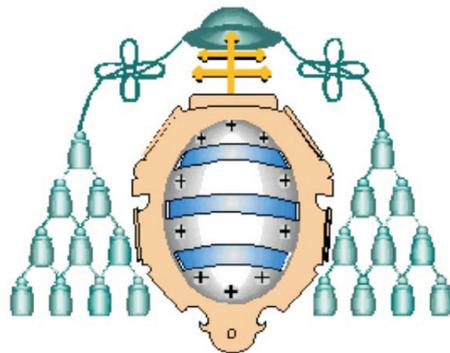


Análisis estadístico comparativo de tres escalas de valoración: Likert, fuzzy-Likert y fuzzy de respuesta libre

Sara de la Rosa de Súa

*Trabajo Fin de Máster dirigido
por Gil González-Rodríguez, María Asunción Lubiano y Gerardo Sanz*

Universidad de Oviedo



Oviedo, 19 de julio de 2012

Análisis estadístico comparativo
de tres escalas de valoración:
Likert, fuzzy-Likert y
fuzzy de respuesta libre

Sara de la Rosa de Súa

Trabajo Fin de Máster

Dirigido por Gil González-Rodríguez,
María Asunción Lubiano y Gerardo Sanz

Máster en Modelización Matemática,
Estadística y Computación

Universidad de Oviedo

19 de julio de 2012

Gil González Rodríguez y María Asunción Lubiano Gómez, profesores de la Universidad de Oviedo, y Gerardo Sanz Sáiz, profesor de la Universidad de Zaragoza, como tutores del Trabajo Fin de Máster de Sara de la Rosa de Saa con el título “Análisis estadístico comparativo de tres escalas de valoración: Likert, fuzzy-Likert y fuzzy de respuesta libre”, autorizan su presentación.

En Oviedo, a 19 de Julio de 2012



Fdo.: Gil González Rodríguez



María Asunción Lubiano Gómez

Gerardo Sanz Sáiz

Índice general

Índice de figuras	III
Prólogo	V
1. Conceptos y resultados preliminares	1
1.1. Las variables lingüísticas y su lógica fuzzy	4
1.2. Las escalas Likert y su uso en cuestionarios de opinión	5
1.3. La escala de números fuzzy y cuestionarios basados en ella	9
1.4. Tratamiento estadístico de respuestas basadas en escalas Likert	12
1.5. Tratamiento estadístico de respuestas basadas en escalas fuzzy	13
1.5.1. Aritmética entre números fuzzy	14
1.5.2. Números fuzzy aleatorios	21
1.5.3. Una metodología para el tratamiento estadístico de datos fuzzy	26
2. Estudio comparativo entre los cuestionarios basados en escalas Likert y los que utilizan la escala de números fuzzy	29
2.1. Preliminares del estudio: simulaciones de números fuzzy, ‘Likertización’ y conversión fuzzy	30
2.1.1. Simulación de números fuzzy	31
2.1.2. ‘Likertización’ de los números fuzzy simulados	33
2.1.3. Conversión fuzzy de las respuestas Likertizadas	34
2.1.4. Conversión de las respuestas fuzzy libres a las respuestas fuzzy prefijadas	35
2.1.5. Transformaciones necesarias de las muestras fuzzy y Likert para su posterior comparación	35
2.1.6. Herramientas de comparación: índice de diversidad y cuantificación del error cuadrático medio	37
2.2. Estudio comparativo empírico	40
2.2.1. Tablas para el caso $k = 4$	45

ÍNDICE GENERAL

2.2.2. Tablas para el caso $k = 5$	51
2.2.3. Tablas para el caso $k = 6$	57
2.2.4. Tablas para el caso $k = 7$	63
2.3. Conclusiones empíricas obtenidas a partir de las simulaciones	69
Bibliografía	71
Índice temático	77

Índice de figuras

1.1.	Ejemplo de variable lingüística y unos números fuzzy asociados	5
1.2.	Ejemplo sencillo de cuestionario Likert	7
1.3.	Ejemplo de cuestionario Likert con codificación entera	7
1.4.	Ejemplo de escala continua	8
1.5.	Ejemplo de conversión a fuzzy de la escala Likert	9
1.6.	Ejemplo de escala fuzzy libre	9
1.7.	Número fuzzy y α -nivel asociado	10
1.8.	Ejemplo de un cuestionario Likert-fuzzy	11
1.9.	Respuesta de un alumno a la valoración de la comida	12
1.10.	Ejemplo de gráficos para datos en escala Likert	13
1.11.	Suma de dos números fuzzy	15
1.12.	Producto de $\gamma = -1,5$ por un número fuzzy	15
1.13.	Ejemplo de suma de cierto número fuzzy \tilde{U} y el valor $(-1) \cdot \tilde{U}$	16
1.14.	Valor medio para las dos preguntas consideradas en el Ejemplo 1.3.1	25
2.1.	Caracterizaciones del número fuzzy trapezoidal	31
2.2.	Densidades de las beta involucradas en las simulaciones	32

Prólogo

Cuando se diseñan cuestionarios para recabar valoraciones sobre aspectos que no son medibles de forma exacta, es habitual recurrir al empleo de escalas Likert. Este tipo de escalas se emplean habitualmente en cuestionarios de calidad, satisfacción, opinión, etc., usualmente vinculados a las Ciencias Sociales y del Comportamiento, aunque también puede verse su uso en problemas de diagnóstico médica, ingeniería de control, etc.

Cada cuestionario de este tipo consta de una serie de preguntas a las que el encuestado debe responder eligiendo la respuesta que mejor representa su opinión o valoración, dentro de un listado preestablecido de posibles respuestas.

Se trata de cuestionarios cuya aplicación es sencilla, no requieren preparación específica previa de los encuestados, ni precisan de un entorno especial. Sin embargo, el análisis estadístico de los datos provenientes de tales cuestionarios resulta bastante pobre. Por un lado, las técnicas estadísticas disponibles para el tratamiento de datos categóricos, como corresponde a las respuestas en escala Likert, son bastante limitadas. Por otro lado, si se recurre a la codificación mediante números enteros de las respuestas, existen más técnicas aplicables, pero se trabaja con variables discretas que toman un número reducido de posibles valores distintos, y a ello hay que añadir la crítica de que la codificación entera no suele reflejar adecuadamente las diferencias reales entre las distintas respuestas, en el sentido de que por ejemplo la mayoría de los individuos no distinguen de la misma forma entre las respuestas codificadas por los enteros 1 y 2 que entre las codificadas por los enteros 2 y 3.

Para soslayar estos inconvenientes, algunos autores han propuesto alternativas al empleo de escalas Likert:

- reemplazar estas escalas por escalas continuas. Más concretamente, la idea consiste en señalar el punto o valor en una barra o intervalo acotado cuyos extremos son opcionales, que quien responde considera que es el que mejor expresa su respuesta a la cuestión específica. Esta alternativa de escala continua recoge muy bien la diversidad, variabilidad y subjetividad en la respuesta, y se adapta muy bien al tratamiento estadístico. Sin embargo, conlleva un

punto crítico importante: la elección del punto que represente fielmente cada respuesta no es sencilla y, de hecho, no parece natural exigir tanta exactitud en la respuesta a cuestiones intrínsecamente imprecisas como son las relativas a opiniones, valoraciones, percepciones, etc.

- realizar una codificación ‘fuzzy’ de las respuestas ofrecidas por las escalas Likert. De este modo, en lugar de recurrir a una codificación entera, el empleo de números fuzzy permite captar mejor la imprecisión inherente a las categorías Likert. Sin embargo, este procedimiento muestra también algunos puntos críticos: esta codificación supondría una escala continua en cuanto al soporte de la imagen pero con un número discreto y pequeño de valores distintos, lo que limitaría su tratamiento estadístico y las conclusiones que pueden derivarse del mismo; además, la diversidad, variabilidad y subjetividad de las respuestas tampoco quedarían bien reflejadas en general.

En resumen, siempre que el cuestionario se aplique con el objeto de tantear tendencias en las valoraciones, que las implicaciones reales o científicas que se deriven no sean de gran calado, y que no se disponga de tiempo para entrenar a los encuestados, su formación sea muy elemental o se realice un sondeo por teléfono o a pie de calle, etc., la sencillez y pocas exigencias del empleo de las escalas Likert pueden hacerlas aconsejables.

Por el contrario, cuando el cuestionario se aplica con una finalidad de mayor trascendencia, conviene tener presente que la elección de respuestas de un listado preestablecido restringe notablemente la subjetividad, diversidad y variabilidad de las respuestas, lo que estadísticamente se traduce en una pérdida importante de información. Además, las técnicas estadísticas aplicables tampoco explotan en su totalidad la información disponible. Por otro lado, a pesar de la riqueza de los diferentes idiomas, resulta prácticamente inviable construir una escala continua o suficientemente amplia a base de matices y modificaciones lingüísticas.

En trabajos previos del Grupo de Investigación SMIRE de la Universidad de Oviedo (<http://bellman.ciencias.uniovi.es/SMIRE>) se ha propuesto un procedimiento que, combinando las ideas directrices de las dos alternativas que hemos comentado anteriormente (reemplazar las escalas Likert por escalas continuas y realizar una codificación fuzzy de las respuestas ofrecidas en la escala Likert), integra las ventajas de ambas y elimina los puntos críticos indicados en el uso de la escala Likert en cuestionarios.

La idea del procedimiento sugerido consiste en utilizar la escala de números fuzzy en forma libre, pues se trata de una escala continua (en cuanto al número de

posibles valores distintos en ella) que permite muchos y variados matices. Además, gracias a la aritmética entre estos números, los valores admiten un buen tratamiento matemático y, más concretamente, estadístico.

La idea del procedimiento sugerido consiste en representar gráficamente el número fuzzy en un intervalo, cuyos extremos son prefijados de antemano, que quien responde considera que es el que mejor expresa su respuesta a la cuestión dada. Por supuesto, no es siempre posible expresar las valoraciones de todos los aspectos en términos de números fuzzy, pero el significado e interpretación intuitiva de estos números se adapta de forma flexible y amigable a un gran número de ellas.

Así, diseñando el cuestionario de acuerdo con un formato de respuesta libre y recurriendo a la escala de números fuzzy, se garantiza que dichas respuestas reflejen la imprecisión intrínseca a las valoraciones a las que se refieren, y que pueda captarse la subjetividad, diversidad y variabilidad propias de las mismas.

Por otro lado, el desarrollo en los últimos años de una metodología estadística para el análisis de datos fuzzy, mayoritariamente llevada a cabo por el Grupo SMIRE, permite un tratamiento muy similar al del análisis de datos numéricos, conservándose la mayoría de las ideas y criterios de este análisis.

Como consecuencia de la conjunción del empleo de la escala fuzzy, el formato de respuesta libre y la aplicación de la metodología para el análisis estadístico de datos fuzzy, se espera una ganancia sustancial de información y fiabilidad en las conclusiones.

Sin embargo, el desarrollo de este último procedimiento también tiene un punto crítico, compartido con las dos alternativas a la codificación entera de las respuestas Likert que se expusieron anteriormente. Ese punto crítico se refiere a las dificultades en su viabilidad práctica en situaciones en las que la aplicación de los cuestionarios requiere una respuesta muy rápida (como ocurre, por ejemplo, cuando las respuestas son el resultado de una entrevista telefónica, a pie de calle, etc.). En líneas generales, para la aplicación de estos cuestionarios se necesita un tiempo para explicar al encuestado el significado e interpretación de un número fuzzy. Sin embargo, y por experiencias realizadas en este sentido, el tiempo imprescindible para familiarizar al encuestado con la técnica para cubrir el cuestionario es más bien corto y usualmente se trata de una explicación grupal y no individual, por lo que la limitación se reduce básicamente a las situaciones antedichas.

Trataremos por tanto de formalizar la intuición que tenemos de que el empleo de cuestionarios de respuesta libre basados en números fuzzy va a ser más informativo desde el punto de vista estadístico que el de las escalas Likert con su codificación entera, y también trataremos de comprobar que utilizar cuestionarios con un formato

de respuesta libre y no con formato cerrado de respuesta, permite explotar mucha más información y recoger de forma más fiable las opiniones y valoraciones de los encuestados.

Es evidente que resulta imposible establecer al respecto conclusiones teóricas de índole general, es decir, que garanticen que el comportamiento es uniformemente mejor en el caso de las escalas fuzzy que en el de las escalas Likert-codificación entera, y en el caso de un formato de respuesta libre fuzzy frente a un formato de respuesta prefijado. Siempre podríamos encontrar contraejemplos (tal vez poco realistas, pero teóricamente posibles) en sentido contrario. Sin embargo, sí puede comprobarse a través de estudios de simulación y desde un punto de vista estadístico que el comportamiento es mayoritariamente mejor en el caso fuzzy y libre.

El presente Trabajo Fin de Máster completa el estudio iniciado con el desarrollo del Trabajo Integrado en el Plan de Estudios y Académicamente Dirigido “Las escalas Likert y fuzzy: comparación empírica por medidas de error estadístico” que se presentó el 8 de Abril de 2011, y la Tesina de Licenciatura “Representatividad de la mediana para datos en escalas Likert y fuzzy: estudio empírico comparativo” presentada el 7 de Julio de 2011.

En el capítulo 1 de este trabajo, se detalla el uso mayoritario de las escalas Likert en cuestionarios de opinión que involucran variables lingüísticas, así como las alternativas al empleo de este tipo de escala, ilustradas todas con ejemplos. Además se explica de forma detallada cómo sería el uso de la escala libre fuzzy en tales cuestionarios, y se definen las nociones de aritmética entre números fuzzy, métricas entre esos valores, números fuzzy aleatorios y parámetros relevantes de su distribución, como son: la media tipo Aumann como medida de tendencia central con valores fuzzy, la diversidad a través del índice de Gini-Simpson, y también se define el error cuadrático medio para estos números fuzzy que se utilizará como indicador de la representatividad de la media tipo Aumann antes mencionada.

En el Capítulo 2 se desarrolla el estudio empírico comparativo llevado a cabo entre el empleo de cuestionarios con escalas Likert con codificación entera y el empleo de cuestionarios de respuesta libre basados en la escala de números fuzzy, y también se realiza una comparación entre las escalas fuzzy usadas en formato libre y en formato preestablecido. Para todo ello se recurre a simulaciones que tratan de imitar en lo posible las respuestas que podrían proporcionar muestras de individuos a los que se les planteara un cuestionario y tuvieran que dar simultáneamente su respuesta en la escala de números fuzzy y en una escala Likert (con posibles respuestas de 4 a 7, que es el número idóneo desde una perspectiva estadística), y se simula también la alternativa que proponen algunos autores de llevar a cabo una conversión fuzzy

de las respuestas Likertizadas, así como una hipotética conversión de los fuzzy libres a unos fuzzy prefijados.

Las simulaciones han considerado números fuzzy trapezoidales y éstos se han generado siguiendo un procedimiento que fue validado por diversos ejemplos reales. Las respuestas fuzzy y numérica no se han simulado independientemente (no tendría sentido), sino haciendo una codificación de cada número fuzzy por un número entero, y también una codificación de cada entero con un fuzzy prefijado, además de una conversión fuzzy libre-fuzzy prefijado. El criterio de codificación empleado que se comentará en dicho capítulo también fue validado recurriendo a ejemplos reales.

Se han generado 1000 muestras de tamaños 10, 30, 50, 100 y 300. Para cada una de las muestras de números fuzzy trapezoidales simuladas, su correspondiente muestra likertizada y su correspondiente muestra fuzzy prefijada, se han determinado:

- el índice de Gini-Simpson (o entropía cuadrática) como medida de la diversidad de valores en cada muestra (también podría interpretarse como medida de la cantidad de información que aportan los datos muestrales);
- el error cuadrático medio respecto de la media tipo Aumann, como medida de la representatividad de la media en las muestras.

Se completa el capítulo con un resumen de las primeras conclusiones derivadas de las simulaciones, a partir de las cuales van a planificarse algunas pautas de actuación en un futuro a corto/medio plazo en esta línea de investigación.

Deseo agradecer a mis tres directores del Trabajo Fin de Máster y a la profesora M^a Ángeles Gil su tiempo y dedicación al desarrollo de este trabajo, así como su esfuerzo por ayudarme y orientarme en su elaboración.

Capítulo 1

Conceptos y resultados preliminares

En este capítulo se recogen las ideas y nociones básicas acerca de las escalas Likert y de la escala de números fuzzy.

Los *conjuntos fuzzy* fueron introducidos por Lotfi A. Zadeh en 1965 como una extensión de los conjuntos clásicos de forma que permitieran una asignación o pertenencia gradual o relativa (es decir, no únicamente plena o nula) de los elementos al conjunto. Esta asignación se corresponde con un grado de compatibilidad, mayor o menor, con la propiedad que caracteriza el conjunto considerado, que se supone no está bien definida a diferencia de los conjuntos clásicos (que se consideran casos especiales de los fuzzy).

Los *números fuzzy* aparecieron por primera vez en trabajos de Zadeh (1975) y han recibido mucha atención investigadora, en parte iniciada por Dubois y Prade (1978) en sus estudios sobre la aritmética basada en el principio de extensión de Zadeh. Los números fuzzy se presentaron como conjuntos fuzzy que extienden los números reales y los valores de intervalo compacto (o las funciones indicador de los conjuntos unitarios e intervalos correspondientes) y se formalizaron mediante funciones semicontinuas superiormente (y, por tanto, continuas o al menos continuas a trozos) y normales (es decir, algún número real debe ser plenamente compatible con el número fuzzy). Algunos autores distinguen entre números e intervalos fuzzy, exigiendo para el primer caso que únicamente exista un número real plenamente compatible con cada número fuzzy; en este trabajo se adopta la definición más amplia que elude esa restricción, y a la que también se ha hecho referencia en la literatura como cantidad fuzzy. Los números fuzzy se han empleado con cierta asiduidad en estudios en Ingeniería, Ciencias Naturales y Ciencias Sociales, entre otros campos. En todos ellos, y de forma muy evidente en los dos últimos, aparecen situaciones en las que los datos experimentales recogidos con fines estadísticos encierran imprecisión y subjetividad.

