_l \left[d_S\left(x_l^*, E(X \mid \mathbf{p})\right)\right]^2,$$

con $E(X \mid \mathbf{p}) = \sum_{l=1}^{r} p_l\, x_l^*$ y $\mathbf{p} = (p_1, \ldots, p_{r-1})$ $(p_r = 1 - \sum_{l=1}^{r-1} p_l)$.

Si se selecciona al azar una muestra de tamaño n de la población y f_{nl} representa la frecuencia relativa del valor x_l^* de X en esa muestra, la *S-DCM central muestral* corresponde al valor:

$$\Delta_S^2(X \mid \mathbf{f}_n) = \sum_{l=1}^{r} f_{nl} \left[d_S\left(x_l^*, \overline{X}_n\right)\right]^2,$$

con $\overline{X}_n = \sum_{l=1}^{r} f_{nl} x_l^*$, y $\mathbf{f}_n = (f_{n1}, \ldots, f_{n(r-1)})$ $(f_{nr} = 1 - \sum_{l=1}^{r-1} f_{nl})$.

La distribución asintótica de la S-DCM central muestral, y sus propiedades asintóticas en la estimación de $\Delta_S^2(X)$, se recogen en el resultado siguiente:

Teorema 3.1.5. *Para cada* $n \in \mathbb{N}$*, se consideran* n *conjuntos aleatorios compactos y convexos independientes e idénticamente distribuidos que* X *(es decir, una muestra aleatoria simple de tamaño* n *a partir de* X*) definido sobre la población finita* $\Omega = \{\omega_1, \ldots, \omega_N\}$ *de forma que* $P(\{\omega \in \Omega \mid X(\omega) = x_l^*\}) = p_l$ *con* $p_l \in (0,1)$*,* $l = 1, \ldots, r$*,* $\sum_{l=1}^{r} p_l = 1$*. Sea* S *una medida normalizada sobre* $[0,1]$*, que satisface las condiciones en la Definición* 0.1.5*. Se cumple entonces que:*

i) *Si* $\mathbf{f}_n = (f_{n1}, \ldots, f_{n(r-1)}) \in [0,1]^{r-1}$*, con* f_{nl} *= frecuencia relativa de* x_l^* *en la correspondiente realización de la muestra aleatoria simple de tamaño* n *(*$l = 1, \ldots, r-1$*,* $f_{nr} = 1 - \sum_{l=1}^{r-1} f_{nl}$*), y* $\Delta_S^2(X \mid \mathbf{f}_n)$ *es la S-DCM central muestral asociada, entonces* $\{\Delta_S^2(X \mid \mathbf{f}_n)\}_n$ *es una sucesión de estimadores de* $\Delta_S^2(X \mid \mathbf{p})$ *fuertemente consistente, es decir, cuando* $n \to \infty$ *se tiene que:*

$$\Delta_S^2(X \mid \mathbf{f}_n) \stackrel{c.s.}{\to} \Delta_S^2(X \mid \mathbf{p})$$

cualquiera que sea $\mathbf{p} = (p_1, \ldots, p_{r-1})$ *(con* $p_1, \ldots, p_{r-1} \in (0,1)$ *y* $\sum_{l=1}^{r-1} p_l < 1$*).*

ii) $\{\sqrt{n}\,(\Delta_S^2(X \mid \mathbf{f}_n) - \Delta_S^2(X \mid \mathbf{p}))\}_n$ *es una sucesión de variables aleatorias que converge en ley hacia una normal unidimensional* $N(0, \sigma^2(\mathbf{p}))$*, con:*

$$\sigma^2(\mathbf{p}) = \mathrm{Var}\left(\left[d_S\left(X, E(X \mid \mathbf{p})\right)\right]^2\right),$$

siempre que $\sigma^2(\mathbf{p}) > 0$*.*

iii) *Si* $\sigma^2(\mathbf{p}) = 0$*, y existe un par* (i,j) *con* $i,j \in \{1, \ldots, r-1\}$ *tal que se cumple que:*

$$h_{ij} = \frac{\partial^2 \Delta_S^2(X \mid \mathbf{p})}{\partial p_i \partial p_j}$$

$$= \frac{\partial}{\partial p_i}\left(\left[d_S\left(x_j^*, E(X \mid \mathbf{p})\right)\right]^2 - \left[d_S\left(x_r^*, E(X \mid \mathbf{p})\right)\right]^2\right) > 0,$$

entonces $\{2n\,(\Delta^2_S(X\,|\,\mathbf{f}_n) - \Delta^2_S(X\,|\,\mathbf{p}))\}_n$ *es una sucesión de variables aleatorias que converge en ley hacia una combinación lineal de, a lo sumo,* $r-1$ *variables chi-cuadrado* χ^2_1 *independientes.*

Cuando $\sigma^2(\mathbf{p})$ se reemplaza por su estimación analógica, $\sigma^2(\mathbf{f}_n)$, se obtiene la conclusión siguiente:

Teorema 3.1.6. *En las condiciones del Teorema* 3.1.5*, se cumple que:*

$$\left\{\frac{\sqrt{n}\,(\Delta^2_S(X\,|\,\mathbf{f}_n) - \Delta^2_S(X\,|\,\mathbf{p}))}{\sqrt{\sigma^2(\mathbf{f}_n)}}\right\}_n$$

converge en ley hacia una distribución normal $N(0,1)$ *cuando* $n \to \infty$*, siempre que* $\sigma^2(\mathbf{p}) > 0$ *y* $\sigma^2(\mathbf{f}_n) > 0$.

El resultado siguiente nos permite establecer de forma aproximada *intervalos de confianza* para $\Delta^2_S(X)$:

Teorema 3.1.7. *En las condiciones de los Teoremas* 3.1.5 *y* 3.1.6*, el intervalo aleatorio:*

$$\left[\Delta^2_S(X\,|\,\mathbf{f}_n) - z_\alpha\sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}},\ \Delta^2_S(X\,|\,\mathbf{f}_n) + z_\alpha\sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}}\right]$$

proporciona para cada muestra de observaciones independientes a partir de X *intervalos de confianza de* $\Delta^2_S(X\,|\,\mathbf{p})$ *con coeficiente aproximadamente igual a* $1-\alpha$ $(\alpha \in [0,1])$.

Así mismo, se derivan los *contrastes de hipótesis* siguientes:

Teorema 3.1.8. *En las condiciones de los Teoremas* 3.1.5 *y* 3.1.6*:*

(i) Para contrastar al nivel de significación $\alpha \in [0,1]$ *la hipótesis nula:*

$$H_0 : \Delta^2_S(X\,|\,\mathbf{p}) = \delta_0$$

frente a la alternativa:

$$H_1 : \Delta_S^2(X \,|\, \mathbf{p}) \neq \delta_0,$$

si $|\Delta_S^2(X \,|\, \mathbf{f}_n) - \delta_0| > z_\alpha \sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$ *debe rechazarse* H_0*. El p-valor del contraste vendrá dado, aproximadamente, por:*

$$p = 2\left[1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\left|\Delta_S^2(X \,|\, \mathbf{f}_n) - \delta_0\right|\right)\right].$$

(ii) Para contrastar al nivel de significación α *la hipótesis nula:*

$$H_0 : \Delta_S^2(X \,|\, \mathbf{p}) \geq \delta_0$$

frente a la alternativa:

$$H_1 : \Delta_S^2(X \,|\, \mathbf{p}) < \delta_0,$$

si $\Delta_S^2(X \,|\, \mathbf{f}_n) - \delta_0 < z_{2\alpha} \sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$ *debe rechazarse* H_0*. El p-valor del contraste vendrá dado, aproximadamente, por:*

$$p = \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\left(\Delta_S^2(X \,|\, \mathbf{f}_n) - \delta_0\right)\right).$$

(iii) Para contrastar al nivel de significación α *la hipótesis nula:*

$$H_0 : \Delta_S^2(X \,|\, \mathbf{p}) \leq \delta_0$$

frente a la alternativa:

$$H_1 : \Delta_S^2(X \,|\, \mathbf{p}) > \delta_0,$$

si $\Delta_S^2(X \,|\, \mathbf{f}_n) - \delta_0 > z_{2\alpha} \sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$ *debe rechazarse* H_0*. El p-valor del contraste vendrá dado, aproximadamente, por:*

$$p = 1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\left(\Delta_S^2(X \,|\, \mathbf{f}_n) - \delta_0\right)\right).$$

3.2 Algunas aplicaciones de la medida de variación absoluta de un conjunto aleatorio compacto y convexo

En esta sección se examinan las particularizaciones de las aplicaciones estudiadas en las Secciones 1.5 y 1.6.

En este sentido, se considera en primer lugar el problema de estimar el valor esperado de un conjunto aleatorio compacto y convexo en los muestreos aleatorios simple o con reposición de poblaciones finitas, y posteriormente se analizan los problemas de regresión lineal y funcional general entre dos conjuntos aleatorios compactos y convexos.

Por lo que se refiere al problema de la *estimación por intervalo del valor esperado poblacional* de un conjunto aleatorio compacto y convexo, si se considera una población finita de N unidades, $\Omega = \{\omega_1, \ldots, \omega_N\}$, para las que el conjunto aleatorio compacto y convexo $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ asociado al espacio de probabilidad $(\Omega, \mathcal{P}(\Omega), P)$ toma los valores $X(\omega_1), \ldots, X(\omega_N)$, y se selecciona una muestra genérica υ de tamaño n al azar y sin reposición a partir de la población Ω, y $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ denotan las unidades en la muestra υ, entonces el *valor esperado muestral* de X en υ, $\overline{X}_n[\upsilon] = [X(\omega_{\upsilon 1}) + \cdots + X(\omega_{\upsilon n})]/n$, define un conjunto aleatorio compacto y convexo asociado con el espacio de probabilidad $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ (ver Sección 1.3).

El valor esperado de $\overline{X}_n$ sobre el espacio $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ viene dado por:

Teorema 3.2.1. *En el muestreo aleatorio simple de tamaño n de la población Ω con N unidades, el estimador con valores en $\mathcal{K}_c(\mathbb{R})$, $\overline{X}_n$, es insesgado para estimar el valor esperado poblacional $\overline{X} = E(X \mid P)$, es decir, se cumple que $E(\overline{X}_n \mid p) = \overline{X}$.*

La precisión de ese estimador con valores en $\mathcal{K}_c(\mathbb{R})$ se discute en los resultados siguientes:

Teorema 3.2.2. *En el muestreo aleatorio simple de tamaño n de una población Ω con N unidades, la S-dispersión cuadrática media de $\overline{X}_n$ respecto a $\overline{X}$ viene dada por:*

$$DCM_S\left(\overline{X}_n, \overline{X}\right) = \Delta_S^2\left(\overline{X}_n \,|\, p\right) = \frac{(1-f)N}{n(N-1)}\Delta_S^2\left(X \,|\, P\right).$$

Teorema 3.2.3. *En el muestreo aleatorio simple de tamaño n de la población Ω con N unidades, el estimador que a la muestra υ le asocia el valor:*

$$\widehat{\Delta_S^2}\left(\overline{X}_n[\upsilon]\right) = \frac{(N-1)n}{N(n-1)}\Delta_S^2\left(\overline{X}_n[\upsilon]\right)$$

$$= \frac{(N-1)n}{N(n-1)}\int_{[0,1]} \mathrm{Var}\left(f_{\overline{X}_n}(\lambda)\right) d\,S(\lambda),$$

es insesgado para estimar $\Delta_S^2\left(\overline{X}_n \,|\, p\right)$.

Con respecto al tamaño muestral adecuado n para estimar $\overline{X}$ mediante $\overline{X}_n$, este tamaño puede aproximarse de acuerdo con la Desigualdad de Tchebychev como sigue:

Teorema 3.2.4. *En el muestreo aleatorio simple de la población Ω con N unidades muestrales, el tamaño muestral*

$$n = \left.\frac{N\Delta_S^2(X \,|\, P)}{(N-1)d^2\alpha + \Delta_S^2(X \,|\, P)}\right]$$

satisface que $\mathrm{P}\left(d_S(\overline{X}_n, \overline{X}) > d\right) \leq \alpha$.

Si se selecciona una muestra de tamaño n al azar y con reposición a partir de toda la población Ω, el *valor esperado muestral* de X, $\overline{X}_n$, define ahora un conjunto aleatorio compacto y convexo asociado con el espacio de probabilidad $(\Upsilon_n^w, \mathcal{P}(\Upsilon_n^w), p^w)$ (ver Sección 1.3), y cumple que:

Teorema 3.2.5. *En el muestreo aleatorio con reposición de tamaño n de la población Ω con N unidades muestrales, $\omega_1, \ldots, \omega_N$, el estimador con valores en $\mathcal{K}_c(\mathbb{R})$, $\overline{X}_n$, es insesgado para estimar el valor esperado poblacional $\overline{X} = E(X \,|\, P)$, es decir, $E(\overline{X}_n \,|\, p^w) = \overline{X}$.*

La discusión sobre la precisión de $\overline{X}_n$ en la estimación de $\overline{X}$, permite concluir que:

Teorema 3.2.6. *En el muestreo aleatorio con reposición de tamaño n de la población Ω con N unidades muestrales, la S-dispersión cuadrática media de $\overline{X}_n$ respecto a $\overline{X}$ viene dada por:*

$$DCM_S\left(\overline{X}_n, \overline{X}\right) = \Delta_S^2\left(\overline{X}_n \,|\, p^w\right) = \frac{\Delta_S^2\,(X \,|\, P)}{n}.$$

Teorema 3.2.7. *En el muestreo aleatorio con reposición de tamaño n de la población Ω con N unidades muestrales, el estimador que a la muestra υ le asocia el valor:*

$$\widehat{\Delta_S^2}^{w}\left(\overline{X}_n[\upsilon]\right) = \frac{n}{n-1}\Delta_S^2\left(\overline{X}_n[\upsilon]\right),$$

es insesgado para estimar $\Delta_S^2\left(\overline{X}_n \,|\, p^w\right)$.

La *elección del tamaño de muestra más apropiado* se aproxima como sigue:

Teorema 3.2.8. *En el muestreo aleatorio con reposición de tamaño n de la población Ω con N unidades muestrales, el tamaño muestral*

$$n = \left.\frac{\Delta_S^2(X \,|\, P)}{d^2\alpha}\right]$$

satisface que $\mathrm{P}\Big(d_S(\overline{X}_n, \overline{X}) > d\Big) \leq \alpha$.

Sean X e Y dos conjuntos aleatorios compactos y convexos integrablemente acotados asociados al espacio de probabilidad $(\Omega, \mathcal{A}, P)$. El objetivo del *Análisis de Regresión Lineal* entre X e Y es la búsqueda de los valores reales a y b para los que la relación $Y = aX + b$ resulta "menos errónea"(ver Figura 3.1 como ejemplo gráfico ilustrativo del efecto de la relación lineal $Y = 2X - 2$ en este caso para cierto $\omega \in \Omega$). El error va a suponerse medido en términos de la distancia d_S y se admitirá que S es una medida que corresponde a una distribución simétrica respecto al valor $\lambda = 0,5$, de modo que el propósito será minimizar la función:

$$\phi(a,b) = E\left([d_S\,(Y, a\,X + b)]^2\right) = E\Big(\int_{[0,1]} [f_Y\,(\lambda) - f_{a\,X+b}\,(\lambda)]^2\; d\,S(\lambda)\Big)$$

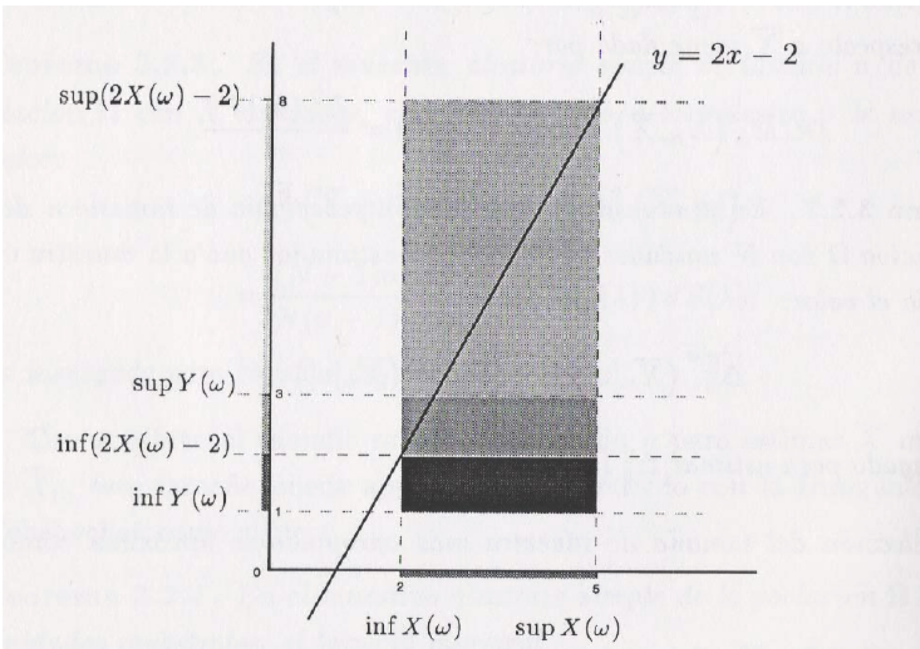


Fig. 3.1: Datos bidimensionales para los valores de Y observado y estimado (por la relación lineal $2X - 2$)

Si g es simétrica respecto a $\lambda = 0,5$ se tiene que si:

$$W_S\,(A,B) = \int_{[0,1]} f_A\,(\lambda)\,f_B\,(\lambda)\;d\,S(\lambda),$$

y:

$$W'_S\,(A,B) = \int_{[0,1]} f_A\,(1-\lambda)\,f_B\,(\lambda)\;d\,S(\lambda),$$

entonces:

$$E\,(W_S\,(X,Y)) - E\,(W'_S\,(X,Y))$$
$$= K(S)E\,((\sup X - \inf X)(\sup Y - \inf Y))\,,$$

con $K(S) = \dfrac{1}{2}\displaystyle\int_{[0,1]} (1-2\lambda)^2\,d\,S(\lambda) > 0.$

Por lo tanto, como $(\sup X - \inf X)(\sup Y - \inf Y) \geq 0$ para todo $\omega \in \Omega$, se tiene que:

$$E\left(W_S\left(X, Y\right)\right) \geq E\left(W_S'\left(X, Y\right)\right),$$

con igualdad si, y sólo si, $\sup X(\,\cdot\,) = \inf X(\,\cdot\,)$ y $\sup Y(\,\cdot\,) = \inf Y(\,\cdot\,)$ c.s. $[P]$. En consecuencia, $E\left(W_S\left(X, Y\right)\right) = E\left(W_S'\left(X, Y\right)\right)$ si, y sólo si, X e $\mathcal{Y}$ toman valores en $\mathbb{R}$ c.s. $[P]$.

La última afirmación indica que, salvo para conjuntos aleatorios compactos y convexos casi seguro reales, se cumple que el valor óptimo del parámetro a no puede anularse, y la solución general del problema es la siguiente:

Teorema 3.2.9. *Si X e Y son dos conjuntos aleatorios compactos y convexos integrablemente acotados que no son casi seguro variables aleatorias reales, y S es una medida cuya densidad asociada g respecto a la medida de Lebesgue es simétrica en $\lambda = 0,5$, entonces la función $\phi(a, b) = E\left(\left[d_S\left(Y, aX + b\right)\right]^2\right)$ alcanza su valor mínimo en el punto (a^*, b^*) con $a^* \neq 0$, tal que:*

$$a^* = \frac{E\left(W_S\left(X,\, Y\right)\right) - E\left(V_S\left(X\right)\right) E\left(V_S\left(Y\right)\right)}{E\left(\left[d_S\left(X, 0\right)\right]^2\right) - \left[E\left(V_S\left(X\right)\right)\right]^2},$$

$$b^* = E\left(V_S\left(Y\right)\right) - a^* E\left(V_S\left(X\right)\right),$$

si o bien se cumple que $E\left(W_S'\left(X,\, Y\right)\right) \geq E\left(V_S\left(X\right)\right)\, E\left(V_S\left(Y\right)\right)$, o bien $E\left(W_S\left(X,\, Y\right)\right) > E\left(V_S\left(X\right)\right)\, E\left(V_S\left(Y\right)\right) > E\left(W_S'\left(X,\, Y\right)\right)$, mientras que:

$$a^* = \frac{E\left(W_S'\left(X, Y\right)\right) - E\left(V_S\left(X\right)\right) E\left(V_S\left(Y\right)\right)}{E\left(\left[d_S\left(X, 0\right)\right]^2\right) - \left[E\left(V_S\left(X\right)\right)\right]^2}$$

$$b^* = E\left(V_S\left(Y\right)\right) - a^* E\left(V_S\left(X\right)\right),$$

si o bien se cumple que $E\left(V_S\left(X\right)\right)\, E\left(V_S\left(Y\right)\right) \geq E\left(W_S\left(X,\, Y\right)\right)$, o bien $E\left(W_S\left(X,\, Y\right)\right) > E\left(V_S\left(X\right)\right)\, E\left(V_S\left(Y\right)\right) > E\left(W_S'\left(X,\, Y\right)\right)$, con:

$$V_S\left(K\right) = \int_{[0,1]} f_K\left(\lambda\right)\, d\,S(\lambda).$$

Por otro lado, dados dos conjuntos aleatorios compactos y convexos integrablemente acotados X e Y asociados al espacio de probabilidad $(\Omega, \mathcal{A}, P)$, el objetivo del *Análisis de Regresión Funcional general* es la búsqueda de la función $h : \mathcal{K}_c(\mathbb{R}) \to \mathcal{K}_c(\mathbb{R})$ que sea $(\mathcal{B}_{d_H}, \mathcal{B}_{d_H})$-medible, y que resulte "menos errónea", admitiéndose que el error se mide a través del valor:

$$F(h) = E\left([d_S(Y, h(X))]^2 \,\Big|\, P \right)$$

$$= E\left(\int_{[0,1]} \left[f_Y(\lambda) - f_{h(X)}(\lambda) \right]^2 d\,S(\lambda) \,\Bigg|\, P \right).$$

La solución de este problema la ofrece el resultado siguiente:

Teorema 3.2.10. *Si X e Y son dos conjuntos aleatorios compactos y convexos integrablemente acotados y S es una medida que satisface las condiciones de la Definición* 0.1.5*, entonces la función F definida sobre el conjunto $\mathcal{H} = \{h : \mathcal{K}_c(\mathbb{R}) \to \mathcal{K}_c(\mathbb{R}) \mid h(X)$ conjunto aleatorio compacto y convexo$\}$ alcanza su valor mínimo para la función $h^* : \mathcal{K}_c(\mathbb{R}) \to \mathcal{K}_c(\mathbb{R})$ tal que:*

$$h^*(x) = E(Y \mid X = x),$$

para cada $x \in X(\Omega)$.

El ejemplo siguiente sirve de ilustración de la aplicación real del Teorema 3.2.9. De nuevo, los datos han sido proporcionados por el Servicio de Nefrología del Hospital Valle del Nalón de Langreo.

Ejemplo 3.2.1. Los datos ligados de la Tabla 3.2 son los valores que para cada uno de los 59 pacientes de la población Ω del Ejemplo 3.1.1 tienen las variables X = rango de la presión arterial sistólica a lo largo de un día e Y = rango de la presión arterial diastólica a lo largo de ese día.

La relación lineal óptima de Y respecto a X, según el criterio del Teorema 3.2.9, y para la medida $S = m$ será:

$$Y = 0,582\, X - 2,096.$$

Si esta relación se usara con fines de predicción, a un paciente con rango de presión sistólica 115-161 le correspondería una predicción del rango de presión diastólica igual a 65-92.

X	Y		X	Y		X	Y
118-173	63-102		119-212	47-93		98-160	47-108
104-161	71-108		122-178	73-105		97-154	60-107
131-186	58-113		127-189	74-125		87-150	47-86
105-157	62-118		113-213	52-112		141-256	77-158
120-179	59-94		141-205	69-133		108-147	62-107
101-194	48-116		99-169	53-109		115-196	65-117
109-174	60-119		126-197	60-98		99-172	42-86
128-210	76-125		99-201	55-121		113-176	57-95
94-145	47-104		88-221	37-94		114-186	46-103
148-201	88-130		113-183	55-85		145-210	100-136
111-192	52-96		94-176	56-121		120-180	59-90
116-201	74-133		102-156	50-94		100-161	54-104
102-167	39-84		103-159	52-95		159-214	99-127
104-161	55-98		102-185	63-118		138-221	70-118
106-167	45-95		111-199	57-113		87-152	50-95
112-162	62-116		130-180	64-121		120-188	53-105
136-201	67-122		103-161	55-97		95-166	54-100
90-177	52-104		125-192	59-101		92-173	45-107
116-168	58-109		97-182	54-104		83-140	45-91
98-157	50-111		127-226	57-101			

Tabla 3.2: Datos sobre los rangos de la presión arterial sistólica (X) y diastólica (Y)

3.3 Los f-índices de desigualdad para conjuntos aleatorios. Resultados inferenciales en poblaciones finitas

Sean Ω una población cualquiera, $(\Omega, \mathcal{A}, P)$ el espacio de probabilidad asociado a ella, y $X : \Omega \to \mathcal{K}_c((0, +\infty))$ un conjunto aleatorio compacto y convexo integrablemente acotado y positivo asociado con $(\Omega, \mathcal{A}, P)$. Sea $f : (0, +\infty) \to \mathbb{R}$ una función monótona y estrictamente cóncava hacia arriba y de clase C^1, que satisface que $f(u) + f(1/u) \geq 0$ para todo $u \in (0, +\infty)$ y $f(1) = 0$.

Definición 3.3.1. *El f-índice de desigualdad asociado a X en la población Ω viene dado por:*

$$\mathrm{I}_f(X) = \frac{1}{2} E\left[f\left(\frac{\inf X}{E(\sup X)} \right) + f\left(\frac{\sup X}{E(\inf X)} \right) \right],$$

siempre que dicho valor exista.

Las condiciones de existencia y las propiedades de las Secciones 2.1 y 2.2 se particularizarían de forma inmediata para $\mathrm{I}_f(X)$.

Cuando se precise especificar la distribución P del espacio probabilístico se denotará $\mathrm{I}_f(X)$ por $\mathrm{I}_f(X \mid P)$.

A continuación se estudian las particularizaciones para conjuntos aleatorios de los resultados de las Secciones 2.3 y 2.4 para el problema de la estimación puntual de un índice de desigualdad poblacional en muestreos aleatorios simple y con reposición de una población finita y para el problema de la determinación bajo ciertas condiciones de la distribución asintótica de los índices muestrales en poblaciones finitas y de los procedimientos inferenciales aproximados derivados de ella.

Sea Ω una población finita de N unidades, $\omega_1, \ldots, \omega_N$, y $X : \Omega \to \mathcal{K}_c((0, +\infty))$ un conjunto aleatorio compacto y convexo asociado al espacio probabilístico definido sobre Ω cuando se supone dotado con la distribución uniforme.

Si se elige una muestra de tamaño n al azar y sin reposición a partir de Ω, υ representa una muestra aleatoria simple genérica de tamaño n, y $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ son las unidades en υ, entonces el *índice de desigualdad hiperbólico muestral* de X, que en la muestra υ viene dado por:

$$\mathrm{I}_H\,(X[\upsilon]) = \frac{1}{2n^2}\sum_{i=1}^{n}\sum_{i'=1}^{n}\left[\frac{\sup X(\omega_{\upsilon i})}{\inf X(\omega_{\upsilon i'})} + \frac{\inf X(\omega_{\upsilon i})}{\sup X(\omega_{\upsilon i'})}\right] - 1,$$

es una variable aleatoria real asociada al espacio probabilístico $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ (ver Sección 1.3, pág. 45), y por lo tanto define un estimador del índice hiperbólico poblacional, que corresponde al valor:

$$\mathrm{I}_H(X\,|\,P) = \frac{1}{2N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}\left[\frac{\sup X(\omega_j)}{\inf X(\omega_{j'})} + \frac{\inf X(\omega_j)}{\sup X(\omega_{j'})}\right] - 1.$$

Se cumple, además, que:

Teorema 3.3.1 . *En el muestreo aleatorio simple de tamaño n a partir de Ω, se cumple que el estimador $\widehat{\mathrm{I}_H}(X[\,\cdot\,])$ tal que:*

$$\widehat{\mathrm{I}_H}(X[\upsilon]) = \frac{n(N-1)}{N(n-1)}\mathrm{I}_H\,(X[\upsilon]) - \frac{1-f}{n}\left\{\frac{1}{n-1}\sum_{i=1}^{n}\mathrm{I}_H\,(\{X(\omega_{\upsilon i})\})\right\}$$

es un estimador insesgado de $\mathrm{I}_H(X\,|\,P)$, donde $\mathrm{I}_H\,(\{X(\omega_{\upsilon i})\})$ representa la (intra)desigualdad hiperbólica del conjunto aleatorio compacto y convexo degenerado en el intervalo $X(\omega_{\upsilon i}) \in \mathcal{K}_c((0,+\infty))$.

La precisión del estimador de $\mathrm{I}_H(X)$ anterior, se discute en el resultado siguiente:

Teorema 3.3.2 . *En el muestreo aleatorio simple de tamaño n a partir de Ω, si $f = n/N$ se cumple que:*

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}(X[\,\cdot\,])\right) = \frac{(1-f)}{4n(n-1)N^2(N-1)(N-2)(N-3)}$$

$$
\begin{aligned}
&\cdot \Bigg\{ 4[N(6-4n)+6(n-1)]N^3(N-1)\left(\mathrm{I}_H(X \mid P)\right)^2 \\
&+4[4N^2(3-2n)+13N(n-1)+3(n-3)]N^2(N-1)\mathrm{I}_H(X \mid P) \\
&+4[N^2(7-3n)+N(5n-7)-4(n-2)]\left(\sum_{j=1}^{N}\mathrm{I}_H(\{X(\omega_j)\})\right)^2 \\
&+4[N^2(n-5)+N(5n+1)-10(n-2)]\sum_{j=1}^{N}\left(\mathrm{I}_H(\{X(\omega_j)\})\right)^2 \\
&+8[N^3(3-n)+N^2(3n-8)+N(n+9)-10(n-2)]\sum_{j=1}^{N}\mathrm{I}_H(\{X(\omega_j)\}) \\
&+(n-2)(N-1)(N-2)\sum_{j=1}^{N}\left(\sum_{l=1}^{N}\left[\frac{\inf X(\omega_j)}{\sup X(\omega_l)}+\frac{\sup X(\omega_j)}{\inf X(\omega_l)}\right.\right. \\
&\left.\left.+\frac{\inf X(\omega_l)}{\sup X(\omega_j)}+\frac{\sup X(\omega_l)}{\inf X(\omega_j)}\right]\right)^2 \\
&+(N-n+1)(N-1)(N-3)\sum_{j=1}^{N}\sum_{l=1}^{N}\left[\frac{\inf X(\omega_j)}{\sup X(\omega_l)}+\frac{\sup X(\omega_j)}{\inf X(\omega_l)}\right]^2 \\
&+2(N-n+1)(N-1)(N-3)\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{I}_H(\{X(\omega_j)\cdot X(\omega_l)\}) \\
&-2[N(3n-7)-3(n-3)](N-1)\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{I}_H(\{X(\omega_l)\}) \\
&\cdot\left[\frac{\inf X(\omega_j)}{\sup X(\omega_l)}+\frac{\sup X(\omega_j)}{\inf X(\omega_l)}\right] \\
&+4[N^5(6-4n)+8N^4n+N^3(n-8)+N^2(9-7n)+9Nn+(23-13n)]\Bigg\}.
\end{aligned}
$$

Si a partir de la población $\Omega = \{\omega_1, \ldots, \omega_N\}$ se considera una selección aleatoria con reposición, υ representa una muestra aleatoria con reposición genérica de tamaño n, y $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ son las unidades en υ, entonces el *índice de desigualdad hiperbólico muestral* de X, $\mathrm{I}_H\left(X[\cdot]\right)$, es una variable aleatoria real asociada al espacio probabilístico $(\Upsilon_n^w, \mathcal{P}(\Upsilon_n^w), p^w)$ (ver Sección

1.3, pág. 51), y por lo tanto define un estimador del *índice hiperbólico poblacional*. Se cumple además que:

Teorema 3.3.3. *En el muestreo aleatorio con reposición de tamaño n a partir de* Ω*, se cumple que el estimador* $\widehat{\mathrm{I}_H}^{w}(X[\,\cdot\,])$ *tal que:*

$$\widehat{\mathrm{I}_H}^{w}(X[v]) = \frac{n}{(n-1)}\mathrm{I}_H\left(X[v]\right) - \frac{1}{n}\left\{\frac{1}{n-1}\sum_{i=1}^{n}\mathrm{I}_H\left(\{X(\omega_{vi})\}\right)\right\},$$

es un estimador insesgado de $\mathrm{I}_H(X\,|\,P)$.

Se satisface, también, que:

Teorema 3.3.4. *En el muestreo aleatorio con reposición de tamaño n a partir de* Ω*, se cumple que:*

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}^{w}(X[\,\cdot\,])\right) = \frac{1}{4n(n-1)N^3}$$

$$\cdot\Bigg\{4(6-4n)N^3\left(\mathrm{I}_H(X\,|\,P)\right)^2 + 8(6-4n)N^3\mathrm{I}_H(X\,|\,P)$$

$$-4[N-4(n-1)]\left(\sum_{j=1}^{N}\mathrm{I}_H(\{X(\omega_j)\})\right)^2 + 4[N-4(n-1)]\sum_{j=1}^{N}\left(\mathrm{I}_H(\{X(\omega_j)\})\right)^2$$

$$-8[N-4(n-1)](N-1)\sum_{j=1}^{N}\mathrm{I}_H(\{X(\omega_j)\})$$

$$+(n-2)\sum_{j=1}^{N}\left(\sum_{l=1}^{N}\left[\frac{\inf X(\omega_j)}{\sup X(\omega_l)} + \frac{\sup X(\omega_j)}{\inf X(\omega_l)} + \frac{\inf X(\omega_l)}{\sup X(\omega_j)} + \frac{\sup X(\omega_l)}{\inf X(\omega_j)}\right]\right)^2$$

$$+N\sum_{j=1}^{N}\sum_{l=1}^{N}\left[\frac{\inf X(\omega_j)}{\sup X(\omega_l)} + \frac{\sup X(\omega_j)}{\inf X(\omega_l)}\right]^2$$

$$+2N\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{I}_H(\{X(\omega_j)\cdot X(\omega_l)\}) + 4[N^2(5-4n) + N(4n-3) + (5-4n)]N\Bigg\}.$$

Por otro lado, el estudio de la distribución asintótica de los índices de desigualdad muestrales va a permitir la estimación asintóticamente insesgada bajo ciertas condiciones, la determinación de la precisión asintótica de las estimaciones, y el desarrollo de aproximaciones de otras inferencias, para la mayoría de los índices poblacionales.

Supongamos que se considera una población finita de tamaño N, $\Omega = \{\omega_1, \ldots, \omega_N\}$, y sea X un conjunto aleatorio compacto y convexo positivo definido sobre Ω. Si sobre la población Ω el conjunto aleatorio X toma r valores distintos, $x_1^*, \ldots, x_r^* \in \mathcal{K}_c((0,+\infty))$, y para $l \in \{1, \ldots, r\}$ se denota por p_l la probabilidad $P(\{\omega \in \Omega \mid X(\omega) = x_l^*\})$, donde P es la medida correspondiente a la distribución uniforme sobre Ω, y $f : (0,+\infty) \to \mathbb{R}$ es una función monótona, cóncava hacia arriba en el sentido estricto y de clase C^1, que satisface que $f(u) + f(1/u) \geq 0$ para todo $u \in (0,+\infty)$ y $f(1) = 0$, el *f-índice de desigualdad poblacional* de X viene dado por el valor:

$$\mathrm{I}_f(X \,|\, \mathbf{p}) = F\left(\sum_{l'=1}^{r} p_{l'} \cdot f\left(x_{l'}^* \Big/ \sum_{l=1}^{r} p_l \cdot x_l^*\right)\right)$$
$$= \sum_{l'=1}^{r} p_{l'} F\left(f\left(x_{l'}^* \Big/ \sum_{l=1}^{r} p_l \cdot x_l^*\right)\right).$$

Si se selecciona al azar una muestra de tamaño n de la población y f_{nl} representa la frecuencia relativa del valor x_l^* de X en esa muestra, el *f-índice de desigualdad muestral* de X corresponde al valor:

$$\mathrm{I}_f(X \,|\, \mathbf{f}_n) = F\left(\sum_{l'=1}^{r} f_{nl'} \cdot f\left(x_{l'}^* \Big/ \sum_{l=1}^{r} f_{nl} \cdot x_l^*\right)\right)$$
$$= \sum_{l'=1}^{r} f_{nl'} F\left(f\left(x_{l'}^* \Big/ \sum_{l=1}^{r} f_{nl} \cdot x_l^*\right)\right).$$

La distribución asintótica del f-índice de desigualdad muestral viene dada por:

Teorema 3.3.5. *Para cada $n \in \mathbb{N}$, se consideran n conjuntos aleatorios compactos y convexos independientes e idénticamente distribuidos que X*

(es decir, una muestra aleatoria simple de tamaño n a partir de X) definida sobre la población finita $\Omega = \{\omega_1, \ldots, \omega_N\}$ *de forma que* $P(\{\omega \in \Omega \mid X(\omega) = x_l^*\}) = p_l$ *con* $p_l \in (0,1)$, $l = 1, \ldots, r$, $\sum_{l=1}^{r} p_l = 1$. *Sea* $f : (0, +\infty) \to \mathbb{R}$ *una función monótona y cóncava hacia arriba en sentido estricto, que satisface que* $f(u) + f(1/u) \geq 0$ *para todo* $u \in (0, +\infty)$ *y que* $f(1) = 0$, *y admite derivadas de tercer orden finitas. Se cumple entonces que:*

i) *Si* $\mathbf{f}_n = (f_{n1}, \ldots, f_{n(r-1)}) \in [0,1]^{r-1}$, *con* f_{nl} = *frecuencia relativa de* x_l^* *en la correspondiente realización de la muestra aleatoria simple de tamaño n* ($l = 1, \ldots, r-1$, $f_{nr} = 1 - \sum_{l=1}^{r-1} f_{nl}$), *e* $\mathrm{I}_f(X \mid \mathbf{f}_n)$ *es el índice de desigualdad muestral asociado, entonces* $\{\mathrm{I}_f(X \mid \mathbf{f}_n)\}_n$ *es una sucesión de estimadores de* $\mathrm{I}_f(X) = \mathrm{I}_f(X \mid \mathbf{p})$ *fuertemente consistente, es decir, cuando* $n \to \infty$:

$$\mathrm{I}_f(X \mid \mathbf{f}_n) \stackrel{c.s.}{\to} \mathrm{I}_f(X \mid \mathbf{p})$$

cualquiera que sea $\mathbf{p} = (p_1, \ldots, p_{r-1})$ *(con* $p_1, \ldots, p_{r-1} \in (0,1)$ *y* $\sum_{l=1}^{r-1} p_l < 1$*).*

ii) $\{\sqrt{n}\,(\mathrm{I}_f(X \mid \mathbf{f}_n) - \mathrm{I}_f(X \mid \mathbf{p}))\}_n$ *es una sucesión de variables aleatorias que converge en ley hacia una normal unidimensional* $N(0, \sigma^2(\mathbf{p}))$, *con:*

$$\sigma^2(\mathbf{p}) = \sum_{l'=1}^{r} p_{l'} (V_{l'}^*)^2 - \left(\sum_{l'=1}^{r} p_{l'} V_{l'}^* \right)^2,$$

donde:

$$V_{l'}^* = \frac{1}{2} \left[f\left(\frac{\sup x_{l'}^*}{E(\inf X \mid \mathbf{p})} \right) + f\left(\frac{\inf x_{l'}^*}{E(\sup X \mid \mathbf{p})} \right) \right]$$

$$-\frac{1}{2} \sum_{l=1}^{r} p_l \Bigg[\inf x_{l'}^* \cdot f'\left(\frac{\sup x_l^*}{E\,(\inf X \mid \mathbf{p})} \right) \cdot \left(\frac{\sup x_l^*}{[E\,(\inf X \mid \mathbf{p})]^2} \right)$$

$$+ \sup x_{l'}^* \cdot f'\left(\frac{\inf x_l^*}{E\,(\sup X \mid \mathbf{p})} \right) \cdot \left(\frac{\inf x_l^*}{[E\,(\sup X \mid \mathbf{p})]^2} \right) \Bigg],$$

siempre que $\sigma^2(\mathbf{p}) > 0$.

iii) *Si* $\sigma^2(\mathbf{p}) = 0$, *y para algún par* (i,j) *con* $i,j \in \{1,\ldots,r-1\}$ *se cumple que:*

$$h_{ij} = \frac{\partial^2 \mathrm{I}_f(X\,|\,\mathbf{p})}{\partial p_i \partial p_j} = \frac{\partial}{\partial p_i}[V_j^* - V_r^*] > 0,$$

entonces $\{2n\,(\mathrm{I}_f(X\,|\,\mathbf{f}_n) - \mathrm{I}_f(X\,|\,\mathbf{p}))\}_n$ *es una sucesión de variables aleatorias que converge en ley hacia una combinación lineal de, a lo sumo,* $r-1$ *variables chi-cuadrado* χ_1^2 *independientes.*

Se cumple, además, que:

Teorema 3.3.6. *En las condiciones del Teorema* 3.3.5*, se cumple que:*

$$\left\{ \frac{\sqrt{n}\,(\mathrm{I}_f(X\,|\,\mathbf{f}_n) - \mathrm{I}_f(X\,|\,\mathbf{p}))}{\sqrt{\sigma^2(\mathbf{f}_n)}} \right\}_n$$

converge en ley hacia una distribución normal $N(0,1)$ *cuando* $n \to \infty$, *siempre que* $\sigma^2(\mathbf{p}) > 0$ *y* $\sigma^2(\mathbf{f}_n) > 0$.

Los procedimientos inferenciales aproximados que se derivan de los dos últimos resultados son los siguientes:

Teorema 3.3.7. *En las condiciones de los Teoremas* 3.3.5 *y* 3.3.6*, el intervalo aleatorio*

$$\left[\mathrm{I}_f(X\,|\,\mathbf{f}_n) - z_\alpha \sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}},\ \mathrm{I}_f(X\,|\,\mathbf{f}_n) + z_\alpha \sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}} \right]$$

proporciona para cada muestra de observaciones independientes a partir de X *intervalos de confianza con coeficiente aproximadamente* $1-\alpha$ *(para* $\alpha \in [0,1]$*).*

Teorema 3.3.8. *En las condiciones de los Teoremas* 3.3.5 *y* 3.3.6*:*

(i) Para contrastar al nivel de significación $\alpha \in [0,1]$ *la hipótesis nula:*

$$H_0 : \mathrm{I}_f(X\,|\,\mathbf{p}) = i_0$$

frente a la alternativa:

$$H_1 : \mathrm{I}_f(X \,|\, \mathbf{p}) \neq i_0,$$

si $|\mathrm{I}_f(X \,|\, \mathbf{f}_n) - i_0| > z_\alpha \sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$ *debe rechazarse* H_0. *El p-valor del contraste vendrá dado, aproximadamente, por:*

$$p = 2\left[1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\,|\mathrm{I}_f(X \,|\, \mathbf{f}_n) - i_0|\right)\right].$$

(ii) Para contrastar al nivel de significación α *la hipótesis nula:*

$$H_0 : \mathrm{I}_f(X \,|\, \mathbf{p}) \geq i_0$$

frente a la alternativa:

$$H_1 : \mathrm{I}_f(X \,|\, \mathbf{p}) < i_0,$$

si $\mathrm{I}_f(X \,|\, \mathbf{f}_n) - i_0 < z_{2\alpha} \sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$ *debe rechazarse* H_0. *El p-valor del contraste vendrá dado, aproximadamente, por:*

$$p = \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\,(\mathrm{I}_f(X \,|\, \mathbf{f}_n) - i_0)\right).$$

(iii) Para contrastar al nivel de significación α *la hipótesis nula:*

$$H_0 : \mathrm{I}_f(X \,|\, \mathbf{p}) \leq i_0$$

frente a la alternativa:

$$H_1 : \mathrm{I}_f(X \,|\, \mathbf{p}) > i_0,$$

si $\mathrm{I}_f(X \,|\, \mathbf{f}_n) - i_0 > z_{2\alpha} \sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$ *debe rechazarse* H_0. *El p-valor del contraste vendrá dado, aproximadamente, por:*

$$p = 1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\,(\mathrm{I}_f(X \,|\, \mathbf{f}_n) - i_0)\right).$$

Como aplicación de los estudios precedentes se considera el siguiente ejemplo.

Ejemplo 3.3.1 . Una gran empresa está interesada en cuantificar la desigualdad de los salarios netos mensuales (excluyendo las pagas extra) de sus empleados en un año. Sin embargo, cada empleado suele cobrar un salario diferente cada mes debido al número de horas extraordinarias realizadas, al pago de atrasos o a otro tipo de causas (variación en el IRPF aplicado, cambios más o menos sustanciales en los descuentos, ...) por lo que la empresa decide elaborar un protocolo de encuesta en el que se pide a los empleados dar, del modo más ajustado posible, el rango a lo largo del año de su salario neto mensual.

Si el protocolo no contempla una lista exhaustiva de posibles respuestas excluyentes, éstas consistirán habitualmente en distintos intervalos de $(0, +\infty)$ que a menudo se solaparán.

325000-335000	267500-275000	515000-530000
125000-130000	185000-200000	240000-255000
405000-420000	400000-410000	190000-200000
200000-210000	370000-385000	440000-457500
260000-267500	210000-217500	245000-257500
185000-200000	222500-230000	215000-225000
347500-352500	410000-422500	310000-317500
295000-310000	150000-160000	265000-280000
415000-425000	272500-277500	295000-305000
255000-266000	360000-370000	280000-290000
310000-317500	225000-235000	420000-432500
190000-200000	167500-180000	197500-205000
410000-422500	280000-290000	255000-266000
205000-215000	267500-275000	370000-385000
330000-340000	430000-445000	237500-250000

Tabla 3.3: Datos sobre rangos de salarios mensuales

Supóngase que, como primer acercamiento al problema, se desea aproximar la desigualdad anterior sobre la base de una muestra de 45 empleados de la empresa, seleccionada al azar y con reposición, y que las respuestas correspondientes son las que aparecen recogidas en la Tabla 3.1.

Si esos 45 intervalos se contemplan como los valores que cierto conjunto aleatorio compacto y convexo X toma sobre la muestra seleccionada, que se representará por $[\,\tau\,]$, una estimación insesgada del índice de desigualdad hiperbólico poblacional valorado en intervalo, $\mathrm{I}_H(X)$ viene dada en virtud del Teorema 3.3.3, por el valor $\widehat{\mathrm{I}_H}^w(X[\,\tau\,]) = 0,105$.

3.4 Valoración final y problemas abiertos

Para justificar el enfoque y los fundamentos que se han adoptado para la cuantificación de la variación de un conjunto aleatorio en este capítulo, debe aducirse como argumento principal el que esa adopción ha obedecido a la suposición de que la variable involucrada es esencialmente valorada en intervalo.

Los estudios desarrollados en este tercer capítulo son aplicables al caso en que se parte de la observación de una variable aleatoria con valores reales, pero por diversas razones (bien por imprecisión en los mecanismos de medición, o por simple tradición o por su utilidad al objeto de buscar modelos de distribuciones que se ajusten bien a la variable original considerada), los datos se presenten agrupados en intervalos. Estos datos agrupados podrían contemplarse como provenientes de un conjunto aleatorio compacto y convexo con valores dos a dos disjuntos.

Sería interesante comprobar mediante diferentes procedimientos la bondad de las medidas Δ_S^2 e I_f como aproximaciones de los verdaderos valores de la varianza y la f-desigualdad de la variable real subyacente, respectivamente. Esta discusión podría realizarse para Δ_S^2 en función de la elección de la medida S, la distribución de la variable real y la agrupación en intervalos que se emplee, y para I_f en términos de la elección de la función f, la distribución de la variable real y la agrupación en intervalos considerada.

No obstante, debe tenerse en cuenta que Δ_S^2 e I_f, como se han definido en las Secciones 3.1 y 3.3, respectivamente, miden en realidad la variación de un elemento aleatorio cuyos valores son los datos agrupados en toda su extensión, y no la de una variable real para la que se ha identificado cada uno de sus valores reales con un intervalo al que pertenecen.

En cuanto a los problemas abiertos que se plantean de forma inmediata en relación con el tema desarrollado en este capítulo, pueden considerarse las particularizaciones de los planteados en las Secciones 1.7 y 2.5. De hecho, muchos de estos problemas se comenzarán estudiando para conjuntos aleatorios, por tratarse de una situación menos compleja y más fácil de interpretar geométricamente (como ocurrirá, por ejemplo, en los estudios sobre correlación o sobre regresión con funciones no lineales). Además, en varios de los problemas el razonamiento llevado a cabo con los conjuntos aleatorios podrá aplicarse directamente sobre las funciones α-nivel de las variables aleatorias difusas, lo que permitirá obtener conclusiones sin dificultad para estas últimas.

Apéndice A

Sobre la métrica que sirve de base para la medida generalizada de variación absoluta

Como se ha señalado en la presentación de la métrica d_∞, Klement *et al.* (1986) han probado que el espacio $(\mathcal{F}_c(\mathbb{R}), d_\infty)$ es no separable. En la no separabilidad de este espacio métrico reside buena parte de su interés teórico, pero también obedecen a ella muchas de las dificultades que surgen en algunos estudios probabilísticos sobre variables aleatorias difusas.

Sin embargo, cuando sobre $\mathcal{F}_c(\mathbb{R})$ se considera la métrica D_S (y, en forma análoga, cuando sobre $\mathcal{K}_c(\mathbb{R})$ se considera d_S), el espacio métrico resultante es separable. Los resultados siguientes formalizan esta afirmación.

Proposición A.1. $(\mathcal{F}_c(\mathbb{R}), D_S)$ es un espacio métrico equivalente a $(\mathcal{F}_c(\mathbb{R}), d_2)$, donde $d_2 : \mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \to [0, +\infty)$ es la métrica que a $\widetilde{V}, \widetilde{W} \in \mathcal{F}_c(\mathbb{R})$ le asocia el valor:

$$d_2(\widetilde{V}, \widetilde{W}) = \sqrt{\int_{(0,1]} \left[d_H(\widetilde{V}_\alpha, \widetilde{W}_\alpha) \right]^2 d\alpha}.$$

Demostración:

Para probar la equivalencia entre D_S y d_2 se tiene en cuenta que para $\widetilde{V}, \widetilde{W} \in \mathcal{F}_c(R)$ se cumple que:

$$D_S(\widetilde{V}, \widetilde{W}) = \sqrt{\int_{(0,1]} \left[d_S\left(\widetilde{V}_\alpha, \widetilde{W}_\alpha\right) \right]^2 d\alpha}$$

$$= \sqrt{\int_{(0,1]} \int_{[0,1]} \left[\lambda(\sup \widetilde{V}_\alpha - \sup \widetilde{W}_\alpha) + (1-\lambda)(\inf \widetilde{V}_\alpha - \inf \widetilde{W}_\alpha) \right]^2 dS(\lambda)\, d\alpha}$$

$$\leq \sqrt{\int_{(0,1]} \left[\text{máx}\,\{|\sup \widetilde{V}_\alpha - \sup \widetilde{W}_\alpha|, |\inf \widetilde{V}_\alpha - \inf \widetilde{W}_\alpha|\}\right]^2 d\alpha} = d_2(\widetilde{V}, \widetilde{W}).$$

Por otro lado:

$$d_2(\widetilde{V}, \widetilde{W}) = \sqrt{\int_{(0,1]} \left[\text{máx}\left\{|\sup \widetilde{V}_\alpha - \sup \widetilde{W}_\alpha|, |\inf \widetilde{V}_\alpha - \inf \widetilde{W}_\alpha|\right\}\right]^2 d\alpha}$$

$$\leq \sqrt{\int_{(0,1]} \left[(\sup \widetilde{V}_\alpha - \sup \widetilde{W}_\alpha)^2 + (\inf \widetilde{V}_\alpha - \inf \widetilde{W}_\alpha)^2\right] d\alpha} = \sqrt{2}\, D_{S_2}(\widetilde{V}, \widetilde{W}),$$

donde S_2 es la medida ponderada tal que la función asociada g satisface que $g(\lambda) = 0$ para todo $\lambda \in (0,1)$ y $g(0) = g(1) = 0,5$.

De acuerdo con Bertoluzza *et al.* (1995a), todas las métricas D_S son topológicamente equivalentes, con lo que se concluye el resultado. □

Corolario A.2. $(\mathcal{F}_c(\mathbb{R}), D_S)$ es un espacio métrico separable.

Demostración:

Obviamente, la separabilidad de $(\mathcal{F}_c(\mathbb{R}), d_2)$ (Diamond & Kloeden 1994) implica la separabilidad de $(\mathcal{F}_c(\mathbb{R}), D_S)$. □

Apéndice B

Sobre la función que sirve de base para la medida generalizada de variación relativa

En el estudio de la extensión de los f-índices de desigualdad para variables aleatorias difusas y conjuntos aleatorios compactos y convexos, se ha supuesto que la función f satisface algunas condiciones adicionales frente a las exigidas en el caso real. La restricción que imponen las nuevas condiciones (y, especialmente, la que imponen dos de las condiciones supuestas en el análisis de la minimalidad o anulación de las medidas de desigualdad) permite establecer ciertas caracterizaciones de las correspondientes funciones f.

El objetivo general que se plantea en este apartado es buscar todas las funciones $f : (0, +\infty) \to \mathbb{R}$ que sean estrictamente cóncavas hacia arriba y monótonas, de clase C^1 y tales que $f(u) + f(1/u) \geq 0$ para todo $u \in (0, +\infty)$ y $f(1) = 0$.

En primer lugar, se va a buscar la subfamilia caracterizada por la condición $f(u) + f(1/u) = 0$ para todo $u \in (0, +\infty)$ con f dos veces derivable, a la que se ha recurrido en el Teorema 2.2.6 del Capítulo 2.

Proposición B.1. El conjunto de las soluciones de la ecuación funcional $f(u) + f(1/u) = 0$ para todo $u \in (0, +\infty)$ con $f : (0, +\infty) \to \mathbb{R}$ derivable dos veces, estrictamente cóncavas hacia arriba y monótonas, y $f(1) = 0$, viene definido por la familia de funciones $f(u) = h(\log u)$ con $h : \mathbb{R} \to \mathbb{R}$ tal que para todo $x \in \mathbb{R}$:

(i) $h(x) = -h(-x)$ (es decir, h impar),

(ii) h estrictamente decreciente, y

(iii) $h''(x) \geq h'(x)$.

Demostración:

Las transformaciones $x = \log u$ y $h(x) = f(e^x)$ permiten concluir que $f(u) + f(1/u) = 0$ para todo $u \in (0, +\infty)$ si, y sólo si, se cumple que $f(u) = h(\log u)$ para todo $u \in (0, +\infty)$ y $h(x) = -h(-x)$ para todo $x \in \mathbb{R}$. Además, como $h(0) = 0$, se satisface que $f(1) = 0$.

Por otro lado, para probar que se requiere la condición *(ii)* debe tenerse en cuenta que como $g(x) = \log x$ es una función estrictamente creciente, $f = h \circ g$ será estrictamente monótona si, y sólo si, h es estrictamente monótona y f y h tienen el mismo sentido de monotonía.

En este caso, h debe ser decreciente, puesto que si h fuera creciente al ser f cóncava hacia arriba se cumpliría que para todo $u \in (0, +\infty)$:

$$0 \leq f''(u) = \frac{1}{u^2} \left(h''[\log u] - h'[\log u] \right),$$

de modo que $h''(\log u)$ debería ser no negativo, y, por lo tanto, h debería ser cóncava hacia arriba. Pero para que h sea creciente, impar y cóncava hacia arriba, necesariamente debe ser $h(x) = ax$ con $a > 0$, lo que contradiría el hecho de que $f(u) = a \log u$ para todo $u \in (0, +\infty)$ (con $a > 0$) fuera cóncava hacia arriba.

En consecuencia, como f es derivable también lo será h y se tendrá que $h'(x) \leq 0$.

Finalmente, como f es derivable dos veces también lo será h, y al ser f cóncava hacia arriba se cumplirá que:

$$0 \leq f''(u) = \frac{1}{u^2} \left(h''[\log u] - h'[\log u] \right),$$

de donde se deduce que debe ser $h''(\log u) \geq h'(\log u)$ para todo $u \in (0, +\infty)$. □

Observación B.1. Conviene señalar que las condiciones *(i)* y *(ii)* son necesarias aun cuando no se exija a f condiciones de derivabilidad. Por otra parte, el índice de tipo Shannon extendido (correspondiente a $h(x) = -x$) va asociado obviamente a una función f que es solución de la ecuación funcional de la Proposición B.1.

En cuanto a las soluciones de la subfamilia caracterizada por la relación $f(u) + f(1/u) \geq 0$ para todo $u \in (0, +\infty)$ con igualdad si, y sólo si, $u = 1$, con f dos veces derivable, razonando en forma análoga a la Proposición B.1 (a la que se ha hecho referencia en el Teorema 2.2.5 del Capítulo 2), y teniendo en cuenta que la suposición $f(1) = 0$ obliga a que $f(u) + f(1/u) > 0$ si $u \in (0, +\infty) \setminus \{1\}$, puede obtenerse el resultado siguiente:

Proposición B.2. El conjunto de las funciones $f : (0, +\infty) \to \mathbb{R}$ derivables dos veces, estrictamente cóncavas hacia arriba y monótonas, y tales que $f(u) + f(1/u) > 0$ para todo $u \in (0, +\infty) \setminus \{1\}$ y $f(1) = 0$, es la familia de funciones $f(u) = h(\log u)$ con $h : \mathbb{R} \to \mathbb{R}$ tal que:

(i) $h(x) > -h(-x)$ para todo $x \in \mathbb{R} \setminus \{0\}$ y $h(0) = 0$,

(ii) h estrictamente monótona, y

(iii) $h''(x) \geq h'(x)$ para todo $x \in \mathbb{R}$.

La extensión de los índices aditivamente descomponibles para $\alpha \neq 0, 1$ (correspondientes a $h(x) = e^{\alpha x} - 1$ para $\alpha < 0$ y $\alpha > 1$ y a $h(x) = 1 - e^{\alpha x}$ para $\alpha \in (0, 1)$) va asociada a una función f que cumple las condiciones de la Proposición B.2.

Si se pretende ahora encontrar soluciones de la desigualdad funcional general $f(u) + f(1/u) \geq 0$ para todo $u \in (0, +\infty)$ con $f : (0, +\infty) \to \mathbb{R}$ dos veces derivable, estrictamente cóncava hacia arriba y monótona, y tal que $f(1) = 0$, las transformaciones consideradas en la demostración de la Proposición B.1 implican que $h(x) \geq -h(-x)$ y $h(0) = 0$.

Teniendo en cuenta que h ha de ser estrictamente monótona (por serlo f), para determinar funciones h que satisfagan la desigualdad anterior podría razonarse de la siguiente forma:

Sean $h_1 : [0, +\infty) \to \mathbb{R}$ y $h_2 : (-\infty, 0] \to \mathbb{R}$ estrictamente decrecientes tales que $h_1(0) = h_2(0) = 0$ y $h_2(x) \geq -h_1(-x)$. Se define la función $h : \mathbb{R} \to \mathbb{R}$, como:

$$h(x) = \begin{cases} h_1(x) & \text{si } x \geq 0 \\ h_2(x) & \text{en caso contrario.} \end{cases}$$

La función h es evidentemente decreciente y satisface la desigualdad $h(x) \geq -h(-x)$ y $h(0) = 0$. Por lo tanto, la función $f(u) = h(\log u)$ es solución del problema planteado, siempre que f sea cóncava hacia arriba. Por otro lado, si se supone que f es derivable dos veces $f''(u) \geq 0$ si, y sólo si, $h'' \geq h'$, y esto ocurre si, y sólo si:

$$\begin{cases} h_1'' \geq h_1' & \text{si } x \geq 0 \\ h_2'' \geq h_2' & \text{en caso contrario,} \end{cases}$$

la función $f(u) = h(\log u)$, con h construida con el procedimiento sugerido, es solución del problema planteado si, y sólo si $h_1'' \geq h_1'$ y $h_2'' \geq h_2'$.

Debe subrayarse que, puesto que $h(0) = 0$ y h ha de ser estrictamente monótona, entonces o bien $h(x) \geq 0$ para todo $x \in (-\infty, 0]$ y $h(x) \leq 0$ para todo $x \in [0, +\infty)$ (caso considerado), o $h(x) \leq 0$ para todo $x \in (-\infty, 0]$ y $h(x) \geq 0$ para todo $x \in [0, +\infty)$. En este último caso, h sería creciente, pero como h ha de satisfacer la condición $h(x) \geq -h(-x)$, si, como en el caso anterior, se construyen funciones $h_1 : [0, +\infty) \to \mathbb{R}$ y $h_2 : (-\infty, 0] \to \mathbb{R}$ estrictamente crecientes tales que $h_1(0) = h_2(0) = 0$ y $h_1(x) \geq -h_2(-x)$, entonces la función $h : \mathbb{R} \to \mathbb{R}$, definida como:

$$h(x) = \begin{cases} h_1(x) & \text{si } x \geq 0 \\ h_2(x) & \text{en caso contrario,} \end{cases}$$

es evidentemente creciente y satisface la desigualdad $h(x) \geq -h(-x)$ y $h(0) = 0$.

Por lo tanto, la función $f(u) = h(\log u)$ es solución del problema planteado siempre que f sea cóncava hacia arriba. Pero siendo h creciente y f cóncava hacia arriba necesariamente obligan a que h sea cóncava hacia arriba. En este sentido, hay que limitarse (dentro de esta forma de construcción) a las funciones:

$$h(x) \; = \; \begin{cases} h_1(x) & \text{si } x \geq 0 \\ h_2(x) & \text{en caso contrario,} \end{cases}$$

crecientes, cóncavas hacia arriba tales que $f(u) = h(\log u)$ sea cóncava hacia arriba. Como el caso anterior, si f es derivable dos veces la función $f(u) = h(\log u)$ es solución del problema planteado si, y sólo si, $h_1'' \geq h_1'$ y $h_2'' \geq h_2'$.

Por último, es conveniente indicar que existen ejemplos de funciones f que satisfacen las condiciones generales que se exigen en los Capítulos 2 y 3, si bien no se ajustan a los modelos recogidos en las Proposiciones B.1 y B.2. En este sentido, la función $f(x) = 1/ex$ si $x \in (0, 1/e]$, $= -\log x$ si $x \in (1/e, +\infty)$ es de clase C^1, monótona y estrictamente cóncava hacia arriba, cumple que $f(1) = 0$, y $f(x) + f(1/x) = 0$ si $x \in [1/e, e]$, > 0 en el resto.

Epílogo

En esta memoria se ha desarrollado un estudio sobre la cuantificación de la variación asociada a un elemento aleatorio cuyos valores no se corresponden con números reales, sino o con intervalos compactos reales o bien con números difusos con clausura del soporte compacto.

Esa cuantificación se ha llevado a cabo en una escala real, puesto que se ha considerado que el interés prioritario era el de poder realizar comparaciones entre variables, poblaciones, estimadores, etc. sobre la base de la variación asociada.

El esquema común seguido tanto en la medición de la variación en sentido absoluto (o desviación), como en la de la variación en sentido relativo (o desigualdad), ha consistido en hallar el promedio de la desviación y desigualdad, respectivamente, de cada valor (difuso o de intervalo) del elemento aleatorio con respecto a su valor esperado.

De este modo, la medida generalizada de la variación absoluta introducida en esta memoria viene definida como el valor esperado (en el sentido clásico) de la S-distancia entre cada valor de la variable aleatoria difusa o del conjunto aleatorio y su valor esperado (según Puri & Ralescu 1986, o Aumann 1965, respectivamente), y la generalización a la que se hace referencia obedece a la posibilidad de optar por distintas medidas S que deben satisfacer condiciones bastante generales.

Por otro lado, la medida generalizada de la variación relativa que se ha presentado viene definida como el valor que la función de ordenación F de Yager de la Definición 0.1.3 asigna al valor esperado de la imagen (difusa) inducida por la función f del "cociente" (difuso o de intervalo) entre cada valor del elemento aleatorio y su valor esperado. En este caso, la generalización a la que se refiere la denominación del concepto se debe a la posibilidad de elegir diferentes funciones f que deben satisfacer ciertas condiciones.

Una cuestión que surge de forma inmediata a la vista de este esquema común, es la de por qué se han definido nuevas medidas para este tipo de datos, y no se ha procedido a adoptar algún criterio de "codificación real" previa de esos datos recurriendo posteriormente al tratamiento de los datos codificados mediante medidas de variación bien conocidas del caso real.

La razón por la que no se ha acudido a la consideración de una función "codificadora" que transforme cada valor impreciso (número difuso o intervalo) en un valor real, es que la aplicación previa de las transformaciones más naturales conduciría a identificar valores diferentes entre sí. En consecuencia, la distancia euclídea, o la imagen por la función f del cociente, entre los valores transformados podría ser nula mientras que los valores imprecisos originales llevarían asociada realmente una variación no nula que la codificación habría ignorado.

Como ejemplo de esta última afirmación, si se admitiera la función $F(\cdot)$ de Yager como transformación "codificadora" de los valores de una variable aleatoria difusa $\mathcal{X}$ que tomara como valores: $\tilde{x}_1 = \text{Tri}(-1, 0, 1)$, $\tilde{x}_2 = \text{Tri}(-2.000, 0, 2.000)$ y $\tilde{x}_3 = \text{Tri}(-3, 1, 1)$, con probabilidades inducidas 1/3, 1/3 y 1/3, se obtendría que $F \circ \mathcal{X}$ sería una variable aleatoria degenerada en el valor 0 a la que le correspondería una varianza nula. Análogamente, si para cierto valor prefijado λ se admitiera la función $f_{(\cdot)}(\lambda) : \mathcal{K}_c(\mathbb{R}) \to \mathbb{R}$ (ver Definiciones 0.1.5 y 3.1.1) como transformación "codificadora" de los valores de un conjunto aleatorio compacto y convexo, ocurriría lo mismo.

Los comentarios precedentes no son igualmente vigentes cuando se trata el caso especial de datos agrupados, puesto que en esta situación dos intervalos de clase distintos (aunque no disjuntos) en dos agrupaciones diferentes pueden corresponder al mismo conjunto de datos reales, con lo que el error cometido en la identificación por la transformación "codificadora" (por ejemplo, la representación de cada intervalo de clase con su punto medio) puede ser inferior al que le asociaría el tomar cada dato agrupado como un valor de intervalo según el Capítulo 3.

Además de los problemas abiertos específicos señalados al final de cada capítulo de la memoria, existen algunos problemas de carácter general que se apuntan brevemente a continuación:

- Generalización de los estudios para elementos aleatorios con valores en $\mathcal{F}(\mathbb{R})$ y $\mathcal{K}(\mathbb{R})$, con $\mathcal{K}(\mathbb{R})$ la clase de los conjuntos compactos y no vacíos de $\mathbb{R}$, y $\mathcal{F}(\mathbb{R})$ la clase de los subconjuntos difusos $\widetilde{V}$ de $\mathbb{R}$, tales que $\widetilde{V}_\alpha \in \mathcal{K}(\mathbb{R})$ para todo $\alpha \in [0, 1]$.

- Generalización de los estudios para elementos aleatorios con valores no necesariamente compactos (o con α-niveles compactos).

- Desarrollo de estudios asintóticos en el problema de Regresión Lineal para su aplicación a contrastes de hipótesis sobre los parámetros a y b de la relación lineal.

- Desarrollo de estudios sobre modelos lineales entre variables aleatorias difusas y conjuntos aleatorios con más de dos elementos (Regresión Múltiple, Análisis de la Varianza).

Referencias y Bibliografía Consultada

Adamo, J.M. (1980). Fuzzy decision trees. *Fuzzy Sets and Systems* **4**, 207-219.

Alonso, M.C. and Gil, M.A. (1995). Testing the convenience of a variate for stratification in estimating the Gini-Simpson diversity. En *Lecture Notes in Computer Sciences* **945**, Advances in Intelligent Computing (B. Bouchon and R. R. Yager, Eds.), 210-216. Springer-Verlag, Berlin.

Alonso, M.C., Bertoluzza, C., Brezmes, T. and Lubiano, M.A. (1998). *A generalized real-valued measure of the inequality associated with a fuzzy random variable.* (enviado para publicación).

Ash, R. (1972). *Real Analysis and Probability.* Academic Press, New York.

Atkinson, A.B. (1983). *The Economics of Inequality.* 2nd ed. Oxford University Press, New York.

Aumann, R.J. (1965). Integrals of set-valued functions. *J. Math. Anal. Appl.* **12**, 1-12.

Bandemer, H. and Gottwald, S. (1995). *Fuzzy Sets, Fuzzy Logic, Fuzzy Methods with Applications.* John Wiley & Sons, Chichester.

Bandemer, H. and Näther, W. (1992). *Fuzzy Data Analysis*. Kluwer Academic Publishers, Dortrecht.

Bárdossy, A. (1990). Note on fuzzy regression. *Fuzzy Sets and Systems* **37**, 65-75.

Barnett, V. (1991). *Sample Survey. Principles and Methods*. Edward Arnold, London.

Basu, K. (1987). Axioms for a fuzzy measure of inequality. *Math. Soc. Sci.* **14**, 275-288.

Bertoluzza, C., Corral, N. and Salas, A. (1995). On a new class of distances between fuzzy numbers. *Mathware & Soft Computing* **2**, 71-84.

Bertoluzza, C., Corral, N. and Salas, A. (1995). Polinomial regression in a fuzzy context. The least squares method, *Proc. 6th IFSA Conference*, 431-434.

Billingsley, P. (1995). *Probability and Measure*. 3rd ed. John Wiley & Sons, New York.

Blackorby, C. and Donaldson, D. (1978). Measures of relative equality and their meaning in terms of social welfare. *J. Econom. Theory* **18**, 59-80.

Bortolan, G. and Degani, P. (1985). A review of some methods for ranking fuzzy subsets. *Fuzzy Sets and Systems* **15**, 1-19.

Bourguignon, F. (1979). Decomposable income inequality measures. *Econometrica* **47**, 901-920.

Breiman, L. (1968). *Probability*. Addison-Wesley, Massachusetts.

Byrne, C. (1978). Remarks on the set-valued integrals of Debreu and Aumann. *J. Math. Anal. Appl.* **62**, 243-246.

Campos, L.M. de and González, A. (1989). A subjective approach for ranking fuzzy numbers. *Fuzzy Sets and Systems* **29**, 145-153.

Casella, G. and Berger, R.L. (1990). *Statistical Inference*. Wadsworth & Brooks/Cole, Pacific Grove.

Celminš, A. (1987). Least squares model fitting to fuzzy vector data. *Fuzzy Sets and Systems* **22**, 245-269.

Cochran, W.G. (1977). *Sampling Techniques*. John Wiley & Sons, New York.

Colubi, A. (1997). *Los f-índices de desigualdad difusos asociados a una variable aleatoria difusa*. Trabajo de Investigación de Tercer ciclo. Universidad de Oviedo.

Colubi, A., Gil, M.A. and López-García, H. (1997). Measuring the inequality associated with a fuzzy random variable in terms of statistical f-divergences. *Proc. EUFIT'97* **3**, 1764-1768.

Colubi, A., López-Díaz, M., Domínguez-Menchero, J.S. and Gil, M.A. (1999). A generalized Strong Law of Large Numbers. *Probab. Theory Relat. Fields* (en prensa).

Corral, N., Gil, M.A., López, M.T., Salas, A. and Bertoluzza, C. (1998). Statistical models. In *Handbook of Fuzzy Computation* Section C: Modeling and Simulation C2.3 (E. Ruspini, P. Bonissone and W. Pedrycz, Eds.), C2.3:1 - C2.3:18. Institute of Physics & Oxford University Press, Bristol.

Cowell, F.A. (1980). On the structure of additive inequality measures. *Rev. Econ. Stud.* **47**, 521-531.

Cowell, F.A. and Kuga, K. (1981). Additivity and the entropy concept. An axiomatic approach. *J. Econom. Theory* **25**, 131-143.

Cowell, F.A. and Mehta, F. (1982). The estimation and interpolation of inequality measures. *Rev. Econ. Stud.* **49**, 273-290.

Cox, E. (1994). *The Fuzzy Systems Handbook.* Academic Press, Cambridge.

Csiszár, I. (1967). Information-type measures of difference of probability distributions and indirect observations. *Studia Scient. Math. Hung.* **2**, 299-318.

Dasgupta, P., Starrett, D. and Sen, A. (1973). Notes on the measurement of inequality. *J. Econom. Theory* **6**, 180-187.

Debreu, G. (1967). Integration of correspondences. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* 1965/66 **2**, Part 1. Univ. of California Press, Berkeley, 351-372.

Delgado, M., Verdegay, J.L. and Vila, M.A. (1988). A procedure for ranking fuzzy numbers using fuzzy relations. *Fuzzy Sets and Systems* **26**, 49-62.

Diamond, P. and Tanaka, H. (1998). Fuzzy regression analysis. In *Fuzzy Sets in Decision Analysis, Operations Research and Statistics* Ch. 11 (R. Słowiński, Eds.), 349-387. Kluwer Acad. Pub., Norwell.

Diamond, P. and Kloeden, P. (1990). Metric spaces of fuzzy sets. *Fuzzy Sets and Systems* **35**, 241-249.

Diamond, P. and Kloeden, P. (1994). *Metric Spaces of Fuzzy Sets: Theory and Applications.* World Scientific, Singapore.

Diamond, P. and Körner, R. (1997). Extended fuzzy linear models and least squares estimates. *Comput. Math. Applic.* **33**, 15-32.

Diamond, P. (1987). Least squares fitting of several fuzzy variables. In *Analysis of Fuzzy Information,* (J.C. Bezdek, Eds.), 329-331. CRC Press, Tokyo.

Diamond, P. (1988). Fuzzy least squares *Inform. Sci.* **46**, 141-157.

Diamond, P. (1990). Least Squares Fitting of Compact Set-Valued Data. *J. Math. Anal. Appl.* **147**, 531-544.

Dubois, D. and Prade, H. (1978). Operations on fuzzy numbers. *Int. J. Syst. Sci.* **9**, 613-626.

Dubois, D. and Prade, H. (1980). *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, New York.

Dubois, D. and Prade, H. (1987). The mean value of a fuzzy number. *Fuzzy Sets and Systems* **24**, 279-300.

Dudewicz, E.J. and Mishra, S.N. (1988). *Modern Mathematical Statistics*. John Wiley & Sons, New York.

Ebert, U. (1984). On a family of inequality measures. Discussion paper No. 146. University of Bonn.

Eichhorn, W. and Gehrig, W. (1982). Measurement of inequality in Economics. In *Applied Mathematics-Optimization and Operations Research* (B. Korte, Eds.), 657-693. North-Holland, Amsterdam.

Gastwirth, J.L. (1975). The estimation of a family of measures of income inequality. *Journal of Econometrics* **3**, 61-70.

Gastwirth, J.L., Nayak, T.K. and Krieger, A.M. (1986). Large sample theory for the bounds on the Gini and related indices of inequality estimated from grouped data. *J. Bus. Econ. Statist.* **4**, 269-273.

Gebhardt, J., Gil, M.A. and Kruse, R. (1998). Fuzzy set-theoretic methods in Statistics. In *Fuzzy Sets in Decision Analysis, Operations Research and Statistics* Ch. 10 (R. Słowiński, Eds.), 311-347. Kluwer Acad. Pub., Norwell.

Gil, M.A. (1979). *Incertidumbre y utilidad.* Tesis Doctoral. Universidad de Oviedo.

Gil, M.A. (1981). Estudio de una medida para la incertidumbre correspondiente a las utilidades. *Trab. de Est. e Inv. Oper.* **32**, 45-66.

Gil, M.A. (1982). On two unquietness measures for the finite experiences. *Proc. 2nd. World Conf. on Math. at the Serv. of man*, 299-304.

Gil, M.A. and Gil, P. (1989). On some information measures of degree $\beta = 2$. Estimation in simple-stage cluster sampling. *Statist. Probab. Lett.* **8**, 157-162.

Gil, M.A., Caso, C. and Gil, P. (1989a). Estudio asintótico de una clase de índices de desigualdad muestrales. *Trab. de Est.* **4**, 95-109.

Gil, M.A., Pérez, R. and Gil, P. (1989b). A family of measures of uncertainty involving utilies: definitions, properties, applications and statistical inferences. *Metrika* **36**, 129-147.

Gil, M.A. and López-Díaz, M. (1996). Fundamentals and Bayesian analyses of decision problems with fuzzy-valued utilities. *Int. J. Approx. Reason.* **15**, 203-224.

Gil, M.A., López-Díaz, M. and López-García, H. (1998a). The fuzzy hyperbolic inequality index associated with fuzzy random variables. *European J. Oper. Res.* **110**, 377-391.

Gil, M.A., López-Díaz, M. and Rodríguez-Muñiz, L.J. (1998b). An improvement of a comparison of experiments in statistical decision problems with fuzzy utilities. *IEEE Trans. Systems, Man, Cybern.* **28A**, 856-864.

Gil, M.A., Gil, P. and Ralescu, D.A. (1999). Fuzzy random variables: modeling linguistic statistical data. In *Computing with Words in Information/Intelligent Systems* (J. Kacprzyk and L.A. Zadeh, Eds.), Springer-Verlag, Berlin (en prensa).

González, A. (1990). A study of the ranking function approach through mean values. *Fuzzy Sets and Systems* **35**, 29-41.

González, A. and Vila, M.A. (1992). Dominance relations on fuzzy numbers. *Inform. Sci.* **64**, 1-16.

Guo, S. and Chen, S. (1992). An approach to monodic fuzzy regression. In *Fuzzy Regression Analysis* (J. Kacprzyk and M. Fedrizzi, Eds.), 81-90. Physica-Verlag, Heidelberg.

Kacprzyk, J. and Fedrizzi, M. (1992). *Fuzzy Regression Analysis.* Physica-Verlag, Heidelberg.

Kaleva, O. (1994). Interpolation of fuzzy data. *Fuzzy Sets and Systems* **61**, 63-70.

Kaufmann, A. and Gupta, M.M. (1991). *Introduction to Fuzzy Arithmetic.* Van Nostrand Reinhold Co., New York.

Kendall, D.G. (1974). Foundations of a theory of random sets. In *Stochastic Geometry* (E.F. Harding and D.G. Kendall Eds.), 3-9. John Wiley & Sons, New York.

Hiai, F. & Umegaki, H. (1977). Integrals, conditional expectations and martingales of multivalued functions. *J. Multivar. Anal.* **7**, 149-182.

Himmelberg, C.J. (1975). Measurable relations. *Fund. Math.* **87**, 53-72.

Klement, E.P., Puri M.L. and Ralescu, D.A. (1986). Limit theorems for fuzzy random variables. *Proc. R. Soc, Lond. A* **407**, 171-182.

Klir, G.J. and Yuan, B. (1995). *Fuzzy sets and Fuzzy Logic. Theory and Applications.* Prentice Hall, New Jersey.

Kölm, S. Ch. (1976). Unequal inequalities I. *J. Econom. Theory* **12**, 416-442.

Kölm, S. Ch. (1976). Unequal inequalities II. *J. Econom. Theory* **13**, 82-111.

Kołodziejczyk, W. (1986). Orlovsky's concept of Decision-Making with fuzzy preference relation-further results. *Fuzzy Sets and Systems* **19**, 11-20.

Körner, R. (1997a). On the variance of fuzzy random variables. *Fuzzy Sets and Systems* **92**, 83-93.

Körner, R. (1997b). *Linear models with random fuzzy variables.* PhD Thesis. Freiberg University of Mining and Technology.

Körner, R. and Näther, W. (1998). Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates. *Inform. Sci.* **109**, 95-118.

Körner, R. and Lubiano, M.A. (1998). A generalized real-valued measure of dispersion for fuzzy random variables. Tech. Rep. Universidad de Oviedo.

Kruse, R. (1987). On the variance of random sets. *J. Math. Anal. Appl.* **122**, 469-473.

Kruse, R., Gebhardt, J. and Gil, M.A. (1999). Fuzzy Statistics. In *Wiley Encyclopedia of Electrical and Electronics Engineering* (J. G. Webster, Eds.), John Wiley & Sons, New York (en prensa).

Kruse, R. and Meyer, K.D. (1987). *Statistics with Vague Data*, Reidel Publ. Co., Dordrecht.

Kullback, S. and Leibler, A. (1951). On information and sufficiency. *Ann. Math. Statist.* **22**, 79-86.

Kwakernaak, H. (1978). Fuzzy random variables. Part I: Definitions and theorems. *Inform. Sci.* **15**, 1-29.

Kwakernaak, H. (1979). Fuzzy random variables. Part II: Algorithms and examples for the discrete case. *Inform. Sci.* **17**, 253-278.

Lehmann, E.H. (1983). *Theory of Point Estimation.* John Wiley & Sons, New York.

Loéve, M. (1963). *Probability Theory.* 3rd ed. Van Nostrand, Princeton.

López Díaz, M. (1996). *Medibilidad e integración de variables aleatorias difusas. Aplicación a problemas de decisión.* Tesis Doctoral. Universidad de Oviedo.

López-Díaz, M. and Gil, M.A. (1998a). The λ-average value and the fuzzy expectation of a fuzzy random variable. *Fuzzy Sets and Systems* **99**, 347-352.

López-Díaz, M. and Gil, M.A. (1998b). Reversing the order of integration in iterated expectations of fuzzy random variables, and statistical applications. *J. Stat. Plan. Infer.* **74**, 11-29.

López García, H. (1997). *Cuantificación de la desigualdad asociada a conjuntos aleatorios y variables aleatorias difusas.* Tesis Doctoral. Universidad de Oviedo.

López-García, H. and Corral, N. (1998). Extension of the additively decomposable inequality indices for random sets and random fuzzy sets. *Statistical Papers* (enviado para publicación).

Lowen, R. (1990). A fuzzy Lagrange interpolation theorem. *Fuzzy Sets and Systems* **34**, 33-38.

Lubiano, M.A. and Gil, M.A. (1999). Estimating the expected value of fuzzy random variables in random samplings from finite populations. *Statistical Papers* (en prensa).

Lubiano, M.A., Gil, M.A., López-Díaz, M. and López, M.T. (1999a). The $\vec{\lambda}$-mean squared dispersion associated with a fuzzy random variable. *Fuzzy Sets and Systems* (en prensa).

Lubiano, M.A.,Gil, M.A. and López-Díaz, M. (1999b). On the Rao-Blackwell Theorem for fuzzy random variable. *Kybernetika* (en prensa).

Lubiano, M.A. Salas, A. and Corral, N. (1998). The $\vec{\lambda}$-mean squared dispersion for the fuzzy linear regression. *Proc. IPMU'98* **I**, 101-107.

Mareš, M. (1994). *Computation over Fuzzy Quantities.* CRC Press, Boca Raton.

Marshall, A.W. and Olkin, I. (1979). *Inequalities.* Academic Press, New York.

Martínez López, I. (1991). *Aproximación de algunas familias de índices de variación en poblaciones finitas.* Tesis Doctoral. Universidad de Oviedo.

Matheron, G. (1975). *Random Sets and Integral Geometry.* John Wiley & Sons, New York.

Molchanov, I.S. (1993). *Limit Theorems for Unions of Random Closed Sets.* Lect. Notes Math. **1561**. Springer-Verlag, Berlin.

Montenegro, M., Casals, M.R. and Gil, M.A. (1998). Estimating fuzzy parameters of fuzzy random variables in a Bayesian context. *Proc. IPMU'98* **II**, 1002-1007.

Moskowitz, H. and Kim, K. (1993). On assessing the H value in fuzzy linear regression. *Fuzzy Sets and Systems* **58**, 303-327.

Nakamura, K. (1986). Preference relations on a set of fuzzy utilities as a basis for decision making. *Fuzzy Sets and Systems* **20**, 147-162.

Näther, W. (1988). On expectation and variance of random fuzzy sets, with application to linear regression with vague observations. *BUSEFAL* **36**, 82-91.

Näther, W. and Albrecht, M. (1990). Linear regression with random fuzzy observations. *Statistics* **21**, 521-531.

Näther, W. (1997). Linear statistical inference for random fuzzy data. *Statistics* **29**, 221-240.

Norwich, A.M. and Turksen, I.B. (1984). A model for the measurement of membership and the consequences of its empirical implementation. *Fuzzy Sets and Systems* **12**, 1-25.

Negoita, C. and Ralescu, D.A. (1987). *Simulation, Knowledge-based Computing, and Fuzzy Statistics*. Van Nostrand Reinhold Co., New York.

Nguyen, H.T. (1978). A note on the extension principle for fuzzy sets. *J. Math. Anal. Appl.* **64**, 369-380.

Ok, E.A. (1994). On the approximation of fuzzy preferences by exact relations. *Fuzzy Sets and Systems* **67**, 173-179.

Ok, E.A. (1995). Fuzzy income inequality measurement: A class of fuzzy inequality measures. *Social Choice and Welfare* **12**, 111-136.

Ok, E.A. (1996). Fuzzy measurement of income inequality: Some possibility results on the fuzzification of the Lorenz ordering. *Econ. Theory* **7**, 513-530.

Okuda, T., Tanaka, H. and Asai, K. (1978). A formulation of fuzzy decision problems with fuzzy information, using probability measures of fuzzy events. *Inform. Contr.* **38**, 135-147.

Puri, M.L. and Ralescu, D. (1981). Différentielle d'une fonction floue. *C.R. Acad. Sci. Paris, Sér. I* **293**, 237-239.

Puri, M.L. and Ralescu, D. (1983). Differentials of fuzzy functions. *J. Math. Annal. Appl.* **91**, 552-558.

Puri, M.L. and Ralescu, D.A. (1985) The concept of normality for fuzzy random variables. *Ann. Probab.* **13**, 1373-1379.

Puri, M.L. and Ralescu, D. (1986). Fuzzy random variables. *J. Math. Anal. Appl.* **114**, 409-422.

Raj, D. and Khamis, S.H. (1958). Some remarks on sampling with replacement. *Ann. Math. Stat.* **29**, 550-557.

Ralescu, A. and Ralescu, D.A. (1984). Probability and fuzziness. *Inf. Sci.* **17**, 85-92.

Ralescu, A. and Ralescu, D.A. (1986). Fuzzy sets in statistical inference. En *The Mathematics of Fuzzy Systems.* (A. Di Nola and A.G.S. Ventre, Eds.), 273-283. Verlag TÜV Rheinland, Köln.

Ralescu, D.A. (1982). Fuzzy logic and statistical estimation. *Proc. 2nd World Conf. on Math. at the Serv. of Man*, 605-606.

Ralescu, D.A. (1995a). Fuzzy probabilities and their applications to statistical inference. En *Lecture Notes in Computer Science* **945**, Advances in Intelligent Computing (B. Bouchon and R. R. Yager, Eds.), 217-222. Springer-Verlag, Berlin.

Ralescu, D.A. (1995b). Inequalities for fuzzy random variables. *Proc. 26th Iranian Mathematical Conference*, 333-335.

Ralescu, D.A. (1995c). Fuzzy random variables revisited. *Proc. IFES'95 and Fuzzy IEEE Joint Conference* **2**, 993-1000.

Ramík, J. and Římánek, J. (1985). Inequality relation between fuzzy numbers and its use in fuzzy optimization. *Fuzzy Sets and Systems* **16**, 123-138.

Rao, C.R. (1973). *Linear Statistical Inference and Its Applications.* John Wiley & Sons, New York.

Rathie, P.N. (1971). On some new measures of uncertainty, inaccuracy and information and their characterizations. *Kybernetika* **7**, 394-403.

Redden, D.T. and Woodall, W.H. (1994). Properties of certain fuzzy linear regression methods. *Fuzzy Sets and Systems* **64**, 361-375.

Redden, D.T. and Woodall, W.H. (1996). Further examination of fuzzy linear regression. *Fuzzy Sets and Systems* **79**, 203-211.

Rudin, W. (1979). *Análisis Real y Complejo.* Alhambra, Madrid.

Salas, A. (1991). *Regresión lineal con observaciones difusas.* Tesis Doctoral. Universidad de Oviedo.

Sánchez de Posada Martínez, C. (1998). *Estudio matemático de los números difusos.* Trabajo Integrado y Académicamente Dirigido del Plan de Estudios de la Licenciatura de Matemáticas. Universidad de Oviedo.

Sancho Royo, A. (1997). *Evaluaciones de estímulos con características difusas.* Tesis Doctoral. Universidad de Granada.

Savic, A.A. and Pedrycz, W. (1991). Evaluation of fuzzy linear regression models. *Fuzzy Sets and Systems* **39**, 51-63.

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* John Wiley & Sons, New York.

Shorrocks, A.F. (1980). The class of additively decomposable inequality measures. *Econometrica* **48**, 613-625.

Stojaković, M. (1994). Fuzzy random variables, expectation and martingales. *J. Math. Anal. Appl.* **184**, 594-606.

Stoyan, D. (1998). Random sets: Models and Statistics. *Int. Stat. Rev.* **66**, 1-27.

Stoyan, D., Kendall, W.S. and Mecke, J. (1995). *Stochastic Geometry and its Applications.* 2nd ed. John Wiley & Sons, Chichester.

Tanaka, H., Uejima, S. and Asai, K. (1980). Fuzzy linear regression model. *Proc. Int. Congress on Applied Systems Research and Cybernetics* **VI**, 2933-2938.

Tanaka, H., Uejima, S. and Asai, K. (1982). Linear regression analysis with fuzzy model. *IEEE Trans. Syst. Man. Cyber.* **12**, 903-907.

Theil, H. (1967). *Economics and Information.* North-Holland, Amsterdam.

Thompson, S.K. (1992). *Sampling.* John Wiley & Sons, New York.

Tseng, T.V. and Klein, C.M. (1989). New algorithm for the ranking procedure in fuzzy decision-making. *IEEE Trans. Syst. Man and Cybernet.* **19**, 1289-1296.

Yager, R.R. (1981) A procedure for ordering fuzzy subsets of the unit interval. *Inf. Sci.* **24**, 143-161.

Zacks, S. (1971). *The Theory of Statistical Inference.* John Wiley, New York.

Zadeh, L.A. (1965). Fuzzy sets. *Inform. Contr.* **8**, 338-353.

Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning. *Inform. Sci.*, Part 1 **8**, 199-249; Part 2 **8**, 301-353; Part 3 **9**, 43-80.

Zadeh, L.A. (1976). A fuzzy-algorithmic approach to the definition of complex or imprecise concepts. *Int. J. Man-Mach. Stud.* **8**, 249-291.

Zagier, D. (1983). On the decomposability of the Gini coefficient and other indices of inequality. Discussion paper No. 108. University of Bonn.

Lista de Símbolos

$\mathcal{K}_c(\mathbb{R})$	clase de los intervalos compactos no vacíos de $\mathbb{R}$
$\mathcal{K}_c((0,+\infty))$	clase de los intervalos compactos no vacíos de $(0,+\infty)$
$\mathcal{F}_c(\mathbb{R})$	clase de los subconjuntos difusos de $\mathbb{R}$ con α-cortes pertenecientes a $\mathcal{K}_c(\mathbb{R})$ (números difusos)
$\mathcal{F}_c((0,+\infty))$	clase de los números difusos de $(0,+\infty)$
$\mathcal{K}_c(\mathbb{R}^d)$	clase de los conjuntos compactos convexos no vacíos de $\mathbb{R}^d$
$\mathcal{F}_c(\mathbb{R}^d)$	clase de los subconjuntos difusos de $\mathbb{R}^d$ con α-cortes y 0-corte pertenecientes a $\mathcal{K}_c(\mathbb{R}^d)$
K	elemento genérico de $\mathcal{K}_c(\mathbb{R})$
$\widetilde{V}$, $\widetilde{A}$	elemento genérico de $\mathcal{F}_c(\mathbb{R})$
$\widetilde{V}_\alpha$	α-corte de $\widetilde{V}$
$\widetilde{V}_0$	envolvente convexa cerrada del soporte de $\widetilde{V}$ (o 0-corte)
$f_K(\lambda)$	combinación lineal convexa de los extremos de K dada por $\lambda \sup K + (1-\lambda) \inf K$
$f_{\widetilde{V}}(\alpha,\lambda)$	combinación lineal convexa de los extremos de $\widetilde{V}_\alpha$ dada por $\lambda \sup \widetilde{V}_\alpha + (1-\lambda) \inf \widetilde{V}_\alpha$
$\oplus$, $\widetilde{\sum}$	suma difusa
$\ominus$	sustracción difusa
$\odot$	producto difuso por escalar

$\otimes$	producto difuso
$\oslash$	cociente difuso
$\succeq_S$	criterio de ordenación de Ramík & Římánek's sobre $\mathcal{F}_c(\mathbb{R})$
$\geq_Y$	criterio de ordenación de Yager sobre $\mathcal{F}_c(\mathbb{R})$
F	función de ordenación en el criterio de Yager
m	medida de Lebesgue
d_H	distancia de Hausdorff sobre $\mathcal{K}_c(\mathbb{R})$ (ver Índice de Materias)
d_∞	distancia generalizada de Hausdorff sobre $\mathcal{K}_c(\mathbb{R})$ (ver Índice de Materias)
δ_p	distancia δ_p (ver Índice de Materias)
d_2	distancia d_2 (ver Índice de Materias)
D_S	S- distancia (ver Índice de Materias)
D_m	(ver S-distancia en Índice de Materias)
$S_{\vec{\lambda}}$	(ver S-distancia en Índice de Materias)
S_2	(see S-distancia en Índice de Materias)
$s_{\widetilde{V}}$	función soporte de $\widetilde{V}$
$\mathcal{B}_{\mathbb{R}}$	σ-álgebra de Borel sobre $\mathbb{R}$
$\mathcal{B}_{[0,1]}$	σ-álgebra de Borel sobre $[0,1]$
$\mathcal{M}_{[0,1]}$	σ-álgebra de Lebesgue sobre $[0,1]$
$\mathcal{B}_{d_H}$	σ-álgebra engendrada por la topología asociada a d_H sobre $\mathcal{K}_c(\mathbb{R})$
$\mathcal{B}_{d_\infty}$	σ-álgebra engendrada por la topología asociada a d_∞ sobre $\mathcal{F}_c(\mathbb{R})$
$\mathcal{P}(\Omega)$	partes de Ω
X	conjunto aleatorio compacto y convexo
$\mathcal{X}$	variable aleatoria difusa
$\mathcal{X}_\alpha$	función α-corte (ver Índice de Materias)
$E(X)$ ó $E(X\|P)$	valor esperado (de Aumann) de X con respecto a P
$\overline{X}$	notación alternativa del valor esperado (de Aumann) de X sobre una población finita
$\widetilde{E}(\mathcal{X})$ ó $\widetilde{E}(\mathcal{X}\|P)$	valor esperado (de Puri & Ralescu) de $\mathcal{X}$ con respecto a P

$\overline{\mathcal{X}}$	notación alternativa del valor esperado (de Puri & Ralescu) de $\mathcal{X}$ sobre una población finita
$\delta(\,\cdot\,)$	distribución de Dirac (ver Índice de Materias)
δ_{ij}	delta de Kronecker
$\mathbf{1}_{\{x\}}$	función indicadora de $x \in \mathbb{R}$
$a\,]$	menor entero mayor o igual que $a \in \mathbb{R}$
$DCM_S(X, K)$	S-dispersión cuadrática media de X con respecto a K
$DCM_S(\mathcal{X}, \widetilde{V})$	S-dispersión cuadrática media de $\mathcal{X}$ con respecto a $\widetilde{V}$
$\Delta^2_S(\,\cdot\,)$ ó $\Delta^2_S(\,\cdot\,\vert P)$	central S-dispersión cuadrática media central de una v.a.d. (c.a.c.c.)
υ	muestra aleatoria genérica de tamaño n
$\overline{X}_n$	valor esperado muestral de X
$\overline{\mathcal{X}}_n$	valor esperado muestral de $\mathcal{X}$
$\Delta^2_S(X[\upsilon])$	S-dispersión cuadrática media central de X en υ
$\Delta^2_S(\mathcal{X}[\upsilon])$	S-dispersión cuadrática media central de $\mathcal{X}$ en υ
$\mathrm{I}^{\alpha}(\,\cdot\,)$ ó $\mathrm{I}^{\alpha}(\,\cdot\,\vert P)$	índice aditivamente descomponible de orden α asociado a una v.a. (v.a.d., c.a.c.c.)
$\mathrm{I}_{\mathrm{NVar}}(\,\cdot\,)$ ó $\mathrm{I}_{\mathrm{NVar}}(\,\cdot\,\vert P)$	índice aditivamente descomponible de orden 2 asociado a una v.a. (v.a.d., c.a.c.c.)
$\mathrm{I}_{\phi}(\,\cdot\,)$ ó $\mathrm{I}_{\phi}(\,\cdot\,\vert P)$	ϕ-índice asociado a una v.a. (v.a.d., c.a.c.c.)
$\mathrm{I}_{H}(\,\cdot\,)$ ó $\mathrm{I}_{H}(\,\cdot\,\vert P)$	índice hiperbólico asociado a una v.a. (v.a.d., c.a.c.c.)
$\mathrm{I}_{H}(\{\,\cdot\,\})$	índice hiperbólico *intra*valores asociado a una v.a.d. (c.a.c.c.)
$\mathrm{I}_{Sh}(\,\cdot\,)$ ó $\mathrm{I}_{Sh}(\,\cdot\,\vert P)$	índice de tipo Shannon asociado a una v.a. (v.a.d., c.a.c.c.)
$\mathrm{I}_{T}(\,\cdot\,)$ ó $\mathrm{I}_{T}(\,\cdot\,\vert P)$	índice de Theil asociado a una v.a. (v.a.d., c.a.c.c.)
$\mathrm{I}_{f}(\,\cdot\,)$ ó $\mathrm{I}_{f}(\,\cdot\,\vert P)$	f-índice asociado a una v.a. (v.a.d., c.a.c.c.)
z_{α}	cuantil de orden $1-\alpha/2$ de la distribución normal estandar
Φ	función de distribución de la distribución normal estandar
$\mathrm{Tri}(\alpha,\beta,\gamma)$	número difuso triangular
$\mathrm{Tra}(\alpha,\beta,\gamma,\delta)$	número difuso trapezoidal
$S(a,b)$	S-curva (ver Índice de Materias)
$\Pi(a,b,c)$	Π-curva (ver Índice de Materias)

UNIVERSIDAD DE OVIEDO
Departamento de Estadística, I.O y D.M

Variation measures for imprecise random elements

PhD Thesis

M. ASUNCIÓN LUBIANO

To my family

Acknowledgements

I would like, most sincerely to express my thanks to the supervisors of this PhD Thesis, María Angeles Gil Alvarez, Miguel López Díaz and María Teresa López García, who have afforded me their unstinted and constant support while I have been working under their guidance.

I want also to thank all my colleagues in the Statistics and Operations Research group in Oviedo, since in many different ways they have helped me to be here, and my special thanks to our Chairman, Professor Pedro Gil.

There have been many people who have have offered me their advice in developing the present work. Among them, I must mention Professors Carlo Bertoluzza (from the University of Pavia, Italy) and María del Carmen Alonso, Teófilo Brezmes, Hortensia López and Antonia Salas from the University of Oviedo for their contributions in some parts of the work. I am also grateful to my colleagues Luis José Rodríguez and Manuel Montenegro for meticulously proof-reading this dissertation.

I must mention Professor Norberto Corral separately, since he has inspired several ideas in this work.

I wish also to express my deep gratitude to Professor Wolfgang Näther for having invited me to visit his research group in the University of Freiberg (Germany) in March 1998. This opportunity was of great value to this Thesis, not only due to Professor Näther's expertise on the subject and the fruitful discussions I was able to have with him and Dr. Ralf Körner, but also due to their suggestions and those of Professor Dietrich Stoyan during my stay in Freiberg.

The real-life data which illustrate some of the studies in the thesis have been supplied, on one hand, by Yolanda Colubi, Orencio Hernández and Teresa Sánchez from the Servicio de Conservación de la Naturaleza, Dirección Regional de Montes y Medio Ambiente of the Consejería de Agricultura del Principado de Asturias and, on the other hand, by Aurora Fonseca, Head of the Servicio de Nefrología, and Miguel de Zárraga, General Medical Manager, from the Hospital Valle del Nalón de Langreo in Asturias. I want to thank them sincerely for their collaboration and support.

I wish also to thank Tomás Aranda and Ana Colubi for their friendship, and for their personal patience and constant scientific support.

I wish to say that throughout these recent years I have been a member of the research group in the Grant FICYT N. 37/PB-TIC97-02 from the Consejería de Cultura of the Principado de Asturias, whose financial support is gratefully acknowledged.

Finally, I want to especially thank my friends, my parents and brothers for their love and for their trust in me which never wavered.

Preface

One of the fundamental aims in Statistics is describing a set of observations in terms of a few measures summarizing this set. Among the main summary measures one can observe the location ones (and especially the expected value), and the variation (or dispersion) ones.

When there is no variation of observations, the statistical methodology has no real interest. Furthermore, the quantification of the variation makes greater sense when we wish to compare populations, samples, variables, estimators, etc.

The variation associated with the random magnitude providing us with the observations, can be quantified by means of either absolute or relative measures. The first ones usually measure the variation in the units (or the squared units) of the magnitude, and they try to achieve an idea of the extent to which the location measures (or more generally, certain reference values) represent the values of the magniude. The measures of the relative variation (more precisely, the inequality measures) are usually dimensionless indices that try to achieve an idea of the extent to which the location measures (or referential values) is above or below the values of the considered magnitude.

The most common measures of the absolute variation are defined in terms of deviations or distances between certain values (usually, between the values of the magnitude and the referential value), so that they are location invariant. Consequently, they are especially appropriate to handle real-valued random variables which are measured on an interval scale.

The most common measures of the relative variation are defined in terms of the ratios between certain values, so that they are scale invariant. Consequently, they are especially suitable to deal with real-valued random variables which are measured in a ratio scale.

In this work we will present the extension of some of the classical variation measures for random elements with imprecise values of a certain type, namely, fuzzy and interval values.

The model employed to characterize random elements taking on fuzzy values is that of fuzzy random variables (also called random fuzzy sets) in the sense formalized by Puri & Ralescu (1986), whereas the model for random elements taking on interval values is that corresponding to convex compact random sets (see Matheron 1975).

The measures of the (absolute and relative) variation introduced in the present work are assumed to take on real values. Nevertheless, some authors have emphasized that the variation (and, in particular, the inequality) is an intrinsically imprecise characteristic, even when it is quantified for real-valued magnitudes. However, since throughout the work we suppose that the main aim of the variation measures is serving as the basis to compare populations, magnitudes, etc., we have chosen to define real-valued indices allowing a direct comparison.

The work starts with an introductory chapter in which the fundamental concepts and results for the studies developed in this work are recalled.

The core of the work is split into three chapters.

Chapter 1 is devoted to the introduction of a generalized measure of the absolute variation (deviation) of a fuzzy random variable. Then, several interesting properties of this measure are examined, and several inferences on the value of the measure in a population are carried out on the basis of samples drawn from it. Later, we analyze two statistical applications of the measure introduced above: the quantification of the sampling error associated with the estimation of the expected value of a fuzzy random variable in the sampling from finite populations, and the quantification of the error associated with either a linear or a general functional relation between two fuzzy random variables, to determine the optimal relations.

In Chapter 2 a generalized measure of the relative variation inequality of a fuzzy random variable is introduced. Properties of this measure are examined to develop two types of statistical inferences: the estimation of a particular measure in random samplings from finite populations, and the study of the asymptotic distribution in finite populations (and under quite general conditions) of the generalized measure associated with a random sample from a fuzzy random variable.

In Chapter 3 the concepts and results in Chapters 1 and 2 are particularized to the special case in which the fuzzy random variable reduces to a convex compact random set.

Each of the three chapters is endowed with a final discussion on the contributions in it and some related open problems, and encloses several illustrative examples.

In the epilogue of the work some general aspects are briefly commented, and a short argument on the considered methodology is concluded.

Finally, some results are gathered together in two appendices concerning on one hand the metric structure of the space on which the generalization of the absolute variation measure is based, and on the other hand, the characterization of the functions on which the generalization of the relative measure is based.

Contents

List of Figures

List of Tables

Preliminaries and foundations

In this chapter some basic concepts and results from the Fuzzy Set and Random Set Theories, along with some other supporting concepts and results on the quantification of the inequality associated with a random variable, are recalled and gathered.

0.1 Concepts and results on fuzzy sets

Throughout this work, the involved experimental data are assumed to be imprecise. Models dealing with this imprecision will be, on one hand, the real intervals, and on the other hand, certain fuzzy sets of the space of real numbers, $\mathbb{R}$.

The fuzzy subsets of $\mathbb{R}$ we will handle in the present work satisfy the conditions indicated in the following definition.

Definition 0.1.1. *$\mathcal{F}_c(\mathbb{R})$ denotes the class of fuzzy subsets of $\mathbb{R}$, $\widetilde{V} : \mathbb{R} \to [0,1]$, satisfying that*

i) the α-level set of $\widetilde{V}$, $\widetilde{V}_\alpha = \{x \in \mathbb{R} \,|\, \widetilde{V}(x) \geq \alpha\}$, is compact for all $\alpha \in (0,1]$,

ii) $\widetilde{V}_1 = \{x \in \mathbb{R} \,|\, \widetilde{V}(x) = 1\} \neq \emptyset$ (i.e., $\widetilde{V}$ is normal), *iii)* $\widetilde{V}$ *is a convex fuzzy subset, that is, for any* $\alpha \in (0,1]$ *the* α*-level set* $\widetilde{V}_\alpha$ *is a convex subset of* $\mathbb{R}$*,*

iv) *the closed convex hull of the support of* $\widetilde{V}$ *(where* $\operatorname{supp} \widetilde{V} = \{x \in \mathbb{R} \,|\, \widetilde{V}(x) > 0\}$*), which in this case coincides with the closure of* $\operatorname{supp} \widetilde{V}$ *and is denoted by* $\widetilde{V}_0$*, is compact.*

Obviously, $\mathcal{F}_c(\mathbb{R})$ can be briefly described as the class of fuzzy subsets $\widetilde{V}$ of $\mathbb{R}$ such that $\widetilde{V}_\alpha \in \mathcal{K}_c(\mathbb{R})$ for all $\alpha \in [0,1]$, with $\mathcal{K}_c(\mathbb{R})$ being the class of nonempty compact intervals contained in $\mathbb{R}$.

From now on, we will refer generically to the elements of $\mathcal{F}_c(\mathbb{R})$ as *fuzzy numbers.*

0.1.1 Operations between fuzzy numbers

The statistical management of imprecise data with interval or fuzzy values usually requires considering elementary operations between such values (especially the sum and the product by a real number). The arithmetic of intervals is well-known and the result of operating two intervals is defined as the image of the Cartesian product of these intervals by the mappings associated with the corresponding operations.

Operations between fuzzy subsets (and, in particular between fuzzy numbers) are defined on the basis of *Zadeh's extension principle* (Zadeh 1975) as follows:

Extension principle Let $f : \mathbb{R}^k \to \mathbb{R}$, and let $\widetilde{V}_i$ be fuzzy subsets of $\mathbb{R}$, for $i \in \{1,\ldots,k\}$. The (fuzzy) *image of* $\left(\widetilde{V}_1, \ldots, \widetilde{V}_k\right)$ *induced* by f is defined as the fuzzy subset of $\mathbb{R}$, $\widetilde{W} = f\left(\widetilde{V}_1, \ldots, \widetilde{V}_k\right)$, given by

$$\widetilde{W}(y) = f\left(\widetilde{V}_1, \ldots, \widetilde{V}_k\right)(y)$$

$$= \begin{cases} \displaystyle\sup_{\substack{(x_1,\ldots,x_k)| \\ f(x_1,\ldots,x_k)=y}} \min\{\widetilde{V}_1(x_1), \ldots, \widetilde{V}_k(x_k)\} & \text{if } f^{-1}(\{y\}) \neq \emptyset \\ 0 & \text{if } f^{-1}(\{y\}) = \emptyset. \end{cases}$$

Remark 0.1.1. On the basis of some results from Nguyen (1978), one can prove (see, for instance, López Díaz 1996) that if f is either an injective or a continuous function, then for all $\alpha \in (0,1]$ we have that

$$\left(f\left(\widetilde{V}_1, \ldots, \widetilde{V}_r\right)\right)_\alpha = f\left(\left(\widetilde{V_1}\right)_\alpha, \ldots, \left(\widetilde{V_r}\right)_\alpha\right).$$

(for a detailed proof of this result see, for instance, Sánchez de Posada Martínez 1998).

When Zadeh's extension principle is employed to establish the algebraic operations between fuzzy subsets (see, for instance, Dubois & Prade 1978, 1987, Kaufmann & Gupta 1991, Mareš 1994), we obtain the notions of fuzzy sum $\oplus$ (or, alternatively, $\widetilde{\sum}$), fuzzy substraction $\ominus$, fuzzy product by a real number $\odot$, fuzzy product $\otimes$, and fuzzy quotient $\oslash$, which for $\widetilde{V}$ and $\widetilde{W} \in \mathcal{F}_c(\mathrm{I\!R})$ ($\widetilde{V}$ and $\widetilde{W}$ assumed to belong to $\mathcal{F}_c((0,+\infty))$ -i.e, $\widetilde{V}$, $\widetilde{W}$ $\in \mathcal{F}_c(\mathrm{I\!R})$ and $\widetilde{V}_0$, $\widetilde{W}_0 \in \mathcal{P}((0,+\infty))$- for the last two operations) are defined as follows:

$$\left(\widetilde{V} \oplus \widetilde{W}\right)(z) = \sup_{(x,y)|x+y=z} \min\{\widetilde{V}(x), \widetilde{W}(y)\},$$

$$\left(\widetilde{V} \ominus \widetilde{W}\right)(z) = \sup_{(x,y)|x-y=z} \min\{\widetilde{V}(x), \widetilde{W}(y)\},$$

$$\left(\lambda \odot \widetilde{V}\right)(z) = \begin{cases} \widetilde{V}(z/\lambda) & \text{if } \lambda \neq 0, \\ \mathbf{1}_{\{0\}}(z) & \text{if } \lambda = 0, \end{cases}$$

$$\left(\widetilde{V} \otimes \widetilde{W}\right)(z) = \sup_{(x,y)|x\times y=z} \min\{\widetilde{V}(x), \widetilde{W}(y)\},$$

$$\left(\widetilde{V} \oslash \widetilde{W}\right)(z) = \sup_{(x,y)|x/y=z} \min\{\widetilde{V}(x), \widetilde{W}(y)\}.$$

Remark 0.1.2. In virtue of the conclusions drawn from Nguyen's results (1978), the arithmetic of fuzzy numbers reduces to the interval arithmetic. More precisely, for $\alpha \in [0,1]$ (the case $\alpha = 0$ can be found in López García 1997), the α-level sets of the operations of fuzzy numbers can be expressed in terms of operations of certain elements in $\mathcal{K}_c(\mathbb{R})$: the α-level sets of the involved elements of $\mathcal{F}_c(\mathbb{R})$. In this way, if $\widetilde{V}$ and $\widetilde{W}$ are in $\mathcal{F}_c(\mathbb{R})$, for all $\alpha \in [0,1]$ we have that $\left(\widetilde{V} \oplus \widetilde{W}\right)_\alpha$ = Minkowski sum of $\widetilde{V}_\alpha$ and $\widetilde{W}_\alpha$, that is,

$$\left(\widetilde{V} \oplus \widetilde{W}\right)_\alpha = \left[\inf \widetilde{V}_\alpha + \inf \widetilde{W}_\alpha, \sup \widetilde{V}_\alpha + \sup \widetilde{W}_\alpha\right],$$

$\left(\widetilde{V} \ominus \widetilde{W}\right)_\alpha$ = Minkowski substraction of $\widetilde{V}_\alpha$ and $\widetilde{W}_\alpha$, that is,

$$\left(\widetilde{V} \ominus \widetilde{W}\right)_\alpha = \left[\inf \widetilde{V}_\alpha - \sup \widetilde{W}_\alpha, \sup \widetilde{V}_\alpha - \inf \widetilde{W}_\alpha\right],$$

and if $\lambda \in \mathbb{R} \setminus \{0\}$, we have that $\left(\lambda \odot \widetilde{V}\right)_\alpha = \lambda \widetilde{V}_\alpha$, that is,

$$\left(\lambda \odot \widetilde{W}\right)_\alpha = \left[\lambda \cdot \inf \widetilde{W}_\alpha, \lambda \cdot \sup \widetilde{W}_\alpha\right], \text{ if } \lambda > 0,$$

$$\left(\lambda \odot \widetilde{W}\right)_\alpha = \left[\lambda \cdot \sup \widetilde{W}_\alpha, \lambda \cdot \inf \widetilde{W}_\alpha\right], \text{ if } \lambda < 0.$$

If $\widetilde{V}$ and $\widetilde{W}$ are in $\mathcal{F}_c((0,+\infty))$ it holds for all $\alpha \in [0,1]$ that $\left(\widetilde{V} \otimes \widetilde{W}\right)_\alpha = \widetilde{V}_\alpha \cdot \widetilde{W}_\alpha$, that is,

$$\left(\widetilde{V} \otimes \widetilde{W}\right)_\alpha = \left[\inf \widetilde{V}_\alpha \cdot \inf \widetilde{W}_\alpha, \sup \widetilde{V}_\alpha \cdot \sup \widetilde{W}_\alpha\right],$$

and $\left(\widetilde{V} \oslash \widetilde{W}\right)_\alpha = \widetilde{V}_\alpha / \widetilde{W}_\alpha$, that is,

$$\left(\widetilde{V} \oslash \widetilde{W}\right)_\alpha = \left[\inf \widetilde{V}_\alpha / \sup \widetilde{W}_\alpha, \sup \widetilde{V}_\alpha / \inf \widetilde{W}_\alpha\right].$$

On the basis of the fuzzy sum and product by a real number above, it is posible to establish in an analogous way to the real-valued case, the matricial operations, when the elements involved in the matrices are fuzzy numbers (see Dubois & Prade 1980).

Some studies and results on the arithmetic of fuzzy numbers can be found, for instance, in Kaufmann & Gupta (1991), Mareš (1994) and in the recent review by Sánchez de Posada Martínez (1998).

0.1.2 Rankings of fuzzy numbers and compact intervals

When we develop a statistical study with imprecise data, and especially whenever our aim is to compare populations, variables, etc., we have often to order imprecise values. There are several criteria to ranking fuzzy numbers and real intervals (see, for instance, Adamo 1980, Yager 1981, Bortolan & Degani 1985, Ramík & Římanek 1985, Kołodziejczyk 1986, Nakamura 1986, Delgado *et al.* 1988, Campos & González 1989, Tseng & Klein 1989, González & Vila 1992, and also Sánchez de Posada Martínez 1998 for a recent review).

Some of these criteria are based on crisp relations, that is, they determine a categorical ordering (or preordering) on the class $\mathcal{F}_c(\mathbb{R})$, and other ones are based on fuzzy relations indicating the degree of truth of the assertion that a given fuzzy number is greater than or equal to another given one. Among the criteria based on crisp relations we can distinguish two groups: the group of the criteria leading to a total ordering (or preordering), and the group of the criteria leading to a partial ordering (or preordering).

In Chapter 2 of this work we will make use of two different criteria, which we now present.

Ramík & Římánek's criterion

The application of this criterion leads to a partial ordering on $\mathcal{F}_c(\mathbb{R})$ which is universally accepted, and is stated (Ramík & Římanek 1985) as follows:

Definition 0.1.2. *If* $\widetilde{V}, \widetilde{W} \in \mathcal{F}_c(\mathbb{R})$*, then* $\widetilde{V} \succeq_S \widetilde{W}$ *if, and only if,*

$$\inf \widetilde{V}_\alpha \geq \inf \widetilde{W}_\alpha \text{ and } \sup \widetilde{V}_\alpha \geq \sup \widetilde{W}_\alpha,$$

for all $\alpha \in [0,1]$.

The relation $\succeq_S$ can be viewed as the one generated by the lattice operations $\widetilde{\max}$ and $\widetilde{\min}$ (both assumed to be based on Zadeh's extension principles) on $\mathcal{F}_c(\mathbb{R})$, so that $\widetilde{V} \succeq_S \widetilde{W}$ if, and only if, $\widetilde{\max}\{\widetilde{V}, \widetilde{W}\} = \widetilde{V}$ and $\widetilde{\min}\{\widetilde{V}, \widetilde{W}\} = \widetilde{W}$. The relation $\succeq_S$ can also be viewed as the extension of the strong dominance criterion by González & Vila (1992) for an ordering system which is dense on $[0, 1]$ (see López García 1997).

The inconveniences of employing this criterion in practice are due to the fact that there are many pairs of fuzzy numbers this criterion cannot compare. However, it leads to an ordering since $\widetilde{V} \succeq_S \widetilde{W}$ and $\widetilde{W} \succeq_S \widetilde{V}$ (that will be denoted by $\widetilde{V} \sim_S \widetilde{W}$) is equivalent to $\widetilde{V} = \widetilde{W}$.

The following criterion avoid the situations which arise when fuzzy numbers cannot be compared, and it agrees with Ramík & Římánek's criterion to pairs for which it is applicable.

Yager's criterion

This criterion is based on one of the ranking functions introduced by Yager (1981) which is stated as follows:

Definition 0.1.3. *If $\widetilde{V}, \widetilde{W} \in \mathcal{F}_c(\mathbb{R})$, $\widetilde{V}$ is said to be preferred or indifferent to $\widetilde{W}$ in accordance with Yager's ranking criterion, and it will be denoted by $\widetilde{V} \geq_Y \widetilde{W}$, if, and only if, $F(\widetilde{V}) \geq F(\widetilde{W})$, where F is the ranking function defined for any $\widetilde{V} \in \mathcal{F}_c(\mathbb{R})$ by*

$$F(\widetilde{V}) = \frac{1}{2} \int_{[0,1]} \left[\sup \widetilde{V}_\alpha + \inf \widetilde{V}_\alpha \right] d\alpha.$$

The ranking function F has an interesting interpretation (see González 1990) in terms of the mean value of a fuzzy number $\widetilde{V}$ as defined by Dubois & Prade (1987) which is the interval $E(\widetilde{V}) = \left[E_*(\widetilde{V}), E^*(\widetilde{V}) \right]$, with $E_*(\widetilde{V}) = \int_{[0,1]} \inf \widetilde{V}_\alpha \, d\alpha$ and $E^*(\widetilde{V}) = \int_{[0,1]} \sup \widetilde{V}_\alpha \, d\alpha$ (which are intended as the infimum and supremum of the expected values of $\widetilde{V}$, respectively). Then, for each $\widetilde{V} \in \mathcal{F}_c(\mathbb{R})$, we have that

$$F(\widetilde{V}) = \frac{E^*(\widetilde{V}) + E_*(\widetilde{V})}{2}.$$

Yager's criterion determines a total preordering on $\mathcal{F}_c(\mathbb{R})$ and it is a special case of the parameterized criterion introduced by Campos & González (1989).

0.1.3 Distances between fuzzy numbers

There are several metrics which can be defined on the class $\mathcal{F}_c(\mathbb{R})$. In Diamond & Kloeden (1994) we can find a broad review on many of these metrics.

In this section we only recall those we will use throughout the work.

Metric d_∞

The first fundamental metric in this work is that which is sometimes referred to as the "generalized Hausdorff metric", introduced by Puri & Ralescu (1981,1983). This metric is the basis for the measurability condition for $\mathcal{F}_c(\mathbb{R})$-valued random elements.

Definition 0.1.4. *If $\widetilde{V}$, $\widetilde{W} \in \mathcal{F}_c(\mathbb{R})$, the "generalized Hausdorff metric" between $\widetilde{V}$ and $\widetilde{W}$, $d_\infty(\widetilde{V}, \widetilde{W})$, is given by*

$$d_\infty(\widetilde{V}, \widetilde{W}) = \sup_{\alpha \in [0,1]} d_H(\widetilde{V}_\alpha, \widetilde{W}_\alpha),$$

where d_H denotes the well-known Hausdorff metric on $\mathcal{K}_c(\mathbb{R})$, which for K, $K' \in \mathcal{K}_c(\mathbb{R})$ is given by

$$d_H(K, K') = \max\left\{\sup_{a \in K} \inf_{b \in K'} |a - b|, \sup_{b \in K'} \inf_{a \in K} |a - b|\right\},$$

and, because of the convexity of elements in $\mathcal{K}_c(\mathbb{R})$, it can be alternatively defined by

$$d_H(K, K') = \max \{ |\inf K - \inf K'| , |\sup K - \sup K'| \}.$$

Klement *et al.* (1986) have proven that $(\mathcal{F}_c(\mathbb{R}), d_\infty)$ is a nonseparable metric space.

Although the use of the metric d_∞ in formalizing the measurability condition for $\mathcal{F}_c(\mathbb{R})$-valued random elements provide them with valuable properties, d_∞ would not be the most operational and suitable metric on $\mathcal{F}_c(\mathbb{R})$ if we adopt it to measure the error in "estimating" certain characteristics.

Metric D_S

For the purpose of this measurement, Bertoluzza *et al.* (1995a) have suggested the metric d_S, based on a previously specified measure S associated with the measurable space $([0,1], \mathcal{B}_{[0,1]})$ (where $\mathcal{B}_{[0,1]}$ is the Borel σ-field on $[0,1]$), and representing a weight normalized measure which can be expressed as the sum of an absolutely continuous measure with respect to the Lebesgue measure m on $[0,1]$ and a finite discretized measure on a set $\lambda_1, \ldots, \lambda_T$, that is,

$$dS = g\, dm \;\; \text{(or, more simply, } dS(\lambda) = g(\lambda)\, d\,\lambda\text{)}$$

with

$$g(\lambda) = \overline{g}(\lambda) + \sum_{t=1}^{T} k_t\, \delta(\lambda - \lambda_t),$$

where $\overline{g}$ denotes a Lebesgue measurable function, and δ is the Dirac distribution (i.e., $\delta(\lambda - \lambda_t) = 1$ if $\lambda = \lambda_t$, $\delta(\lambda - \lambda_t) = 0$ otherwise), and the function g satisfies that,

$$g(\lambda) \geq 0,$$
$$\int_{[0,1]} g(\lambda)\, d\lambda = 1,$$

$$g(0) > 0, \quad g(1) > 0,$$
$$\lambda_1 = 0, \quad \lambda_T = 1 \quad \text{if } T > 1.$$

Then,

Definition 0.1.5. *If S is a measure satisfying the preceding conditions, then the S-distance between two fuzzy numbers is defined as the mapping $D_S : \mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \rightarrow [0, +\infty)$ such that for $\widetilde{V}$, $\widetilde{W} \in \mathcal{F}_c(\mathbb{R})$*

$$D_S(\widetilde{V}, \widetilde{W}) = \sqrt{\int_{[0,1]} \left[d_S(\widetilde{V}_\alpha, \widetilde{W}_\alpha)\right]^2 d\,\alpha},$$

where

$$d_S(\widetilde{V}_\alpha, \widetilde{W}_\alpha) = \sqrt{\int_{[0,1]} \left[f_{\widetilde{V}}(\alpha,\lambda) - f_{\widetilde{W}}(\alpha,\lambda)\right]^2 \, d\,S(\lambda)}$$
$$= \sqrt{\int_{[0,1]} \left[f_{\widetilde{V}_\alpha}(\lambda) - f_{\widetilde{W}_\alpha}(\lambda)\right]^2 \, d\,S(\lambda)},$$

with

$$f_{\widetilde{V}_\alpha}(\lambda) = \lambda \sup \widetilde{V}_\alpha + (1-\lambda) \inf \widetilde{V}_\alpha = f_{\widetilde{V}}(\alpha, \lambda).$$

Conditions assumed for the measure S have been considered in Definition 0.1.5 in order to ensure that D_S is in fact a metric, since otherwise $D_S(\widetilde{V}, \widetilde{W}) = 0$ would not necessarily entail that $\widetilde{V} = \widetilde{W}$. Thus, for instance, if we assume that $g(1) = 1$ and $g(\lambda) = 0$ for all $\lambda \in [0, 1)$ and $\widetilde{V} = \text{Tri}(0, 1, 2)$, $\widetilde{W} = \text{Tra}(0, 0, 1, 2)$, then $D_S(\widetilde{V}, \widetilde{W}) = 0$ although $\widetilde{V}$ and $\widetilde{W}$ are not equal.

In contrast to the nonseparability of $(\mathcal{F}_c(\mathbb{R}), d_\infty)$, when $\mathcal{F}_c(\mathbb{R})$ is endowed with the metric D_S the corresponding metric space is separable (see Appendix A).

The metric D_S can be particularized to the case $S = S_{\vec{\lambda}}$ (with $\vec{\lambda} = (k_1, k_2, k_3)$, $k_1 + k_2 + k_3 = 1$ and $k_1, k_2, k_3 \in [0, 1)$), where $S_{\vec{\lambda}}$ is the parameterized weight measure such that the associated density g satisfies that

$g(0) = k_1$, $g(.5) = k_2$, $g(1) = k_3$ and $g(\lambda) = 0$ otherwise, which leads to the distance introduced by Salas (1991) and used by Lubiano *et al.* (1999a).

If we denotes by S_2 the metric $S = S_{\vec{\lambda}}$ corresponding to $\vec{\lambda} = (.5, .0, .5)$ we obtain that D_{S_2} coincides in the case of $\mathcal{F}_c(\mathbb{R})$ with the metric δ_2 considered by Näther (1997) and Körner (1997ab).

The metric δ_2 is a special case of the metric δ_p defined (see Diamond & Kloeden 1994) for all $\widetilde{V}$, $\widetilde{W} \in \mathcal{F}_c(\mathbb{R}^d)$ (class of fuzzy subsets of $\mathbb{R}^d$, i.e., mappings from $\mathbb{R}^d$ to $[0, 1]$, satisfying conditions *(i)-(iv)* in Definition 0.1.1) and $p \in [1, +\infty)$ by

$$\delta_p(\widetilde{V}, \widetilde{W}) = \left(\int_{[0,1]} \int_{S^{d-1}} \left| s_{\widetilde{V}}(u, \alpha) - s_{\widetilde{W}}(u, \alpha) \right|^p \nu(d\,u)\, d\alpha \right)^{\frac{1}{p}},$$

where ν is the normalized Lebesgue measure on the unit sphere S^{d-1}, $s_{\widetilde{V}}$ denotes the support function of $\widetilde{V}$ (Puri & Ralescu 1985), that is,

$$s_{\widetilde{V}}(u, \alpha) = \sup\{\langle u, v\rangle \mid v \in \widetilde{V}_\alpha\} \text{ for all } u \in S^{d-1} \text{ and } \alpha \in [0, 1],$$

where $\langle \cdot, \cdot \rangle$ is the inner product on $\mathbb{R}^d$.

Arguments supporting the use of the generalized metric D_S, and in particular the priority conferred on certain choices of S, are founded on the brief discussion we now present and which will be illustrated by an example.

In this way, Bertoluzza *et al.* (1995a) have pointed out some criticisms (in the case of $S = S_{\vec{\lambda}}$) of the choice of $\vec{\lambda} = (.5, .0, .5)$ (i.e., of the choice of S_2) to motivate the introduction of a more general weight measure. In this respect, the most common choices for $\vec{\lambda}$ in this case are those corresponding to $k_2 \leq k_1 = k_3$, and especially $\vec{\lambda} = (.375, .25, .375)$ or $\vec{\lambda} = (1/3, 1/3, 1/3)$.

A simple example illustrates the advantage of choosing either $\vec{\lambda} = (.375, .25,.375)$ or S = the Lebesgue measure m on $[0, 1]$, instead of S_2, when D_S^2 is applied to quantify a squared error in estimating. Consider the four pairs of triangular fuzzy numbers $(\widetilde{S}_1, \widetilde{S}_2)$, $(\widetilde{U}_1, \widetilde{U}_2)$, $(\widetilde{V}_1, \widetilde{V}_2)$ and $(\widetilde{W}_1, \widetilde{W}_2)$, where $\widetilde{S}_1 = \text{Tri}(-2, 0, 2)$ and $\widetilde{S}_2 = \text{Tri}(-1, 0, 1)$ represent two (more or less “narrow”) characterizations of the value or property ‘AROUND 0’, $\widetilde{U}_1 = \text{Tri}(-\sqrt{2},0,\ \sqrt{2})$ is another description of ‘AROUND 0’, $\widetilde{U}_2 = \text{Tri}(0,0,\sqrt{2})$ characterizes the value ‘AROUND 0 and positive’, $\widetilde{V}_1 = \text{Tri}(-1,0,0)$ is a description of ‘AROUND 0 and negative’, $\widetilde{V}_2 = \text{Tri}(0,0,1)$ is another characterization of ‘AROUND 0 and positive’, $\widetilde{W}_1 = \text{Tri}(-2,0,1)$ characterizes ‘AROUND 0 and TENDENTIALLY negative’, and $\widetilde{W}_2 = \text{Tri}(-1,0,2)$ characterizes ‘AROUND 0 and TENDENTIALLY positive’ (see Figure 0.1).

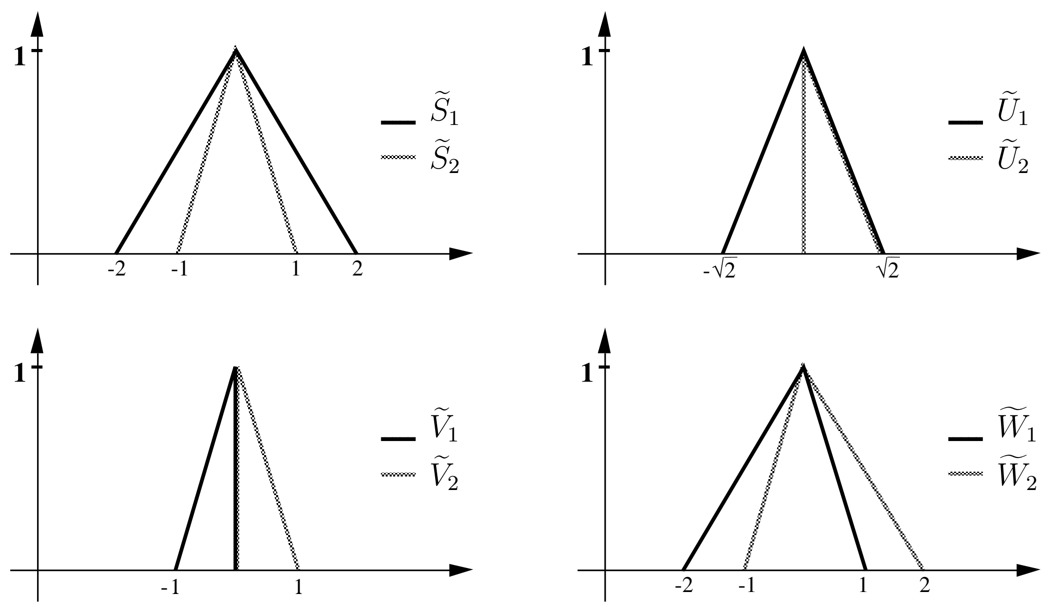


Fig. 0.1: Graphical representation of ($\widetilde{S}_1$,$\widetilde{S}_2$), ($\widetilde{U}_1$,$\widetilde{U}_2$), ($\widetilde{V}_1$,$\widetilde{V}_2$) y ($\widetilde{W}_1$,$\widetilde{W}_2$).

Then, we have that $D^2_{S_2}(\widetilde{S}_1,\widetilde{S}_2)= D^2_{S_2}(\widetilde{U}_1,\widetilde{U}_2)= D^2_{S_2}(\widetilde{V}_1,\widetilde{V}_2) = D^2_{S_2}(\widetilde{W}_1,\ \widetilde{W}_2) = .33333$, whereas $D^2_{(.375,.25,.375)}(\widetilde{S}_1,\ \widetilde{S}_2) = .25$, $D^2_{(.375,.25,.375)}(\widetilde{U}_1,\widetilde{U}_2) = .29167$, $D^2_{(.375,.25,.375)}(\widetilde{V}_1,\widetilde{V}_2) = D^2_{(.375,.25,.375)}(\widetilde{W}_1,\ \widetilde{W}_2) = .33333$ and $D^2_m(\widetilde{S}_1,\ \widetilde{S}_2) = .11111, D^2_m(\widetilde{U}_1,\widetilde{U}_2) = .22222, D^2_m(\widetilde{V}_1,\widetilde{V}_2) = D^2_m(\widetilde{W}_1,\ \widetilde{W}_2) = .33333$. Thus, when we try to estimate the first component of each pair by means of the second component in the same pair, the error incurred in estimating is much better quantified by either $D^2_{(.375,.25,.375)}$ or D^2_m, than by $D^2_{S_2}$.

On the other hand, the same example gives an idea of the convenience of using D_S instead of d_∞ in the considered problem. Obviously, $D_S\left(\widetilde{V},\widetilde{W}\right) \leq d_\infty\left(\widetilde{V},\widetilde{W}\right)$, for all $\widetilde{V}$, $\widetilde{W} \in \mathcal{F}_c(\mathbb{R})$ and for all S, and hence most of the results related to convergence and bounding based on d_∞ will be stronger than those based on D_S. Nevertheless, the separability of $(\mathcal{F}_c(\mathbb{R}), D_S)$ contrary to the nonseparability of $(\mathcal{F}_c(\mathbb{R}), d_\infty)$, implies that some results we could obtain for the first one are not necessarily true for the second one. However, in the preceding example, we have that $d_\infty(\widetilde{S}_1, \widetilde{S}_2) = d_\infty(\widetilde{V}_1, \widetilde{V}_2) = d_\infty(\widetilde{W}_1, \widetilde{W}_2) = 1$ and $d_\infty(\widetilde{U}_1, \widetilde{U}_2) = \sqrt{2}$, so that d_∞ will be valuable for purposes of formalizing measurability, integrability and probabilistic results on fuzzy-valued random elements, but it is not quite appropiate if it is employed for purposes of estimating a fuzzy number by another one. The reason justifying the behavior we have just observed lies in the use of the supremum considered in d_∞ in contrast to the "average" which is considered in D_S.

In this work we will consider random experiments involving quantification processes which cannot be identified with a real-valued random variable (that is, with a real-valued Borel measurable function), but rather with either interval-valued or fuzzy-valued quantification processes (more precisely, taking on values on either $\mathcal{K}_c(\mathbb{R})$ or $\mathcal{F}_c(\mathbb{R})$, respectively).

It should be emphasized that the quantification processes will be assumed to be intrinsically imprecise, so that imprecision does not come from an imprecise perception or report of an existing real-valued quantification.

Concepts and results in the following two sections concern the quantification processes converting the outcomes from a random experiment into elements of either $\mathcal{K}_c(\mathbb{R})$ or $\mathcal{F}_c(\mathbb{R})$. Throughout these two sections, we will assume that the original random experiment is mathematically modeled by means of a probability space $(\Omega, \mathcal{A}, P)$. A real-valued random variable associated with this space (that is, an $(\mathcal{A}, \mathcal{B}_{\mathbb{R}})$-measurable function from Ω to

$\mathbb{R}$, where $\mathcal{B}_{\mathbb{R}}$ is the Borel σ-field on $\mathbb{R}$) would represent the highest level of precision, but the lowest level of generalization from a theoretical viewpoint.

0.2 Concepts and results on random sets

The lowest level of precision, although associated with an intermediate level of generalization, corresponds to the concept of random set. In this section we will consider the special case of $\mathcal{K}_c(\mathbb{R})$-valued *convex compact random sets* (although a more general definition can be easily stated -see, for instance, Kendall 1974, Matheron 1975, Molchanov 1993-).

Given a probability space $(\Omega, \mathcal{A}, P)$, then

Definition 0.2.1. *A mapping $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ is said to be a convex compact random set associated with $(\Omega, \mathcal{A})$ if, and only if, X is $(\mathcal{A}, \mathcal{B}_{d_H})$-measurable, $\mathcal{B}_{d_H}$ being the σ-field generated by the topology induced by d_H on $\mathcal{K}_c(\mathbb{R})$.*

The *expected value* of a convex compact random set is defined in terms of the notion of integral of a measurable set-valued mapping with respect to a probability measure, which was introduced by Aumann (1965) as follows:

Definition 0.2.2. *If $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ is a convex compact random set, the expected value of X is the Aumann integral of X over Ω with respect to P, that is, the value*

$$E(X) = \Big\{ \textstyle\int_\Omega f(\omega)\, dP(\omega) \Big| f : \Omega \to \mathbb{R}, \\ f \in L^1(\Omega, \mathcal{A}, P),\ f \in X \, a.s.\, [P] \Big\}.$$

When we have to specify the probability measure P, we will alternatively denote $E(X)$ by $E(X|P)$.

The following notion establishes a sufficient condition for $E(X)$ not to be empty and furthermore to belong to $\mathcal{K}_c(\mathbb{R})$.

Definition 0.2.3. *A convex compact random set $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ is said to be integrably bounded if there exists a function $h : \Omega \to \mathbb{R}$, $h \in L^1(\Omega, \mathcal{A}, P)$, such that $|X(\omega)| \leq h(\omega)$ for all $\omega \in \Omega$, where $|X(\omega)| = \sup_{x \in X(\omega)} |x| = \max\{|\inf X(\omega)|, |\sup X(\omega)|\}$.*

To assume that X is $\mathcal{K}_c(\mathbb{R})$-valued, and more specifically the convexity of the valuesof X, allows us to express $E(X)$ in a very simple way (see, for instance, López-Díaz & Gil 1998a) as follows:

$$
\begin{aligned}
E(X) &= [E(\inf X), E(\sup X)] \\
&= \Big[\int_\Omega \inf X(\omega) d\,P(\omega), \int_\Omega \sup X(\omega) d\,P(\omega)\Big],
\end{aligned}
$$

since under the assumption above $\inf X$ and $\sup X$ become real-valued random variables.

Some special cases of convex compact random sets, we will sometimes refer to in the present work, are the following:

Definition 0.2.4. *A convex compact random set $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ is said to be degenerate, if there exists an interval $K \in \mathcal{K}_c(\mathbb{R})$ such that $X = K$ almost surely $[P]$. In particular, if K reduces to a singleton in $\mathbb{R}$, X is said to be a convex compact random set degenerate at a real value.*

Definition 0.2.5. *A convex compact random set $X : \Omega \to \mathcal{K}_c((0, +\infty))$ (i.e, $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ with $X(\omega) \subset \mathcal{P}((0, +\infty))$), is said to be a positive convex compact random set.*

Remark 0.2.1. In the case of a positive convex compact random set X, the integrable boundedness of X is equivalent to the condition $E(\sup X) < +\infty$.

Finally, the concept of *independence* of random sets is formalized as follows:

Definition 0.2.6. *Two convex compact random sets $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ and $Y : \Omega \to \mathcal{K}_c(\mathbb{R})$ are said to be independent if, and only if, $P(X \in K, Y \in K') = P(X \in K) \cdot P(Y \in K')$ whatever K, $K' \in \mathcal{B}_{d_H}$ may be.*

0.3 Concepts and results on fuzzy random variables

An intermediate level of precision, associated with the highest level of mathematical generalization, corresponds to the concept of fuzzy random variable. In this section we are going to consider the special case of $\mathcal{F}_c(\mathbb{R})$-valued fuzzy random variables (although Puri & Ralescu -1986- have established, in fact, a more general notion concerning a Euclidean space of finite dimension and the convexity of variable values is not necessarily assumed).

Given a probability space $(\Omega, \mathcal{A}, P)$, then

Definition 0.3.1. *A mapping $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ is said to be a fuzzy random variable (also called random fuzzy set) associated with $(\Omega, \mathcal{A})$ if, and only if, $\mathcal{X}$ is $(\mathcal{A}, \mathcal{B}_{d_\infty})$-measurable, $\mathcal{B}_{d_\infty}$ being the σ-field generated by the topology induced from d_∞ on $\mathcal{F}_c(\mathbb{R})$.*

Remark 0.3.1. An important fact to be emphasized in working with fuzzy random variables is the connection of this concept with that of random sets. Thus, if $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ is a fuzzy random variable, then for all $\alpha \in [0,1]$ the α-level function $\mathcal{X}_\alpha$, where $\mathcal{X}_\alpha : \Omega \to \mathcal{K}_c(\mathbb{R})$ is defined so that $\mathcal{X}_\alpha(\omega) = \left(\mathcal{X}(\omega)\right)_\alpha$ for all $\omega \in \Omega$, is a convex compact random set (see Puri & Ralescu 1986, Klement *et al.* 1986).

The *expected value* of a fuzzy random variable has been introduced by Puri & Ralescu (1986) as follows:

Definition 0.3.2. *If $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ is a fuzzy random variable, the expected value of $\mathcal{X}$ is the unique fuzzy subset of $\mathbb{R}$ (if it exists), $\widetilde{E}(\mathcal{X})$, such that for all $\alpha \in (0,1]$ we have that $\left(\widetilde{E}(\mathcal{X})\right)_\alpha = E\left(\mathcal{X}_\alpha\right)$, that is, $\left(\widetilde{E}(\mathcal{X})\right)_\alpha$ equals the Aumann integral of the convex compact random set $\mathcal{X}_\alpha$.*

When we have to specify the probability measure P, we will alternatively denote $\widetilde{E}(\mathcal{X})$ by $\widetilde{E}(\mathcal{X}|P)$.

The following notion (Puri & Ralescu 1986) establishes a sufficient condition for $\widetilde{E}(\mathcal{X})$ being well-defined and belonging to $\mathcal{F}_c(\mathbb{R})$.

Definition 0.3.3. *A fuzzy random variable* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *is said to be integrably bounded if, and only if,* $\mathcal{X}_0$ *is an integrably bounded convex compact random set.*

Remark 0.3.2. Stojaković (1994) proved that $\left(\widetilde{E}(\mathcal{X})\right)_0 = E\left(\mathcal{X}_0\right)$.

Remark 0.3.3. Since $\mathcal{X}$ is $\mathcal{F}_c(\mathbb{R})$-valued, then $\left(\widetilde{E}(\mathcal{X})\right)_\alpha = \left[E\left(\inf \mathcal{X}_\alpha\right), E\left(\sup \mathcal{X}_\alpha\right)\right]$ for all $\alpha \in [0,1]$.

It should be emphasized that Yager's ranking criterion is especially operational when it is combined with the (fuzzy) expected value of a fuzzy random variable. More precisely, and on the basis of the results from López-Díaz & Gil (1998a) (which have been really established for the more general criterion of Campos & González 1989) one can conclude that if $\mathcal{X}$ is an integrably bounded fuzzy random variable with expected value $\widetilde{E}(\mathcal{X})$, then

$$F\left(\widetilde{E}(\mathcal{X})\right) = E(F \circ \mathcal{X}),$$

so that the value of the ranking function F for the (fuzzy) expected value of the fuzzy random variable $\mathcal{X}$ reduces to the expected value of the real-valued random variable $F \circ \mathcal{X}$.

Some special cases of fuzzy random variables, we will refer sometimes to in the present work, are the following:

Definition 0.3.4. *A fuzzy random variable* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *is said to be degenerate, if there exists a fuzzy number* $\widetilde{V} \in \mathcal{F}_c(\mathbb{R})$ *such that* $\mathcal{X} = \widetilde{V}$ *almost surely* $[P]$*. In particular, if* $\widetilde{V}$ *reduces to the indicator function of an interval of* $\mathcal{K}_c(\mathbb{R})$*,* $\mathcal{X}$ *is said to be a fuzzy random variable degenerate at an interval value, and if* $\widetilde{V}$ *reduces to the indicator function of a singleton in* $\mathbb{R}$*,* $\mathcal{X}$ *is said to be a fuzzy random variable degenerate at a real value.*

Definition 0.3.5. *A fuzzy random variable* $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ *(i.e.,* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *with* $\mathcal{X}_0(\omega) \subset \mathcal{P}((0,+\infty))$*), is said to be a positive fuzzy random variable.*

Remark 0.3.4. In the case of a positive fuzzy random variable $\mathcal{X}$, the integrable boundedness of $\mathcal{X}$ is equivalent to $E(\sup \mathcal{X}_0) < +\infty$.

The notion of *independence* of fuzzy random variables is formalized as follows:

Definition 0.3.6. *Two fuzzy random variables* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *and* $\mathcal{Y} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *are said to be independent if, and only if,* $P(\mathcal{X} \in A, \mathcal{Y} \in B) = P(\mathcal{X} \in A)P(\mathcal{Y} \in B)$ *whatever* A, $B \in \mathcal{B}_{d_\infty}$ *may be.*

Finally, a supporting result for some conclusions to be drawn in this work is that corresponding to the extension of the *double expectation Theorem* which has been proved by López-Díaz & Gil (1998b):

Theorem 0.3.1. *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space and let* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *and* $\mathcal{Y} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *be two integrably bounded fuzzy random variables. Let* $\sigma_{\mathcal{X}}$ *and* $\sigma_{\mathcal{Y}}$ *be the* σ*-fields on* $\mathcal{F}_c(\mathbb{R})$ *induced from* $\mathcal{A}$ *by* $\mathcal{X}$ *and* $\mathcal{Y}$*, respectively (that is,* $\sigma_{\mathcal{X}} = \{B \subset \mathcal{F}_c(\mathbb{R}) \,|\, \mathcal{X}^{-1}(B) \in \mathcal{A}\}$, $\sigma_{\mathcal{Y}} = \{B \subset \mathcal{F}_c(\mathbb{R}) \,|\, \mathcal{Y}^{-1}(B) \in \mathcal{A}\}$*). Let* $P_{\mathcal{X}}$ *and* $P_{\mathcal{Y}}$ *be the probability measures induced from* P *by* $\mathcal{X}$ *and* $\mathcal{Y}$*, respectively.*

Let $(\mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{X}} \otimes \sigma_{\mathcal{Y}}, P_{\mathcal{X}} \otimes P_{\mathcal{Y}})$ *be the product probability space, and let* $\mathcal{Y}^* : \mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \to \mathcal{F}_c(\mathbb{R})$ *be the integrably bounded fuzzy random variable defined such that* $\mathcal{Y}^*(\tilde{x}, \tilde{y}) = \tilde{y}$ *whatever* $\tilde{x}, \tilde{y} \in \mathcal{F}_c(\mathbb{R})$ *may be. Assume that when* $\mathcal{X} = \tilde{x}$ *the conditional probability distribution induced by* $\mathcal{Y}$ *corresponds to a regular conditional distribution on* $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{Y}})$ *denoted by* $P_{\tilde{x}}$*, that is,*

- *$P_{\tilde{x}}$ is a probability measure on $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{Y}})$ for each $\tilde{x} \in \mathcal{X}(\Omega)$, and*

- *for each $B \in \sigma_{\mathcal{Y}}$ the mapping $h_B : \mathcal{X}(\Omega) \to [0,1]$ such that $h_B(\tilde{x}) = P_{\tilde{x}}(B)$ is a real-valued random variable associated with the measurable space $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{X}})$ and satisfying that for all $A \in \sigma_{\mathcal{X}}$*

$$P(\mathcal{X} \in A, \mathcal{Y} \in B) = \int_A P_{\tilde{x}}(B)\, d\, P_{\mathcal{X}}.$$

If we define the mapping $\varphi : \mathcal{X}(\Omega) \to \mathcal{F}_c(\mathbb{R})$ such that $\varphi(\tilde{x}) = \widetilde{E}\,(\mathcal{Y} \,|\, \mathcal{X} = \tilde{x})$ where it denotes by $\widetilde{E}\,(\mathcal{Y} \,|\, \mathcal{X} = \tilde{x}) = \widetilde{E}\,(\mathcal{Y}^*_{\tilde{x}} \,|\, P_{\tilde{x}})$ for each $\tilde{x} \in \mathcal{X}(\Omega)$, then,

a) φ is a integrably bounded fuzzy random variable associated with $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{X}}, P_{\mathcal{X}})$;

b) $\widetilde{E}\,(\mathcal{Y} \,|\, P) = \widetilde{E}\left(\varphi \,\middle|\, P_{\mathcal{X}}\right)$, that is, $\widetilde{E}\,(\mathcal{Y} \,|\, P) = \widetilde{E}\left(\widetilde{E}\,(\mathcal{Y}^*_{\tilde{x}} \,|\, P_{\tilde{x}}) \,\middle|\, P_{\mathcal{X}}\right)$.