Los *números fuzzy aleatorios* (que son el caso unidimensional de los conjuntos fuzzy aleatorios, originalmente denominadas variables aleatorias fuzzy según el enfoque de Puri y Ralescu, 1986) constituyen un modelo para los mecanismos que generan aleatoriamente datos imprecisos que puedan asimilarse adecuadamente con números fuzzy. Se trata de un modelo bien formalizado dentro del contexto probabilístico como elemento aleatorio con valores en el espacio de los números fuzzy y de manera que las nociones fundamentales de distribución inducida, independencia, muestra aleatoria simple, etc. son aplicables de forma directa.

Los aspectos probabilísticos y métricos relativos a los números fuzzy aleatorios (y, en general, a los conjuntos fuzzy aleatorios) han recibido bastante atención desde su introducción en 1986. Entre algunos de estos aspectos destacan los concernientes a las condiciones de medibilidad y equivalencias, medidas resumen relevantes de su distribución, métricas entre números fuzzy y leyes de los grandes números basadas en ellas, así como otros resultados límite, procesos estocásticos, etc. Sin embargo, hay aspectos probabilísticos para los que aún no se han establecido conceptos ni resultados equiparables a los del caso real, como son: modelos de distribuciones realistas y manejables para las distribuciones de los números fuzzy aleatorios, y teoremas del límite central que proporcionen siempre un límite dentro de la clase de los números fuzzy aleatorios.

Los aspectos estadísticos no se han estudiado en la misma profundidad que los probabilísticos, si bien en la última década y en buena parte dentro del Grupo SMIRE, se está desarrollando una metodología estadística, mayoritariamente inferencial, que subsana las dificultades surgidas (debidas a aspectos probabilísticos aún deficientes, a la falta de una diferencia entre números fuzzy bien definida y con propiedades análogas a las del caso real y de un orden completo universalmente aceptable entre tales números) mediante, entre otros, la incorporación de métricas idóneas y las aproximaciones bootstrap de teoremas límite.

Aunque los números fuzzy aleatorios son un modelo bien establecido y con certeza aplicable a un buen número de problemas del mundo real, son escasas las aplicaciones que se han hecho de los mismos y casi siempre han partido de los investigadores ‘teóricos’. Es posible que esta situación se deba a las posibles dificultades conceptuales, propiamente matemáticas/probabilísticas/estadísticas del modelo y de la metodología estadística con datos fuzzy que está empezando a desarrollarse. Pero, se trata de dificultades compartidas con la estadística de datos reales y con la única complejidad añadida de la descripción de los datos imprecisos mediante números fuzzy.

Cuando, como matemáticos-probabilistas-estadísticos, se estudia por primera vez el concepto de número fuzzy aleatorio, uno de los primeros ejemplos reales que viene a la mente es el de los datos asociados con opiniones, valoraciones cualitativas, juicios o percepciones subjetivas, que con tanta asiduidad se presentan en estudios de Ciencias Sociales, Biomédicas, etc. De hecho, y de forma más concreta, cuando se observan cuestionarios basados en escalas Likert es fácil relacionar e identificar las respuestas posibles con ciertos números fuzzy. Estos procedimientos los podemos encontrar en cuestionarios fuzzy SERVQUAL (ver, por ejemplo, Aydin y Pakdil, 2008, Chou *et al.*, 2011 y Hu *et al.*, 2010) y en otros trabajos como los de Bharadwaj (2007), Bocklisch *et al.* (2010), Bocklisch (2011), Herrera (2008), Lalla *et al.* (2004), Lazim y Osman (2009), Li (2010) o Turksen y Willson (1994) .

Un procedimiento diferente al anterior es considerar cuestionarios con un formato de respuesta libre basado en la escala de números fuzzy. Esta técnica fue iniciada por Hesketh, Hesketh y colaboradores (ver, por ejemplo, Hesketh y Hesketh, 1994, Hesketh *et al.*, 1988, 1995, 2011). Por este motivo, en trabajos previos del Grupo SMIRE (ver, entre los más recientes, González-Rodríguez *et al.*, 2011, Gil y González-Rodríguez, 2011, el Trabajo Integrado en el Plan de Estudios y la Tesina de Licenciatura referidos en el Prólogo -de la Rosa de Súa, 2011a, 2011b-) se ha presentado y estudiado este tipo de cuestionario alternativo a los basados en escalas Likert. El fundamento de este cuestionario alternativo radica en la consideración de la escala de números fuzzy (que permite un manejo matemático y una explotación estadística de la información subyacente muy superiores a los de las escalas Likert) unida a una libertad, diversidad y variabilidad completa a la hora de dar las respuestas (inspirada en parte en la elección de puntuaciones reales en intervalos acotados, en trabajos como los de van Laerhoven *et al.*, 2004, o Kambaki-Vougioukli *et al.*, 2011). De este modo, en lugar de considerar formatos de respuesta fija, a elegir entre un número reducido de respuestas posibles, los encuestados pueden representar su opinión, valoración, juicio o percepción personal sin estar sujetos a un conjunto preestablecido de respuestas.

En consecuencia, a través del cuestionario alternativo es posible captar en una escala continua la subjetividad, diversidad y variabilidad que no permitirían las escalas Likert. Por ejemplo, cuando dos personas consideran que están ‘bastante en desacuerdo’ con cierta política, la escala de números fuzzy permite describir su grado de desacuerdo de forma matizada, de manera que pueden existir ligeras diferencias en su percepción; así mismo, un encuestado puede considerar que está ‘bastante en desacuerdo’ con ciertas iniciativas, pero con matices diferenciadores que la escala de números fuzzy le permiten reflejar. Además, la imprecisión intrínseca a las respuestas

de este tipo de cuestionarios queda recogida y bien formalizada.

Con el fin de analizar las ventajas del nuevo cuestionario frente a los basados en escalas Likert, desde una perspectiva estadística, en el Trabajo Integrado en el Plan de Estudios y Académicamente Dirigido “Las escalas Likert y fuzzy: comparación empírica por medidas de error estadístico” que se presentó el 8 de Abril de 2011, se desarrolló un estudio introductorio al respecto. En dicho trabajo, se analizó empíricamente, a través del error cuadrático medio correspondiente, el comportamiento de la media como representante de la tendencia central de la población o muestra de datos, habida cuenta de que tanto para datos fuzzy como numéricos la media es el valor (fuzzy o real, respectivamente) para el que dicho error cuadrático medio es menor.

Además, en la Tesina de Licenciatura “Representatividad de la mediana para datos en escalas Likert y fuzzy: estudio empírico comparativo” presentada el 7 de Julio de 2011, se siguió con este estudio comparativo entre las dos escalas, analizándose en esta ocasión la diversidad de las respuestas, la variabilidad relativa entre las mismas y la representatividad de la mediana, además de la media, como medidas de tendencia central a través del error absoluto medio y cuadrático medio respectivamente.

En este Trabajo Fin de Máster vamos a realizar modificaciones en la forma de simular las respuestas fuzzy, trabajaremos con una métrica distinta a las empleadas anteriormente, y realizaremos además un estudio comparativo entre respuestas fuzzy según el cuestionario sea con formato libre o preestablecido, además de seguir con el estudio comparativo entre escalas Likert y fuzzy.

En este capítulo se recopilan las nociones básicas y las herramientas matemáticas que se aplicarán para los estudios comparativos del próximo capítulo. Una parte de este capítulo se presentó en forma esencialmente análoga dentro del Trabajo Integrado y de la Tesina.

1.1. Las variables lingüísticas y su lógica fuzzy

Una **variable lingüística** es una variable cuyos valores son palabras o sentencias del lenguaje natural, en vez de valores numéricos. Los valores de una variable de este tipo se denominan términos o etiquetas lingüísticas, y ayudan a caracterizar o describir fenómenos que están mal definidos o son complejos de definir. Por ejemplo, la *temperatura* es una variable lingüística, y ‘muy fría’, ‘fría’, ‘normal’, ‘caliente’ y ‘muy caliente’ son etiquetas lingüísticas.

Toda variable lingüística posee una regla semántica que asocia a cada etiqueta

lingüística un número fuzzy que viene caracterizado por una aplicación $\tilde{U} : \mathbb{R} \rightarrow [0, 1]$ tal que para cada $x \in \mathbb{R}$, $\tilde{U}(x)$ denota el ‘grado de compatibilidad’ de x con la propiedad que define \tilde{U} (más adelante volveremos a definir de forma más detallada este concepto).

Ejemplo 1.1.1. En la Figura 1.1 podemos ver un ejemplo. En este caso, la variable lingüística sería el *rendimiento*, y sus etiquetas lingüísticas ‘muy bajo’, ‘bajo’, ‘medio’, ‘alto’ y ‘muy alto’. A través de una regla semántica estos cinco términos lingüísticos se asocian con un número fuzzy (restricción difusa) cada uno. Notar que una característica importante de esta lógica fuzzy es que un mismo valor puede pertenecer a varias etiquetas lingüísticas a la vez. Por ejemplo, un valor numérico para el rendimiento puede ser considerado como ‘bajo’ y ‘medio’ al mismo tiempo.

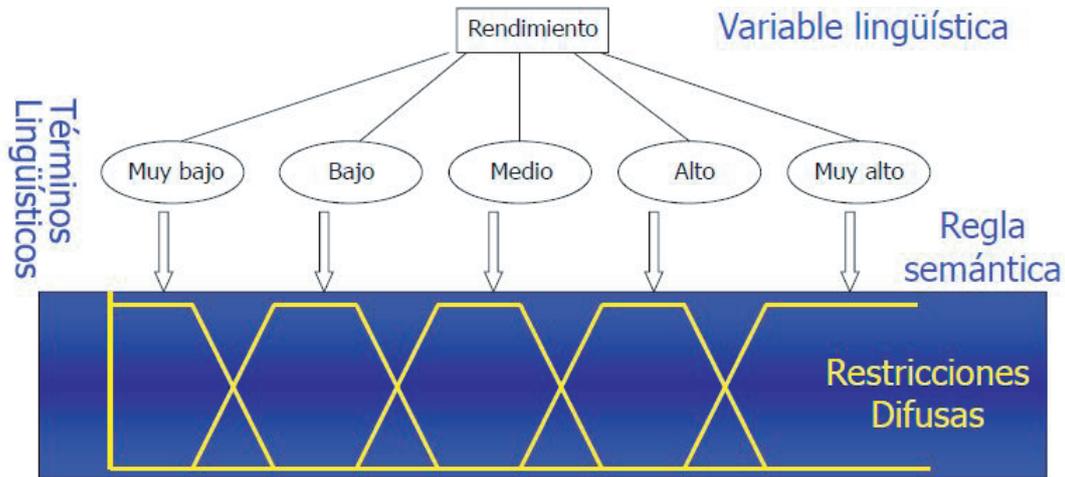


Figura 1.1: Ejemplo de variable lingüística y unos números fuzzy asociados

Cuando nos referimos a cuestionarios de opinión y valoración, son este tipo de variables lingüísticas las que suelen involucrar. Más concretamente, para definir sus etiquetas lingüísticas se suele recurrir a las escalas tipo Likert, que pasamos a definir a continuación.

1.2. Las escalas Likert y su uso en cuestionarios de opinión

Una *escala Likert* es un tipo de escala ordinal comúnmente utilizado en cuestionarios de opinión y valoración, y de uso muy frecuente en encuestas para la investigación.

Desarrollada en 1932 por el sociólogo Rensis Likert, con ella se pretende medir la actitud de los encuestados ante un tema determinado. Para ello, se elaboran una serie de enunciados (ítems) sobre el tema, ante los cuales los individuos tienen que manifestar su grado de acuerdo o desacuerdo eligiendo una de las k respuestas posibles de la escala dentro un listado preestablecido (normalmente entre 3 y 10, si bien y como argumentan Lozano *et al.*, 2008, desde un punto de vista psicométrico y estadístico las elecciones más convenientes son $4 \leq k \leq 7$). Normalmente el listado de respuestas es el mismo para todas las cuestiones formuladas.

Es frecuente codificar cada respuesta mediante un número entero para poder realizar posteriormente un análisis estadístico más rico que si tratáramos las variables como datos categóricos, aunque el número de técnicas que podemos aplicar sigue siendo más escaso que si trabajáramos con variables continuas.

Como **ventajas** del uso de este tipo de escala se pueden citar:

- es una escala fácil y rápida de construir, a la vez que sencilla de comprender y rápida de contestar;
- la verbalización forzada que supone responder a cuestionarios que utilizan este tipo de escala conlleva menos ambigüedad en las respuestas que la obtenida con otro tipo de cuestionarios.

y entre las **desventajas** se encuentran:

- discretiza la variable en un número reducido de valores distintos;
- la codificación entera de las respuestas no suele reflejar adecuadamente las diferencias reales entre cada respuesta;
- la obligación de elegir entre una lista previa de etiquetas categóricas resta libertad de descripción, y la diversidad y subjetividad que llevan implícitas estas valoraciones;
- como ya citamos anteriormente, el análisis estadístico de las respuestas es bastante pobre.

Ejemplo 1.2.1. La Figura 1.2 muestra un ejemplo de un cuestionario cuya lista de respuestas posibles es tipo Likert.

Se trata de una encuesta realizada a los alumnos de una clase para que evalúen el trabajo de un profesor. La encuesta consta de siete preguntas, y los encuestados deben elegir para cada una de ellas la respuesta que mejor se adecúe a su opinión dentro del siguiente listado de respuestas posibles: EXCELENTE, MUY BUENO, BUENO, INDIFERENTE y MALO.

	EXCELENTE	MUY BUENO	BUENO	INDIFERENTE	MALO
Ha utilizado explicaciones teóricas	<input type="radio"/>				
Ha realizado prácticas	<input type="radio"/>				
Ha fomentado el trabajo en equipo	<input type="radio"/>				
Tiene buena capacitación técnica	<input type="radio"/>				
Ha propiciado buena comunicación	<input type="radio"/>				
Ha sabido motivar al grupo	<input type="radio"/>				
Ha cumplido los objetivos propuestos	<input type="radio"/>				

Figura 1.2: Ejemplo sencillo de cuestionario Likert

Como hemos indicado anteriormente, es frecuente codificar numéricamente estas respuestas y tratarlas como si fueran datos numéricos enteros a la hora de analizarlos con fines estadísticos. En ocasiones el propio cuestionario incluye dicha codificación como puede verse en el ejemplo siguiente:

MUY DEFICIENTE	1	DEFICIENTE	2	REGULAR	3	BIEN	4	MUY BIEN	5	NO SABE/ NO CONTESTA
I.- SOBRE LA LABOR DEL PROFESOR										
A.- SOBRE LA INFORMACION FACILITADA POR ESTE PROFESOR AL COMENZAR EL CURSO										
1.- Informa de los objetivos, contenidos, bibliografía y materiales recomendados										
	<input type="checkbox"/>									
2 - Informa de las pruebas y criterios de evaluación que se seguirá										
	<input type="checkbox"/>									
3.- Informa de los fines y horario de las tutorías										
	<input type="checkbox"/>									
B.- SOBRE EL CUMPLIMIENTO DE OBLIGACIONES DE ESTE PROFESOR										
4.- Asiste a sus clases y, en caso contrario, se justifica y se sustituye o recupera										
	<input type="checkbox"/>									
5.- Es puntual al comenzar y al finalizar la actividad docente										
	<input type="checkbox"/>									
6.- El profesor atiende las tutorías										
	<input type="checkbox"/>									
C.- SOBRE LAS RELACIONES DE ESTE PROFESOR CON EL ESTUDIANTE										
7.- Es correcto y respetuoso con el estudiante										
	<input type="checkbox"/>									
8.- Tiene una actitud receptiva y muestra disposición para el diálogo										
	<input type="checkbox"/>									
9.- Promueve el interés por la materia										
	<input type="checkbox"/>									
10.- Durante las tutorías ayuda a la comprensión y estudio de la asignatura										
	<input type="checkbox"/>									
D.- SOBRE EL DESARROLLO DE LA ACTIVIDAD DOCENTE DE ESTE PROFESOR										
11.- Explica de manera clara y ordenada, destacando los aspectos mas importantes										
	<input type="checkbox"/>									
12.- Relaciona unos temas con otros de la materia										
	<input type="checkbox"/>									
13.- Relaciona los conceptos de la materia con sus aplicaciones										
	<input type="checkbox"/>									
14.- La labor de este profesor hace que la asistencia a clase facilite la comprensión de la materia										
	<input type="checkbox"/>									
15.- Realiza el seguimiento y asesora sobre las actividades o trabajos										
	<input type="checkbox"/>									

Figura 1.3: Ejemplo de cuestionario Likert con codificación entera

Ejemplo 1.2.2. En la Figura 1.3 tenemos un ejemplo de una encuesta que utiliza la escala Likert y que además incluye dentro del mismo cuestionario la codificación de cada respuesta mediante un número entero.

Como el ejemplo anterior, vuelve a tratarse de una encuesta de valoración acerca de la labor de un profesor, y cada una de las cinco etiquetas lingüísticas disponibles para cada respuesta está codificada con un número entero del 1 al 5, cuya codificación viene recogida en la parte superior del cuestionario.

Para subsanar las desventajas que citamos anteriormente del empleo de este tipo de escala, algunos autores han propuesto alternativas para ayudar a mejorar estos inconvenientes. Entre ellas destacamos:

- elegir un punto concreto dentro de un intervalo cuyos extremos estén prefijados (habitualmente 0 y 1, 0 y 10 ó 0 y 100 -este último con una traducción intuitiva en términos de porcentajes-), como puede verse en la Figura 1.4. Se trata de una alternativa que trabaja con una escala continua, y que por tanto capta muy bien la diversidad y variabilidad en la respuesta, además de adaptarse muy bien al tratamiento estadístico. Sin embargo, no parece natural exigir al encuestado tanta exactitud para responder a cuestiones intrínsecamente imprecisas como son las relativas a opiniones, valoraciones, juicios, etc.

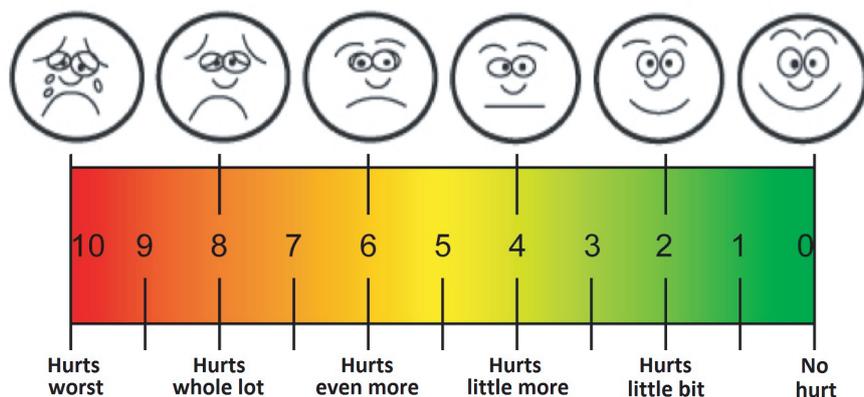


Figura 1.4: Ejemplo de escala continua

- realizar una conversión fuzzy de las respuestas ofrecidas por la escala Likert, como vemos en la Figura 1.5. De este modo, en lugar de recurrir a una codificación entera, el empleo de números fuzzy permite captar mejor la imprecisión inherente a las categorías Likert. Sin embargo, sigue tratándose de un tipo de escala discreta en cuanto al número de posibles valores distintos, y la diversidad y variabilidad no quedan bien reflejadas en general. En este trabajo consideraremos diferentes conversiones fuzzy a partir de la escala Likert que se han propuesto en la literatura.

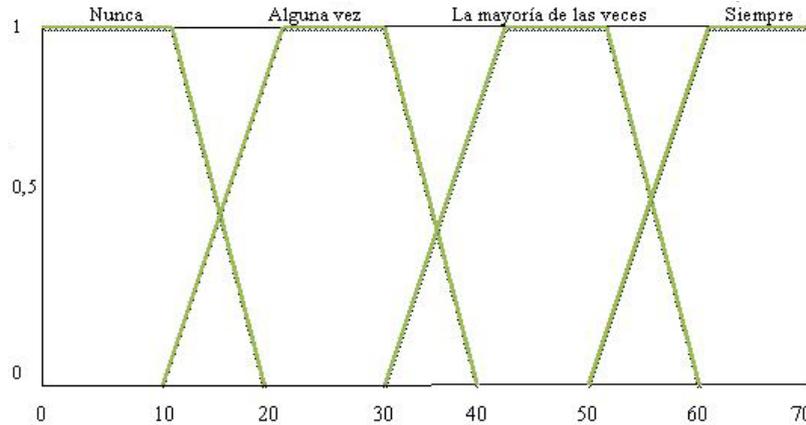


Figura 1.5: Ejemplo de conversión a fuzzy de la escala Likert

- utilizar una escala de números fuzzy libre, en la que el encuestado represente él mismo el número fuzzy que más se acerque a su valoración dentro de un intervalo (ver Figura 1.6). Esta alternativa es la que proponemos en este trabajo y trataremos de comprobar que en determinadas situaciones resulta más apropiada por captar más información que una escala Likert.

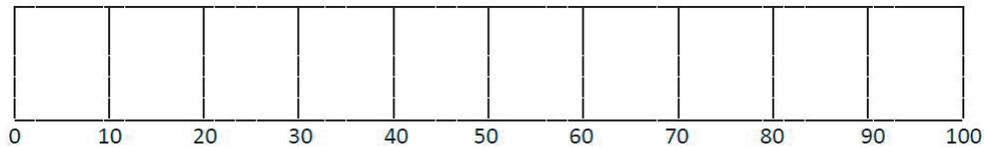


Figura 1.6: Ejemplo de escala fuzzy libre

1.3. La escala de números fuzzy y cuestionarios basados en ella

La *escala de números fuzzy* es una escala con valores funcionales, y de uso muy frecuente en campos muy diversos, desde la Ingeniería de Control a la Diagnósis Médica.

El concepto básico es el de número fuzzy, que es un caso especial de la noción de conjunto fuzzy introducida por Zadeh (1965). Concretamente:

Definición 1.3.1. *Un número fuzzy es una cantidad o valor ‘mal definido’ que se supone caracterizado por una aplicación $\tilde{U} : \mathbb{R} \rightarrow [0, 1]$ tal que*

- por un lado, captar mediante herramientas matemáticas la imprecisión intrínseca a este tipo de valoraciones, y la diversidad, subjetividad y variabilidad de las mismas;
- y por otro lado, concebir una metodología estadística que haga posible explorar y explotar exhaustivamente la información involucrada en esas respuestas.

Es en este último aspecto en el que Hesketh y Hesketh señalaban que las técnicas disponibles al respecto resultaban muy escasas y más bien poco potentes, lo que limitó sus estudios a la consideración de números fuzzy trapezoidales y al análisis estadístico de las variables que caracterizaban tal respuesta aleatoria.

Se trata por tanto de diseñar cuestionarios con formato de respuesta libre basados en la escala de números fuzzy (ver Figura 1.6), y realizar comparaciones empíricas con el uso de cuestionarios que emplean la escala Likert o la escala fuzzy en formato preestablecido.

Ejemplo 1.3.1. El ejemplo que se ilustra en la Figura 1.8 muestra una pequeña parte del cuestionario llevado a cabo durante el curso 2011/2012 en el Máster en Soft Computing and Intelligent Data Analysis realizado en Mieres y coordinado por miembros del grupo SMIRE.

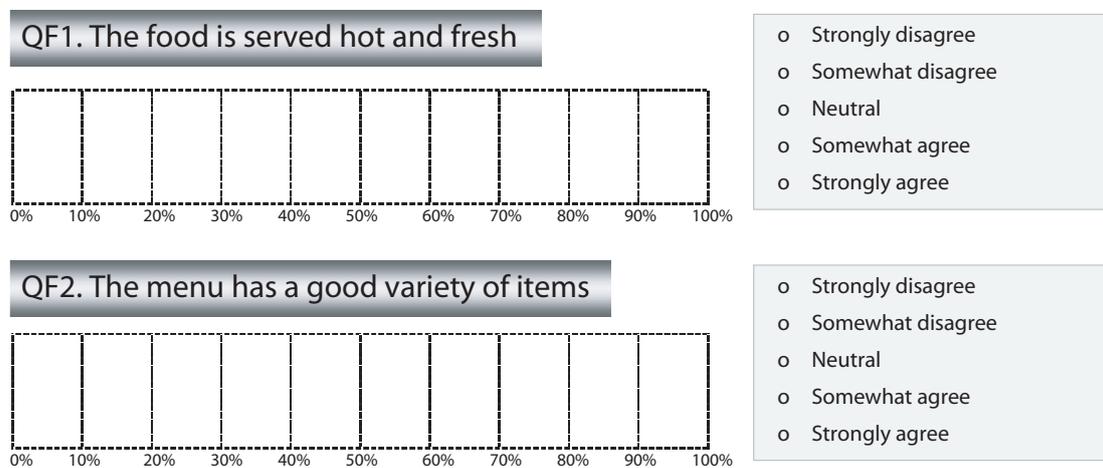


Figura 1.8: Ejemplo de un cuestionario Likert-fuzzy

Dicho cuestionario se realizó con el objeto de que los encuestados representaran su valoración y opinión acerca de un restaurante elegido por cada uno de ellos utilizando dos tipos de escala: la escala libre de números fuzzy y la escala Likert.

Debían valorar varios aspectos del restaurante elegido, entre ellos el estado de la comida al ser servida y la variedad de platos que ofrece el menú, como muestra

la Figura 1.8. El cuestionario constaba de 14 preguntas, y para cada una de ellas debían dar su respuesta empleando números fuzzy trapezoidales, por su simplicidad en la representación y posterior tratamiento, y también eligiendo una de las cinco posibles respuestas de la escala Likert.

En el caso de la representación de los números fuzzy, se tomó como conjunto de referencia (es decir, de valores potencialmente compatibles con cualquiera de las posibles respuestas) el intervalo $[0, 100]$ (interpretando las valoraciones como una especie de porcentajes, indicando con 0% el grado de acuerdo mínimo y con 100% el máximo).

La Figura 1.9 muestra un ejemplo de la respuesta que dio uno de los encuestados a la primera pregunta. La valoración que consideró más apropiada en la escala Likert fue “Strongly agree”. En cuanto a la interpretación de la respuesta según el número fuzzy representado, indica que ese estudiante considera que la valoración de la comida no es inferior al 70%, y es plenamente compatible con los valores entre el 75 y el 95% (es decir, el 1-nivel será $[75, 95]$), interpolándose por último estos intervalos mediante interpolación lineal para obtener un número fuzzy trapezoidal.

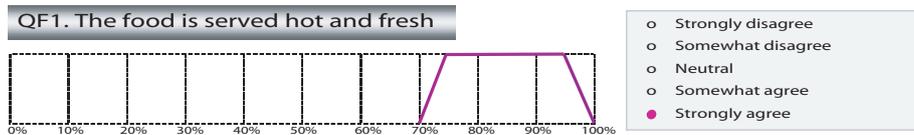


Figura 1.9: Respuesta de un alumno a la valoración de la comida

1.4. Tratamiento estadístico de respuestas basadas en escalas Likert

Un rasgo diferencial esencial entre el empleo de los cuestionarios basados en escalas Likert y el de los basados en la escala de números fuzzy es el correspondiente a su análisis estadístico.

Las escalas Likert discretizan las variables a las que se refieren en un número reducido de valores potenciales. El análisis estadístico de respuestas de un cuestionario basado en una escala Likert consistiría, en principio, en un estudio de datos categóricos. En ocasiones, y con el objeto de ampliar el conjunto de los posibles métodos que puedan aplicarse para tal análisis, se codifican tales datos categóricos mediante números enteros, y se aplican algunos métodos estadísticos para el tratamiento de

datos numéricos. Unos y otros resultan bastante limitados, puesto que el número de valores distintos es muy pequeño.

Además, como ya hemos apuntado anteriormente, una codificación entera no suele reflejar las diferencias reales entre los diferentes valores, es abrupta y difícilmente puede captar la imprecisión subyacente a las respuestas. Además, al corresponder a una lista prefijada de respuestas, la subjetividad apenas se recoge, de modo que desde el punto de vista estadístico, suele perderse y, por tanto, no explotarse mucha información relevante.

Volviendo al Ejemplo 1.3.1, podemos analizar los datos obtenidos con la escala Likert haciendo gráficos de barra o de sectores, como vemos en la Figura 1.10.

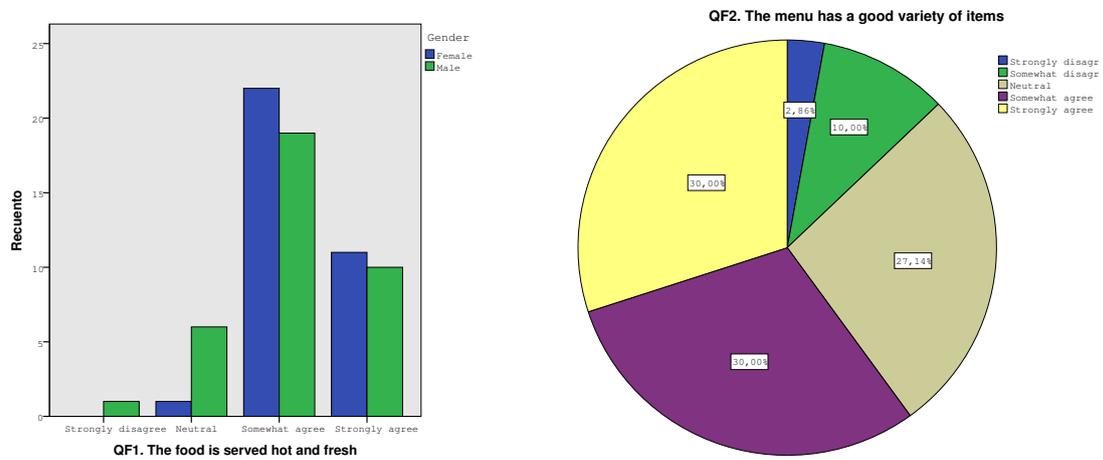


Figura 1.10: Ejemplo de gráficos para datos en escala Likert

1.5. Tratamiento estadístico de respuestas basadas en escalas fuzzy

A diferencia de las escalas Likert, la escala de números fuzzy es una escala continua y su uso permite soslayar los inconvenientes críticos que acaban de exponerse en la sección anterior.

El Grupo de Métodos Estadísticos con Elementos Aleatorios Imprecisos, SMIRE, lleva algo más de una década investigando sobre el análisis estadístico de datos fuzzy basándose en los conjuntos aleatorios fuzzy (Puri y Ralescu, 1986). El grupo es pionero en este tema y ha desarrollado y continúa desarrollando una metodología estadística basada esencialmente, además de en la aritmética fuzzy usual, en el uso de:

- métricas operativas entre datos fuzzy,
- resultados y procedimientos asintóticos para elementos aleatorios con valores en espacios generalizados,
- y la aproximación bootstrap para algunos de estos resultados.

Al ser los datos reales/vectoriales/intervalares un caso especial de los datos fuzzy que vamos a analizar, la aritmética y la distancia consideradas deben extender las de los primeros, persiguiendo a un tiempo que las extensiones sean operativas, naturales y con interpretación intuitiva.

1.5.1. Aritmética entre números fuzzy

Las operaciones aritméticas elementales que se precisan para el tratamiento estadístico de datos son esencialmente dos: la suma y el producto por escalares. La extensión usual de estas dos operaciones a la clase $\mathcal{F}_c(\mathbb{R})$ es la que se define sobre la base del *principio de extensión de Zadeh* (1975). De este modo:

Definición 1.5.1. *Dados dos números fuzzy $\tilde{U}, \tilde{V} \in \mathcal{F}_c(\mathbb{R})$, la **suma de \tilde{U} y \tilde{V}** es el número fuzzy $\tilde{U} + \tilde{V} \in \mathcal{F}_c(\mathbb{R})$ tal que*

$$(\tilde{U} + \tilde{V})(x) = \sup_{(y,z): y+z=x} \min \{ \tilde{U}(y), \tilde{V}(z) \}$$

lo que, por involucrar elementos de $\mathcal{F}_c(\mathbb{R})$, resulta equivalente a que para cada $\alpha \in (0, 1]$ se considere la suma intervalar usual, es decir, la suma de Minkowski entre intervalos, de manera que:

$$(\tilde{U} + \tilde{V})_\alpha = \tilde{U}_\alpha + \tilde{V}_\alpha = \{y + z : y \in \tilde{U}_\alpha, z \in \tilde{V}_\alpha\}.$$

La Figura 1.11 muestra la representación gráfica de la suma de dos números fuzzy.

Definición 1.5.2. *Dado un número fuzzy $\tilde{U} \in \mathcal{F}_c(\mathbb{R})$ y un escalar $\gamma \in \mathbb{R}$, el **producto de γ por \tilde{U}** es el número fuzzy $\gamma \cdot \tilde{U} \in \mathcal{F}_c(\mathbb{R})$ tal que*

$$(\gamma \cdot \tilde{U})(x) = \sup_{y: y=\gamma x} \tilde{U}(y) = \begin{cases} \tilde{U}\left(\frac{x}{\gamma}\right) & \text{si } \gamma \neq 0 \\ \mathbf{1}_{\{0\}}(t) & \text{si } \gamma = 0 \end{cases}$$

(con $\mathbf{1}_A =$ función indicador del conjunto A) lo que, por involucrar elementos de $\mathcal{F}_c(\mathbb{R})$, resulta equivalente a considerar nivel-a-nivel el producto natural de un escalar por un intervalo, es decir:

$$(\gamma \cdot \tilde{U})_\alpha = \gamma \cdot \tilde{U}_\alpha = \{\gamma \cdot y : y \in \tilde{U}_\alpha\}$$

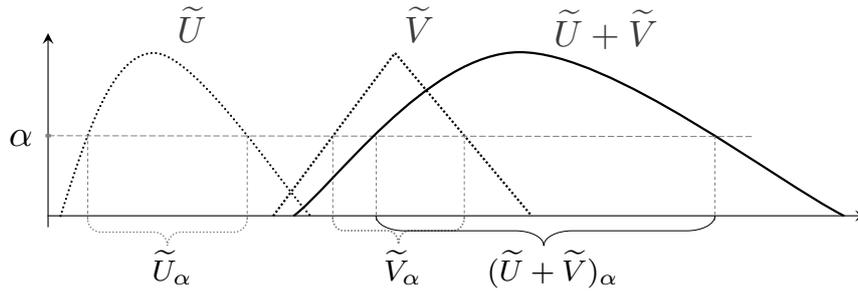


Figura 1.11: Suma de dos números fuzzy

$$= \begin{cases} [\gamma \cdot \inf \tilde{U}_\alpha, \gamma \cdot \sup \tilde{U}_\alpha] & \text{si } \gamma \geq 0 \\ [\gamma \cdot \sup \tilde{U}_\alpha, \gamma \cdot \inf \tilde{U}_\alpha] & \text{si } \gamma < 0 \end{cases}$$

La Figura 1.12 muestra la representación gráfica del producto de $\gamma = -1,5$ por un número fuzzy.

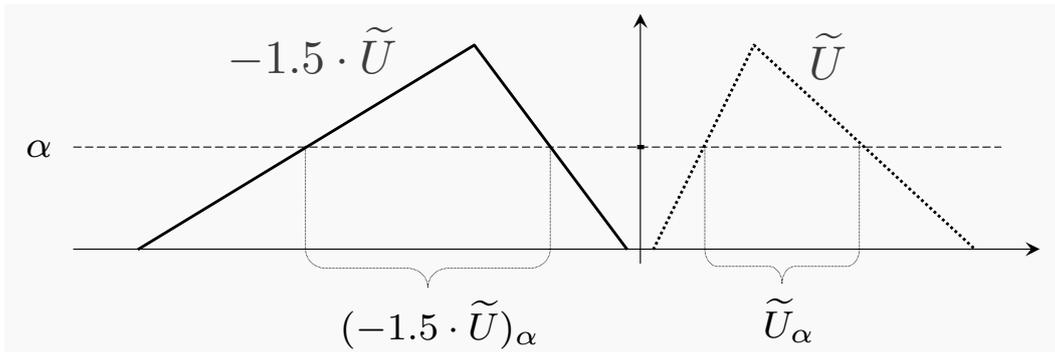


Figura 1.12: Producto de $\gamma = -1,5$ por un número fuzzy

Conviene realizar algunas puntualizaciones en relación con la aritmética de números fuzzy, como son las siguientes:

Observación 1.5.1. Aunque los números fuzzy son valores funcionales (funciones que tienen su imagen en $[0, 1]$), la aritmética usual entre números fuzzy no coincide con la usual entre funciones. De hecho, la aplicación de la aritmética funcional sobre los números fuzzy no produciría en general un número fuzzy salvo en situaciones esporádicas. En consecuencia, los datos con valores fuzzy no pueden tratarse directamente como datos funcionales, si bien sí puede establecerse una identificación indirecta que haga viable dicho tratamiento.