0.4 Inequality measures for real-valued random variables

In this section we will recall the basic notions concerning the measurement of the inequality of a population associated with a real-valued random variable.

The *inequality* of a population with respect to a quantitative attribute which can be formalized by means of a random variable (which is usually assumed to be positive), is a characteristic measuring the relative variation of the attribute in the population. The analysis of the inequality, its numerical quantification, and the properties of the measures suggested for this quantification give rise to a topic having many applications to fields like Economics (income inequality, wealth inequality, etc.) and Industry (industrial concentration, etc.).

In the literature on the quantification of the inequality of a population with respect to a certain random variable, many indices have been proposed.

Several of them have been widely accepted by both the international mathematical community and the international scientific communities in the fields of application. During recent years, some of the best known and most used inequality measures are those coinciding with (or being increasing functions of) the additively decomposable indices of order α.

The studies on inequality usually assume that the values of the considered attribute are positive, since the most common attributes are monetary (like income, wealth, etc.) or correspond to the size of a subpopulation. In addition, variables for which inequality is measured are usually ratio scale ones.

If X is a positive random variable, the *additively decomposable indices* of order α (see, for instance, Bourguignon 1979, Cowell 1980, Shorrocks 1980, Cowell & Kuga 1981, Eichhorn & Gehrig 1982, Zagier 1983, Gil *et al.* 1989b), are defined (if they exist) as follows:

$$\mathrm{I}^{\alpha}(X) = [\alpha(\alpha-1)]^{-1}\left[E\left(\left(\frac{X}{E(X)}\right)^{\alpha}\right)-1\right] \text{ if } \alpha \neq 0, 1,$$

$$\mathrm{I}^{0}(X) = -E\left(\log\left(\frac{X}{E(X)}\right)\right),$$

$$\mathrm{I}^{1}(X) = E\left(\frac{X}{E(X)}\log\left(\frac{X}{E(X)}\right)\right),$$

although often the factor $[\alpha(\alpha-1)]^{-1}$ is ignored but we take into account the associated sign.

The parameter α plays the role of a weighting element of the "degree of aversion to inequality", or "degree of sensitivity to transfers between classes".

Indices I^0 and I^1 satisfy that $\mathrm{I}^0(X) = \lim_{\alpha\to 0}\mathrm{I}^{\alpha}(X)$ and $\mathrm{I}^1(X) = \lim_{\alpha\to 1}\mathrm{I}^{\alpha}(X)$, whatever the variable X may be. I^1 is the well-known *Theil index* (1967). I^0 (or *index of the Shannon type*) has received special attention in some studies (c.f. Bourgignon 1979, Gil 1979, 1981, 1982), and I^{-1} (or *hyperbolic index*), and its practical interest has been examined in detail in previous works (see Gil & Gil 1989, Gil *et al.* 1989ab, Martínez 1991).

A more general family of indices which includes the additively decomposable indices, except by the Theil and the Shannon type indices, is that introduced by Gastwirth (1975) (see also Gastwirth *et al.* 1986) or family of the inequality ϕ*-indices* which for a random variable X taking on values in $(0, +\infty)$ are generically given by

$$\mathrm{I}_{\phi}(X) = E\left(\frac{\phi(X)}{\phi\left(E(X)\right)} - 1\right),$$

where $\phi : (0, +\infty) \to \mathrm{I\!R}$, $\phi \in C^1$ and ϕ is *convex* (where this convexity is usually assumed to be *strict*, that is, for all $\lambda \in (0, 1)$ and x, $y \in (0, +\infty)$ with $x \neq y$ we have that $f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y)$).

Recently (see Alonso *et al.* 1998), we have defined a generalized family of inequality indices, including all the additively decomposable indices (even for Theil's and Shannon's type cases). These indices are based on the generalized family of the directed-divergence measures stated by Csiszár (1967).

In previous works (see, for instance, Gil *et al.* 1989b) we have pointed out the existence of a formal connection between the quantification of a population with respect to a real-valued random variable and the Information Theory. In this way, the best known measures of inequality can be obtained as particular versions, or derived from, measures of the directed-divergence between two probability distributions.

More specifically, each of the additively decomposable inequality indices of order α coincides, except may be by a constant factor, with a nonadditive directed-divergence measure of order α (Rathie 1971) for $\alpha \neq 0, 1$, or with the directed-divergence measure of Kullback-Leibler (see Kullback-Leibler 1951) for $\alpha = 0$ or 1, for two special probability distributions.

This fact, along with the obvious interest in using families of measures rich enough to allow us to find a proper measure in the family to handle each problem, support the idea of constructing *generalized inequality measures.*

Divergence measures between two distributions are defined as functions of these distributions quantifying the "degree of discrepancy" between them (where the easier it is to discriminate between two distributions, the higher the discrepancy between them).

In 1967 Csiszár introduced a family of measures extending Kullback-Leibler's one, and also including as special cases the non-additive directed-divergence measures by Rathie (1971). This family has been established as follows: given two probability measures P and Q on the measurable space $(\Omega, \mathcal{A})$, and having density functions p and q, respectively, with respect to a σ-finite measure ν on $(\Omega, \mathcal{A})$, the *Csiszár f-divergence* between P and Q with respect to P is defined by the value (if it exists)

$$\mathrm{D}_f(P;Q) = \int_\Omega p(\omega)\, f\left(\frac{q(\omega)}{p(\omega)}\right) d\,\nu(\omega),$$

where $f : (0, +\infty) \to \mathbb{R}$ is a convex function satisfying that

$$f(0) = \lim_{u \downarrow 0} f(u), \quad 0 \cdot f\left(\frac{0}{0}\right) = 0, \quad 0 \cdot f\left(\frac{a}{0}\right) = \lim_{\varepsilon \downarrow 0} \varepsilon \cdot f\left(\frac{a}{\varepsilon}\right) = a \cdot \lim_{u \uparrow \infty} \frac{f(u)}{u}$$

(with $a \geq 0$).

On the basis of the Csiszár f-divergence it is possible to construct a new generalized family of inequality indices including the hyperbolic index, the normalized variance, the index of the Shannon type and Theil's index (the two latter not being included in the Gastwirth family -1975- as pointed out before).

The hyperbolic index has shown a valuable behavior in dealing wih finite populations, since it allows us to define unbiased estimators in probabilistic samplings. Furthermore, and on the basis of simulations studies, we could also conclude that the associated relative sampling error is lower for the hyperbolic index than for other more frequently used ones (see, for instance, Martínez 1991, Gil *et al* 1989b, Gil & Gil 1989).

The index of the Shannon type has an operational additive decomposability and a very intuitive interpretation.

Although Theil's index is included in this new generalized family, we will not consider it when the indices are extended to the fuzzy environment, since we will be compelled to assume the function f to be monotonic in order to get operational expressions for the extended indices, and Theil's index is associated with a nonmonotonic function f.

Thus, if X is a random variable associated with $(\Omega, \mathcal{A})$ and we consider the probability measures P and Q which are related such that, if p and q are the respective densities associated with P and Q with respect to ν, $q(\omega) = p(\omega)X(\omega)/E(X)$ for all $\omega \in \Omega$, and we assume that $f(x) = x^{-1} - 1$ for all $x \in (0, +\infty) = \mathrm{Dom} f$, we obtain the *hyperbolic index*

$$\mathrm{I}_H(X) = E\left(\frac{E(X)}{X} - 1\right).$$

For the same distributions and the function $f(x) = -\log x$ for all $x \in (0, +\infty) = \mathrm{Dom} f$, we obtain the *index of the Shannon type*

$$\mathrm{I}_{Sh}(X) = -E\left(\log\left(\frac{X}{E(X)}\right)\right).$$

Finally, to achieve Theil's index we have just considered the probability measures P and Q above mentioned and the function $f(x) = x\log(x)$ for all $x \in (0, +\infty) = \mathrm{Dom} f$, whence

$$\mathrm{I}_T(X) = E\left(\frac{X}{E(X)}\log\left(\frac{X}{E(X)}\right)\right).$$

On the basis of the preceding measures we can state the following:

Definition 0.4.1. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, and let $X : \Omega \to (0, +\infty)$ be a positive random variable associated with $(\Omega, \mathcal{A})$. Let $f : (0, +\infty) \to \mathbb{R}$ be a strictly convex and continuous function satisfying that $f(1) = 0$. The f-inequality index associated with X is given by the value (if it exists)*

$$\mathrm{I}_f(X) = E\left(f\left(\frac{X}{E(X)}\right)\right).$$

The f-inequality indices satisfy the following suitable and desirable properties. Let $(\Omega, \mathcal{A}, P)$ be a probability space, and let $X : \Omega \rightarrow (0, +\infty)$ be a positive integrable random variable. Then,

- **(Mean independence)** *For all* $k \in (0, +\infty)$ *we have that* $\mathrm{I}_f(k \cdot X) = \mathrm{I}_f(X)$.
- **(Nonnegativeness)** $\mathrm{I}_f(X) \geq 0$.
- **(Minimality)** $\mathrm{I}_f(X) = 0$ *if, and only if,* X *is a degenerate random variable.*

If X is defined on a population Ω of N individuals, $\omega_1, \ldots, \omega_N$, and $\mathcal{A} = \mathcal{P}(\Omega)$, then

- **(Symmetry)** *If* σ *is a permutation on* Ω, *then* $\mathrm{I}_f(X \circ \sigma) = \mathrm{I}_f(X)$.
- **(Population homogeneity)** *If* X^{*r} *is the positive random variable defined as the extension of* X *to the population of* $N \times r$ *individuals, obtained from* Ω *by replicating it* r *times, then* $\mathrm{I}_f(X^{*r}) = \mathrm{I}_f(X)$.
- **(Continuity)** *If* $X_{h,\varepsilon}$ *is a positive random variable associated with* $(\Omega, \mathcal{A})$ *such that* $X_{h,\varepsilon}(\omega_h) = X(\omega_h) + \varepsilon$ *and* $X_{h,\varepsilon}(\omega_j) = X(\omega_j)$ *for all* $j \in \{1, \ldots, N\} \setminus \{h\}$ *for some* $\varepsilon \in \mathbb{R}$ *such that* $X(\omega_h) + \varepsilon > 0$, *then*

$$\lim_{\varepsilon \to 0} \mathrm{I}_f(X_{h,\varepsilon}) = \mathrm{I}_f(X).$$

- **(Schur-convexity)** *If* (μ_{jl}) *is an* $N \times N$ *doubly stochastic matrix and* X' *is a positive random variable such that*

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_N \end{pmatrix} = \begin{pmatrix} \mu_{11} & \cdots & \mu_{1N} \\ \vdots & \ddots & \vdots \\ \mu_{N1} & \cdots & \mu_{NN} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix},$$

then we have that $\mathrm{I}_f(X) \geq \mathrm{I}_f(X')$ *with equality if, and only if,* (μ_{jl}) *is equivalent to a permutation on* Ω.

- **(Compatibility with Lorenz criterion)** *If we consider two positive random variables* $X : \Omega \to (0, +\infty)$ *and* $X' : \Omega \to (0, +\infty)$, *such that* $X(\omega_N) \geq \ldots \geq X(\omega_1), X'(\omega_N) \geq \ldots \geq X'(\omega_1)$, $E(X) = E(X')$ *and* $X(\omega_1)+\cdots+X(\omega_k) \geq X'(\omega_1)+\cdots+X'(\omega_k)$ *for all* $k \in \{1, \ldots, N\}$ *with strict inequality for at least one* k, *then we have that* $\mathrm{I}_f(X) < \mathrm{I}_f(X')$.

- **(Progressive principle of transfers)** *Assume that there exist* $h, l \in \{1, \ldots, N\}$, *with* $X(\omega_h) \geq X(\omega_l) > 0$, *and consider a value* $\varepsilon \in [0, X(\omega_h) - X(\omega_l)]$ *so that we define a new positive random variable* X' *such that*

$$X'(\omega_h) = X(\omega_h) - \varepsilon\,,\ X'(\omega_l) = X(\omega_l) + \varepsilon,$$

$$X'(\omega_j) = X(\omega_j),\ \textit{for } j \in \{1, \ldots, N\} \setminus \{h, l\},$$

then we have that $\mathrm{I}_f(X) \geq \mathrm{I}_f(X')$, *with equality if, and only if,* $X' = X$ *(i.e.,* $\varepsilon = 0$*) or* $X' = X \circ \sigma_{hl}$ *where* σ_{hl} *denotes the permutation on* Ω *exchanging the* h*-th and* l*-th individuals (i.e.,* $\varepsilon = X(\omega_h) - X(\omega_l)$*).*

- **(Regressive principle of transfers)** *Assume that there exist* $h, l \in \{1, \ldots, N\}$, *with* $X(\omega_h) \geq X(\omega_l) > 0$, *and consider a value* $\varepsilon \in [0, X(\omega_l)]$, *so that we define a new positive random variable* X' *such that*

$$X'(\omega_h) = X(\omega_h) + \varepsilon\,,\ X'(\omega_l) = X(\omega_l) - \varepsilon,$$

$$X'(\omega_j) = X(\omega_j),\ \textit{for } j \in \{1, \ldots, N\} \setminus \{h, l\},$$

then we have that $\mathrm{I}_f(X') \geq \mathrm{I}_f(X)$, *with equality if, and only if,* $X' = X$ *(i.e.,* $\varepsilon = 0$*).*

- **(Grouping effects)** *If* Ω *is divided in accordance with the partition* $\mathcal{P} = \{\Omega_m\}_{m=1}^M$ *with* $N_1, \ldots, N_M$ *(*$N = N_1 + \cdots + N_M$*) the cardinals of* $\Omega_1, \ldots, \Omega_M$, *respectively, and being* $\Omega_m = \{\omega_{m1}, \ldots, \omega_{mN_m}\}$, $m = 1, \ldots, M$, *and we consider the uniform dstribution* P *on the probability* *space associated with* Ω *. If* $X : \Omega \to (0, +\infty)$ *is a positive random variable associated with* $(\Omega, \mathcal{P}(\Omega), P)$*, and* $X_{\mathcal{P}} : \mathcal{P} \to (0, +\infty)$ *is the random variable such that* $X_{\mathcal{P}}(\Omega_m)$ = *the expected value of* X *on* Ω_m $(m = 1, \ldots, M)$*, and* X_{Ω_m} *denotes the restriction of* X *from* Ω *to* Ω_m $(m = 1, \ldots, M)$*, then*

$$\mathrm{I}_f(X) \geq \mathrm{I}_f(X_{\mathcal{P}}).$$

Furthermore, $\mathrm{I}_f(X)$ *equals* $\mathrm{I}_f(X_{\mathcal{P}})$ *if, and only if,* X_{Ω_m} *is a degenerate variable for each* $m \in \{1, \ldots, M\}$.

In the case of the additively decomposable indices, the grouping effects property above can be strengthened as follows:

- **(Additive decomposability)** *If* Ω *is divided in accordance with the partition* $\mathcal{P}$ *in the latter property, then we have that*

$$\mathrm{I}^{\alpha}(X) = \mathrm{I}^{\alpha}(X_{\mathcal{P}}) + \sum_{m=1}^{M} \left(\frac{N_m}{N}\right)^{\alpha} \left(\frac{N_m X_{\mathcal{P}}(\Omega_m)}{N E(X)}\right)^{1-\alpha} \mathrm{I}^{\alpha}(X_{\Omega_m}).$$

Moreover, $\mathrm{I}^{\alpha}(X) = \mathrm{I}^{\alpha}(X_{\mathcal{P}})$ *if, and only if,* X_{Ω_m} *is a degenerate variable for each* $m \in \{1, \ldots, M\}$.

As we have already commented, when we extend the f-inequality indices to the case of fuzzy random variables and convex compact random sets, we will assume some extra conditions for the function f in order to get operational indices satisfying suitable properties. As compensation for this extra requirement, we will have only to be deprived of the extension of a few indices, like Theil's one. The family of functions f satisfying the conditions assumed in Chapters 2 and 3 to define the generalized f-inequality measure associated with fuzzy random variables and convex compact random sets is very wide.

Nevertheless, and under some quite general conditions, one can state a certain characterization of this family (see Alonso *et al.* 1998), as well as of some relevant subfamilies we will refer to in Chapters 2 and 3 to extend the minimality property from the real-valued case (see Appendix B).

Chapter 1

Generalized measure of the absolute variation (deviation) of a fuzzy random variable

In the study of fuzzy random variables it is worth describing their behavior by means of certain measures summarizing some of the most relevant characteristics. In this way, the expected value of a fuzzy random variable has been introduced (Puri & Ralescu 1986) as a summary (fuzzy-valued) measure of the "central tendency" of the variable values.

Another useful characteristic in the case of real-valued random variables is the "deviation" or absolute variation. This characteristic allows the comparison between different variables or populations, usually through the measurement of the deviation of variable values with respect to a concrete value.

Since the absolute variation measures are mainly defined with the purpose of establishing comparisons, it would be valuable (even when working with fuzzy-valued random variables) that the measures for this variation were real-valued, so that these comparisons would reduce to the standard ranking of real numbers.

A measure of the average deviation of a fuzzy random variable with respect to a fuzzy number, should express how much "in error" this number is expected to be as a description of variable values. This error can be quantified in a natural way in terms of an increasing function of a suitable distance between fuzzy numbers.

In the case of real-valued random variables, the error above mentioned is usually measured by the squared Euclidean distance in $\mathbb{R}$. Reasons justifying the use of the expected value of the squared Euclidean distance to quantify the mean deviation against other possible measures are diverse, namely: all the variable values are involved in its computation; it can be easily estimated in probabilistic random samplings in finite or infinite populations; it behaves operationally in dealing with linear combinations of random variables (especially when we consider pairwise independent and identically distributed ones); it is properly adapted to statistical inferential management.

In this chapter we introduce a measure of the absolute variation or deviation, which is an extension of the second moment (and, in particular, of the notion of variance) of the classical case to a fuzzy random variable with respect to a fuzzy number. This measure is stated on the basis of the generalized metric defined on the space of fuzzy numbers by Bertoluzza *et al.* (1995a) and presented in Definition 0.1.5.

We later examine conditions for the existence of such an extension, and analyze properties of the variation measure we have introduced, as well as the extension of some outstanding results having subsequent interest.

The unbiased estimation of the absolute variation measure is developed in random samplings with and without replacement from finite populations and an exact quantification of the associated sampling error is given. This study is complemented with that of the asymptotic distribution of the sample variation measure, which allows us to state approximate inferential techniques on the population measure.

Finally, two applications of the measure of absolute variation are considered: the quantification of the sampling error associated with the fuzzy estimation of the fuzzy parameter corresponding to the expected value of the fuzzy random variable, in random samplings with and without replacement; the quantification of the error associated with a functional (and, in particular, a linear) relation between two fuzzy random variables.

1.1 The S-mean squared dispersion associated with a fuzzy random variable

In this section we first introduce a measure of the absolute variation associated with a fuzzy random variable with respect to an element in $\mathcal{F}_c(\mathbb{R})$, and in particular with respect to the variable expected value. We also examine a sufficient condition for the existence of the value of the introduced measure.

Let $(\Omega, \mathcal{A}, P)$ be a probability space and let $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ be an integrably bounded fuzzy random variable associated with it.

Definition 1.1.1. *The S-mean squared dispersion of $\mathcal{X}$ about $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$, is given by the value (if it exists)*

$$MSD_S\left(\mathcal{X}, \widetilde{A}\right) = E\left(\left[D_S\left(\mathcal{X}(\,\cdot\,), \widetilde{A}\right)\right]^2\right) = \int_\Omega \left[D_S\left(\mathcal{X}(\omega), \widetilde{A}\right)\right]^2 dP(\omega).$$

The S-mean squared dispersion (or, briefly, S-MSD) does not necessarily exist for a fuzzy random variable. We are now going to study some conditions about $\mathcal{X}$ to ensure that $MSD_S\left(\mathcal{X}, \widetilde{A}\right) \in \mathbb{R}$ whatever $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ may be.

In the first result, we will prove that the measurability of $\mathcal{X}$ ensures the $(\mathcal{A}, \mathcal{B}_{\mathbb{R}})$-measurability of $\left[D_S\left(\mathcal{X}(\,\cdot\,), \widetilde{A}\right)\right]^2$.

Theorem 1.1.1. *Let $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ be a fuzzy random variable associated with the probability space $(\Omega, \mathcal{A}, P)$, and let $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$. Then, the function $\left[D_S\left(\mathcal{X}(\,\cdot\,), \widetilde{A}\right)\right]^2 : \Omega \to \mathbb{R}$ is $(\mathcal{A}, \mathcal{B}_{\mathbb{R}})$-measurable.*

Proof.

Let $\left([0,1], \mathcal{M}_{[0,1]}, m\right)$ denote the Lebesgue measure space on $[0,1]$ and let $([0,1],\ \mathcal{B}_{[0,1]}, S)$ be a measure space on $[0,1]$, where S is a measure satisfying the conditions in Definition 0.1.5. Consider the mapping $h_{\mathcal{X}} : \Omega\times[0,1]\times[0,1] \to \mathbb{R}$ given by

$$h_{\mathcal{X}}(\omega,\alpha,\lambda) = f_{\mathcal{X}(\omega)}(\alpha,\lambda) = \lambda \sup \mathcal{X}_\alpha(\omega) + (1-\lambda)\inf \mathcal{X}_\alpha(\omega).$$

Since the functions $h_1 : \Omega\times[0,1] \to \mathbb{R}$ and $h_2 : \Omega\times[0,1] \to \mathbb{R}$, where $h_1(\omega,\alpha) = \inf \mathcal{X}_\alpha(\omega)$ and $h_2(\omega,\alpha) = \sup \mathcal{X}_\alpha(\omega)$, are $\mathcal{A}\otimes\mathcal{M}_{[0,1]}$-measurable (see López-Díaz & Gil 1998b), then $h_{\mathcal{X}}$ is $\mathcal{A}\otimes\mathcal{M}_{[0,1]}\otimes\ \mathcal{B}_{[0,1]}$-measurable.

On the other hand, the mappings $h_1^* : \Omega\times[0,1] \to \mathbb{R}$ and $h_2^* : \Omega\times[0,1] \to \mathbb{R}$ given by $h_1^*(\omega,\alpha) = \inf \widetilde{A}_\alpha$ and $h_2^*(\omega,\alpha) = \sup \widetilde{A}_\alpha$, are left-continuous with respect to α on $[0,1]$ and do not depend on ω, so that they are $\mathcal{A}\otimes\mathcal{M}_{[0,1]}$-measurable. Consequently, the function $f_{\widetilde{A}} :\ [0,1]\times[0,1] \to \mathbb{R}$, where $f_{\widetilde{A}}(\alpha,\lambda) = \lambda\sup\widetilde{A}_\alpha + (1-\lambda)\inf\widetilde{A}_\alpha$ is $\mathcal{A}\otimes\mathcal{M}_{[0,1]}\otimes\mathcal{B}_{[0,1]}$-measurable, and hence $\left[f_{\mathcal{X}(\cdot)}(\cdot,\cdot) - f_{\widetilde{A}}(\cdot,\cdot)\right]^2$ is also $\mathcal{A}\otimes\mathcal{M}_{[0,1]}\otimes\mathcal{B}_{[0,1]}$-measurable.

Since S and m are finite measures, on the basis of the Fubini Theorem we can conclude that $\left[D_S\left(\mathcal{X}(\cdot),\widetilde{A}\right)\right]^2$ is $(\mathcal{A},\mathcal{B}_{\mathbb{R}})$-measurable. □

Now, we establish an extra condition with respect to the integrable boundedness of $\mathcal{X}$ to ensure the existence of the S-mean squared dispersion of $\mathcal{X}$ about $\widetilde{A}$.

Theorem 1.1.2. *Let $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ be a fuzzy random variable associated with $(\Omega,\mathcal{A},P)$, and let $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$. If $|\mathcal{X}_0| \in L^2(\Omega,\mathcal{A},P)$, the function $\left[D_S\left(\mathcal{X}(\cdot),\widetilde{A}\right)\right]^2 :\ \Omega \to \mathbb{R}$ belongs to $L^1(\Omega,\mathcal{A},P)$.*

Proof.

Indeed, since for all $\omega\in\Omega$ and $\alpha\in[0,1]$ we have that

$$|\sup \mathcal{X}_\alpha(\omega)| \le |\mathcal{X}_0|(\omega),\ \ |\inf \mathcal{X}_\alpha(\omega)| \le |\mathcal{X}_0|(\omega),$$

and $|\mathcal{X}_0| \in L^2(\Omega,\mathcal{A},P)$, then the functions h_1 and h_2 defined in the proof of Theorem 1.1.1 belong to $L^2(\Omega\times[0,1],\mathcal{A}\otimes\mathcal{M}_{[0,1]}, P\otimes m)$.

On the other hand, as $\widetilde{A}_0$ does not depend on ω and α, and the function $\max\{|\sup \widetilde{A}_0|, |\inf \widetilde{A}_0|\} \in L^2(\Omega, \mathcal{A}, P)$ and dominates to $|\sup \widetilde{A}_\alpha|$ and $|\inf \widetilde{A}_\alpha|$, then h_1^* and h_2^* belong to $L^2(\Omega \times [0,1], \mathcal{A} \otimes \mathcal{M}_{[0,1]}, P \otimes m)$. Consequently, $\left[f_{\mathcal{X}(\cdot)}(\cdot,\cdot) - f_{\widetilde{A}}(\cdot,\cdot)\right]^2 \in L^1(\Omega \times [0,1] \times [0,1], \mathcal{A} \otimes \mathcal{M}_{[0,1]} \otimes \mathcal{B}_{[0,1]}, P \otimes m \otimes S)$, whence on the basis of the Fubini Theorem we can conclude that the mapping $\left[D_S\left(\mathcal{X}(\cdot), \widetilde{A}\right)\right]^2$ belongs to $L^1(\Omega, \mathcal{A}, P)$. □

The S-mean squared dispersion of a fuzzy random variable can be defined about the expected value of this variable. This measure will extend the notion of variance of a real-valued random variable, and it is formalized as follows:

Definition 1.1.2. *The central S-mean squared dispersion of $\mathcal{X}$ is given by the value (if it exists)*

$$\Delta_S^2(\mathcal{X}) = MSD_S\left(\mathcal{X}, \widetilde{E}(\mathcal{X})\right) = \int_\Omega \left[D_S\left(\mathcal{X}(\omega), \widetilde{E}(\mathcal{X})\right)\right]^2 dP(\omega),$$

where $\widetilde{E}(\mathcal{X})$ represents the (fuzzy) expected value of $\mathcal{X}$ with respect to P.

Occasionally, and when we have to specify the probability measure P, we will alternatively denote $\Delta_S^2(\mathcal{X})$ by $\Delta_S^2(\mathcal{X} \mid P)$.

Obviously, conditions in Theorem 1.1.2 also guarantee the existence of $\Delta_S^2(\mathcal{X})$.

1.2 Properties of the S-mean squared dispersion associated with a fuzzy random variable

In this section we are going to examine some of the most valuable properties of the S-MSD. Most of these properties extend those for the second moment and variance of real-valued random variables.

Throughout this section, we will assume that the involved fuzzy random variables satisfy the conditions guaranteeing the existence of the S-MSD about the considered fuzzy value.

First of all, we will verify that *the S-MSD extends* to the fuzzy-valued case the well-known *second moment* of a real-valued random variable about a real number.

Theorem 1.2.1 (Extension of the real-valued case). *If $\mathcal{X}$ is a real-valued random variable (that is, $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ is a fuzzy random variable with $\mathcal{X}(\Omega) \subset \{\mathbf{1}_{\{x\}} \mid x \in \mathbb{R}\}$) and $a \in \mathbb{R}$, then $MSD_S(\mathcal{X}, \mathbf{1}_{\{a\}})$ corresponds to the second moment of $\mathcal{X}$ about a. In particular, $\Delta_S^2(\mathcal{X}) = \mathrm{Var}(\mathcal{X})$, where* Var *denotes the variance of a real-valued random variable.*

Proof.

Indeed, under the above conditions we have that $\mathcal{X}_\alpha(\omega) = [\mathcal{X}(\omega), \mathcal{X}(\omega)]$ and $(\mathbf{1}_{\{a\}})_\alpha = \{a\}$ for all $\alpha \in [0,1]$, whence $\left[f_{\mathcal{X}(\omega)}(\alpha,\lambda) - f_{\mathbf{1}_{\{a\}}}(\alpha,\lambda)\right]^2 = [\mathcal{X}(\omega) - a]^2$ for all $\omega \in \Omega$. Consequently,

$$MSD_S(\mathcal{X}, \mathbf{1}_{\{a\}}) = \mathrm{E}\left[(\mathcal{X} - a)^2\right] = \int_\Omega (\mathcal{X}(\omega) - a)^2 \, dP(\omega),$$

which coincides with the second moment of $\mathcal{X}$ about a. □

The *nonnegativeness* of the S-mean squared dispersion is established in the following theorem:

Theorem 1.2.2 (Nonnegativeness). *The S-MSD of $\mathcal{X}$ about $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ is nonnegative, whatever $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ and S satisfying the conditions of Definition* 0.1.5 *may be.*

The following property *characterizes degenerate fuzzy random variables* in terms of the "insensitiveness" of the S-MSD.

Theorem 1.2.3 (Minimality). $MSD_S\left(\mathcal{X}, \widetilde{A}\right) = 0$ *if, and only if, $\mathcal{X}$ is a degenerate fuzzy random variable at $\widetilde{A}$, that is, $\mathcal{X} = \widetilde{A}$ a.s. $[P]$, whatever $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ and S satisfying the conditions of Definition* 0.1.5 *may be.*

Proof.

Indeed, the nonnegativeness of $\left[D_S\left(\mathcal{X}(\,\cdot\,),\widetilde{A}\right)\right]^2$ on Ω ensures that the measure $MSD_S\left(\mathcal{X},\widetilde{A}\right)$ equals 0 if, and only if, $\left[D_S\left(\mathcal{X}(\omega),\widetilde{A}\right)\right]^2=0$ almost surely $[P]$, and this happens if, and only if, $\mathcal{X}=\widetilde{A}$ a.s. $[P]$ since D_S is a metric on $\mathcal{F}_c(\mathbb{R})$. □

The following result offers an *operational way to compute the S-MSD* (and, hence *the central S-MSD*) associated with a fuzzy random variable $\mathcal{X}$, in terms of the second moment (and variance) of certain real-valued random variables (variables $f_{\mathcal{X}}(\alpha,\lambda)$).

Theorem 1.2.4 (Simplified computation). *Whatever* $\widetilde{A}\in\mathcal{F}_c(\mathbb{R})$ *and* S *satisfying the conditions of Definition* 0.1.5 *may be, we have that*

$$MSD_S\left(\mathcal{X},\widetilde{A}\right)=\int_{[0,1]}\int_{[0,1]}\left(\int_{\Omega}\left[f_{\mathcal{X}(\omega)}(\alpha,\lambda)-f_{\widetilde{A}}(\alpha,\lambda)\right]^2\,d\,P(\omega)\right)\,d\,S(\lambda)\,d\,\alpha.$$

In particular,

$$\Delta^2_S(\mathcal{X})=\int_{[0,1]}\int_{[0,1]}\mathrm{Var}\left[f_{\mathcal{X}}(\alpha,\lambda)\right]\,d\,S(\lambda)\,d\,\alpha.$$

Proof.

In virtue of Theorem 1.1.2, we have that $\left[f_{\mathcal{X}(\cdot)}(\,\cdot\,,\,\cdot\,)\ -f_{\widetilde{A}}(\,\cdot\,,\,\cdot\,)\right]^2\in L^1(\Omega\times[0,1]\times[0,1],\mathcal{A}\otimes\mathcal{M}_{[0,1]}\otimes\mathcal{B}_{[0,1]},P\otimes m\otimes S)$, and the application of the Fubini Theorem proves the first result.

The particularization of this result to $\Delta^2_S(\mathcal{X})$ is concluded immediately by taking into account that for a fuzzy random variable $\mathcal{X}$ taking on values on $\mathcal{F}_c(\mathbb{R})$, we can assert for all $\alpha\in[0,1]$ that $\sup\widetilde{E}(\mathcal{X})_\alpha=E\left(\sup\mathcal{X}_\alpha\right)$, and $\inf\widetilde{E}(\mathcal{X})_\alpha=E\left(\inf\mathcal{X}_\alpha\right)$ (see López-Díaz & Gil 1998b). □

In the following lemma, we present a supporting result which will be used several times throughout this work:

Lemma 1.2.5. *Let $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ be an integrably bounded fuzzy random variable associated with the probability space $(\Omega, \mathcal{A}, P)$. Then,*

$$E\left(f_{\mathcal{X}}(\alpha,\lambda)\right) = f_{\widetilde{E}(\mathcal{X})}(\alpha,\lambda).$$

Proof.

Since $\mathcal{X}$ takes on values on $\mathcal{F}_c(\mathbb{R})$, we have that $\sup \widetilde{E}(\mathcal{X})_\alpha = E\left(\sup \mathcal{X}_\alpha\right)$ and $\inf \widetilde{E}(\mathcal{X})_\alpha = E\left(\inf \mathcal{X}_\alpha\right)$, and hence,

$$E\left(f_{\mathcal{X}}(\alpha,\lambda)\right) = \int_\Omega f_{\mathcal{X}(\omega)}(\alpha,\lambda)\, d\,P(\omega)$$

$$= \lambda E\left(\sup \mathcal{X}_\alpha\right) + (1-\lambda) E\left(\inf \mathcal{X}_\alpha\right) = f_{\widetilde{E}(\mathcal{X})}(\alpha,\lambda).$$

□

The following property states a *connection between the S-mean squared dispersion* associated with a fuzzy random variable about an element in $\mathcal{F}_c(\mathbb{R})$ *and the central S-mean squared dispersion.*

Theorem 1.2.6 (Connection between the S-MSD and the central S-MSD). *Whatever $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ and S satisfying the conditions of Definition* 0.1.5 *may be, we have that*

$$MSD_S\left(\mathcal{X}, \widetilde{A}\right) = \Delta_S^2(\mathcal{X}) + \left[D_S\left(\widetilde{E}(\mathcal{X}), \widetilde{A}\right)\right]^2.$$

Proof.

Indeed, for fixed arbitrary values of α and λ, we have that $f_{\mathcal{X}}(\alpha,\lambda)$ is a real-valued random variable and $f_{\widetilde{A}}(\alpha,\lambda) \in \mathbb{R}$, and since $\int_\Omega [f_{\mathcal{X}(\omega)}(\alpha,\lambda) - f_{\widetilde{A}}(\alpha,\lambda)]^2\, d\,P(\omega)$ is the mean squared error associated with $f_{\mathcal{X}}(\alpha,\lambda)$ about $f_{\widetilde{A}}(\alpha,\lambda)$, this integral coincides with $\operatorname{Var}\left[f_{\mathcal{X}}(\alpha,\lambda)\right] + [E\left(f_{\mathcal{X}}(\alpha,\lambda)\right) - f_{\widetilde{A}}(\alpha,\lambda)]^2$.

The previous lemma, and the application of Theorem 1.2.4 and the Fubini Theorem prove the result in this theorem. □

As a consequence from the latter theorem, we can formalize the result asserting that the $MSD_S\left(\mathcal{X}, \widetilde{A}\right)$ attains the minimum value over $\mathcal{F}_c(\mathbb{R})$ for the expected value of $\mathcal{X}$. In this way, the extension suggested in this work along with the definition of expected value by Puri & Ralescu *agree with Fréchet's approach.*

Theorem 1.2.7 (Agreement with Fréchet's approach). *Whatever* S *satisfying the conditions of Definition* 0.1.5 *may be, the function* $\mathcal{G}_{\mathcal{X},S} : \mathcal{F}_c(\mathbb{R}) \to [0,+\infty)$ *defined so that* $\mathcal{G}_{\mathcal{X},S}\left(\widetilde{A}\right) = MSD_S\left(\mathcal{X},\widetilde{A}\right)$ *for all* $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$*, is minimized for* $\widetilde{A} = \widetilde{E}(\mathcal{X})$.

Proof.

Indeed, on the basis of Theorem 1.2.6 we can conclude that $\mathcal{G}_{\mathcal{X},S}$ is minimized as $\left[D_S\left(\widetilde{E}(\mathcal{X}),\widetilde{A}\right)\right]^2$ is also minimized, and this happens as $\widetilde{A} = \widetilde{E}(\mathcal{X})$, since D_S is a metric on $\mathcal{F}_c(\mathbb{R})$. □

The next properties will concern the central S-MSD. First of all, on the basis of Theorem 1.2.6, we obtain an *alternative computation* for the central S-MSD as follows:

Theorem 1.2.8 (Alternative expression). *Whatever* S *satisfying the conditions of Definition* 0.1.5 *may be, we have that*

$$\Delta_S^2(\mathcal{X}) = MSD_S\left(\mathcal{X},\mathbf{1}_{\{0\}}\right) - \left[D_S\left(\widetilde{E}(\mathcal{X}),\mathbf{1}_{\{0\}}\right)\right]^2.$$

In statistical problems in which fuzzy random variables are used to model the processes leading to fuzzy experimental data, the most common operations between elements in $\mathcal{F}_c(\mathbb{R})$ will be the fuzzy addition $\oplus$ and the fuzzy product by a real number $\odot$. The properties we next establish analyze the behavior of the central S-MSD in connection with these operations.

The following result states the *"fuzzy" location invariance* of the central S-MSD, that is, that adding a constant fuzzy value to each value of a fuzzy random variable (i.e., shifting the fuzzy location of this variable), leaves the central S-MSD unchanged. Thus,

Theorem 1.2.9 (Location invariance). *Whatever* $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ *and* S *satisfying the conditions of Definition* 0.1.5 *may be, we have that*

$$\Delta_S^2(\mathcal{X} \oplus \widetilde{A}) = \Delta_S^2(\mathcal{X}).$$

Proof.

Indeed, for each $\alpha \in [0,1]$ and $\omega \in \Omega$ we have that $\left(\mathcal{X}(\omega) \oplus \widetilde{A}\right)_\alpha = \mathcal{X}_\alpha(\omega) + \widetilde{A}_\alpha$, and because of the properties of the variance of real-valued random variables and the definition of $f_{\mathcal{X}}(\alpha, \lambda)$, we conclude that

$$\mathrm{Var}\left[f_{\mathcal{X} \oplus \widetilde{A}}(\alpha, \lambda)\right] = \mathrm{Var}\left[f_{\mathcal{X}}(\alpha, \lambda) + f_{\widetilde{A}}(\alpha, \lambda)\right] = \mathrm{Var}\left[f_{\mathcal{X}}(\alpha, \lambda)\right]$$

whatever $\alpha \in [0,1]$ and $\lambda \in [0,1]$ may be.

Consequently, on the basis of Theorem 1.2.4 we obtain the result in the present one. □

The following result presents a discussion about the effects of multiplying a fuzzy random variable by a nonnegative constant real value (that is, performing a *change in the scale* of variable values); in this case, the central S-MSD is multiplied by the square of this constant. Thus,

Theorem 1.2.10 (Effects of a scale change). *Whatever* S *satisfying the conditions of Definition* 0.1.5 *and* $a \in [0, +\infty)$ *may be, we have that*

$$\Delta_S^2(a \odot \mathcal{X}) = a^2 \Delta_S^2(\mathcal{X}).$$

Proof.

Indeed, for each $\alpha \in (0,1]$ we have that $(a \odot \mathcal{X})_\alpha = a\,\mathcal{X}_\alpha$ and because of the properties of the variance of real-valued random variables, we conclude that

$$\mathrm{Var}\left[f_{a \odot \mathcal{X}}(\alpha, \lambda)\right] = \mathrm{Var}\left[a \cdot f_{\mathcal{X}}(\alpha, \lambda)\right] = a^2\,\mathrm{Var}\left[f_{\mathcal{X}}(\alpha, \lambda)\right],$$

whence on the basis of Theorem 1.2.4 we obtain the result in the present one. □

Remark 1.2.1. For $a \in (-\infty, 0)$, and if S satisfies the conditions assumed in Definition 0.1.5 and corresponds to a symmetrical distribution with respect to $\lambda = .5$, then the same conclusion of the latter theorem can be obtained, but unlike the classical case, the result does not remain valid for any $a \in \mathbb{R}$ when g is not symmetrical with respect to $\lambda = .5$.

The following result relates the *central S-MSD of the sum of two fuzzy random variables* defined on the same probability space with the central S-MSD of these variables. Thus,

Theorem 1.2.11 (Central S-MSD of the sum of two fuzzy random variables). *Whatever S satisfying the conditions of Definition* 0.1.5 *may be, we have that*

$$\Delta_S^2(\mathcal{X} \oplus \mathcal{Y}) = \Delta_S^2(\mathcal{X}) + \Delta_S^2(\mathcal{Y}) + 2\Delta_S(\mathcal{X}, \mathcal{Y}),$$

where

$$\Delta_S(\mathcal{X}, \mathcal{Y}) = \int_{[0,1]} \int_{[0,1]} \mathrm{Cov}\left[f_{\mathcal{X}}(\alpha, \lambda), f_{\mathcal{Y}}(\alpha, \lambda)\right] \, d\,S(\lambda)\, d\,\alpha,$$

where Cov *denotes the covariance of the corresponding real-valued random variables.*

Proof.

Indeed, in virtue of Theorem 1.2.4, we have that

$$\Delta_S^2(\mathcal{X} \oplus \mathcal{Y}) = \int_{[0,1]} \int_{[0,1]} \mathrm{Var}\left[f_{\mathcal{X} \oplus \mathcal{Y}}(\alpha, \lambda)\right] \, d\,S(\lambda)\, d\,\alpha,$$

and because of the properties of the variance of a sum of real-valued random variables, we have that

$$\mathrm{Var}\left[f_{\mathcal{X} \oplus \mathcal{Y}}(\alpha, \lambda)\right] = \mathrm{Var}\left[\lambda\left(\sup \mathcal{X}_\alpha + \sup \mathcal{Y}_\alpha\right) + (1-\lambda)\left(\inf \mathcal{X}_\alpha + \inf \mathcal{Y}_\alpha\right)\right]$$

$$= \mathrm{Var}\left[\lambda \sup \mathcal{X}_\alpha + (1-\lambda) \inf \mathcal{X}_\alpha\right] + \mathrm{Var}\left[\lambda \sup \mathcal{Y}_\alpha + (1-\lambda) + \inf \mathcal{Y}_\alpha\right]$$

$$+2\,\mathrm{Cov}\left[\lambda \sup \mathcal{X}_\alpha + (1-\lambda)\inf \mathcal{X}_\alpha, \lambda \sup \mathcal{Y}_\alpha + (1-\lambda) + \inf \mathcal{Y}_\alpha\right]$$
$$= \mathrm{Var}\left[f_{\mathcal{X}}\left(\alpha,\lambda\right)\right] + \mathrm{Var}\left[f_{\mathcal{Y}}\left(\alpha,\lambda\right)\right] + 2\,\mathrm{Cov}\left[f_{\mathcal{X}}\left(\alpha,\lambda\right), f_{\mathcal{Y}}\left(\alpha,\lambda\right)\right].$$

The linearity of the Lebesgue integral with respect to a finite measure on $[0,1]$ entails the result in this theorem. □

In particular, the *central S-MSD of the sum of two independent fuzzy random variables* equals the sum of the central S-MSD of these variables, that is,

Theorem 1.2.12 (Central S-MSD of the sum of two independent fuzzy random variables). *Whatever S satisfying the conditions of Definition* 0.1.5 *may be, if $\mathcal{X}$ and $\mathcal{Y}$ are independent fuzzy random variables, then*

$$\Delta_S^2(\mathcal{X} \oplus \mathcal{Y}) = \Delta_S^2(\mathcal{X}) + \Delta_S^2(\mathcal{Y}).$$

Proof.

Of course, if $\mathcal{X}$ and $\mathcal{Y}$ are independent fuzzy random variables, then for each $\alpha \in [0,1]$ the convex compact random sets $\mathcal{X}_\alpha$ and $\mathcal{Y}_\alpha$ are independent, whence $\sup \mathcal{X}_\alpha$ and $\sup \mathcal{Y}_\alpha$ are independent real-valued random variables, and $\inf \mathcal{X}_\alpha$ and $\inf \mathcal{Y}_\alpha$ are also independent real-valued random variables, and this fact guarantees the independence of the real-valued random variables $f_{\mathcal{X}}\left(\alpha,\lambda\right)$ and $f_{\mathcal{Y}}\left(\alpha,\lambda\right)$, so that on the basis of Theorem 1.2.11, we obtain the result in the present one. □

The following result is an *extension of the Tchebychev Inequality* for fuzzy random variables.

Theorem 1.2.13 (Extension of the Tchebychev Inequality). *Whatever $\varepsilon > 0$, $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ and S satisfying the conditions of Definition* 0.1.5 *may be, we have that*

$$P\left(\left\{\omega \in \Omega \,\Big|\, D_S\left(\mathcal{X}(\omega), \widetilde{A}\right) \leq \varepsilon\right\}\right) \geq 1 - \frac{MSD_S(\mathcal{X}, \widetilde{A})}{\varepsilon^2}.$$

Proof.

Indeed, in Theorem 1.1.1 we proved that $D_S\left(\mathcal{X}(\,\cdot\,), \widetilde{A}\right)$ is a real-valued random variable. By applying the Tchebychev inequality to this variable, we obtain the result. □

Finally, the last property in this section presents a result on the basis of which we can obtain a fuzzy random variable from another given one having the same fuzzy expected value and lower central S-MSD.

This method *extends the ideas in the Rao-Blackwell Theorem* (see, for instance, Dudewicz & Mishra 1988, Casella & Berger 1990) in the simplest version (that is, without using the notion of sufficiency), and its application to the fuzzy estimation of a fuzzy parameter is immediate.

Theorem 1.2.14 (Extension of the Rao-Blackwell Theorem). *Let $(\Omega, \mathcal{A}, P)$ be a probability space and let $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ and $\mathcal{Y} : \Omega \to \mathcal{F}_c(\mathbb{R})$ be two integrably bounded fuzzy random variables. Let $\sigma_{\mathcal{X}}$ and $\sigma_{\mathcal{Y}}$ be the σ-fields in $\mathcal{F}_c(\mathbb{R})$ induced from $\mathcal{A}$ by $\mathcal{X}$ and $\mathcal{Y}$, respectively (that is, $\sigma_{\mathcal{X}} = \{B \subset \mathcal{F}_c(\mathbb{R}) \,|\, \mathcal{X}^{-1}(B) \in \mathcal{A}\}$, $\sigma_{\mathcal{Y}} = \{B \subset \mathcal{F}_c(\mathbb{R}) \,|\, \mathcal{Y}^{-1}(B) \in \mathcal{A}\}$). Let $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ be the probability measures induced from P by $\mathcal{X}$ and $\mathcal{Y}$, respectively.*

Consider the product probability space $(\mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{X}} \otimes \sigma_{\mathcal{Y}}, P_{\mathcal{X}} \otimes P_{\mathcal{Y}})$, and let $\mathcal{Y}^* : \mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \to \mathcal{F}_c(\mathbb{R})$ be the integrably bounded fuzzy random variable such that $\mathcal{Y}^*(\tilde{x}, \tilde{y}) = \tilde{y}$ for all $\tilde{x}, \tilde{y} \in \mathcal{F}_c(\mathbb{R})$. Assume that when $\mathcal{X} = \tilde{x}$ the conditional probability distribution induced by $\mathcal{Y}$ is given by a regular conditional probability distribution in $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{Y}})$ denoted by $P_{\tilde{x}}$, that is,

- *$P_{\tilde{x}}$ is a probability measure on $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{Y}})$ for each $\tilde{x} \in \mathcal{X}(\Omega)$, and*
- *for each $B \in \sigma_{\mathcal{Y}}$ the mapping $g_B : \mathcal{X}(\Omega) \to [0,1]$ such that $g_B(\tilde{x}) = P_{\tilde{x}}(B)$ is a real-valued random variable associated with the measurable space $(\mathcal{F}_c(\mathbb{R}), \sigma_{\mathcal{X}})$ and satisfying that for all $A \in \sigma_{\mathcal{X}}$*

$$P(\mathcal{X} \in A,\, \mathcal{Y} \in B) = \int_A P_{\tilde{x}}(B)\, d\, P_{\mathcal{X}}.$$

Assume that $\widetilde{V} \in \mathcal{F}_c(\mathbb{R})$ *is a fuzzy parameter and* $\widetilde{E}(\mathcal{Y} \,|\, P) = \widetilde{V}$ *with* $\Delta^2_S(\mathcal{Y} \,|\, P) < \infty$, *and* $\varphi : \mathcal{X}(\Omega) \to \mathcal{F}_c(\mathbb{R})$ *is defined so that* $\varphi(\tilde{x}) = \widetilde{E}\,(\mathcal{Y} \,|\, \mathcal{X} = \tilde{x}) = \widetilde{E}\,(\mathcal{Y}^*_{\tilde{x}} \,|\, P_{\tilde{x}})$ *for all* $\tilde{x} \in \mathcal{X}(\Omega)$, *then* $\widetilde{E}(\varphi(\mathcal{X}) \,|\, P) = \widetilde{V}$, *and* $\Delta^2_S(\varphi(\mathcal{X}) \,|\, P) \leq \Delta^2_S(\mathcal{Y} \,|\, P)$, *with equality if, and only if,* $f_{\mathcal{Y}(\omega)}(\alpha, \lambda) = f_{\varphi(\mathcal{X}(\omega))}(\alpha, \lambda)$ *a.s.* $[m \otimes S \otimes P]$.

Proof.

Indeed, in accordance with the results of Theorem 0.3.1 in connection with the computation of iterated expectations of fuzzy random variables, irrespectively of the order of integration, we have that

$$\widetilde{E}(\mathcal{Y} \,|\, P) = \widetilde{E}\left(\widetilde{E}\,(\mathcal{Y}^*_{\tilde{x}} \,|\, P_{\tilde{x}}) \;\Big|\; P_{\mathcal{X}}\right).$$

If $\varphi(\tilde{x}) = \widetilde{E}\,(\mathcal{Y}^*_{\tilde{x}} \,|\, P_{\tilde{x}})$, then

$$\widetilde{V} = \widetilde{E}(\mathcal{Y} \,|\, P) = \widetilde{E}\left(\widetilde{E}\,(\mathcal{Y}^*_{\tilde{x}} \,|\, P_{\tilde{x}}) \;\Big|\; P_{\mathcal{X}}\right) = \widetilde{E}(\varphi(\mathcal{X}) \,|\, P).$$

On the other hand,

$$\Delta^2_S(\mathcal{Y} \,|\, P) = \int_\Omega \int_{[0,1]} \int_{[0,1]} \left[f_{\mathcal{Y}(\omega)}(\alpha, \lambda) - f_{\widetilde{V}}(\alpha, \lambda)\right]^2 d\, S(\lambda)\, d\, \alpha\, d\, P(\omega).$$

In virtue of the Fubini Theorem, we have that

$$\Delta^2_S(\mathcal{Y} \,|\, P) = \int_{[0,1]} \int_{[0,1]} \left(\int_\Omega \left[f_{\mathcal{Y}(\omega)}(\alpha, \lambda) - f_{\widetilde{V}}(\alpha, \lambda)\right]^2 d\, P(\omega)\right) d\, S(\lambda)\, d\, \alpha$$

$$= \int_{[0,1]} \int_{[0,1]} \left(\int_\Omega \left[f_{\mathcal{Y}(\omega)}(\alpha, \lambda) - f_{\varphi(\tilde{x})}(\alpha, \lambda)\right]^2 d\, P(\omega)\right) d\, S(\lambda)\, d\, \alpha$$

$$+ \int_{[0,1]} \int_{[0,1]} \left(\int_\Omega \left[f_{\varphi(\tilde{x})}(\alpha, \lambda) - f_{\widetilde{V}}(\alpha, \lambda)\right]^2 d\, P(\omega)\right) d\, S(\lambda)\, d\, \alpha$$

$$+\,2\int_{[0,1]}\int_{[0,1]}\left(\int_{\Omega}\left[f_{\mathcal{Y}(\omega)}(\alpha,\lambda)-f_{\varphi(\tilde{x})}(\alpha,\lambda)\right]\right.$$

$$\left.\cdot\left[f_{\varphi(\tilde{x})}(\alpha,\lambda)-f_{\widetilde{V}}(\alpha,\lambda)\right]\,d\,P(\omega)\right)\,d\,S(\lambda)\,d\,\alpha.$$

Since $\varphi(\tilde{x})=\widetilde{E}\left(\mathcal{Y}^*_{\tilde{x}}\,|\,P_{\tilde{x}}\right)$, then for all $\alpha\in[0,1]$ and $\lambda\in[0,1]$:

$$\int_{\Omega}\left[f_{\mathcal{Y}(\omega)}(\alpha,\lambda)-f_{\varphi(\tilde{x})}(\alpha,\lambda)\right]\left[f_{\varphi(\tilde{x})}(\alpha,\lambda)-f_{\widetilde{V}}(\alpha,\lambda)\right]\,d\,P(\omega)$$

$$=\int_{\mathcal{X}(\Omega)}\left[f_{\varphi(\tilde{x})}(\alpha,\lambda)-f_{\widetilde{V}}(\alpha,\lambda)\right]$$

$$\left(\int_{\mathcal{Y}(\Omega)}\left[f_{\tilde{y}}(\alpha,\lambda)-f_{\varphi(\tilde{x})}(\alpha,\lambda)\right]\,d\,P_{\tilde{x}}(\tilde{y})\right)\,d\,P_{\mathcal{X}}(\tilde{x}),$$

whence for all $\alpha\in[0,1]$ and $\lambda\in[0,1]$, we have that

$$\int_{\mathcal{Y}(\Omega)}\left[f_{\tilde{y}}(\alpha,\lambda)-f_{\varphi(\tilde{x})}(\alpha,\lambda)\right]\,d\,P_{\tilde{x}}(\tilde{y})$$

$$=\int_{\mathcal{Y}(\Omega)}f_{\tilde{y}}(\alpha,\lambda)\,d\,P_{\tilde{x}}(\tilde{y})-f_{\varphi(\tilde{x})}(\alpha,\lambda)$$

$$=\int_{\mathcal{Y}(\Omega)}\left[\lambda\sup\tilde{y}_\alpha+(1-\lambda)\inf\tilde{y}_\alpha\right]\,d\,P_{\tilde{x}}(\tilde{y})$$

$$-\left[\lambda\sup\left(\widetilde{E}\left(\mathcal{Y}^*_{\tilde{x}}\,|\,P_{\tilde{x}}\right)\right)_\alpha+(1-\lambda)\inf\left(\widetilde{E}\left(\mathcal{Y}^*_{\tilde{x}}\,|\,P_{\tilde{x}}\right)\right)_\alpha\right].$$

Since

$$\left(\widetilde{E}\left(\mathcal{Y}^*_{\tilde{x}}\,|\,P_{\tilde{x}}\right)\right)_\alpha=\left[\int_{\mathcal{Y}(\Omega)}\inf\tilde{y}_\alpha\,d\,P_{\tilde{x}}(\tilde{y}),\int_{\mathcal{Y}(\Omega)}\sup\tilde{y}_\alpha\,d\,P_{\tilde{x}}(\tilde{y})\right],$$

then for all $\alpha\in[0,1]$ and $\lambda\in[0,1]$ we can conclude that

$$\int_{\mathcal{Y}(\Omega)}\left[f_{\tilde{y}}(\alpha,\lambda)-f_{\varphi(\tilde{x})}(\alpha,\lambda)\right]\,d\,P_{\tilde{x}}(\tilde{y})=0$$

whatever $\tilde{x}\in\mathcal{X}(\Omega)$ may be.

Consequently,

$$\Delta_S^2(\mathcal{Y}\,|\,P) = \int_{[0,1]}\int_{[0,1]}\left(\int_\Omega \left[f_{\mathcal{Y}(\omega)}(\alpha,\lambda) - f_{\varphi(\tilde{x})}(\alpha,\lambda)\right]^2 d\,P(\omega)\right) d\,S(\lambda)\,d\,\alpha$$

$$+\int_{[0,1]}\int_{[0,1]}\left(\int_\Omega \left[f_{\varphi(\tilde{x})}(\alpha,\lambda) - f_{\widetilde{V}}(\alpha,\lambda)\right]^2 d\,P(\omega)\right) d\,S(\lambda)\,d\,\alpha.$$

Furthermore,

$$\Delta_S^2(\varphi(\mathcal{X})\,|\,P) = \int_{[0,1]}\int_{[0,1]}\left(\int_\Omega \left[f_{\varphi(\tilde{x})}(\alpha,\lambda) - f_{\widetilde{V}}(\alpha,\lambda)\right]^2 d\,P(\omega)\right) d\,S(\lambda)\,d\,\alpha,$$

so that

$$\Delta_S^2(\mathcal{Y}\,|\,P) \geq \Delta_S^2(\varphi(\mathcal{X})\,|\,P).$$

In addition, we have that $\Delta_S^2(\mathcal{Y}\,|\,P) = \Delta_S^2(\varphi(\mathcal{X})\,|\,P)$ if, and only if, $f_{\mathcal{Y}(\omega)}(\alpha,\lambda) = f_{\varphi(\mathcal{X}(\omega))}(\alpha,\lambda)$ a.s. $[m \otimes S \otimes P]$. □

In the following examples we illustrate the computation of the S-MSD and the central S-MSD, as well as their use to compare certain populations.

Example 1.2.1. When one wants to classify days in accordance with their temperature, usual classes correspond to linguistic "values" like COLD, COOL, NORMAL, WARM and HOT. According to this classification, the type of day in a given area, and during the summer, could be viewed as a fuzzy random variable $\mathcal{X}$ whose values are the preceding linguistic ones, which could be identified by means of some fuzzy numbers like those with support contained in $[8, 40]$ (measured in $^\circ$ C) and described in terms of S- and Π-curves (see, for instance, Cox 1994) as follows: COLD $= 1 - S(14, 19)$, COOL $= \Pi(14, 19, 24)$, NORMAL $= \Pi(19, 24, 29)$, WARM $= \Pi(24, 29, 34)$ and HOT $= S(29, 34)$ (see Figure 1.1), where

$$S(a,b)(t) = \begin{cases} 0 & \text{if } t \leq a \\ 2\left(\dfrac{t-a}{b-a}\right)^2 & \text{if } t \in \left[a, \dfrac{a+b}{2}\right] \\ 1 - 2\left(\dfrac{t-b}{b-a}\right)^2 & \text{if } t \in \left[\dfrac{a+b}{2}, b\right] \\ 1 & \text{otherwise,} \end{cases}$$

$$\Pi(a,(a+b)/2,b)(t) = \begin{cases} S(a,(a+b)/2) & \text{if } t \leq (a+b)/2 \\ 1 - S((a+b)/2, a) & \text{otherwise,} \end{cases}$$

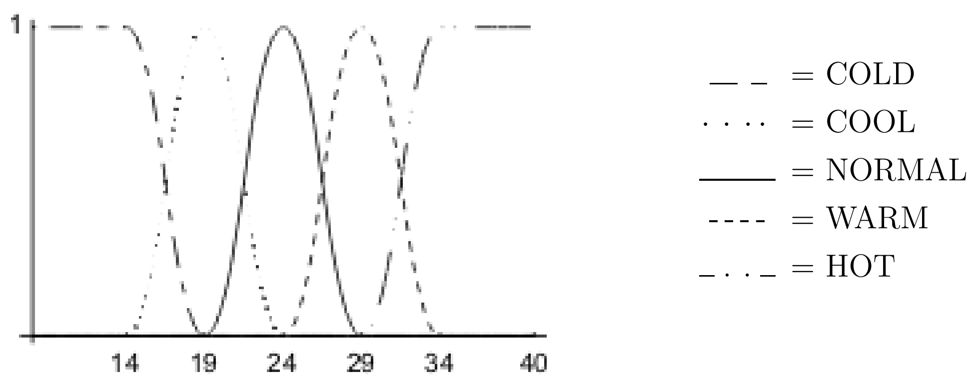


Fig. 1.1: Fuzzy values of the type of day

In this first example, we examine the S-mean squared dispersion associated with $\mathcal{X}$ by choosing the Lebesgue measure m on $[0,1]$, such that $L=0$, and $\overline{g}(\lambda) = 1$ if $\lambda \in [0,1]$. Consider the population Ω_1 of the 31 days of July of a given year and in certain area, and assume that during this month there are 2 COLD, 4 COOL, 7 NORMAL, 17 WARM and 1 HOT days. We want to compute the m-mean squared dispersion associated with $\mathcal{X}$ with respect to $\widetilde{A}_1 = \widetilde{E}(\mathcal{X})$, $\widetilde{A}_2 = 24$ and $\widetilde{A}_3 =$ NORMAL.

Since $\left[D_m\left(\text{COLD}, \widetilde{A}_1\right)\right]^2 = 67.084$, $\left[D_m\left(\text{COOL}, \widetilde{A}_1\right)\right]^2 = 47.549$, $\left[D_m\left(\text{NORMAL}, \widetilde{A}_1\right)\right]^2 = 3.598$, $\left[D_m\left(\text{WARM}, \widetilde{A}_1\right)\right]^2 = 9.646$, and $\left[D_m\left(\text{HOT}, \widetilde{A}_1\right)\right]^2 = 19.738$. Then,

$$\Delta_m^2(\mathcal{X}) = MSD_m\left(\mathcal{X}, \widetilde{A}_1\right) = 17.202.$$

In a similar way, we have that

$$MSD_m\left(\mathcal{X}, \widetilde{A}_2\right) = 22.995, \ MSD_m\left(\mathcal{X}, \widetilde{A}_3\right) = 20.8.$$

Example 1.2.2. In this second example, we compare the variability throughout the m-mean squared dispersion associated with $\mathcal{X}$ with respect to $\widetilde{E}(\mathcal{X})$, of the distribution of the type of days during July in the given year and area considered in Example 1.2.1 (population Ω_1) with a hypothetical uniform distribution during a month of 30 days (and which we will denote by Ω_2), such that there are 6 COLD, 6 COOL, 6 NORMAL, 6 WARM and 6 HOT days.

In this case, by proceding as in the Example 1.2.1, we obtain that in Ω_2

$$\Delta_m^2(\mathcal{X}) = MSD_m\left(\mathcal{X}, \widetilde{E}(\mathcal{X})\right) = 25.875,$$

so that $\mathcal{X}$ is clearly less "variable" with respect to its expected value in Ω_1 than in Ω_2.

Example 1.2.3. In this third example, we study the comparison throughout the central m-mean squared dispersion associated with the "distribution" of the type of days $\mathcal{X}$ during July (population Ω_1) and August (population Ω_3) of the given year and area considered in Example 1.2.1. Assume that during August there are 3 COOL, 6 NORMAL, 12 WARM and 10 HOT days.

Following a development similar to that in Example 1.2.1, we obtain that in Ω_3

$$\Delta_m^2(\mathcal{X}) = MSD_m\left(\mathcal{X}, \widetilde{E}(\mathcal{X})\right) = 12.862,$$

so that $\mathcal{X}$ is slightly more "variable" in July than in August.

1.3 Estimating the absolute variation in random samplings from finite populations

This section is focussed on the study of the problem of estimating the generalized measure of the absolute variation introduced in Section 1.1, in random samplings from finite populations.

To this purpose, we are going to check that it is possible to construct an unbiased estimator of this measure in random samplings with and without replacement, and we are going to determine the accuracy associated with this estimator in both samplings.

Consider a finite population Ω of N units, $\omega_1, \ldots, \omega_N,$ and let $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathrm{I\!R})$ be a fuzzy random variable associated with the measurable space defined on Ω, $(\Omega, \mathcal{P}(\Omega))$, and suppose that it is endowed with the uniform distribution.

Assume that a sample of size n is chosen at random and without replacement from Ω, υ denotes a generic simple random sample of size $n,$ and $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ are the units in it. Then, the *sample central S-mean squared dispersion* of $\mathcal{X}$ in υ is given by

$$\Delta_S^2 \left(\mathcal{X}[\upsilon]\right) = \frac{1}{n} \sum_{i=1}^{n} \left[D_S\left(\mathcal{X}(\omega_{\upsilon i}), \overline{\mathcal{X}}_n[\upsilon]\right)\right]^2,$$

where $\overline{\mathcal{X}}_n[\upsilon] = \dfrac{1}{n} \odot [\mathcal{X}(\omega_{\upsilon 1}) \oplus \cdots \oplus \mathcal{X}(\omega_{\upsilon n})]$ is the sample expected value of $\mathcal{X}$ in υ.

$\Delta_S^2 \left(\mathcal{X}[\,\cdot\,]\right)$ is a real-valued random variable associated with the probability space $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ (Υ_n being the space of the $C_{N,n} = \binom{N}{n}$ distinct possible random samples without replacement of size n from the given population, $\mathcal{P}(\Upsilon_n)$ being the associated power set, and $p[\upsilon] = 1/C_{N,n}$ for all $\upsilon \in \Upsilon_n$), and hence define a real-valued estimator of the *population central*

S-mean squared dispersion, which is given by

$$\Delta_S^2(\mathcal{X} \mid P) = \frac{1}{N} \sum_{j=1}^{N} \left[D_S \left(\mathcal{X}(\omega_j), \overline{\mathcal{X}} \right) \right]^2 ,$$

with $\overline{\mathcal{X}} = \dfrac{1}{N} \odot [\mathcal{X}(\omega_1) \oplus \cdots \oplus \mathcal{X}(\omega_N)]$.

To obtain from the sample measure an unbiased estimator of the population measure, the first one has to be revised in an immediate way, by taking into account the results for real-valued random variables and the property in Theorem 1.2.4. Thus,

Theorem 1.3.1. *In random sampling without replacement of size n from the population Ω, we have that the estimator $\widehat{\Delta_S^2}(\mathcal{X}[\,\cdot\,])$ which for a sample υ takes on the value*

$$\widehat{\Delta_S^2}(\mathcal{X}[\upsilon]) = \frac{(N-1)n}{N(n-1)} \Delta_S^2 \left(\mathcal{X}[\upsilon] \right) ,$$

is an unbiased estimator of $\Delta_S^2(\mathcal{X} \mid P)$.

Proof.

Indeed, since $f_{\mathcal{X}}(\alpha, \lambda)$ is a real-valued random variable for each $\alpha \in [0,1]$ and $\lambda \in [0,1]$, then $\mathrm{Var}\left[f_{\mathcal{X}}(\alpha, \lambda) \right]$ can be unbiasedly estimated in random sampling without replacement by means of the estimator $\dfrac{N-1}{N} s_n^2(f_{\mathcal{X}}(\alpha, \lambda)[\,\cdot\,])$, which for a sample υ takes on the value

$$\frac{N-1}{N} s_n^2 \left(f_{\mathcal{X}}(\alpha, \lambda) [\upsilon] \right) = \frac{N-1}{N(n-1)} \sum_{i=1}^{n} \left[f_{\mathcal{X}(\omega_{\upsilon i})}(\alpha, \lambda) - \overline{f_{\mathcal{X}}(\alpha, \lambda) [\upsilon]} \right]^2 ,$$

where

$$\overline{f_{\mathcal{X}}(\alpha, \lambda) [\upsilon]} = \frac{1}{n} \sum_{i=1}^{n} f_{\mathcal{X}(\omega_{\upsilon i})}(\alpha, \lambda) = f_{\overline{\mathcal{X}}_{n[\upsilon]}}(\alpha, \lambda) .$$

Consequently, in virtue of Theorem 1.2.4 $\Delta_S^2(\mathcal{X} \mid P)$ can be unbiasedly estimated in this sampling by means of the estimator which for the sample

υ takes on the value

$$\widehat{\Delta^2_S}\left(\mathcal{X}[\upsilon]\right) = \int_{[0,1]}\int_{[0,1]} \frac{N-1}{N} s_n^2\left(f_{\mathcal{X}}\left(\alpha,\lambda\right)[\upsilon]\right)\, d\,S(\lambda)\, d\,\alpha$$

$$= \frac{(N-1)n}{N(n-1)}\,\frac{1}{n}\sum_{i=1}^{n}\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{X}(\omega_{\upsilon i})}\left(\alpha,\lambda\right) - f_{\overline{\mathcal{X}}_n[\upsilon]}\left(\alpha,\lambda\right)\right]^2\, d\,S(\lambda)\, d\,\alpha$$

$$= \frac{(N-1)n}{N(n-1)}\Delta^2_S\left(\mathcal{X}[\upsilon]\right).$$

□

To assess the precision of the estimator $\widehat{\Delta^2_S}(\mathcal{X}[\,\cdot\,])$, we are going to obtain the associated mean squared error which coincides with its variance.

Theorem 1.3.2. *In random sampling without replacement of size n from the population Ω, we can conclude that, if $f = n/N$ denotes the sampling fraction, then*

$$\mathrm{Var}\left(\widehat{\Delta^2_S}(\mathcal{X}[\,\cdot\,])\right) = \frac{(1-f)}{n(n-1)N^2(N-2)(N-3)}$$

$$\cdot\Bigg\{2(N-n-1)(N-1)\sum_{j=1}^{N}\sum_{l=1}^{N}\left(\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{X}(\omega_j)}\left(\alpha,\lambda\right) - f_{\overline{\mathcal{X}}}\left(\alpha,\lambda\right)\right]\right.$$

$$\left.\cdot\left[f_{\mathcal{X}(\omega_l)}\left(\alpha,\lambda\right) - f_{\overline{\mathcal{X}}}\left(\alpha,\lambda\right)\right]\, d\,S(\lambda)\, d\,\alpha\right)^2$$

$$-[4(n+2)N-(6n-1)]N\,\left(\Delta^2_S(\mathcal{X}\,|\,P)\right)^2$$

$$+[(n-1)N^3 - 2N^2 - 3(n-3)N + (6n-8)]N\ \mathrm{Var}\left(\left[D_S(\mathcal{X},\overline{\mathcal{X}})\right]^2\,\Big|\,P\right)\Bigg\}.$$

Proof.

Indeed,

$$\mathrm{Var}\left(\widehat{\Delta^2_S}(\mathcal{X}[\,\cdot\,])\right) = \frac{(N-1)^2 n^2}{N^2(n-1)^2}$$

$$\cdot\,\mathrm{Var}\left(\frac{1}{n}\sum_{j=1}^{N}\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{X}(\omega_j)}\left(\alpha,\lambda\right) - f_{\overline{\mathcal{X}_n}}\left(\alpha,\lambda\right)\right]^2\, a_j\, d\,S(\lambda)\, d\,\alpha\right)$$

$$= \frac{(N-1)^2}{N^2(n-1)^2} \operatorname{Var}\left(\sum_{j=1}^{N} \int_{[0,1]} \int_{[0,1]} \left\{ \left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda) \right]^2 a_j \right.\right.$$

$$+ \left[f_{\overline{\mathcal{X}}}(\alpha,\lambda) - f_{\overline{\mathcal{X}}_n}(\alpha,\lambda) \right]^2 a_j - 2 \left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda) \right]$$

$$\left.\left. \cdot \left[f_{\overline{\mathcal{X}}}(\alpha,\lambda) - f_{\overline{\mathcal{X}}_n}(\alpha,\lambda) \right] a_j \right\} d\,S(\lambda)\, d\,\alpha \right)$$

$$= \frac{(N-1)^2}{N^2(n-1)^2} \operatorname{Var}\left(\sum_{j=1}^{N} \left[D_S\left(\mathcal{X}(\omega_j), \overline{\mathcal{X}} \right) \right]^2 a_j - n \left[D_S\left(\overline{\mathcal{X}}_n, \overline{\mathcal{X}} \right) \right]^2 \right),$$

where a_j is the Bernoulli variable associated with the presence of ω_j in each sample, and defined on the probability space $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$.

If we use the following notation,

$$\operatorname{Var}\left(\widehat{\Delta^2_S}(\mathcal{X}[\,\cdot\,]) \right) = \frac{(N-1)^2}{N^2(n-1)^2} \operatorname{Var}\left(\sum_{j=1}^{N} \mathrm{W}_{jj}\, a_j - \frac{1}{n} \sum_{j=1}^{N} \sum_{l=1}^{N} \mathrm{W}_{jl}\, a_j a_l \right),$$

where

$$\mathrm{W}_{jl} = \mathrm{W}_{lj} = \int_{[0,1]} \int_{[0,1]} \left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda) \right]$$

$$\cdot \left[f_{\mathcal{X}(\omega_l)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda) \right] d\,S(\lambda)\, d\,\alpha,$$

for $j, l \in \{1, \ldots, N\}$, then we have that

$$\operatorname{Var}\left(\widehat{\Delta^2_S}(\mathcal{X}[\,\cdot\,]) \right) = \frac{(N-1)^2}{N^2(n-1)^2} \left[\sum_{j=1}^{N} \sum_{l=1}^{N} \mathrm{W}_{jj} \mathrm{W}_{ll} \operatorname{Cov}(a_j,\, a_l) \right.$$

$$+ \frac{1}{n^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} \sum_{l=1}^{N} \sum_{l'=1}^{N} \mathrm{W}_{jj'} \mathrm{W}_{ll'} \operatorname{Cov}(a_j a_{j'},\, a_l a_{l'})$$

$$\left. - \frac{2}{n} \sum_{j=1}^{N} \sum_{j'=1}^{N} \sum_{l=1}^{N} \mathrm{W}_{jj'} \mathrm{W}_{ll} \operatorname{Cov}(a_l, a_j\, a_{j'}) \right]$$

$$= \frac{(N-1)^2}{N^2(n-1)^2} \left[\sum_{j=1}^{N} \mathrm{W}_{jj}^2 \operatorname{Var}(a_j) + \sum_{j=1}^{N} \sum_{\substack{l=1 \\ l \neq j}}^{N} \mathrm{W}_{jj} \mathrm{W}_{ll} \operatorname{Cov}(a_j, a_l) \right.$$

$$+\frac{1}{n^2}\Bigg\{\sum_{j=1}^{N} \mathrm{W}_{jj}^2 \,\mathrm{Var}(a_j^2) + 4\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jl}\mathrm{W}_{jj}\ \mathrm{Cov}(a_j\,a_l, a_j^2)$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jj}\mathrm{W}_{ll}\ \mathrm{Cov}(a_j^2,\ a_l^2) +\ 2\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jl}^2\ \mathrm{Var}(a_j\,a_l)$$

$$+\,2\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N} \mathrm{W}_{jj'}\mathrm{W}_{ll}\ \mathrm{Cov}(a_j\,a_{j'}, a_l^2)$$

$$+\,4\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N} \mathrm{W}_{jj'}\mathrm{W}_{jl}\ \mathrm{Cov}(a_j\,a_{j'}, a_j\,a_l)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\sum_{\substack{l'=1\\ l'\neq l,j,j'}}^{N} \mathrm{W}_{jj'}\mathrm{W}_{ll'}\ \mathrm{Cov}(a_j\,a_{j'}, a_l\,a_{l'})\Bigg\}$$

$$-\frac{2}{n}\Bigg\{\sum_{j=1}^{N}\mathrm{W}_{jj}^2\ \mathrm{Cov}(a_j, a_j^2) + 2\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jl}\mathrm{W}_{jj}\ \mathrm{Cov}(a_j, a_j\,a_l)$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jj}\mathrm{W}_{ll}\ \mathrm{Cov}(a_l, a_j^2) + \sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N} \mathrm{W}_{jj'}\mathrm{W}_{ll}\ \mathrm{Cov}(a_l, a_j\,a_{j'})\Bigg\}\Bigg].$$

Since, under the distribution p we have that

$$\mathrm{Var}(a_j) = \mathrm{Var}(a_j^2) = \mathrm{Cov}(a_j, a_j^2) = (1-f)\frac{n}{N}\ \text{ for each }\ j\in\{1,\,\ldots,\,N\},$$

$$\mathrm{Var}(a_j\,a_l) = \left(1-\frac{n(n-1)}{N(N-1)}\right)\frac{n(n-1)}{N(N-1)},$$

$$\mathrm{Cov}(a_j, a_j a_l) = \mathrm{Cov}(a_j\,a_l, a_j^2) = (1-f)\frac{n(n-1)}{N(N-1)},$$

$$\mathrm{Cov}(a_l, a_j^2) = \mathrm{Cov}(a_j^2, a_l^2) = \mathrm{Cov}(a_j, a_l) = -(1-f)\frac{n}{N(N-1)}$$

$$\text{for each }\ j\in\{1,\,\ldots,\,N\},\, l\in\{1,\,\ldots,\,N\}\setminus\{j\},$$

$$\mathrm{Cov}(a_j\,a_{j'}, a_j\,a_l) = \frac{(1-f)n(n-1)}{N(N-1)^2(N-2)}[(n-1)(N-2)-N],$$

$$\mathrm{Cov}(a_l, a_j\, a_{j'}) = \mathrm{Cov}(a_j\, a_{j'}, a_l^2) = \frac{-2(1-f)n(n-1)}{N(N-1)(N-2)}$$

for each $j \in \{1, \ldots, N\}$, $j' \in \{1, \ldots, N\}\backslash\{j\}$, $l \in \{1, \ldots, N\}\backslash\{j, j'\}$, and

$$\mathrm{Cov}(a_j\, a_{j'}, a_l\, a_{l'}) = \frac{(1-f)n(n-1)}{N(N-1)^2(N-2)(N-3)}[N(6-4n)+6(n-1)]$$

for each $j \in \{1, \ldots, N\}$, $j' \in \{1, \ldots, N\} \setminus \{j\}$, $l \in \{1, \ldots, N\} \setminus \{j, j'\}$, $l' \in \{1, \ldots, N\} \setminus \{j, j', l\}$, we obtain that

$$\begin{aligned}
\mathrm{Var}\left(\widehat{\Delta^2_S}(\mathcal{X}[\,\cdot\,])\right) &= \frac{(1-f)(N-1)^2}{nN^3}\Bigg\{\sum_{j=1}^{N}\mathrm{W}_{jj}^2 \\
&-\frac{1}{N-1}\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jj}\mathrm{W}_{ll} - \frac{4}{N-1}\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jl}\mathrm{W}_{jj} \\
&+\frac{2(N+n-1)}{(n-1)(N-1)^2}\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jl}^2 + \frac{4}{(N-1)(N-2)}\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll} \\
&+\frac{4}{(n-1)(N-1)^2(N-2)}\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{jl} \\
&+\frac{[N(6-4n)+6(n-1)]}{(n-1)(N-1)^2(N-2)(N-3)}\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\sum_{\substack{l'=1\\ l'\neq l,j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll'}\Bigg\}
\end{aligned}$$

$$\begin{aligned}
&= \frac{(1-f)}{n(n-1)N^2(N-2)(N-3)}\Bigg\{\frac{[N(6-4n)+6(n-1)]}{N}\left(\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jl}\right)^2 \\
&+4(n-2)(N-1)\sum_{j=1}^{N}\left(\sum_{l=1}^{N}\mathrm{W}_{jl}\right)^2 + 4[(n-1)(N-1)-n]\sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll} \\
&+2(N-n-1)(N-1)\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jl}^2 - (n-1)(N-1)^2\left(\sum_{j=1}^{N}\mathrm{W}_{jj}\right)^2 \\
&-4(N-1)[(n-1)(N-1)-n]\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jl}\mathrm{W}_{jj}
\end{aligned}$$

$$+[n(N-2)(N^2-3)-(N-1)(N^2+N-8)]\sum_{j=1}^{N}\mathrm{W}_{jj}^2\Bigg\},$$

and since $\sum_{l=1}^{N}\mathrm{W}_{jl}=0$ for all $j\in\{1,\ldots,N\}$, $\sum_{j=1}^{N}\mathrm{W}_{jj}=N\,\Delta_S^2(\mathcal{X}\,|\,P)$, and $\sum_{j=1}^{N}\mathrm{W}_{jj}^2=N\,E\left(\left[D_S(\mathcal{X},\overline{\mathcal{X}})\right]^4\,\Big|\,P\right)$, then we have the result in the present theorem. □

If, instead of adopting a choice at random and without replacement, we consider a choice at random and with replacement of n units from $\Omega=\{\omega_1,\ldots,\omega_N\}$, and υ represents a generic random sample with replacement of size n, and $\omega_{\upsilon 1},\ldots,\omega_{\upsilon n}$ are the units in υ, the *sample central S-mean squared dispersion* of $\mathcal{X}$ in υ is given by

$$\Delta_S^2\left(\mathcal{X}[\upsilon]\right)=\frac{1}{n}\sum_{i=1}^{n}\left[D_S\left(\mathcal{X}(\omega_{\upsilon i}),\overline{\mathcal{X}}_n[\upsilon]\right)\right]^2$$

$$=\frac{1}{n}\sum_{j=1}^{N}\left[D_S\left(\mathcal{X}(\omega_j),\overline{\mathcal{X}}_n[\upsilon]\right)\right]^2 t_j[\upsilon]=\frac{1}{n}\Bigg\{\sum_{j=1}^{N}\left[D_S\left(\mathcal{X}(\omega_j),\overline{\mathcal{X}}\right)\right]^2 t_j[\upsilon]$$

$$-\frac{1}{n}\sum_{j=1}^{N}\sum_{l=1}^{N}\Bigg(\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda)-f_{\overline{\mathcal{X}}}(\alpha,\lambda)\right]$$

$$\cdot\left[f_{\mathcal{X}(\omega_l)}(\alpha,\lambda)-f_{\overline{\mathcal{X}}}(\alpha,\lambda)\right]\,d\,S(\lambda)\,d\,\alpha\Bigg)t_j[\upsilon]t_l[\upsilon]\Bigg\},$$

where t_j is the real-valued random variable defined on the probabilistic space $(\Upsilon_n^w,\ \mathcal{P}(\Upsilon_n^w),\ p^w)$ (Υ_n^w being the space of the $CR_{N,n}=\binom{N+n-1}{n}$ distinct possible random samples with replacement of size n from the considered population, and $p^w[\upsilon]$ is the probability of choosing the sample $\upsilon\in\Upsilon_n^w$, which does not determine a uniform distribution on Υ_n^w for the considered sample) so that $t_j[\upsilon]$ is the "number of times that ω_j appears in υ".

$\Delta_S^2(\mathcal{X}[\,\cdot\,])$ is a real-valued random variable associated with the probability space $(\Upsilon_n^w,\mathcal{P}(\Upsilon_n^w),p^w)$, and hence it defines an estimator of the population central S-MSD.

As for the simple random sampling, to obtain an unbiased estimator of $\Delta^2_S(\mathcal{X} \mid P)$ from the sampling one, we need only to correct the latter as follows:

Theorem 1.3.3. *In random sampling with replacement of size n from the population Ω, we have that the estimator $\widehat{\Delta^2_S}^w(\mathcal{X}[\,\cdot\,])$, which for each sample υ takes on the value*

$$\widehat{\Delta^2_S}^w(\mathcal{X}[\upsilon]) = \frac{n}{(n-1)}\Delta^2_S(\mathcal{X}[\upsilon]),$$

is an unbiased estimator of $\Delta^2_S(\mathcal{X} \mid P)$.

Proof.

Indeed, as in Theorem 1.3.1, $\Delta^2_S(\mathcal{X})$ can be unbiasedly estimated in the random sampling with replacement of size n by means of the estimator

$$\widehat{\Delta^2_S}^w(\mathcal{X}[\upsilon]) = \int_{[0,1]}\int_{[0,1]} s^2_n\left(f_{\mathcal{X}}(\alpha,\lambda)[\upsilon]\right) d\,S(\lambda)\,d\,\alpha = \frac{n}{n-1}\Delta^2_S(\mathcal{X}[\upsilon]).$$

□

The mean squared error associated with $\widehat{\Delta^2_S}^w(\mathcal{X}[\,\cdot\,])$ in this sampling will be given by

Theorem 1.3.4. *In random sampling with replacement of size n from the population Ω, we have that*

$$\mathrm{Var}\left(\widehat{\Delta^2_S}^w(\mathcal{X}[\,\cdot\,])\right) = \frac{1}{n(n-1)^2N^3}$$

$$\cdot\Bigg\{2(n-1)(N-2)\sum_{j=1}^{N}\sum_{l=1}^{N}\left(\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda)\right]\right.$$

$$\left.\cdot\left[f_{\mathcal{X}(\omega_l)}(\alpha,\lambda) - f_{\overline{\mathcal{X}}}(\alpha,\lambda)\right]\,d\,S(\lambda)\,d\,\alpha\right)^2$$

$$+2(nN+n^2-2)N^2\left(\Delta^2_S(\mathcal{X} \mid P)\right)^2$$

$$+[(n-1)^2N - 2(n^2-5n+5)]N^2\;\mathrm{Var}\left(\left[D_S(\mathcal{X},\overline{\mathcal{X}})\right]^2 \,\Big|\, P\right)\Bigg\}.$$

Proof.

Indeed, by using the notation in Theorem 1.3.2,

$$\operatorname{Var}\left(\widehat{\Delta_S^2}^w(\mathcal{X}[\,\cdot\,])\right) = \frac{1}{(n-1)^2}\operatorname{Var}\left(\sum_{j=1}^N \mathrm{W}_{jj}t_j - \frac{1}{n}\sum_{j=1}^N\sum_{l=1}^N \mathrm{W}_{jl}t_jt_l\right)$$

$$= \frac{1}{(n-1)^2}\left[\sum_{j=1}^N \mathrm{W}_{jj}^2 \operatorname{Var}(t_j) + \sum_{j=1}^N \sum_{\substack{l=1\\ l\neq j}}^N \mathrm{W}_{jj}\mathrm{W}_{ll}\operatorname{Cov}(t_j, t_l)\right.$$

$$+\frac{1}{n^2}\left\{\sum_{j=1}^N \mathrm{W}_{jj}^2 \operatorname{Var}(t_j^2) + 4\sum_{j=1}^N\sum_{\substack{l=1\\ l\neq j}}^N \mathrm{W}_{jl}\mathrm{W}_{jj}\operatorname{Cov}(t_j\, t_l, t_j^2)\right.$$

$$+\sum_{j=1}^N\sum_{\substack{l=1\\ l\neq j}}^N \mathrm{W}_{jj}\mathrm{W}_{ll}\operatorname{Cov}(t_j^2,\, t_l^2) + 2\sum_{j=1}^N\sum_{\substack{l=1\\ l\neq j}}^N \mathrm{W}_{jl}^2\operatorname{Var}(t_j\, t_l)$$

$$+2\sum_{j=1}^N\sum_{\substack{j'=1\\ j'\neq j}}^N\sum_{\substack{l=1\\ l\neq j,j'}}^N \mathrm{W}_{jj'}\mathrm{W}_{ll}\operatorname{Cov}(t_j\, t_{j'}, t_l^2)$$

$$+4\sum_{j=1}^N\sum_{\substack{j'=1\\ j'\neq j}}^N\sum_{\substack{l=1\\ l\neq j,j'}}^N \mathrm{W}_{jj'}\mathrm{W}_{jl}\operatorname{Cov}(t_j\, t_{j'}, t_j\, t_l)$$

$$\left.+\sum_{j=1}^N\sum_{\substack{j'=1\\ j'\neq j}}^N\sum_{\substack{l=1\\ l\neq j,j'}}^N\sum_{\substack{l'=1\\ l'\neq l,j,j'}}^N \mathrm{W}_{jj'}\mathrm{W}_{ll'}\operatorname{Cov}(t_j\, t_{j'}, t_l\, t_{l'})\right\}$$

$$-\frac{2}{n}\left\{\sum_{j=1}^N \mathrm{W}_{jj}^2\operatorname{Cov}(t_j,\, t_j^2) + 2\sum_{j=1}^N\sum_{\substack{l=1\\ l\neq j}}^N \mathrm{W}_{jl}\mathrm{W}_{jj}\operatorname{Cov}(t_j, t_j\, t_l)\right.$$

$$\left.\left.+\sum_{j=1}^N\sum_{\substack{l=1\\ l\neq j}}^N \mathrm{W}_{jj}\mathrm{W}_{ll}\operatorname{Cov}(t_l, t_j^2) + \sum_{j=1}^N\sum_{\substack{j'=1\\ j'\neq j}}^N\sum_{\substack{l=1\\ l\neq j,j'}}^N \mathrm{W}_{jj'}\mathrm{W}_{ll}\operatorname{Cov}(t_j, t_j\, t_{j'})\right\}\right].$$

Since under the distribution p^w we have that

$$\operatorname{Var}(t_j) = \frac{n(N-1)}{N^2},$$

$$\mathrm{Var}(t_j^2) = \frac{n}{N^4}\{N^2(N-1) + (n-1)[2N(3N+2n-6) + (6-4n)]\}$$

$$\mathrm{Cov}(t_j, t_j^2) = \frac{n(N-1)(N+2(n-1))}{N^3}$$

for each $j \in \{1, \ldots, N\}$,

$$\mathrm{Var}(t_j\, t_l) = \frac{n(n-1)}{N^4}[N(N+2n-4) + (6-4n)],$$

$$\mathrm{Cov}(t_j, t_j\, t_l) = \frac{n(n-1)(N-2)}{N^3},$$

$$\mathrm{Cov}(t_j\, t_l, t_j^2) = \frac{n(n-1)}{N^4}[N(N+2n-6) + (6-4n)],$$

$$\mathrm{Cov}(t_j^2, t_l^2) = \frac{n[(n-1)(6-4n-4N) - N^2]}{N^4},$$

$$\mathrm{Cov}(t_l, t_j^2) = -\frac{2n(n-1)}{N^3},\ \mathrm{Cov}(t_j,\, t_l) = -\frac{n}{N^2}$$

for each $j \in \{1, \ldots, N\},\ l \in \{1, \ldots, N\} \setminus \{j\}$,

$$\mathrm{Cov}(t_j\, t_{j'}, t_j\, t_l) = \frac{n(n-1)[N(n-2) + (6-4n)]}{N^4},$$

$$\mathrm{Cov}(t_l, t_j\, t_{j'}) = \frac{-2(n-1)}{N^3},$$

$$\mathrm{Cov}(t_j\, t_{j'}, t_l^2) = \frac{n(n-1)[(6-4n) - 2N]}{N^4}$$

for each $j \in \{1, \ldots, N\},\ j' \in \{1, \ldots, N\}\setminus\{j\},\ l \in \{1, \ldots, N\}\setminus\{j, j'\}$, and

$$\mathrm{Cov}(t_j\, t_{j'}, t_l\, t_{l'}) = \frac{n(n-1)[(6-4n) - 2N]}{N^4}$$

for each $j \in \{1, \ldots, N\},\ j' \in \{1, \ldots, N\} \setminus \{j\},\ l \in \{1, \ldots, N\} \setminus \{j, j'\}$, $l' \in \{1, \ldots, N\} \setminus \{j, j', l\}$, we conclude that

$$\mathrm{Var}\left(\widehat{\Delta_S^2}^{w}(\mathcal{X}[\,\cdot\,])\right) = \frac{1}{n(n-1)^2 N^4}\Bigg[(n-1)(N-1)[(n-1)N^2$$

$$+(6-4n)(N-1)] \sum_{j=1}^{N} \mathrm{W}_{jj}^2$$

$$+ [(n-1)[2N-(6-4n)(N-1)] - N^2(n^2+1)]\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jj}\mathrm{W}_{ll}$$

$$-4(N-1)[N^2(n-1)+(6-4n)(N-1)]\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jl}\mathrm{W}_{jj}$$

$$+2(n-1)[N(N+2n-4)+(6-4n)]\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{W}_{jl}^2$$

$$+2(n-1)(6-4n)\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll}$$

$$+4(n-1)[N(n-2)+(6-4n)]\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{jl}$$

$$+(n-1)[(6-4n)-2N]\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\sum_{\substack{l'=1\\ l'\neq l,j,j'}}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll'}\Bigg]$$

$$= \frac{1}{n(n-1)N^3}\Bigg\{\frac{(6-4n)-2N}{N}\left(\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jl}\right)^2$$

$$+4n\sum_{j=1}^{N}\left(\sum_{l=1}^{N}\mathrm{W}_{jl}\right)^2 + 4\sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jj'}\mathrm{W}_{ll}$$

$$+2(N-2)\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jl}^2 - \frac{(6-4n)(n-1)+N(n^2+1)}{(n-1)}\left(\sum_{j=1}^{N}\mathrm{W}_{jj}\right)^2$$

$$-4[(N-2)(n+1)+6]\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{W}_{jl}\mathrm{W}_{jj}$$

$$+\frac{N[N(n-1)^2-2(n^2-5n+5)]}{n-1}\sum_{j=1}^{N}\mathrm{W}_{jj}^2\Bigg\},$$

and since $\sum_{l=1}^{N} \mathrm{W}_{jl} = 0$ for all $j \in \{1,\ldots,N\}$, $\sum_{j=1}^{N} \mathrm{W}_{jj} = N\,\Delta_S^2(\mathcal{X}\,|\,P)$, and $\sum_{j=1}^{N} \mathrm{W}_{jj}^2 = N\,E\left(\left[D_S(\mathcal{X},\overline{\mathcal{X}})\right]^4 \,\Big|\, P\right)$, then we have the result in the present theorem. □

The results we have just stated, can be used to compare the accuracy in the estimation of the population central S-MSD associated with different variables, to derive confidence intervals and testing hypotheses, and determining adequate sample sizes in the estimation.

However, since we do not know the exact distribution of the estimators $\widehat{\Delta_S^2}$ and $\widehat{\Delta_S^2}^w$, we will be forced to construct techniques based on Tchebychev's inequality or similar approximations, which would lead in most of the situations to very conservative procedures. On the other hand, these methods would involve unknown population values. Although in the present work we will not undertake the task, one could analyze the effect of the substitution of these population values by some unbiased estimates on the inferences derived from the application of these methods. These estimates would be easily defined from the values

$$E(a_j) = \frac{n}{N} \text{ for each } j \in \{1, \ldots, N\},$$

$$E(a_j\,a_l) = \frac{n(n-1)}{N(N-1)} \text{ for each } j \in \{1, \ldots, N\},\, l \in \{1, \ldots, N\} \setminus \{j\}.$$

$$E(t_j) = \frac{n}{N} \text{ for each } j \in \{1, \ldots, N\},$$

$$E(t_j\,t_l) = \frac{n(n-1)}{N^2} \text{ for each } j \in \{1, \ldots, N\},\; l \in \{1, \ldots, N\} \setminus \{j\}.$$

Some of the most important results in Section 1.3 will be useful for other later ones we will present in Section 1.5, in which we will illustrate their application.

1.4 Asymptotic behavior of the sample central S-mean squared dispersion in finite populations. Application to the development of approximate inferential procedures for the population value

In the preceding section we have shown that the population central S-MSD admits an unbiased estimator in the sampling of finite populations, even though the development of other exact inferential procedures concerning this population value are not feasible. In addition, as we have already pointed out, the techniques that could be constructed on the basis of the results in Section 1.3 and other complementary ones would be very conservative. However, the study we are now going to develop in this section will allow us to construct approximate inferential procedures which will give us very valuable conclusions.

As implications from this study and the inferential procedures derived from it, one should remark that it could be applied when we consider larger samples either chosen at random and with replacement from any population, or chosen at random and without replacement from a large population whose size is substantially bigger than the sample one. Under these conditions, the assumptions of independence and distribution identity for fuzzy random variables in the involved large random sample are (respectively) exact or approximately valid.

Let $\Omega = \{\omega_1, \ldots, \omega_N\}$ be a finite population and let $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ be a fuzzy random variable. On Ω we can define the probability space $(\Omega, \mathcal{P}(\Omega), P)$, where P denotes the probability measure associated with the uniform distribution on Ω (that is, with a random choice of units from Ω).

Assume that the variable $\mathcal{X}$ takes on r different values on Ω, $\tilde{x}_1^*, \ldots, \tilde{x}_r^*$, let $p_l = P(\{\omega \in \Omega \mid \mathcal{X}(\omega) = \tilde{x}_l^*\})$ for $l \in \{1, \ldots, r\}$, and let S be a measure on $[0, 1]$ satisfying the conditions in Definition 0.1.5. Then, the *population central S-MSD* of $\mathcal{X}$ on Ω can be expressed as follows:

$$\Delta_S^2(\mathcal{X} \mid \mathbf{p}) = \sum_{l=1}^{r} p_l \left[D_S \left(\tilde{x}_l^*, \widetilde{E}(\mathcal{X} \mid \mathbf{p}) \right) \right]^2 ,$$

with $\widetilde{E}(\mathcal{X} \mid \mathbf{p}) = \bigodot\!\!\!\!\!\sum_{l=1}^{r} p_l \odot \tilde{x}_l^*$ and $\mathbf{p} = (p_1, \ldots, p_{r-1})$ $(p_r = 1 - \sum_{l=1}^{r-1} p_l)$.

If a sample of size n is chosen at random from Ω and f_{nl} is the relative frequency of the value $\tilde{x}_l^*$ of $\mathcal{X}$ in the sample, the *sample central S-MSD* is given by

$$\Delta_S^2(\mathcal{X} \mid \mathbf{f}_n) = \sum_{l=1}^{r} f_{nl} \left[D_S \left(\tilde{x}_l^*, \overline{\mathcal{X}}_n \right) \right]^2 ,$$

with $\overline{\mathcal{X}}_n = \bigodot\!\!\!\!\!\sum_{l=1}^{r} f_{nl} \odot \tilde{x}_l^*$, and $\mathbf{f}_n = (f_{n1}, \ldots, f_{n(r-1)})$ $(f_{nr} = 1 - \sum_{l=1}^{r-1} f_{nl})$.

The following result establishes the asymptotic distribution of the sample central S-MSD, and its consistency (in fact, we can conclude that the sample central S-MSD is the best asymptotically normal estimator -see, for instance, Zacks 1971, pp. 248-249-) in estimating $\Delta_S^2(\mathcal{X})$.

Theorem 1.4.1. *For each $n \in \mathbb{N}$, consider n independent fuzzy random variables being identically distributed as $\mathcal{X}$ (that is, a simple random sample of size n from $\mathcal{X}$), defined over the finite population $\Omega = \{\omega_1, \ldots, \omega_N\}$ such that $P(\{\omega \in \Omega \mid \mathcal{X}(\omega) = \tilde{x}_l^*\}) = p_l$ with $p_l \in (0, 1)$, $l = 1, \ldots, r$, $\sum_{l=1}^{r} p_l = 1$. Let S be a normalized measure on $[0, 1]$ satisfying the conditions in Definition 0.1.5. Then,*

i) If $\mathbf{f}_n = (f_{n1}, \ldots, f_{n(r-1)}) \in [0, 1]^{r-1}$, with f_{nl} = the relative frequency of $\tilde{x}_l^*$ in the performance of the simple random sample of size n ($l =$

$1,\dots,r-1,\ f_{nr}=1-\sum_{l=1}^{r-1} f_{nl}),$ *and* $\Delta^2_S(\mathcal{X}\,|\,\mathbf{f}_n)$ *is the associated central S-MSD, then* $\{\Delta^2_S(\mathcal{X}\,|\,\mathbf{f}_n)\}_n$ *is a sequence of estimators of* $\Delta^2_S(\mathcal{X}\,|\,\mathbf{p})$, *which is strongly consistent, that is, as* $n\to\infty$ *we have that*

$$\Delta^2_S(\mathcal{X}\,|\,\mathbf{f}_n)\stackrel{a.s.}{\to}\Delta^2_S(\mathcal{X}\,|\,\mathbf{p}),$$

whatever $\mathbf{p}=(p_1,\dots,p_{r-1})$ $(p_1,\dots,p_{r-1}\in(0,1)$ *and* $\sum_{l=1}^{r-1}p_l<1)$ *may be.*

ii) $\{\sqrt{n}\,(\Delta^2_S(\mathcal{X}\,|\,\mathbf{f}_n)-\Delta^2_S(\mathcal{X}\,|\,\mathbf{p}))\}_n$ *is a sequence of real-valued random variables which converges in law as* $n\to\infty$ *to a one-dimensional normal distribution* $N(0,\sigma^2(\mathbf{p}))$, *with*

$$\sigma^2(\mathbf{p})=\mathrm{Var}\left(\left[D_S\left(\mathcal{X},\widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2\right),$$

whenever $\sigma^2(\mathbf{p})>0$.

iii) If $\sigma^2(\mathbf{p})=0$, *and there is a pair* (i,j) *with* $i,j\in\{1,\dots,r-1\}$ *such that*

$$h_{ij}=\frac{\partial^2\Delta^2_S(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i\partial p_j}$$

$$=\frac{\partial}{\partial p_i}\left(\left[D_S\left(\tilde{x}^*_j,\widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2-\left[D_S\left(\tilde{x}^*_r,\widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2\right)>0,$$

then $\{2n\,(\Delta^2_S(\mathcal{X}\,|\,\mathbf{f}_n)-\Delta^2_S(\mathcal{X}\,|\,\mathbf{p}))\}_n$ *is a sequence of real-valued random variables which converges in law as* $n\to\infty$ *to a linear combination of, at most,* $r-1$ *chi-square* χ^2_1 *independent random variables.*

Proof.

The fuzzy random variable $\mathcal{X}$ depends over Ω on the "parameter" $\mathbf{p}=(p_1,\dots,p_{r-1})$, with $p_1,\dots,p_{r-1}\in(0,1)$ and $p_r=1-\sum_{l=1}^{r-1}p_l\in(0,1)$, and the following conditions are satisfied

- The corresponding parameter space is $\mathbb{P} = [0,1]^{r-1}$, and the true parameter value of $\mathbf{p}$ is an interior point (that is, $\mathbf{p} \in (0,1)^{r-1}$). Furthermore, for different values of $\mathbf{p}$ the corresponding probability measures are obviously different.

- The "range" of $\mathcal{X}$ over Ω, $\{\tilde{x}_1^*, \ldots, \tilde{x}_r^*\}$, does not depend on $\mathbf{p}$.

- For each parameter value $\mathbf{p} \in (0,1)^{r-1}$, we have that if $A_l = \{\omega \in \Omega \mid \mathcal{X}(\omega) = \tilde{x}_l^*\}$ and $P(A_l \,|\, \mathbf{p}) = p_l$ for each $l \in \{1, \ldots, r\}$, then $\log P(A_l \,|\, \mathbf{p})$ can be derived three times with respect to the components p_i of $\mathbf{p}$, and in a neighborhood of $\mathbf{p}$ it satisfies that

$$\sum_{l=1}^{r} p_l \left| \frac{\partial^3}{\partial p_i \partial p_j \partial p_k} \log P(A_l \,|\, \mathbf{p}) \right| < \infty$$

whatever $i, j, k \in \{1, \ldots, r-1\}$ may be.

- The Fisher information matrix associated with the family $\wp$ (family of the probability measures corresponding to different values of $\mathbf{p}$ in $\mathbb{P}$) in $\mathbf{p}$ is given by

$$\mathrm{I}_{\mathcal{X}}^F(\mathbf{p}) = \left[\mathrm{I}_{ij}^F(\mathbf{p})\right] = \left[\sum_{l=1}^{r} p_l \, \frac{\partial \log p_l}{\partial p_i} \cdot \frac{\partial \log p_l}{\partial p_j}\right]$$

$$= \left[\frac{\delta_{ij}}{p_i} + \frac{1}{p_r}\right] \quad \text{with } \delta_{ij} = \text{Kronecker delta}$$

$(i, j \in \{1, \ldots, r-1\})$. $\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})$ is definite (since $\mathrm{I}_{ij}^F(\mathbf{p})$ is finite for any pair (i,j)), and it is positive definite (that is, the associated quadratic form is positive definite) for each $\mathbf{p} \in (0,1)^{r-1}$, since for any $i \in \{1, \ldots, r-1\}$ we have that

$$\begin{vmatrix} \mathrm{I}_{11}^F(\mathbf{p}) & \cdots & \mathrm{I}_{1i}^F(\mathbf{p}) \\ \vdots & \ddots & \vdots \\ \mathrm{I}_{i1}^F(\mathbf{p}) & \cdots & \mathrm{I}_{ii}^F(\mathbf{p}) \end{vmatrix} = \frac{p_1 + \cdots + p_i + p_r}{p_1 \cdots p_i \cdot p_r} > 0$$

(and this entails -see, for instance, Rao 1973, p. 36- that the Fisher information matrix is positive definite).

On the other hand, $\{\mathbf{f}_n\}_n$ is a sequence of estimators of $\mathbf{p}$ for the considered random sample, that is, a solution of the system of likelihood equations, since they are the maximum-likelihood estimators of $\mathbf{p}$. The sequence $\{\mathbf{f}_n\}_n$ is strongly consistent and it is asymptotically distributed in accordance with an (r-1)-dimensional $N\left(\mathbf{p}, \dfrac{\left[\mathrm{I}_{ij}^{F}(\mathbf{p})\right]^{-1}}{n}\right)$ as $n \to \infty$, where

$$\left[\mathrm{I}_{ij}^{F}(\mathbf{p})\right]^{-1} = [p_i(\delta_{ij} - p_j)],$$

since $\left|\left[\mathrm{I}_{ij}^{F}(\mathbf{p})\right]\right| = \left(\prod_{l=1}^{r} p_l\right)^{-1}$, and for $i, j \in \{1, \ldots, r-1\}$, the adjoint of the element $\mathrm{I}_{ij}^{F}(\mathbf{p})$ is ${}^{\alpha}\mathrm{I}_{ij}^{F}(\mathbf{p}) = p_i(\delta_{ij} - p_j)\left(\prod_{l=1}^{r} p_l\right)^{-1}$.

Moreover, $n(\mathbf{f}_n - \mathbf{p})\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})(\mathbf{f}_n - \mathbf{p})^t$ converges in law to a chi-squared distribution χ^2_{r-1} as $n \to \infty$ (see, for instance, Rao 1973, pp. 359-363, Serfling 1980, pp. 144-153, Lehmann 1983, pp. 403-436).

From these conditions we are going to prove the main results in this theorem.

i) Under the above conditions one can guarantee that $\mathbf{f}_n \overset{a.s.}{\to} \mathbf{p}$, whence according to the results in the Asymptotic Theory in Parametric Inference (see, for instance, Serfling 1980, p. 24) and since $\Delta_S^2(\mathcal{X} \mid \mathbf{p})$ is continuous in a neighborhood of $\mathbf{p}$, we can conclude that

$$\Delta_S^2(\mathcal{X} \mid \mathbf{f}_n) \overset{a.s.}{\to} \Delta_S^2(\mathcal{X} \mid \mathbf{p})$$

as $n \to \infty$, that is, $\{\Delta_S^2(\mathcal{X} \mid \mathbf{f}_n)\}_n$ is a strongly consistent sequence of estimators of $\Delta_S^2(\mathcal{X} \mid \mathbf{p})$.

ii) Because of the assumed conditions, if n is large enough one can expand $\Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n)$ in a neighborhood of $\mathbf{p}$. In this way, the first Taylor expansion is given by

$$\Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n) = \Delta_S^2(\mathcal{X} \,|\, \mathbf{p}) + \nabla\Delta_S^2(\mathbf{p})(\mathbf{f}_n - \mathbf{p})^t$$
$$+\frac{1}{2}(\mathbf{f}_n - \mathbf{p})H\Big(\Delta_S^2(\mathbf{p}_n^*)\Big)\,(\mathbf{f}_n - \mathbf{p})^t,$$

with $\nabla\Delta_S^2(\mathbf{p})$ = gradient vector of $\Delta_S^2(\mathcal{X} \,|\, \cdot)$ in $\mathbf{p}$, and $H(\Delta_S^2(\mathbf{p})) =$ $(r-1)\times(r-1)$ Hessian matrix such that

$$H\Big(\Delta_S^2(\mathbf{p}_n^*)\Big) = [h_{ij}] = \left[\left(\frac{\partial^2\Delta_S^2(\mathcal{X} \,|\, \mathbf{p})}{\partial p_i \partial p_j}\right)_{\mathbf{p}=\mathbf{p}_n^*}\right],$$

and with $\mathbf{p}_n^* \in \mathbb{P}$ such that $\|\mathbf{p}_n^* - \mathbf{p}\| \le \|\mathbf{f}_n - \mathbf{p}\|$.

Consequently,

$$\sqrt{n}\left[\Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n) - \Delta_S^2(\mathcal{X} \,|\, \mathbf{p})\right]$$
$$= \nabla\Delta_S^2(\mathbf{p})\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t + \frac{1}{2}(\mathbf{f}_n - \mathbf{p})H\Big(\Delta_S^2(\mathbf{p}_n^*)\Big)\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t.$$

The gradient vector is given by the $1 \times (r-1)$ matrix

$$\nabla\Delta_S^2(\mathbf{p}) = \left(\frac{\partial}{\partial p_1}\Delta_S^2(\mathcal{X} \,|\, \mathbf{p}) \;\cdots\; \frac{\partial}{\partial p_{r-1}}\Delta_S^2(\mathcal{X} \,|\, \mathbf{p})\right),$$

and for $i \in \{1, \ldots, r-1\}$ we have that

$$\frac{\partial}{\partial p_i}\Delta_S^2(\mathcal{X} \,|\, \mathbf{p}) = \frac{\partial}{\partial p_i}\left[\sum_{l=1}^{r-1} p_l \left\{\left[D_S\left(\tilde{x}_l^*, \widetilde{E}(\mathcal{X} \,|\, \mathbf{p})\right)\right]^2\right.\right.$$
$$\left.\left.-\left[D_S\left(\tilde{x}_r^*, \widetilde{E}(\mathcal{X} \,|\, \mathbf{p})\right)\right]^2\right\} + \left[D_S\left(\tilde{x}_r^*, \widetilde{E}(\mathcal{X} \,|\, \mathbf{p})\right)\right]^2\right]$$
$$= \left[D_S\left(\tilde{x}_i^*, \widetilde{E}(\mathcal{X} \,|\, \mathbf{p})\right)\right]^2 - \left[D_S\left(\tilde{x}_r^*, \widetilde{E}(\mathcal{X} \,|\, \mathbf{p})\right)\right]^2$$
$$+\sum_{l=1}^{r-1} p_l \left\{\frac{\partial}{\partial p_i}\left[D_S\left(\tilde{x}_l^*, \widetilde{E}(\mathcal{X} \,|\, \mathbf{p})\right)\right]^2\right.$$

$$-\frac{\partial}{\partial p_i}\left[D_S\left(\tilde{x}_r^*, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2\Bigg\} + \frac{\partial}{\partial p_i}\left[D_S\left(\tilde{x}_r^*, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2.$$

If $l \in \{1, \ldots, r\}$, then

$$\frac{\partial}{\partial p_i}\left[D_S\left(\tilde{x}_l^*, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2$$

$$= -2\int_{[0,1]}\int_{[0,1]}\left[f_{\tilde{x}_l^*}(\alpha,\lambda) - f_{\widetilde{E}(\mathcal{X}\,|\,\mathbf{p})}(\alpha,\lambda)\right] \cdot \left(\frac{\partial}{\partial p_i} f_{\widetilde{E}(\mathcal{X}\,|\,\mathbf{p})}(\alpha,\lambda)\right) d\,S(\lambda)\,d\,\alpha$$

$$= -2\int_{[0,1]}\int_{[0,1]}\left[f_{\tilde{x}_l^*}(\alpha,\lambda) - f_{\widetilde{E}(\mathcal{X}\,|\,\mathbf{p})}(\alpha,\lambda)\right] \cdot \left[f_{\tilde{x}_i^*}(\alpha,\lambda) - f_{\tilde{x}_r^*}(\alpha,\lambda)\right] d\,S(\lambda)\,d\,\alpha.$$

As $\sqrt{n}(\mathbf{f}_n - \mathbf{p})$ is asymptotically distributed as an $(r-1)$-dimensional normal $N\Big(\mathbf{0}, \left[\mathrm{I}\,^F_{\mathcal{X}}(\mathbf{p})\right]^{-1}\Big)$, because of the properties of the convergence in law (see, for instance, Serfling 1980, p. 26), we have that as $n \to \infty$

$$\nabla\Delta_S^2(\mathbf{p})\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \xrightarrow{\mathcal{L}} N\left(0,\, \nabla\Delta_S^2(\mathbf{p})\left[\mathrm{I}\,^F_{\mathcal{X}}(\mathbf{p})\right]^{-1}\left(\nabla\Delta_S^2(\mathbf{p})\right)^t\right),$$

whenever $\nabla\Delta_S^2(\mathbf{p})\left[\mathrm{I}\,^F_{\mathcal{X}}(\mathbf{p})\right]^{-1}(\nabla\Delta_S^2(\mathbf{p}))^t > 0$.

In virtue of the properties of the convergences in law and in probability (see, for instance, Serfling 1980, pp. 19, 24, 26), since $\|\mathbf{p}_n^* - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$ we have that as $n \to \infty$:

$$\left(\frac{\partial^2\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i \partial p_j}\right)_{\mathbf{p}=\mathbf{p}_n^*} \xrightarrow{p} \frac{\partial^2\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i \partial p_j},$$

for any $i, j \in \{1, \ldots, r-1\}$. On the other hand, $\mathbf{f}_n \xrightarrow{a.s.} \mathbf{p}$, and hence $\mathbf{f}_n \xrightarrow{p} \mathbf{p}$, as $n \to \infty$, whence

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p})H\Big(\Delta_S^2(\mathbf{p}_n^*)\Big) \xrightarrow{p} \mathbf{0},$$

and due to the fact that $\sqrt{n}(\mathbf{f}_n - \mathbf{p}) \xrightarrow{\mathcal{L}} N\left(\mathbf{0},\, \left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}\right)$, we obtain that

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p})H\Big(\Delta_S^2(\mathbf{p}_n^*)\Big)\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \xrightarrow{\mathcal{L}} 0,$$

and, consequently,

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p})H\Big(\Delta_S^2(\mathbf{p}_n^*)\Big)\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \xrightarrow{p} 0.$$

The above results guarantee that

$$\sqrt{n}\left[\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \Delta_S^2(\mathcal{X}\,|\,\mathbf{p})\right]$$

$$\xrightarrow{\mathcal{L}} N\left(0,\, \nabla\Delta_S^2(\mathbf{p})\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}\left(\nabla\Delta_S^2(\mathbf{p})\right)^t\right)$$

as $n \to \infty$, and therefore the result in *ii)* holds whenever $\sigma^2(\mathbf{p}) = \nabla\Delta_S^2(\mathbf{p})\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}\left(\nabla\Delta_S^2(\mathbf{p})\right)^t = \mathrm{W}^*\Sigma\mathrm{W}^{*t} > 0$, with W^* the $1 \times r$ matrix $(\mathrm{W}_1^*, \ldots, \mathrm{W}_r^*)$, where $\mathrm{W}_l^* = \left[D_S\left(\tilde{x}_l^*, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2$ and

$$\Sigma = \begin{pmatrix} p_1(1-p_1) & \cdots & -p_1p_r \\ \vdots & \ddots & \vdots \\ -p_rp_1 & \cdots & p_r(1-p_r) \end{pmatrix}$$

since

$$\sigma^2(\mathbf{p}) = \sum_{i=1}^{r-1}\sum_{j=1}^{r-1} p_i(\delta_{ij} - p_j)\left(\frac{\partial}{\partial p_i}\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})\right)\left(\frac{\partial}{\partial p_j}\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})\right)$$

$$= \mathrm{W}^*\Sigma\mathrm{W}^{*t} = \mathrm{Var}\left(\left[D_S\left(\mathcal{X}, \widetilde{E}(\mathcal{X}\,|\,\mathbf{p})\right)\right]^2\right).$$

iii) Matrix $\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})$ is positive definite, whence the quadratic forms associated with $\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})$ and $\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}$ are positive definite. Consequently, if $\sigma^2(\mathbf{p}) = \nabla\Delta_S^2(\mathbf{p})\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}\left(\nabla\Delta_S^2(\mathbf{p})\right)^t = 0$, then $\nabla\Delta_S^2(\mathbf{p}) = 0$.

If we now consider the second Taylor expansion of $\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n)$, in a neighborhood of $\mathbf{p}$, one can ensure that for n large enough we have that

$$\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) = \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}) + \frac{1}{2}(\mathbf{f}_n - \mathbf{p})H\Big(\Delta_S^2(\mathbf{p})\Big)\,(\mathbf{f}_n - \mathbf{p})^t$$

$$+\frac{1}{6}\sum_{i=1}^{r-1}\sum_{j=1}^{r-1}\sum_{k=1}^{r-1}\left(\frac{\partial^3\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i\partial p_j\partial p_k}\right)_{\mathbf{p}=\mathbf{p}_n^{**}}(f_{ni}-p_i)(f_{nj}-p_j)(f_{nk}-p_k),$$

with $\mathbf{p}_n^{**} \in \mathbb{P}$ such that $\|\mathbf{p}_n^{**} - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$. Therefore,

$$2n\left[\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \Delta_S^2(\mathcal{X}\,|\,\mathbf{p})\right]$$

$$= (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))H\Big(\Delta_S^2(\mathbf{p})\Big)\,\Big(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\Big)^t$$

$$+\frac{1}{3}\sum_{i=1}^{r-1}\sum_{j=1}^{r-1}\sum_{k=1}^{r-1}\left(\frac{\partial^3\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i\partial p_j\partial p_k}\right)_{\mathbf{p}=\mathbf{p}_n^{**}}$$

$$\cdot\,(f_{ni}-p_i)\left(\sqrt{n}(f_{nj}-p_j)\right)\left(\sqrt{n}(f_{nk}-p_k)\right).$$

Following the arguments used to prove *ii)*, we can conclude that

$$(f_{ni}-p_i)\left(\sqrt{n}(f_{nj}-p_j)\right)\left(\sqrt{n}(f_{nk}-p_k)\right) \xrightarrow{p} 0.$$

Furthermore (see, for instance, Serfling 1980, p. 25) we have that

$$\Big(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\Big)\,H\Big(\Delta_S^2(\mathbf{p})\Big)\,\Big(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\Big)^t \xrightarrow{\mathcal{L}} YH\Big(\Delta_S^2(\mathbf{p})\Big)\,Y^t,$$

where Y is a random vector with an $(r-1)$-dimensional normal distribution $N\Big(\mathbf{0},\Big[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\Big]^{-1}\Big)$.

Because of the assumption that some of the second derivatives $h_{ij}= \dfrac{\partial^2\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i\partial p_j}$ are positive, and since the rank of $H(\Delta_S^2(\mathbf{p}))$ is $r-1$, then vector Y can be expressed as $Y = ZB$ where Z is the $(r-1)$-dimensional random vector whose components are $r-1$ independent and identically distributed real-valued random variables, all of them having a standard normal distribution $N(0,1)$, and B is an $(r-1)\times(r-1)$ matrix such that $BB^t = \left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}$. Moreover, there exists a transformation $Z = UC$ in which C is an orthogonal matrix such that

$$Y H\Big(\Delta_S^2(\mathbf{p})\Big)\, Y^t = Z B H\Big(\Delta_S^2(\mathbf{p})\Big)\, B^t Z^t$$

$$= UCBH\Big(\Delta_S^2(\mathbf{p})\Big)\, B^t C^t U^t = \lambda_1 U_1^2 + \cdots + \lambda_q U_q^2,$$

$\lambda_1, \ldots, \lambda_q$ $(q \le r-1)$ being the nonnull eigenvalues of the $(r-1)\times(r-1)$ matrix $BH(\Delta_S^2(\mathbf{p}))\, B^t$, and where $U_1, \ldots, U_q$ are independent and identically distributed $N(0,1)$ variables, that is, a linear combination of independent and identically distributed chi-squared χ_1^2 variables (see, for instance, Rao 1973, pp. 186-188 and Serfling 1980, pp. 25, 128-130, for a review on the basic results for the preceding assertions).

□

Remark 1.4.1. It should be emphasized that (see, for instance, Rao 1973, pp. 186-188 and Serfling 1980, pp. 25, 128-130), since $\left|\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}\right| = p_1 \cdots p_r \neq 0$, the variable $YH(\Delta_S^2(\mathbf{p}))\, Y^t$ has a chi-squared distribution if, and only if,

$$H\Big(\Delta_S^2(\mathbf{p})\Big) \left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1} H\Big(\Delta_S^2(\mathbf{p})\Big) = H\Big(\Delta_S^2(\mathbf{p})\Big),$$

and should this be the case, the number of degrees of freedom for this distribution would coincide with the rank of $H(\Delta_S^2(\mathbf{p}))\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}$, and the regularity of $\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}$ entails that this rank equals that of $H(\Delta_S^2(\mathbf{p}))$.

Remark 1.4.2. The preceding results, along with the remaining results and procedures we will develop in this section, can be also directly applied over infinite populations, whenever the number of different variable values in these populations is finite, which is the most common situation in practice.

Remark 1.4.3. From Theorem 1.3.4 one can easily verify that

$$\sigma^2(\mathbf{p}) = \lim_{N\to\infty} \lim_{n\to\infty} \mathrm{Var}\left(\sqrt{n}\, \widehat{\Delta^2_S}^{w}(\mathcal{X}[\,\cdot\,]) \,\middle|\, p^w\right).$$

The property *ii)* in the latter theorem can be revised to calculate easily in practice the asymptotic variance of the estimator $\Delta^2_S(\mathcal{X}\,|\,\mathbf{f}_n)$ and to develop (although in an approximate way) inferences as the interval estimation and hypotheses testing. More precisely, when $\sigma^2(\mathbf{p})$ is replaced by its analogue estimate, $\sigma^2(\mathbf{f}_n)$, we obtain the following conclusion:

Theorem 1.4.2. *Under the conditions in Theorem* 1.4.1*, we have that*

$$\left\{\frac{\sqrt{n}\,(\Delta^2_S(\mathcal{X}\,|\,\mathbf{f}_n) - \Delta^2_S(\mathcal{X}\,|\,\mathbf{p}))}{\sqrt{\sigma^2(\mathbf{f}_n)}}\right\}_n$$

converges in law to a standard normal distribution $N(0,1)$ *as* $n \to \infty$, *whenever* $\sigma^2(\mathbf{p}) > 0$ *and* $\sigma^2(\mathbf{f}_n) > 0$.

Proof.

Since the components of $\nabla\Delta^2_S(\mathbf{p})$ and $\left[\mathrm{I}^F_{\mathcal{X}}(\mathbf{p})\right]^{-1}$ are continuous in a neighborhood of $\mathbf{p}$, then as $n \to \infty$ we have that

$$\nabla\Delta^2_S(\mathbf{f}_n) \xrightarrow{p} \nabla\Delta^2_S(\mathbf{p}),$$

and

$$\left[\mathrm{I}^F_{\mathcal{X}}(\mathbf{f}_n)\right]^{-1} \xrightarrow{p} \left[\mathrm{I}^F_{\mathcal{X}}(\mathbf{p})\right]^{-1}.$$

Consequently,

$$\sqrt{\sigma^2(\mathbf{f}_n)} \xrightarrow{p} \sqrt{\sigma^2(\mathbf{p})}$$

as $n \to \infty$, and since

$$\sqrt{n}\left(\Delta^2_S(\mathcal{X}\,|\,\mathbf{f}_n) - \Delta^2_S(\mathcal{X}\,|\,\mathbf{p})\right) \xrightarrow{\mathcal{L}} N(0,\sigma^2(\mathbf{p})),$$

then we have (see, for instance, Serfling 1980, p. 19) that

$$\frac{\sqrt{n}\,(\Delta^2_S(\mathcal{X}\,|\,\mathbf{f}_n) - \Delta^2_S(\mathcal{X}\,|\,\mathbf{p}))}{\sqrt{\sigma^2(\mathbf{f}_n)}} \xrightarrow{\mathcal{L}} N(0,1). \qquad \square$$

The study we have just developed allows us to perform some additional inferences on the population measure of the absolute variation.

Concerning the *interval estimation*, the procedure we now present will provide us with a range of possible values for the population central S-MSD which will cover the true value of this measure with (in this case approximately) a prescribed probability.

Theorem 1.4.2 allows us to state approximately the limits defining the above range when large samples are available, as follows:

Theorem 1.4.3. *Under the conditions in Theorems* 1.4.1 *and* 1.4.2*, the random interval*

$$\left[\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - z_\alpha\sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}},\ \Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) + z_\alpha\sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}}\right],$$

with $z_\alpha = (1-\alpha/2)$ *fractile of the* $N(0,1)$ *distribution, supplies for each sample of* n *independent observations from* $\mathcal{X}$ *confidence intervals of* $\Delta_S^2(\mathcal{X}\,|\,\mathbf{p})$ *with coefficient approximately equals to* $1-\alpha$ *(with* $\alpha \in [0,1]$*).*

Finally, and in an immediate way, we can derive the following *tests of hypotheses* for large samples

Theorem 1.4.4. *Under the conditions in Theorems* 1.4.1 *and* 1.4.2,

(i) To test at the significance level $\alpha \in [0,1]$ *the null hypothesis*

$$H_0 : \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}) = \delta_0$$

against the alternative one

$$H_1 : \Delta_S^2(\mathcal{X}\,|\,\mathbf{p}) \neq \delta_0,$$

H_0 *must be rejected whenever* $|\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \delta_0| > z_\alpha\sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$*. The p-value of the test is given, approximately, by*

$$p = 2\left[1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\left|\Delta_S^2(\mathcal{X}\,|\,\mathbf{f}_n) - \delta_0\right|\right)\right]$$

($\Phi =$ *being the distribution function of the* $N(0,\,1)$*).*

(ii) To test at the significance level α *the null hypothesis*

$$H_0 : \Delta_S^2(\mathcal{X} \,|\, \mathbf{p}) \geq \delta_0$$

against the alternative one

$$H_1 : \Delta_S^2(\mathcal{X} \,|\, \mathbf{p}) < \delta_0,$$

H_0 *must be rejected whenever* $\Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n) - \delta_0 < z_{2\alpha}\sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$. *The p-value of the test is given, approximately, by*

$$p = \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\left(\Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n) - \delta_0\right)\right).$$

(iii) To test at the significance level α *the null hypothesis*

$$H_0 : \Delta_S^2(\mathcal{X} \,|\, \mathbf{p}) \leq \delta_0$$

against the alternative one

$$H_1 : \Delta_S^2(\mathcal{X} \,|\, \mathbf{p}) > \delta_0,$$

H_0 *must be rejected whenever* $\Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n) - \delta_0 > z_{2\alpha}\sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$. *The p-value of the test is given, approximately, by*

$$p = 1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\left(\Delta_S^2(\mathcal{X} \,|\, \mathbf{f}_n) - \delta_0\right)\right).$$

In the next two sections in this chapter, we are going to examine two statistical applications of the generalized measure of the absolute variation of a fuzzy random variable introduced and analysed in the preceding sections. The S-MSD's role in these applications is similar to that of the mean squared error in the case of real-valued random variables. In Section 1.5, we will examine the problem of estimating a fuzzy parameter (the expected value) associated with a fuzzy random variable in random samplings with and without replacement from finite populations. In Section 1.6, we will analyze the problem of linear and functional regression between fuzzy random variables and we solve the problem under quite general conditions.

In Section 1.5 we will illustrate the use of some of the procedures presented in Theorem 1.4.4, combined with the results in Sections 1.3 and 1.5.

1.5 Application of the measure of absolute variation to estimating the expected value of fuzzy random variables in random samplings from finite populations

In this section we first consider the problem of estimating the expected value of a fuzzy random variable in a finite population, when the available information is that supplied by a sample chosen at random and with or without replacement from this population. This study will be complemented with the analysis of the precision of the estimation process.

In the Theory of Statistical Mathematics, we often assume that the population distribution of the involved real-valued random variables belongs to a known parameter family, and populations are not supposed to be necessarily finite, whereas Sampling Theory usually considers finite populations and the population distribution is completely unknown.

In the literature on fuzzy random variables, the study of models for the "distributions" of a fuzzy random variable is practically reduced to that of normal fuzzy random variables (see Puri & Ralescu 1985, Ralescu 1995c), so that in the framework of fuzzy random variables the development of studies is similar to those in Sampling Theory.

Regarding the estimation of the population expected value of a fuzzy random variable in random samplings with and without replacement, we are going to prove that the sample expected value of this variable is an unbiased estimator (the unbiasedness intended in the sense of the expected value of Puri & Ralescu 1986).

In the present study the S-MSD will be employed to measure the precision of the "fuzzy estimator" sample expected value of the "fuzzy parameter" population expected value.

If we consider a finite population of N units, $\Omega = \{\omega_1, \ldots, \omega_N\}$, on which the fuzzy random variable $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ associated with the probability space $(\Omega, \mathcal{P}(\Omega), P)$ (P being the uniform distribution on Ω) takes on the values $\mathcal{X}(\omega_1), \ldots, \mathcal{X}(\omega_N)$, then, in accordance with the expected value properties for $\mathcal{F}_c(\mathbb{R})$-valued random variables, the *population expected value of* $\mathcal{X}$ is given by

$$\overline{\mathcal{X}} = \widetilde{E}(\mathcal{X} \mid P) = \frac{1}{N} \odot \bigoplus_{j=1}^{N} \mathcal{X}(\omega_j),$$

that is, for each $\alpha \in [0,1]$:

$$\overline{\mathcal{X}}_\alpha = \left[\frac{1}{N} \sum_{j=1}^{N} \inf \mathcal{X}_\alpha(\omega_j), \frac{1}{N} \sum_{j=1}^{N} \sup \mathcal{X}_\alpha(\omega_j) \right].$$

Assume that a sample of size n is chosen at random and without replacement from Ω. Let υ denote a generic simple random sample of size n and let $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ be the units in υ. Then, the *sample expected value of* $\mathcal{X}$ in υ is given by

$$\overline{\mathcal{X}}_n[\upsilon] = \frac{1}{n} \odot \bigoplus_{i=1}^{n} \mathcal{X}(\omega_{\upsilon i}),$$

and $\overline{\mathcal{X}}_n$ defines a fuzzy random variable associated with the probability space $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ (see Section 1.3).

The fuzzy random variable $\overline{\mathcal{X}}_n$ represents a *fuzzy estimator* of $\overline{\mathcal{X}}$ and, the computation of the expected value of $\overline{\mathcal{X}}_n$ over the space $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ leads to $\overline{\mathcal{X}}$, so that we can conclude that $\overline{\mathcal{X}}_n$ is an *unbiased fuzzy estimator of* $\overline{\mathcal{X}}$ *in simple random sampling* (the unbiasedness being intended in the sense of the expectation defined by Puri & Ralescu 1986). In this way,

Theorem 1.5.1. *In random sampling without replacement of size n from the population Ω of N units, the $\mathcal{F}_c(\mathbb{R})$-valued estimator $\overline{\mathcal{X}}_n$ is unbiased to estimate $\overline{\mathcal{X}}$, that is, $\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p) = \overline{\mathcal{X}}$.*

Proof.

Indeed,

$$\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p) = \sum_{\upsilon \in \Upsilon_n} p[\upsilon] \odot \overline{\mathcal{X}}_n[\upsilon].$$

On the other hand,

$$\overline{\mathcal{X}}_n[\upsilon] = \frac{1}{n} \odot \sum_{j=1}^{N} a_j[\upsilon] \odot \mathcal{X}(\omega_j),$$

where a_j is the Bernoulli variable defined on $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ and associated with the presence of ω_j in each sample.

In virtue of the properties of the expected value of an $\mathcal{F}_c(\mathbb{R})$-valued fuzzy random variable (see, for instance, Puri & Ralescu 1986), we have that

$$\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p) = \frac{1}{n} \odot \sum_{j=1}^{N} E(a_j) \odot \mathcal{X}(\omega_j),$$

where $E(a_j) = n/N$ and, hence, $\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p) = \overline{\mathcal{X}}$. □

To *assess the accuracy or precision of a fuzzy estimator* of a fuzzy parameter, we will actually quantify the error of the estimation and will use for this purpose the S-mean squared dispersion of this estimator with respect to the parameter. The interest of this quantification lies in the possibility of comparing (with no need of rankings of fuzzy numbers) different sampling procedures or different estimators, and in the formalization of the problem of choosing the sample size to yield a prescribed precision.

In connection with the estimation of $\overline{\mathcal{X}}$ by means of $\overline{\mathcal{X}}_n$, we can establish that

Theorem 1.5.2. *In random sampling without replacement of size n from the population Ω of N units, the S-mean squared dispersion of $\overline{\mathcal{X}}_n$ with respect to $\overline{\mathcal{X}}$ is given by (if $f = n/N$ denotes the sampling fraction)*

$$MSD_S\left(\overline{\mathcal{X}}_n, \overline{\mathcal{X}}\right) = \Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p\right) = \frac{(1-f)N}{n(N-1)}\Delta_S^2\left(\mathcal{X} \,|\, P\right).$$

Proof.

Indeed, the properties of the S-MSD ensure that

$$\Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p\right) = \Delta_S^2\left(\frac{1}{n} \odot \sum_{j=1}^{N} a_j \odot \mathcal{X}(\omega_j)\right)$$

$$= \frac{1}{n^2}\sum_{j=1}^{N} \Delta_S^2\left(a_j \odot \mathcal{X}(\omega_j)\right) + \frac{2}{n^2}\sum_{j=1}^{N-1}\sum_{l=j+1}^{N} \Delta_S\left(a_j \odot \mathcal{X}(\omega_j), a_l \odot \mathcal{X}(\omega_l)\right).$$

For each $j \in \{1, \ldots, N\}$ we have (see Theorem 1.2.4) that

$$\Delta_S^2\left(a_j \odot \mathcal{X}(\omega_j)\right) = \int_{[0,1]}\int_{[0,1]} \mathrm{Var}\left[f_{a_j\odot\mathcal{X}(\omega_j)}(\alpha,\lambda)\right] d\,S(\lambda)\,d\,\alpha,$$

where for all $\alpha \in [0,1]$ and $\lambda \in [0,1]$

$$f_{a_j\odot\mathcal{X}(\omega_j)}(\alpha,\lambda) = \lambda \sup\left(a_j \odot \mathcal{X}(\omega_j)\right)_\alpha + (1-\lambda)\inf\left(a_j \odot \mathcal{X}(\omega_j)\right)_\alpha$$

$$= \lambda a_j \sup \mathcal{X}(\omega_j)_\alpha + (1-\lambda)a_j \inf \mathcal{X}(\omega_j)_\alpha = a_j f_{\mathcal{X}(\omega_j)}(\alpha,\lambda),$$

and, hence,

$$\mathrm{Var}\left[f_{a_j\odot\mathcal{X}(\omega_j)}(\alpha,\lambda)\right] = \mathrm{Var}(a_j)\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda)\right]^2,$$

whence

$$\Delta_S^2\left(a_j \odot \mathcal{X}(\omega_j)\right) = \mathrm{Var}(a_j)\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{X}(\omega_j)}(\alpha,\lambda)\right]^2 d\,S(\lambda)\,d\,\alpha.$$

On the other hand, for each $j, l \in \{1, \ldots, N\}$ with $l \neq j$, we have that

$$\begin{aligned}&\Delta_S\left(a_j \odot \mathcal{X}(\omega_j), a_l \odot \mathcal{X}(\omega_l)\right)\\ &= \int_{[0,1]} \int_{[0,1]} \operatorname{Cov}\left[f_{a_j \odot \mathcal{X}(\omega_j)}(\alpha, \lambda), f_{a_l \odot \mathcal{X}(\omega_l)}(\alpha, \lambda)\right] d\,S(\lambda)\,d\,\alpha,\end{aligned}$$

with

$$\begin{aligned}&\operatorname{Cov}\left[f_{a_j \odot \mathcal{X}(\omega_j)}(\alpha, \lambda), f_{a_l \odot \mathcal{X}(\omega_l)}(\alpha, \lambda)\right]\\ &= \operatorname{Cov}\left[a_j f_{\mathcal{X}(\omega_j)}(\alpha, \lambda), a_l f_{\mathcal{X}(\omega_l)}(\alpha, \lambda)\right]\\ &= \left[f_{\mathcal{X}(\omega_j)}(\alpha, \lambda)\right]\left[f_{\mathcal{X}(\omega_l)}(\alpha, \lambda)\right] \operatorname{Cov}(a_j, a_l),\end{aligned}$$

so that

$$\begin{aligned}&\Delta_S\left(a_j \odot \mathcal{X}(\omega_j), a_l \odot \mathcal{X}(\omega_l)\right)\\ &= \operatorname{Cov}(a_j, a_l) \int_{[0,1]} \int_{[0,1]} \left[f_{\mathcal{X}(\omega_j)}(\alpha, \lambda)\right]\left[f_{\mathcal{X}(\omega_l)}(\alpha, \lambda)\right] d\,S(\lambda)\,d\,\alpha.\end{aligned}$$

The substitution of the expressions for $\operatorname{Var}(a_j)$ and $\operatorname{Cov}(a_j, a_l)$ (see Theorem 1.3.2) indicates that

$$\begin{aligned}\Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p\right) &= \frac{1-f}{n} \int_{[0,1]} \int_{[0,1]} \Bigg[\frac{1}{N} \sum_{j=1}^{N} \left[f_{\mathcal{X}(\omega_j)}(\alpha, \lambda)\right]^2\\ &-\frac{2}{N(N-1)} \sum_{j=1}^{N-1} \sum_{l=j+1}^{N} \left[f_{\mathcal{X}(\omega_j)}(\alpha, \lambda)\right]\left[f_{\mathcal{X}(\omega_l)}(\alpha, \lambda)\right]\Bigg] d\,S(\lambda)\,d\,\alpha,\\ &= \frac{(1-f)N}{n(N-1)} \int_{[0,1]} \int_{[0,1]} \Bigg[\frac{1}{N} \sum_{j=1}^{N} \left[f_{\mathcal{X}(\omega_j)}(\alpha, \lambda)\right]^2\\ &-\frac{1}{N^2} \sum_{j=1}^{N} \sum_{l=1}^{N} \left[f_{\mathcal{X}(\omega_j)}(\alpha, \lambda)\right]\left[f_{\mathcal{X}(\omega_l)}(\alpha, \lambda)\right]\Bigg] d\,S(\lambda)\,d\,\alpha\\ &= \frac{(1-f)N}{n(N-1)} \int_{[0,1]} \int_{[0,1]} \operatorname{Var}\left[f_{\mathcal{X}(\omega)}(\alpha, \lambda)\right] d\,S(\lambda)\,d\,\alpha\\ &\quad \frac{(1-f)N}{n(N-1)} \Delta_S^2(\mathcal{X} \,|\, P).\end{aligned}$$

□

The result in Theorem 1.5.2 cannot help us usually in computing the sampling error in estimating $\overline{\mathcal{X}}$ by means of $\overline{\mathcal{X}}_n$, since $\Delta^2_S\left(\mathcal{X}\,|\,P\right)$ are commonly unknown. As a consequence, in most of the problems the above measure of the error is computed for (either in the approximation of the magnitude of the population central S-mean squared dispersion of $\mathcal{X}$, or in the determination of the sample size required to yield a desired precision), $\Delta^2_S\left(\mathcal{X}\,|\,P\right)$ must be also estimated from fuzzy sample data.

Theorem 1.3.1 allows us to state such an estimate as follows:

Theorem 1.5.3. *In random sampling without replacement of size n from the population Ω of N units, the estimator associating to sample υ the value*

$$\widehat{\Delta^2_S}\left(\overline{\mathcal{X}}_n[\upsilon]\right) = \frac{(N-1)n}{N(n-1)}\Delta^2_S\left(\overline{\mathcal{X}}_n[\upsilon]\right)$$

$$= \frac{(N-1)n}{N(n-1)}\int_{[0,1]}\int_{[0,1]} \mathrm{Var}\left(f_{\overline{\mathcal{X}}_n}(\alpha,\lambda)\right) d\,S(\lambda)d\,\alpha,$$

is unbiased to estimate $\Delta^2_S\left(\overline{\mathcal{X}}_n\,|\,p\right)$.

Obviously, an increase in sample size entails an increase in the precision of $\overline{\mathcal{X}}_n$ as an estimator of $\overline{\mathcal{X}}$, although the sampling costs will also increase. Assume that taking into account both aspects (precision and costs), we wish to estimate $\overline{\mathcal{X}}$ by means of the expected value of $\mathcal{X}$ in a simple random sample of size n from Ω, such that we are restricting to an *acceptable risk level* $\alpha \in [0,1]$, the probability that the error of estimating $\overline{\mathcal{X}}$ by means of $\overline{\mathcal{X}}_n$ is greater than some specified value or *tolerance* d. Thus, the aim is to seek the *minimum sample size* $n \in \mathbb{N}$ with which we can guarantee that

$$\mathrm{P}\Big(D_S(\overline{\mathcal{X}}_n,\overline{\mathcal{X}}) > d\Big) \leq \alpha.$$

To this purpose we are going to use the extension of Tchebychev Inequality in Theorem 1.2.13, in accordance with which

Theorem 1.5.4. *In random sampling without replacement of size n from the population Ω of N units, the sample size*

$$n = \left.\frac{N\Delta_S^2(\mathcal{X}\,|\,P)}{(N-1)d^2\alpha + \Delta_S^2(\mathcal{X}\,|\,P)}\right]$$

satisfies that $\mathrm{P}\Big(D_S(\overline{\mathcal{X}}_n,\, \overline{\mathcal{X}}) > d\Big) \leq \alpha$, *(where* $a]$ = *greatest integer part of the value* $a \in \mathbb{R}$*).*

Proof.

Indeed, on the basis of Theorem 1.2.13, a sufficient condition for $\mathrm{P}\Big(D_S(\overline{\mathcal{X}}_n, \overline{\mathcal{X}}) > d\Big) \leq \alpha$ is given by

$$\frac{\Delta_S^2(\overline{\mathcal{X}}_n\,|\,p)}{d^2} \leq \alpha,$$

that is,

$$\frac{1}{n} - \frac{1}{N} \leq \frac{(N-1)d^2\alpha}{N\Delta_S^2(\mathcal{X}\,|\,P)},$$

the minimum value in $\mathbb{N}$ satisfying the latter condition being

$$n = \left.\frac{N\Delta_S^2(\mathcal{X}\,|\,P)}{(N-1)d^2\alpha + \Delta_S^2(\mathcal{X}\,|\,P)}\right].$$

□

The sample size suggested in Theorem 1.5.4 is not really the minimum sample size satisfying our aim, since it has been obtained from a sufficient condition for this aim. In fact, and as in the real-valued case, Theorem 1.2.13 determines usually a quite conservative procedure for the choice of the sample size. Anyway, the size in Theorem 1.5.4 will depend on $\Delta_S^2(\mathcal{X}\,|\,P)$, which is commonly unknown. A proper way (whenever it is feasible) to approximate this unknown value, is to consider a preliminary simple random sample of a moderate size n_1, and using this sample to estimate $\Delta_S^2(\mathcal{X}\,|\,P)$ by means of the real-valued estimator $\widehat{\Delta_S^2}\Big(\overline{\mathcal{X}}_{n_1}[\,\cdot\,]\Big)$. If

$$n = \left.\frac{N\widehat{\Delta_S^2}\Big(\overline{\mathcal{X}}_{n_1}[\,\cdot\,]\Big)}{(N-1)d^2\alpha + \widehat{\Delta_S^2}\Big(\overline{\mathcal{X}}_{n_1}[\,\cdot\,]\Big)}\right] > n_1,$$

then we will complete the preliminary sample by choosing an additional simple random sample of size $n - n_1$.

Assume that a sample of size n at random and with replacement from Ω. The *sample expected value* of $\mathcal{X}$, $\overline{\mathcal{X}}_n$, defines now a fuzzy random variable associated with the probability space $(\Upsilon_n^w, \mathcal{P}(\Upsilon_n^w), p^w)$ (see Section 1.3).

As the simple random sampling, $\overline{\mathcal{X}}_n$ is an *unbiased fuzzy estimator of* $\overline{\mathcal{X}}$ in random sampling with replacement. In this way,

Theorem 1.5.5. *In random sampling with replacement of size n from the population* Ω *of* N *units, the* $\mathcal{F}_c(\mathbb{R})$*-valued estimator* $\overline{\mathcal{X}}_n$ *is unbiased to estimate* $\overline{\mathcal{X}}$*, that is,* $\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p^w) = \overline{\mathcal{X}}$*.*

Proof.

Indeed,

$$\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p^w) = \sum_{\upsilon \in \Upsilon_n^w} p^w[\upsilon] \odot \overline{\mathcal{X}}_n[\upsilon],$$

and

$$\overline{\mathcal{X}}_n[\upsilon] = \frac{1}{n} \odot \sum_{j=1}^{N} t_j[\upsilon] \odot \mathcal{X}(\omega_j),$$

where $t_j[\upsilon]$ is the number of times that ω_j appears in $\upsilon \in \Upsilon_n^w$, $j = 1, \ldots, N$ (see Theorem 1.3.4).

In virtue of the properties of the expected value of a fuzzy random variable, we have that

$$\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p^w) = \frac{1}{n} \odot \sum_{j=1}^{N} E(t_j) \odot \mathcal{X}(\omega_j),$$

where $E(t_j) = n/N$, and therefore $\widetilde{E}(\overline{\mathcal{X}}_n \,|\, p^w) = \overline{\mathcal{X}}$. □

Regarding the *precision* of $\overline{\mathcal{X}}_n$ in estimating $\overline{\mathcal{X}}$, we can now establish that

Theorem 1.5.6. *In random sampling with replacement of size n from the population Ω of N units, the S-mean squared dispersion of $\overline{\mathcal{X}}_n$ with respect to $\overline{\mathcal{X}}$ is given by*

$$MSD_S\left(\overline{\mathcal{X}}_n, \overline{\mathcal{X}}\right) = \Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p^w\right) = \frac{\Delta_S^2\,(\mathcal{X} \,|\, P)}{n}.$$

Proof.

Indeed, following arguments similar to those in Theorem 1.5.2, we have that

$$\Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p^w\right) = \Delta_S^2\left(\frac{1}{n} \odot \sum_{j=1}^{N} t_j \odot \mathcal{X}(\omega_j)\right),$$

and, the substitution of the expressions for $\mathrm{Var}(t_j)$ and $\mathrm{Cov}(t_j, t_l)$ (see Theorem 1.3.4), indicates that

$$\Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p^w\right) = \frac{1}{n}\int_{[0,1]}\int_{[0,1]}\left(\frac{N-1}{N^2}\sum_{j=1}^{N}\left[f_{\mathcal{X}(\omega_j)}\,(\alpha,\lambda)\right]^2\right.$$

$$\left.-\frac{2}{N^2}\sum_{j=1}^{N-1}\sum_{l=j+1}^{N}\left[f_{\mathcal{X}(\omega_j)}\,(\alpha,\lambda)\right]\left[f_{\mathcal{X}(\omega_l)}\,(\alpha,\lambda)\right]\right)d\,S(\lambda)\,d\,\alpha$$

$$= \frac{\Delta_S^2\,(\mathcal{X} \,|\, P)}{n}.$$

□

As an immediate consequence of Theorems 1.5.2 and 1.5.7 we can conclude that the simple random sampling is more precise than random sampling with replacement in estimating $\overline{\mathcal{X}}$.