Observación 1.5.2. La clase de los números fuzzy dotada con la aritmética introducida mediante las dos operaciones precedentes, es decir, $(\mathcal{F}_c(\mathbb{R}), +, \cdot)$, no tiene

estructura lineal sino semilineal y, más concretamente, cónica. La falta de linealidad es consecuencia directa del hecho de que $\tilde{U} + (-1) \cdot \tilde{U} \neq \mathbf{1}_{\{0\}}$ (elemento neutro de la suma de Minkowski de números fuzzy), de manera que no existe en general el simétrico respecto de la suma de un elemento cualquiera de $\mathcal{F}_c(\mathbb{R})$.

La Figura 1.13 muestra gráficamente la suma de un número fuzzy \tilde{U} y el valor $(-1) \cdot \tilde{U}$, que da como resultado un número fuzzy simétrico respecto de 0, pero que no se reduce a la función indicador de $\{0\}$.

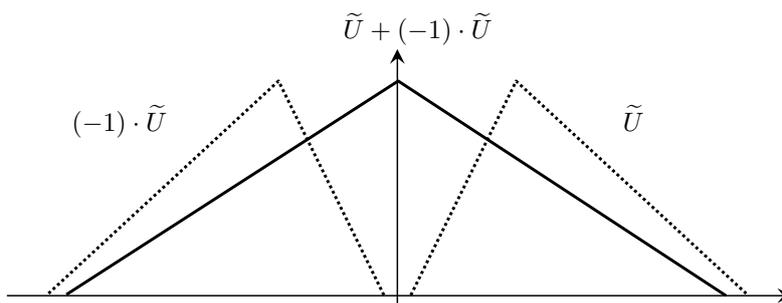


Figura 1.13: Ejemplo de suma de cierto número fuzzy \tilde{U} y el valor $(-1) \cdot \tilde{U}$

La suma usual de números fuzzy y el producto por un escalar desempeñan un papel central en el análisis estadístico de datos fuzzy. Esta aritmética depende básicamente de las ‘fronteras’ de los intervalos α -nivel que intervienen en las operaciones, es decir, de la caracterización inf/sup a la que haremos referencia más adelante.

Por otro lado, la ‘localización’ y la ‘forma’ (o ‘imprecisión’ o ‘extensión’) de un número fuzzy son dos rasgos clave en la caracterización de ese valor, como se formalizará después mediante la caracterización mid/spr.

El objetivo de la subsección siguiente es establecer una conexión entre dos vías de caracterización de los números fuzzy: las funciones que los definen y las funciones soporte. Esta conexión permitirá reformular la clase de los números fuzzy con la aritmética usual dentro de un espacio de Hilbert de funciones.

La función soporte: una caracterización funcional alternativa de los números fuzzy

Como hemos señalado anteriormente, la aritmética usual entre los elementos de $\mathcal{F}_c(\mathbb{R})$ depende de las fronteras de sus α -niveles. Por ello es interesante establecer una conexión entre $\mathcal{F}_c(\mathbb{R})$ y \mathcal{H} el espacio de las funciones reales tipo L^2 definidas sobre $\{-1, 1\} \times (0, 1]$ con respecto a la medida de probabilidad uniforme en $\{-1, 1\}$ (i.e., la que asigna pesos igual a $1/2$ a -1 y 1), y la medida de Lebesgue sobre $(0, 1]$, que denotaremos por λ_1 y λ , respectivamente. Esta conexión va a realizarse a través de la llamada función soporte que se formaliza como sigue:

Definición 1.5.3. La *función soporte (de Minkowski)* de un número fuzzy $\tilde{U} \in \mathcal{F}_c(\mathbb{R})$ (ver Puri y Ralescu, 1985) extiende nivel-a-nivel la noción de función soporte de un intervalo (ver, por ejemplo, Castaing y Valadier, 1977, para la definición más general de función soporte de un conjunto) y viene dada por la aplicación $s_{\tilde{U}} : \{-1, 1\} \times (0, 1] \rightarrow \mathbb{R}$ definida de forma que

$$s_{\tilde{U}}(u, \alpha) = \sup_{v \in \tilde{U}_\alpha} \langle u, v \rangle$$

para $u \in \{-1, 1\}, \alpha \in (0, 1]$, y con $\langle \cdot, \cdot \rangle$ el producto escalar en \mathbb{R} .

Es importante observar que cualquiera que sea $\alpha \in (0, 1]$ y $\tilde{U} \in \mathcal{F}_c(\mathbb{R})$ se cumple que

$$s_{\tilde{U}}(-1, \alpha) = -\inf \tilde{U}_\alpha, \quad s_{\tilde{U}}(1, \alpha) = \sup \tilde{U}_\alpha.$$

Si denotamos por $\mathcal{F}_c^2(\mathbb{R}) = \{\tilde{U} \in \mathcal{F}_c(\mathbb{R}) : s_{\tilde{U}} \in \mathcal{H}\}$, $(\mathcal{F}_c^2(\mathbb{R}), +, \cdot)$ tiene también estructura semilineal. De hecho, la función soporte satisface la propiedad siguiente que establece una correspondencia entre la aritmética usual en $\mathcal{F}_c^2(\mathbb{R})$ y la funcional en \mathcal{H} (ver, por ejemplo, Puri y Ralescu, 1985, Diamond y Kloeden, 1999, Körner, 1997, Trutschnig *et al.*, 2009):

Proposición 1.5.1. La aplicación $s : \mathcal{F}_c^2(\mathbb{R}) \rightarrow \mathcal{H}$, tal que $s(\tilde{U}) = s_{\tilde{U}}$ para todo $\tilde{U} \in \mathcal{F}_c^2(\mathbb{R})$, respeta la estructura semilineal de $\mathcal{F}_c^2(\mathbb{R})$, y

$$s(\tilde{U} + \tilde{V}) = s(\tilde{U}) + s(\tilde{V}), \quad s(\gamma \cdot \tilde{U}) = \gamma \cdot s(\tilde{U})$$

cualesquiera que sean $\tilde{U}, \tilde{V} \in \mathcal{F}_c^2(\mathbb{R})$ y $\gamma \in [0, +\infty)$.

Además (ver, por ejemplo, Körner, 1997):

Proposición 1.5.2. Para cualquier función $f \in \mathcal{H}$ que es continua, sublineal (es decir, positivamente homogénea -equivariante por escala positiva- y subaditiva) y no creciente, existe un único número fuzzy $\tilde{U} \in \mathcal{F}_c^2(\mathbb{R})$ tal que $f = s_{\tilde{U}}$ con

$$\tilde{U}_\alpha = \{x \in \mathbb{R} : \langle u, x \rangle \leq f(u, \alpha) \text{ para todo } u \in \mathbb{R}\}.$$

Como recordaremos más tarde, la función soporte va a permitir una serie de identificaciones entre números fuzzy y valores funcionales, con sus correspondientes aritméticas y métricas, que ayudan a su vez a determinar de forma más específica la estructura semilineal del espacio de los números fuzzy con su aritmética usual.

Se consideran a continuación dos caracterizaciones alternativas de los números fuzzy, que guardan una relación biunívoca con su función soporte.

**Las representaciones inf/sup, mid/spr
y mid₁/lspr/rspr de un número fuzzy**

Si consideramos el espacio de los números fuzzy con función soporte perteneciente a \mathcal{H} , $\mathcal{F}_c^2(\mathbb{R})$, se cumple como ya se ha señalado que:

$$s_{\tilde{U}}(-1, \alpha) = -\inf \tilde{U}_\alpha, \quad s_{\tilde{U}}(1, \alpha) = \sup \tilde{U}_\alpha.$$

Además:

Definición 1.5.4. *La función soporte de $\tilde{U} \in \mathcal{F}_c^2(\mathbb{R})$ puede expresarse siguiendo la llamada **descomposición ortogonal mid/spr**, que viene dada por*

$$s_{\tilde{U}} = \text{mid } s_{\tilde{U}} + \text{spr } s_{\tilde{U}}$$

donde $\text{mid } s_{\tilde{U}} : \{-1, 1\} \times (0, 1] \rightarrow \mathbb{R}$ y $\text{spr } s_{\tilde{U}} : \{-1, 1\} \times (0, 1] \rightarrow [0, \infty)$ son funciones tales que

$$\begin{aligned} \text{mid } s_{\tilde{U}}(1, \alpha) &= -\text{mid } s_{\tilde{U}}(-1, \alpha) = \frac{s_{\tilde{U}}(1, \alpha) - s_{\tilde{U}}(-1, \alpha)}{2} \\ &= \text{punto medio (mid-point) de } \tilde{U}_\alpha = \text{mid } \tilde{U}_\alpha, \end{aligned}$$

y

$$\begin{aligned} \text{spr } s_{\tilde{U}}(1, \alpha) &= \text{spr } s_{\tilde{U}}(-1, \alpha) = \frac{s_{\tilde{U}}(1, \alpha) + s_{\tilde{U}}(-1, \alpha)}{2} \\ &= \text{semiamplitud (spread) de } \tilde{U}_\alpha = \text{spr } \tilde{U}_\alpha. \end{aligned}$$

Por otro lado, la función soporte de $\tilde{U} \in \mathcal{F}_c^2(\mathbb{R})$ puede expresarse siguiendo la **descomposición mid₁/lspr/rspr**, que viene dada por

$$s_{\tilde{U}}(u) = \text{mid}_1 s_{\tilde{U}}(u) + v_u \text{lspr } s_{\tilde{U}}(u) + (1 - v_u) \text{rspr } s_{\tilde{U}}(u)$$

donde, para todo $u \in \{-1, 1\}$, $\alpha \in (0, 1]$, se denota $v_u = (1 - u)/2$ y

$$\text{mid}_1 s_{\tilde{U}}(u) = \text{mid } s_{\tilde{U}}(u, 1) = \frac{s_{\tilde{U}}(u, 1) - s_{\tilde{U}}(-u, 1)}{2},$$

$$\text{lspr } s_{\tilde{U}}(u, \alpha) = \text{rspr } s_{\tilde{U}}(-u, \alpha) = \frac{s_{\tilde{U}}(u, 1) - s_{\tilde{U}}(-u, 1)}{2} + s_{\tilde{U}}(-u, \alpha).$$

La función mid (y, en particular, mid₁) de un número fuzzy puede interpretarse como una especie de cuantificación funcional de su ‘localización’ y las funciones spr, lspr y rspr pueden interpretarse como una especie de cuantificación funcional de su ‘forma’, y satisfacen las siguientes propiedades:

Proposición 1.5.3. *Si $\tilde{U} \in \mathcal{F}_c^2(\mathbb{R})$, se cumple que:*

- i) $\text{mid } s_{\tilde{v}} \in \mathcal{H}$, $\text{spr } s_{\tilde{v}} \in \mathcal{H}$, $\text{lspr } s_{\tilde{v}} \in \mathcal{H}$, $\text{rspr } s_{\tilde{v}} \in \mathcal{H}$, siendo $\text{mid } s_{\tilde{v}}(\cdot, \alpha)$ una función impar, $\text{spr } s_{\tilde{v}}(\cdot, \alpha)$ una función par, y $\text{lspr } s_{\tilde{v}}$ simétrica respecto de $\text{rspr } s_{\tilde{v}}$, cualquiera que sea $\alpha \in (0, 1]$;
- ii) $\text{mid } s_{\tilde{u}+\tilde{v}} = \text{mid } s_{\tilde{u}} + \text{mid } s_{\tilde{v}}$; $\text{spr } s_{\tilde{u}+\tilde{v}} = \text{spr } s_{\tilde{u}} + \text{spr } s_{\tilde{v}}$;
 $\text{lspr } s_{\tilde{u}+\tilde{v}} = \text{lspr } s_{\tilde{u}} + \text{lspr } s_{\tilde{v}}$; $\text{rspr } s_{\tilde{u}+\tilde{v}} = \text{rspr } s_{\tilde{u}} + \text{rspr } s_{\tilde{v}}$;
- iii) $\text{mid } s_{\gamma \cdot \tilde{u}} = \gamma \cdot \text{mid } s_{\tilde{u}}$; $\text{spr } s_{\gamma \cdot \tilde{u}} = |\gamma| \cdot \text{spr } s_{\tilde{u}}$;
 $\text{lspr } s_{\gamma \cdot \tilde{u}} = \begin{cases} |\gamma| \cdot \text{lspr } s_{\tilde{u}} & \text{si } \gamma \geq 0 \\ |\gamma| \cdot \text{rspr } s_{\tilde{u}} & \text{si } \gamma < 0 \end{cases}$; $\text{rspr } s_{\gamma \cdot \tilde{u}} = \begin{cases} |\gamma| \cdot \text{rspr } s_{\tilde{u}} & \text{si } \gamma \geq 0 \\ |\gamma| \cdot \text{lspr } s_{\tilde{u}} & \text{si } \gamma < 0 \end{cases}$

Como consecuencia, cualquier número fuzzy $\tilde{U} \in \mathcal{F}_c^2(\mathbb{R})$ puede caracterizarse por su **representación inf/sup**, $\{(\text{inf } \tilde{U}_\alpha, \text{sup } \tilde{U}_\alpha)\}_{\alpha \in (0,1]}$.

Alternativamente, cualquier número fuzzy $\tilde{U} \in \mathcal{F}_c^2(\mathbb{R})$ puede caracterizarse por su **representación mid/spr**, $\{(\text{mid } \tilde{U}_\alpha, \text{spr } \tilde{U}_\alpha)\}_{\alpha \in (0,1]}$, o por su **representación mid₁/lspr/rspr**, $\{(\text{mid}_1 \tilde{U}, \text{lspr } \tilde{U}_\alpha, \text{rspr } \tilde{U}_\alpha)\}_{\alpha \in (0,1]}$

Métricas en $\mathcal{F}_c^2(\mathbb{R})$ basadas en representaciones de números fuzzy.

Identificación de números fuzzy con valores funcionales

Otro de los rasgos diferenciadores más importantes del empleo de la escala de números fuzzy frente a las escalas Likert reside en la posibilidad de establecer distancias entre los datos con misión e interpretación similares a las distancias entre valores numéricos, cosa que no ocurría cuando se manejaban datos categóricos o codificaciones enteras de las respuestas Likert.

En la cuantificación de la distancia entre dos valores de conjunto (entre ellos, valores de intervalo) es usual considerar la distancia más corta entre dos puntos de los mismos. Sin embargo, si se considerara una distancia definida de esa forma, se ignorarían frecuentemente la localización y la extensión de tales valores (ver, por ejemplo, Grégoire y Bouillot, 1998).

Una métrica que evita este inconveniente y posee propiedades relevantes en la formalización, el tratamiento y la aplicación de los elementos aleatorios con valores de conjunto es la conocida métrica de Hausdorff, que se define en el caso intervalar como la distancia máxima entre uno de los dos conjuntos y el punto más próximo del otro, es decir, dados dos intervalos A y B compactos no vacíos se cuantifica su distancia como $\delta_H(A, B) = \max\{|\text{inf } A - \text{inf } B|, |\text{sup } A - \text{sup } B|\} = |\text{mid } A - \text{mid } B| + |\text{spr } A - \text{spr } B|$.

Al extender esta misma idea al caso de números fuzzy se han propuesto diversas extensiones (ver, por ejemplo, Klement *et al.*, 1986, Diamond y Kloeden, 1999, Kör-

ner, 1997), si bien alguna de ellas (como la basada en la norma del supremo), dan lugar a espacios métricos no separables, lo que dificulta su tratamiento matemático.

Para trabajar con números fuzzy dentro de un contexto estadístico, a menudo es conveniente recurrir a distancias tipo L^2 . La métrica que va a presentarse a continuación surge con una doble misión: definir distancias tipo L^2 en $\mathcal{F}_c^2(\mathbb{R})$ que sean versátiles e intuitivas, e incorporar la posibilidad de ponderar separadamente la diferencia en la localización de los valores de conjunto y la diferencia en la forma/imprecisión/extensión.

Trutschnig *et al.* (2009) han introducido recientemente una métrica en $\mathcal{F}_c^2(\mathbb{R})$ inspirada en la métrica de Bertoluzza *et al.* (1995) y en artículos posteriores como el de Gil *et al.* (2002) (en el que se mencionó por primera vez la caracterización mid/spr de un número fuzzy). Esta métrica se basa en la ponderación del producto escalar de los spr's frente al de los mid's y recurre a la representación mid/spr de las funciones soporte, como sigue:

Definición 1.5.5. Si \tilde{U} y \tilde{V} son números fuzzy pertenecientes a $\mathcal{F}_c^2(\mathbb{R})$, se define la τ -distancia generalizada tipo L^2 entre \tilde{U} y \tilde{V} como:

$$\begin{aligned} D_\tau(\tilde{U}, \tilde{V}) &= \sqrt{\|s_{\tilde{U}} - s_{\tilde{V}}\|_\tau} = \sqrt{\langle s_{\tilde{U}} - s_{\tilde{V}}, s_{\tilde{U}} - s_{\tilde{V}} \rangle_\tau} \\ &= \sqrt{(1 - \tau) \int_{(0,1]} (\text{mid } \tilde{U}_\alpha - \text{mid } \tilde{V}_\alpha)^2 d\alpha + \tau \int_{(0,1]} (\text{spr } \tilde{U}_\alpha - \text{spr } \tilde{V}_\alpha)^2 d\alpha} \end{aligned}$$

donde $\tau \in (0, 1]$ es un valor de ponderación y $\langle \cdot, \cdot \rangle_\tau$ y $\|\cdot\|_\tau$ son un producto escalar y su norma asociada definidas en \mathcal{H} tales que:

$$\langle f, g \rangle_\tau = (1 - \tau)[\text{mid } f, \text{mid } g] + \tau[\text{spr } f, \text{spr } g]$$

con

$$[f, g] = \int_{(0,1]} f(u, \alpha)g(u, \alpha) d\alpha.$$

La completitud y separabilidad del espacio $\mathcal{F}_c^2(\mathbb{R})$ al dotarse de esta métrica es conocida por estudios anteriores (ver Körner, 1997, y Trutschnig *et al.*, 2009).