The following result which states the estimation of $\Delta_S^2\left(\overline{\mathcal{X}}_n \,|\, p^w\right)$ in the random sampling with replacement is based in Theorem 1.3.3:

Theorem 1.5.7. *In random sampling with replacement of size n from the population Ω of N units, the estimator associating to sample υ the value*

$$\widehat{\Delta_S^2}^{w}\left(\overline{\mathcal{X}}_n[\upsilon]\right) = \frac{n}{n-1}\Delta_S^2\left(\overline{\mathcal{X}}_n[\upsilon]\right),$$

is unbiased to estimate $\Delta_S^2\left(\overline{\mathcal{X}} \,|\, p^w\right)$.

The problem of *choosing an appropriate sample size* can be now solved as follows:

Theorem 1.5.8. *In random sampling with replacement of size* n *from the population* Ω *of* N *units, the sample size*

$$n = \frac{\Delta^2_S(\mathcal{X}\,|\,P)}{d^2\alpha}\Bigg]$$

satisfies that $\mathrm{P}\Big(D_S(\overline{\mathcal{X}}_n, \overline{\mathcal{X}}) > d\Big) \leq \alpha.$

Before concluding this section, we are going to examine a result concerning the estimation of the population expected value $\overline{\mathcal{X}}$ in the random sampling with replacement from finite populations. This result illustrates the application of the extension of the Rao-Blackwell Theorem studied in Theorem 1.2.14 (see also Lubiano *et al.* 1999b) to the fuzzy estimation of a fuzzy parameter.

It should be pointed out that variable $\mathcal{Y}$ in Theorem 1.2.14 corresponds in this case to a real-valued random variable.

Assume that a sample of size n is chosen at random and with replacement from the population $\Omega = \{\omega_1, \ldots, \omega_N\}$. In accordance with the results we have obtained in the present section, $\overline{\mathcal{X}}_n$ defines a fuzzy unbiased estimator of $\overline{\mathcal{X}}$ in this sampling and $\Delta^2_S(\overline{\mathcal{X}}_n\,|\,p^w) = \Delta^2_S(\mathcal{X}\,|\,P)/n$.

The use of the extension of the Rao-Blackwell Theorem developed in Section 1.2, allows us to construct an unbiased estimator of $\overline{\mathcal{X}}$ with a lower central S-mean squared dispersion.

More precisely, consider an arbitrary ordering on the set $\bigcup_{k=1}^{n} \Upsilon_k$ of all simple random samples from Ω of size lower than or equal to n from Ω. Let $M = \mathrm{card}(\bigcup_{k=1}^{n} \Upsilon_k) = \sum_{i=1}^{n} \binom{N}{i}$, and let $\mathbb{M} = \{1, \ldots, M\}$.

One can state a real-valued random variable $Y : \Upsilon^w_n \to \mathbb{M}$ associated with the probability space $(\Upsilon^w_n, \mathcal{P}(\Upsilon^w_n), p^w)$ and defined so that, for each sample $\upsilon \in \Upsilon_n^w$, $Y(\upsilon)$ means the rank of the simple random sample of the distinct units in υ in the ordering considered on Ω.

If $m \in Y(\Upsilon_n^w)$ and m is the rank corresponding to a simple random sample y_m having k distinct units $\omega_1^*(y_m), \ldots, \omega_k^*(y_m)$, the (conditional given m) probability of $\omega_i^*(y_m)$ to belong to a sample in Υ_n^w for which Y takes on the value m equals $1/k$, $i = 1, \ldots, k$, so that

$$\widetilde{E}(\overline{\mathcal{X}}_n \,|\, m) = \frac{1}{k} \odot \sum_{i=1}^{k} \mathcal{X}(\omega_i^*(y_m)).$$

Consequently, $\widetilde{E}(\overline{\mathcal{X}}_n|m)$ is equivalent to the (conditional given m) expected value of the fuzzy estimator $\overline{\mathcal{X}}_\nu$ associating with each sample in Υ_n^w the expected value of $\mathcal{X}$ over the distinct units in this sample (which is a fuzzy random variable defined on the space $(\Upsilon_n^w, \mathcal{P}(\Upsilon_n^w), p^w)$, and whose distribution depends on the real-valued random variable *effective sample size* ν -see, for instance, Thompson 1992-), that is, we have that

$$\widetilde{E}(\overline{\mathcal{X}}_\nu \,|\, m) = \widetilde{E}(\overline{\mathcal{X}}_n \,|\, m).$$

Therefore, if we consider $\varphi(Y)$ such that $\varphi(m) = \widetilde{E}(\overline{\mathcal{X}}_n|m) = \widetilde{E}(\overline{\mathcal{X}}_\nu|m)$, and $\#y_m$ denotes the number of distinct units in y_m, we have that

$$\widetilde{E}(\varphi(Y) \,|\, p^w) = \sum_{k=1}^{n} \left[\sum_{m \in Y(\Upsilon_n^w) \,|\, \#y_m = k} \varphi(m) \odot P(Y = m \,|\, \nu = k) \right] \odot P_\nu(k)$$
$$= \sum_{k=1}^{n} \widetilde{E}(\overline{\mathcal{X}}_\nu \,|\, \nu = k) \odot P_\nu(k),$$

which in accordance with López-Díaz & Gil (1998b) coincides with $\widetilde{E}(\overline{\mathcal{X}}_\nu|p^w)$.

Since, in virtue of Theorem 1.5.1 we have that $\widetilde{E}(\overline{\mathcal{X}}_\nu|\nu = k) = \widetilde{E}(\mathcal{X})$ for $k = 1, \cdots, n$, then

$$\widetilde{E}(\varphi(Y) \,|\, p^w) = \overline{\mathcal{X}}.$$

Furthermore, Theorem 1.2.14 guarantees that

$$\Delta_S^2(\varphi(Y) \,|\, p^w) \leq \Delta_S^2 \left(\overline{\mathcal{X}}_\nu \,|\, p^w \right),$$

and

$$\Delta_S^2\left(\overline{\mathcal{X}}_\nu \,|\, p^w\right) = \int_{[0,1]}\int_{[0,1]} \operatorname{Var}\left[f_{\overline{\mathcal{X}}_\nu}(\alpha,\lambda)\right]\, d\,S(\lambda)\, d\,\alpha$$

$$= \int_{[0,1]}\int_{[0,1]} \left[\sum_{k=1}^{n} \operatorname{Var}\left[f_{\overline{\mathcal{X}}_\nu}(\alpha,\lambda) \,|\, \nu = k\right] P_\nu(k) \right.$$

$$\left. + \operatorname{Var}\left(E\left[f_{\overline{\mathcal{X}}_\nu}(\alpha,\lambda) \,|\, \nu\right]\right)\right] d\,S(\lambda)\, d\,\alpha$$

$$= \int_{[0,1]}\int_{[0,1]} \Big[\sum_{k=1}^{n} \operatorname{Var}\left(\overline{(f_{\mathcal{X}}(\alpha,\lambda))}_\nu \,|\, \nu = k\right) P_\nu(k)$$

$$+ \operatorname{Var}\left(E\left(\overline{(f_{\mathcal{X}}(\alpha,\lambda))}_\nu \,|\, \nu\right)\right)\Big]\, d\,S(\lambda)\, d\,\alpha,$$

where $\overline{(f_{\mathcal{X}}(\alpha,\lambda))}_\nu$ represents the sample mean of the real-valued random variable $f_{\mathcal{X}}(\alpha,\lambda)$ for the distinct units in the sample. Following the conclusions in Sampling Theory for real-valued random variables (see, for instance, Raj and Khamis 1958, Thompson 1992, pp. 20, 90), $\overline{(f_{\mathcal{X}}(\alpha,\lambda))}_\nu$ is an unbiased estimator of $\overline{f_{\mathcal{X}}(\alpha,\lambda)} = f_{\overline{\mathcal{X}}}(\alpha,\lambda)$ for any value of ν, whence $\operatorname{Var}\left(E\left(\overline{(f_{\mathcal{X}}(\alpha,\lambda))}_\nu|\nu\right)\right) = 0$.

On the other hand, and also in virtue of the results in Sampling Theory for real-valued random variables, we can conclude that

$$\operatorname{Var}\left(\overline{(f_{\mathcal{X}}(\alpha,\lambda))}_\nu \,|\, \nu = k\right) = \left(\frac{1}{k} - \frac{1}{N}\right) \operatorname{Var}\left(f_{\mathcal{X}}(\alpha,\lambda)\right),$$

and, hence,

$$\Delta_S^2\left(\overline{\mathcal{X}}_\nu \,|\, p^w\right) = \int_{[0,1]}\int_{[0,1]} \left[\sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{N}\right) \operatorname{Var}\left(f_{\mathcal{X}}(\alpha,\lambda)\right) P_\nu(k)\right] d\,S(\lambda)\, d\,\alpha$$

$$= \left[E\left(\frac{1}{\nu}\right) - \frac{1}{N}\right] \frac{N\Delta_S^2(\mathcal{X})}{N-1}.$$

Since, in accordance with Raj and Kharmis (1958), we have that $E\left(\dfrac{1}{\nu}\right) \leq \dfrac{1}{N} + \dfrac{N-1}{nN}$, with equality if, and only if, $n = 2$, then

$$\Delta_S^2(\varphi(Y)|p^w) \leq \Delta_S^2\left(\overline{\mathcal{X}}_n|p^w\right),$$

with equality if, and only if, $n = 2$.

Remark 1.5.1. The studies developed in this section on the fuzzy estimation of the fuzzy parameter $\overline{\mathcal{X}} = \widetilde{E}(\mathcal{X} \mid P)$ by means of the sample expected value $\overline{\mathcal{X}}_n$, can be completed with the particularization to finite populations of the result stated by Colubi *et al.* (1999), in accordance with which $\{\overline{\mathcal{X}}_n\}_n$ is a sequence of "strongly consistent" fuzzy estimators (this strong consistency being intended in the sense of the convergence with respect to the metric d_∞).

Some of the studies developed in the last three sections will be next illustrated by means of an example.

Example 1.5.1. A psychologist wants to make a survey of the favourite age for the inhabitants of a given city of $25,000$, to estimate the "mean preference".

For this purpose, the psychologist selects a preliminary simple random sample of $n_1 = 100$ inhabitants of this city, and asks them for the time of their life they consider has been/is/will be the best (assuming that they are not suffering from any illness during this period).

Assume the answers of the people polled are as follows:

15 'when I am/was YOUNG' ($\tilde{x}_1$),
2 'when I am/was VERY YOUNG' ($\tilde{x}_2$),
2 'when I am/was EXTREMELY YOUNG' ($\tilde{x}_3$),
3 'when I am/was FAIRLY YOUNG' ($\tilde{x}_4$),
10 'when I am/was MIDDLE-AGED' ($\tilde{x}_5$),
8 'when I am/was ABOVE MIDDLE-AGED' ($\tilde{x}_6$),
15 'when I am/was BELOW MIDDLE-AGED' ($\tilde{x}_7$),
10 'when I am/was AROUND MIDDLE-AGED' ($\tilde{x}_8$),
4 'when I am/was VERY MIDDLE-AGED' ($\tilde{x}_9$),
7 'when I am/was OLD' ($\tilde{x}_{10}$),
2 'when I am VERY OLD' ($\tilde{x}_{11}$),

1 'when I am EXTREMELY OLD' ($\tilde{x}_{12}$),
15 'when I am/was NOT OLD' ($\tilde{x}_{13}$),
6 'when I am/was FAIRLY OLD' ($\tilde{x}_{14}$).

The preceding values for the variable $\mathcal{X}$ = "time of life" over the population Ω of the $25,000$ inhabitants of the considered city, are clearly ill-defined and no exact boundaries are universally established for them. Suppose that to express the answers above, we use for instance the fuzzy regions with supports contained in $[0, 100]$ based on S-curves (see, for instance, Cox 1994), where along with linguistic modifiers (VERY, ABOVE, AROUND, etc.) which have been defined by using ideas in Bandemer and Gottwald (1995) (so that modifiers combine a "translation" and a functional change of the fuzzy value shape). In particular, we have assumed that $\tilde{x}_i$ are modeled as follows:

$$\tilde{x}_1 = 1 - S(25, 40),$$
$$\tilde{x}_2(t) = (\tilde{x}_1(t + 5))^2,$$
$$\tilde{x}_3(t) = (\tilde{x}_1(t + 10))^3,$$
$$\tilde{x}_4(t) = \sqrt{\tilde{x}_1(t - 5)},$$

$$\tilde{x}_5 = \begin{cases} S(25, 35) & \text{en } [0, 35] \\ 1 & \text{en } [35, 55] \\ 1 - S(55, 65) & \text{en otro caso} \end{cases}$$

$$\tilde{x}_6 = S(55, 65),$$
$$\tilde{x}_7 = 1 - S(25, 35),$$

$$\tilde{x}_8(t) = \begin{cases} (\tilde{x}_5(t - 5))^{0.7} & \text{si } t \geq 45 \\ (\tilde{x}_5(t + 5))^{0.7} & \text{en otro caso} \end{cases}$$

$$\tilde{x}_9(t) = \begin{cases} (\tilde{x}_5(t + 5))^{2} & \text{si } t \geq 45 \\ (\tilde{x}_5(t - 5))^{2} & \text{en otro caso} \end{cases}$$

$$\tilde{x}_{10} = S(60, 85),$$

$$\tilde{x}_{11}(t) = (\tilde{x}_{10}(t-5))^2,$$
$$\tilde{x}_{12}(t) = (\tilde{x}_{10}(t-10))^3,$$
$$\tilde{x}_{13}(t) = 1 - \tilde{x}_{10}(t),$$
$$\tilde{x}_{14}(t) = \sqrt{\tilde{x}_{10}(t+5)},$$

on [0,100], and $\tilde{x}_i$ are null on $\mathbb{R} \setminus [0,100]$ (see Figures 1.2, 1.3 and 1.4).

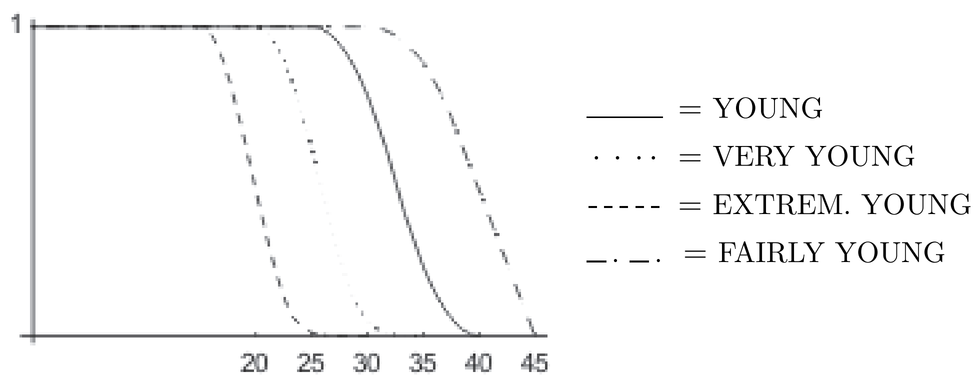


Fig. 1.2: Fuzzy value YOUNG and related linguistic modifiers

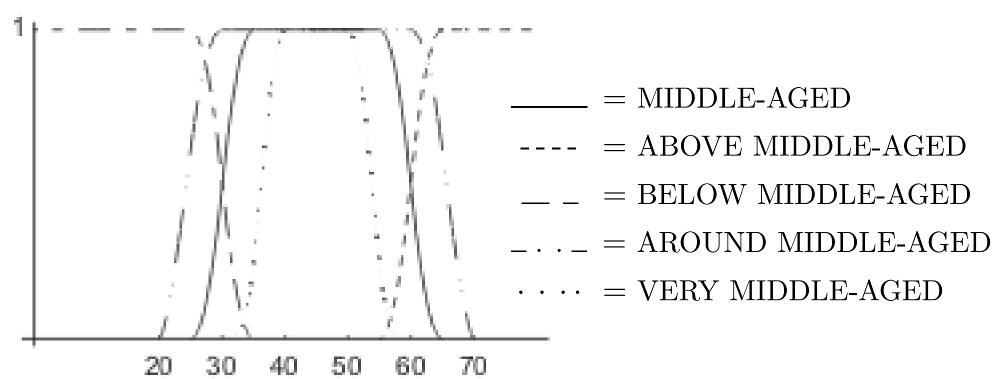


Fig. 1.3: Fuzzy value MIDDLE-AGED and related linguistic modifiers

On the basis of the available sample fuzzy information above, and in virtue of Theorem 1.3.1, we can compute $\Delta^2_{\mathcal{S}}(\mathcal{X})$ for the discretized weight

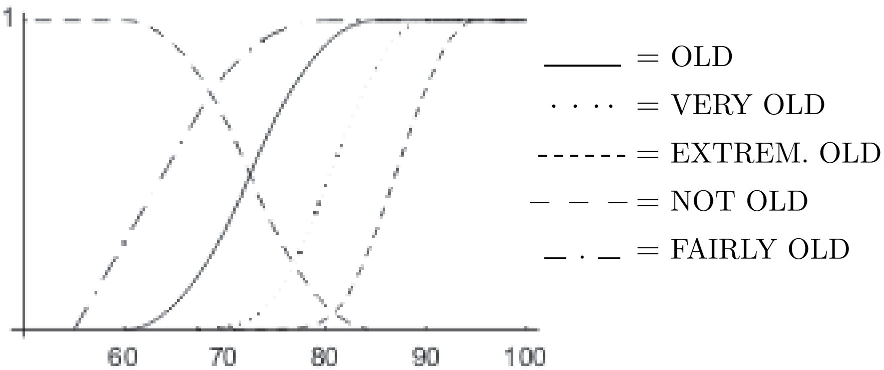


Fig. 1.4: Fuzzy value OLD and related linguistic modifiers

measure S_0 such that $L = 3,\ \lambda_1 = 0,\ \lambda_2 = .5,\ \lambda_3 = 1,\ k_1 = k_2 = k_3 = 1/3$, and $\overline{g}(\lambda) = 0$ if $\lambda \in (0,1) \setminus \{.5\}$ by means of the value

$$\widehat{\Delta^2_{S_0}}(\mathcal{X}[\upsilon_1]) = 371.075.$$

This preliminary estimate of $\Delta^2_{S_0}(\mathcal{X})$ allows us to determine as a proper sample size to estimate the population average preferred time of life, with an acceptable level $\alpha = .05$ and a tolerance $d =$'5 years', the value

$$n = \frac{25,000 \cdot 371.075}{24,999 \cdot 25 \cdot 0.05 + 371.075}\Bigg] = 294.$$

Assume now that a new simple random sample of size 194 is chosen from the population, it is added to the previous sample, and that answers to the question about the preferred time of their life are given for the whole sample υ by: 64 'when I am/was YOUNG', 10 'when I am/was VERY YOUNG', 4 'when I am/was EXTREMELY YOUNG', 39 'when I am/was FAIRLY YOUNG', 30 'when I am/was MIDDLE-AGED', 15 'when I am/was ABOVE MIDDLE-AGED', 17 'when I am/was BELOW MIDDLE-AGED', 34 'when I am/was AROUND MIDDLE-AGED', 19 'when I am/was APPROXIMATELY 40 TO 50 YEARS OLD', 18 'when I am/was

OLD', 4 'when I am VERY OLD', 3 'when I am EXTREMELY OLD', 26 'when I am/was NOT OLD', and 11 'when I am/was FAIRLY OLD'. The estimate of the mean preference in the population, which is given by the sample expected value for the whole sample of $n = 294$ inhabitants, is presented in Figure 1.5.

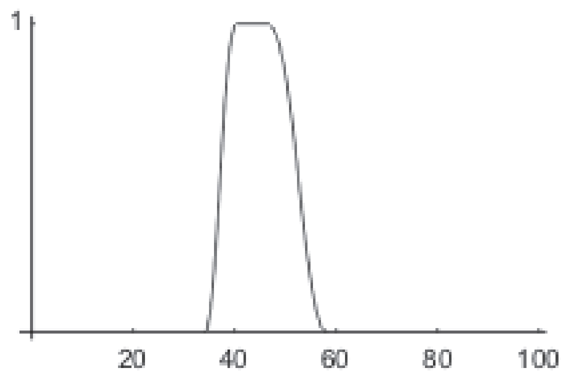


Fig. 1.5: Sample mean of the "favourite age"

Furthermore, "the variety of preferences" can be approximated by the final estimate of $\Delta^2_{S_0}(\mathcal{X})$ from the whole simple random sample, which is given by

$$\widehat{\Delta^2_{S_0}}(\mathcal{X}[\upsilon]) = 315.317.$$

In order to complete the study of the estimation of the population absolute variation of the preferred age, we can use the result in Theorem 1.4.3 to obtain a $100(1-\alpha)\%$ confidence interval, which will be (approximately) given for the considered sample υ by

$$\left[\Delta^2_{S_0}(\mathcal{X}[\upsilon]) - z_\alpha\sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}},\ \Delta^2_{S_0}(\mathcal{X}[\upsilon]) + z_\alpha\sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}}\right].$$

If we consider, for instance, $\alpha = .5$, the confidence interval would be given by

$$\left[315.317 - 1.96\sqrt{\frac{115978.04585}{294}},\ 315.317 + 1.96\sqrt{\frac{115978.04585}{294}}\right]$$

$$= [276.388,\ 354.245].$$

1.6 Applying the measure of the absolute variation to the linear and general regression with fuzzy random variables

Another interesting problem to which a measure finding inspiration in the S-mean squared dispersion can be applied, is that concerning the *prediction* of the value of a fuzzy random variable, given the value of another one by means of a certain transformation.

This problem is the aim of the so-called Regression Analysis with fuzzy data, which is a topic widely studied. In the literature on this topic we can distinguish different approaches, depending on the elements in the problem being affected by fuzziness, on the model to deal with the problem, and on the way considered to solve it.

Most of the research attention on the subject has been focussed on the approach which handles the fuzzy linear regression problem as a Linear Programming problem, and on the one based on a extension of the traditional Least Squares criterion.

In connection with the first approach, one should note the works of Tanaka *et al.* (1980, 1982, and others), in which the involved random variables are assumed to be real- or vectorial-valued, and the parameters of the linear relation are supposed to be fuzzy. The choice of the values for these fuzzy parameters, and the functions in the Linear Programming problem formalizing the aims of the regression problem, are crucial in solving it. Moskowitz & Kim (1993), and Redden & Woodall (1994, 1996) analyze the effects of such a choice on the solution of the corresponding Linear Programming problem (see, for instance, Corral *et al.* 1998, Kruse *et al.* 1999, for a short review on this and other problems).

The approach extending the Least Squares method, is mainly dependent on the extension considered for the distance between the observed values and the ones expected by the relation of the assumed model.

Diamond (1987, 1988, 1990) has developed a Least Squares method using the D_{S_2} metric when both the prediction and response variables take on values given by triangular fuzzy numbers, and the parameters of the fitting relation are real values. The projection on a cone or closed linear subspace, permits the use of the techniques of the Least Squares method for the real-valued case.

Näther (1997) and Körner (1997) have faced the linear regression problem in which the prediction variable is a real- (or vectorial-) valued random variable, and the response variable is a fuzzy random variable, the intercept parameter of the linear relation being assumed to be also a fuzzy random variable, whereas the slope parameter is real- or vectorial-valued.

Salas (1991) and Bertoluzza *et al.* (1995ab) have introduced the metric D_S and developed the extension of the Least Squares criterion when a finite number of pairs of fuzzy data and a certain choice of S is considered.

Guo & Chen (1992) have dealt with the fuzzy multiple regression problem by using triangular fuzzy numbers and the optimal relation being the one obtained by minimizing the squares of the differences of the functions characterizing the corresponding fuzzy data.

For a deeper and more detailed survey of the existing models and methods to manage the Regression Analysis with fuzzy elements, we can see the text by Kacprzyk & Fedrizzi (1992) and the recent review by Diamond & Tanaka (1998).

In this section we develop a study generalizing the one by Salas (1991) and Bertoluzza *et al.* (1995ab), in which fuzziness is assumed to arise in both the prediction and the response variables, and such a fuzziness will be represented by considering the notion of fuzzy random variables to model these variables and either a linear (with real-valued parameters) or a general functional relation. In this study, we present an extension of the Least Squares method, where the error between the observed and estimated fuzzy values will be measured by means of the mean value of the S-mean squared dispersion associated with the response fuzzy random variable about each of the values of the prediction fuzzy random variable.

For purposes of simplifying notations, we will employ the mappings $V_S : \mathcal{F}_c(\mathbb{R}) \to \mathbb{R}$, $W_S : \mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \to \mathbb{R}$ and $W'_S : \mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \to \mathbb{R}$ defined so that

$$V_S\left(\widetilde{A}\right) = \int_{[0,1]} \int_{[0,1]} f_{\widetilde{A}}(\alpha, \lambda)\; d\, S(\lambda)\, d\,\alpha,$$

$$W_S\left(\widetilde{A}, \widetilde{B}\right) = \int_{[0,1]} \int_{[0,1]} f_{\widetilde{A}}(\alpha, \lambda)\, f_{\widetilde{B}}(\alpha, \lambda)\; d\, S(\lambda)\, d\alpha,$$

and

$$W'_S\left(\widetilde{A}, \widetilde{B}\right) = \int_{[0,1]} \int_{[0,1]} f_{\widetilde{A}}(\alpha, 1-\lambda)\, f_{\widetilde{B}}(\alpha, \lambda)\; d\, S(\lambda)\, d\alpha.$$

Let $\mathcal{X}$ and $\mathcal{Y}$ be two integrably bounded fuzzy random variables associated with a probability space $(\Omega, \mathcal{A}, P)$, and firstly consider a *linear relation* of variable $\mathcal{X}$ with respect to $\mathcal{Y}$ as follows:

$$\mathcal{Y} = (a \odot \mathcal{X}) \oplus b,$$

where a and b are real numbers.

To explain the meaning of such a linear relation, we can consider the example in Figures 1.6 and 1.7.

Figure 1.6 shows the case in which we consider on one hand the two-dimensional fuzzy datum corresponding to the Cartesian product of the

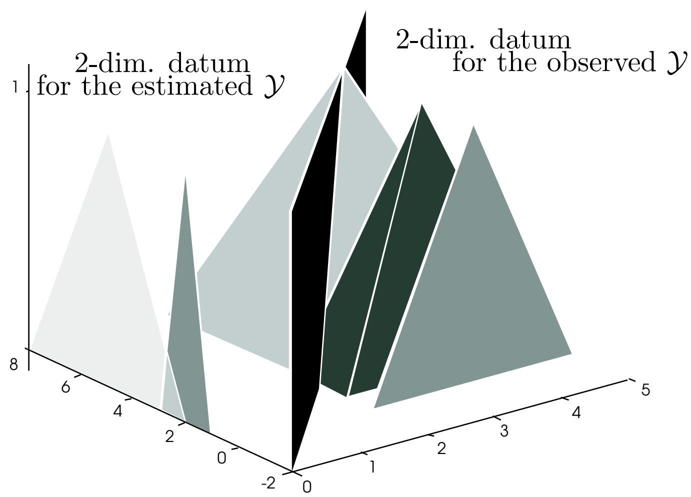


Fig. 1.6: Two-dimensional fuzzy data for the observed and estimated (by the linear relation $y = 2x - 2$) values of $\mathcal{Y}$

(observed) value of $\mathcal{X}$ (given by the triangular fuzzy number Tri(2,3.5,5)) and the observed value of $\mathcal{Y}$ (Tri(1,2,3)) and, on the other hand, the two-dimensional fuzzy datum corresponding to the Cartesian product of the same value of $\mathcal{X}$ and the estimated value of $\mathcal{Y}$ from the linear relation $\mathcal{Y} = (2 \odot \mathcal{X}) \oplus (-2)$ (which is given by the fuzzy number Tri(2,5,8)).

Figure 1.7 shows the above situation at the α-level with $\alpha = .5$.

The objective of the *Linear Regression Analysis* for this situation is to determine the values of the real-valued parameters a and b minimizing the error associated with the linear relation $\mathcal{Y} = (a \odot \mathcal{X}) \oplus b$. This error will be quantified in this section through the S-MSD or, more precisely, a measure defined in an analogous way, since randomness is now involved in the two variables affecting this error, and we will usually suppose that the weight measure S is symmetric in the sense that the associated density g is a symmetric function with respect to $\lambda = .5$ (that is, $g(\lambda) = g(1 - \lambda)$).

The latter assumption does not entail a significant loss of generality, since the error associated with the linear relation $\mathcal{Y} = (a \odot \mathcal{X}) \oplus b$ corresponds in the regression problem to a measure of the "distance" between the observed

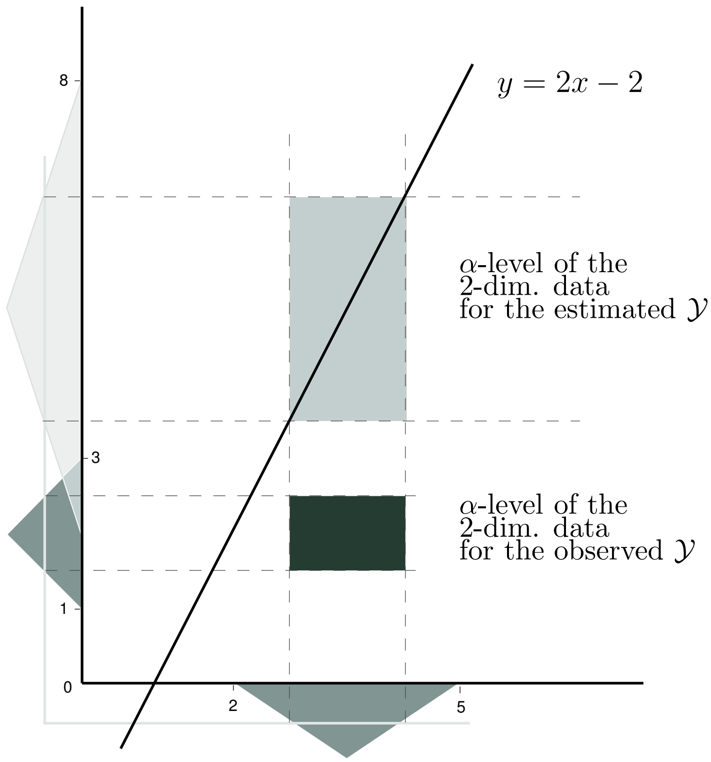


Fig. 1.7: α-level of the observed and estimated two-dimensional fuzzy data

values of $\mathcal{Y}$ and those observed by the linear relation $(a \odot \mathcal{X}) \oplus b$. In this sense, there are no arguments to discriminate (by the assigned weight) between $[f_{\mathcal{Y}}(\alpha,\lambda) - f_{(a\odot\mathcal{X})\oplus b}(\alpha,\lambda)]^2$ and $[f_{\mathcal{Y}}(\alpha,1-\lambda) - f_{(a\odot\mathcal{X})\oplus b}(\alpha,1-\lambda)]^2$ in quantifying the error of estimation.

$\mathcal{X}$ and $\mathcal{Y}$ will be also assumed not to be almost surely real-valued.

The aim is then given by the minimization of the function

$$\phi(a,b) = E\left([D_S(\mathcal{Y},(a\odot\mathcal{X})\oplus b)]^2\right)$$

$$= E\Big(\int_{[0,1]}\int_{[0,1]}\Big[f_{\mathcal{Y}}(\alpha,\lambda) - f_{(a\odot\mathcal{X})\oplus b}(\alpha,\lambda)\Big]^2\, d\,S(\lambda)\, d\,\alpha\Big),$$

with $a,\, b \in \mathbb{R}$.

Since the product of a fuzzy number by a real number admits different expressions for different signs of this real number, the minimization will be carried out in three stages, depending on whether $a > 0$, $a < 0$ or $a = 0$.

If $a \geq 0$, the function to be minimized is given by

$$\phi_1(a,b) = E\Big(\int_{[0,1]} \int_{[0,1]} \Big[f_{\mathcal{Y}}(\alpha,\lambda) - f_{(a\odot\mathcal{X})\oplus b}(\alpha,\lambda)\Big]^2 \, d\,S(\lambda)\, d\,\alpha \Big)$$

$$= E\Big(\int_{[0,1]} \int_{[0,1]} [f_{\mathcal{Y}}(\alpha,\lambda) - a\, f_{\mathcal{X}}(\alpha,\lambda) \,-\, b]^2 \, d\,S(\lambda)\, d\,\alpha \Big)$$

$$= E\left([D_S(\mathcal{Y},0)]^2 \right) + a^2 E\left([D_S(\mathcal{X},0)]^2 \right)$$

$$+ b^2 + 2abE\left(V_S(\mathcal{X})\right) - 2bE\left(V_S(\mathcal{Y})\right) - 2aE\left(W_S(\mathcal{X},\mathcal{Y})\right).$$

To minimize ϕ_1 as $a > 0$ we will consider

$$\frac{\partial \phi_1(a,b)}{\partial a} = 2aE\left([D_S(\mathcal{X},0)]^2 \right) + 2bE\left(V_S(\mathcal{X})\right) - 2E\left(W_S(\mathcal{X},\mathcal{Y})\right) = 0,$$

$$\frac{\partial \phi_1(a,b)}{\partial b} = 2b + 2aE\left(V_S(\mathcal{X})\right) - 2E\left(V_S(\mathcal{Y})\right) = 0,$$

which leads to a system of equations whose solution is given by

$$a_1 = \frac{E\left(W_S(\mathcal{X},\,\mathcal{Y})\right) - E\left(V_S(\mathcal{X})\right) E\left(V_S(\mathcal{Y})\right)}{E\left([D_S(\mathcal{X},0)]^2 \right) - [E\left(V_S(\mathcal{X})\right)]^2},$$

$$b_1 = E\left(V_S(\mathcal{Y})\right) - a_1 E\left(V_S(\mathcal{X})\right),$$

and will be accepted if $a_1 > 0$.

The second derivatives will be

$$\frac{\partial^2 \phi_1(a,b)}{\partial a^2} = 2E\left([D_S(\mathcal{X},0)]^2 \right),$$

$$\frac{\partial^2 \phi_1(a,b)}{\partial a \partial b} = \frac{\partial^2 \phi_1(a,b)}{\partial b \partial a} = 2E\left(V_S(\mathcal{X})\right),$$

$$\frac{\partial^2 \phi_1(a,b)}{\partial b^2} = 2,$$

and, consequently, the Hessian associated with ϕ_1 in (a_1, b_1) will be given by

$$H = \begin{vmatrix} 2E\left([D_S(\mathcal{X},0)]^2\right) & 2E\left(V_S(\mathcal{X})\right) \\ & \\ 2E\left(V_S(\mathcal{X})\right) & 2 \end{vmatrix}$$

$$= 4\left\{E\left([D_S(\mathcal{X},0)]^2\right) - [E\left(V_S(\mathcal{X})\right)]^2\right\}.$$

The solution (a_1, b_1) will provides us with an absolute minimum, whenever the Hessian H is positive. The convexity of $h(x) = x^2$ and the application of Jensen's inequality allows us to conclude that

$$[E\left(V_S(\mathcal{X})\right)]^2 = \left[E\left(\int_{[0,1]}\int_{[0,1]} f_{\mathcal{X}}(\alpha,\lambda)\; d\,S(\lambda)\,d\,\alpha\right)\right]^2$$

$$\leq E\left[\left(\int_{[0,1]}\int_{[0,1]} f_{\mathcal{X}}(\alpha,\lambda)\; d\,S(\lambda)\,d\,\alpha\right)^2\right]$$

$$\leq E\left(\int_{[0,1]}\left[\int_{[0,1]} f_{\mathcal{X}}(\alpha,\lambda)\; d\,S(\lambda)\right]^2 d\,\alpha\right)$$

$$\leq E\left(\int_{[0,1]}\int_{[0,1]} [f_{\mathcal{X}}(\alpha,\lambda)]^2\, d\,S(\lambda)\,d\,\alpha\right) = E\left([D_S(\mathcal{X},0)]^2\right).$$

The Hessian H would vanish if all the preceding inequalities reduce to equalities, and in virtue of Jensen's inequality these equalities are achieved if, and only if, $f_{\mathcal{X}(\cdot)}(\cdot\,,\,\cdot)$ is almost surely $[P \otimes m \otimes S]$ constant, which is equivalent to the fact $\mathcal{X}$ is a fuzzy random variable degenerate at a real value.

The solution for the case $a = 0$ is $(0, E\left(V_S(\mathcal{Y})\right).$

Similarly, if $a < 0$ the function to be minimized is given by

$$\phi_2(a,b) = E\left(\int_{[0,1]}\int_{[0,1]}\left[f_{\mathcal{Y}}(\alpha,\lambda) - f_{(a\odot\mathcal{X})\oplus b}(\alpha,\lambda)\right]^2 d\,S(\lambda)\,d\,\alpha\right)$$

$$= E\left(\int_{[0,1]}\int_{[0,1]} [f_{\mathcal{Y}}(\alpha,\lambda) - a\,f_{\mathcal{X}}(\alpha,1-\lambda)\; -\; b]^2\; d\,S(\lambda)\,d\,\alpha\right).$$

The symmetry of the measure S implies that

$$E\left(V_S\left(\mathcal{X}\right)\right) = E\Big(\int_{[0,1]}\int_{[0,1]} f_{\mathcal{X}}\left(\alpha,\lambda\right)\, d\,S(\lambda)\, d\,\alpha\Big)$$

$$= E\Big(\int_{[0,1]}\int_{[0,1]} f_{\mathcal{X}}\left(\alpha,1-\lambda\right)\, d\,S(\lambda)\, d\,\alpha\Big),$$

$$E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right) = E\Big(\int_{[0,1]}\int_{[0,1]} \left[f_{\mathcal{X}}\left(\alpha,\lambda\right)\right]^2\, d\,S(\lambda)\, d\,\alpha\Big)$$

$$= E\Big(\int_{[0,1]}\int_{[0,1]} \left[f_{\mathcal{X}}\left(\alpha,1-\lambda\right)\right]^2\, d\,S(\lambda)\, d\,\alpha\Big),$$

whence

$$\phi_2(a,b) = E\left(\left[D_S\left(\mathcal{Y},0\right)\right]^2\right) + a^2 E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right)$$

$$+ b^2 + 2abE\left(V_S\left(\mathcal{X}\right)\right) - 2bE\left(V_S\left(\mathcal{Y}\right)\right) - 2aE\left(W'_S\left(\mathcal{X},\mathcal{Y}\right)\right).$$

The first partial derivatives of ϕ_2 are given by

$$\frac{\partial\phi_2(a,b)}{\partial a} = 2aE\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right) + 2bE\left(V_S\left(\mathcal{X}\right)\right) - 2E\left(W'_S\left(\mathcal{X},\mathcal{Y}\right)\right),$$

$$\frac{\partial\phi_2(a,b)}{\partial b} = 2b + 2aE\left(V_s\left(\mathcal{X}\right)\right) - 2E\left(V_S\left(\mathcal{Y}\right)\right),$$

and the solution of the system determined by these two equations is

$$a_2 = \frac{E\left(W'_S\left(\mathcal{X},\mathcal{Y}\right)\right) - E\left(V_S\left(\mathcal{X}\right)\right)E\left(V_S\left(\mathcal{Y}\right)\right)}{E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right) - \left[E\left(V_S\left(\mathcal{X}\right)\right)\right]^2},$$

$$b_2 = E\left(V_S\left(\mathcal{Y}\right)\right) - a_2 E\left(V_S\left(\mathcal{X}\right)\right),$$

and it is acceptable if $a_2 < 0$.

The second derivatives are given by

$$\frac{\partial^2\phi_2(a,b)}{\partial a^2} = 2E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right),$$

$$\frac{\partial^2\phi_2(a,b)}{\partial a\partial b} = \frac{\partial^2\phi_2(a,b)}{\partial b\partial a} = 2E\left(V_S\left(\mathcal{X}\right)\right),$$

$$\frac{\partial^2\phi_2(a,b)}{\partial b^2} = 2,$$

whence the Hessian corresponds to

$$H = 4\left\{E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right) - \left[E\left(V_S\left(\mathcal{X}\right)\right)\right]^2\right\},$$

which is positive whenever $\mathcal{X}$ is not a fuzzy random variable degenerate at a real value.

If the density g is symmetrical with respect to $\lambda = .5$, we have that

$$E\left(W_S\left(\mathcal{X},\mathcal{Y}\right)\right) - E\left(W'_S\left(\mathcal{X},\mathcal{Y}\right)\right)$$

$$= E\left(\int_{[0,1]}\int_{[0,1]} g(\lambda) f_{\mathcal{Y}}\left(\alpha,\lambda\right)\left[f_{\mathcal{X}}\left(\alpha,\lambda\right) - f_{\mathcal{X}}\left(\alpha,1-\lambda\right)\right]\,d\,\lambda\,d\alpha\right)$$

$$= E\left(\int_{[0,1]}\int_{[0,0.5]} g(\lambda) f_{\mathcal{Y}}\left(\alpha,\lambda\right)\left[f_{\mathcal{X}}\left(\alpha,\lambda\right) - f_{\mathcal{X}}\left(\alpha,1-\lambda\right)\right]\,d\,\lambda\,d\alpha\right.$$

$$\left. + \int_{[0,1]}\int_{[0.5,1]} g(1-\lambda) f_{\mathcal{Y}}\left(\alpha,\lambda\right)\left[f_{\mathcal{X}}\left(\alpha,\lambda\right) - f_{\mathcal{X}}\left(\alpha,1-\lambda\right)\right]\,d\,\lambda\,d\alpha\right)$$

$$= E\Bigg(\int_{[0,1]}\int_{[0,0.5]} g(\lambda)\Big[f_{\mathcal{Y}}\left(\alpha,\lambda\right) f_{\mathcal{X}}\left(\alpha,\lambda\right) + f_{\mathcal{Y}}\left(\alpha,1-\lambda\right) f_{\mathcal{X}}\left(\alpha,1-\lambda\right)$$

$$- f_{\mathcal{Y}}\left(\alpha,\lambda\right) f_{\mathcal{X}}\left(\alpha,1-\lambda\right) - f_{\mathcal{Y}}\left(\alpha,1-\lambda\right) f_{\mathcal{X}}\left(\alpha,\lambda\right)\Big]\,d\,\lambda\,d\alpha\Bigg)$$

$$= K(S)\int_{[0,1]} E\left((\sup \mathcal{X}_\alpha - \inf \mathcal{X}_\alpha)(\sup \mathcal{Y}_\alpha - \inf \mathcal{Y}_\alpha)\right)\,d\,\alpha,$$

with $K(S) = \dfrac{1}{2}\displaystyle\int_{[0,1]}(1-2\lambda)^2\,d\,S(\lambda) > 0$.

Therefore, since $(\sup \mathcal{X}_\alpha - \inf \mathcal{X}_\alpha)(\sup \mathcal{Y}_\alpha - \inf \mathcal{Y}_\alpha) \geq 0$ for all $\alpha \in [0,1]$ y $\omega \in \Omega$, then

$$E\left(W_S\left(\mathcal{X},\mathcal{Y}\right)\right) \geq E\left(W'_S\left(\mathcal{X},\mathcal{Y}\right)\right),$$

with equality if, and only if, $\sup \mathcal{X}_{(\,\cdot\,)}(\,\cdot\,) = \inf \mathcal{X}_{(\,\cdot\,)}(\,\cdot\,)$ y $\sup \mathcal{Y}_{(\,\cdot\,)}(\,\cdot\,) = \inf \mathcal{Y}_{(\,\cdot\,)}(\,\cdot\,)$ almost surely $[m \otimes P]$.

In virtue of the left-continuity of $\inf \mathcal{X}_{(\,\cdot\,)}(\omega)$ and $\sup \mathcal{X}_{(\,\cdot\,)}(\omega)$ with respect to α in $[0,1]$, whatever $\omega \in \Omega$ may be, for any $\omega \in \Omega$ such that $\sup \mathcal{X}_{(\,\cdot\,)}(\omega) = \inf \mathcal{X}_{(\,\cdot\,)}(\omega)$ almost surely $[m]$ we have that $\sup \mathcal{X}_\alpha(\omega) = \inf \mathcal{X}_\alpha(\omega)$ for all

$\alpha \in [0,1]$ and the same happens for $\mathcal{Y}$. Consequently, $E\left(W_S\left(\mathcal{X},\mathcal{Y}\right)\right) = E\left(W'_S\left(\mathcal{X},\mathcal{Y}\right)\right)$ if, and only if, $\mathcal{X}$ and $\mathcal{Y}$ take on values on $\mathbb{R}$ almost surely $[P]$.

The latter assertion indicates that, except for fuzzy random variables almost surely real-valued, the value a_1 is greater than a_2, and hence and because of the continuity of $\phi(a,b)$ we can conclude that $\phi(a,b)$ cannot be minimized at a point (a^*,b^*) with $a^*=0$, since $a_1 \leq 0$ would force a_2 to be negative and $a_2 \geq 0$ would force a_1 to be positive, whence $\phi(a,b)$ would be minimized as $a^* = a_2$ if $a_1 \leq 0$ and as $a^* = a_1$ if $a_2 \geq 0$.

The above conclusions are summarized in the following result:

Theorem 1.6.1. *If $\mathcal{X}$ and $\mathcal{Y}$ are two integrably bounded fuzzy random variables which are not almost surely real-valued, and S is a measure whose associated density g with respect to the Lebesgue measure m on $[0,1]$ is symmetrical with respect to $\lambda = .5$, then the function $\phi(a,b) = E\Big([D_S(\mathcal{Y},(a\odot\mathcal{X}) \oplus b)]^2\Big)$ is minimized in (a^*,b^*) with $a^* \neq 0$ such that*

$$\phi(a^*,b^*) = \min\{\phi(a_1,b_1),\phi(a_2,b_2)\},$$

where

$$a^* = \frac{E\left(W_S\left(\mathcal{X},\,\mathcal{Y}\right)\right) - E\left(V_S\left(\mathcal{X}\right)\right)E\left(V_S\left(\mathcal{Y}\right)\right)}{E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right) - \left[E\left(V_S\left(\mathcal{X}\right)\right)\right]^2},$$

$$b^* = E\left(V_S\left(\mathcal{Y}\right)\right) - a^*E\left(V_S\left(\mathcal{X}\right)\right),$$

if either $E\left(W'_S\left(\mathcal{X},\,\mathcal{Y}\right)\right) \geq E\left(V_S\left(\mathcal{X}\right)\right)\ E\left(V_S\left(\mathcal{Y}\right)\right)$ *or* $E\left(W_S\left(\mathcal{X},\,\mathcal{Y}\right)\right) > E\left(V_S\left(\mathcal{X}\right)\right)E\left(V_S\left(\mathcal{Y}\right)\right) > E\left(W'_S\left(\mathcal{X},\,\mathcal{Y}\right)\right)$, *and*

$$a^* = \frac{E\left(W'_S\left(\mathcal{X},\mathcal{Y}\right)\right) - E\left(V_S\left(\mathcal{X}\right)\right)E\left(V_S\left(\mathcal{Y}\right)\right)}{E\left(\left[D_S\left(\mathcal{X},0\right)\right]^2\right) - \left[E\left(V_S\left(\mathcal{X}\right)\right)\right]^2}$$

$$b^* = E\left(V_S\left(\mathcal{Y}\right)\right) - a^*E\left(V_S\left(\mathcal{X}\right)\right),$$

if either $E\left(V_S\left(\mathcal{X}\right)\right)\ E\left(V_S\left(\mathcal{Y}\right)\right) \geq E\left(W_S\left(\mathcal{X},\,\mathcal{Y}\right)\right)$ *or* $E\left(W_S\left(\mathcal{X},\,\mathcal{Y}\right)\right) > E\left(V_S\left(\mathcal{X}\right)\right)\ E\left(V_S\left(\mathcal{Y}\right)\right) > E\left(W'_S\left(\mathcal{X},\,\mathcal{Y}\right)\right).$

Assume now that $\mathcal{X}$ and $\mathcal{Y}$ are two integrably bounded fuzzy random variables associated with the probability space $(\Omega, \mathcal{A}, P)$ and consider a general functional relation between $\mathcal{X}$ and $\mathcal{Y}$ such that

$$\mathcal{Y} = h(\mathcal{X})$$

with $h : \mathcal{F}_c(\mathbb{R}) \to \mathcal{F}_c(\mathbb{R})$ being $(\mathcal{B}_{d_\infty}, \mathcal{B}_{d_\infty})$-measurable, so that $h(\mathcal{X})$ is hence a fuzzy random variable.

The objective of the *general Functional Regression Analysis* for this situation is to determine the function h^* minimizing the error associated with the relation $\mathcal{Y} = h(\mathcal{X})$ with respect to h, and assuming that this error is measured by

$$F(h) = E\left([D_S\left(\mathcal{Y}, h(\mathcal{X})\right)]^2 \,\Big|\, P \right)$$
$$= E\left(\int_{[0,1]} \int_{[0,1]} \left[f_{\mathcal{Y}}(\alpha,\lambda) - f_{h(\mathcal{X})}(\alpha,\lambda) \right]^2 d\,S(\lambda)\, d\,\alpha \,\Bigg|\, P \right).$$

If we apply Theorem 0.3.1, then for all $\alpha \in [0,1]$ and $\lambda \in [0,1]$ we have that

$$E\left(\left[f_{\mathcal{Y}}(\alpha,\lambda) - f_{h(\mathcal{X})}(\alpha,\lambda) \right]^2 \,\Big|\, P \right)$$
$$= E\left(E\left(\left[f_{\mathcal{Y}}(\alpha,\lambda) - f_{h(\tilde{x})}(\alpha,\lambda) \right]^2 \,\Big|\, P_{\tilde{x}} \right) \,\Big|\, P_{\mathcal{X}} \right),$$

and hence

$$F(h) = E\left(MSD_S(\mathcal{Y}, h(\tilde{x})) \,\big|\, P_{\mathcal{X}} \right),$$

where $MSD_S(\mathcal{Y}, h(\tilde{x}))$ is the S-mean squared dispersion associated with $\mathcal{Y}$ with respect to $h(\tilde{x})$ on the probability space conditioned by $\tilde{x}$.

Given $\tilde{x} \in \mathcal{X}(\Omega)$, in virtue of Theorem 1.2.7, we have that $MSD_S(\mathcal{Y}, h(\tilde{x}))$ is minimized for $h^*(\tilde{x}) = \widetilde{E}(\mathcal{Y} \,|\, \mathcal{X} = \tilde{x})$, whence we conclude that

Theorem 1.6.2. *If $\mathcal{X}$ and $\mathcal{Y}$ are two integrably bounded fuzzy random variables and S is a measure satisfying the conditions in Definition* 0.1.5*,* *then the above function* F*, defined on the set* $\mathcal{H} = \{h : \mathcal{F}_c(\mathbb{R}) \to \mathcal{F}_c(\mathbb{R}) \mid h(\mathcal{X})$ *fuzzy random variable*$\}$*, is minimized for the function* $h^* : \mathcal{F}_c(\mathbb{R}) \to \mathcal{F}_c(\mathbb{R})$ *such that*

$$h^*(\tilde{x}) = \widetilde{E}(\mathcal{Y} \mid \mathcal{X} = \tilde{x}),$$

for all $\tilde{x} \in \mathcal{X}(\Omega)$*.*

Obviously, the functions $h_1 : \mathcal{F}_c(\mathbb{R}) \to \mathcal{F}_c(\mathbb{R})$ such that $h_1(\widetilde{V}) = (a \odot \widetilde{V}) \oplus b$ for all $\widetilde{V} \in \mathcal{F}_c(\mathbb{R})$, where a and b are arbitrary real values, satisfy that $h_1 \in \mathcal{H}$, so that the optimal solution in Theorem 1.6.2 cannot be "worse" than that in Theorem 1.6.1.

Remark 1.6.1. The optimal general relation does not depend on the choice of the measure S.

Remark 1.6.2. In principle, the latter study only allows us to predict the value of $\mathcal{Y}$ for observed values of $\mathcal{X}$ which have appeared in the population (or sample) we have considered. In Section 1.7 we include a few comments on possible ways of solving this inconvenience.

The results in Theorems 1.6.1 and 1.6.2, as well as their practical worth, are now illustrated by means of a real-life example, in which data have been supplied by members of the Departamento de Medio Ambiente of the Consejería de Agricultura in the Principado de Asturias in Spain.

Example 1.6.1. Consider the population Ω of 50 days of a certain year, in which CLOUDINESS (variable $\mathcal{X}$) and VISIBILITY (variable $\mathcal{Y}$) have been observed.

Variable $\mathcal{X}$ takes on the values BRIGHT ($\tilde{x}_1$), CLEAR ($\tilde{x}_2$), DULL ($\tilde{x}_3$), CLOUDY ($\tilde{x}_4$), and OVERCAST ($\tilde{x}_5$), and variable $\mathcal{Y}$ takes on the values PERFECT ($\tilde{y}_1$), GOOD ($\tilde{y}_2$), MEDIUM ($\tilde{y}_3$), POOR ($\tilde{y}_4$), and BAD ($\tilde{y}_5$).

Experts in the measurement of these values have described them in terms of the fuzzy sets (meaning fuzzy percentages) and based on S- and Π-curves, and triangular and trapezoidal fuzzy numbers, whose support is strictly contained in $[0, 100]$ as follows (see Figures 1.8 and 1.9):

$$\tilde{x}_1 = S(0, 20),$$
$$\tilde{x}_2 = \Pi(10, 20, 40),$$

$$\tilde{x}_3 = \begin{cases} S(20, 40) & \text{in } [20, 40] \\ 1 & \text{in } [40, 50] \\ 1 - S(50, 70) & \text{in } [50, 70] \\ 0 & \text{otherwise,} \end{cases}$$

$$\tilde{x}_4 = \begin{cases} S(60, 70) & \text{in } [60, 70] \\ 1 & \text{in } [70, 80] \\ 1 - S(80, 90) & \text{in } [80, 90] \\ 0 & \text{otherwise,} \end{cases}$$

$$\tilde{x}_5 = 1 - S(80, 100),$$
$$\tilde{y}_1 = \text{Tra}(90, 95, 100, 100),$$
$$\tilde{y}_2 = \text{Tri}(70, 90, 100),$$

$$\tilde{y}_3 = \begin{cases} S(40, 50) & \text{in } [40, 50] \\ 1 & \text{in } [50, 70] \\ 1 - S(70, 80) & \text{in } [70, 80] \\ 0 & \text{otherwise,} \end{cases}$$

$$\tilde{y}_4 = \begin{cases} S(20, 30) & \text{in } [20, 30] \\ 1 & \text{in } [30, 40] \\ 1 - S(40, 50) & \text{in } [40, 50] \\ 0 & \text{otherwise,} \end{cases}$$

$$\tilde{y}_5 = S(0, 20).$$

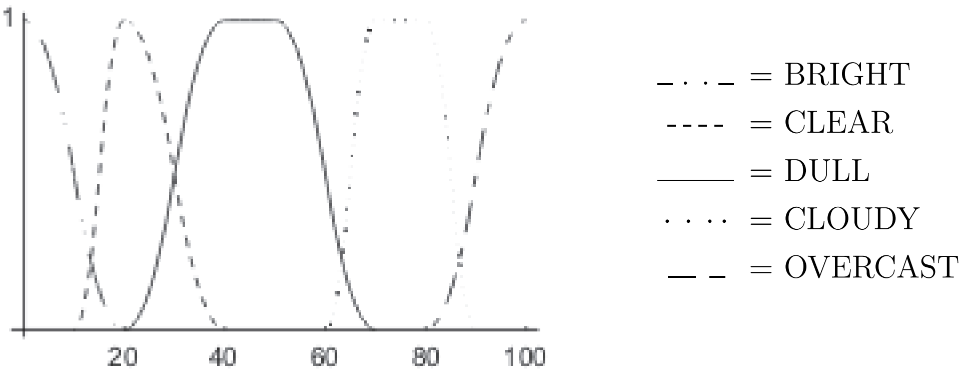


Fig. 1.8: Values of the variable CLOUDINESS

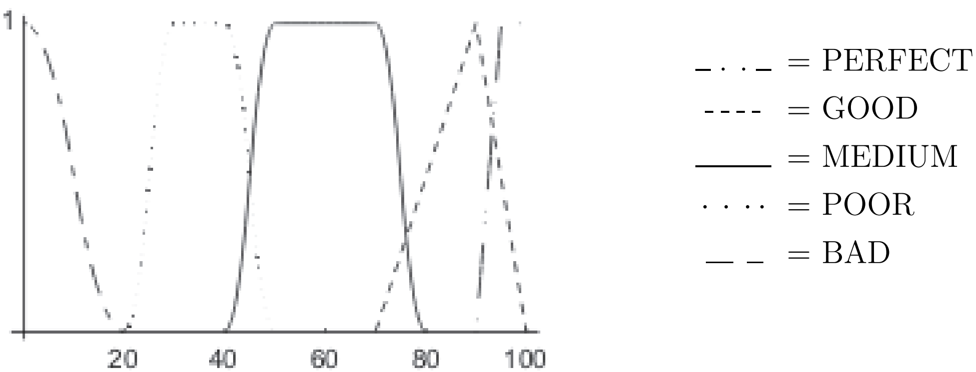


Fig. 1.9: Values of the variable VISIBILITY

For the considered population Ω the observed two-dimensional fuzzy data have been collected in Table 1.1.

If we want to look for the optimal relation between VISIBILITY and CLOUDINESS, in which the first one is expressed as a fuzzy linear function of the second one, and we use the result in Theorem 1.6.1 by choosing the measure S as a discretized one such that $L = 3$, $\lambda_1 = 0$, $\lambda_2 = .5$, $\lambda_3 = 1$, $k_1 = k_3 = .375$, $k_2 = .25$ and $\overline{g}(\lambda) = 0$ if $\lambda \in (0,1) \setminus \{.5\}$, then the function $\phi(a,b)$ is the one represented in Figure 1.10, and the optimal relation is the one whose slope parameter corresponds to $a^* = -.95$ and the intercept parameter is given by $b^* = 111.77$, that is, $\mathcal{Y} = (-.95 \odot \mathcal{X}) \oplus 111.77$.

$\mathcal{X} \setminus \mathcal{Y}$	$\tilde{y}_1$	$\tilde{y}_2$	$\tilde{y}_3$	$\tilde{y}_4$	$\tilde{y}_5$
$\tilde{x}_1$	3	7			
$\tilde{x}_2$	2	7	1		
$\tilde{x}_3$		5	4	1	
$\tilde{x}_4$		2	5	3	
$\tilde{x}_5$			2	5	3

Table 1.1: Contingency table for variables CLOUDINESS and VISIBILITY on Ω

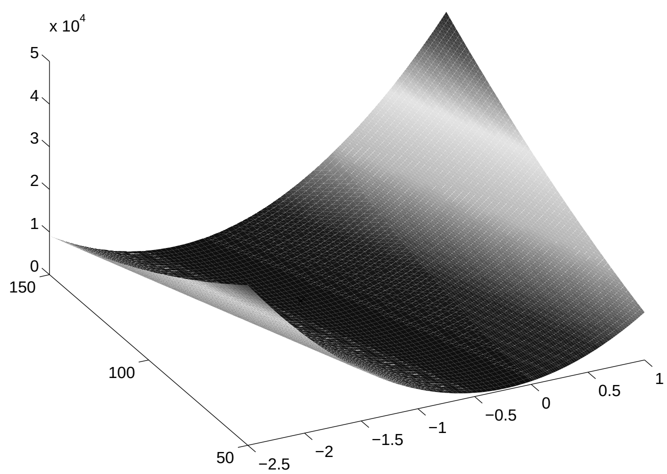


Fig. 1.10: Graphical representation of $\phi(a,b)$ in Example 1.6.1

On the basis of this relation, we can make a prediction for the VISIBILITY of a day which does not belong to Ω but has been classified as DULL for CLOUDINESS variable, by means of the value

$$\tilde{y}^* = (-.95 \odot \tilde{x}_3) \oplus 111.77 = \begin{cases} S(45.27, 64.27) & \text{in } [45.27, 64.27] \\ 1 & \text{in } [64.27, 73.77] \\ 1 - S(73.77, 92.77) & \text{in } [73.77, 92.77] \\ 0 & \text{otherwise,} \end{cases}$$

which would correspond to a VISIBILITY that can be (more or less) interpreted as RATHER GOOD (Figure 1.11).

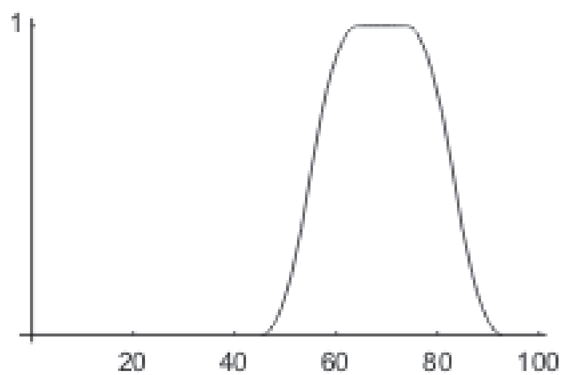


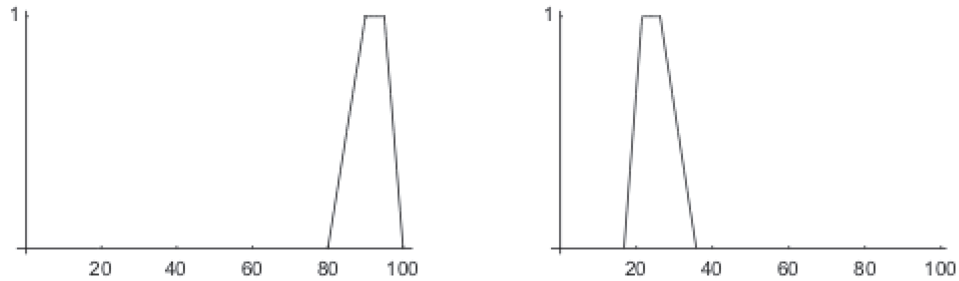
Fig. 1.11: Predicted VISIBILITY of a DULL day in accordance with the optimal linear regression

On the other hand, we could also predict the VISIBILITY of a day which has been classified as VERY CLOUDY for CLOUDINESS variable (that is, for an observed value which would not appear in elements of Ω), and described by the experts by means of the trapezoidal fuzzy set $\tilde{x} = \text{Tra}(80, 90, 95, 100))$. The estimated value of $\mathcal{Y}$ is given by

$$\tilde{y} = (-.95 \odot \tilde{x}) \oplus 111.77 = \text{Tra}(16.77, 21.52, 26.27, 35.77),$$

which would correspond to a VISIBILITY that can be (more or less) interpreted as QUITE BAD (Figure 1.12).

If we want to look for the optimal relation between VISIBILITY and CLOUDINESS, in which the first one is expressed as a fuzzy general function of the second one, and we use the result in Theorem 1.6.2, then (whatever the measure S may be) the prediction for a VERY CLOUDY day could not be given, but rather only the prediction for the values of $\mathcal{X}$ which have been taken on by elements in Ω would make sense. Thus, if we wish to predict



(a) (b)

Fig. 1.12: (a) Graphical representation of a VERY CLOUDY day (b) Predicted VISIBILITY for a VERY CLOUDY day in accordance with the optimal linear regression

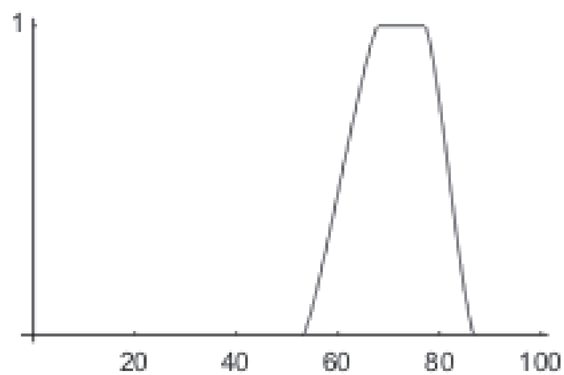


Fig. 1.13: Predicted VISIBILITY for a DULL day in accordance with the optimal general regression

the VISIBILITY of a DULL day, we will obtain (see Figure 1.13)

$$\tilde{y} = E(\mathcal{Y} \,|\, \mathcal{X} = \tilde{x}_3) = \sum_{i=1}^{5} P(\tilde{y}_i \,|\, \mathcal{X} = \tilde{x}_3) \odot \tilde{y}_i$$

$$= \left(\frac{5}{10} \odot \tilde{y}_2\right) \oplus \left(\frac{4}{10} \odot \tilde{y}_3\right) \oplus \left(\frac{1}{10} \odot \tilde{y}_4\right),$$

which is narrower than the predicted value of $\mathcal{Y}$ by the linear relation.

1.7 Final discussion and future directions

In evaluating Chapter 1 we conclude that the extension of the second moment for fuzzy random variables on the basis of the metric D_S by Bertoluzza *et al.* (1995a), preserves all the relevant properties from the real-valued case and it is adapted in a similar way to the statistical applications in this chapter.

Nevertheless, other studies based on the same extension would not preserve such analogies. Some of them are indicated in this section.

From an operational viewpoint, difficulties in the extension developed in this chapter reduce basically to the characterization of variable values and to the computation of several integrals which will be solved by means of an appropriate software (developed on the basis of that by López García 1997, and in the usual approximations for the integrals).

In connection with the research in this chapter, there are several open problems which are currently being studied or are in line to be so in the near future.

- Properties in Section 1.2 can be automatically completed with the development of some Laws of Large Numbers. Since there exists a Strong Law for fuzzy random variable pairwise independent and identically distributed based on the convergence in the sense of the d_∞ (see Colubi *et al.* 1999), which would imply a similar one for the D_S metric, it would be interesting to state laws which do not require the above conditions for the fuzzy random variables in the sequence (in this sense, we can consider as a reference guide the results by Körner 1997b).

- The so-called "improved" Rao-Blackwell Theorem could be extended for fuzzy random variables, even though to establish it, the notion of sufficiency of a fuzzy estimator of a fuzzy parameter should be formalized, and this formalization is not at all immediate.

- Results in Sections 1.3 to 1.5 could be extended to more complex samplings. In this way, it would be interesting to pay attention to the stratified sampling which in the asymptotical case would correspond to using finite mixtures of probability distributions. It would be also convenient to examine the suitability of a criterion for stratification by means of the introduction of an appropriate measure of the relative gain of precision associated with the stratification (by following the ideas by Alonso & Gil 1995). At this point, an inconvenience arises in contrast to the results of the real-valued case: how to determine approximately an adequate sample size to estimate the population S-MSD on the basis of the asymptotic distribution, since with the results available at present we could not state a result similar to that in the Stone Theorem to be used with such a purpose. It would be then valuable to look for procedures "refining" the sizes obtained in Theorems 1.5.4 and 1.5.8.

- The asymptotic studies in Section 1.4 can be carried out when we consider two populations and analyze the difference between their S-MSD, with the aim of comparing them.

- On the other hand, it is necessary to develop a discussion on the sample sizes providing us with good approximations from the asymptotic results in Section 1.4. This discussion could be performed in terms of simulation studies from fuzzy random variables, which lead to two new open problems: to consider operational models describing the "usual behavior" for fuzzy random variables, and to simulate data from these models.

- The study of the Linear Regression developed in Section 1.6 assumes that neither $\mathcal{X}$ nor $\mathcal{Y}$ are variables taking on almost surely real values. If one of the two variables is almost surely real-valued, and in particular if it is a real-valued random variable, the study in this chapter cannot be applied. It would be useful to examine the problem in this particular situation, by discussing the cases in which the slope parameter of the linear relation vanishes. However, this latter problem becomes more challenging as the parameters of the linear relation are allowed to take on fuzzy values, and in the case in which $\mathcal{Y}$ is the fuzzy-valued variable and $\mathcal{X}$ is the real-valued one would lead to the problem considered by Näther (1997) and Körner (1997ab).

- One of the most attractive open problems from this chapter is that of measuring the linear (and general) correlation between two fuzzy random variables, especially because of the differences that this problem will show in comparison with the problem for real-valued random variable. In this respect, one can point out that given two two-dimensional fuzzy data, there would not exist in general a linear relation describing exactly the relation between these two data. It should be also valuable to study the relationship between the two possible "linear regression planes"representing the optimal linear relations of $\mathcal{Y}$ with respect to $\mathcal{X}$, and of $\mathcal{X}$ with respect to $\mathcal{Y}$.

- Another interesting related question is that of the search of methods for predicting the response variable $\mathcal{Y}$ from the value of the observed one $\mathcal{X}$ in the General Regression Analysis, when we consider a value of $\mathcal{X}$ which has not appeared in the population/sample on which the optimal relation has been based. To this purpose, we could try to adapt proper interpolation methods of fuzzy data (see, for instance, Lowen 1990, Kaleva 1994, Gal 1996). Furthermore, some of the fuzzy mappings involved in the interpolation studies could inspire the search of nonlinear regression models which satisfy certain conditions of "continuity" and "regularity".

- Finally, the studies carried out in this chapter require a deep analysis on the robustness from different perspectives, namely, the sensitiveness of the values and comparisons based on the S-MSD with respect to the choice of the measure S, and the sensitiveness of the values, comparisons and descriptive and inferential results with respect to the shape of the fuzzy sets chosen to model the values of the considered variable.

Chapter 2

Generalized measure of the relative variation (inequality) of a fuzzy random variable

Just like the deviation, the "inequality" or relative variation of a random variable allows us to compare variables or populations. However, whereas the deviation is usually quantified in terms of the distance between variable values and a certain referential value (commonly the expected value), the inequality is usually quantified in terms of the ratio of these values and the referential value. Moreover, in the inequality one must consider "how many times each variable value has as much as the referential value", and hence, it distinguishes between values being above and below the referential value (unlike the deviation that does not distinguish between values equidistant from the referential one).

Another distinctive characteristic of the inequality, opposite to the deviation, is that the "contribution" of variable values to the inequality associated with this variable is not always nonnegative. More precisely, the contribution for values above the referential one is negative, for values below the referential one it is positive, and for values being equal to the referential the contribution is null.

In the introductory chapter, we have pointed out that variables for which inequality indices are defined are usually assumed to be positive random variables, and in most cases (although it is not always explicited) it is admitted that variable values are measured on a ratio scale.

In the present chapter, one first takes into account that there exist many aspects of a quantitative attribute (like social/political repercussion, added reputation/damage, imprecision in a quantitative report or evaluation, etc.), for which a numerical scale cannot reflect the true value. Aspects of this type have motivated the introduction of the concept of the fuzzy random variable and, in particular, that of the utility function in decision problems (see, for instance, Gil & López-Díaz 1996, Gil *et al.* 1998b).

Atkinson (1983, p. 53) has emphasized that the inequality measures "...are widely used for two purposes:

(i) to compare (income, wealth, etc.) distributions;

(ii) to attach some measure to the degree of inequality, or to give some idea whether the inequality is 'large' or 'small'..."

To achieve the purpose *(ii)* some studies have been already developed (Basu 1987, Ok 1994, 1995, 1996, López García 1997, Colubi *et al.* 1997, Colubi Cervero 1997, Gil *et al.* 1998a, López-García & Corral 1998) in which several fuzzy-valued measures of the degree of inequality of a population with respect to a (real- or fuzzy-valued) random variable have been stated.

Nevertheless, if we want to give an appropriate answer to the purpose *(i)*, one can either make use of a real-valued mesure of inequality, or apply a ranking of fuzzy numbers after a fuzzy-valued measurement of inequality. In the family of measures we will present in this chapter, we combine both options, by defining a real-valued measure which is obtained by composing a ranking function on the space $\mathcal{F}_c(\mathbb{R})$ and a generalized fuzzy-valued inequality measure.

In this way, we will introduce a generalized family of real-valued indices of inequality with respect to a fuzzy random variable. This family extends that in Definition 0.4.1 and includes the extension to fuzzy random variables of the additively decomposable indices to all but Theil's one. In particular, the hyperbolic and the Shannon type indices are both extended in this generalized family defined in Section 2.1.

We next examine conditions to guarantee the existence of the extension, and we analyze properties of the relative variation defined.

The unbiased estimation of the population extended hyperbolic index in random samplings with and without replacement from finite popultions, as well as the exact quantification of the associated sampling error of the estimation, are developed. This study is complemented with that of the asymptotic distribution for most of the extended indices in the defined family, which allows us to construct approximated inferential techniques on the population measures.

2.1 The f-inequality indices for fuzzy random variables

Assume that we consider a general population Ω and let $(\Omega, \mathcal{A}, P)$ be a probability space defined on it. Let $\mathcal{X} : \Omega \to \mathcal{F}_c\left((0, +\infty)\right)$ be an integrably bounded fuzzy random variable associated with $(\Omega, \mathcal{A}, P)$ (that is, $\mathcal{X}$ being a positive fuzzy random variable such that $\sup \mathcal{X}_0 \in L^1(\Omega, \mathcal{A}, P)$).

Let $f : (0, +\infty) \to \mathbb{R}$ be a strictly convex (intended as convex downward) and monotonic function satisfying that $f(u)+f(1/u) \geq 0$ for all $u \in (0, +\infty)$ and $f(1) = 0$, and let F be the ranking function in $\mathcal{F}_c(\mathbb{R})$ introduced by Yager (1981) (see Subsection 0.1).

To quantify the inequality associated with $\mathcal{X}$ in Ω by means of a real-valued measurement, we suggest the following indices:

Definition 2.1.1. *The f-inequality index associated with $\mathcal{X}$ in the population Ω given by the value (if it exists)*

$$\mathrm{I}_f(\mathcal{X}) = F\left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right),$$

where

$$\widetilde{\mathrm{I}}_f(\mathcal{X}) = \widetilde{E}\left[f\left(\mathcal{X} \oslash \widetilde{E}(\mathcal{X})\right)\right],$$

*is the fuzzy f-inequality index associated with $\mathcal{X}$ in Ω (*Colubi *et al.* 1997, Colubi Cervero 1997*), where $f\left(\mathcal{X} \oslash \widetilde{E}(\mathcal{X})\right)$ denotes the image of $\mathcal{X} \oslash \widetilde{E}(\mathcal{X})$ induced from f on the basis of Zadeh's extension principle (that is,*

$$\left(f\left(\mathcal{X} \oslash \widetilde{E}(\mathcal{X})\right)\right)_\alpha = \left[\min\left\{f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right), f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right\},\right.$$
$$\left.\max\left\{f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right), f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right\}\right]$$

for all $\alpha \in [0,1]$).