Observación 1.5.3. El papel que juega el peso τ en la definición de la métrica D_τ es el de ponderar la relevancia o influencia de la diferencia en imprecisión/extensión/forma de los valores de conjunto frente a la diferencia en su localización, por lo que parece natural considerar que $\tau \in (0, 1]$. También debe señalarse que la medida de Lebesgue en $(0, 1]$ podría reemplazarse por otra medida que establezca una ponderación de la importancia que quiere asignársele a los distintos α -niveles o grados de compatibilidad de los datos involucrados en la distancia.

Por otro lado, para trabajar con números fuzzy dentro de un contexto probabilístico, resulta útil identificar dichos valores con elementos de ciertos espacios de funciones, haciendo corresponder también las correspondientes aritméticas. Formalmente, esta identificación viene garantizada por el encaje isométrico que establece la función soporte (Bobylev, 1985, Klement *et al.*, 1986, Diamond y Kloeden, 1999, Körner, 1997), según el cual:

Teorema 1.5.1. *La función soporte $s : \mathcal{F}_c^2(\mathbb{R}) \rightarrow \mathcal{H}$ establece un encaje isométrico de $\mathcal{F}_c^2(\mathbb{R})$ (con la aritmética de conjuntos usual y la métrica D_τ) en un cono convexo cerrado del espacio de Hilbert \mathcal{H} (con la aritmética funcional usual y la métrica basada en la τ -norma).*

Como implicación inmediata de este resultado, un número fuzzy puede tratarse como un valor funcional (el asociado a su función soporte) y este tratamiento involucra el paso de la aritmética usual entre números fuzzy a la aritmética funcional y de las métricas introducidas en esta subsección a las normas referidas entre funciones.

1.5.2. Números fuzzy aleatorios

Cuando se consideran datos fuzzy asociados con experimentos aleatorios y se quiere proceder a su tratamiento estadístico, es necesario considerar un modelo matemático idóneo para los mecanismos aleatorios que dan lugar a tales datos.

Los *números fuzzy aleatorios* cumplen este cometido, y su formalización está justificada por diversos argumentos que se exponen a continuación.

Números fuzzy aleatorios.

Definiciones y distribución inducida

Los números fuzzy aleatorios son elementos aleatorios cuyos valores son números fuzzy. Por cubrir la gran mayoría de situaciones reales y por su operatividad matemática, supondremos que tales números fuzzy pertenecen a la clase $\mathcal{F}_c^2(\mathbb{R})$. En realidad, y tras los desarrollos sobre $\mathcal{F}_c^2(\mathbb{R})$ examinados en la sección anterior, es prácticamente inmediato definir un número fuzzy aleatorio sin más que considerar sobre $\mathcal{F}_c^2(\mathbb{R})$ una σ -álgebra oportuna.

Con esta idea en mente, se introdujo la noción de variable aleatoria fuzzy por Puri y Ralescu (1985, 1986) y Klement *et al.* (1986). En este trabajo, al considerar elementos aleatorios con valores en $\mathcal{F}_c^2(\mathbb{R})$, nos referiremos a ellas como números fuzzy aleatorios. Una definición inmediata sería la inducida a partir de la de elemento aleatorio con valores funcionales por medio de la función soporte y las isometrías

establecidas en el Teorema 1.5.1. No obstante, recurriremos aquí en primer lugar a la definición original dada por Puri y Ralescu (1986).

Definición 1.5.6. *Dado un espacio de probabilidad (Ω, \mathcal{A}, P) , se dice que una aplicación $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R})$ es un **número fuzzy aleatorio** (abreviadamente NFA) asociado a ese espacio si para todo $\alpha \in (0, 1]$ la función α -nivel, \mathcal{X}_α , es un conjunto aleatorio compacto y convexo (es decir, medible Borel), con la métrica de Hausdorff o equivalentes sobre el espacio de los conjuntos compactos convexos no vacíos.*

Puede comprobarse que:

Teorema 1.5.2. *Dado un espacio de probabilidad (Ω, \mathcal{A}, P) , y un número fuzzy aleatorio $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R})$, entonces se cumple que:*

- i) \mathcal{X} es medible Borel respecto de \mathcal{A} y la σ -álgebra de Borel generada en $\mathcal{F}_c^2(\mathbb{R})$ por la topología asociada a la métrica D_τ .*
- ii) Para todo $\alpha \in (0, 1]$, las funciones reales $\text{mid } \mathcal{X}_\alpha : \Omega \rightarrow \mathbb{R}$, $\text{spr } \mathcal{X}_\alpha : \Omega \rightarrow [0, +\infty)$ son variables aleatorias.*
- iii) Para todo $\alpha \in (0, 1]$, las funciones reales $\text{mid}_1 \mathcal{X} : \Omega \rightarrow \mathbb{R}$ (que no depende de α), $\text{lspr } \mathcal{X}_\alpha : \Omega \rightarrow [0, +\infty)$ y $\text{rspr } \mathcal{X}_\alpha : \Omega \rightarrow [0, +\infty)$ son variables aleatorias.*
- iv) Para todo $\alpha \in (0, 1]$, las funciones reales $\text{inf } \mathcal{X}_\alpha : \Omega \rightarrow \mathbb{R}$, $\text{sup } \mathcal{X}_\alpha : \Omega \rightarrow \mathbb{R}$ son variables aleatorias.*

Observación 1.5.4. Una de las ventajas más relevantes de este enfoque en el contexto probabilístico es que la medibilidad Borel garantiza que pueda hablarse propiamente de la distribución inducida por un NFA, la independencia estocástica de NFAs, etc., sin tener que introducir nociones distintas de las conocidas sobre espacios generalizados.

Parámetros relevantes de la distribución de un número fuzzy aleatorio

En el análisis de datos fuzzy pueden distinguirse dos tipos de *medidas resumen/parámetros*:

- medidas resumen con valores fuzzy, como son el valor medio o la mediana, en su mayoría medidas de ‘tendencia central’ del conjunto de datos;
- medidas resumen con valores reales o vectoriales, como el índice de diversidad de Gini-Simpson (al que realmente no afecta la naturaleza fuzzy de los datos), el índice de desigualdad hiperbólico de Lubiano (1999) o el D_τ -error cuadrático medio que constituye una medida de la representatividad de la media tipo Aumann.

Como $(\mathcal{F}_c^2(\mathbb{R}), +, \cdot)$ no es un espacio lineal, los conceptos convencionales de esperanza en espacios lineales no son aplicables directamente para NFAs. Sin embargo, la identificación a la que conducen las isometrías del Teorema 1.5.1, así como la definición de NFA en la Definición 1.5.6, permiten inducir una noción del concepto de *esperanza o valor medio* de un NFA que a su vez es equivalente a la esperanza a través de las selecciones de las funciones α -nivel del NFA (o esperanzas de Aumann de esas funciones) que es la más y mejor investigada para este concepto (ver, Puri y Ralescu, 1986).

La media tipo Aumann de un NFA se define como sigue:

Definición 1.5.7. *Dado un espacio de probabilidad (Ω, \mathcal{A}, P) y un NFA asociado \mathcal{X} se define el **valor medio o esperanza** de \mathcal{X} como el número fuzzy $\tilde{E}(\mathcal{X}) \in \mathcal{F}_c^2(\mathbb{R})$, si existe, tal que para todo $\alpha \in (0, 1]$:*

$$\left(\tilde{E}(\mathcal{X})\right)_\alpha = \left\{ \int_{\mathbb{R}} h(\omega) dP(\omega) \text{ para todo } h : \Omega \rightarrow \mathbb{R}, h \in L^1(\Omega, \mathcal{A}, P), h \in \mathcal{X}_\alpha \text{ c.s. } [P] \right\}.$$

Si \mathcal{X} es un NFA tal que $\max\{|\inf \mathcal{X}_\alpha|, |\sup \mathcal{X}_\alpha|\} \in L^1(\Omega, \mathcal{A}, P)$, se satisface para todo $\alpha \in (0, 1]$ que:

$$\begin{aligned} \left(\tilde{E}(\mathcal{X})\right)_\alpha &= [E(\inf \mathcal{X}_\alpha), E(\sup \mathcal{X}_\alpha)] \\ &= [E(\text{mid } \mathcal{X}_\alpha) - E(\text{spr } \mathcal{X}_\alpha), E(\text{mid } \mathcal{X}_\alpha) + E(\text{spr } \mathcal{X}_\alpha)] \\ &= [E(\text{mid}_1 \mathcal{X}) - E(\text{lspr } \mathcal{X}_\alpha), E(\text{mid}_1 \mathcal{X}) + E(\text{rspr } \mathcal{X}_\alpha)]. \end{aligned}$$

Si $s_{\mathcal{X}} \in L^1(\Omega, \mathcal{A}, P)$, se cumple también que $s_{\tilde{E}(\mathcal{X})} = E(s_{\mathcal{X}})$.

Sobre la base de las propiedades de la función soporte y de los elementos aleatorios hilbertianos, el valor medio de un NFA cumple las propiedades usuales de linealidad y es la esperanza tipo Fréchet con respecto a la métrica D_τ , es decir, con independencia del valor de τ considerado, $\tilde{E}(\mathcal{X})$ verifica que:

$$E\left(\left[D_\tau(\mathcal{X}, \tilde{E}(\mathcal{X}))\right]^2\right) = \arg \min_{\tilde{U} \in \mathcal{F}_c^2(\mathbb{R})} E\left(\left[D_\tau(\mathcal{X}, \tilde{U})\right]^2\right),$$

lo que corrobora su papel como medida de tendencia central de la distribución del NFA.

Volviendo al Ejemplo 1.3.1, algunos de los datos obtenidos para las preguntas QF1 y QF2 del cuestionario vienen recogidos en la Tabla 1.1. Se trata de una muestra de 70 individuos de ambos sexos, diferentes edades y diversas profesiones. En la tabla aparecen las respuestas dadas en las dos escalas: Likert y fuzzy. En cuanto a las respuestas fuzzy trapezoidales aparecen, de izquierda a derecha, el extremo inferior del 0-nivel (inf_0), el extremo inferior del 1-nivel (inf_1), el extremo superior del 1-nivel (sup_1) y el extremo superior del 0-nivel (sup_0).

Tabla 1.1: Algunas de las respuestas Likert y trapezoidales a las cuestiones QF1 y QF2 del cuestionario sobre la valoración de los restaurantes

Ind.	QF1. The food is served hot and fresh					QF2. The menu has a good variety of items				
	Likert	ínf ₀	ínf ₁	sup ₁	sup ₀	Likert	ínf ₀	ínf ₁	sup ₁	sup ₀
1	4	70	80	90	100	5	90	95	100	100
2	5	80	90	90	100	5	80	80	80	90
3	4	70	80	80	90	4	80	85	85	90
4	4	70	75	80	80	3	30	35	35	50
5	5	80	95	95	100	4	60	65	70	80
6	4	60	65	75	80	5	70	80	85	90
7	5	90	98	100	100	4	70	75	80	90
8	5	80	90	100	100	5	75	80	90	100
9	4	60	70	80	90	3	30	40	60	70
10	4	60	70	90	100	4	50	60	70	80
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
67	4	75	80	90	100	5	50	80	90	100
68	5	85	90	95	100	4	60	67	73	80
69	4	75	90	90	100	3	50	60	60	60
70	3	37	47	56	62	4	51	62	69	72

El valor medio de todos los datos para las dos preguntas en la escala de números fuzzy son los dos números fuzzy trapezoidales que aparecen representados en la Figura 1.14.

Para la primera pregunta, el valor medio muestral es el número fuzzy trapezoidal caracterizado por la cuaterna (62.25714, 71.75714, 81.87143, 89.52857), mientras que la respuesta media a la segunda pregunta se expresa como (53.08571, 62.32857, 71.94286, 80.81429).

En cuanto a las medidas resumen con valores reales es interesante considerar un índice de diversidad de los valores que servirá para formalizar la impresión intuitiva de que la subjetividad/diversidad es muy superior si se emplean cuestionarios de respuesta libre basados en la escala de números fuzzy, que si se recurre a escalas Likert.

En la cuantificación de la diversidad debe notarse que no se trata de extender un índice para variables aleatorias reales, puesto que los índices de diversidad (o entropía, o cantidad de información) sólo tienen en cuenta la distribución de los valores distintos del elemento aleatorio y no su orden de magnitud. El índice que va

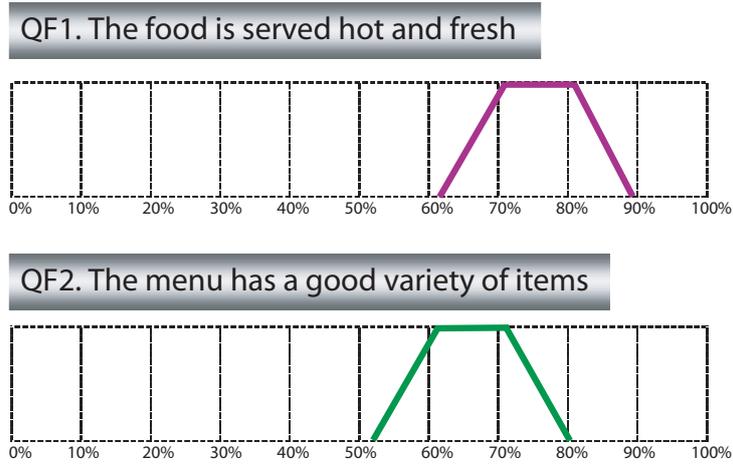


Figura 1.14: Valor medio para las dos preguntas consideradas en el Ejemplo 1.3.1

a considerarse en los estudios del Capítulo 2 es el siguiente:

Definición 1.5.8. *Dado un espacio de probabilidad (Ω, \mathcal{A}, P) y un NFA asociado \mathcal{X} se define el **índice de diversidad de Gini-Simpson** de \mathcal{X} como el valor real:*

$$\mathfrak{D}(\mathcal{X}) = 1 - \sum_{j=1}^m [P(\mathcal{X} = \tilde{x}_j)]^2 = 1 - \sum_{j=1}^m p_j^2,$$

supuesto que \mathcal{X} tiene una distribución $\{p_1, \dots, p_m\}$ (con $p_j \geq 0, p_1 + \dots + p_m = 1$).

Esta definición podría ampliarse al caso en el que el NFA tomara una infinidad de posibles valores distintos, si bien en los estudios de este trabajo al manejar únicamente muestras no será necesario.

Para el Ejemplo 1.3.1, el valor del índice de Gini-Simpson para la pregunta QF1, teniendo en cuenta que de los 70 números fuzzy de la muestra, hay 48 valores distintos, de los cuales 38 se repiten sólo una vez, seis se repiten 2 veces, uno se repite 3 veces, otro valor se repite 5 veces, y dos valores se repiten 6 veces, tenemos que:

$$\begin{aligned} \mathfrak{D}(\mathcal{X}) &= 1 - \sum_{j=1}^{48} p_j^2 = 1 - 38 \cdot \left(\frac{1}{70}\right)^2 - 6 \cdot \left(\frac{1}{35}\right)^2 - \left(\frac{3}{70}\right)^2 - \left(\frac{1}{14}\right)^2 \\ &\quad - 2 \cdot \left(\frac{3}{35}\right)^2 = 0,9657, \end{aligned}$$

y para la pregunta QF2, teniendo en cuenta que de los 70 números fuzzy de la muestra, hay 61 valores distintos, de los cuales 54 se repiten sólo una vez, cinco se

repiten 2 veces y dos se repiten 3 veces, tenemos que:

$$\mathfrak{D}(\mathcal{X}) = 1 - \sum_{j=1}^{61} p_j^2 = 1 - 54 \cdot \left(\frac{1}{70}\right)^2 - 5 \cdot \left(\frac{1}{35}\right)^2 - 2 \cdot \left(\frac{3}{70}\right)^2 = 0,9812.$$

Para las dos preguntas la diversidad es muy grande en la muestra de los 70 encuestados, teniendo en cuenta que el valor máximo que puede alcanzar el índice de Gini-Simpson es $1 - 1/70 = 0,9857$.

En el Capítulo 2, además de considerar el índice de diversidad referido para comparar desde una perspectiva estadística el empleo de cuestionarios basados en escalas Likert frente al de cuestionarios de respuesta libre basados en la escala de números fuzzy, se van a desarrollar comparaciones relativas a la *representatividad de la medida de centralización* considerada: la media tipo Aumann. Dicha representatividad se cuantificará a través del correspondiente error que hace mínimo, es decir:

Definición 1.5.9. *Dado un espacio de probabilidad (Ω, \mathcal{A}, P) y un NFA asociado \mathcal{X} , se define el D_τ -error cuadrático medio de \mathcal{X} respecto del número fuzzy \tilde{U} como el valor real:*

$$\text{ECM}(\mathcal{X}, \tilde{U}) = E \left(\left[D_\tau(\mathcal{X}, \tilde{U}) \right]^2 \right).$$

Es obvio que, cuando un NFA se reduce a una variable aleatoria real, el error anterior se corresponde con el error cuadrático medio.

1.5.3. Una metodología para el tratamiento estadístico de datos fuzzy

Para analizar estadísticamente las respuestas de un cuestionario basado en la escala de números fuzzy puede recurrirse a una metodología desarrollada en los últimos años mayoritariamente por el grupo de investigación SMIRE coordinado por la Universidad de Oviedo (<http://bellman.ciencias.uniovi.es/SMIRE>). Esta metodología tiene en cuenta algunos rasgos distintivos cruciales para el tratamiento de estos datos, como son:

- la falta de linealidad en $(\mathcal{F}_c^2(\mathbb{R}), +, \cdot)$, lo que previene de considerar una extensión general de la diferencia entre números fuzzy que conserve las propiedades elementales de la misma;
- la falta de orden completo universalmente aceptable entre los números fuzzy;
- la carencia de modelos realistas y manejables para las distribuciones de los NFAs;

- la ausencia de Teoremas del Límite Central para NFAs que puedan aplicarse directamente con fines inferenciales; aunque existen algunos de estos resultados según los cuales la distancia normalizada entre la media fuzzy muestral y la poblacional converge en ley a la norma de un elemento aleatorio gaussiano, éste toma valores que a menudo están fuera del cono de los números fuzzy.