Occasionally, and when the probability measure P has to be specified, we will denote $\mathrm{I}_f(\mathcal{X})$ alternatively by $\mathrm{I}_f(\mathcal{X} \mid P)$.

The conditions assumed for f are satisfied by the functions associated with the additively decomposable indices for $\alpha \neq 0, 1$, and the function $f(x) = -\log x$ associated with the index of the Shannon type (see Appendix B). However, the function associated with Theil's index ($f(x) = x\log(x)$) is nonmonotonic and it will be removed from the present study.

The f-inequality index is not necessarily defined for a fuzzy random variable in any population, and conditions guaranteeing the existence of $\widetilde{\mathrm{I}}_f$ depend on the function f.

The following conditions, which ensure the existence of $\widetilde{\mathrm{I}}_f(\mathcal{X})$ (see Colubi Cervero 1997), would also guarantee that $\mathrm{I}_f(\mathcal{X}) \in \mathbb{R}$.

Consider a probabilistic space $(\Omega, \mathcal{A}, P)$, and assume that $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ is a positive integrably bounded fuzzy random variable associated with $(\Omega, \mathcal{A}, P)$. Let $f : (0, +\infty) \to \mathbb{R}$ be a strictly convex and monotonic function belonging to C^1 and satisfying that $f(u) + f(1/u) \geq 0$ for all $u \in (0, +\infty)$ and $f(1) = 0$.

Then, the fuzzy f-inequality index $\widetilde{\mathrm{I}}_f(\mathcal{X})$ is well-defined and belongs to $\mathcal{F}_c(\mathbb{R})$ if, and only if,

(1) $f\left(\inf \mathcal{X}_0 / E(\sup \mathcal{X}_0)\right) \in L^1(\Omega, \mathcal{A}, P)$ if f is nonincreasing;

(2) $f\left(\sup \mathcal{X}_0 / E(\inf \mathcal{X}_0)\right) \in L^1(\Omega, \mathcal{A}, P)$ if f is nondecreasing.

On the other hand, the fuzzy f-inequality index associated with $\mathcal{X}$ in the population (when it exists under the conditions above assumed for f), is the fuzzy number in $\mathcal{F}_c(\mathbb{R})$ such that for all $\alpha \in [0,1]$

$$\left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha = \left[\inf\left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha, \sup\left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha\right],$$

where

$$\inf\left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha = E\left(\min\left\{f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right), f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right\}\right),$$

$$\sup\left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha = E\left(\max\left\{f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right), f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right\}\right).$$

On the basis of the latter assertions, and because of the properties of the ranking function F, we have that

Theorem 2.1.1. *Let $(\Omega, \mathcal{A}, P)$ be a probabilistic space and let $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ be an integrably bounded fuzzy random variable. Consider a mapping $f : (0, +\infty) \to \mathbb{R}$ strictly convex and monotonic, belonging to C^1, and satisfying $f(u) + f(1/u) \geq 0$ for all $u \in (0, +\infty)$ and $f(1) = 0$.*

If either (1) or (2) are also satisfied, then,

$$\mathrm{I}_f(\mathcal{X}) = \frac{1}{2}\int_{[0,1]} E\left[f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right) + f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right] d\alpha.$$

Obviously, if we forget about either condition for the values of $\mathcal{X}$ or the monotonicity of the function f, we could not characterize the f-indices as presented in Theorem 2.1.1, which would mean an important (both, practical and theoretical) inconvenience for the later studies in this chapter.

In the following section we will examine several properties being convenient and desirable in the measurement of the inequality of a population wih respect to a fuzzy random variable.

2.2 Properties of the f-inequality indices

From now on, and without mentioning it for each property, we will consider a probability space $(\Omega, \mathcal{A}, P)$, and a function $f : (0, +\infty) \to \mathbb{R}$ which is strictly convex and monotonic, belongs to C^1, and satisfies that $f(u)$ $+f(1/u) \geq 0$ for all $u \in (0, +\infty)$ and $f(1) = 0$. We also will suppose that if f is nonincreasing Condition (1) in Section 2.1 is satisfied, and if f is nondecreasing Condition (2) in Section 2.1 is satisfied. The strict convexity of f could be weakened by assuming f is convex, but in such a case we could not establish the conditions under which the equality would hold in several of the properties below.

The indices introduced in Section 2.1 are not changed by an equiproportional real-valued variation in the values of the fuzzy random variables. In other words, and in accordance with Kölm's terminology (1976ab) these indices are "rightist measures", and following Blackorby & Donaldson (1978) they are measures of "relative inequality". In this way, the following result extends to fuzzy random variables, the *mean independence property* (also referred to as *scale invariance* or *homogeneity of degree 0*) of most of the classical inequality indices.

Theorem 2.2.1 (Mean independence). *If* $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ *is an integrably bounded fuzzy random variable, then for all* $k \in (0,+\infty)$ *we have that* $\mathrm{I}_f(k \odot \mathcal{X}) = \mathrm{I}_f(\mathcal{X})$.

Proof.

Since $k \odot \mathcal{X}$ is a positive integrably bounded fuzzy random variable, with $k \odot \mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ defined so that $(k \odot \mathcal{X})(\omega) = k \odot \mathcal{X}(\omega)$ for all $\omega \in \Omega$, then for all $\alpha \in [0,1]$ and $k \in (0,+\infty)$ we have that

$$\frac{\inf(k \odot \mathcal{X})_\alpha}{E\left(\sup(k \odot \mathcal{X})_\alpha\right)} = \frac{k \inf \mathcal{X}_\alpha}{kE(\sup \mathcal{X}_\alpha)} = \frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)},$$

and

$$\frac{\sup(k \odot \mathcal{X})_\alpha}{E\left(\inf(k \odot \mathcal{X})_\alpha\right)} = \frac{k \sup \mathcal{X}_\alpha}{kE(\inf \mathcal{X}_\alpha)} = \frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)},$$

whence

$$\mathrm{I}_f(k \odot \mathcal{X}) = \mathrm{I}_f(\mathcal{X}).$$ □

The *sign-preserving* (or *nonnegativeness*) holds for all the indices in Definition 2.1.1. Thus,

Theorem 2.2.2 (Nonnegativeness). *Let* $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ *be an integrably bounded fuzzy random variable. Then, we have that* $\mathrm{I}_f(\mathcal{X}) \geq 0$.

Proof.

Indeed, because of the convexity of f and in virtue of Jensen's inequality and the conditions assumed for f, we have that

$$E\left(f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right)\right) + E\left(f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right)$$

$$\geq f\left(E\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right)\right) + f\left(E\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right)$$

$$= f\left(\frac{E(\inf \mathcal{X}_\alpha)}{E(\sup \mathcal{X}_\alpha)}\right) + f\left(\frac{E(\sup \mathcal{X}_\alpha)}{E(\inf \mathcal{X}_\alpha)}\right) \geq 0.$$ □

The *positiveness* (or *sensitivity out of equality*) is formalized in the following result:

Theorem 2.2.3 (Sensitivity out of equality). *Let $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ be an integrably bounded fuzzy random variable. If $\mathrm{I}_f(\mathcal{X}) = 0$, then $\mathcal{X}$ has to be a degenerate fuzzy random variable (that is, if $\mathcal{X}$ is nondegenerate we can ensure that $\mathrm{I}_f(\mathcal{X})$ is positive).*

Proof.

In virtue of Theorem 2.1.1, and the left-continuity of $\inf\left(f\left(\mathcal{X} \oslash \widetilde{E}(\mathcal{X})\right)\right)_\alpha$ and $\sup\left(f\left(\mathcal{X} \oslash \widetilde{E}(\mathcal{X})\right)\right)_\alpha$ with respect to α, then $\mathrm{I}_f(\mathcal{X}) = 0$ if, and only if, for all $\alpha \in [0,1]$

$$\mathrm{E}\left[f\left(\frac{\inf \mathcal{X}_\alpha}{E(\sup \mathcal{X}_\alpha)}\right)\right] + \mathrm{E}\left[f\left(\frac{\sup \mathcal{X}_\alpha}{E(\inf \mathcal{X}_\alpha)}\right)\right] = 0.$$

To obtain this condition the inequality relation obtained by applying Jensen's inequality in Theorem 2.2.2 should be an equality. This happens if, and only if, for all $\alpha \in [0,1]$ the real-valued random variables $\inf \mathcal{X}_\alpha$ and $\sup \mathcal{X}_\alpha$ are degenerate, and, therefore, all the α-level functions $\mathcal{X}_\alpha$ must be degenerate convex compact random sets. Consequently, (see López García 1997, p. 191), $\mathcal{X}$ has to be a degenerate fuzzy random variable. □

The *insensitivity* or *nullity* of the f-inequality indices cannot be guaranteed for a degenerate fuzzy random variable whatever f may be. The following result states that for fuzzy random variables degenerate at a positive real number this insensitivity always holds.

Theorem 2.2.4 (Insensitivity). *Let $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ be an integrably bounded fuzzy random variable. If $\mathcal{X}$ is degenerate at a positive real value, then $\mathrm{I}_f(\mathcal{X}) = 0$.*

Proof.

Indeed, if $\mathcal{X}$ is a fuzzy random variable degenerate at a positive real value, then for all $\alpha \in [0,1]$ we have that $\inf \mathcal{X}_\alpha = \sup \mathcal{X}_\alpha$ a.s. $[P]$ and, hence, $E(\inf \mathcal{X}_\alpha) = E(\sup \mathcal{X}_\alpha)$, so that $\inf \left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha = \sup \left(\widetilde{\mathrm{I}}_f(\mathcal{X})\right)_\alpha = 0$, whence $\mathrm{I}_f(\mathcal{X}) = 0$. □

The next results present two different *minimality* properties for the f-indices, depending on certain conditions the function f satisfies.

Theorem 2.2.5 (Minimality I). *Let $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ be an integrably bounded fuzzy random variable. If f satisfies that $f(u) + f(1/u) = 0$ if, and only if, $u = 1$, then $\mathrm{I}_f(\mathcal{X}) = 0$ if, and only if, $\mathcal{X}$ is a fuzzy random variable degenerate at a positive real number.*

Proof.

In virtue of Theorems 2.2.2 and 2.2.3, $\mathrm{I}_f(\mathcal{X}) = 0$ if, and only if, $\inf \mathcal{X}_\alpha$ and $\sup \mathcal{X}_\alpha$ are degenerate real-valued random variables, and (because of the extra condition assumed for f) $E(\inf \mathcal{X}_\alpha) = E(\sup \mathcal{X}_\alpha)$, that is, if, and only if, $\inf \mathcal{X}_\alpha$ and $\sup \mathcal{X}_\alpha$ are real-valued random variables degenerate at the same value, and this forces $\mathcal{X}$ to be a fuzzy random variable degenerate at this value. □

As indicated in Appendix B, the additional condition $f(u) + f(1/u) = 0$ if, and only if, $u = 1$, is satisfied by many functions f (in particular, for those serving to extend the additively decomposable indices of order $\alpha \neq 0, 1$), although other valuable functions like $f(x) = -\log x$ (which is the basis of the Shannon type index) satisfy that $f(u) + f(1/u) = 0$ for all $u \in (0,+\infty)$ (see also Appendix B). In this latter case, the necessary and sufficient condition for $\mathrm{I}_f(\mathcal{X})$ being null is gathered in the following result:

Theorem 2.2.6 (Minimality II). *If $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ is an integrably bounded fuzzy random variable, and $f(u) + f(1/u) = 0$ for all*

$u \in (0, +\infty)$, then $\mathrm{I}_f(\mathcal{X}) = 0$ if, and only if, $\mathcal{X}$ is degenerate at an element in $\mathcal{F}_c((0, +\infty))$.

Proof.

The necessary condition has been already proved in Theorem 2.2.3.

Conversely, if $\mathcal{X}$ is a degenerate fuzzy random variable, $\inf \mathcal{X}_\alpha$ and $\sup \mathcal{X}_\alpha$ are degenerate real-valued random variables, whence because of the extra condition assumed for f

$$\frac{1}{2}\left[f\left(\frac{E(\inf \mathcal{X}_\alpha)}{E(\sup \mathcal{X}_\alpha)}\right) + f\left(\frac{E(\sup \mathcal{X}_\alpha)}{E(\inf \mathcal{X}_\alpha)}\right)\right] = 0.$$

□

Remark 2.2.1. In accordance with Theorems 2.2.5 and 2.2.6, if $\mathcal{X}$ is a fuzzy random variable degenerate at a fuzzy number in $\mathcal{F}_c((0, +\infty))$, the f-inequality index does not necessarily equal 0. Thus, for instance, if $\mathcal{X}$ equals almost surely the value $\tilde{x} = \mathrm{Tri}(1, 2, 3)$ on Ω and we consider $f(x) = x^{-1} - 1$ for all $x \in (0, 1)$, then we obtain that $\mathrm{I}_f(\mathcal{X}) = .099$.

Reasons justifying that some f-indices do not vanish for degenerate fuzzy random variables lie in the fact that several of these indices (in particular, those associated with functions f such that $f(u) + f(1/u) = 0$ if, and only if, $u = 1$), in addition to quantifying the *inter*values inequality also measures the *intra*values inequality. In this sense, and as a special case, one can prove the following *additive descomposition* property for the hyperbolic index I_H (that is, I_f with $f(x) = x^{-1} - 1$ for all $x \in (0, +\infty)$), in accordance with which if $\Omega = \{\omega_1, \ldots, \omega_N\}$, $\mathcal{X}(\Omega) = \{\tilde{x}_1^*, \ldots, \tilde{x}_r^*\}$ and $p_l = P(\{\omega \in \Omega \mid \mathcal{X}(\omega) = \tilde{x}_l^*\})$, $l = 1, \ldots, r$, we have that

$$\mathrm{I}_H(\mathcal{X}) = \sum_{l=1}^{r} p_l^2\, \mathrm{I}_H(\{\tilde{x}_l^*\}) + \mathrm{I}_H^{bv}(\mathcal{X})$$

(where $\mathrm{I}_H^{bv}(\mathcal{X})$ represents a kind of *inter*values inequality index and $\mathrm{I}_H(\{\tilde{x}_l^*\})$ denotes the *intra*value inequality index for value $\tilde{x}_l^*$) with $\mathrm{I}_H^{bv}(\mathcal{X}) = 0$ if, and only if, $\mathcal{X}$ is a degenerate fuzzy random variable. Nevertheless, this property is not true for all f.

When we deal with finite populations, there are more properties having a clear meaning and application. For this reason, for the remaining properties in this section we will consider a probability space $(\Omega, \mathcal{P}(\Omega), P)$ where Ω is the finite population $\{\omega_1, \ldots, \omega_N\}$ and P is the uniform distribution on Ω.

Theorem 2.2.7 (Expression for finite populations). *Let $\mathcal{X}$ be a positive fuzzy random variable defined on the population $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$. Then, we have that*

$$\mathrm{I}_f(\mathcal{X}) = \frac{1}{2}\int_{[0,1]} \left[\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right) + \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right)\right] d\alpha.$$

The *symmetry* of the f-inequality indices formalizes the fact that they do not depend on the identity or numbering of individuals in the population (that is, the f-inequality indices are objective measures). In this way,

Theorem 2.2.8 (Symmetry). *Let $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ be a fuzzy random variable. Then, $\mathrm{I}_f(\mathcal{X} \circ \sigma) = \mathrm{I}_f(\mathcal{X})$ for any permutation σ on Ω.*

Proof.

Indeed, $\widetilde{E}(\mathcal{X} \circ \sigma) = \widetilde{E}(\mathcal{X})$. Furthermore, for all $\alpha \in [0, 1]$, we obtain that

$$E\left(f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{\sigma(j)})}{\sup E((\mathcal{X} \circ \sigma)_\alpha)}\right)\right) = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{\sigma(j)})}{\sup E((\mathcal{X} \circ \sigma)_\alpha)}\right)$$

$$= \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right),$$

$$E\left(f\left(\frac{\sup \mathcal{X}_\alpha(\omega_{\sigma(j)})}{\inf E((\mathcal{X} \circ \sigma)_\alpha)}\right)\right) = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_{\sigma(j)})}{\inf E((\mathcal{X} \circ \sigma)_\alpha)}\right)$$

$$= \frac{1}{N} \sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right),$$

whence $\mathrm{I}_f(\mathcal{X} \circ \sigma) = \mathrm{I}_f(\mathcal{X})$. □

The *principle of population* (or *population homogeneity*) underlines the fact that the value of an f-inequality index of a given population Ω with respect to a positive fuzzy random variable coincides with that of the population $\Omega^{(r)} = \{\omega_{11}, \ldots, \omega_{1N}, \ldots, \omega_{r1}, \ldots, \omega_{rN}\}$ obtained from Ω by replicating it an arbitrary finite number r of times (i.e., $\omega_{ij} = \omega_j$ for all i, j), with respect to the immediate extension of this variable. In a more concise way, the f-inequality index only depends on the frequency of each possible value $\mathcal{X}$ in the population (that is, on the population structure) irrespectively of its size.

Theorem 2.2.9 (Population homogeneity). *If $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ is a fuzzy random variable and $\mathcal{X}^{(r)} : \Omega^{(r)} \to \mathcal{F}_c((0, +\infty))$ is the fuzzy random variable extending $\mathcal{X}$ to $\Omega^{(r)}$, then $\mathrm{I}_f(\mathcal{X}^{(r)}) = \mathrm{I}_f(\mathcal{X})$.*

Proof.

Since $\mathcal{X}^{(r)}$ is also defined on a finite population, then for all $\alpha \in [0, 1]$ we have that

$$\frac{1}{N\,r} \sum_{i=1}^{r} \sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha^{(r)}(\omega_{ij})}{\frac{1}{N\,r}\sum_{i=1}^{r}\sum_{j=1}^{N} \sup \mathcal{X}_\alpha^{(r)}(\omega_{ij})}\right)$$

$$= \frac{1}{N\,r} \sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\frac{1}{N\,r}\sum_{j=1}^{N} r \sup \mathcal{X}_\alpha(\omega_j)}\right) \cdot r = \frac{1}{N} \sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\frac{1}{N}\sum_{j=1}^{N} \sup \mathcal{X}_\alpha(\omega_j)}\right),$$

$$\frac{1}{N\,r}\sum_{i=1}^{r}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_{\alpha}^{(r)}(\omega_{ij})}{\dfrac{1}{N\,r}\displaystyle\sum_{i=1}^{r}\sum_{j=1}^{N}\inf \mathcal{X}_{\alpha}^{(r)}(\omega_{ij})}\right)$$

$$= \frac{1}{N\,r}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_{\alpha}(\omega_{j})}{\dfrac{1}{N\,r}\displaystyle\sum_{j=1}^{N} r\inf \mathcal{X}_{\alpha}(\omega_{j})}\right)\cdot r = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_{\alpha}(\omega_{j})}{\dfrac{1}{N}\displaystyle\sum_{j=1}^{N}\inf \mathcal{X}_{\alpha}(\omega_{j})}\right),$$

and this guarantees the equality of the f-inequality index for both populations. □

Another relevant property of the family of f-inequality indices is the *continuity,* in accordance with which "small" (fuzzy) changes in the values of the fuzzy random variables entail "small" (real-valued) variations in the f-inequality indices. To formalize this property we will make use of the metric on the basis of which the measurability of fuzzy random variables have been stated. Thus,

Theorem 2.2.10 (Continuity). *Let $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ be a fuzzy random variable and let $\mathcal{X}_{l,\widetilde{\mathcal{E}}}$ be a fuzzy random variable defined on Ω such that $\mathcal{X}_{l,\widetilde{\mathcal{E}}}(\omega_j) = \mathcal{X}(\omega_j)$ for all $j \in \{1,\ldots,N\}\setminus\{l\}$ and $\mathcal{X}_{l,\widetilde{\mathcal{E}}}(\omega_l) = \mathcal{X}(\omega_l)\oplus\widetilde{\mathcal{E}}$ with $\widetilde{\mathcal{E}} \in \mathcal{F}_c\big(\mathbb{R}\big)$ such that $\mathcal{X}(\omega_l)\oplus\widetilde{\mathcal{E}} \in \mathcal{F}_c((0,+\infty))$. Then,*

$$\lim_{d_\infty(\widetilde{\mathcal{E}},\mathbf{1}_{\{0\}})\to 0} \mathrm{I}_f(\mathcal{X}_{l,\widetilde{\mathcal{E}}}) = \mathrm{I}_f(\mathcal{X}).$$

Proof.

Indeed,

$$d_\infty(\widetilde{\mathcal{E}},\mathbf{1}_{\{0\}}) = \sup_{\alpha\in[0,1]} \max\{|\sup\widetilde{\mathcal{E}}_\alpha|, |\inf\widetilde{\mathcal{E}}_\alpha|\},$$

and, furthermore,

$$|\mathrm{I}_f(\mathcal{X}_{l,\widetilde{\mathcal{E}}}) - \mathrm{I}_f(\mathcal{X})| = \left|\frac{1}{N}\sum_{j=1}^{N}\int_{[0,1]}\frac{1}{2}\left[f\left(\frac{\inf(\mathcal{X}_{l,\widetilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\widetilde{\mathcal{E}}})_\alpha)}\right) - f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right)\right]\right.$$

$$+\frac{1}{2}\left[f\left(\frac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}\right)-f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right)\right]d\,\alpha\Bigg|.$$

For all $\alpha \in [0,1]$ and $j \in \{1,\ldots,N\}$, one can immediately prove that

$$\left|\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha - \inf \mathcal{X}_\alpha(\omega_j)\right| \leq |\inf \widetilde{\mathcal{E}}_\alpha| \leq d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}}),$$

$$\left|\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha - \sup \mathcal{X}_\alpha(\omega_j)\right| \leq |\sup \widetilde{\mathcal{E}}_\alpha| \leq d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}}),$$

whence $d_\infty(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j), \mathcal{X}(\omega_j)) \leq d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}})$, and

$$\left|\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha) - \inf E(\mathcal{X}_\alpha)\right| \leq |\inf \widetilde{\mathcal{E}}_\alpha|/N \leq d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}})/N,$$

$$\left|\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha) - \sup E(\mathcal{X}_\alpha)\right| \leq |\sup \widetilde{\mathcal{E}}_\alpha|/N \leq d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}})/N,$$

and, therefore, $d_\infty(\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}}), \widetilde{E}(\mathcal{X})) \leq d_\infty(\widetilde{\mathcal{E}}, \mathbf{1}_{\{0\}})/N$.

Consequently, we have that for all $j \in \{1, \ldots, N\}$

$$\lim_{d_\infty(\tilde{\mathcal{E}},\mathbf{1}_{\{0\}})\to 0} d_\infty\left(\frac{\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j)}{\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}})}, \frac{\mathcal{X}(\omega_j)}{\widetilde{E}(\mathcal{X})}\right) = 0.$$

Moreover, in virtue of the Mean Value Theorem of the Differential Calculus, and because of the strict convexity and monotonicity of f and $f \in C^1$, we have that

$$\left|f\left(\frac{\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}\right)-f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right)\right|$$

$$= |f'(c_{1j}(\alpha))|\left|\frac{\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)} - \frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right|,$$

and, similarly,

$$\left|f\left(\frac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}\right)-f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right)\right|$$

$$= |f'(c_{2j}(\alpha))|\left|\frac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)} - \frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right|,$$

where c_{1j} and c_{2j} are in between $\dfrac{\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}$ and $\dfrac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}$, and in between $\dfrac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}$ and $\dfrac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}$, respectively.

Since $f \in C^1$, if $[a_j, b_j]$ is an element in $\mathcal{K}_c((0, +\infty))$ containing the values $\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_0 / \inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_0)$ and $\sup \mathcal{X}_0(\omega_j)/\inf E(\mathcal{X}_0)$, if k_j is the maximum of the derivative function f' on $[a_j, b_j]$ one has that

$$\left| f\left(\frac{\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}\right) - f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right)\right|$$

$$\leq k_j \left|\frac{\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)} - \frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right|,$$

and

$$\left| f\left(\frac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}\right) - f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right)\right|$$

$$\leq k_j \left|\frac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)} - \frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right|.$$

Therefore, for all $j \in \{1, \ldots, N\}$

$$d_H\left(f\left(\frac{(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{(\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}}))_\alpha}\right), f\left(\frac{\mathcal{X}_\alpha(\omega_j)}{E(\mathcal{X}_\alpha)}\right)\right)$$

$$\leq d_\infty\left(f\left(\frac{\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j)}{\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}})}\right), f\left(\frac{\mathcal{X}(\omega_j)}{\widetilde{E}(\mathcal{X})}\right)\right)$$

$$\leq k_j d_\infty\left(\frac{\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j)}{\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}})}, \frac{\mathcal{X}(\omega_j)}{\widetilde{E}(\mathcal{X})}\right),$$

and, hence, for all $j \in \{1, \ldots, N\}$ we have that

$$\lim_{d_\infty(\tilde{\mathcal{E}}, \mathbf{1}_{\{0\}}) \to 0} d_\infty\left(f\left(\frac{\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j)}{\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}})}\right), f\left(\frac{\mathcal{X}(\omega_j)}{\widetilde{E}(\mathcal{X})}\right)\right) = 0.$$

Consequently, as $\alpha \in [0,1]$ and $j \in \{1, \ldots, N\}$ it satisfies that

$$\frac{1}{2}\left[f\left(\frac{\inf(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\sup E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}\right) - f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right)\right]$$
$$+\frac{1}{2}\left[f\left(\frac{\sup(\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j))_\alpha}{\inf E((\mathcal{X}_{l,\tilde{\mathcal{E}}})_\alpha)}\right) - f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right)\right]$$
$$\leq d_\infty\left(f\left(\frac{\mathcal{X}_{l,\tilde{\mathcal{E}}}(\omega_j)}{\widetilde{E}(\mathcal{X}_{l,\tilde{\mathcal{E}}})}\right), f\left(\frac{\mathcal{X}(\omega_j)}{\widetilde{E}(\mathcal{X})}\right)\right),$$

whence we conclude that

$$\lim_{d_\infty(\tilde{\mathcal{E}},\mathbf{1}_{\{0\}})\to 0} \mathrm{I}_f(\mathcal{X}_{l,\tilde{\mathcal{E}}}) = \mathrm{I}_f(\mathcal{X}).$$

□

Some of the main and desirable properties of the inequality indices are those concerning the reaction of these indices to a "redistribution" of the values of the considered attribute.

The *strict Schur-convexity* of I_f, formalizes the fact that when the redistribution is carried out by considering convex linear combinations, the f-inequality index of the fuzzy random variable cannot increase in the population. Thus,

Theorem 2.2.11 (Strict Schur-convexity). *If $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ is a fuzzy random variable, (μ_{jl}) is an $N \times N$ doubly stochastic matrix and $\mathcal{X}' : \Omega \to \mathcal{F}_c((0,+\infty))$ is another fuzzy random variable defined from $\mathcal{X}$ as follows:*

$$\begin{pmatrix} \mathcal{X}'(\omega_1) \\ \mathcal{X}'(\omega_2) \\ \vdots \\ \mathcal{X}'(\omega_N) \end{pmatrix} = \begin{pmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1N} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{N1} & \mu_{N2} & \cdots & \mu_{NN} \end{pmatrix} \odot \begin{pmatrix} \mathcal{X}(\omega_1) \\ \mathcal{X}(\omega_2) \\ \vdots \\ \mathcal{X}(\omega_N) \end{pmatrix},$$

then, we have that $\mathrm{I}_f(\mathcal{X}) \geq \mathrm{I}_f(\mathcal{X}')$ with equality if, and only if, $\mathcal{X}' = \mathcal{X} \circ \sigma$ for certain permutation σ on Ω.

Proof.

First of all, we are going to prove that under the assumed conditions $\widetilde{E}(\mathcal{X}') = \widetilde{E}(\mathcal{X})$. To this purpose, we just consider that for all $j \in \{1, \ldots, N\}$ and $\alpha \in [0,1]$ we have that

$$\inf \mathcal{X}'_\alpha(\omega_j) = \sum_{l=1}^{N} \mu_{jl} \inf \mathcal{X}_\alpha(\omega_j),$$

$$\sup \mathcal{X}'_\alpha(\omega_j) = \sum_{l=1}^{N} \mu_{jl} \sup \mathcal{X}_\alpha(\omega_j).$$

Consequently, for all $\alpha \in [0,1]$ we have that

$$E\left(\inf \mathcal{X}'_\alpha\right) = \frac{1}{N}\sum_{j=1}^{N} \inf \mathcal{X}'_\alpha(\omega_j) = \frac{1}{N}\sum_{j=1}^{N}\sum_{l=1}^{N} \mu_{jl} \inf \mathcal{X}_\alpha(\omega_l)$$

$$= \frac{1}{N}\sum_{l=1}^{N}\left(\inf \mathcal{X}_\alpha(\omega_l)\sum_{j=1}^{N}\mu_{jl}\right) = \frac{1}{N}\sum_{l=1}^{N} \inf \mathcal{X}_\alpha(\omega_l) = E\left(\inf \mathcal{X}_\alpha\right),$$

and, analogously, for all $\alpha \in [0,1]$

$$\sup E\left(\mathcal{X}'_\alpha\right) = \sup E\left(\mathcal{X}_\alpha\right),$$

whence $\widetilde{E}(\mathcal{X}') = \widetilde{E}(\mathcal{X})$.

Furthermore, for all $\alpha \in [0,1]$ the application of Jensen's inequality along with the double stochasticity of the matrix (μ_{jl}) guarantee that

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_j)}{\sup E(\mathcal{X}'_\alpha)}\right) = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\displaystyle\sum_{l=1}^{N}\mu_{jl}\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)}\right)$$

$$\leq \frac{1}{N}\sum_{j=1}^{N}\sum_{l=1}^{N}\mu_{jl}\, f\left(\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)}\right) = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right),$$

and, in the same way,

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}'_\alpha(\omega_j)}{\inf E(\mathcal{X}'_\alpha)}\right) \leq \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right),$$

whence $\mathrm{I}_f(\mathcal{X}) \geq \mathrm{I}_f(\mathcal{X}')$.

On the other hand, the equality is achieved if, and only if, after properly permuting the elements in Ω (what does not entail any change in the value of the f-inequality index, because of the symmetry stated in Theorem 2.2.8) and the corresponding columns in the matrix (μ_{jl}), we have that either this matrix reduces to the identity one, or $\mu_{ll} \in (0,1)$ corresponds to the individuals $\omega_l \in \Omega$ for which $\mathcal{X}(\omega_l)$ coincides with $\mathcal{X}(\omega_j)$ for all the $j \neq l$ such that $\mu_{jl} \in (0,1)$ or $\mu_{lj} \in (0,1)$, whence $\mathcal{X}$ and $\mathcal{X}'$ will take on the same values in ω_l and ω_j (may be permuted from the original population). □

The well-known Lorenz criterion of the real-valued case could be extended to ordering the vectors of the values of a fuzzy random variable on a finite population by considering the Ramík and Římánek (1985) ranking. Thus, if we denote $V \succ_S W$ if, and only if, $V \succeq_S W$ but not $V \preceq_S W$, the *compatibility with Lorenz's criterion* of the f-inequality indices can be stated as follows:

Theorem 2.2.12 (Compatibility with Lorenz's criterion). *Consider two fuzzy random variables $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ and $\mathcal{X}' : \Omega \to \mathcal{F}_c((0,+\infty))$ such that $\mathcal{X}(\omega_N) \succeq_S \ldots \succeq_S \mathcal{X}(\omega_1), \mathcal{X}'(\omega_N) \succeq_S \ldots \succeq_S \mathcal{X}'(\omega_1)$, $\widetilde{E}(\mathcal{X}) = \widetilde{E}(\mathcal{X}')$ and $\mathcal{X} L \mathcal{X}'$ (which will mean that $\mathcal{X}(\omega_1) \oplus \cdots \oplus \mathcal{X}(\omega_k) \succeq_S \mathcal{X}'(\omega_1) \oplus \cdots \oplus \mathcal{X}'(\omega_k)$ for all $k \in \{1,\ldots,N\}$ with $\succ_S$ for at least one k). Then, $\mathrm{I}_f(\mathcal{X}) < \mathrm{I}_f(\mathcal{X}')$.*

Proof.

If $\mathcal{X} L\, \mathcal{X}'$, we have that the real-valued random variables $\sup \mathcal{X}_\alpha$ and $\sup \mathcal{X}'_\alpha$ preserve the same relation (with the classical Lorenz criterion), that is $(\sup \mathcal{X}_\alpha)\, L\, (\sup \mathcal{X}'_\alpha)$ and, analogously, $(\inf \mathcal{X}_\alpha)\, L\, (\inf \mathcal{X}'_\alpha)$ for all $\alpha \in [0,1]$.

In virtue of a well-known result (see, for instance, Dasgupta *et al.* 1973, Marshall & Olkin 1979, and Eichhorn & Gehrig 1982), we then have that for all $\alpha \in [0,1]$ there exist two doubly stochastic matrices $(\mu_{jl}^{\sup}(\alpha))$ and $(\mu_{jl}^{\inf}(\alpha))$ such that

$$\begin{pmatrix} \sup \mathcal{X}_\alpha(\omega_1) \\ \vdots \\ \sup \mathcal{X}_\alpha(\omega_N) \end{pmatrix} = \begin{pmatrix} \mu_{11}^{\sup}(\alpha) & \cdots & \mu_{1N}^{\sup}(\alpha) \\ \vdots & \ddots & \vdots \\ \mu_{N1}^{\sup}(\alpha) & \cdots & \mu_{NN}^{\sup}(\alpha) \end{pmatrix} \odot \begin{pmatrix} \sup \mathcal{X}'_\alpha(\omega_1) \\ \vdots \\ \sup \mathcal{X}'_\alpha(\omega_N) \end{pmatrix},$$

and

$$\begin{pmatrix} \inf \mathcal{X}_\alpha(\omega_1) \\ \vdots \\ \inf \mathcal{X}_\alpha(\omega_N) \end{pmatrix} = \begin{pmatrix} \mu_{11}^{\inf}(\alpha) & \cdots & \mu_{1N}^{\inf}(\alpha) \\ \vdots & \ddots & \vdots \\ \mu_{N1}^{\inf}(\alpha) & \cdots & \mu_{NN}^{\inf}(\alpha) \end{pmatrix} \odot \begin{pmatrix} \inf \mathcal{X}'_\alpha(\omega_1) \\ \vdots \\ \inf \mathcal{X}'_\alpha(\omega_N) \end{pmatrix}.$$

Consequently, and by applying Jensen's inequality, we have that

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right) = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\displaystyle\sum_{l=1}^{N} \mu_{jl}^{inf}(\alpha) \inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)}\right)$$

$$\leq \frac{1}{N}\sum_{j=1}^{N}\sum_{l=1}^{N} \mu_{jl}^{\inf}(\alpha) f\left(\frac{inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)}\right) = \frac{1}{N}\sum_{l=1}^{N} f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)}\right),$$

with equality if, and only if, either $\left(\mu_{jl}^{\inf}(\alpha)\right)$ is the $N \times N$ identity matrix or $\mu_{ll}^{\inf}(\alpha) \in (0,1)$ corresponds to the individuals $\omega_l \in \Omega$ for which $\mathcal{X}(\omega_l)$ coincides with $\mathcal{X}(\omega_j)$ for all the $j \neq l$ such that $\mu_{jl}^{\inf}(\alpha) \in (0,1)$ or $\mu_{lj}^{\inf}(\alpha) \in (0,1)$, whence $\inf \mathcal{X}_\alpha$ and $\inf \mathcal{X}'_\alpha$ will take on the same values in ω_l and ω_j (may be permuted from the original population).

Therefore, for all $\alpha \in [0,1]$ we have that

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right) < \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_j)}{\sup E(\mathcal{X}'_\alpha)}\right).$$

Similarly, we have that

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right) < \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}'_\alpha(\omega_j)}{\inf E(\mathcal{X}'_\alpha)}\right),$$

whence $\mathrm{I}_f(\mathcal{X}) < \mathrm{I}_f(\mathcal{X}')$. □

The *progressive and regressive principles of transfers* formalize the reaction to redistributions which are carried out by means of transfers from a higher value to a lower one (or conversely) when preserving the expected value of the fuzzy random variable. More precisely, the progressive principle of transfers indicates that

Theorem 2.2.13 (Progressive principle of transfers). *If* $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ *is a fuzzy random variable, and* $\mathcal{X}' : \Omega \to \mathcal{F}_c((0,+\infty))$ *is another fuzzy random variable such that* $\mathcal{X}(\omega_j) = \mathcal{X}'(\omega_j)$ *for all* $j \in \{1,\ldots,N\} \setminus \{l,l'\}$, $\mathcal{X}(\omega_l) \succeq_S \mathcal{X}'(\omega_l) \succeq_S \mathcal{X}(\omega_{l'})$, $\mathcal{X}(\omega_l) \succeq_S \mathcal{X}'(\omega_{l'}) \succeq_S \mathcal{X}(\omega_{l'})$ *and* $\widetilde{E}(\mathcal{X}) = \widetilde{E}(\mathcal{X}')$. *Then, we have that* $\mathrm{I}_f(\mathcal{X}) \geq \mathrm{I}_f(\mathcal{X}')$.

Furthermore, $\mathrm{I}_f(\mathcal{X}) = \mathrm{I}_f(\mathcal{X}')$ *if, and only if, either* $\mathcal{X} = \mathcal{X}'$ *on* Ω, *or* $\mathcal{X}' = \mathcal{X} \circ \sigma_{ll'}$ *for the permutation* $\sigma_{ll'}$ *on* Ω *which exchanges* ω_l *and* $\omega_{l'}$, *or for all* $\alpha \in [0,1]$ *it happens that* $\inf \mathcal{X}_\alpha(\omega_l) = \inf \mathcal{X}'_\alpha(\omega_{l'})$, $\inf \mathcal{X}_\alpha(\omega_{l'}) = \inf \mathcal{X}'_\alpha(\omega_l)$, $\sup \mathcal{X}_\alpha(\omega_l) = \sup \mathcal{X}'_\alpha(\omega_l)$ *and* $\sup \mathcal{X}_\alpha(\omega_{l'}) = \sup \mathcal{X}'_\alpha(\omega_{l'})$, *or* $\inf \mathcal{X}_\alpha(\omega_l) = \inf \mathcal{X}'_\alpha(\omega_l)$, $\inf \mathcal{X}_\alpha(\omega_{l'}) = \inf \mathcal{X}'_\alpha(\omega_{l'})$, $\sup \mathcal{X}_\alpha(\omega_l) = \sup \mathcal{X}'_\alpha(\omega_{l'})$ *and* $\sup \mathcal{X}_\alpha(\omega_{l'}) = \sup \mathcal{X}'_\alpha(\omega_l)$.

Proof.

Indeed, under the assumed hypotheses, $\sup E(\mathcal{X}'_\alpha) = \sup E(\mathcal{X}_\alpha)$, and we have that

$$\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)} \geq \frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)} \geq \frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)},$$

$$\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)} \geq \frac{\inf \mathcal{X}'_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)} \geq \frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)},$$

$$\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)} + \frac{\inf \mathcal{X}'_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)} = \frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)} + \frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)},$$

whence there will be $\lambda_1(\alpha) \in [0,1]$ such that

$$\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)} = \lambda_1(\alpha)\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)} + (1-\lambda_1(\alpha))\frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)},$$

$$\frac{\inf \mathcal{X}'_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)} = (1-\lambda_1(\alpha))\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)} + \lambda_1(\alpha)\frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)}.$$

In virtue of Jensen's inequality, we can conclude that

$$f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)}\right) \le \lambda_1(\alpha) f\left(\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)}\right) + (1-\lambda_1(\alpha)) f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)}\right),$$

$$f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)}\right) \le (1-\lambda_1(\alpha)) f\left(\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)}\right) + \lambda_1(\alpha) f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)}\right),$$

whence

$$f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)}\right) + f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)}\right) \le f\left(\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)}\right) + f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}_\alpha)}\right),$$

and, because of the preceding inequalities, we have that

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_j)}{\sup E(\mathcal{X}'_\alpha)}\right) = \frac{1}{N}\left[\left(\sum_{\substack{j=1\\ j\neq l,l'}}^{N} f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_j)}{\sup E(\mathcal{X}'_\alpha)}\right)\right)\right.$$

$$\left. + f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_l)}{\sup E(\mathcal{X}'_\alpha)}\right) + f\left(\frac{\inf \mathcal{X}'_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)}\right)\right]$$

$$\le \frac{1}{N}\left[\left(\sum_{\substack{j=1\\ j\neq l,l'}}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right)\right) + f\left(\frac{\inf \mathcal{X}_\alpha(\omega_l)}{\sup E(\mathcal{X}_\alpha)}\right) + f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{l'})}{\sup E(\mathcal{X}'_\alpha)}\right)\right]$$

$$= \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right).$$

Analogously, and by using another value $\lambda_2(\alpha) \in [0,1]$, we obtain

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}'_\alpha(\omega_j)}{\inf E(\mathcal{X}'_\alpha)}\right) \le \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right),$$

and, therefore, $\mathrm{I}_f(\mathcal{X}) \ge \mathrm{I}_f(\mathcal{X}')$.

The sufficiency of the condition announced for the equality $\mathrm{I}_f(\mathcal{X}) = \mathrm{I}_f(\mathcal{X}')$ is obvious. The necessity of this condition can be immediately deduced by assuming that it is not satisfied, whence (because of the left-continuity of $\lambda_1(\alpha)$ and $\lambda_2(\alpha)$ in $[0,1]$) there exists $\alpha \in [0,1]$ such that either $\lambda_1(\alpha) \in (0,1)$ or $\lambda_2(\alpha) \in (0,1)$, and hence $\mathrm{I}_f(\mathcal{X}) > \mathrm{I}_f(\mathcal{X}')$. □

On the other hand, the regressive principle of transfers can be derived from the last one by exchanging the roles of $\mathcal{X}$ and $\mathcal{X}'$ in it.

Theorem 2.2.14 (Regressive principle of transfers). *If* $\mathcal{X} : \Omega \to \mathcal{F}_c((0,+\infty))$ *is a fuzzy random variable, and* $\mathcal{X}' : \Omega \to \mathcal{F}_c((0,+\infty))$ *is another fuzzy random variable such that* $\mathcal{X}(\omega_j) = \mathcal{X}'(\omega_j)$ *for all* $j \in \{1,\dots,N\} \setminus \{l,l'\}$, $\mathcal{X}'(\omega_l) \succeq_S \mathcal{X}(\omega_l) \succeq_S \mathcal{X}'(\omega_{l'})$, $\mathcal{X}'(\omega_l) \succeq_S \mathcal{X}(\omega_{l'}) \succeq_S \mathcal{X}'(\omega_{l'})$ *and* $\widetilde{E}(\mathcal{X}) = \widetilde{E}(\mathcal{X}')$. *Then, we have that* $\mathrm{I}_f(\mathcal{X}') \ge \mathrm{I}_f(\mathcal{X})$.

Furthermore, $\mathrm{I}_f(\mathcal{X}) = \mathrm{I}_f(\mathcal{X}')$ *if, and only if, either* $\mathcal{X} = \mathcal{X}'$ *on* Ω, *or* $\mathcal{X}' = \mathcal{X} \circ \sigma_{ll'}$ *(with the notation in Theorem* 2.2.13*), or for all* $\alpha \in [0,1]$ *we have that either* $\inf \mathcal{X}'_\alpha(\omega_l) = \inf \mathcal{X}_\alpha(\omega_{l'})$, $\inf \mathcal{X}'_\alpha(\omega_{l'}) = \inf \mathcal{X}_\alpha(\omega_l)$, $\sup \mathcal{X}'_\alpha(\omega_l) = \sup \mathcal{X}_\alpha(\omega_l)$ *and* $\sup \mathcal{X}'_\alpha(\omega_{l'}) = \sup \mathcal{X}_\alpha(\omega_{l'})$, *or* $\inf \mathcal{X}'_\alpha(\omega_l) = \inf \mathcal{X}_\alpha(\omega_l)$, $\inf \mathcal{X}'_\alpha(\omega_{l'}) = \inf \mathcal{X}_\alpha(\omega_{l'})$, $\sup \mathcal{X}'_\alpha(\omega_l) = \sup \mathcal{X}_\alpha(\omega_{l'})$ *and* $\sup \mathcal{X}'_\alpha(\omega_{l'}) = \sup \mathcal{X}_\alpha(\omega_l)$.

The following result corresponds to a property formalizing the *effects* of the *"grouping" of fuzzy data* in quantifying the f-inequality index. More precisely, this property expresses the "ordering relation" between the inequality of the population and the inequality between the groups of a given (classical) partition of the population, when each of the groups is represented by the expected value of the fuzzy random variable in it. From this result we can conclude that grouping entails an increase in inequality.

Theorem 2.2.15 (Grouping effects). *Consider a finite population* $\Omega = \{\omega_{11}, \ldots, \omega_{1N_1}, \ldots\ldots, \omega_{M1}, \ldots, \omega_{MN_M}\}$ *(with* $N = N_1 + \cdots + N_M$*) which is divided into* M *subpopulations* $\Omega_m = \{\omega_{m1}, \ldots, \omega_{mN_m}\}$, $m = 1, \ldots, M$, *and assume that* $(\Omega, \mathcal{P}(\Omega))$ *is endowed with the uniform distribution* P *and that* $\mathcal{P} = \{\Omega_m\}_{m=1}^{M}$ *denotes the above partition. If* $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ *is a fuzzy random variable associated with* $(\Omega, \mathcal{P}(\Omega), P)$*, and* $\mathcal{X}_{\mathcal{P}} : \mathcal{P} \to \mathcal{F}_c((0, +\infty))$ *is the fuzzy random variable such that* $\mathcal{X}_{\mathcal{P}}(\Omega_m)$ = *expected value of* $\mathcal{X}$ *on* Ω_m *(*$m = 1, \ldots, M$*), and* $\mathcal{X}_{\Omega_m}$ *denotes the restriction of* $\mathcal{X}$ *from* Ω *to* Ω_m *(*$m = 1, \ldots, M$*), then we have that*

$$\mathrm{I}_f(\mathcal{X}) \geq \mathrm{I}_f(\mathcal{X}_{\mathcal{P}}).$$

On the other hand, $\mathrm{I}_f(\mathcal{X}) = \mathrm{I}_f(\mathcal{X}_{\mathcal{P}})$ *if, and only if, for each* $m \in \{1, \ldots, M\}$ *the fuzzy random variable* $\mathcal{X}_{\Omega_m}$ *is degenerate in* Ω_m*.*

Proof.

Obviously, $\widetilde{E}(\mathcal{X}) = \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)$. On the other hand, by applying Jensen's inequality, we obtain that

$$\sum_{m=1}^{M} \frac{N_m}{N} f\left(\frac{\inf \mathcal{X}_{\mathcal{P}}(\Omega_m)_\alpha}{\sup \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)_\alpha}\right) = \sum_{m=1}^{M} \frac{N_m}{N} f\left(\frac{\frac{1}{N_m}\sum_{i=1}^{N_m} \inf \mathcal{X}_\alpha(\omega_{mi})}{\sup E(\mathcal{X}_\alpha)}\right)$$

$$\leq \frac{1}{N}\sum_{m=1}^{M}\sum_{i=1}^{N_m} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_{mi})}{\sup E(\mathcal{X}_\alpha)}\right) = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right).$$

Analogously,

$$\sum_{m=1}^{M} \frac{N_m}{N} f\left(\frac{\sup \mathcal{X}_{\mathcal{P}}(\Omega_m)_\alpha}{\inf \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)_\alpha}\right) \leq \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right),$$

whence

$$\mathrm{I}_f(\mathcal{X}) = \frac{1}{2}\int_{[0,1]}\left[\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right) + \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right)\right] d\,\alpha$$

$$\geq \frac{1}{2}\int_{[0,1]}\left[\sum_{m=1}^{M}\frac{N_m}{N} f\left(\frac{\sup \mathcal{X}_{\mathcal{P}}(\Omega_m)_\alpha}{\inf \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)_\alpha}\right)\right.$$

$$\left. + \sum_{m=1}^{M}\frac{N_m}{N} f\left(\frac{\inf \mathcal{X}_{\mathcal{P}}(\Omega_m)_\alpha}{\sup \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)_\alpha}\right)\right] d\,\alpha = \mathrm{I}_f(\mathcal{X}_{\mathcal{P}}).$$

The equality will be achieved if, and only if,

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\sup \mathcal{X}_\alpha(\omega_j)}{\inf E(\mathcal{X}_\alpha)}\right) + \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{\inf \mathcal{X}_\alpha(\omega_j)}{\sup E(\mathcal{X}_\alpha)}\right)$$

$$= \sum_{m=1}^{M}\frac{N_m}{N} f\left(\frac{\sup \mathcal{X}_{\mathcal{P}}(\Omega_m)_\alpha}{\inf \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)_\alpha}\right) + \sum_{m=1}^{M}\frac{N_m}{N} f\left(\frac{\inf \mathcal{X}_{\mathcal{P}}(\Omega_m)_\alpha}{\sup \widetilde{E}\left(\mathcal{X}_{\mathcal{P}}(\Omega_m)\right)_\alpha}\right),$$

for all $\alpha \in [0,1]$, and this happens if, and only if, all the expressions to which we have applied Jensen's inequality become equalities. This condition is equivalent to the fact that for each $\alpha \in [0,1]$, and whatever Ω_m may be $m \in \{1,\ldots,M\}$, the values $\inf \mathcal{X}_\alpha(\omega_{mi})$ coincide for all different values of $i = 1,\ldots,N_m$ and the values $\sup \mathcal{X}_\alpha(\omega_{mi})$ also coincide for different $i = 1,\ldots,N_m$, that is, if the fuzzy random variable is degenerate in Ω_m. □

Remark 2.2.2. The *additive decomposability* of the indices in Page 19 in the introduction of the present work is lost in the extension to the fuzzy case, except for Shannon's index. In this way, if $f(x) = -\log x$ for all $x \in (0,+\infty)$, it can be easily proven for $\mathrm{I}_{Sh}(\mathcal{X})$ (or extended Shannon's type index) that

$$\mathrm{I}_{Sh}(\mathcal{X}) = \mathrm{I}_{Sh}(\mathcal{X}_{\mathcal{P}}) + \sum_{m=1}^{M}\frac{N_m \mathrm{I}_{Sh}(\mathcal{X}_{\Omega_m})}{N},$$

that is, the inequality in the population coincides with the sum of the inequality between groups (more concretely, between the expected values of $\mathcal{X}$ in different groups) and the average of the inequality within groups.

In the following examples we illustrate the computation and use of certain f-inequality indices in comparing populations.

Example 2.2.1. A phone poll is developed on the population Ω of the 105 male members of a sports center who are requested to classify themselves into one of the following four groups: SHORT, NOT TALL, TALL and VERY TALL. Assume that the obtained answers are 12 'SHORT', 23 'NOT TALL', 57 'TALL' and 13 'VERY TALL'. This type of classification can be identified with a fuzzy random variable whose values are the preceding four groups. Suppose that to describe these values we use the characterization given by Norwich & Turksen (1984), which is based on mean direct rating for certain referential points in the interval $[54, 88]$ (under the assumption that units considered are inches) and on a linear interpolation for the remaining points in $\mathbb{R}$, that supplies us the poligonals in Figure 2.1.

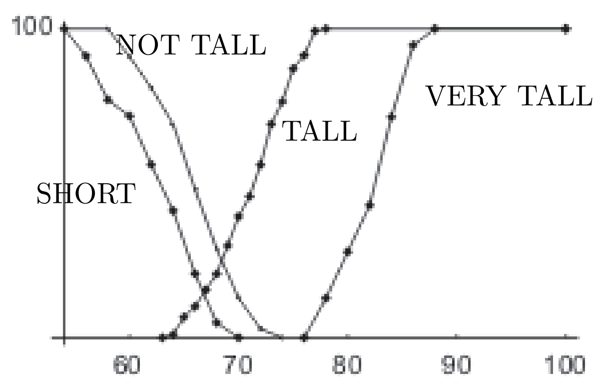


Fig. 2.1: Graphical representation of the value TALL and related values

If we want to quantify the inequality of heights in this population on the basis of the performed poll, we can consider, for instance, the f-inequality index extending the normalized variance for the real-value case (corresponding to $f(x) = x^2 - 1$) which takes on the value

$$\mathrm{I}_{\mathrm{NVar}}(\mathcal{X}_\Omega) = .04.$$

Example 2.2.2. Consider the variable ANNUAL INCOME, $\mathcal{X}$, in accordance with the classification which is adopted in some credit assessment systems. Following Cox (1994), this variable can be viewed as a variable whose (fuzzy) values are $\tilde{x}_1$ = SOMEWHAT HIGH, $\tilde{x}_2$ = MODERATELY HIGH, $\tilde{x}_3$ = HIGH and $\tilde{x}_4$ = VERY HIGH, where $\tilde{x}_1$, $\tilde{x}_2$, $\tilde{x}_3$ and $\tilde{x}_4$ are described (in US thousand dollars) by means of the following S- and Π-curves:

$$\tilde{x}_1 = 1 - S(100, 125),$$
$$\tilde{x}_2 = \Pi(100, 125, 150),$$
$$\tilde{x}_3 = \Pi(125, 147.5, 170)$$

and

$$\tilde{x}_4 = S(147.5, 170),$$

and $\operatorname{supp} \tilde{x}_i \subset [90, 180]$, $i = 1, 2, 3, 4$, for all candidates for a credit in the considered system (see Figure 2.2).

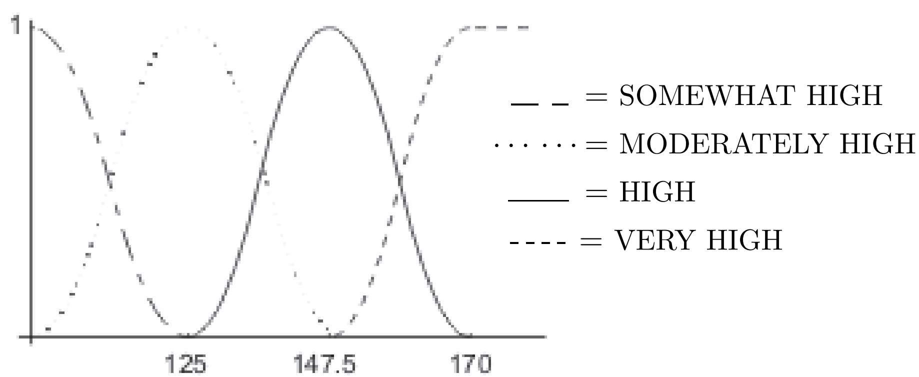


Fig. 2.2: Fuzzy values of the variable ANNUAL INCOME

Assume that a bank adopting the above system wishes to compare two different towns by means of the income inequality, and to this purpose we observe the values of $\mathcal{X}$ in the central offices of these two towns.

If there are 125 candidates for a credit in one of the offices (Ω_1) during a certain period, 28 of them having a SOMEWHAT HIGH annual income, 43 MODERATELY HIGH, 31 HIGH and 23 VERY HIGH, whereas there are 178 candidates for a credit in the other office (Ω_2) during the same period, 63 of them having SOMEWHAT HIGH, 79 MODERATELY HIGH, 27 HIGH and 9 VERY HIGH, and we employ the f-inequality index with $f(x) = -\log x$, we obtain that

$$\mathrm{I}_{Sh}(\mathcal{X}_{\Omega_1}) = .01224,$$
$$\mathrm{I}_{Sh}(\mathcal{X}_{\Omega_2}) = .009,$$

whence we can conclude that the two towns have a close inequality of annual income.

2.3 Estimating the hyperbolic index in random samplings from finite populations

In this section we consider the problem of estimating the population inequality index associated with a fuzzy random variable in a finite population in random samplings from finite populations.

To this purpose, we are going to check that it is possible to construct an unbiased estimator of hyperbolic index (associated with the function $f(x) = x^{-1} - 1$), for samples of any size in samplings with and without replacement. However, for the other inequality indices introduced in this chapter, the construction of unbiased estimators is very complex and often is not feasible.

Consider a finite population Ω of N units, $\omega_1, \ldots, \omega_N$, and let $\mathcal{X} : \Omega \to \mathcal{F}_c((0, +\infty))$ be a fuzzy random variable associated with a probabilistic space defined on Ω which is endowed with the uniform distribution.

Assume that a sample of size n is chosen at random and without replacement from Ω, υ denotes a generic simple random sample of size n, and $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ are the units in it. Then, the *sample hyperbolic index* of $\mathcal{X}$ in υ is given by

$$\mathrm{I}_H(\mathcal{X}[\upsilon]) = F\left(\frac{1}{n^2} \odot \sum_{i=1}^{n}\sum_{i'=1}^{n} (\mathcal{X}(\omega_{\upsilon i}) \oslash \mathcal{X}(\omega_{\upsilon i'})) \ominus 1\right)$$

$$= \frac{1}{n^2} \sum_{i=1}^{n}\sum_{i'=1}^{n} F(\mathcal{X}(\omega_{\upsilon i}) \oslash \mathcal{X}(\omega_{\upsilon i'})) - 1.$$

$\mathrm{I}_H(\mathcal{X}[\cdot])$ is a real-valued random variable associated with the probability space $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ (see Chapter 1, p. 45), and hence, defines a real-valued estimator of the *population hyperbolic index*, which is given by

$$\mathrm{I}_H(\mathcal{X}\,|\,P) = F\left(\frac{1}{N^2} \odot \sum_{j=1}^{N}\sum_{j'=1}^{N} (\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_{j'})) \ominus 1\right)$$

$$= \frac{1}{N^2} \sum_{j=1}^{N}\sum_{j'=1}^{N} F(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_{j'})) - 1.$$

To obtain from the sample index an unbiased estimator of the population index, the first one has to be revised. The correction which we will apply is based on the following result about the expected value of the sample index:

Theorem 2.3.1. *In random sampling without replacement of size n from the population Ω, if $f = n/N$ we have that*

$$E(\mathrm{I}_H(\mathcal{X}[\cdot])) = \frac{(n-1)N}{n(N-1)}\mathrm{I}_H(\mathcal{X}\,|\,P) + \frac{1-f}{n(N-1)}\sum_{j=1}^{N}\mathrm{I}_H(\{\mathcal{X}(\omega_j)\}),$$

where $\mathrm{I}_H(\{\mathcal{X}(\omega_j)\})$ represents the (intra)hyperbolic inequality index of the fuzzy random variable degenerate at the value $\mathcal{X}(\omega_j) \in \mathcal{F}_c((0,+\infty))$.

Proof.

Indeed,

$$E(\mathrm{I}_H(\mathcal{X}[\cdot])) = \sum_{\upsilon\in\Upsilon_n} p[\upsilon]\,\mathrm{I}_H(\mathcal{X}[\upsilon])$$

$$= \sum_{\upsilon\in\Upsilon_n} p[\upsilon]\frac{1}{n^2}\sum_{j=1}^{N}\sum_{j'=1}^{N} F(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_{j'})) \cdot a_j[\upsilon]a_{j'}[\upsilon] - 1,$$

where a_j is the Bernoulli variable introduced in Chapter 1, p. 48.

Consequently,

$$E\left(\mathrm{I}_H\left(\mathcal{X}[\,\cdot\,]\right)\right) = \frac{1}{n^2}\sum_{j=1}^{N}\sum_{j'=1}^{N} F\left(\mathcal{X}(\omega_j)\,\oslash\,\mathcal{X}(\omega_{j'})\right)\,E\left(a_j\,a_{j'}\right) - 1$$

$$= \frac{1}{n^2}\sum_{j=1}^{N} F\left(\mathcal{X}(\omega_j)\,\oslash\,\mathcal{X}(\omega_j)\right)\,\frac{n}{N}$$

$$+\frac{1}{n^2}\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N} F\left(\mathcal{X}(\omega_j)\,\oslash\,\mathcal{X}(\omega_{j'})\right)\,\frac{n(n-1)}{N(N-1)} - 1$$

$$= \frac{(n-1)N}{n(N-1)}\,\mathrm{I}_H(\mathcal{X}) + \frac{1-f}{n(N-1)}\Big[\sum_{j=1}^{N} F\left(\mathcal{X}(\omega_j)\,\oslash\,\mathcal{X}(\omega_j)\right)\,-1\Big]$$

$$= \frac{(n-1)N}{n(N-1)}\,\mathrm{I}_H(\mathcal{X}\,|\,P) + \frac{1-f}{n(N-1)}\,\mathrm{I}_H(\mathcal{X}(\omega_j)).$$

□

On the basis of the above result, we can conclude that

Theorem 2.3.2. *In random sampling without replacement of size n from the population Ω, if $f = n/N$ we have that the estimator $\widehat{\mathrm{I}_H}(\mathcal{X}[\,\cdot\,])$ such that*

$$\widehat{\mathrm{I}_H}(\mathcal{X}[v]) = \frac{n(N-1)}{N(n-1)}\mathrm{I}_H\left(\mathcal{X}[v]\right) - \frac{1-f}{n}\left\{\frac{1}{n-1}\sum_{i=1}^{n}\mathrm{I}_H\left(\{\mathcal{X}(\omega_{vi})\}\right)\right\},$$

is an unbiased estimator of $\mathrm{I}_H(\mathcal{X}\,|\,P)$.

Proof.

Indeed, since the sample expected value of the hyperbolic *intra*values inequality indices is an unbiased estimator of the corresponding population expected value, that is,

$$E\left(\frac{1}{n}\sum_{i=1}^{n}\mathrm{I}_H\left(\{\mathcal{X}(\omega_{\cdot i})\}\right)\right) = \frac{1}{N}\sum_{j=1}^{N}\mathrm{I}_H\left(\{\mathcal{X}(\omega_j)\}\right),$$

then, we have that

$$E\left(\frac{1-f}{n}\left(\frac{1}{n}\sum_{i=1}^{n}\mathrm{I}_{H}\left(\{\mathcal{X}(\omega_{\cdot i})\}\right)\right)\right)=\frac{1-f}{N(n-1)}\sum_{j=1}^{N}\mathrm{I}_{H}\left(\{\mathcal{X}(\omega_{j})\}\right),$$

and in virtue of Theorem 2.3.1 we have that

$$E\left(\widehat{\mathrm{I}_{H}}(\mathcal{X}[\,\cdot\,])\right)=\mathrm{I}_{H}(\mathcal{X}\,|\,P).$$

□

To establish the accuracy of the preceding estimator of $\mathrm{I}_{H}(\mathcal{X}\,|\,P)$, we determine the associated "mean squared error", which in this case coincides with the variance $\mathrm{Var}\left(\widehat{\mathrm{I}_{H}}(\mathcal{X}[\,\cdot\,])\right)$, in the following result:

Theorem 2.3.3. *In random sampling without replacement of size n from* Ω, *if* $f=n/N$ *we have that*

$$\mathrm{Var}\left(\widehat{\mathrm{I}_{H}}(\mathcal{X}[\,\cdot\,])\right)=\frac{(1-f)}{n(n-1)N^{2}(N-1)(N-2)(N-3)}$$

$$\cdot\Big\{[N(6-4n)+6(n-1)]N^{3}(N-1)\left(\mathrm{I}_{H}(\mathcal{X}\,|\,P)\right)^{2}$$

$$+[4N^{2}(3-2n)+13N(n-1)+3(n-3)]N^{2}(N-1)\,\mathrm{I}_{H}(\mathcal{X}\,|\,P)$$

$$+[N^{2}(7-3n)+N(5n-7)-4(n-2)]\left(\sum_{j=1}^{N}\mathrm{I}_{H}(\{\mathcal{X}(\omega_{j})\})\right)^{2}$$

$$+[N^{2}(n-5)+N(5n+1)-10(n-2)]\sum_{j=1}^{N}\left(\mathrm{I}_{H}(\{\mathcal{X}(\omega_{j})\})\right)^{2}$$

$$+2[N^{3}(3-n)+N^{2}(3n-8)+N(n+9)-10(n-2)]\sum_{j=1}^{N}\mathrm{I}_{H}(\{\mathcal{X}(\omega_{j})\})$$

$$+4(n-2)N^{2}(N-1)\mathrm{I}_{H}(\mathcal{X}\,|\,P)\sum_{j=1}^{N}\mathrm{I}_{H}(\{\mathcal{X}(\omega_{j})\})$$

$$+(n-2)(N-1)(N-2)\sum_{j=1}^{N}\left(\sum_{l=1}^{N}\left[F\left(\mathcal{X}(\omega_{j})\oslash\mathcal{X}(\omega_{l})\right)+F\left(\mathcal{X}(\omega_{l})\oslash\mathcal{X}(\omega_{j})\right)\right]\right)^{2}$$

$$+(N-n+1)(N-1)(N-3)\sum_{j=1}^{N}\sum_{l=1}^{N} F\left(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_l)\right)$$

$$\cdot \left[F\left(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_l)\right) + F\left(\mathcal{X}(\omega_l) \oslash \mathcal{X}(\omega_j)\right)\right]$$

$$-[N(3n-7)-3(n-3)](N-1)\sum_{j=1}^{N}\sum_{l=1}^{N} \mathrm{I}_H(\{\mathcal{X}(\omega_l)\}) F\left(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_l)\right)$$

$$+[N^5(6-4n)+8N^4n+N^3(n-9)-6N^2(n-2)+N(5n+1)-10(n-2)]\Big\}.$$

Proof.

To simplify the formulae, we will use the following notation for all $j,\, l \in \{1, \ldots, N\}$:

$$F\left(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_l)\right) = \mathrm{V}_{jl}.$$

Consequently,

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}(\mathcal{X}[\,\cdot\,])\right) = \frac{1}{n^2(n-1)^2}\Bigg\{ \left(\frac{N-1}{N}\right)^2$$

$$\cdot \sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N}\sum_{l'=1}^{N} \mathrm{V}_{jj'}\mathrm{V}_{ll'}\ \mathrm{Cov}(a_j\, a_{j'},\, a_l\, a_{l'})$$

$$+\left(\frac{N-n}{N}\right)^2 \sum_{j=1}^{N}\sum_{l=1}^{N} \mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(a_j,\, a_l)$$

$$-\,2\frac{(N-1)(N-n)}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N} \mathrm{V}_{jj'}\mathrm{V}_{ll}\,\mathrm{Cov}(a_j\, a_{j'},\, a_l)\Bigg\}$$

$$= \frac{N-1}{nN(n-1)}\Bigg\{ \sum_{j=1}^{N} \mathrm{V}_{jj}^2\,\mathrm{Var}(a_j^2) \,+\, 2\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{V}_{lj}\mathrm{V}_{jj}\ \mathrm{Cov}(a_l\, a_j,\, a_j^2)$$

$$+\,2\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{V}_{jl}\mathrm{V}_{jj}\ \mathrm{Cov}(a_j\, a_l,\, a_j^2) \,+\, \sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(a_j^2,\, a_l^2)$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{V}_{jl}\mathrm{V}_{lj}\ \mathrm{Var}(a_j\, a_l) + \sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N} \mathrm{V}_{jl}^2\ \mathrm{Var}(a_j\, a_l)$$

$$+2\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll}\ \mathrm{Cov}(a_j\,a_{j'},\,a_l^2)$$

$$+2\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{lj}\ \mathrm{Cov}(a_j\,a_{j'},\,a_l\,a_j)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{lj'}\ \mathrm{Cov}(a_j\,a_{j'},\,a_l\,a_{j'})$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{jl}\ \mathrm{Cov}(a_j\,a_{j'},\,a_j\,a_l)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\sum_{\substack{l'=1\\ l'\neq l,j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll'}\ \mathrm{Cov}(a_j\,a_{j'},\,a_l\,a_{l'})\Bigg\}$$

$$+\left(\frac{N-n}{nN(n-1)}\right)^2\left\{\sum_{j=1}^{N}\mathrm{V}_{jj}^2\ \mathrm{Var}(a_j)\ +\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(a_j,\,a_l)\right\}$$

$$-2\frac{(N-1)(N-n)}{[nN(n-1)]^2}\Bigg\{\sum_{j=1}^{N}\mathrm{V}_{jj}^2\ \mathrm{Cov}(a_j^2,\,a_j)\ +\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{V}_{jl}\mathrm{V}_{jj}\ \mathrm{Cov}(a_j\,a_l,\,a_j)$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{V}_{jl}\mathrm{V}_{ll}\ \mathrm{Cov}(a_j\,a_l,\,a_l)\ +\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(a_j^2,\,a_l)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll}\ \mathrm{Cov}(a_j\,a_{j'},\,a_l)\Bigg\},$$

and, in virtue of the moments of the variables a_j (see Cap. 1, p. 48), we have that

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}(\mathcal{X}[\,\cdot\,])\right)=\frac{1-f}{nN^3}\Bigg\{\sum_{j=1}^{N}\mathrm{V}_{jj}^2\ +\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\Big[2\,(\mathrm{V}_{jl}\mathrm{V}_{ll}+\mathrm{V}_{jl}\mathrm{V}_{jj})$$

$$-\frac{1}{N-1}\mathrm{V}_{jj}\mathrm{V}_{ll}+\frac{N+n-1}{n-1}\left(\mathrm{V}_{jl}\mathrm{V}_{lj}+\mathrm{V}_{jl}^2\right)\Bigg]+\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\left[\frac{-4}{(N-2)}\mathrm{V}_{jj'}\mathrm{V}_{ll}\right.$$

$$\left.+\frac{[N(n-2)-2(n-1)]}{(n-1)(N-2)}\left(\mathrm{V}_{jj'}\mathrm{V}_{lj'}+\mathrm{V}_{jj'}\mathrm{V}_{jl}+2\mathrm{V}_{jj'}\mathrm{V}_{lj}\right)\right]$$

$$\left.+\frac{[N(6-4n)+6(n-1)]}{(n-1)(N-2)(N-3)}\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N}\sum_{\substack{l=1\\ l\neq j,j'}}^{N}\sum_{\substack{l'=1\\ l'\neq l,j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll'}\right\}$$

$$=\frac{(1-f)}{n(n-1)N^2}\left\{\frac{[N(6-4n)+6(n-1)]}{N(N-2)(N-3)}\left(\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jl}\right)^2\right.$$

$$+\frac{n-2}{N-3}\left[\sum_{j=1}^{N}\left(\sum_{l=1}^{N}(\mathrm{V}_{jl}+\mathrm{V}_{lj})\right)^2\right]$$

$$+\frac{N-n+1}{N-2}\sum_{j=1}^{N}\sum_{l=1}^{N}\left(\mathrm{V}_{jl}^2+\mathrm{V}_{jl}\mathrm{V}_{lj}\right)+\frac{4(n-2)}{(N-2)(N-3)}\left(\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jl}\right)\left(\sum_{j=1}^{N}\mathrm{V}_{jj}\right)$$

$$-\frac{N(3n-7)-3(n-3)}{(N-2)(N-3)}\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jl}\mathrm{V}_{ll}$$

$$+\frac{N^2(7-3n)+N(5n-7)-4(n-2)}{(N-1)(N-2)(N-3)}\left(\sum_{j=1}^{N}\mathrm{V}_{jj}\right)^2$$

$$\left.+\frac{N^2(n-5)+N(5n+1)-10(n-2)}{(N-1)(N-2)(N-3)}\sum_{j=1}^{N}\mathrm{V}_{jj}^2\right\}.$$

Since $\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jl} = N^2\left(\mathrm{I}_H(\mathcal{X}\,|\,P)+1\right)$ and $\mathrm{V}_{jj} = \mathrm{I}_H(\{\mathcal{X}(\omega_j)\})$ for all $j\in\{1,\ldots,N\}$, we can conclude the result in the present theorem. □

If, instead of adopting a choice at random and without replacement, we consider a choice at random and with replacement of n units from $\Omega = \{\omega_1,\ldots,\omega_N\}$, and υ represents a generic random sample with replacement of size n, and $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ are the units in υ, then, the *sample hyperbolic index* of $\mathcal{X}$ in υ is given now by

$$\mathrm{I}_H(\mathcal{X}[\upsilon]) = \frac{1}{N^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} F(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_{j'})) \cdot t_j[\upsilon] t_{j'}[\upsilon] - 1,$$

where t_j are the real-valued random variables introduced in Chapter 1, p. 51.

$\mathrm{I}_H(\mathcal{X}[\cdot])$ is a real-valued random variable associated with the probability space$(\Upsilon_n^w, \mathcal{P}(\Upsilon_n^w), p^w)$ (see Chapter 1, p. 51), and, hence, it defines an estimator of the population hyperbolic index.

As for the simple random sampling, to obtain an unbiased estimator of $\mathrm{I}_H(\mathcal{X})$ from the sampling one, we first examine the expected value of the sample hyperbolic index in the sampling with replacement.

Theorem 2.3.4. *In random sampling with replacement of size n from the population Ω, we have that*

$$E(\mathrm{I}_H(\mathcal{X}[\cdot])) = \frac{(n-1)}{n} \mathrm{I}_H(\mathcal{X} \mid P) + \frac{1}{nN} \sum_{j=1}^{N} \mathrm{I}_H(\{\mathcal{X}(\omega_j)\}).$$

Proof.

Indeed,

$$E(\mathrm{I}_H(\mathcal{X}[\cdot])) = \sum_{\upsilon \in \Upsilon_n^w} p^w[\upsilon] \, \mathrm{I}_H(\mathcal{X}[\upsilon])$$

$$= \sum_{\upsilon \in \Upsilon_n^w} p^w[\upsilon] \frac{1}{n^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} F(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_{j'})) \cdot t_j[\upsilon] t_{j'}[\upsilon] - 1.$$

Consequently,

$$E(\mathrm{I}_H(\mathcal{X}[\cdot])) = \frac{1}{n^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} F(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_{j'})) \, E(t_j \, t_{j'}) - 1$$

$$= \frac{1}{n^2} \sum_{j=1}^{N} F(\mathcal{X}(\omega_j) \oslash \mathcal{X}(\omega_j)) \, \frac{n(N+n-1)}{N^2}$$

$$+\frac{1}{n^2}\sum_{j=1}^{N}\sum_{\substack{j'=1\\ j'\neq j}}^{N} F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_{j'})\right)\frac{n(n-1)}{N^2}-1$$

$$=\frac{(n-1)}{n}\mathrm{I}_H(\mathcal{X}\,|\,P)+\frac{1}{nN}\sum_{j=1}^{N}\mathrm{I}_H\left(\{\mathcal{X}(\omega_j)\}\right).$$

□

On the basis of the above result, we can conclude that

Theorem 2.3.5. *In random sampling with replacement of size n from the population* Ω, *we have that the estimator* $\widehat{\mathrm{I}_H}^{w}(\mathcal{X}[\cdot])$ *which for each sample* υ *takes on the value*

$$\widehat{\mathrm{I}_H}^{w}(\mathcal{X}[\upsilon])=\frac{n}{(n-1)}\mathrm{I}_H\left(\mathcal{X}[\upsilon]\right)-\frac{1}{n}\left\{\frac{1}{n-1}\sum_{i=1}^{n}\mathrm{I}_H\left(\{\mathcal{X}(\omega_{\upsilon i})\}\right)\right\},$$

is an unbiased estimator of $\mathrm{I}_H(\mathcal{X}\,|\,P)$.

Proof.

Indeed, as in the simple random sampling, the sample expected value of the hyperbolic *intra*values inequality indices is an unbiased estimator of the corresponding population one, then, we have that

$$E\left(\frac{1}{n}\left\{\frac{1}{n-1}\sum_{i=1}^{n}\mathrm{I}_H\left(\{\mathcal{X}(\omega_{\cdot i})\}\right)\right\}\right)=\frac{1}{N(n-1)}\sum_{j=1}^{N}\mathrm{I}_H\left(\{\mathcal{X}(\omega_j)\}\right),$$

and in virtue of Theorem 2.3.4 it satisfies that

$$E\left(\widehat{\mathrm{I}_H}^{w}(\mathcal{X}[\cdot])\right)=\mathrm{I}_H(\mathcal{X}\,|\,P).$$

□

The mean squared error associated with the estimator in Theorem 2.3.5, coincides with the variance $\mathrm{Var}\left(\widehat{\mathrm{I}_H}^{w}(\mathcal{X}[\cdot])\right)$, and will be given by

Theorem 2.3.6. *In random sampling with replacement of size n from the population* Ω, *we have that*

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}^{w}(\mathcal{X}[\,\cdot\,])\right) = \frac{1}{n(n-1)N^3}$$

$$\cdot\Big\{(6-4n)N^3\left(\mathrm{I}_H(\mathcal{X}\,|\,P)\right)^2 + 2(6-4n)N^3\mathrm{I}_H(\mathcal{X}\,|\,P)$$

$$-[N-4(n-1)]\left(\sum_{j=1}^{N}\mathrm{I}_H(\{\mathcal{X}(\omega_j)\})\right)^2 + [N-4(n-1)]\sum_{j=1}^{N}\left(\mathrm{I}_H(\{\mathcal{X}(\omega_j)\})\right)^2$$

$$-2[N-4(n-1)](N-1)\sum_{j=1}^{N}\mathrm{I}_H(\{\mathcal{X}(\omega_j)\})$$

$$+(n-2)\sum_{j=1}^{N}\left(\sum_{l=1}^{N}\left[F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_l)\right)+F\left(\mathcal{X}(\omega_l)\oslash\mathcal{X}(\omega_j)\right)\right]\right)^2$$

$$+N\sum_{j=1}^{N}\sum_{l=1}^{N}F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_l)\right)\left[F\left(\mathcal{X}(\omega_j)\oslash\mathcal{X}(\omega_l)\right)+F\left(\mathcal{X}(\omega_l)\oslash\mathcal{X}(\omega_j)\right)\right]$$

$$+[N^2(5-4n)+N(4n-3)-4(n-1)]N\Big\}.$$

Proof.

To simplify, we will use the same notation V_{jl} as that in Theorem 2.3.3.

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}^{w}(\mathcal{X}[\,\cdot\,])\right) = \frac{1}{n^2(n-1)^2}\Big\{\sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N}\sum_{l'=1}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll'}\ \mathrm{Cov}(t_j\,t_{j'},\,t_l\,t_{l'})$$

$$+\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j,\,t_l) - 2\sum_{j=1}^{N}\sum_{j'=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll}\,\mathrm{Cov}(t_j\,t_{j'},\,t_l)\Big\}$$

$$= \frac{1}{n^2(n-1)^2}\Big\{\sum_{j=1}^{N}\mathrm{V}_{jj}^2\,\mathrm{Var}(t_j^2) + 2\sum_{j=1}^{N}\sum_{\substack{l=1\\ l\neq j}}^{N}\mathrm{V}_{lj}\mathrm{V}_{jj}\ \mathrm{Cov}(t_l\,t_j,\,t_j^2)$$

$$+\,2\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jl}\mathrm{V}_{jj}\ \mathrm{Cov}(t_j\,t_l,\ t_j^2)\ +\ \sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j^2,\ t_l^2)$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jl}\mathrm{V}_{lj}\ \mathrm{Var}(t_j\,t_l)+\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jl}^2\ \mathrm{Var}(t_j\,t_l)$$

$$+\,2\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j\,t_{j'},\ t_l^2)+2\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{lj}\ \mathrm{Cov}(t_j\,t_{j'},\ t_l\,t_j)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{lj'}\ \mathrm{Cov}(t_j\,t_{j'},\ t_l\,t_{j'})+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{jl}\ \mathrm{Cov}(t_j\,t_{j'},\ t_j\,t_l)$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\sum_{\substack{l'=1\\l'\neq l,j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll'}\ \mathrm{Cov}(t_j\,t_{j'},\ t_l\,t_{l'})$$

$$+\sum_{j=1}^{N}\mathrm{V}_{jj}^2\ \mathrm{Var}(t_j)\ +\ \sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j,\ t_l)$$

$$-\,2\sum_{j=1}^{N}\mathrm{V}_{jj}^2\ \mathrm{Cov}(t_j^2,\ t_j)\ -\ 2\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jl}\mathrm{V}_{jj}\ \mathrm{Cov}(t_j\,t_l,\ t_j)$$

$$-2\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jl}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j\,t_l,\ t_l)\ -\ 2\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\mathrm{V}_{jj}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j^2,\ t_l)$$

$$-2\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll}\ \mathrm{Cov}(t_j\,t_{j'},\ t_l)\Bigg\}.$$

In virtue of the expressions of the moments of the variables t_j (see Ch. 1, p. 51), we have that

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}^{\,w}(\mathcal{X}[\,\cdot\,])\right)=\frac{1-f}{n^2(n-1)^2}\Bigg\{\frac{n(n-1)(N-1)[2N+(4n-6)]}{N^4}\sum_{j=1}^{N}\mathrm{V}_{jj}^2$$

$$+\sum_{j=1}^{N}\sum_{\substack{l=1\\l\neq j}}^{N}\left[\frac{2n(n-1)}{N^4}[2N(n-2)+(6-4n)]\left(\mathrm{V}_{jl}\mathrm{V}_{ll}+\mathrm{V}_{jl}\mathrm{V}_{jj}\right)\right.$$

$$+\frac{n}{N^4}[-N^2(n+1)+4N(n-1)^2+(6-4n)(n-1)]\mathrm{V}_{jj}\mathrm{V}_{ll}$$

$$\left.+\frac{n(n-1)}{N^4}[N(N+2n-4)+(6-4n)]\left(\mathrm{V}_{jl}\mathrm{V}_{lj}+\mathrm{V}_{jl}^2\right)\right]$$

$$+\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\left[\frac{2n(n-1)(6-4n)}{(N^4)}\mathrm{V}_{jj'}\mathrm{V}_{ll}\right.$$

$$\left.+\frac{n(n-1)}{N^4}[N(n-2)+(6-4n)]\left(\mathrm{V}_{jj'}\mathrm{V}_{lj'}+\mathrm{V}_{jj'}\mathrm{V}_{jl}+2\mathrm{V}_{jj'}\mathrm{V}_{lj}\right)\right]$$

$$\left.+\frac{n(n-1)(6-4n)}{N^4}\sum_{j=1}^{N}\sum_{\substack{j'=1\\j'\neq j}}^{N}\sum_{\substack{l=1\\l\neq j,j'}}^{N}\sum_{\substack{l'=1\\l'\neq l,j,j'}}^{N}\mathrm{V}_{jj'}\mathrm{V}_{ll'}\right\}$$

$$=\frac{1}{n(n-1)N^2}\left\{\frac{6-4n}{N^2}\left(\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jl}\right)^2\right.$$

$$+\frac{n-2}{N}\sum_{j=1}^{N}\left(\sum_{l=1}^{N}(\mathrm{V}_{jl}+\mathrm{V}_{lj})\right)^2+\sum_{j=1}^{N}\sum_{l=1}^{N}\left(\mathrm{V}_{jl}^2+\mathrm{V}_{jl}\mathrm{V}_{lj}\right)$$

$$\left.-\frac{N(n-1)-4(n-1)^2}{N(n-1)}\left(\left(\sum_{j=1}^{N}\mathrm{V}_{jj}\right)^2-\sum_{j=1}^{N}\mathrm{V}_{jj}^2\right)\right\},$$

and by reasoning as in Theorem 2.3.3 we obtain the result in the present one. □

As in Section 1.3, the results we have just stated can be used to compare the accuracy in the estimation of the population hyperbolic index associated with different variables, to derive confidence intervals and testing hypotheses, and determining adequate sample sizes, although this would lead to conservative procedures and we would need to estimate the population parameters in the variance of the estimators.

2.4 Asymptotic behavior of the sample inequality indices in finite populations. Applications to the development of approximate inferential procedures for the population inequality indices

In the preceding section we have shown that the population hyperbolic index admits an unbiased estimator in the sampling of finite populations. This assertion cannot be extended immediately to other indices, either because their unbiased estimation is not possible, or because this estimation is very complex. Similar conclusions can be obtained in relation to determining the accuracy of the estimates and with respect to other exact inferential procedures.

However, the study of the asymptotic behavior of the sample inequality indices allows us to get, under certain conditions, the asymptotically unbiased estimation, to determine the asymptotic accuracy of the estimates, and the development of approximations of other inferential procedures for most of the population indices.

Regarding implications from this study and the approximate inferential procedures derived from it, one should remark that they could be applied when we consider large samples either chosen at random, and with replacement of any population, or chosen at random and without replacement from a large population whose size is substantially bigger than the sample one.

Let $\Omega = \{\omega_1, \ldots, \omega_N\}$ be a finite population of size N, and let $\mathcal{X}$ be a positive fuzzy random variable. Assume that the variable $\mathcal{X}$ takes on r different values on Ω, $\tilde{x}_1^*, \ldots, \tilde{x}_r^* \in \mathcal{F}_c((0, +\infty))$, and for any $l \in \{1, \ldots, r\}$ it denotes by p_l the probability $P(\{\omega \in \Omega \mid \mathcal{X}(\omega) = \tilde{x}_l^*\})$, where P denotes the probability measure associated with the uniform distribution on Ω. If

$f : (0, +\infty) \to \mathbb{R}$ is a strictly convex and monotonic function, $f \in C^1$ and satisfying that $f(u) + f(1/u) \geq 0$ for all $u \in (0, +\infty)$ and $f(1) = 0$, the population f-inequality index of $\mathcal{X}$ is given by

$$\begin{aligned} \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) &= F\left(\bigoplus_{l'=1}^{r} p_{l'} \odot f\left(\tilde{x}^*_{l'} \oslash \left(\bigoplus_{l=1}^{r} p_l \odot \tilde{x}^*_l\right)\right)\right) \\ &= \sum_{l'=1}^{r} p_{l'} F\left(f\left(\tilde{x}^*_{l'} \oslash \left(\bigoplus_{l=1}^{r} p_l \odot \tilde{x}^*_l\right)\right)\right). \end{aligned}$$

If a sample of size n is chosen at random from Ω and f_{nl} is the relative frequency of the value $\tilde{x}^*_l$ of $\mathcal{X}$ in the sample, the *sample f-inequality index* of $\mathcal{X}$ corresponds to the value

$$\begin{aligned} \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) &= F\left(\bigoplus_{l'=1}^{r} f_{nl'} \odot f\left(\tilde{x}^*_{l'} \oslash \left(\bigoplus_{l=1}^{r} f_{nl} \odot \tilde{x}^*_l\right)\right)\right) \\ &= \sum_{l'=1}^{r} f_{nl'} F\left(f\left(\tilde{x}^*_{l'} \oslash \left(\bigoplus_{l=1}^{r} f_{nl} \odot \tilde{x}^*_l\right)\right)\right). \end{aligned}$$

The following result establishes the asymptotic distribution of the sample inequality index.

Theorem 2.4.1. *For each $n \in \mathbb{N}$, consider n independent fuzzy random variables being identically distributed as $\mathcal{X}$ (that is, a simple random sampling from $\mathcal{X}$) defined over the finite population $\Omega = \{\omega_1, \ldots, \omega_N\}$ such that $P(\{\omega \in \Omega \mid \mathcal{X}(\omega) = \tilde{x}^*_l\}) = p_l$ with $p_l \in (0,1)$, $l = 1, \ldots, r$, $\sum_{l=1}^{r} p_l = 1$. Let $f : (0, +\infty) \to \mathbb{R}$ be a strictly convex and monotonic function, and satisfying that $f(u) + f(1/u) \geq 0$ for all $u \in (0, +\infty)$ and $f(1) = 0$, and f admits finite third-order derivates. Then, we have that*

i) If $\mathbf{f}_n = (f_{n1}, \ldots, f_{n(r-1)}) \in [0,1]^{r-1}$, with $f_{nl} =$ the relative frequency of $\tilde{x}^*_l$ in the performance of the simple random sample of size n ($l =$

$1,\ldots,r-1,\ f_{nr}=1-\sum_{l=1}^{r-1} f_{nl}),$ and $\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)$ is the associated sample inequality index, then $\{\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)\}_n$ is a sequence of estimators of $\mathrm{I}_f(\mathcal{X})=\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})$, which is strongly consistent, that is, as $n\to\infty$ we have that

$$\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)\stackrel{c.s.}{\to}\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}),$$

whatever $\mathbf{p}=(p_1,\ldots,p_{r-1})$ $(p_1,\ldots,p_{r-1}\in(0,1)$ and $\sum_{l=1}^{r-1} p_l<1)$ may be.

ii) *$\{\sqrt{n}\,(\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)-\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}))\}_n$ is a sequence of real-valued random variables which converges in law as $n\to\infty$ to a one-dimensional normal distribution $N(0,\sigma^2(\mathbf{p}))$, with*

$$\sigma^2(\mathbf{p})=\sum_{l'=1}^{r} p_{l'}(V_{l'}^*)^2-\left(\sum_{l'=1}^{r} p_{l'}\,V_{l'}^*\right)^2,$$

where

$$V_{l'}^*=F\left(f\left(\tilde{x}_{l'}^*\oslash E(\mathcal{X}\,|\,\mathbf{p})\right)\right)-\sum_{l=1}^{r} p_l\left[\frac{1}{2}\int_{[0,1]}\left\{\inf(\tilde{x}_{l'}^*)_\alpha\right.\right.$$
$$\cdot f'\left(\frac{\sup(\tilde{x}_l^*)_\alpha}{E\left(\inf\mathcal{X}_\alpha\,|\,\mathbf{p}\right)}\right)\cdot\left(\frac{\sup(\tilde{x}_l^*)_\alpha}{\left[E\left(\inf\mathcal{X}_\alpha\,|\,\mathbf{p}\right)\right]^2}\right)$$
$$\left.\left.+\sup(\tilde{x}_{l'}^*)_\alpha\cdot f'\left(\frac{\inf(\tilde{x}_l^*)_\alpha}{E\left(\sup\mathcal{X}_\alpha\,|\,\mathbf{p}\right)}\right)\cdot\left(\frac{\inf(\tilde{x}_l^*)_\alpha}{\left[E\left(\sup\mathcal{X}_\alpha\,|\,\mathbf{p}\right)\right]^2}\right)\right\}d\,\alpha\right],$$

whenever $\sigma^2(\mathbf{p})>0$.

iii) *If $\sigma^2(\mathbf{p})=0$, and there is a pair (i,j) with $i,j\in\{1,\ldots,r-1\}$ such that*

$$h_{ij}=\frac{\partial^2\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i\partial p_j}=\frac{\partial}{\partial p_i}[\,V_j^*-V_r^*]>0,$$

then $\{2n\,(\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)-\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}))\}_n$ is a sequence of real-valued random variables which converges in law as $n\to\infty$ to a linear combination of, at most, $r-1$ chi-squared χ_1^2 independent random variables.

Proof.

If the introduction of the proof in Theorem 1.4.1 is revised, we conclude that the conditions satisfied by $\mathbf{f}_n$, $\mathbf{p}$, the parameter space and the set of values of the variable $\mathcal{X}$, guarantee that

i) $\mathbf{f}_n \overset{c.s.}{\to} \mathbf{p}$ and since $\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})$ is continuous in a neighborhood of $\mathbf{p}$, we can conclude that

$$\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) \overset{c.s.}{\to} \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}),$$

as $n \to \infty$, that is, $\{\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)\}_n$ is a strongly consistent sequence of esimators of $\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})$.

ii) Because of the assumed conditions, if n is large enough one can expand $\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)$ in a neighborhood of $\mathbf{p}$. In this way, the first Taylor expansion is given by

$$\begin{aligned}\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) &= \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}) + \nabla \mathrm{I}_f(\mathbf{p})(\mathbf{f}_n-\mathbf{p})^t \\ &+\frac{1}{2}(\mathbf{f}_n-\mathbf{p})H(\mathrm{I}_f(\mathbf{p}_n^*))\,(\mathbf{f}_n-\mathbf{p})^t,\end{aligned}$$

with $\nabla \mathrm{I}_f(\mathbf{p})$ = gradient vector of $\mathrm{I}_f(\mathcal{X}\,|\,\cdot)$ in $\mathbf{p}$, and $H(\mathrm{I}_f(\mathbf{p}))$ = $(r-1)\times(r-1)$ hessian matrix such that

$$H(\mathrm{I}_f(\mathbf{p}_n^*)) = \left[\left(\frac{\partial^2 \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})}{\partial p_i \partial p_j}\right)_{\mathbf{p}=\mathbf{p}_n^*}\right],$$

and with $\mathbf{p}_n^* \in \mathbb{P}$ such that $\|\mathbf{p}_n^*-\mathbf{p}\| \le \|\mathbf{f}_n-\mathbf{p}\|$.

Consequently,

$$\begin{aligned}&\sqrt{n}\,[\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})] \\ &= \nabla \mathrm{I}_f(\mathbf{p})\left(\sqrt{n}(\mathbf{f}_n-\mathbf{p})\right)^t + \frac{1}{2}(\mathbf{f}_n-\mathbf{p})H(\mathrm{I}_f(\mathbf{p}_n^*))\left(\sqrt{n}(\mathbf{f}_n-\mathbf{p})\right)^t.\end{aligned}$$

The gradient vector is given by the $1\times(r-1)$ matrix

$$\nabla \mathrm{I}_f(\mathbf{p}) = \left(\frac{\partial}{\partial p_1}\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}) \;\cdots\; \frac{\partial}{\partial p_{r-1}}\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})\right),$$

and for $i \in \{1, \ldots, r-1\}$ we have that

$$\frac{\partial}{\partial p_i} \mathrm{I}_f(\mathcal{X} \mid \mathbf{p}) = \frac{\partial}{\partial p_i} \left[\sum_{l=1}^{r-1} p_l \left\{ F\Big(f\left(\tilde{x}_l^* \oslash E(\mathcal{X} \mid \mathbf{p})\right)\Big) \right. \right.$$

$$\left. \left. -F\Big(f\left(\tilde{x}_r^* \oslash E(\mathcal{X} \mid \mathbf{p})\right)\Big) \right\} + F\Big(f\left(\tilde{x}_r^* \oslash E(\mathcal{X} \mid \mathbf{p})\right)\Big) \right]$$

$$= F\Big(f\left(\tilde{x}_i^* \oslash E(\mathcal{X} \mid \mathbf{p})\right)\Big) - F\Big(f\left(\tilde{x}_r^* \oslash E(\mathcal{X} \mid \mathbf{p})\right)\Big)$$

$$+ \sum_{l=1}^{r-1} p_l \left\{ \frac{\partial}{\partial p_i} F\Big(f\left(\tilde{x}_l^* \oslash E(\mathcal{X} \mid \mathbf{p})\right)\Big) \right.$$

$$\left. - \frac{\partial}{\partial p_i} F\Big(f\left(\tilde{x}_r^* \oslash E(\mathcal{X} \mid \mathbf{p})\right)\Big) \right\} + \frac{\partial}{\partial p_i} F\Big(f\left(\tilde{x}_r^* \oslash E(\mathcal{X} \mid \mathbf{p})\right)\Big).$$

If $l \in \{1, \ldots, r\}$, then

$$\frac{\partial}{\partial p_i} F\Big(f\left(\tilde{x}_l^* \oslash E(\mathcal{X} \mid \mathbf{p})\right)\Big)$$

$$= \frac{1}{2} \int_{[0,1]} \left\{ \frac{\partial}{\partial p_i} f\left(\sup(\tilde{x}_l^*)_\alpha \Big/ \sum_{l'=1}^{r-1} p_{l'} \left[\inf(\tilde{x}_{l'}^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right] \right) \right.$$

$$\left. + \frac{\partial}{\partial p_i} f\left(\inf(\tilde{x}_l^*)_\alpha \Big/ \sum_{l'=1}^{r-1} p_{l'} \left[\sup(\tilde{x}_{l'}^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right] \right) \right\} d\,\alpha$$

$$= \frac{1}{2} \int_{[0,1]} \left\{ \left[f'\left(\sup(\tilde{x}_l^*)_\alpha \Big/ \sum_{l'=1}^{r-1} p_{l'} \left[\inf(\tilde{x}_{l'}^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right] \right) \right. \right.$$

$$\cdot \left(- \sup(\tilde{x}_l^*)_\alpha \Big/ \left[\sum_{l'=1}^{r-1} p_{l'} \left(\inf(\tilde{x}_{l'}^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right) \right]^2 \right)$$

$$\left. \cdot \left(\inf(\tilde{x}_i^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right) \right] + \left[f'\left(\inf(\tilde{x}_l^*)_\alpha \Big/ \sum_{l'=1}^{r-1} p_{l'} \left[\sup(\tilde{x}_{l'}^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right] \right) \right.$$

$$\cdot\left(-\inf(\tilde{x}_l^*)_\alpha \Bigg/ \left[\sum_{l'=1}^{r-1} p_{l'}\left(\sup(\tilde{x}_{l'}^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right)\right]^2\right)$$

$$\cdot\left(\sup(\tilde{x}_i^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right)\Bigg]\Bigg\}d\,\alpha.$$

Consequently, if $i \in \{1, \ldots, r-1\}$, we have that

$$\frac{\partial}{\partial p_i}\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}) = F\Big(f\left(\tilde{x}_i^* \oslash E(\mathcal{X}\,|\,\mathbf{p})\right)\Big) - F\Big(f\left(\tilde{x}_r^* \oslash E(\mathcal{X}\,|\,\mathbf{p})\right)\Big)$$

$$+\frac{1}{2}\sum_{l=1}^{r} p_l \cdot \int_{[0,1]}\Bigg\{f'\left(\sup(\tilde{x}_l^*)_\alpha \Bigg/ \sum_{l'=1}^{r-1} p_{l'}\left[\inf(\tilde{x}_{l'}^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right]\right)$$

$$\cdot\left(-\sup(\tilde{x}_l^*)_\alpha \Bigg/ \left[\sum_{l'=1}^{r-1} p_{l'}\left(\inf(\tilde{x}_{l'}^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right)\right]^2\right)\left(\inf(\tilde{x}_i^*)_\alpha - \inf(\tilde{x}_r^*)_\alpha\right)$$

$$+f'\left(\inf(\tilde{x}_l^*)_\alpha \Bigg/ \sum_{l'=1}^{r-1} p_{l'}\left[\sup(\tilde{x}_{l'}^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right]\right)$$

$$\cdot\left(-\inf(\tilde{x}_l^*)_\alpha \Bigg/ \left[\sum_{l'=1}^{r-1} p_{l'}\left(\sup(\tilde{x}_{l'}^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right)\right]^2\right)$$

$$\cdot\left(\sup(\tilde{x}_i^*)_\alpha - \sup(\tilde{x}_r^*)_\alpha\right)\Bigg\}d\,\alpha = \mathrm{V}_i^* - \mathrm{V}_r^*,$$

with

$$\mathrm{V}_{l'}^* = F\left(f\left(\tilde{x}_{l'}^* \oslash E(\mathcal{X}\,|\,\mathbf{p})\right)\right)$$

$$-\sum_{l=1}^{r} p_l\Bigg[\frac{1}{2}\int_{[0,1]}\Bigg\{\inf(\tilde{x}_{l'}^*)_\alpha f'\left(\frac{\sup(\tilde{x}_l^*)_\alpha}{E\left(\inf\mathcal{X}_\alpha\,|\,\mathbf{p}\right)}\right)\cdot\left(\frac{\sup(\tilde{x}_l^*)_\alpha}{\left[E\left(\inf\mathcal{X}_\alpha\,|\,\mathbf{p}\right)\right]^2}\right)$$

$$+\sup(\tilde{x}_{l'}^*)_\alpha \cdot f'\left(\frac{\inf(\tilde{x}_l^*)_\alpha}{E\left(\sup\mathcal{X}_\alpha\,|\,\mathbf{p}\right)}\right)\cdot\left(\frac{\inf(\tilde{x}_l^*)_\alpha}{\left[E\left(\sup\mathcal{X}_\alpha\,|\,\mathbf{p}\right)\right]^2}\right)\Bigg\}\,d\,\alpha\Bigg]$$

for $l' \in \{1, \ldots, r\}$.

As $\sqrt{n}(\mathbf{f}_n - \mathbf{p})$ is asymptotically distributed as an $(r-1)$-dimensional normal $N\Big(\mathbf{0},\, \left[\mathrm{I}\,_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}\Big)$, because of the properties of the convergence in law, we have that as $n \to \infty$

$$\nabla \mathrm{I}\,_f(\mathbf{p}) \left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \xrightarrow{\mathcal{L}} N\left(0,\, \nabla \mathrm{I}\,_f(\mathbf{p}) \left[\mathrm{I}\,_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1} (\nabla \mathrm{I}\,_f(\mathbf{p}))^t\right),$$

whenever $\nabla \mathrm{I}\,_f(\mathbf{p}) \left[\mathrm{I}\,_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1} (\nabla \mathrm{I}\,_f(\mathbf{p}))^t > 0$.

In virtue of the properties of the convergences in law and in probability, since $\|\mathbf{p}_n^* - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$ we have that as $n \to \infty$

$$\left(\frac{\partial^2 \mathrm{I}\,_f(\mathcal{X} \mid \mathbf{p})}{\partial p_i \partial p_j}\right)_{\mathbf{p}=\mathbf{p}_n^*} \xrightarrow{p} \frac{\partial^2 \mathrm{I}\,_f(\mathcal{X} \mid \mathbf{p})}{\partial p_i \partial p_j}$$

for any $i, j \in \{1, \ldots, r-1\}$. On the other hand, $\mathbf{f}_n \xrightarrow{c.s.} \mathbf{p}$, and hence $\mathbf{f}_n \xrightarrow{p} \mathbf{p}$, as $n \to \infty$, whence

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p}) H(\mathrm{I}\,_f(\mathbf{p}_n^*)) \xrightarrow{p} \mathbf{0},$$

and due to the fact that $\sqrt{n}(\mathbf{f}_n - \mathbf{p}) \xrightarrow{\mathcal{L}} N\left(\mathbf{0},\, \left[\mathrm{I}\,_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}\right)$, we obtain that

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p}) H(\mathrm{I}\,_f(\mathbf{p}_n^*)) \left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \xrightarrow{\mathcal{L}} 0,$$

and, consequently,

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p}) H(\mathrm{I}\,_f(\mathbf{p}_n^*)) \left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \xrightarrow{p} 0.$$

The above results guarantee that

$$\sqrt{n}\left[\mathrm{I}\,_f(\mathcal{X} \mid \mathbf{f}_n) - \mathrm{I}\,_f(\mathcal{X} \mid \mathbf{p})\right]$$

$$\xrightarrow{\mathcal{L}} N\left(0, \nabla \mathrm{I}\,_f(\mathbf{p}) \left[\mathrm{I}\,_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1} (\nabla \mathrm{I}\,_f(\mathbf{p}))^t\right)$$

as $n \to \infty$, and whenever $\sigma^2(\mathbf{p}) = \nabla \mathrm{I}_f(\mathbf{p}) \left[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\right]^{-1} (\nabla \mathrm{I}_f(\mathbf{p}))^t = \mathrm{V}^*\Sigma\mathrm{V}^{*t} > 0$, with V^* = the $1 \times r$ matrix $(\mathrm{V}_1^*, \ldots, \mathrm{V}_r^*)$ and

$$\Sigma = \begin{pmatrix} p_1(1-p_1) & \cdots & -p_1p_r \\ \vdots & \ddots & \vdots \\ -p_rp_1 & \cdots & p_r(1-p_r) \end{pmatrix}$$

since

$$\sigma^2(\mathbf{p}) = \sum_{i=1}^{r-1}\sum_{j=1}^{r-1} p_i(\delta_{ij} - p_j) \left(\frac{\partial}{\partial p_i}\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})\right)\left(\frac{\partial}{\partial p_j}\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})\right) = \mathrm{V}^*\Sigma\mathrm{V}^{*t}.$$

iii) Matrix $\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})$ is positive definite, whence the quadratic forms associated with $\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})$ and $\left[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\right]^{-1}$ are positive definite. Consequently, if $\sigma^2(\mathbf{p}) = \nabla \mathrm{I}_f(\mathbf{p}) \left[\mathrm{I}_{\mathcal{X}}^F(\mathbf{p})\right]^{-1} (\nabla \mathrm{I}_f(\mathbf{p}))^t = 0$, then, $\nabla \mathrm{I}_{\mathcal{X}}^F(\mathbf{p}) = 0$.

If we now consider the second Taylor expansion of $\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n)$, one can ensure that for n large enough we have that

$$\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) = \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) + \frac{1}{2}(\mathbf{f}_n - \mathbf{p})H(\mathrm{I}_f(\mathbf{p}))\,(\mathbf{f}_n - \mathbf{p})^t$$

$$+\frac{1}{6}\sum_{i=1}^{r-1}\sum_{j=1}^{r-1}\sum_{k=1}^{r-1}\left(\frac{\partial^3 \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})}{\partial p_i \partial p_j \partial p_k}\right)_{\mathbf{p}=\mathbf{p}_n^{**}} (f_{ni} - p_i)(f_{nj} - p_j)(f_{nk} - p_k),$$

with $\mathbf{p}_n^{**} \in \mathbb{P}$ such that $\|\mathbf{p}_n^{**} - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$. Therefore,

$$\begin{aligned} &2n\left[\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) - \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})\right] \\ &= \sqrt{n}(\mathbf{f}_n - \mathbf{p})H(\mathrm{I}_f(\mathbf{p}))\left(\sqrt{n}(\mathbf{f}_n - \mathbf{p})\right)^t \\ &+ \frac{1}{3}\sum_{i=1}^{r-1}\sum_{j=1}^{r-1}\sum_{k=1}^{r-1}\left(\frac{\partial^3 \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p})}{\partial p_i \partial p_j \partial p_k}\right)_{\mathbf{p}=\mathbf{p}_n^{**}} \\ &\cdot (f_{ni} - p_i)\left(\sqrt{n}(f_{nj} - p_j)\right)\left(\sqrt{n}(f_{nk} - p_k)\right). \end{aligned}$$

Following the arguments used to prove *ii)*, we can conclude that

$$(f_{ni} - p_i)\left(\sqrt{n}(f_{nj} - p_j)\right)\left(\sqrt{n}(f_{nk} - p_k)\right) \xrightarrow{p} 0.$$

Furthermore, we have that

$$\left(\sqrt{n}(\mathbf{f}_n-\mathbf{p})\right) H(\mathrm{I}_f(\mathbf{p}))\left(\sqrt{n}(\mathbf{f}_n-\mathbf{p})\right)^t \xrightarrow{\mathcal{L}} Y H(\mathrm{I}_f(\mathbf{p}))\, Y^t,$$

where Y is a random vector with an $(r-1)$-dimensional normal distribution $N\Big(\mathbf{0},\Big[\mathrm{I}^F_{\mathcal{X}}(\mathbf{p})\Big]^{-1}\Big)$.

Because of the assumption that some of the second derivatives $H(\mathrm{I}_f(\mathbf{p}))$ are positive, and since the rank of $\Big[\mathrm{I}^F_{\mathcal{X}}(\mathbf{p})\Big]^{-1}$ is $r-1$, the vector Y can be expressed as $Y = ZB^*$ where Z is the $(r-1)$-dimensional random vector whose components are $r-1$ independent and identically distributed real-valued random variables, all of them having a standard normal distribution $N(0,1)$, and B^* is an $(r-1)\times(r-1)$ matrix such that $B^*B^{*t} = \Big[\mathrm{I}^F_{\mathcal{X}}(\mathbf{p})\Big]^{-1}$. Moreover, there exists a transformation $Z = UC^*$ in which C^* is an orthogonal matrix such that

$$Y H(\mathrm{I}_f(\mathbf{p}))\, Y^t = ZB^* H(\mathrm{I}_f(\mathbf{p}))\, B^{*t} Z^t$$

$$= UC^*B^* H(\mathrm{I}_f(\mathbf{p}))\, B^{*t}C^{*t}U^t = \lambda_1^* U_1^2 + \cdots + \lambda_q^* U_q^2,$$

$\lambda_1^*, \ldots, \lambda_q^*$ $(q \leq r-1)$ being the nonnull eigenvalues of the $(r-1)\times(r-1)$ matrix $B^* H(\mathrm{I}_f(\mathbf{p}))\, B^{*t}$, and where $U_1, \ldots, U_q$ are independent and identically distributed $N(0,1)$ variables, that is, a linear combination of independent and identically distributed chi-square χ^2_1 variables. □

Remark 2.4.1. When we particularize the asymptotic variance $\sigma^2(\mathbf{p})$ in Theorem 2.4.1 for the hyperbolic index (i.e., we assume $f(x) = x^{-1} - 1$ for all $x \in (0,+\infty) = \mathrm{Dom} f$), then we obtain in terms of the notations in the proof of Theorem 2.3.6 that $\sigma^2(\mathbf{p})$ is equivalent to the variance of the real-valued random variable which takes on values

$$\mathrm{T}_j = \frac{1}{N}\sum_{l=1}^{N}(\mathrm{V}_{jl} + \mathrm{V}_{lj})$$

with probabilities equal to $1/N$, for $j \in \{1, \ldots, N\}$, so that

$$\sigma^2(\mathbf{p}) = \frac{1}{N}\sum_{j=1}^{N} \mathrm{T}_j^2 - \left(\frac{1}{N}\sum_{j=1}^{N} \mathrm{T}_j\right)^2$$

$$= \frac{1}{N}\sum_{j=1}^{N}\left(\frac{1}{N}\sum_{l=1}^{N}[\mathrm{V}_{jl} + \mathrm{V}_{lj}]\right)^2 - 4\left(\frac{1}{N^2}\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{V}_{jl}\right)^2$$

$$= \lim_{n\to\infty} \mathrm{Var}\left(\sqrt{n}\widehat{\mathrm{I}_H}^{w}(\mathcal{X}[\,\cdot\,])\right).$$

The property *ii)* in the latter theorem can be revised to calculate easily in practice the asymptotic variance of the estimator $\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n)$ and to develop (although in an approximate way) inferences as the interval estimation and hypotheses testing. More precisely, when $\sigma^2(\mathbf{p})$ is replaced by its analogue estimate, $\sigma^2(\mathbf{f}_n)$, we obtain the following conclusion:

Theorem 2.4.2. *Under the conditions in Theorem* 2.4.1*, we have that*

$$\left\{\frac{\sqrt{n}\,(\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}))}{\sqrt{\sigma^2(\mathbf{f}_n)}}\right\}_n$$

converges in law to a standard normal distribution $N(0,1)$ *as* $n \to \infty$, *whenever* $\sigma^2(\mathbf{p}) > 0$ *and* $\sigma^2(\mathbf{f}_n) > 0$.

Proof.

Since the components of $\nabla\mathrm{I}_f(\mathbf{p})$ and $\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}$ are continuous in a neighborhood of $\mathbf{p}$, then as $n \to \infty$ we have that

$$\nabla\mathrm{I}_f(\mathbf{f}_n) \xrightarrow{p} \nabla\mathrm{I}_f(\mathbf{p}),$$

and

$$\left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{f}_n)\right]^{-1} \xrightarrow{p} \left[\mathrm{I}_{\mathcal{X}}^{F}(\mathbf{p})\right]^{-1}.$$

Consequently,

$$\sqrt{\sigma^2(\mathbf{f}_n)} \xrightarrow{p} \sqrt{\sigma^2(\mathbf{p})}$$

as $n \to \infty$, and since

$$\sqrt{n}\,(\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})) \xrightarrow{\mathcal{L}} N(0,\,\sigma^2(\mathbf{p})),$$

then we have that

$$\frac{\sqrt{n}\,(\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p}))}{\sqrt{\sigma^2(\mathbf{f}_n)}} \xrightarrow{\mathcal{L}} N(0,\,1).$$

□

The study we have just developed allows us to perform some additional inferences on the population relative variation.

Concerning the *interval estimation*, the procedure we now present will provide us with a range of possible values for the population relative variation which will cover the true value of this measure with (in this case approximately) a prescribed probability.

Theorem 2.4.2 allows us to state approximately the limits defining the above range, as follows:

Theorem 2.4.3. *Under the conditions in Theorems* 2.4.1 *and* 2.4.2*, the random interval*

$$\left[\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - z_\alpha\sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}},\ \mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) + z_\alpha\sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}}\right],$$

supplies for each sample of n *independent observations from* $\mathcal{X}$ *confidence intervals of* $\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{p})$ *with coefficient approximately equal to* $1-\alpha$ $(\alpha \in [0,1])$.

Finally, and in an immediate way, we can derive the following *tests of hypotheses*

Theorem 2.4.4. *Under the conditions in Theorems* 2.4.1 *and* 2.4.2*,*

(i) To test at the significance level $\alpha \in [0,1]$ *the null hypothesis*

$$H_0 : \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) = i_0$$

against the alternative one

$$H_1 : \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) \neq i_0,$$

H_0 *must be rejected whenever* $|\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) - i_0| > z_\alpha \sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}}$*. The p-value of the test is given, approximately, by*

$$p = 2\left[1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\,|\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) - i_0|\right)\right].$$

(ii) To test at the significance level α *the null hypothesis*

$$H_0 : \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) \geq i_0$$

against the alternative one

$$H_1 : \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) < i_0,$$

H_0 *must be rejected whenever* $\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) - i_0 < z_{2\alpha} \sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}}$*. The p-value of the test is given, approximately, by*

$$p = \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\,(\mathrm{I}_f(\mathcal{X} \,|\, \mathbf{f}_n) - i_0)\right).$$

(iii) To test at the significance level α *the null hypothesis*

$$H_0 : \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) \leq i_0$$

against the alternative one

$$H_1 : \mathrm{I}_f(\mathcal{X} \,|\, \mathbf{p}) > i_0,$$

H_0 must be rejected whenever $\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - i_0 > z_{2\alpha}\sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$. The p-value of the test is given, approximately, by

$$p = 1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\,(\mathrm{I}_f(\mathcal{X}\,|\,\mathbf{f}_n) - i_0)\right).$$

Some of the developed studies in the two latter sections will be illustated by means of the following example.

Example 2.4.1. In Klir & Yuan (1995) it has been pointed out that there is a large number of situations in Civil Engineering to which Fuzzy Set Theory has already proven to be especially valuable, like those consisting of problems of assessing or evaluating existing constructions. Typical examples of these problems are the assessment of fatigue in metal structure, the assessment of quality of highway pavements, the assessment of damage to a building after an earthquake, etc.

Klir & Yuan have mentioned as an example in the study of the physical conditions of highway bridges, the variable CURRENT CONDITION OF THE PIERS ($\mathcal{X}$) of a bridge, whose values are POOR ($\tilde{x}_1$), FAIR ($\tilde{x}_2$) and GOOD ($\tilde{x}_3$), and which have been assumed to be characterized (Klir & Yuan 1995) by means of the triangular fuzzy numbers, $\tilde{x}_1 = \mathrm{Tri}(1,2,3)$, $\tilde{x}_2 = \mathrm{Tri}(2,3,5)$ and $\tilde{x}_3 = \mathrm{Tri}(3,5,5)$ (see Figure 2.3).

In the United States there are approximately 600,000 highway bridges, about one half of which were built before 1940. To estimate by point and testing the value of the f-inequality index associated with $\mathcal{X}$ in the population Ω of the 600,000 US bridges, with $f = x^{-1} - 1$, we can consider a simple random sample v of $n = 400$ bridges in the US. Assume that the above sample provides us with the following data

$\tilde{x}_l$	POOR	FAIR	GOOD
f_{nl}	15	56	329

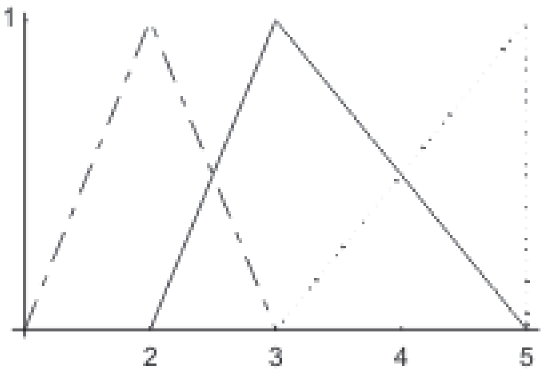


Fig. 2.3: Fuzzy values of the CURRENT CONDITION OF THE PIERS of the highway bridges

In virtue of Theorem 2.3.2, $\mathrm{I}_H(\mathcal{X})$ can be estimated by means of the value

$$\widehat{\mathrm{I}_H}(\mathcal{X}[v]) = \frac{400(599,999)}{600,000(399)} \mathrm{I}_H(\mathcal{X}[v])$$

$$-\frac{1-\dfrac{400}{600,000}}{400}\left\{\frac{1}{399}\sum_{l=1}^{3} \mathrm{I}_H(\{\tilde{x}_l\})\, f_{nl}\right\} = .099.$$

On the other hand, if we want to test the hypothesis

$$H_0 : \mathrm{I}_H(\mathcal{X}) \leq .05$$

against the alternative

$$H_1 : \mathrm{I}_H(\mathcal{X}) > .05,$$

we obtain for the sample information the p-value

$$p = 1 - \Phi\left(\sqrt{\frac{399}{\sigma^2(\mathbf{f}_n)}}\,(\mathrm{I}_H(\mathcal{X}[v]) - i_0)\right) = .19$$

so that we can conclude that the null hypothesis H_0 is quite supported by the sample information.

2.5 Final discussion and future directions

To conclude this chapter we can indicate that in order to preserve some of the fundamental properties of the inequality indices in the real-valued case, as those concerning the principles of transfers and the agreement with Lorenz's criterion, we have considered the strong ordering $\succeq_S$ by Ramík & Římánek (1985) to formalize the assumptions on the ranking of the values of the fuzzy random variable. Although it is a restrictive condition in the fuzzy case (because $\succeq_S$ establishes a partial ordering on $\mathcal{F}_c(\mathbb{R})$), it allows us to obtain strong assertions (the same obtained in the real-valued case) in comparing the values of each inequality index.

The additive decomposability cannot be preserved for the indices I^α, except for the index of the Shannon type.

In contrast with the computation of the fuzzy-valued inequality indices (see López García 1997 and Colubi Cervero 1997), the f-inequality indices in the present work are definitely much easier to be computed, and we do not need to supply a fuzzy number (usually its graphical representation) as the answer to the problem of quantifying the inequality. Nevertheless, most of the software developed by López García and Colubi Cervero, in what concerns the characterization of fuzzy data, and the calculus of the induced images of the functions f and the expected values of fuzzy random variables, are useful for the aim of the present work because of the complementary information it provides us with, although the final calculus of the f-inequality indices only requires quite simple approximation techniques.

In connection with the research developed in this chapter, the following problems are similar to those in Chapter 1:

- Results in Sections 2.3 and 2.4 could be extended to more complex samplings, like the stratified one.

- The asymptotic properties and the inferences based on them in Section 2.4 can be developed in a similar way to compare the inequality of two populations in terms of two independent random samples from them. For both the two-sample, and the one-sample problem studied in the present work, suitable sample sizes guaranteeing the adequacy of the assymptotic approximations should be analyzed.

- Most of the studies in this chapter could be immediately extended by considering a more general index, obtained by composing the λ-average ranking function by Campos & González (1989) with $\lambda \in [0, 1]$, which when particularized to $\lambda = .5$ leads to the function F by Yager (1981). The only (even though determinant) inconvenience to deal with this more general ranking is that for values $\lambda \in [0, .5)$ the nonnegativeness and minimality properties could be lost (actually, Theorem 2.2.6 could not be stated for $\lambda \neq .5$). Consequently, the comparison of populations/samples and/or variables through indices which could take on nonnegative values would often lead to wrong conclusions.

Chapter 3

Generalized measures of the absolute and relative variation of a random set

In quantifying the extent of the absolute and relative variation of a quantitative attribute, one usually assumes that it is modelled by means of a real-valued random variable. Nevertheless, some variables (like, for instance, the numerical ranges of certain characteristics) are not properly modelled in terms of real-valued random variables, but rather in terms of a special kind of fuzzy random variables. These special variables are those being essentially interval-valued, and some of their values may overlap.

These random elements can be formally identified with particular cases of random sets (more precisely, the convex compact random sets taking on values belonging to $\mathcal{K}_c(\mathbb{R})$).

Since these random sets are particular $\mathcal{F}_c(\mathbb{R})$-valued fuzzy random variables, concepts and conclusions in Chapters 1 and 2 can be immediately applied to these random sets.

In the present chapter we will develop such an application, we will pin down some aspects characterizing this interesting situation and we will pay attention to the particular expressions for estimators, parameters of the asymptotic distributions, etc.

Firstly, a generalized measure of the absolute variation of a convex compact random set about a compact real interval is introduced. We determine an unbiased estimate of this measure and the corresponding associated error in random samplings with and without replacement from finite populations, the asymptotic distribution of the sample absolute variation measure, and inferential procedures based on this distribution. The applications of the absolute variation measure to quantifying the sampling error associated with the estimation of the (population) Aumann integral of the convex compact random set in random samplings with and without replacement from finite populations, and to the regression analysis between two convex compact random sets in general populations, are carried out.

Secondly, a generalized measure of the relative variation of a convex compact random set is introduced. We determine an unbiased estimate of a particular measure (the hyperbolic index), as well as the asymptotic distribution of the measures satisfying quite general regularity conditions (in fact, all the measures being of practical interest) and some related inferential methods.

3.1 The S-mean squared dispersion associated with a random set. Inferential results for finite populations

In this section we present the particularization of the generalized measure of absolute variation in Chapter 1, to the case in which we consider a convex compact random set and an element of $\mathcal{K}_c(\mathbb{R})$.

Let $(\Omega, \mathcal{A}, P)$ be a probability space and let $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ be an integrably bounded convex compact random set associated with $(\Omega, \mathcal{A}, P)$.

Definition 3.1.1. *The S-mean squared dispersion associated with X about the interval $K \in \mathcal{K}_c(\mathbb{R})$, is the real value (if it exists) given by*

$$MSD_S(X,K) = E\left([d_S(X(\cdot),K)]^2\right) = \int_\Omega [d_S(X(\omega),K)]^2 dP(\omega)$$
$$= \int_\Omega \int_{[0,1]} \left[f_{X(\omega)}(\lambda) - f_K(\lambda)\right]^2 dS(\lambda)\, dP(\omega),$$

where $f_K(\lambda) = \lambda \sup K + (1-\lambda) \inf K$ for all $\lambda \in [0,1]$.

The central S-mean squared dispersion associated with a convex compact random set will be defined as follows:

Definition 3.1.2. *The central S-mean squared dispersion associated with X is defined to be the real value (if it exists) given by*

$$\Delta_S^2(X) = MSD_S(X, E(X)) = \int_\Omega [d_S(X(\omega), E(X))]^2 dP(\omega).$$

When we need to specify the probability distribution P in the probability space, we will denote $\Delta_S^2(X)$ by $\Delta_S^2(X \mid P)$.

The existence conditions and the properties in Sections 1.1 and 1.2 could be particularized in an immediate way for the measures $MSD_S(X,K)$ and $\Delta_S^2(X)$.

The following example illustrates the computation of the absolute variation for a convex compact random set on a real-life example, whose data have been supplied by the Servicio de Nefrología of the Hospital Valle del Nalón in Langreo (Asturias).

Example 3.1.1. Data in Table 3.3 correspond to the "range of the cardiac frequency over a day", X, observed in a population Ω of 59 patients who are hospitalized. Values of X are obtained from several measures of the cardiac frequency of each patient at different moments (usually 60 to 70) over a concrete day.

58-90	64-107	52-78	56-133	54-78	75-124
47-68	54-84	55-84	37-75	53-103	58-99
32-114	47-95	61-101	61-94	47-86	59-78
61-110	56-90	65-92	44-110	70-132	55-89
62-89	44-108	38-66	46-83	63-115	55-80
63-119	63-109	48-73	52-98	47-83	70-105
51-95	62-95	59-98	56-84	56-103	40-80
49-78	48-107	59-87	54-92	71-121	56-97
43-67	26-109	49-82	53-120	68-91	37-86
55-102	61-108	48-77	49-88	62-100	

Table 3.1: Data on ranges of cardiac frequency

If we wish to measure the variation of X in Ω, and we consider the central S-MSD associated with X for the discretized measure S such that $L=3$, $\lambda_1=0$, $\lambda_2=.5$, $\lambda_3=1$, $k_1=k_2=k_3=1/3$, and $\overline{g}(\lambda)=0$ if $\lambda\in(0,1)\setminus\{.5\}$, we obtain that

$$\Delta_S^2(X)=146.833.$$

We are now going to study the particularization to random sets of the results obtained in Sections 1.3 and 1.4 in connection with the problem of estimating the generalized measure of the absolute variation in random samplings from finite populations, and with the problem of determining under certain conditions the asymptotic distribution of the sample measure in finite populations and the approximate procedures based on them.

Consider a finite population Ω of N units, $\omega_1,\ldots,\omega_N$, and let $X:\Omega\to\mathcal{K}_c(\mathbb{R})$ be a convex compact random set associated with a measurable space defined on Ω and being endowed with the uniform distribution.

Assume that a sample of size n is chosen at random and without replacement from Ω, υ denotes a generic simple random sample of size n, and

$\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ are the units in υ. Then the *sample central S-mean squared dispersion* of X in υ is given by

$$\Delta_S^2\left(X[\upsilon]\right) = \frac{1}{n}\sum_{i=1}^{n}\left[d_S\left(X(\omega_{\upsilon i}), \overline{X}_n[\upsilon]\right)\right]^2,$$

where $\overline{X}_n[\upsilon] = \dfrac{1}{n}\left[X(\omega_{\upsilon 1}) + \cdots + X(\omega_{\upsilon n})\right]$ is the sample expected value of X in υ.

$\Delta_S^2\left(X[\,\cdot\,]\right)$ is a real-valued random variable associated with the probability space $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ (see Section 1.3, p. 45), and therefore it defines an estimator of the *population central S-mean squared dispersion* which is now given by

$$\Delta_S^2(X \mid P) = \frac{1}{N}\sum_{j=1}^{N}\left[d_S\left(X(\omega_j), \overline{X}\right)\right]^2,$$

with $\overline{X} = \dfrac{1}{N}\left[X(\omega_1) + \cdots + X(\omega_N)\right]$.

Concerning the unbiased estimation (in the sense of the Aumann integral over Υ_n), the conclusions we can obtain in this particular case are the following:

Theorem 3.1.1. *In random sampling without replacement of size n from the population Ω, we have that the estimator $\widehat{\Delta_S^2}(X[\,\cdot\,])$ which for sample υ takes on the value*

$$\widehat{\Delta_S^2}(X[\upsilon]) = \frac{(N-1)n}{N(n-1)}\Delta_S^2\left(X[\upsilon]\right),$$

is an unbiased estimator of $\Delta_S^2(X \mid P)$.

Theorem 3.1.2. *In random sampling without replacement of size n from the population Ω, we have that if $f = n/N$ then*

$$\operatorname{Var}\left(\widehat{\Delta_S^2}(X[\,\cdot\,])\right) = \frac{(1-f)}{n(n-1)N^2(N-2)(N-3)}$$

$$\cdot\Bigg\{2(N-n-1)(N-1)\sum_{j=1}^{N}\sum_{l=1}^{N}\Bigg(\int_{[0,1]}\left[f_{X(\omega_j)}(\lambda)-f_{\overline{X}}(\lambda)\right]$$

$$\cdot\left[f_{X(\omega_l)}(\lambda)-f_{\overline{X}}(\lambda)\right]\,d\,S(\lambda)\Bigg)^2+[(6-4N)n-8(N-1)]\left(\Delta_S^2(X\,|\,P)\right)^2$$

$$+[n(N-2)(N^2-3)-(N-1)(N^2+N-8)]N\ \mathrm{Var}\left(\left[d_S(X,\overline{X})\right]^2\,\Big|\,P\right)\Bigg\}.$$

If we consider a random choice with replacement of n units from the population Ω, υ denotes a generic random sample with reposition of size n, and $\omega_{\upsilon 1},\ldots,\omega_{\upsilon n}$ are the units in υ, then $\Delta_S^2(X[\,\cdot\,])$ is a real-valued random variable associated with the probability space $(\Upsilon_n^w,\mathcal{P}(\Upsilon_n^w),p^w)$ (see Section 1.3, p. 51), and therefore it defines an estimator of the population central S-MSD. Conclusions are now the following:

Theorem 3.1.3. *In random sampling with replacement of size n from the population Ω, we have that the estimator $\widehat{\Delta_S^2}^w(X[\,\cdot\,])$ which for sample υ takes on the value*

$$\widehat{\Delta_S^2}^w(X[\upsilon])=\frac{n}{(n-1)}\Delta_S^2(X[\upsilon]),$$

in an unbiased estimator of $\Delta_S^2(X\,|\,P)$.

Theorem 3.1.4. *In random sampling with replacement of size n from the population Ω, we have that*

$$\mathrm{Var}\left(\widehat{\Delta_S^2}^w(X[\,\cdot\,])\right)=\frac{1}{n(n-1)N^3}$$

$$\cdot\Bigg\{2(n-1)(N-2)\sum_{j=1}^{N}\sum_{l=1}^{N}\Bigg(\int_{[0,1]}\left[f_{X(\omega_j)}(\lambda)-f_{\overline{X}}(\lambda)\right]$$

$$\cdot\left[f_{X(\omega_l)}(\lambda)-f_{\overline{X}}(\lambda)\right]\,d\,S(\lambda)\Bigg)^2+\ 2[n^2-Nn-2]N^2\ \left(\Delta_S^2(X\,|\,P)\right)^2$$

$$+[N(n-1)^2-2(n^2-5n+5)]N^2\ \mathrm{Var}\left(\left[d_S(X,\overline{X})\right]^2\,\Big|\,P\right)\Bigg\}.$$

To develop other inferential procedures on the population measure of the absolute variation of a convex compact random set, we should consider large samples either chosen at random and with replacement from any population, or chosen at random and without replacement from a large population whose size is substantially bigger than the sample one, and later applying the next results concerning the asymptotic distribution of the sample central S-MSD.

Let $\Omega = \{\omega_1, \ldots, \omega_N\}$ be a finite population and let $X : \Omega \to \mathcal{K}_c(\mathbb{R})$ be a convex compact random set. On Ω we can define a probability space $(\Omega, \mathcal{P}(\Omega), P)$, where P means the probability measure corresponding to the uniform distribution on Ω.

If on the population Ω variable X takes on r different values, $x_1^*, \ldots, x_r^*$, and, for $l \in \{1, \ldots, r\}$, p_l denotes the probability $P(\{\omega \in \Omega \mid X(\omega) = x_l^*\})$, and S is a measure on $[0,1]$ satisfying the conditions in Definition 0.1.5, then the *population central S-MSD* associated with X in Ω can be expressed as

$$\Delta_S^2(X \,|\, \mathbf{p}) = \sum_{l=1}^{r} p_l \left[d_S\left(x_l^*, E(X \,|\, \mathbf{p})\right)\right]^2,$$

with $E(X \,|\, \mathbf{p}) = \sum_{l=1}^{r} p_l\, x_l^*$ y $\mathbf{p} = (p_1, \ldots, p_{r-1})$ $(p_r = 1 - \sum_{l=1}^{r-1} p_l)$.

If a sample of size n is chosen at random from Ω and f_{nl} is the relative frequency of the value x_l^* of X in the sample, the *sample central S-MSD* is given by

$$\Delta_S^2(X \,|\, \mathbf{f}_n) = \sum_{l=1}^{r} f_{nl} \left[d_S\left(x_l^*, \overline{X}_n\right)\right]^2,$$

with $\overline{X}_n = \sum_{l=1}^{r} f_{nl} x_l^*$, y $\mathbf{f}_n = (f_{n1}, \ldots, f_{n(r-1)})$ $(f_{nr} = 1 - \sum_{l=1}^{r-1} f_{nl})$.

The asymptotic distribution of the sample central S-MSD, and its asymptotic properties in estimating $\Delta_S^2(X)$, are gathered in the following result:

Theorem 3.1.5. *For each $n \in \mathbb{N}$, consider n independent convex compact random sets being identically distributed as X (that is, a simple random sample of size n from X), defined over the finite population $\Omega =$*

$\{\omega_1, \ldots, \omega_N\}$ *such that* $P(\{\omega \in \Omega \mid X(\omega) = x_l^*\}) = p_l$ *with* $p_l \in (0,1)$, $l = 1, \ldots, r$, $\sum_{l=1}^{r} p_l = 1$. *Let* S *be a normalized measure on* $[0,1]$, *satisfying the conditions in Definition* 0.1.5. *Then,*

i) If $\mathbf{f}_n = (f_{n1}, \ldots, f_{n(r-1)}) \in [0,1]^{r-1}$, *with* f_{nl} = *relative frequency of* x_l^* *in the performance of the simple random sample of size* n ($l = 1, \ldots, r-1$, $f_{nr} = 1 - \sum_{l=1}^{r-1} f_{nl}$), *and* $\Delta_S^2(X \mid \mathbf{f}_n)$ *is the associated sample central* S*-MSD, then* $\{\Delta_S^2(X \mid \mathbf{f}_n)\}_n$ *is a sequence of estimators of* $\Delta_S^2(X \mid \mathbf{p})$ *which is strongly consistent, that is, as* $n \to \infty$ *we have that*

$$\Delta_S^2(X \mid \mathbf{f}_n) \stackrel{a.s.}{\to} \Delta_S^2(X \mid \mathbf{p})$$

whatever $\mathbf{p} = (p_1, \ldots, p_{r-1})$ *(with* $p_1, \ldots, p_{r-1} \in (0,1)$ *y* $\sum_{l=1}^{r-1} p_l < 1$*) may be.*

ii) $\{\sqrt{n}\,(\Delta_S^2(X \mid \mathbf{f}_n) - \Delta_S^2(X \mid \mathbf{p}))\}_n$ *is a sequence of real-valued random variables converging in law as* $n \to \infty$ *to a one-dimensional normal distribution* $N(0, \sigma^2(\mathbf{p}))$, *with*

$$\sigma^2(\mathbf{p}) = \operatorname{Var}\left([d_S\,(X, E(X \mid \mathbf{p}))]^2\right),$$

whenever $\sigma^2(\mathbf{p}) > 0$.

iii) If $\sigma^2(\mathbf{p}) = 0$, *and there is a pair* (i,j) *with* $i, j \in \{1, \ldots, r-1\}$ *such that*

$$h_{ij} = \frac{\partial^2 \Delta_S^2(X \mid \mathbf{p})}{\partial p_i \partial p_j}$$
$$= \frac{\partial}{\partial p_i}\left(\left[d_S\left(x_j^*, E(X \mid \mathbf{p})\right)\right]^2 - [d_S\,(x_r^*, E(X \mid \mathbf{p}))]^2\right) > 0,$$

then $\{2n\,(\Delta_S^2(X \mid \mathbf{f}_n) - \Delta_S^2(X \mid \mathbf{p}))\}_n$ *is a sequence of real-valued random variables converging in law as* $n \to \infty$ *to a linear combination of, at most,* $r-1$ *chi-squared* χ_1^2 *independent random variables.*

When $\sigma^2(\mathbf{p})$ is substituted by its analogue estimate, $\sigma^2(\mathbf{f}_n)$, we can conclude that

Theorem 3.1.6. *Under the conditions in Theorem* 3.1.5, *we have that*

$$\left\{ \frac{\sqrt{n}\,(\Delta_S^2(X\,|\,\mathbf{f}_n) - \Delta_S^2(X\,|\,\mathbf{p}))}{\sqrt{\sigma^2(\mathbf{f}_n)}} \right\}_n$$

converges in law to a standard normal distribution $N(0,1)$ *as* $n \to \infty$, *whenever* $\sigma^2(\mathbf{p}) > 0$ *and* $\sigma^2(\mathbf{f}_n) > 0$.

The following result allows us to establish in an approximate way *confidence intervals* for $\Delta_S^2(X)$:

Theorem 3.1.7. *Under the conditions in Theorems* 3.1.5 *and* 3.1.6, *the random interval*

$$\left[\Delta_S^2(X\,|\,\mathbf{f}_n) - z_\alpha \sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}},\ \Delta_S^2(X\,|\,\mathbf{f}_n) + z_\alpha \sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}} \right]$$

supplies for each sample of n *independent observations from* X *confidence intervals of* $\Delta_S^2(X\,|\,\mathbf{p})$ *with coefficient approximately equal to* $1-\alpha$ $(\alpha \in [0,1])$.

In the same way, we can derive the following *hypotheses testing*:

Theorem 3.1.8. *Under the conditions in Theorems* 3.1.5 *and* 3.1.6,

(i) To test at the significance level $\alpha \in [0,1]$ *the null hypothesis*

$$H_0 : \Delta_S^2(X\,|\,\mathbf{p}) = \delta_0$$

against the alternative hypothesis

$$H_1 : \Delta_S^2(X\,|\,\mathbf{p}) \neq \delta_0,$$

H_0 *must be rejected whenever* $|\Delta_S^2(X\,|\,\mathbf{f}_n) - \delta_0| > z_\alpha \sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}}$. *The p-value of the test is given, approximately, by*

$$p = 2\left[1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\left|\Delta_S^2(X\,|\,\mathbf{f}_n) - \delta_0\right|\right)\right].$$

(ii) To test at the significance level α *the null hypothesis*

$$H_0 : \Delta_S^2(X \mid \mathbf{p}) \geq \delta_0$$

against the alternative

$$H_1 : \Delta_S^2(X \mid \mathbf{p}) < \delta_0,$$

H_0 *must be rejected whenever* $\Delta_S^2(X \mid \mathbf{f}_n) - \delta_0 < z_{2\alpha}\sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$. *The p-value of the test is given, approximately, by*

$$p = \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\left(\Delta_S^2(X \mid \mathbf{f}_n) - \delta_0\right)\right).$$

(iii) To test at the significance level α *the null hypothesis*

$$H_0 : \Delta_S^2(X \mid \mathbf{p}) \leq \delta_0$$

against the alternative

$$H_1 : \Delta_S^2(X \mid \mathbf{p}) > \delta_0,$$

H_0 *must be rejected whenever* $\Delta_S^2(X \mid \mathbf{f}_n) - \delta_0 > z_{2\alpha}\sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$. *The p-value of the test is given, approximately, by*

$$p = 1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\left(\Delta_S^2(X \mid \mathbf{f}_n) - \delta_0\right)\right).$$

3.2 Some applications of the measure of the absolute variation of a convex compact random set

In this section we are going to examine the particularizations of the applications studied in Sections 1.5 and 1.6.

In this way, we first consider the problem of estimating the expected value of a convex compact random set in random samplings with and without replacement from finite populations, and we later analyze the problems of linear and general functional regression between two convex compact random sets.

Concerning the problem of the *interval-valued estimation of the population expected value* of a convex compact random set, if we consider a finite population of N units, $\Omega = \{\omega_1, \ldots, \omega_N\}$, on which the convex compact random set $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ associated with the probability space $(\Omega, \mathcal{P}(\Omega), P)$ takes on the values $X(\omega_1), \ldots, X(\omega_N)$, and a generic sample υ of size n chosen at random and without replacement from Ω, and $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ denote the units in sample υ, then the *sample expected value* of X in υ, $\overline{X}_n[\upsilon] = [X(\omega_{\upsilon 1}) + \cdots + X(\omega_{\upsilon n})] / n$, defines a convex compact random set associated with the probability space $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ (see Section 1.3).

The expected value of $\overline{X}_n$ over the space $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ is given by

Theorem 3.2.1. *In random sampling without replacement of size n from the population Ω of N units, the $\mathcal{K}_c(\mathbb{R})$-valued estimator $\overline{X}_n$ is unbiased to estimate the population expected value $\overline{X} = E(X \mid P)$, that is, $E(\overline{X}_n \mid p) = \overline{X}$.*

The accuracy of this $\mathcal{K}_c(\mathbb{R})$-valued estimator is discussed in the following results:

Theorem 3.2.2. *In random sampling without replacement of size n from the population Ω of N units, the S-mean squared dispersion of $\overline{X}_n$ about $\overline{X}$ is given by*

$$MSD_S\left(\overline{X}_n, \overline{X}\right) = \Delta_S^2\left(\overline{X}_n \mid p\right) = \frac{(1-f)N}{n(N-1)} \Delta_S^2\left(X \mid P\right).$$

Theorem 3.2.3. *In random sampling without replacement of size n from the population Ω of N units, the estimator associating to sample υ the value*

$$\widehat{\Delta^2_S}\left(\overline{X}_n[\upsilon]\right) = \frac{(N-1)n}{N(n-1)}\Delta^2_S\left(\overline{X}_n[\upsilon]\right)$$

$$= \frac{(N-1)n}{N(n-1)}\int_{[0,1]} \mathrm{Var}\,(f_{\overline{X}_n}(\lambda))\,d\,S(\lambda),$$

is unbiased to estimate $\Delta^2_S\left(\overline{X}_n \,|\, p\right)$.

Regarding the suitable sample size n in estimating $\overline{X}$ by means of $\overline{X}_n$, this size can be approximated in accordance with Tchebychev's Inequality as follows:

Theorem 3.2.4. *In random sampling without replacement of size n from the population Ω of N units, the sample size*

$$n = \left.\frac{N\Delta^2_S(X\,|\,P)}{(N-1)d^2\alpha + \Delta^2_S(X\,|\,P)}\right]$$

satisfies that $\mathrm{P}\Big(d_S(\overline{X}_n, \overline{X}) > d\Big) \leq \alpha.$

If we choose a sample of size n at random and with replacement from the population Ω, the *sample expected value* of X, $\overline{X}_n$, defines a convex compact random set associated with the probability space $(\Upsilon^w_n, \mathcal{P}(\Upsilon^w_n), p^w)$ (see Section 1.3), and holding that

Theorem 3.2.5. *In random sampling without replacement of size n from the population Ω of N units, $\omega_1, \ldots, \omega_N$, the $\mathcal{K}_c(\mathbb{R})$-valued estimator $\overline{X}_n$ is unbiased to estimate the population expected value $\overline{X} = E(X\,|\,P)$, that is, $E(\overline{X}_n\,|\,p^w) = \overline{X}$.*

The discussion on the accuracy of $\overline{X}_n$ in estimating $\overline{X}$, allows us to conclude that

Theorem 3.2.6. *In random sampling without replacement of size n from the population Ω of N units, the S-mean squared dispersion of $\overline{X}_n$ about $\overline{X}$ is given by*

$$MSD_S\left(\overline{X}_n, \overline{X}\right) = \Delta_S^2\left(\overline{X}_n \,|\, p^w\right) = \frac{\Delta_S^2\,(X \,|\, P)}{n}.$$

Theorem 3.2.7. *In random sampling without replacement of size n from the population Ω of N units, the estimator associating to sample υ the value*

$$\widehat{\Delta_S^2}^{w}\left(\overline{X}_n[\upsilon]\right) = \frac{n}{n-1}\Delta_S^2\left(\overline{X}_n[\upsilon]\right),$$

is unbiased to estimate $\Delta_S^2\left(\overline{X}_n \,|\, p^w\right)$.

The *choice of the suitable sample size* is now approximated as follows:

Theorem 3.2.8. *In random sampling without replacement of size n from the population Ω of N units, the sample size*

$$n = \left.\frac{\Delta_S^2(X \,|\, P)}{d^2\alpha}\right\rceil$$

satisfies that $\mathrm{P}\Big(d_S(\overline{X}_n, \overline{X}) > d\Big) \leq \alpha$.

Let X and Y be two integrably bounded convex compact random sets associated with the probability space $(\Omega, \mathcal{A}, P)$. The aim of the *Linear Regression Analysis* between X and Y is to seek for the real values a and b for which $Y = aX + b$ is the "least erroneous" linear relation of Y with respect to X (see Figure 3.1 as a graphical illustrative example of the effect of the linear relation $Y = 2X - 2$, in this case for a certain $\omega \in \Omega$). The error will be assumed to be quantified in terms of the metric d_S on $\mathcal{K}_c(\mathrm{I\!R})$, and we will admit that S is a measure corresponding to a symmetrical distribution

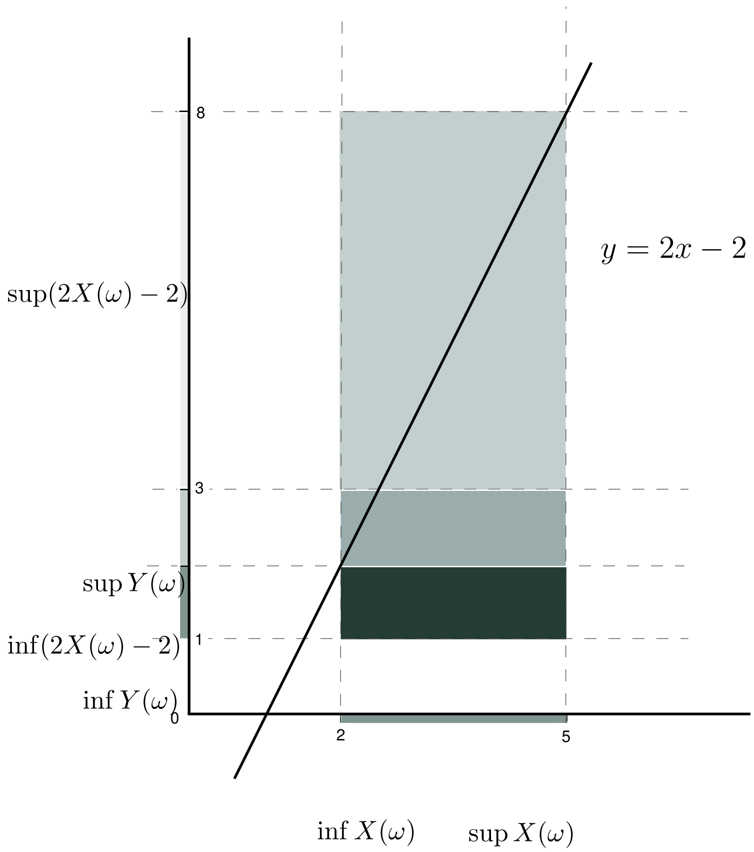


Fig. 3.1: Two-dimensional set-valued data for the observed and estimated (by the linear relation $Y = 2X - 2$) values of Y

with respect to the value $\lambda = .5$, such that our purpose will be minimizing the function

$$\phi(a,b) = E\left([d_S\,(Y, a\,X + b)]^2\right) = E\Big(\int_{[0,1]} [f_Y\,(\lambda) - f_{a\,X+b}\,(\lambda)]^2 \; d\,S(\lambda)\Big)$$

If g is symmetrical with respect to $\lambda = .5$, we have that if

$$W_S\,(A,B) = \int_{[0,1]} f_A\,(\lambda)\, f_B\,(\lambda)\; d\,S(\lambda),$$

and

$$W'_S\,(A,B) = \int_{[0,1]} f_A\,(1-\lambda)\, f_B\,(\lambda)\; d\,S(\lambda),$$

then

$$E\left(W_S\left(X,Y\right)\right)-E\left(W_S'\left(X,Y\right)\right)$$

$$=K(S)E\left((\sup X-\inf X)(\sup Y-\inf Y)\right),$$

with $K(S)=\dfrac{1}{2}\displaystyle\int_{[0,1]}(1-2\lambda)^2\,d\,S(\lambda)>0.$

Therefore, since $(\sup X-\inf X)(\sup Y-\inf Y)\geq 0$ for all $\omega\in\Omega$, we have that

$$E\left(W_S\left(X,Y\right)\right)\geq E\left(W_S'\left(X,Y\right)\right),$$

with equality if, and only if, $\sup X(\,\cdot\,)=\inf X(\,\cdot\,)$ y $\sup Y(\,\cdot\,)=\inf Y(\,\cdot\,)$ a.s. $[P]$. Consequently, $E\left(W_S\left(X,Y\right)\right)=E\left(W_S'\left(X,Y\right)\right)$ if, and only if, X and $\mathcal{Y}$ take on values in $\mathbb{R}$ a.s. $[P]$.