Para evitar estos graves inconvenientes, van a ser determinantes el empleo de la distancia D_τ y, para el caso de las inferencias en la existencia de aproximaciones bootstrap de Teoremas del Límite Central para elementos aleatorios con valores en espacios de Hilbert (ver Giné y Zinn, 1990).

Como para el caso de datos numéricos, el objetivo de la Estadística con datos fuzzy será (desde una perspectiva inferencial) extraer conclusiones sobre la distribución de NFAs sobre la base de la información suministrada por muestras de observaciones a partir de ellos.

Buena parte de la metodología estadística desarrollada hasta el momento para datos fuzzy se ha centrado en examinar el *comportamiento de la media fuzzy muestral como estimador de la media fuzzy poblacional* (ver, por ejemplo, Lubiano y Gil, 1999) y en desarrollar *contrastos de hipótesis sobre las medias poblacionales de NFAs* en el caso de una muestra (cf. Montenegro *et al.*, 2004, González-Rodríguez *et al.*, 2006), y dos muestras independientes o dependientes (cf. Montenegro *et al.*, 2001, Montenegro, 2003), y tests ANOVA para varias muestras independientes o dependientes (ver, por ejemplo, Montenegro, 2003, Gil *et al.*, 2006, González-Rodríguez *et al.*, 2011). Recientemente (Sinova *et al.*, 2012a, 2012b) se han llevado a cabo algunos primeros estudios inferenciales sobre la estimación de la mediana de un NFA. También se han establecido algunos resultados metodológicos para la regresión y correlación entre NFAs donde las relaciones consideradas se han basado en la aritmética de números fuzzy (Ferraro *et al.*, 2010, González-Rodríguez *et al.*, 2009).

Capítulo 2

Estudio comparativo entre los cuestionarios basados en escalas Likert y los que utilizan la escala de números fuzzy

Este segundo capítulo constituye un estudio preliminar sobre la comparación del análisis estadístico que puede llevarse a cabo para las respuestas de los cuestionarios, según estén basados en escalas Likert o admitan respuesta libre en la escala de números fuzzy. También se hará una comparación entre el análisis estadístico para respuestas en la escala de números fuzzy en formato libre (el encuestado representa el número fuzzy que él considere para cada pregunta de la encuesta) y la escala de números fuzzy en cuestionarios con respuestas prefijadas (el encuestado elige para cada pregunta un número fuzzy de entre un conjunto de fuzzy prefijados). En realidad, la envergadura de tales comparaciones estimamos que superaría con creces el alcance, nivel y características de un trabajo fin de máster, tratándose más bien de un tema que correspondería a un proyecto de investigación.

A la hora de realizar un estudio comparativo entre las respuestas que utilizan la escala Likert y las respuestas que utilizan la escala fuzzy, es difícil y a menudo inviable extraer conclusiones de índole general. Sin embargo, el desarrollo de comparaciones empíricas que recurran a simulaciones de situaciones realistas imitando el comportamiento humano parece factible. En este sentido surgen varios problemas claves:

- cómo realizar simulaciones realistas de las respuestas que daría un encuestado si simultáneamente se le aplicaran cuestionarios en escala Likert y cuestionarios de respuesta libre en escala fuzzy;

- cómo simular la conversión fuzzy que a veces se realiza de las respuestas en escala Likert;
- cómo simular las respuestas que daría un encuestado si simultáneamente se le aplicaran cuestionarios fuzzy en formato libre y prefijado;
- y cómo queremos comparar las series de respuestas simuladas.

Para abordar los tres primeros problemas, van a realizarse simulaciones de respuestas a un cuestionario de respuesta libre basado en la escala de números fuzzy. El procedimiento seguido para llevar a cabo estas simulaciones se abordará más adelante. Posteriormente, estos números van a asociarse con respuestas enteras (‘Likertización’) y también con respuestas fuzzy prefijadas: para realizar esa asociación fuzzy-fuzzy utilizaremos dos métodos distintos: pasaremos de las respuestas Likert a las respuestas fuzzy prefijadas, imitando así el procedimiento que emplean algunos encuestadores para mejorar algo el limitado análisis estadístico que se puede hacer a las respuestas en escala Likert, y también pasaremos de las respuestas libres fuzzy a las respuestas fuzzy prefijadas. Más avanzado el capítulo explicaremos cómo hemos llevado a cabo estas asociaciones.

En cuanto al último problema, como herramientas estadísticas de comparación se emplearán:

- el índice de diversidad de Gini-Simpson;
- el error cuadrático medio respecto de la media, y su D_τ -extensión al caso fuzzy.

Se trata sólo de un primer paso en la comparación de los cuestionarios. De este primer paso se esperan obtener ciertas conclusiones que permitan verificar la adecuación del procedimiento de simulación y, si ello es posible, realizar algunas propuestas para su mejora. Así mismo, en los problemas abiertos se sugerirán algunas de las herramientas estadísticas adicionales para efectuar la comparación.

2.1. Preliminares del estudio: simulaciones de números fuzzy, ‘Likertización’ y conversión fuzzy

Como ya se ha indicado en el Capítulo 1, el número de respuestas o etiquetas lingüísticas posibles para cada pregunta de un cuestionario que emplea la escala Likert varía normalmente entre cuatro y siete, y rara vez supera dicha cifra. Además, existen argumentos psicométricos y estadísticos que sustentan esta consideración (ver, por ejemplo, Lozano *et al.*, 2008). Por este motivo, hemos realizado simulaciones

de respuesta a un cuestionario libre fuzzy que posteriormente serán identificadas con números enteros que irán desde 1 hasta k , con $k \in \{4, 5, 6, 7\}$, y con números fuzzy prefijados.

2.1.1. Simulación de números fuzzy

Para cada individuo se ha simulado la respuesta que daría a una pregunta de un cuestionario de respuesta libre fuzzy, es decir, un número fuzzy que hemos supuesto trapezoidal por su simplicidad en el posterior tratamiento, facilidad de interpretación y adaptabilidad para describir de forma bastante fiel el tipo de valoraciones que se consideran.

Como se muestra en la Figura 2.1, todo número fuzzy trapezoidal \tilde{U} queda caracterizado por la cuaterna de valores $(\text{inf } \tilde{U}_0, \text{inf } \tilde{U}_1, \text{sup } \tilde{U}_1, \text{sup } \tilde{U}_0)$. Alternativamente, también puede caracterizarse por la representación $(\text{mid } \tilde{U}_1, \text{spr } \tilde{U}_1, \text{lspr } \tilde{U}_0, \text{rspr } \tilde{U}_0)$.

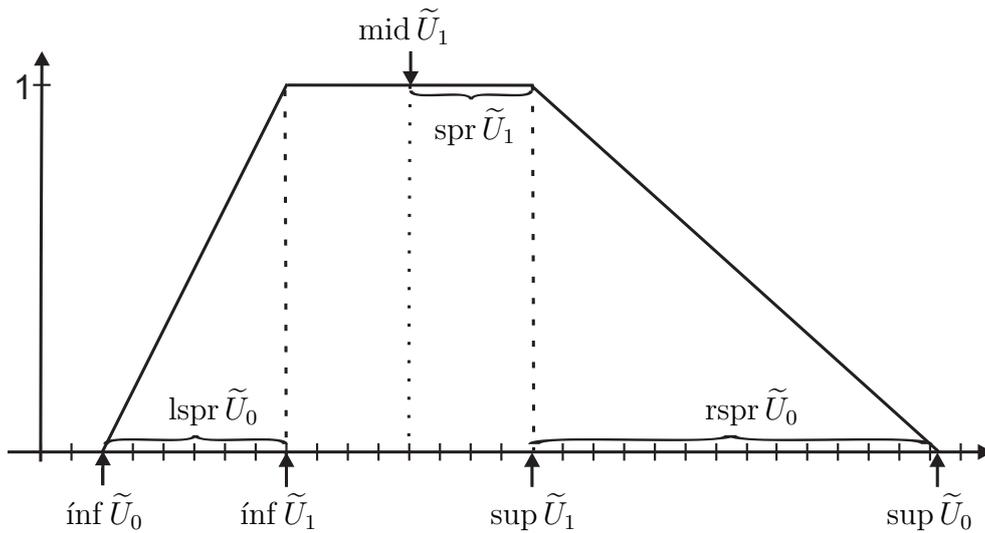


Figura 2.1: Caracterizaciones del número fuzzy trapezoidal

Teniendo en cuenta estas caracterizaciones, la forma de generar los números fuzzy trapezoidales ha sido la siguiente:

- En primer lugar, hemos simulado el mid_1 de cada número fuzzy trapezoidal. Para ello, hemos recurrido a la extensa clase paramétrica de las distribuciones beta, que nos permiten abarcar un gran número de situaciones diferentes. Puesto que éstas sólo toman valores en el intervalo $[0, 1]$, se han reescalado y trasladado los mid_1 generados con esta distribución al intervalo $[0, 100]$, lo que equivale a considerar una simulación a partir de una distribución $100 \cdot \text{beta}$.

En la Figura 2.2 pueden visualizarse las densidades consideradas. Entre ellas aparece la distribución uniforme, una distribución simétrica con concentración mayoritaria en valores centrales, una distribución simétrica con concentración mayoritaria repartida hacia los valores extremos, una distribución marcada-mente asimétrica y dos distribuciones asimétricas.

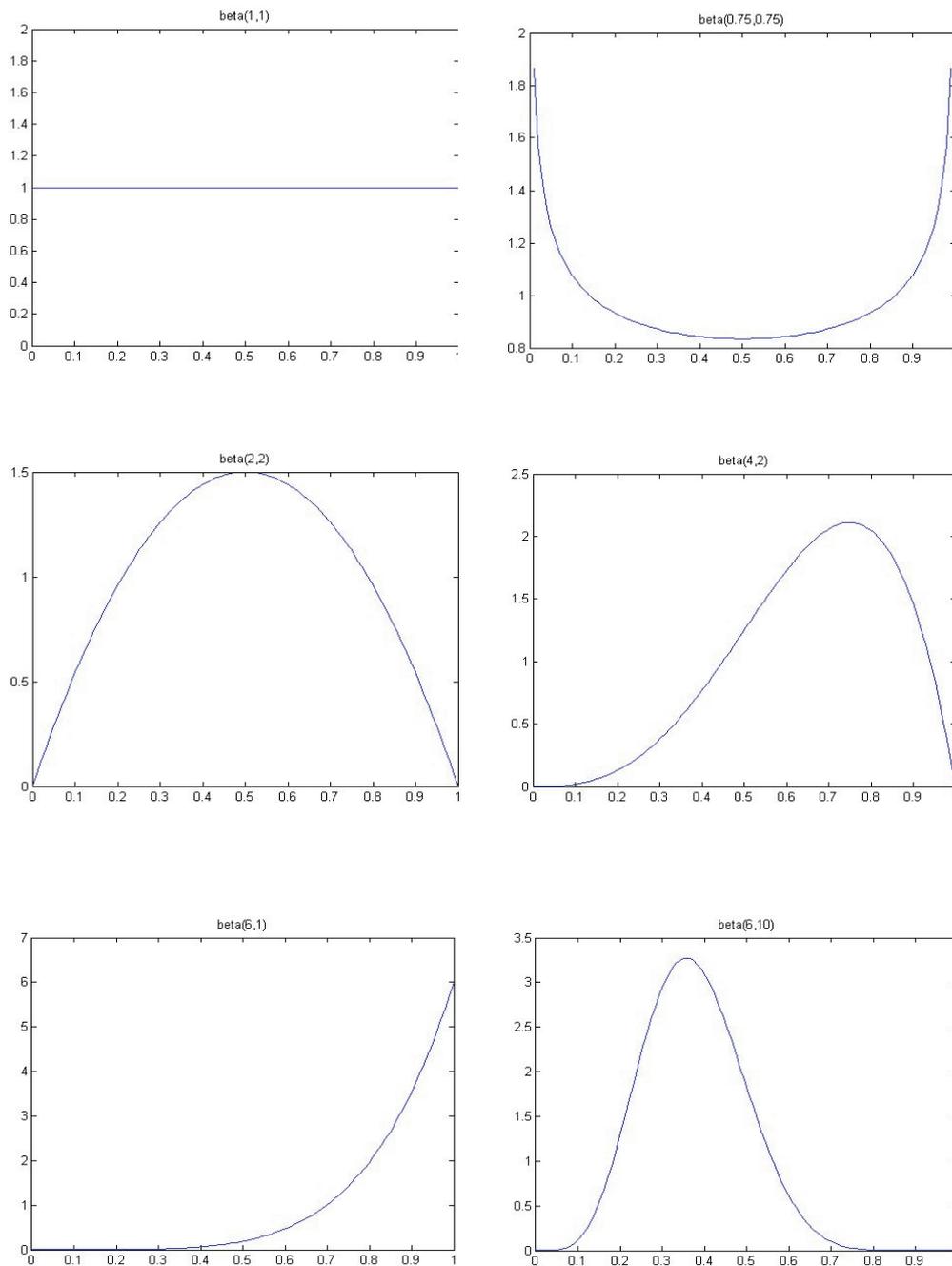


Figura 2.2: Densidades de las beta involucradas en las simulaciones

- Una vez obtenidos los mid_1 , los spr_1 de cada número fuzzy trapezoidal se han simulado de la siguiente forma:

- Si $\text{mid}_1 \in [25, 75]$, el valor de spr_1 se genera a partir de la distribución $\text{Exp}(2)$.
- Si $\text{mid}_1 < 25$, generamos el valor de spr_1 a partir de la distribución $\text{Exp}\left(1 + \frac{\text{mid}_1}{25}\right)$.
- Si $\text{mid}_1 > 75$, generamos el valor de spr_1 a partir de la distribución $\text{Exp}\left(2 - \frac{\text{mid}_1 - 75}{25}\right)$.

- A continuación generamos los lspr_0 de cada número fuzzy de la siguiente forma:

- Si $\text{inf}_1 \in [25, 100]$, generamos el lspr_0 a partir de la distribución $\gamma(4, 1)$.
- Si $\text{inf}_1 < 25$, generamos el lspr_0 a partir de la distribución $\gamma\left(4, 1 + \frac{\text{mid}_1}{25}\right)$.

- Por último, simulamos los rspr_0 de cada número fuzzy trapezoidal como sigue:

- Si $\text{sup}_1 \in [0, 75]$, generamos el rspr_0 a partir de la distribución $\gamma(4, 1)$.
- Si $\text{sup}_1 > 75$, generamos el valor de rspr_0 a partir de la distribución $\gamma\left(4, 2 - \frac{\text{mid}_1 - 75}{25}\right)$.

De esta forma, obtenemos los cuatro valores que caracterizan a un número fuzzy trapezoidal en el intervalo $[0, 100]$ (si el ínfimo del 0-nivel o el ínfimo del 1-nivel se encontraban por debajo de 0 les dábamos el valor 0 automáticamente. De forma análoga, si el supremo del 1-nivel o el supremo del 0-nivel resultaban estar por encima de 100, les asignábamos el valor 100). Para poder asociar posteriormente a cada número fuzzy generado un número entero de la escala Likert desde el 1 hasta el k , con $k \in \{4, 5, 6, 7\}$, reescalamos y trasladamos los cuatro números generados en el intervalo $[0, 100]$ que caracterizan a cada número fuzzy trapezoidal al intervalo $[1, k]$.

2.1.2. ‘Likertización’ de los números fuzzy simulados

Puesto que tratamos de hacer una comparación empírica entre el uso de escalas Likert frente al uso de escalas de valoración fuzzy en cuestionarios, una vez simuladas las respuestas de números fuzzy en formato libre, asociamos a cada una de ellas un número entero del 1 al k (con $k \in \{4, 5, 6, 7\}$) para pasar así a la escala Likert.

Como criterio de Likertización se ha hallado la distancia D_τ de cada uno de los n números fuzzy trapezoidales generados (correspondientes a los n individuos de la muestra) a los números enteros $1, 2, \dots, k$, con $k \in \{4, 5, 6, 7\}$. Así, se ha asociado a cada número fuzzy \tilde{U} el número entero $\iota(\tilde{U})$ cuya distancia con él resultase mínima, es decir:

$$\iota(\tilde{U}) = \arg \min_{j \in \{1, \dots, k\}} D_\tau(\tilde{U}, \mathbf{1}_{\{j\}})$$

Como se ha señalado, la distancia utilizada ha sido la τ -distancia generalizada tipo L^2 definida en el Capítulo 1, con $\tau \in \{0,0476, 0,0909, 0,2, 0,25, 0,5\}$.

Observación 2.1.1. En caso de empate entre las distancias mínimas se asoció al número fuzzy trapezoidal uno de los dos enteros implicados elegido al azar.

2.1.3. Conversión fuzzy de las respuestas Likertizadas

Puesto que también queremos comparar el uso de la escala libre de números fuzzy frente a la conversión fuzzy que a veces se realiza de las respuestas Likert en cuestionarios, una vez obtenidos los enteros $\iota(\tilde{U})$ asociados a los números fuzzy simulados ('Likertización' de los números fuzzy), identificamos cada uno ellos con un número fuzzy de entre k números fuzzy prefijados (con $k \in \{4, 5, 6, 7\}$).

Para cada valor de k , hemos considerado dos conjuntos diferentes C_1 y C_2 con k números fuzzy trapezoidales prefijados cada uno, que hemos recogido de los estudios llevados a cabo en diferentes artículos. Mientras los k números fuzzy que forman el conjunto C_1 están dispuestos de forma simétrica respecto a la posición intermedia, los que forman el C_2 no verifican esta restricción. Más adelante detallaremos cuáles han sido estos conjuntos elegidos para cada valor de k . Además, todos los números fuzzy prefijados han sido reescalados y trasladados al intervalo $[1, k]$ para poder realizar la conversión.

Así, hemos hallado la distancia D_τ de cada uno de los n números enteros de la escala Likert a los k números fuzzy prefijados de cada uno de los dos conjuntos. Posteriormente, hemos asociado a cada número entero $\iota(\tilde{U})$ el número fuzzy trapezoidal $C_1(\iota(\tilde{U}))$ del conjunto C_1 cuya distancia con él resultase mínima, y análogamente, asociamos el número fuzzy $C_2(\iota(\tilde{U}))$ del conjunto C_2 con distancia mínima. Al igual que para el proceso de Likertización, la métrica empleada ha sido la τ -distancia generalizada tipo L^2 , con $\tau \in \{0,0476, 0,0909, 0,2, 0,25, 0,5\}$.