The last assertion idicates that, except for convex compact random sets almost surely real-valued, the optimal value of a cannot be null, and the general solution of the problem is the following:

Theorem 3.2.9. *If X and Y are two integrably bounded convex compact random sets which are not almost surely real-valued, and S is a measure whose associated density g with respect to the Lebesgue measure m on $[0,1]$ is symmetrical with respect to $\lambda=.5$, then the function $\phi(a,b)=E\left(\left[d_S\left(Y,aX+b\right)\right]^2\right)$ is minimized in (a^*,b^*) with $a^*\neq 0$ such that*

$$a^*=\frac{E\left(W_S\left(X,\,Y\right)\right)-E\left(V_S\left(X\right)\right)E\left(V_S\left(Y\right)\right)}{E\left(\left[d_S\left(X,0\right)\right]^2\right)-\left[E\left(V_S\left(X\right)\right)\right]^2},$$

$$b^*=E\left(V_S\left(Y\right)\right)-a^*E\left(V_S\left(X\right)\right),$$

whenever either $E\left(W_S'\left(X,\,Y\right)\right)\geq E\left(V_S\left(X\right)\right)E\left(V_S\left(Y\right)\right)$, or $E\left(W_S\left(X,\,Y\right)\right)>E\left(V_S\left(X\right)\right)\,E\left(V_S\left(Y\right)\right)>E\left(W_S'\left(X,\,Y\right)\right)$, whereas

$$a^*=\frac{E\left(W_S'\left(X,Y\right)\right)-E\left(V_S\left(X\right)\right)E\left(V_S\left(Y\right)\right)}{E\left(\left[d_S\left(X,0\right)\right]^2\right)-\left[E\left(V_S\left(X\right)\right)\right]^2}$$

$$b^*=E\left(V_S\left(Y\right)\right)-a^*E\left(V_S\left(X\right)\right),$$

whenever $E(V_S(X))\ E(V_S(Y)) \geq E(W_S(X,Y))$, *or* $E(W_S(X,Y)) > E(V_S(X))\ E(V_S(Y)) > E(W'_S(X,Y))$, *with*

$$V_S(K) = \int_{[0,1]} f_K(\lambda)\, d\,S(\lambda).$$

On the other hand, given two integrably bounded convex compact random sets X and Y associated with the probability space $(\Omega, \mathcal{A}, P)$, the aim of the *general Functional Regression Analysis* is searching for the $(\mathcal{B}_{d_H}, \mathcal{B}_{d_H})$-measurable function $h : \mathcal{K}_c(\mathbb{R}) \to \mathcal{K}_c(\mathbb{R})$ being the "least erroneous", when the error is assumed to be quantified by means of

$$F(h) = E\left([d_S(Y, h(X))]^2 \,\middle|\, P\right)$$

$$= E\left(\int_{[0,1]} \left[f_Y(\lambda) - f_{h(X)}(\lambda)\right]^2 d\,S(\lambda) \middle| P\right).$$

The solution for this problem is presented in the following result:

Theorem 3.2.10. *If X and Y are two integrably bounded convex compact random sets, and S is a measure satisfying the conditions in Definition* 0.1.5*, then the function F defined on the set $\mathcal{H} = \{h : \mathcal{K}_c(\mathbb{R}) \to \mathcal{K}_c(\mathbb{R}) \mid h(X)$ convex compact random set$\}$ is minimized for the function $h^* : \mathcal{K}_c(\mathbb{R}) \to \mathcal{K}_c(\mathbb{R})$ such that*

$$h^*(x) = E(Y \mid X = x),$$

for each $x \in X(\Omega)$.

The following example illustrates the application of Theorem 3.2.9 by means of a real-life case. Data have been again supplied by the Servicio de Nefrología of the Hospital Valle del Nalón in Langreo.

Example 3.2.1. The paired data in Table 3.2 are the values that for the 59 patients in the population Ω in Example 3.1.1 take on variables X = "range of systolic blood pressure over a day" and Y = "range of diastolic blood pressure over the same day".

X	Y		X	Y		X	Y
118-173	63-102		119-212	47-93		98-160	47-108
104-161	71-108		122-178	73-105		97-154	60-107
131-186	58-113		127-189	74-125		87-150	47-86
105-157	62-118		113-213	52-112		141-256	77-158
120-179	59-94		141-205	69-133		108-147	62-107
101-194	48-116		99-169	53-109		115-196	65-117
109-174	60-119		126-197	60-98		99-172	42-86
128-210	76-125		99-201	55-121		113-176	57-95
94-145	47-104		88-221	37-94		114-186	46-103
148-201	88-130		113-183	55-85		145-210	100-136
111-192	52-96		94-176	56-121		120-180	59-90
116-201	74-133		102-156	50-94		100-161	54-104
102-167	39-84		103-159	52-95		159-214	99-127
104-161	55-98		102-185	63-118		138-221	70-118
106-167	45-95		111-199	57-113		87-152	50-95
112-162	62-116		130-180	64-121		120-188	53-105
136-201	67-122		103-161	55-97		95-166	54-100
90-177	52-104		125-192	59-101		92-173	45-107
116-168	58-109		97-182	54-104		83-140	45-91
98-157	50-111		127-226	57-101			

Table 3.2: Data on the ranges of the systolic (X) and diastolic (Y) blood pressure

The optimal linear relation of Y with respect to X, in accordance with the criterion in Theorem 3.2.9, and for the Lebesgue measure on $[0,1]$ (i.e., $S = m$) is given by

$$Y = .582\, X - 2.096.$$

If the above relation is considered to predict values of Y from values of X, a patient having a range for the systolic blood pressure of 115-161 will have a prediction of 65-92 for his/her range of diastolic blood pressure.

3.3 The f-inequality indices for random sets. Inferential results on finite populations

Let Ω be a population, let $(\Omega, \mathcal{A}, P)$ denote a probability space associated with Ω, and let $X : \Omega \to \mathcal{K}_c((0,+\infty))$ be a positive integrably bounded convex compact random set associated with $(\Omega, \mathcal{A}, P)$. Let $f : (0,+\infty) \to \mathbb{R}$ be a strictly convex and monotonic function belonging to C^1, and satisfying that $f(u) + f(1/u) \geq 0$ for all $u \in (0,+\infty)$ and $f(1) = 0$.

Definition 3.3.1. *The f-inequality index associated with X in the population Ω is given by*

$$\mathrm{I}_f(X) = \frac{1}{2} E\left[f\left(\frac{\inf X}{E(\sup X)} \right) + f\left(\frac{\sup X}{E(\inf X)} \right) \right],$$

whenever this value exists.

Existence conditions and properties in Sections 2.1 and 2.2 could be immediately particularized to indices $\mathrm{I}_f(X)$.

When we need to specify the probability distribution P in the probability space, we will denote $\mathrm{I}_f(X)$ by $\mathrm{I}_f(X \mid P)$.

We are now going to particularize to random sets, the results in Sections 2.3 and 2.4 for the problem of point estimation of the population inequality index in random samplings with and without replacement from finite populations, and for the problem of determining under certain conditions the asymptotic distribution of the sample inequality indices in finite populations and the approximate inferential procedures derived from it.

Let Ω be a finite population of N units, $\omega_1, \ldots, \omega_N$, and let $X : \Omega \to \mathcal{K}_c((0, +\infty))$ be a positive convex compact random set associated with a probability space defined on Ω when it is assumed to be endowed with the uniform distribution.

Assume that a sample of size n is chosen at random and without replacement from Ω, v denotes a generic random sample without replacement of size n, and $\omega_{v1}, \ldots, \omega_{vn}$ are the units in v. Then, the *sample hyperbolic inequality index* of X, which for v is given by

$$\mathrm{I}_H\left(X[v]\right) = \frac{1}{2n^2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \left[\frac{\sup X(\omega_{vi})}{\inf X(\omega_{vi'})} + \frac{\inf X(\omega_{vi})}{\sup X(\omega_{vi'})} \right] - 1,$$

is a real-valued random variable associated with the probability space $(\Upsilon_n, \mathcal{P}(\Upsilon_n), p)$ (see Section 1.3, p. 45), and hence it defines an estimator of the population hyperbolic index, which is given by the value

$$\mathrm{I}_H(X \mid P) = \frac{1}{2N^2} \sum_{j=1}^{N} \sum_{j'=1}^{N} \left[\frac{\sup X(\omega_j)}{\inf X(\omega_{j'})} + \frac{\inf X(\omega_j)}{\sup X(\omega_{j'})} \right] - 1.$$

Furthermore, we have that

Theorem 3.3.1. *In random sampling without replacement of size n from the population Ω, the estimator $\widehat{\mathrm{I}_H}(X[\,\cdot\,])$ such that*

$$\widehat{\mathrm{I}_H}(X[v]) = \frac{n(N-1)}{N(n-1)} \mathrm{I}_H\left(X[v]\right) - \frac{1-f}{n} \left\{ \frac{1}{n-1} \sum_{i=1}^{n} \mathrm{I}_H\left(\{X(\omega_{vi})\}\right) \right\}$$

is unbiased to estimate $\mathrm{I}_H(X \mid P)$, where $\mathrm{I}_H\left(\{X(\omega_{vi})\}\right)$ represents the value of the (intra)hyperbolic inequality index for the convex compact random set degenerate at the interval $X(\omega_{vi}) \in \mathcal{K}_c((0, +\infty))$.

The accuracy of the estimator $\mathrm{I}_H(X)$ above introduced, is discussed in the following result:

Theorem 3.3.2. *In random sampling without replacement of size n from the population* Ω*, we have that*

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}(X[\cdot])\right) = \frac{(1-f)}{4n(n-1)N^2(N-1)(N-2)(N-3)}$$

$$\cdot\Big\{4[N(6-4n)+6(n-1)]N^3(N-1)\left(\mathrm{I}_H(X\mid P)\right)^2$$

$$+4[4N^2(3-2n)+13N(n-1)+3(n-3)]N^2(N-1)\mathrm{I}_H(X\mid P)$$

$$+4[N^2(7-3n)+N(5n-7)-4(n-2)]\left(\sum_{j=1}^{N}\mathrm{I}_H(\{X(\omega_j)\})\right)^2$$

$$+4[N^2(n-5)+N(5n+1)-10(n-2)]\sum_{j=1}^{N}\left(\mathrm{I}_H(\{X(\omega_j)\})\right)^2$$

$$+8[N^3(3-n)+N^2(3n-8)+N(n+9)-10(n-2)]\sum_{j=1}^{N}\mathrm{I}_H(\{X(\omega_j)\})$$

$$+(n-2)(N-1)(N-2)\sum_{j=1}^{N}\left(\sum_{l=1}^{N}\left[\frac{\inf X(\omega_j)}{\sup X(\omega_l)}+\frac{\sup X(\omega_j)}{\inf X(\omega_l)}\right.\right.$$

$$\left.\left.+\frac{\inf X(\omega_l)}{\sup X(\omega_j)}+\frac{\sup X(\omega_l)}{\inf X(\omega_j)}\right]\right)^2$$

$$+(N-n+1)(N-1)(N-3)\sum_{j=1}^{N}\sum_{l=1}^{N}\left[\frac{\inf X(\omega_j)}{\sup X(\omega_l)}+\frac{\sup X(\omega_j)}{\inf X(\omega_l)}\right]^2$$

$$+2(N-n+1)(N-1)(N-3)\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{I}_H(\{X(\omega_j)\cdot X(\omega_l)\})$$

$$-2[N(3n-7)-3(n-3)](N-1)\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{I}_H(\{X(\omega_l)\})$$

$$\cdot\left[\frac{\inf X(\omega_j)}{\sup X(\omega_l)}+\frac{\sup X(\omega_j)}{\inf X(\omega_l)}\right]$$

$$+4[N^5(6-4n)+8N^4n+N^3(n-8)+N^2(9-7n)+9Nn+(23-13n)]\Big\}.$$

If we now consider a random choice with replacement of size n from $\Omega = \{\omega_1, \ldots, \omega_N\}$, υ represents a generic random sample of size n and $\omega_{\upsilon 1}, \ldots, \omega_{\upsilon n}$ are the units in υ, then the *sample hyperbolic inequality index* of X, $\mathrm{I}_H(X[\cdot])$, is a real-valued random variable associated with the probability space $(\Upsilon_n^w, \mathcal{P}(\Upsilon_n^w), p^w)$ (see Section 1.3, p. 51), so that it defines an estimator of the *population hyperbolic index*. Furthermore, we have that

Theorem 3.3.3. *In random sampling with replacement of size n from the population Ω, the estimator $\widehat{\mathrm{I}_H}^w(X[\cdot])$ such that*

$$\widehat{\mathrm{I}_H}^w(X[\upsilon]) = \frac{n}{(n-1)}\mathrm{I}_H(X[\upsilon]) - \frac{1}{n}\left\{\frac{1}{n-1}\sum_{i=1}^{n}\mathrm{I}_H(\{X(\omega_{\upsilon i})\})\right\},$$

is unbiased to estimate $\mathrm{I}_H(X\,|\,P)$.

Moreover, we have that

Theorem 3.3.4. *In random sampling with replacement of size n from the population Ω, we have that*

$$\mathrm{Var}\left(\widehat{\mathrm{I}_H}^w(X[\cdot])\right) = \frac{1}{4n(n-1)N^3}$$

$$\cdot\Big\{4(6-4n)N^3\left(\mathrm{I}_H(X\,|\,P)\right)^2 + 8(6-4n)N^3\mathrm{I}_H(X\,|\,P)$$

$$-4[N-4(n-1)]\left(\sum_{j=1}^{N}\mathrm{I}_H(\{X(\omega_j)\})\right)^2 + 4[N-4(n-1)]\sum_{j=1}^{N}\left(\mathrm{I}_H(\{X(\omega_j)\})\right)^2$$

$$-8[N-4(n-1)](N-1)\sum_{j=1}^{N}\mathrm{I}_H(\{X(\omega_j)\})$$

$$+(n-2)\sum_{j=1}^{N}\left(\sum_{l=1}^{N}\left[\frac{\inf X(\omega_j)}{\sup X(\omega_l)} + \frac{\sup X(\omega_j)}{\inf X(\omega_l)} + \frac{\inf X(\omega_l)}{\sup X(\omega_j)} + \frac{\sup X(\omega_l)}{\inf X(\omega_j)}\right]\right)^2$$

$$+N\sum_{j=1}^{N}\sum_{l=1}^{N}\left[\frac{\inf X(\omega_j)}{\sup X(\omega_l)}+\frac{\sup X(\omega_j)}{\inf X(\omega_l)}\right]^2$$

$$+2N\sum_{j=1}^{N}\sum_{l=1}^{N}\mathrm{I}_H(\{X(\omega_j)\cdot X(\omega_l)\})+4[N^2(5-4n)+N(4n-3)+(5-4n)]N\Bigg\}.$$

On the other hand, the study of the asymptotic distribution of the sample inequality indices will allow us to give an asymptotically unbiased estimate of the population indices under certain conditions, to discuss the asymptotic accuracy of the estimates, and to develop approximations to other inferential techniques for most of the population indices.

Assume that we consider a finite population of size N, $\Omega = \{\omega_1, \ldots, \omega_N\}$, and let X be a positive convex compact random set defined on Ω. If on the population Ω variable X takes on r different values, $x_1^*, \ldots, x_r^*$, and, for $l \in \{1, \ldots, r\}$, p_l denotes the probability $P(\{\omega \in \Omega \mid X(\omega) = x_l^*\})$, and $f : (0, +\infty) \to \mathbb{R}$ is a strictly convex and monotonic function belonging to C^1, and satisfying that $f(u) + f(1/u) \geq 0$ for all $u \in (0, +\infty)$ and $f(1) = 0$, the *population f-inequality index* of X is given by the value

$$\mathrm{I}_f(X\,|\,\mathbf{p}) = F\left(\sum_{l'=1}^{r} p_{l'}\cdot f\left(x_{l'}^* \Big/ \sum_{l=1}^{r} p_l\cdot x_l^*\right)\right)$$

$$= \sum_{l'=1}^{r} p_{l'}F\left(f\left(x_{l'}^* \Big/ \sum_{l=1}^{r} p_l\cdot x_l^*\right)\right).$$

If a sample of size n is chosen at random and from the population and f_{nl} denotes the relative frequency of the value x_l^* of X in this sample, the *sample f-inequality index* of X corresponds to the value

$$\mathrm{I}_f(X\,|\,\mathbf{f}_n) = F\left(\sum_{l'=1}^{r} f_{nl'}\cdot f\left(x_{l'}^* \Big/ \sum_{l=1}^{r} f_{nl}\cdot x_l^*\right)\right)$$

$$= \sum_{l'=1}^{r} f_{nl'}F\left(f\left(x_{l'}^* \Big/ \sum_{l=1}^{r} f_{nl}\cdot x_l^*\right)\right).$$

The asymptotic distribution of the sample f-inequality index is now given as follows:

Theorem 3.3.5. *For each $n \in \mathbb{N}$, consider n independent convex compact random sets being identically distributed as X (that is, a simple random sample of size n from X), defined over the finite population $\Omega = \{\omega_1, \ldots, \omega_N\}$ such that $P(\{\omega \in \Omega \mid X(\omega) = x_l^*\}) = p_l$ with $p_l \in (0,1)$, $l = 1, \ldots, r$, $\sum_{l=1}^{r} p_l = 1$. Let $f : (0, +\infty) \to \mathbb{R}$ be a strictly convex and monotonic function satisfying that $f(u) + f(1/u) \geq 0$ for all $u \in (0, +\infty)$ and $f(1) = 0$, and having finite third derivatives. We have then that*

i) If $\mathbf{f}_n = (f_{n1}, \ldots, f_{n(r-1)}) \in [0,1]^{r-1}$, with f_{nl} = relative frequency of x_l^* in the performance of the simple random sample of size n ($l = 1, \ldots, r-1$, $f_{nr} = 1 - \sum_{l=1}^{r-1} f_{nl}$), and $\mathrm{I}_f(X \mid \mathbf{f}_n)$ is the associated sample inequality index, then $\{\mathrm{I}_f(X \mid \mathbf{f}_n)\}_n$ is a sequence of estimators of $\mathrm{I}_f(X) = \mathrm{I}_f(X \mid \mathbf{p})$, which is strongly consistent, that is, as $n \to \infty$

$$\mathrm{I}_f(X \mid \mathbf{f}_n) \stackrel{a.s.}{\to} \mathrm{I}_f(X \mid \mathbf{p})$$

whatever $\mathbf{p} = (p_1, \ldots, p_{r-1})$ (con $p_1, \ldots, p_{r-1} \in (0,1)$ y $\sum_{l=1}^{r-1} p_l < 1$) may be.

ii) $\{\sqrt{n}\,(\mathrm{I}_f(X \mid \mathbf{f}_n) - \mathrm{I}_f(X \mid \mathbf{p}))\}_n$ is a sequence of random variables converging in law to a one-dimensional normal distribution $N(0, \sigma^2(\mathbf{p}))$, with

$$\sigma^2(\mathbf{p}) = \sum_{l'=1}^{r} p_{l'} (V_{l'}^*)^2 - \left(\sum_{l'=1}^{r} p_{l'} V_{l'}^* \right)^2,$$

where

$$V_{l'}^* = \frac{1}{2} \left[f\left(\frac{\sup x_{l'}^*}{E(\inf X \mid \mathbf{p})} \right) + f\left(\frac{\inf x_{l'}^*}{E(\sup X \mid \mathbf{p})} \right) \right]$$

$$-\frac{1}{2}\sum_{l=1}^{r} p_l \Bigg[\inf x_{l'}^{*} \cdot f'\left(\frac{\sup x_l^{*}}{E\,(\inf X \,|\, \mathbf{p})}\right) \cdot \left(\frac{\sup x_l^{*}}{[E\,(\inf X \,|\, \mathbf{p})]^2}\right) + \sup x_{l'}^{*} \cdot f'\left(\frac{\inf x_l^{*}}{E\,(\sup X \,|\, \mathbf{p})}\right) \cdot \left(\frac{\inf x_l^{*}}{[E\,(\sup X \,|\, \mathbf{p})]^2}\right) \Bigg],$$

whenever $\sigma^2(\mathbf{p}) > 0$.

iii) If $\sigma^2(\mathbf{p}) = 0$, *for some pair* (i,j), *with* $i,j \in \{1,\ldots,r-1\}$, *we have that*

$$h_{ij} = \frac{\partial^2 \mathrm{I}_f(X \,|\, \mathbf{p})}{\partial p_i \partial p_j} = \frac{\partial}{\partial p_i}[V_j^{*} - V_r^{*}] > 0,$$

then $\{2n\,(\mathrm{I}_f(X \,|\, \mathbf{f}_n) - \mathrm{I}_f(X \,|\, \mathbf{p}))\}_n$ *is a sequence of real-valued random variables converging in law as* $n \to \infty$ *to a linear combination of, at most,* $r-1$ *chi-squared* χ_1^2 *independent random variables.*

Furthermore, we have that

Theorem 3.3.6. *Under the conditions in Theorem* 3.3.5*, we have that*

$$\left\{ \frac{\sqrt{n}\,(\mathrm{I}_f(X \,|\, \mathbf{f}_n) - \mathrm{I}_f(X \,|\, \mathbf{p}))}{\sqrt{\sigma^2(\mathbf{f}_n)}} \right\}_n$$

converges in law to a normal distribution $N(0,1)$ *as* $n \to \infty$, *whenever* $\sigma^2(\mathbf{p}) > 0$ *and* $\sigma^2(\mathbf{f}_n) > 0$.

The approximate inferential procedures derived from the last two results are the following:

Theorem 3.3.7. *Under the conditions in Theorems* 3.3.5 *and* 3.3.6*, the random interval*

$$\left[\mathrm{I}_f(X \,|\, \mathbf{f}_n) - z_\alpha \sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}},\ \mathrm{I}_f(X \,|\, \mathbf{f}_n) + z_\alpha \sqrt{\frac{\sigma^2(\mathbf{f}_n)}{n}} \right]$$

supplies for each sample of n *independent observations from* X *confidence intervals of* $\mathrm{I}_f(X \,|\, \mathbf{p})$ *with coefficient approximately equal to* $1-\alpha$ $(\alpha \in [0,1])$.

Theorem 3.3.8. *Under the conditions in Theorems* 3.3.5 *and* 3.3.6*:*

(i) *To test at the significance level* $\alpha \in [0,1]$ *the null hypothesis*

$$H_0 : \mathrm{I}_f(X \mid \mathbf{p}) = i_0$$

against the alternative

$$H_1 : \mathrm{I}_f(X \mid \mathbf{p}) \neq i_0,$$

H_0 *must be rejected whenever* $|\mathrm{I}_f(X \mid \mathbf{f}_n) - i_0| > z_\alpha \sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$. *The p-value of this test will be given, approximately, by*

$$p = 2\left[1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\,|\mathrm{I}_f(X \mid \mathbf{f}_n) - i_0|\right)\right].$$

(ii) *To test at the significance level* $\alpha \in [0,1]$ *the null hypothesis*

$$H_0 : \mathrm{I}_f(X \mid \mathbf{p}) \geq i_0$$

against the alternative

$$H_1 : \mathrm{I}_f(X \mid \mathbf{p}) < i_0,$$

H_0 *must be rejected whenever* $\mathrm{I}_f(X \mid \mathbf{f}_n) - i_0 < z_{2\alpha} \sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$. *The p-value of this test will be given, approximately, by*

$$p = \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\,(\mathrm{I}_f(X \mid \mathbf{f}_n) - i_0)\right).$$

(iii) *To test at the significance level* $\alpha \in [0,1]$ *the null hypothesis*

$$H_0 : \mathrm{I}_f(X \mid \mathbf{p}) \leq i_0$$

against the alternative

$$H_1 : \mathrm{I}_f(X \mid \mathbf{p}) > i_0,$$

H_0 must be rejected whenever $\mathrm{I}_f(X \mid \mathbf{f}_n) - i_0 > z_{2\alpha}\sqrt{\dfrac{\sigma^2(\mathbf{f}_n)}{n}}$. *The p-value of this test will be given, approximately, by*

$$p = 1 - \Phi\left(\sqrt{\frac{n}{\sigma^2(\mathbf{f}_n)}}\,(\mathrm{I}_f(X \mid \mathbf{f}_n) - i_0)\right).$$

As an application of the preceding studies we consider the following example.

Example 3.3.1. A big firm is interested in estimating the inequality of the "monthly salary after deductions" of their employees during a given year. However, and due to the work developed in this firm, the net salary for each employee varies from month to month in the same year, so we can actually talk about the "range of monthly net salaries" of employees, which is an interval-valued random element X whose values can obviously overlap.

Assume that in order to estimate $\mathrm{I}_H(X)$ in the population of all employees of this firm we consider a sample $[\tau]$ of 45 employees, chosen at random and with replacement from the population, and that these employees are requested to indicate their "range of the monthly net salary" in the given year, their answers being those gathered in Table 3.3.

Then, on the basis of Theorem 3.3.3 the estimate $\widehat{\mathrm{I}_H}^w(X[\tau]) = .105$ corresponds to the sample value of an unbiased estimator of $\mathrm{I}_H(X)$.

3.4 Final discussion and future directions

To justify the approach and bases used in quantifying the variation of a random set in this chapter, we must adduce as a main argument that this has been due to the assumption that the involved variable is essentially interval-valued.

The studies developed in this chapter would be applicable in principle to the case in which we consider the observation of an existing real-valued

325000-335000	267500-275000	515000-530000
125000-130000	185000-200000	240000-255000
405000-420000	400000-410000	190000-200000
200000-210000	370000-385000	440000-457500
260000-267500	210000-217500	245000-257500
185000-200000	222500-230000	215000-225000
347500-352500	410000-422500	310000-317500
295000-310000	150000-160000	265000-280000
415000-425000	272500-277500	295000-305000
255000-266000	360000-370000	280000-290000
310000-317500	225000-235000	420000-432500
190000-200000	167500-180000	197500-205000
410000-422500	280000-290000	255000-266000
205000-215000	267500-275000	370000-385000
330000-340000	430000-445000	237500-250000

Table 3.3: Ranges of net salaries per month

random variable, but for several reasons (either because of imprecision in the mechanisms of measurement/observation/report, or simply because of the tradition or because of its interest in seeking for suitable operational models for the distributions of the original real-valued random variable), data are grouped in intervals. These grouped data could be regarded as values of a convex compact random set whose values must be mutually exclusive.

It would be valuable to verify by means of different procedures the "goodness" of the measures Δ_S^2 and I_f in approximating the exact true value of the variance and f-inequality of the underlying original real-valued random variable, respectively. This discussion could be developed for Δ_S^2 in terms of the choice of measure S, the distribution of the original random variable and the considered grouping, and for I_f in terms of the choice of function f, the distribution of the original variable and the considered grouping.

However, it should be taken into account that Δ_S^2 and I_f, as defined in Sections 3.1 and 3.3, respectively, actually measure the variation of a random element whose values are the grouped data in all their range, instead of that of a real-valued random variable for which each of the values has been identified with an interval containing it.

Concerning the open problems in connection with the study in the present chapter, we can consider those particularizing the ones suggested in Sections 1.7 and 2.5. In fact, some of these problems will be first analyzed for random sets, since it is usually a less complex situation and it is easier to interpret from a geometrical viewpoint (as it will happen, for instance, in the studies on the correlation or the regression with nonlinear functions). Furthermore, in some of these problems, the study carried out for random sets could be directly applied to the α-level functions of the fuzzy random variables, and this will allow us to draw immediate conclusions for the latter ones.

Appendix A

About the metric on which the generalized absolute variation measure is based

As we remarked in the presentation of the metric d_∞, Klement *et al.* (1986) have proved that the space $(\mathcal{F}_c(\mathbb{R}), d_\infty)$ is nonseparable. The theoretical interest is mainly due to its nonseparability, but also most of the difficulties which arise in some probabilistic studies about fuzzy random variables are due to the nonseparability.

Nevertheless, when we consider the metric D_S on $\mathcal{F}_c(\mathbb{R})$ (and, in a similar way, when we consider the metric d_S on $\mathcal{K}_c(\mathbb{R})$), the metric space is separable.

Proposition A.1. $(\mathcal{F}_c(\mathbb{R}), D_S)$ is a metric space which is equivalent to $(\mathcal{F}_c(\mathbb{R}),\, d_2)$, where $d_2 : \mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \to [0, +\infty)$ is the metric such that for $\widetilde{V}$, $\widetilde{W} \in \mathcal{F}_c(\mathbb{R})$

$$d_2(\widetilde{V}, \widetilde{W}) = \sqrt{\int_{(0,1]} \left[d_H(\widetilde{V}_\alpha, \widetilde{W}_\alpha)\right]^2 d\alpha}.$$

Proof.

To prove that D_S and d_2 are equivalent metrics we have to take into account that for $\widetilde{V}, \widetilde{W} \in \mathcal{F}_c(R)$ we have that

$$D_S(\widetilde{V}, \widetilde{W}) = \sqrt{\int_{(0,1]} \left[d_S\left(\widetilde{V}_\alpha, \widetilde{W}_\alpha\right)\right]^2 d\alpha}$$

$$= \sqrt{\int_{(0,1]} \int_{[0,1]} \left[\lambda(\sup \widetilde{V}_\alpha - \sup \widetilde{W}_\alpha) + (1-\lambda)(\inf \widetilde{V}_\alpha - \inf \widetilde{W}_\alpha)\right]^2 dS(\lambda)\, d\alpha}$$

$$\leq \sqrt{\int_{(0,1]} \left[\max\{|\sup \widetilde{V}_\alpha - \sup \widetilde{W}_\alpha|, |\inf \widetilde{V}_\alpha - \inf \widetilde{W}_\alpha|\}\right]^2 d\alpha} = d_2(\widetilde{V}, \widetilde{W}).$$

On the other hand,

$$d_2(\widetilde{V}, \widetilde{W}) = \sqrt{\int_{(0,1]} \left[\max\left\{|\sup \widetilde{V}_\alpha - \sup \widetilde{W}_\alpha|, |\inf \widetilde{V}_\alpha - \inf \widetilde{W}_\alpha|\right\}\right]^2 d\alpha}$$

$$\leq \sqrt{\int_{(0,1]} \left[(\sup \widetilde{V}_\alpha - \sup \widetilde{W}_\alpha)^2 + (\inf \widetilde{V}_\alpha - \inf \widetilde{W}_\alpha)^2\right] d\alpha} = \sqrt{2}\, D_{S_2}(\widetilde{V}, \widetilde{W}),$$

where S_2 is the weight measure (see Introductory Chapter) such that the associated density g satisfies that $g(\lambda) = 0$ for all $\lambda \in (0,1)$ and $g(0) = g(1) = .5$.

In accordance with Bertoluzza *et al.* (1995), all metrics D_S are topologically equivalent, whence the assertion in the present proposition is easily concluded. □

Corollary A.2. $(\mathcal{F}_c(\mathbb{R}), D_S)$ is a separable metric space.

Proof.

Obviously, the separability of $(\mathcal{F}_c(\mathbb{R}), d_2)$ (see, for instance, Diamond & Kloeden 1994) entails that of $(\mathcal{F}_c(\mathbb{R}), D_S)$. □

Appendix B

About the function on which the generalized relative variation measure is based

In the study of the extension of the f-inequality indices to fuzzy random variables and convex compact random sets, it is supposed that the function f satisfies some additional conditions (in comparison with the real-valued case). The constraints imposed by the new conditions (and, especially, by the ones assumed in the analysis of the minimality properties of the inequality measures), allow us to establish certain characterizations of the corresponding functions f.

The general aim of this section is to look for all the functions $f : (0, +\infty) \to \mathbb{R}$ being strictly convex and monotonic, $f \in C^1$, $f(1) = 0$, and such that $f(u) + f(1/u) \geq 0$ for all $u \in (0, +\infty)$.

First of all, we are going to look for the subfamily of the twice-differentiable functions f satisfying that $f(u) + f(1/u) = 0$ for all $u \in (0, +\infty)$. This last condition will be considered in Theorem 2.2.6 in Chapter 2.

Proposition B.1. The set of the solutions of the functional equation $f(u) + f(1/u) = 0$ for all $u \in (0, +\infty)$ with $f : (0, +\infty) \to \mathbb{R}$ twice-differentiable, strictly convex and monotonic and satisfying that $f(1) = 0$, is given by the family of functions $f(u) = h(\log u)$ with $h : \mathbb{R} \to \mathbb{R}$ such that for all $x \in \mathbb{R}$ we have that *(i)* $h(x) = -h(-x)$,

(ii) h (strictly) decreasing, and

(iii) $h''(x) \geq h'(x)$.

Proof.

The transformations $x = \log u$ and $h(x) = f(e^x)$ allows us to conclude that $f(u) + f(1/u) = 0$ for all $u \in (0, +\infty)$ if, and only if, we have that $f(u) = h(\log u)$ for all $u \in (0, +\infty)$ and $h(x) = -h(-x)$ for all $x \in \mathbb{R}$. Furthermore, since $h(0) = 0$, we then ensure that $f(1) = 0$.

On the other hand, to prove that condition *(ii)* must hold we should take into account that because of the (strict) increasing of the function g such that $g(x) = \log x$, then $f = h \circ g$ is strictly monotonic if, and only if, h is also strictly monotonic and f and h are both increasing or decreasing.

In this case, h must be decreasing since otherwise the convexity of f would entail that all $u \in (0, +\infty)$:

$$0 \leq f''(u) = \frac{1}{u^2} \left(h''[\log u] - h'[\log u] \right),$$

whence $h''(\log u) \leq 0$, so that h should be convex. If h should be convex, increasing and satisfying that $h(x) = -h(-x)$ for all $x \in \mathbb{R}$, h has to be defined as $h(x) = ax$ with $a > 0$, and hence $f(u) = a \log u$ for all $u \in (0, +\infty)$ with $a > 0$ could not be convex.

Consequently, the differentiability of f guarantees that of h, and therefore $h'(x) \leq 0$ for all $x \in \mathbb{R}$.

Finally, since f is twice-differentiable, h must be also twice-differentiable, and in virtue of the convexity of f we will have that

$$0 \leq f''(u) = \frac{1}{u^2} \left(h''[\log u] - h'[\log u] \right),$$

so that we can conclude that $h''(\log u) \geq h'(\log u)$ for all $u \in (0, +\infty)$ must be satisfied. $\square$

Remark B.1. It can be pointed out that the conditions *(i)* and *(ii)* are necessary even if f is not assumed to be differentiable. On the other hand, the Shannon type index (corresponding to $h(x) = -x$) is obviously associated with a function f being a solution of the functional equation in Proposition B.1.

Regarding the solutions of the subfamily of the twice-differentiable functions f satisfying that $f(u) + f(1/u) \geq 0$ for all $u \in (0, +\infty)$ with equality if, and only if, $u = 1$, (that is, $f(u) + f(1/u) = 0$ if $u = 1$, > 0 otherwise), we can obtain by following arguments similar to those in Proposition B.1 (to which we will refer in Theorem 2.2.5 of Chapter 2) that

Proposition B.2. The set of the functions $f : (0, +\infty) \to \mathbb{R}$ being twice-differentiable, strictly convex and monotonic and satisfying that $f(u) + f(1/u) > 0$ for all $u \in (0, +\infty) \setminus \{1\}$ and $f(1) = 0$, is the family of functions $f(u) = h(\log u)$ with $h : \mathbb{R} \to \mathbb{R}$ such that

(i) $h(x) > -h(-x)$ for all $x \in \mathbb{R} \setminus \{0\}$ and $h(0) = 0$,

(ii) h strictly monotonic, and

(iii) $h''(x) \geq h'(x)$ for all $x \in \mathbb{R}$.

The extension of the additive decomposable indices for $\alpha \neq 0, 1$ (corresponding to $h(x) = e^{\alpha x} - 1$ for $\alpha < 0$ and $\alpha > 1$ and to $h(x) = 1 - e^{\alpha x}$ for $\alpha \in (0, 1)$) is associated with a function f being a solution of the functional equation in Proposition B.2.

If we want now to find the solutions of the general functional inequation $f(u)+f(1/u) \geq 0$ for all $u \in (0, +\infty)$ with $f : (0, +\infty) \to \mathbb{R}$ being twice-differentiable, strictly convex and monotonic and satisfying that $f(1) = 0$, the transformations in the proof of Proposition B.1 indicate that $h(x) \geq -h(-x)$ and $h(0) = 0$.

Since h is strictly monotonic, then to determine functions h satisfying the latter inequality we can reason as follows:

Let $h_1 : [0, +\infty) \to \mathbb{R}$ and $h_2 : (-\infty, 0] \to \mathbb{R}$ be (strictly) decreasing functions such that $h_1(0) = h_2(0) = 0$ and $h_2(x) \geq -h_1(-x)$. Consider $h : \mathbb{R} \to \mathbb{R}$, defined so that

$$h(x) = \begin{cases} h_1(x) & \text{if } x \geq 0 \\ h_2(x) & \text{otherwise.} \end{cases}$$

Function h is clearly (strictly) decreasing and satisfies that $h(x) \geq -h(-x)$ and $h(0) = 0$. Therefore, the function $f(u) = h(\log u)$ is a solution for the considered problem, whenever f is convex. On the other hand, if f is assumed to be twice-differentiable, then $f''(u) \geq 0$ if, and only if, $h'' \geq h'$, and this happens if, and only if,

$$\begin{cases} h_1'' \geq h_1' & \text{if } x \geq 0 \\ h_2'' \geq h_2' & \text{otherwise,} \end{cases}$$

the function $f(u) = h(\log u)$, where h is constructed in accordance with the suggested procedure, is a solution of the considered problem if, and only if, $h_1'' \geq h_1'$ and $h_2'' \geq h_2'$.

It should be emphasized that, since $h(0) = 0$ and h must be strictly monotonic, then either $h(x) \geq 0$ for all $x \in (-\infty, 0]$ and $h(x) \leq 0$ for all $x \in [0, +\infty)$ (which is the considered case), or $h(x) \leq 0$ for all $x \in (-\infty, 0]$ and $h(x) \geq 0$ for all $x \in [0, +\infty)$. In this latter case, h would be increasing, but since h must hold that $h(x) \geq -h(-x)$, if as in the preceding case one constructs (strictly) increasing functions $h_1 : [0, +\infty) \to \mathbb{R}$ and $h_2 : (-\infty, 0] \to \mathbb{R}$ such that $h_1(0) = h_2(0) = 0$ and $h_1(x) \geq -h_2(-x)$, then the function $h : \mathbb{R} \to \mathbb{R}$, defined so that

$$h(x) = \begin{cases} h_1(x) & \text{if } x \geq 0 \\ h_2(x) & \text{otherwise,} \end{cases}$$

is clearly increasing and satisfies that $h(x) \geq -h(-x)$ and $h(0) = 0$.

Consequently, the function $f(u) = h(\log u)$ is a solution of the considered problem, whenever f is convex. But since h is increasing and f is convex, then h should be convex. In this way we are constrained (under the construction we have just suggested) to the functions

$$h(x) = \begin{cases} h_1(x) & \text{if } x \geq 0 \\ h_2(x) & \text{otherwise,} \end{cases}$$

being increasing, convex and such that $f(u) = h(\log u)$ is convex. As in the preceding case, if f is twice-differentiable, the function $f(u) = h(\log u)$ is a solution of the considered problem if, and only if, $h_1'' \geq h_1'$ and $h_2'' \geq h_2'$.

Finally, it should be remarked that there exist functions f satisfying the general conditions assumed in Chapters 2 and 3 for the extended inequality indices, although they do not fit the models in Propositions B.1 and B.2. In this way, the function $f(x) = 1/ex$ if $x \in (0, 1/e]$, $= -\log x$ if $x \in (1/e, +\infty)$, is such that $f \in C^1$, f is strictly convex and monotonic, satisfies that $f(1) = 0$, and furthermore $f(x) + f(1/x) = 0$ if $x \in [1/e, e]$, > 0 otherwise.

Epilogue

In this work we have developed a study on the quantification of the variation associated with a random element whose values are not real numbers, but either compact real intervals or fuzzy numbers with compact closure of the support.

The above quantification has been carried out in a numerical scale, since it has been assumed that the main interest is performing comparisons between variables, populations, estimators, etc. on the basis of the associated variation.

The common scheme in both the measurement of the absolute variation (or deviation) and the measurement of the relative variation (or inequality), has been using the average deviation and average inequality, respectively, of each value taken on by the random element with respect to its expected value.

In this way, the generalized measure of the absolute variation introduced in the present work has been defined as the (classical) expected value of the S-distance between each value from the fuzzy random variable or the random set and its expected value (in Puri & Ralescu's sense -1986-, or Aumann's sense -1965-, respectively). The generalization qualifying the measure, refers to the possibility of choosing among different measures S satisfying quite general conditions.

On the other hand, the generalized measure of the relative variation introduced in this work has been defined as the value that the ranking function F by Yager in Definition 0.1.3 assigns to the expected value of the (fuzzy) image induced from f, of the (fuzzy or interval) "quotient" of each value of the random element and its expected value. In this case, the generalization in the name of the measure is due to the possibility of choosing among different functions f satisfying certain conditions.

In view of this common scheme, the question which immediately arises is why we have introduced new measures to the considered purpose, instead of using a prior real-valued "code" (or defuzzification process) to the available imprecise data, and later handling the coded data by means of well-known variation measures of the real-valued case.

The reason for not having used a "codifier" function converting each imprecise (fuzzy or interval) value into a real number, is that the most natural codifier functions would lead to identifying different values of the random elements. Consequently, either the Euclidean distance, or the image by f of the quotient, between the converted values could be equal to 0, whereas the original imprecise values would have associated really a nonnull variation, the code would make us ignore.

As an example for the last assertion, if we admit Yager's function $F(\,\cdot\,)$ to code the values of a fuzzy random variable $\mathcal{X}$ taking on values $\tilde{x}_1 = \mathrm{Tri}(-1, 0, 1)$, $\tilde{x}_2 = \mathrm{Tri}(-2.000, 0, 2.000)$ and $\tilde{x}_3 = \mathrm{Tri}(-3, 1, 1)$, with induced probabilities 1/3, 1/3 and 1/3, we would conclude that $F \circ \mathcal{X}$ would be a real-valued random variable degenerate at the value 0, so that $\mathrm{Var}(F \circ \mathcal{X}) = 0$. In an analogous way, if we admit the function $f_{(\,\cdot\,)}(\lambda) : \mathcal{K}_c(\mathbb{R}) \to \mathbb{R}$ (see Definitions 0.1.5 and 3.1.1) to code the values of a convex compact random set, for an arbitrary fixed $\lambda \in [0, 1]$, we can achieve similar conclusions.

The preceding comments do not necessarily prevail when we deal with the special case of grouped data, since in such a situation two different (though not exclusive) intervals in two different groupings can correspond to the same set of real-valued data, so that the error incurred in identifying them by the "codifier" function (say, coding each interval in the grouping by means of its mid-point) can be lower than that incurred when each grouped data is treated as an interval value in the way proposed in Chapter 3.

In addition to the specific open problems indicated in each chapter, there are some general problems which are subsequently shortly summarized:

- Generalizing the studies for random elements with values in $\mathcal{F}(\mathbb{R})$ and $\mathcal{K}(\mathbb{R})$, with $\mathcal{K}(\mathbb{R})$ the class of the nonempty compact sets of $\mathbb{R}$, and $\mathcal{F}(\mathbb{R})$ the class of the fuzzy subsets $\widetilde{V}$ of $\mathbb{R}$, such that $\widetilde{V}_\alpha \in \mathcal{K}(\mathbb{R})$ for all $\alpha \in [0, 1]$.

- Generalizing the studies for random elements whose values are not necessarily compact (or necessarily having compact α-level sets).

- Developing asymptotic studies on the Linear Regression problem to apply them to testing hypotheses on the slope and intercept parameters of the linear relation.

- Developing studies on linear models involving more than two random elements (namely, Multiple Regression, Analysis of "Variance").

References and Consulted Bibliography

Adamo, J.M. (1980) Fuzzy decision trees. *Fuzzy Sets and Systems* **4**, 207-219.

Alonso, M.C. and Gil, M.A. (1995). Testing the convenience of a variate for stratification in estimating the Gini-Simpson diversity. In *Lecture Notes in Computer Sciences* **945**, Advances in Intelligent Computing (B. Bouchon and R. R. Yager, Eds.), 210-216. Springer-Verlag, Berlin.

Alonso, M.C., Bertoluzza, C., Brezmes, T. and Lubiano, M.A. (1998). *A generalized real-valued measure of the inequality associated with a fuzzy random variable.* (submitted).

Ash, R. (1972). *Real Analysis and Probability.* Academic Press, New York.

Atkinson, A.B. (1983). *The Economics of Inequality.* 2nd ed. Oxford University Press, New York.

Aumann, R.J. (1965). Integrals of set-valued functions. *J. Math. Anal. Appl.* **12**, 1-12.

Bandemer, H. and Gottwald, S. (1995). *Fuzzy Sets, Fuzzy Logic, Fuzzy Methods with Applications.* John Wiley & Sons, Chichester.

Bandemer, H. and Näther, W. (1992). *Fuzzy Data Analysis*. Kluwer Academic Publishers, Dortrecht.

Bárdossy, A. (1990). Note on fuzzy regression. *Fuzzy Sets and Systems* **37**, 65-75.

Barnett, V. (1991). *Sample Survey. Principles and Methods*. Edward Arnold, London.

Basu, K. (1987). Axioms for a fuzzy measure of inequality. *Math. Soc. Sci.* **14**, 275-288.

Bertoluzza, C., Corral, N. and Salas, A. (1995). On a new class of distances between fuzzy numbers. *Mathware & Soft Computing* **2**, 71-84.

Bertoluzza, C., Corral, N. and Salas, A. (1995). Polinomial regression in a fuzzy context. The least squares method, *Proc. 6th IFSA Conference*, 431-434.

Billingsley, P. (1995). *Probability and Measure* 3rd ed. John Wiley & Sons, New York.

Blackorby, C. and Donaldson, D. (1978). Measures of relative equality and their meaning in terms of social welfare. *J. Econom. Theory* **18**, 59-80.

Bortolan, G. and Degani, P. (1985). A review of some methods for ranking fuzzy subsets. *Fuzzy Sets and Systems* **15**, 1-19.

Bourguignon, F. (1979). Decomposable income inequality measures. *Econometrica* **47**, 901-920.

Breiman, L. (1968). *Probability*. Addison-Wesley, Massachusetts.

Byrne, C. (1978). Remarks on the set-valued integrals of Debreu and Aumann. *J. Math. Anal. Appl.* **62**, 243-246.

Campos, L.M. de and González, A. (1989). A subjective approach for ranking fuzzy numbers. *Fuzzy Sets and Systems* **29**, 145-153.

Casella, G. and Berger, R.L. (1990). *Statistical Inference*. Wadsworth & Brooks/Cole, Pacific Grove.

Celminš, A. (1987). Least squares model fitting to fuzzy vector data. *Fuzzy Sets and Systems* **22**, 245-269.

Cochran, W.G. (1977). *Sampling Techniques*. John Wiley & Sons, New York.

Colubi, A. (1997). *Los f-índices de desigualdad difusos asociados a una variable aleatoria difusa*. Trabajo de Investigación de Tercer ciclo. Universidad de Oviedo.

Colubi, A., Gil, M.A. and López-García, H. (1997). Measuring the inequality associated with a fuzzy random variable in terms of statistical f-divergences. *Proc. EUFIT'97* **3**, 1764-1768.

Colubi, A., López-Díaz, M., Domínguez-Menchero, J.S. and Gil, M.A. (1999). A generalized Strong Law of Large Numbers. *Probab. Theory Relat. Fields* (in press).

Corral, N., Gil, M.A., López, M.T., Salas, A. and Bertoluzza, C. (1998). Statistical models. In *Handbook of Fuzzy Computation* Section C: Modeling and Simulation C2.3 (E. Ruspini, P. Bonissone and W. Pedrycz, Eds.), C2.3:1 - C2.3:18. Institute of Physics & Oxford University Press, Bristol.

Cowell, F.A. (1980). On the structure of additive inequality measures. *Rev. Econ. Stud.* **47**, 521-531.

Cowell, F.A. and Kuga, K. (1981). Additivity and the entropy concept. An axiomatic approach. *J. Econom. Theory* **25**, 131-143.

Cowell, F.A. and Mehta, F. (1982). The estimation and interpolation of inequality measures. *Rev. Econ. Stud.* **49**, 273-290.

Cox, E. (1994). *The Fuzzy Systems Handbook.* Academic Press, Cambridge.

Csiszár, I. (1967). Information-type measures of difference of probability distributions and indirect observations. *Studia Scient. Math. Hung.* **2**, 299-318.

Dasgupta, P., Starrett, D. and Sen, A. (1973). Notes on the measurement of inequality. *J. Econom. Theory* **6**, 180-187.

Debreu, G. (1967). Integration of correspondences. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* 1965/66 **2**, Part 1. Univ. of California Press, Berkeley, 351-372.

Delgado, M., Verdegay, J.L. and Vila, M.A. (1988). A procedure for ranking fuzzy numbers using fuzzy relations. *Fuzzy Sets and Systems* **26**, 49-62.

Diamond, P. and Tanaka, H. (1998). Fuzzy regression analysis. In *Fuzzy Sets in Decision Analysis, Operations Research and Statistics* Ch. 11 (R. Słowiński, Eds.), 349-387. Kluwer Acad. Pub., Norwell.

Diamond, P. and Kloeden, P. (1990). Metric spaces of fuzzy sets. *Fuzzy Sets and Systems* **35**, 241-249.

Diamond, P. and Kloeden, P. (1994). *Metric Spaces of Fuzzy Sets: Theory and Applications.* World Scientific, Singapore.

Diamond, P. and Körner, R. (1997). Extended fuzzy linear models and least squares estimates. *Comput. Math. Applic.* **33**, 15-32.

Diamond, P. (1987). Least squares fitting of several fuzzy variables. In *Analysis of Fuzzy Information,* (J.C. Bezdek, Eds.), 329-331. CRC Press, Tokyo.

Diamond, P. (1988). Fuzzy least squares. *Inform. Sci.* **46**, 141-157.

Diamond, P. (1990). Least Squares Fitting of Compact Set-Valued Data. *J. Math. Anal. Appl.* **147**, 531-544.

Dubois, D. and Prade, H. (1978). Operations on fuzzy numbers. *Int. J. Syst. Sci.* **9**, 613-626.

Dubois, D. and Prade, H. (1980). *Fuzzy Sets and Systems: Theory and applications*. Academic Press, New York.

Dubois, D. and Prade, H. (1987). The mean value of a fuzzy number. *Fuzzy Sets and Systems* **24**, 279-300.

Dudewicz, E.J. and Mishra, S.N. (1988). *Modern Mathematical Statistics*. John Wiley & Sons, New York.

Ebert, U. (1984). On a family of inequality measures. Discussion paper No. 146, University of Bonn.

Eichhorn, W. and Gehrig, W. (1982). Measurement of inequality in Economics. In *Applied Mathematics-Optimization and Operations Research* (B. Korte, Eds.), 657-693. North-Holland, Amsterdam.

Gastwirth, J.L. (1975). The estimation of a family of measures of income inequality. *Journal of Econometrics* **3**, 61-70.

Gastwirth, J.L., Nayak, T.K. and Krieger, A.M. (1986). Large sample theory for the bounds on the Gini and related indices of inequality estimated from grouped data. *J. Bus. Econ. Statist.* **4**, 269-273.

Gebhardt, J., Gil, M.A. and Kruse, R. (1998). Fuzzy set-theoretic methods in Statistics. In *Fuzzy Sets in Decision Analysis, Operations Research and Statistics* Ch. 10 (R. Słowiński, Eds.), 311-347. Kluwer Acad. Pub., Norwell.

Gil, M.A. (1979). *Incertidumbre y utilidad*. PhD Thesis. Universidad de Oviedo.

Gil, M.A. (1981). Estudio de una medida para la incertidumbre correspondiente a las utilidades. *Trab. de Est. e Inv. Oper.* **32**, 45-66.

Gil, M.A. (1982). On two unquietness measures for the finite experiences. *Proc. 2nd World Conf. on Math. at the Serv. of Man*, 299-304.

Gil, M.A. and Gil, P. (1989). On some information measures of degree $\beta = 2$. Estimation in simple-stage cluster sampling. *Statist. Probab. Lett.* **8**, 157-162.

Gil, M.A., Caso, C. and Gil, P. (1989a). Estudio asintótico de una clase de índices de desigualdad muestrales. *Trab. de Est.* **4**, 95-109.

Gil, M.A., Pérez, R. and Gil, P. (1989b). A family of measures of uncertainty involving utilies: definitions, properties, applications and statistical inferences. *Metrika* **36**, 129-147.

Gil, M.A. and López-Díaz, M. (1996). Fundamentals and Bayesian analyses of decision problems with fuzzy-valued utilities. *Int. J. Approx. Reason.* **15**, 203-224.

Gil, M.A., López-Díaz, M. and López-García, H. (1998a). The fuzzy hyperbolic inequality index associated with fuzzy random variables. *European J. Oper. Res.* **110**, 377-391.

Gil, M.A., López-Díaz, M. and Rodríguez-Muñiz, L.J. (1998b). An improvement of a comparison of experiments in statistical decision problems with fuzzy utilities. *IEEE Trans. Systems, Man, Cybern.* **28A**, 856-864.

Gil, M.A., Gil, P. and Ralescu, D.A. (1999). Fuzzy random variables: modeling linguistic statistical data. In *Computing with Words in Information/Intelligent Systems* (J. Kacprzyk and L.A. Zadeh, Eds.), Springer-Verlag, Berlin (in press).

González, A. (1990). A study of the ranking function approach through mean values. *Fuzzy Sets and Systems* **35**, 29-41.

González, A. and Vila, M.A. (1992). Dominance relations on fuzzy numbers. *Inform. Sci.* **64**, 1-16.

Guo, S. and Chen, S. (1992). An approach to monodic fuzzy regression. In *Fuzzy Regression Analysis* (J. Kacprzyk and M. Fedrizzi, Eds.), 81-90. Physica-Verlag, Heidelberg.

Kacprzyk, J. and Fedrizzi, M. (1992). *Fuzzy Regression Analysis.* Physica-Verlag, Heidelberg.

Kaleva, O. (1994). Interpolation of fuzzy data. *Fuzzy Sets and Systems* **61**, 63-70.

Kaufmann, A. and Gupta, M.M. (1991). *Introduction to Fuzzy Arithmetic.* Van Nostrand Reinhold Co., New York.

Kendall, D.G. (1974). Foundations of a theory of random sets. In *Stochastic Geometry* (E.F. Harding and D.G. Kendall Eds.), 3-9. John Wiley & Sons, New York.

Hiai, F. & Umegaki, H. (1977). Integrals, conditional expectations and martingales of multivalued functions. *J. Multivar. Anal.* **7**, 149-182.

Himmelberg, C.J. (1975). Measurable relations. *Fund. Math.* **87**, 53-72.

Klement, E.P., Puri M.L. and Ralescu, D.A. (1986). Limit theorems for fuzzy random variables. *Proc. R. Soc. Lond. A* **407**, 171-182.

Klir, G.J. and Yuan, B. (1995). *Fuzzy sets and Fuzzy Logic. Theory and Applications.* Prentice Hall, New Jersey.

Kölm, S. Ch. (1976). Unequal inequalities I. *J. Econom. Theory* **12**, 416-442.

Kölm, S. Ch. (1976). Unequal inequalities II. *J. Econom. Theory* **13**, 82-111.

Kołodziejczyk, W. (1986). Orlovsky's concept of Decision-Making with fuzzy preference relation-further results. *Fuzzy Sets and Systems* **19**, 11-20.

Körner, R. (1997a). On the variance of fuzzy random variables. *Fuzzy Sets and Systems* **92**, 83-93.

Körner, R. (1997b). *Linear Models with Random Fuzzy Variables.* PhD Thesis. Freiberg University of Mining and Technology.

Körner, R. and Näther, W. (1998). Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates. *Inform. Sci.* **109**, 95-118.

Körner, R. and Lubiano, M.A. (1998). A generalized real-valued measure of dispersion for fuzzy random variables. Tech. Rep. Universidad de Oviedo.

Kruse, R. (1987). On the variance of random sets. *J. Math. Anal. Appl.* **122**, 469-473.

Kruse, R., Gebhardt, J. and Gil, M.A. (1999). Fuzzy Statistics. In *Wiley Encyclopedia of Electrical and Electronics Engineering* (John G. Webster, Ed.), John Wiley & Sons, New York (in press).

Kruse, R. and Meyer, K.D. (1987). *Statistics with Vague Data*, Reidel Publ. Co., Dordrecht.

Kullback, S. and Leibler, A. (1951). On information and sufficiency. *Ann. Math. Statist.* **22**, 79-86.

Kwakernaak, H. (1978). Fuzzy random variables. Part I: Definitions and theorems. *Inform. Sci.* **15**, 1-29.

Kwakernaak, H. (1979). Fuzzy random variables. Part II: Algorithms and examples for the discrete case. *Inform. Sci.* **17**, 253-278.

Lehmann, E.H. (1983). *Theory of Point Estimation.* John Wiley & Sons, New York.

Loéve, M. (1963). *Probability Theory.* 3rd ed. Van Nostrand, Princeton.

López Díaz, M. (1996). *Medibilidad e integración de variables aleatorias difusas. Aplicación a problemas de decisión.* PhD Thesis. Universidad de Oviedo.

López-Díaz, M. and Gil, M.A. (1998a). The λ-average value and the fuzzy expectation of a fuzzy random variable. *Fuzzy Sets and Systems* **99**, 347-352.

López-Díaz, M. and Gil, M.A. (1998b). Reversing the order of integration in iterated expectations of fuzzy random variables, and statistical applications. *J. Stat. Plan. Infer.* **74**, 11-29.

López García, H. (1997). *Cuantificación de la desigualdad asociada a conjuntos aleatorios y variables aleatorias difusas.* PhD Thesis. Universidad de Oviedo.

López-García, H. and Corral, N. (1998). Extension of the additively decomposable inequality indices for random sets and random fuzzy sets. *Statistical Papers* (submitted).

Lowen, R. (1990). A fuzzy Lagrange interpolation theorem. *Fuzzy Sets and Systems* **34**, 33-38.

Lubiano, M.A. and Gil, M.A. (1999). Estimating the expected value of fuzzy random variables in random samplings from finite populations. *Statistical Papers* (in press).

Lubiano, M.A., Gil, M.A., López-Díaz, M. and López, M.T. (1999a). The $\vec{\lambda}$-mean squared dispersion associated with a fuzzy random variable. *Fuzzy Sets and Systems* (in press).

Lubiano, M.A., Gil, M.A. and López-Díaz, M. (1999b). On the Rao-Blackwell Theorem for fuzzy random variable. *Kybernetika* (in press).

Lubiano, M.A. Salas, A. and Corral, N. (1998). The $\vec{\lambda}$-mean squared dispersion for the fuzzy linear regression. *Proc. IPMU'98* **I**, 101-107.

Mareš, M. (1994). *Computation over Fuzzy Quantities*. CRC Press, Boca Raton.

Marshall, A.W. and Olkin, I. (1979). *Inequalities*. Academic Press, New York.

Martínez López, I. (1991). *Aproximación de algunas familias de índices de variación en poblaciones finitas*. PhD. Thesis. Universidad de Oviedo.

Matheron, G. (1975). *Random Sets and Integral Geometry.* John Wiley & Sons, New York.

Molchanov, I.S. (1993). *Limit Theorems for Unions of Random Closed Sets.* Lect. Notes Math. **1561**. Springer-Verlag, Berlin.

Montenegro, M., Casals, M.R. and Gil, M.A. (1998). Estimating fuzzy parameters of fuzzy random variables in a Bayesian context. *Proc. IPMU'98* **II**, 1002-1007.

Moskowitz, H. and Kim, K. (1993). On assessing the H value in fuzzy linear regression. *Fuzzy Sets and Systems* **58**, 303-327.

Nakamura, K. (1986). Preference relations on a set of fuzzy utilities as a basis for decision making. *Fuzzy Sets and Systems* **20**, 147-162.

Näther, W. (1988). On expectation and variance of random fuzzy sets, with application to linear regression with vague observations. *BUSEFAL* **36**, 82-91.

Näther, W. and Albrecht, M. (1990). Linear regression with random fuzzy observations. *Statistics* **21**, 521-531.

Näther, W. (1997). Linear statistical inference for random fuzzy data. *Statistics* **29**, 221-240.

Norwich, A.M. and Turksen, I.B. (1984). A model for the measurement of membership and the consequences of its empirical implementation. *Fuzzy Sets and Systems* **12**, 1-25.

Negoita, C. and Ralescu, D.A. (1987). *Simulation, Knowledge-based Computing, and Fuzzy Statistics*. Van Nostrand Reinhold Co., New York.

Nguyen, H.T. (1978). A note on the extension principle for fuzzy sets. *J. Math. Anal. Appl.* **64**, 369-380.

Ok, E.A. (1994). On the approximation of fuzzy preferences by exact relations. *Fuzzy Sets and Systems* **67**, 173-179.

Ok, E.A. (1995). Fuzzy income inequality measurement: A class of fuzzy inequality measures. *Social Choice and Welfare* **12**, 111-136.

Ok, E.A. (1996). Fuzzy measurement of income inequality: Some possibility results on the fuzzification of the Lorenz ordering. *Econ. Theory* **7**, 513-530.

Okuda, T., Tanaka, H. and Asai, K. (1978). A formulation of fuzzy decision problems with fuzzy information, using probability measures of fuzzy events. *Inform. Contr.* **38**, 135-147.

Puri, M.L. and Ralescu, D. (1981). Différentielle d'une fonction floue. *C.R. Acad. Sci. Paris, Sér. I* **293**, 237-239.

Puri, M.L. and Ralescu, D. (1983). Differentials of fuzzy functions. *J. Math. Annal. Appl.* **91**, 552-558.

Puri, M.L. and Ralescu, D.A. (1985) The concept of normality for fuzzy random variables. *Ann. Probab.* **13**, 1373-1379.

Puri, M.L. and Ralescu, D. (1986). Fuzzy random variables. *J. Math. Anal. Appl.* **114**, 409-422.

Raj, D. and Khamis, S.H. (1958). Some remarks on sampling with replacement. *Ann. Math. Stat.* **29**, 550-557.

Ralescu, A. and Ralescu, D.A. (1984). Probability and fuzziness. *Inf. Sci.* **17**, 85-92.

Ralescu, A. and Ralescu, D.A. (1986). Fuzzy sets in statistical inference. In The Mathematics of Fuzzy Systems. (A. Di Nola and A.G.S. Ventre, Eds.), 273-283. Verlag TÜV Rheinland, Köln.

Ralescu, D.A. (1982). Fuzzy logic and statistical estimation. *Proc. 2nd World Conf. on Math. at the Serv. of Man*, 605-606.

Ralescu, D.A. (1995a). Fuzzy probabilities and their applications to statistical inference. In *Lecture Notes in Computer Sciences* **945**, Advances in Intelligent Computing (B. Bouchon and R. R. Yager, Eds.), 217-222. Springer-Verlag, Berlin.

Ralescu, D.A. (1995b). Inequalities for fuzzy random variables. *Proc. 26th Iranian Mathematical Conference*, 333-335.

Ralescu, D.A. (1995c). Fuzzy random variables revisited. *Proc. IFES'95 and Fuzzy IEEE Joint Conference* **2**, 993-1000.

Ramík, J. and Římánek, J. (1985). Inequality relation between fuzzy numbers and its use in fuzzy optimization. *Fuzzy Sets and Systems* **16**, 123-138.

Rao, C.R. (1973). *Linear Statistical Inference and Its Applications*. John Wiley & Sons, New York.

Rathie, P.N. (1971). On some new measures of uncertainty, inaccuracy and information and their characterizations. *Kybernetika* **7**, 394-403.

Redden, D.T. and Woodall, W.H. (1994). Properties of certain fuzzy linear regression methods. *Fuzzy Sets and Systems* **64**, 361-375.

Redden, D.T. and Woodall, W.H. (1996). Further examination of fuzzy linear regression. *Fuzzy Sets and Systems* **79**, 203-211.

Rudin, W. (1979). *Análisis Real y Complejo*. Alhambra, Madrid.

Salas, A. (1991). *Regresión lineal con observaciones difusas*. PhD Thesis. Universidad de Oviedo.

Sánchez de Posada Martínez, C. (1998). *Estudio matemático de los números difusos*. Trabajo Integrado y Académicamente Dirigido del Plan de Estudios de la Licenciatura de Matemáticas. Universidad de Oviedo.

Sancho Royo, A. (1997). *Evaluaciones de estímulos con características difusas*. PhD Thesis. Universidad de Granada.

Savic, A.A. and Pedrycz, W. (1991). Evaluation of fuzzy linear regression models. *Fuzzy Sets and Systems* **39**, 51-63.

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, New York.

Shorrocks, A.F. (1980). The class of additively decomposable inequality measures. *Econometrica* **48**, 613-625.

Stojaković, M. (1994). Fuzzy random variables, expectation and martingales. *J. Math. Anal. Appl.* **184**, 594-606.

Stoyan, D. (1998). Random sets: Models and Statistics. *Int. Stat. Rev.* **66**, 1-27.

Stoyan, D., Kendall, W.S. and Mecke, J. (1995). *Stochastic Geometry and its Applications.* 2nd ed. John Wiley & Sons, Chichester.

Tanaka, H., Uejima, S. and Asai, K. (1980). Fuzzy linear regression model. *Proc. Int. Congress on Applied Systems Research and Cybernetics* **VI**, 2933-2938.

Tanaka, H., Uejima, S. and Asai, K. (1982). Linear regression analysis with fuzzy model. *IEEE Trans. Syst. Man. Cyber.* **12**, 903-907.

Theil, H. (1967). *Economics and Information.* North-Holland, Amsterdam.

Thompson, S.K. (1992). *Sampling.* John Wiley & Sons, New York.

Tseng, T.V. and Klein, C.M. (1989). New algorithm for the ranking procedure in fuzzy decision-making. *IEEE Trans. Syst. Man and Cybernet.* **19**, 1289-1296.

Yager, R.R. (1981) A procedure for ordering fuzzy subsets of the unit interval. *Inf. Sci.* **24**, 143-161.

Zacks, S. (1971). *The Theory of Statistical Inference.* John Wiley, New York.

Zadeh, L.A. (1965). Fuzzy sets. *Inform. Contr.* **8**, 338-353.

Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning. *Inform. Sci.*, Part 1 **8**, 199-249; Part 2 **8**, 301-353; Part 3 **9**, 43-80.

Zadeh, L.A. (1976). A fuzzy-algorithmic approach to the definition of complex or imprecise concepts. *Int. J. Man-Mach. Stud.* **8**, 249-291.

Zagier, D. (1983). On the decomposability of the Gini coefficient and other indices of inequality. Discussion paper No. 108, University of Bonn.

Glossary of Symbols

$\mathcal{K}_c(\mathbb{R})$	class of nonempty compact intervals contained in $\mathbb{R}$
$\mathcal{K}_c((0,+\infty))$	class of nonempty compact intervals contained in $(0,+\infty)$
$\mathcal{F}_c(\mathbb{R})$	class of fuzzy subsets of $\mathbb{R}$ with α-level sets in $\mathcal{K}_c(\mathbb{R})$ (fuzzy numbers)
$\mathcal{F}_c((0,+\infty))$	class of fuzzy numbers contained in $(0,+\infty)$
$\mathcal{K}_c(\mathbb{R}^d)$	class of nonempty convex compact sets contained in $\mathbb{R}^d$
$\mathcal{F}_c(\mathbb{R}^d)$	class of fuzzy subsets of $\mathbb{R}^d$ with α-level sets and 0-level in $\mathcal{K}_c(\mathbb{R}^d)$
K	generic element of $\mathcal{K}_c(\mathbb{R})$
$\widetilde{V}$, $\widetilde{A}$	generic elements of $\mathcal{F}_c(\mathbb{R})$
$\widetilde{V}_\alpha$	α-level set of $\widetilde{V}$
$\widetilde{V}_0$	closed convex hull of the support of $\widetilde{V}$ (or 0-level)
$f_K(\lambda)$	convex linear combination of the extreme points of K given by $\lambda \sup K + (1-\lambda) \inf K$
$f_{\widetilde{V}}(\alpha,\lambda)$	convex linear combination of the extreme points of $\widetilde{V}_\alpha$ given by $\lambda \sup \widetilde{V}_\alpha + (1-\lambda) \inf \widetilde{V}_\alpha$
$\oplus$, $\bigoplus$	fuzzy addition
$\ominus$	fuzzy substraction
$\odot$	fuzzy product by a real number

$\otimes$	fuzzy product
$\oslash$	fuzzy quotient
$\succeq_S$	Ramík & Římánek's ranking criterion on $\mathcal{F}_c(\mathbb{R})$
$\geq_Y$	Yager's ranking criterion on $\mathcal{F}_c(\mathbb{R})$
F	ranking function in Yager's criterion
m	Lebesgue measure
d_H	Hausdorff's metric on $\mathcal{K}_c(\mathbb{R})$ (see Subject Index)
d_∞	generalized Hausdorff metric on $\mathcal{F}_c(\mathbb{R})$ (see Subject Index)
δ_p	δ_p metric (see Subject Index)
d_2	d_2 metric (see Subject Index)
D_S	S-distance (see Subject Index)
D_m	(see S-distance in Subject Index)
$S_{\vec{\lambda}}$	(see S-distance in Subject Index)
S_2	(see S-distance in Subject Index)
$s_{\widetilde{V}}$	support function of $\widetilde{V}$
$\mathcal{B}_{\mathbb{R}}$	Borel σ-field on $\mathbb{R}$
$\mathcal{B}_{[0,1]}$	Borel σ-field on $[0,1]$
$\mathcal{M}_{[0,1]}$	Lebesgue σ-field on $[0,1]$
$\mathcal{B}_{d_H}$	σ-field generated by the topology induced by d_H on $\mathcal{K}_c(\mathbb{R})$
$\mathcal{B}_{d_\infty}$	σ-field generated by the topology induced by d_∞ on $\mathcal{F}_c(\mathbb{R})$
$\mathcal{P}(\Omega)$	power set of Ω
X	convex compact random set
$\mathcal{X}$	fuzzy random variable
$\mathcal{X}_\alpha$	α-level function (see Subject Index)
$E(X)$ or $E(X\|P)$	(Aumann's) expected value of X w.r.t. P
$\overline{X}$	alternative notation to (Aumann's) expected value of X on a finite population
$\widetilde{E}(\mathcal{X})$ or $\widetilde{E}(\mathcal{X}\|P)$	(Puri & Ralescu's) expected value of $\mathcal{X}$ w.r.t. P

$\overline{\mathcal{X}}$	alternative notation to (Puri & Ralescu's) expected value of $\mathcal{X}$ on a finite population
$\delta(\,\cdot\,)$	Dirac distribution (see Subject Index)
δ_{ij}	Kronecker delta
$\mathbf{1}_{\{x\}}$	indicator function of $x \in \mathbb{R}$
$a\,]$	greatest integer part of $a \in \mathbb{R}$
$DCM_S(X, K)$	S-mean squared dispersion of X about K
$DCM_S(\mathcal{X}, \widetilde{V})$	S-mean squared dispersion of $\mathcal{X}$ about $\widetilde{V}$
$\Delta_S^2(\,\cdot\,)$ or $\Delta_S^2(\,\cdot\,\|P)$	central S-mean squared dispersion of a FRV (CCRS)
υ	generic random sample of size n
$\overline{X}_n$	sample expected value of X
$\overline{\mathcal{X}}_n$	sample expected value of $\mathcal{X}$
$\Delta_S^2(X[\upsilon])$	central S-mean squared dispersion of X in υ
$\Delta_S^2(\mathcal{X}[\upsilon])$	central S-mean squared dispersion of $\mathcal{X}$ in υ
$\mathrm{I}^{\alpha}(\,\cdot\,)$ or $\mathrm{I}^{\alpha}(\,\cdot\,\|P)$	additively decomposable inequality index of order α associated with a RV (FRV, CCRS)
$\mathrm{I}_{\mathrm{NVar}}(\,\cdot\,)$ or $\mathrm{I}_{\mathrm{NVar}}(\,\cdot\,\|P)$	additively decomposable inequality index of order 2 associated with a RV (FRV, CCRS)
$\mathrm{I}_{\phi}(\,\cdot\,)$ or $\mathrm{I}_{\phi}(\,\cdot\,\|P)$	ϕ-inequality index associated with a RV (FRV, CCRS)
$\mathrm{I}_{H}(\,\cdot\,)$ or $\mathrm{I}_{H}(\,\cdot\,\|P)$	hyperbolic inequality index associated with a RV (FRV, CCRS)
$\mathrm{I}_{H}(\{\,\cdot\,\})$	hyperbolic *intra*value inequality index associated with a FRV (CCRS)
$\mathrm{I}_{Sh}(\,\cdot\,)$ or $\mathrm{I}_{Sh}(\,\cdot\,\|P)$	inequality index of the Shannon type associated with a RV (FRV, CCRS)
$\mathrm{I}_{T}(\,\cdot\,)$ or $\mathrm{I}_{T}(\,\cdot\,\|P)$	Theil index associated with a RV (FRV, CCRS)
$\mathrm{I}_{f}(\,\cdot\,)$ or $\mathrm{I}_{f}(\,\cdot\,\|P)$	f-inequality index associated with a RV (FRV, CCRS)
z_{α}	$1 - \alpha/2$ fractile of the standard normal distribution
Φ	distribution function of the standard normal distribution
$\mathrm{Tri}(\alpha, \beta, \gamma)$	triangular fuzzy number
$\mathrm{Tra}(\alpha, \beta, \gamma, \delta)$	trapezoidal fuzzy number
$S(a, b)$	S-curve (see Subject Index)
$\Pi(a, b, c)$	Π-curve (see Subject Index)

Subject Index