Observación 2.1.2. En caso de empate entre las distancias mínimas volvemos a asociar al número entero de la escala Likert uno de los dos número fuzzy trapezoidales implicados elegido al azar.

2.1.4. Conversión de las respuestas fuzzy libres a las respuestas fuzzy prefijadas

Para comparar el empleo en cuestionarios de opinión de la escala de números fuzzy en formato libre frente a la escala de números fuzzy con formato de respuesta prefijado, además de realizar una conversión fuzzy de las respuestas Likertizadas (como acabamos de ver en el apartado anterior), también realizamos una conversión fuzzy a partir de los fuzzy simulados, es decir, identificamos cada uno de los n números fuzzy trapezoidales simulados con otro número fuzzy de un conjunto de k números fuzzy prefijados.

Al igual que para el caso anterior, hemos considerado los dos mismos conjuntos C_1 y C_2 de k números fuzzy prefijados cada uno, y hallado la distancia D_τ de cada uno de los n números fuzzy simulados a los k números fuzzy de cada conjunto. Posteriormente asociamos a cada número fuzzy \tilde{U} de la escala libre los números fuzzy $C_1(\tilde{U})$ y $C_2(\tilde{U})$ de los conjuntos C_1 y C_2 respectivamente, cuyas distancias con él fuesen mínimas dentro de cada conjunto.

Observación 2.1.3. En caso de empate entre las distancias mínimas volvemos a resolver de la misma manera, eligiendo al azar entre los dos números fuzzy implicados.

2.1.5. Transformaciones necesarias de las muestras fuzzy y Likert para su posterior comparación

Recordemos que disponemos en total de seis muestras de tamaño n para cada valor de k formadas cada una de ellas por números fuzzy trapezoidales, números enteros, y números fuzzy trapezoidales de los conjuntos C_1 y C_2 . Las muestras son las siguientes:

- Muestra 1: formada por n números fuzzy trapezoidales \tilde{U}_i con $i = 1, \dots, n$, que simulan las respuestas de n individuos a un cuestionario que utiliza la escala de números fuzzy con formato de respuesta libre.
- Muestra 2: formada por n números enteros $\iota(\tilde{U}_i)$ con valores del 1 al k asociados a los números fuzzy \tilde{U}_i .
- Muestra 3: formada por n números fuzzy $C_1(\iota(\tilde{U}_i))$ del conjunto C_1 asociados a los números enteros $\iota(\tilde{U}_i)$.
- Muestra 4: formada por n números fuzzy $C_2(\iota(\tilde{U}_i))$ del conjunto C_2 asociados a los números enteros $\iota(\tilde{U}_i)$.

- Muestra 5: formada por n números fuzzy $C_1(\tilde{U}_i)$ del conjunto C_1 asociados a los números fuzzy \tilde{U}_i .
- Muestra 6: formada por n números fuzzy $C_2(\tilde{U}_i)$ del conjunto C_2 asociados a los números fuzzy \tilde{U}_i .

Con el propósito de poder comparar de forma conveniente estas muestras reescalaremos todos los valores para trabajar en condiciones similares de dispersión. Para ello realizaremos la siguiente transformación para cada uno de los n valores de las seis muestras:

$$\text{nuevo valor} = \frac{\text{valor en la muestra} - 1}{\text{desviación típica}} + 1$$

1. En el caso de la muestra de números fuzzy libres tomados en el intervalo $[1, k]$, hemos considerado como desviación típica la correspondiente a una distribución uniforme continua en el intervalo $[1, k]$, es decir:

$$\sqrt{\frac{(k-1)^2}{12}}$$

2. Para la muestra Likert codificada en los valores enteros del 1 al k , se toma como desviación típica la de la siguiente distribución discreta:

$$p_j = \begin{cases} \frac{1}{2k-2} & \text{si } j = 1 \\ \frac{1}{k-1} & \text{si } j = 2, \dots, k-1 \\ \frac{1}{2k-2} & \text{si } j = k \end{cases}$$

3. Para las muestras formadas por los números fuzzy prefijados de los conjuntos C_1 y C_2 , se ha considerado la siguiente distribución de probabilidad y se ha hallado la desviación típica utilizando D_τ cuando $\tau = 0$:

$$p_j = \begin{cases} \frac{1}{2k-2} & \text{para } C_i^1 \text{ con } i = 1, 2 \\ \frac{1}{k-1} & \text{para } C_i^j \text{ si } j = 2, \dots, k-1 \text{ con } i = 1, 2 \\ \frac{1}{2k-2} & \text{para } C_i^k \text{ con } i = 1, 2 \end{cases}$$

Observación 2.1.4. En la práctica, la percepción y valoración de las respuestas extremas no es la misma que la de las intermedias. Este hecho queda reflejado en la ‘Likertización’ de los números fuzzy. Más concretamente, la codificación entera que hemos empleado tiende a asociar enteros ‘próximos’ en sentido bilateral, a excepción de en el caso de los valores extremos en que la asociación sigue una pauta de aproximación unilateral. Lo mismo ocurre con la conversión fuzzy que hemos realizado. Es por este motivo por el que al hallar la desviación típica hemos considerado muestras cuyos valores extremos tienen la mitad del peso que tienen el resto de valores de la muestra.

Observación 2.1.5. Para hallar las desviaciones típicas en el caso de muestras fuzzy prefijadas, hemos utilizado D_τ con $\tau = 0$, ya que así sólo intervienen las localizaciones y no la forma de estos números fuzzy, equiparándose así a una escala Likert.

Observación 2.1.6. En la transformación que realizamos para cada uno de los valores de la muestra, se le suma 1 al cociente para que así las muestras empiecen en el 1 y no en el 0. Asimismo, restamos 1 a cada valor en el cociente para que el valor final no sea superior a k .

2.1.6. Herramientas de comparación: índice de diversidad y cuantificación del error cuadrático medio

Una vez considerados los datos fuzzy simulados y asociado a los mismos el entero con el que presumiblemente el encuestado identificaría tal dato si en lugar de la escala de números fuzzy se le pidiera optar por la codificación usual de una escala Likert, se procederá a examinar y comparar los resultados de ambas escalas. Además, puesto que hemos identificado cada número fuzzy simulado con otro número fuzzy de un conjunto de números fuzzy prefijados, también podremos comparar el uso de la escala de números fuzzy en cuestionarios con dos tipos de formato: libre (el encuestado representa el número fuzzy que él considere más oportuno dentro de un intervalo) y prefijado (el encuestado elige uno de los k números fuzzy prefijados como respuesta a cada pregunta del cuestionario).

Como herramientas de comparación entre los dos tipos de escala y entre los dos tipos de formato de encuesta utilizaremos las dos medidas siguientes:

- El índice de diversidad de Gini-Simpson.
- El error cuadrático medio respecto de la media. En el caso fuzzy emplearemos su D_τ -extensión.

Así, para una muestra $(\omega_1, \dots, \omega_n)$ para la que $\tilde{U}_i = \mathcal{X}(\omega_i)$ las medidas anteriores vendrán dadas de la forma siguiente:

El *índice de diversidad de Gini-Simpson muestral del NFA \mathcal{X}* será:

$$\mathfrak{D}\mathbb{F} = \widehat{\mathfrak{D}(\mathcal{X})} = 1 - \frac{1}{n^2} \sum_{j=1}^m [\text{card } \mathcal{X}^{-1}(\{\tilde{x}_j\})]^2$$

(con $\mathcal{X}^{-1}(\{\tilde{x}_j\}) = \{\omega_i : \mathcal{X}(\omega_i) = \tilde{x}_j\}$).

El *D_τ -error cuadrático medio muestral del NFA \mathcal{X} respecto de su media* será:

$$\text{ECMF} = \text{ECM}[\widehat{\mathcal{X}}, \tilde{E}(\mathcal{X})] = \frac{1}{n} \sum_{i=1}^n [D_\tau(\tilde{U}_i, \bar{\mathcal{X}})]^2,$$

con $\bar{\mathcal{X}} = \frac{1}{n} \cdot (\tilde{U}_1 + \dots + \tilde{U}_n)$. Al ser

$$\text{mid}(\tilde{U}_i)_\alpha = \frac{u_i^{(2)} + u_i^{(3)}}{2} \cdot \alpha + \frac{u_i^{(1)} + u_i^{(4)}}{2} \cdot (1 - \alpha),$$

$$\text{spr}(\tilde{U}_i)_\alpha = \frac{u_i^{(3)} - u_i^{(2)}}{2} \cdot \alpha + \frac{u_i^{(4)} - u_i^{(1)}}{2} \cdot (1 - \alpha),$$

realizando las operaciones pertinentes se llega a que:

$$\begin{aligned} [D_\tau(\tilde{U}_i, \bar{\mathcal{X}})]^2 &= \frac{(1 - \tau)}{12} \left[\left((u_i^{(2)} - \bar{u}^{(2)}) + (u_i^{(3)} - \bar{u}^{(3)}) \right)^2 + \left((u_i^{(1)} - \bar{u}^{(1)}) + (u_i^{(4)} - \bar{u}^{(4)}) \right)^2 \right. \\ &\quad \left. + \left((u_i^{(2)} - \bar{u}^{(2)}) + (u_i^{(3)} - \bar{u}^{(3)}) \right) \cdot \left((u_i^{(1)} - \bar{u}^{(1)}) + (u_i^{(4)} - \bar{u}^{(4)}) \right) \right] \\ &+ \frac{\tau}{12} \left[\left((u_i^{(3)} - \bar{u}^{(3)}) - (u_i^{(2)} - \bar{u}^{(2)}) \right)^2 + \left((u_i^{(4)} - \bar{u}^{(4)}) - (u_i^{(1)} - \bar{u}^{(1)}) \right)^2 \right. \\ &\quad \left. + \left((u_i^{(3)} - \bar{u}^{(3)}) - (u_i^{(2)} - \bar{u}^{(2)}) \right) \cdot \left((u_i^{(4)} - \bar{u}^{(4)}) - (u_i^{(1)} - \bar{u}^{(1)}) \right) \right]. \end{aligned}$$

De manera análoga se definen los índices de diversidad de Gini-Simpson $\mathfrak{D}\mathbb{C}_1(\mathbb{L})$, $\mathfrak{D}\mathbb{C}_2(\mathbb{L})$, $\mathfrak{D}\mathbb{C}_1(\mathbb{F})$ y $\mathfrak{D}\mathbb{C}_2(\mathbb{F})$, y los D_τ -errores cuadrático medio $\text{ECMC}_1(\mathbb{L})$, $\text{ECMC}_2(\mathbb{L})$, $\text{ECMC}_1(\mathbb{F})$ y $\text{ECMC}_2(\mathbb{F})$, correspondientes a la conversión fuzzy a través de las respuestas Likertizadas y a través de los fuzzy libres simulados.

Siguiendo una idea similar, las medidas respectivas para los datos Likertizados serán:

El *índice de diversidad de Gini-Simpson muestral de $\iota(\mathcal{X})$* será:

$$\mathfrak{D}\mathbb{L} = \widehat{\mathfrak{D}(\iota(\mathcal{X}))} = 1 - \frac{1}{n^2} \sum_{j=1}^k [\text{card } \iota(\mathcal{X})^{-1}(\{j\})]^2$$

(con $\iota(\mathcal{X})^{-1}(\{j\}) = \{\omega_i : \iota(\mathcal{X}(\omega_i)) = j\}$).

El *error cuadrático medio muestral* de $\imath(\mathcal{X})$ será:

$$\text{ECML} = \text{ECM}[\widehat{\imath(\mathcal{X})}, E(\imath(\mathcal{X}))] = \frac{1}{n} \sum_{i=1}^n \left[\imath(\tilde{U}_i) - \overline{\imath(\mathcal{X})} \right]^2.$$

Observación 2.1.7. Obviamente, los índices \mathfrak{DL} y ECML son la particularización del \mathfrak{DF} y ECMF, respectivamente, cuando \mathcal{X} se reduce a la función indicador $\mathbf{1}_{\{\imath(\mathcal{X})\}}$, de manera que las comparaciones se van a llevar a cabo mediante los mismos procedimientos.

Observación 2.1.8. Es muy importante tener en cuenta (especialmente cuando tengan que interpretarse las conclusiones derivadas de las simulaciones) que aunque la métrica D_τ extiende la euclídea en \mathbb{R} , involucra las distancias euclídeas entre ‘localizaciones’ de los números fuzzy (en un orden de magnitud similar al de sus ‘Likertizaciones’), pero añade además las distancias provenientes de la diferencia en imprecisión/forma. En consecuencia, la distancia entre dos números fuzzy estará acotada inferiormente por la distancia entre sus ‘localizaciones’ correspondientes. De este modo, la comparación mediante los errores no es inmediatamente interpretable, aunque sí es plausible colegir que si el error en el caso fuzzy es inferior al Likert la reducción obedecerá mayoritariamente a una mayor representatividad de las medidas de tendencia central consideradas.

En la siguiente proposición se comprueba que el índice de diversidad de Gini-Simpson es siempre mayor para los números fuzzy simulados que para sus asociados enteros de la escala Likert. De manera análoga, este índice también es siempre superior en el caso de los datos fuzzy libres que en sus correspondientes datos fuzzy prefijados.

Proposición 2.1.1. *Si se considera como medida de la diversidad el índice de Shannon, o el de Gini-Simpson o, de forma más general, cualquier índice de diversidad \mathfrak{D} que cumpla la propiedad de recursividad o ramificación, se satisfará lo siguiente:*

Dada una muestra de tamaño n de simulaciones de números fuzzy trapezoidales, consideradas k categorías Likert, supongamos que se denota por n_1, \dots, n_k el cardinal del conjunto de números fuzzy a los que el criterio de Likertización lleva sobre el valor $1, \dots, k$, respectivamente; supongamos además que para cada $i \in \{1, \dots, k\}$, n_{i1} de los n_i números fuzzy Likertizados por i son iguales entre sí y a cierto número fuzzy, \dots , n_{ir_i} son iguales entre sí y a otro número fuzzy diferente (con $n_{i1} + \dots + n_{ir_i} = n_i$). Entonces, al ser el proceso de Likertización una aplicación, la recursividad de la diversidad garantiza que (con la notación usual para los índices de diversidad):

Diversidad asociada a la muestra basada en la escala fuzzy libre

$$\begin{aligned}
 &= \mathfrak{D} \left(\frac{n_{11}}{n}, \dots, \frac{n_{1r_1}}{n}, \dots, \frac{n_{k1}}{n}, \dots, \frac{n_{kr_k}}{n} \right) \\
 &= \mathfrak{D} \left(\frac{n_1}{n}, \dots, \frac{n_k}{n} \right) + \sum_{i=1}^k g \left(\frac{n_i}{n} \right) \cdot \mathfrak{D} \left(\frac{n_{i1}}{n_i}, \dots, \frac{n_{ir_i}}{n_i} \right)
 \end{aligned}$$

= diversidad asociada a la muestra basada en la escala Likert (i.e., entre categorías)
 $+ \sum_{i=1}^k g \left(\frac{n_i}{n} \right) \cdot$ *diversidad asociada a la submuestra basada en la escala fuzzy libre a la que se Likertiza por i (i.e., intra categoría i)*

donde g es una función no negativa cuya expresión depende del índice \mathfrak{D} elegido. Por ejemplo, si \mathfrak{D} es el índice basado en la entropía de Shannon $g(x) = x$; si es el basado en la entropía cuadrática (es decir, es el índice de Gini-Simpson), entonces $g(x) = x^2$, etc. En consecuencia, con independencia del índice elegido:

“Diversidad asociada a la muestra basada en la escala fuzzy libre” \geq “diversidad asociada a la muestra basada en la escala Likert”, siendo en la mayoría de los casos sustancialmente superior la primera que la segunda.

2.2. Estudio comparativo empírico

Para el estudio comparativo que llevamos a cabo consideramos muestras de diferente tamaño n : 10, 30, 50, 100 y 300. Así mismo, hemos desarrollado el estudio para diferentes valores del peso τ que interviene en la distancia D_τ : 0.0476, 0.0909, 0.2, 0.25 y 0.5.

En cada uno de los 5×5 casos se han simulado 1000 muestras y se han determinado:

- el porcentaje de muestras para las que el ECMF ha resultado inferior al ECML;
- el porcentaje de muestras para las que el ECMF ha resultado inferior al $\text{ECMC}_1(L)$;
- el porcentaje de muestras para las que el ECMF ha resultado inferior al $\text{ECMC}_2(L)$;
- el porcentaje de muestras para las que el ECMF ha resultado inferior al $\text{ECMC}_1(F)$;
- el porcentaje de muestras para las que el ECMF ha resultado inferior al $\text{ECMC}_2(F)$.

Además, se han determinado:

- el cociente entre la media del ECMF a lo largo de las 1000 muestras y la media del ECML;
- el cociente entre la media del ECMF a lo largo de las 1000 muestras y la media del $ECMC_1(L)$;
- el cociente entre la media del ECMF a lo largo de las 1000 muestras y la media del $ECMC_2(L)$;
- el cociente entre la media del ECMF a lo largo de las 1000 muestras y la media del $ECMC_1(F)$;
- el cociente entre la media del ECMF a lo largo de las 1000 muestras y la media del $ECMC_2(F)$.

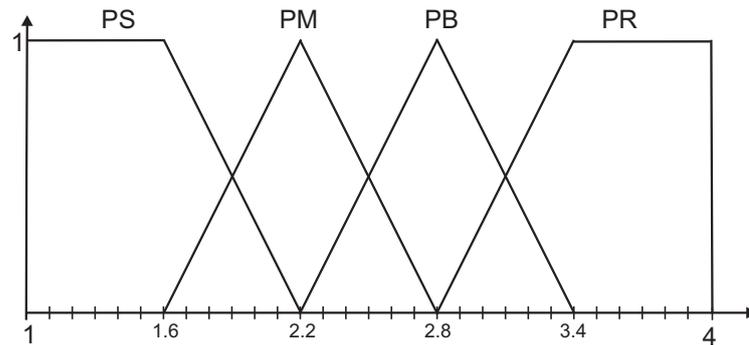
El estudio se ha realizado para cuatro valores de k diferentes: 4, 5, 6 y 7, y como ya se explicó anteriormente, las simulaciones de los n números fuzzy trapezoidales se empezaron con la simulación de sus mid_1 a partir de distribuciones $\beta(p, q)$, haciendo variar este parámetro vectorial (p, q) de modo que se considerasen diferentes situaciones de simetría y asimetría.

Para comparar el formato de respuesta libre fuzzy con el formato de respuesta fuzzy prefijado en cuestionarios, utilizamos dos conjuntos diferentes de k números fuzzy trapezoidales prefijados cada uno para cada valor de k . Todos ellos fueron obtenidos de diferentes artículos a los que remitiremos a continuación, eligiéndose que uno de esos conjuntos estuviera formado por números fuzzy cuya disposición en el plano fuese simétrica, y el otro conjunto contuviese números fuzzy con disposición asimétrica, para así poder abarcar todos los casos posibles del uso de etiquetas lingüísticas en cuestionarios, y analizar en un futuro si la disposición de los números fuzzy en forma simétrica o en forma asimétrica influye en nuestro estudio.

Recordando que todo número fuzzy trapezoidal \tilde{U} queda completamente determinado por la cuaterna de valores $(\inf \tilde{U}_0, \inf \tilde{U}_1, \sup \tilde{U}_1, \sup \tilde{U}_0)$, los conjuntos de etiquetas lingüísticas fuzzy trapezoidales que hemos utilizado fueron los siguientes, ya reescalados y trasladados estos números fuzzy al intervalo $[1, k]$:

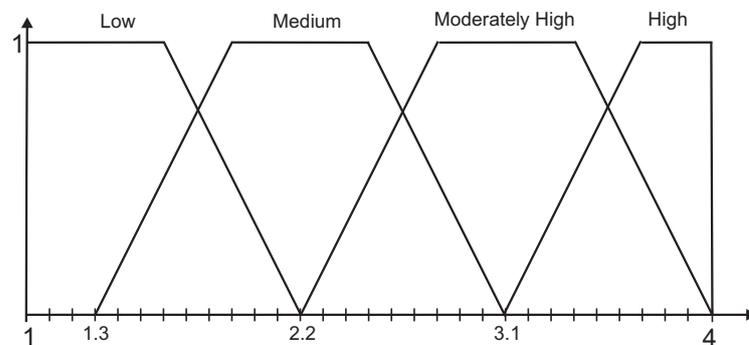
- Caso $k = 4$: se eligió como conjunto simétrico (ver Jiejun Cai *et al.*, 2012):

$$C_1 = \{(1; 1; 1,6; 2,2), (1,6; 2,2; 2,2; 2,8), (2,2; 2,8; 2,8; 3,4), (2,8; 3,4; 4; 4)\}$$



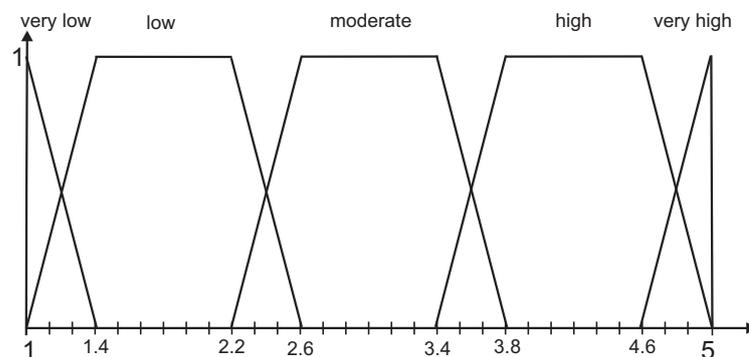
y como conjunto asimétrico (ver Shailendra Bajpai *et al.*, 2010):

$$C_2 = \{(1; 1; 1,6; 2,2), (1,3; 1,9; 2,5; 3,1), (2,2; 2,8; 3,4; 4), (3,1; 3,7; 4; 4)\}$$



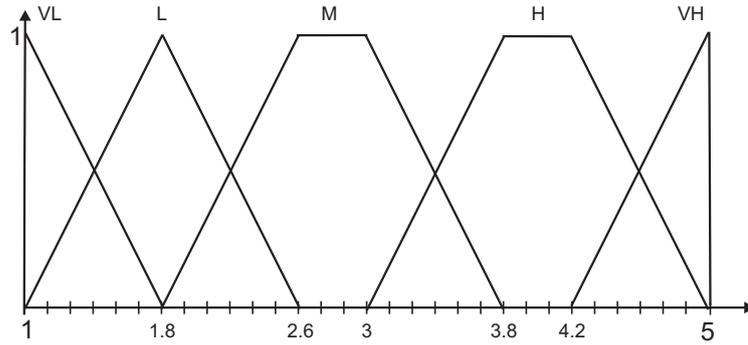
- Caso $k = 5$: se eligió como conjunto simétrico (ver Ibrahim A. Motawa *et al.*, 2006):

$$C_1 = \{(1; 1; 1; 1,4), (1; 1,4; 2,2; 2,6), (2,2; 2,6; 3,4; 3,8), (3,4; 3,8; 4,6; 5), (4,6; 5; 5; 5)\}$$



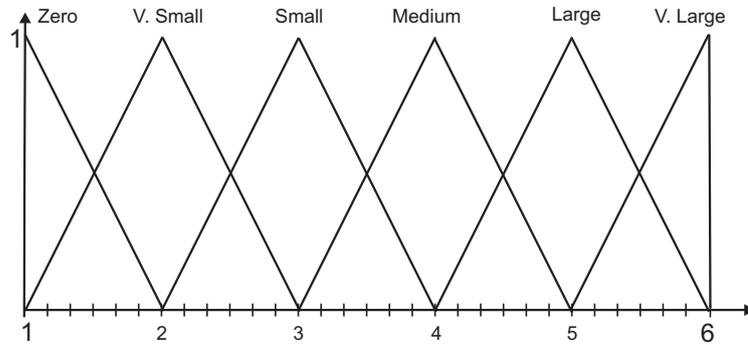
y como conjunto asimétrico (ver Chung-Hsing Yeh *et al.*, 1999):

$$C_2 = \{(1; 1; 1; 1,8), (1; 1,8; 1,8; 2,6), (1,8; 2,6; 3; 3,8), (3; 3,8; 4,2; 5), (4,2; 5; 5; 5)\}$$



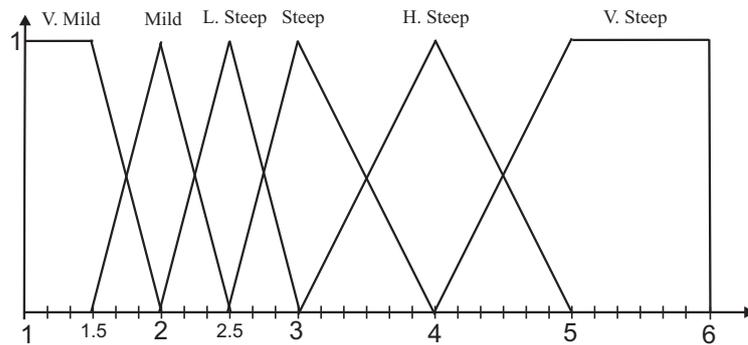
- Caso $k = 6$: se eligió como conjunto simétrico (ver Mehmet Saltan *et al.*, 2007):

$$C_1 = \{(1; 1; 1; 2), (1; 2; 2; 3), (2; 3; 3; 4), (3; 4; 4; 5), (4; 5; 5; 6), (5; 6; 6; 6)\}$$



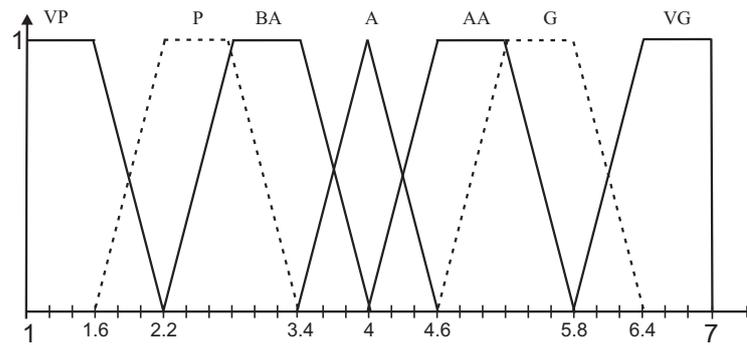
y como conjunto asimétrico (ver Gokmen Tayfur *et al.*, 2003):

$$C_2 = \{(1; 1; 1,5; 2), (1,5; 2; 2; 2,5), (2; 2,5; 2,5; 3), (2,5; 3; 3; 4), (3; 4; 4; 5), (4; 5; 6; 6)\}$$



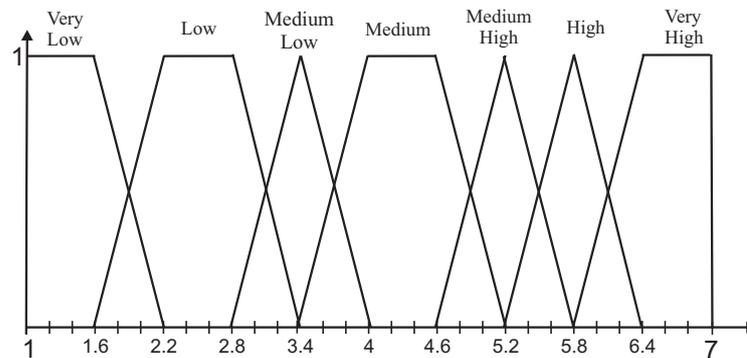
- Caso $k = 7$: se eligió como conjunto simétrico (ver Edyta Plebankiewicz, 2012):

$$C_1 = \{(1; 1; 1,6; 2,2), (1,6; 2,2; 2,8; 3,4), (2,2; 2,8; 3,4; 4), (3,4; 4; 4; 4,6), (4; 4,6; 5,2; 5,8), (4,6; 5,2; 5,8; 6,4), (5,8; 6,4; 7; 7)\}$$



y como conjunto asimétrico (ver Atakan Yücel y Ali Fuat Güneri, 2011):

$$C_2 = \{(1; 1; 1,6; 2,2), (1,6; 2,2; 2,8; 3,4), (2,8; 3,4; 3,4; 4), (3,4; 4; 4,6; 5,2), (4,6; 5,2; 5,2; 5,8), (5,2; 5,8; 5,8; 6,4), (5,8; 6,4; 7; 7)\}$$



Los resultados de las simulaciones realizadas se han recogido en las tablas siguientes:

2.2.1. Tablas para el caso $k = 4$ Tabla 2.1: Simulaciones de los mid_1 a partir de $\beta(1, 1)$

τ	n	ECML	%ECMF <			
			ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	64.9	85.7	81	99.7	99.4
	30	76.3	98.1	94.8	100	100
	50	83.7	99.6	98.6	100	100
	100	91	100	99.8	100	100
	300	99.4	100	100	100	100
0.0909	10	73.1	85.7	81.2	99.7	99.4
	30	87.8	98.2	95	100	100
	50	95.4	99.6	98.7	100	100
	100	97.6	100	99.8	100	100
	300	100	100	100	100	100
0.2	10	88	85.7	81.9	99.7	99.5
	30	99	98.2	95.5	100	100
	50	99.9	99.6	98.8	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.25	10	93.2	85.7	82.1	99.7	99.6
	30	99.9	98.2	95.7	100	100
	50	100	99.6	98.9	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	99.1	85.9	84.3	99.7	99.8
	30	100	98.3	96.8	100	100
	50	100	99.6	99.3	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.2: Simulaciones de los mid_1 a partir de $\beta(1, 1)$

τ	n	ECMF	ECMF	ECMF	ECMF	ECMF
		$\overline{\text{ECML}}$	$\overline{\text{ECMC}_1(\text{L})}$	$\overline{\text{ECMC}_2(\text{L})}$	$\overline{\text{ECMC}_1(\text{F})}$	$\overline{\text{ECMC}_2(\text{F})}$
0.0476	10	0.9355	0.7863	0.866	0.533	0.688
	30	0.9336	0.7844	0.8644	0.5346	0.6902
	50	0.9346	0.7849	0.8659	0.5354	0.6923
	100	0.934	0.7842	0.8655	0.5357	0.6922
	300	0.9333	0.7833	0.8658	0.5361	0.6932
0.0909	10	0.8932	0.7859	0.8651	0.5343	0.6834
	30	0.8915	0.784	0.8635	0.5353	0.6861
	50	0.8924	0.7845	0.865	0.536	0.6875
	100	0.8918	0.7838	0.8646	0.5365	0.6874
	300	0.8911	0.783	0.8649	0.5369	0.6886
0.2	10	0.7866	0.7847	0.8624	0.537	0.67
	30	0.7851	0.7828	0.8608	0.5376	0.6719
	50	0.7859	0.7833	0.8623	0.5387	0.6733
	100	0.7854	0.7826	0.8619	0.539	0.6735
	300	0.7848	0.7818	0.8622	0.5393	0.6745
0.25	10	0.7378	0.7841	0.8609	0.5381	0.6618
	30	0.7364	0.7822	0.8593	0.5392	0.6643
	50	0.7372	0.7827	0.8609	0.5401	0.6659
	100	0.7367	0.782	0.8604	0.5402	0.666
	300	0.7361	0.7812	0.8607	0.5406	0.6669
0.5	10	0.4936	0.7789	0.8491	0.547	0.6079
	30	0.4927	0.7771	0.8477	0.5497	0.6112
	50	0.4932	0.7776	0.8492	0.5497	0.6132
	100	0.4929	0.777	0.8489	0.5493	0.6128
	300	0.4925	0.7761	0.849	0.55	0.6131

Tabla 2.3: Simulaciones de los mid_1 a partir de $\beta(0,75; 0,75)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	65.3	90.5	80.3	99.9	98.8
	30	78.9	98.8	94.5	100	100
	50	82.7	100	97.9	100	100
	100	91.5	100	100	100	100
	300	99.3	100	100	100	100
0.0909	10	75.8	90.5	80.5	99.9	99.1
	30	91	98.8	94.6	100	100
	50	95.6	100	97.9	100	100
	100	99.6	100	100	100	100
	300	100	100	100	100	100
0.2	10	91.7	90.5	80.9	99.9	99.1
	30	99	98.8	95.1	100	100
	50	100	100	98.1	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.25	10	95.4	90.6	81.2	99.9	99.3
	30	100	98.8	95.4	100	100
	50	100	100	98.2	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	99.9	91	83.6	100	99.6
	30	100	99.1	96.8	100	100
	50	100	100	99.5	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.4: Simulaciones de los mid_1 a partir de $\beta(0,75; 0,75)$

τ	n	$\frac{\text{ECMF}}{\text{ECML}}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{F})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{F})}$
		0.0476	10	0.9433	0.7825	0.8902
	30	0.9404	0.7796	0.8879	0.5736	0.7401
	50	0.9468	0.7848	0.8942	0.5776	0.7426
	100	0.9432	0.7817	0.8904	0.576	0.7437
	300	0.9427	0.7816	0.89	0.5752	0.741
0.0909	10	0.9006	0.7821	0.8893	0.5759	0.7361
	30	0.8978	0.7792	0.887	0.5743	0.736
	50	0.904	0.7844	0.8933	0.5784	0.7378
	100	0.9005	0.7813	0.8895	0.5767	0.7394
	300	0.9	0.7812	0.8891	0.5758	0.7367
0.2	10	0.793	0.781	0.8866	0.5776	0.7238
	30	0.7906	0.7781	0.8843	0.5765	0.7236
	50	0.796	0.7833	0.8905	0.5799	0.7253
	100	0.793	0.7802	0.8868	0.5787	0.7264
	300	0.7925	0.7801	0.8864	0.5779	0.7243
0.25	10	0.7437	0.7804	0.885	0.5792	0.7166
	30	0.7415	0.7775	0.8828	0.5777	0.7169
	50	0.7465	0.7827	0.889	0.5811	0.719
	100	0.7437	0.7796	0.8853	0.58	0.719
	300	0.7433	0.7795	0.8849	0.579	0.7175
0.5	10	0.4973	0.7755	0.8731	0.5884	0.6651
	30	0.4957	0.7726	0.8709	0.5859	0.6679
	50	0.4991	0.7778	0.8771	0.5889	0.6712
	100	0.4972	0.7748	0.8735	0.5873	0.6702
	300	0.497	0.7747	0.873	0.5868	0.6689

Tabla 2.5: Simulaciones de los mid_1 a partir de $\beta(2, 2)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	59.2	69.5	81.5	99.9	99.9
	30	67.5	83.3	95.3	100	100
	50	70.6	90.2	98.3	100	100
	100	75.1	96	99.9	100	100
	300	90.3	99.9	100	100	100
0.0909	10	66.5	69.5	81.6	99.9	99.9
	30	76.4	83.3	95.3	100	100
	50	82.5	90.2	98.3	100	100
	100	90	96	99.9	100	100
	300	99.3	99.9	100	100	100
0.2	10	78.4	69.4	81.7	99.9	99.9
	30	92.8	83.4	95.6	100	100
	50	97.2	90.3	98.3	100	100
	100	99.7	96	99.9	100	100
	300	100	99.9	100	100	100
0.25	10	83.5	69.5	81.7	99.9	99.9
	30	97.4	83.5	95.7	100	100
	50	98.9	90.5	98.3	100	100
	100	100	96	99.9	100	100
	300	100	99.9	100	100	100
0.5	10	98.3	69.9	82.4	99.6	99.9
	30	100	83.9	96	100	100
	50	100	91	98.6	100	100
	100	100	96.3	100	100	100
	300	100	99.9	100	100	100

Tabla 2.6: Simulaciones de los mid_1 a partir de $\beta(2, 2)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.9441	0.8381	0.8123
	30	0.9393	0.8341	0.8075	0.4541	0.5853
	50	0.944	0.8375	0.8144	0.4542	0.5876
	100	0.9482	0.8432	0.8152	0.4536	0.5873
	300	0.9438	0.839	0.8114	0.4538	0.5864
0.0909	10	0.9016	0.8378	0.8118	0.456	0.5794
	30	0.897	0.8338	0.807	0.4554	0.5795
	50	0.9015	0.8372	0.8139	0.4552	0.5818
	100	0.9055	0.8429	0.8146	0.4549	0.5817
	300	0.9013	0.8388	0.8108	0.4551	0.5804
0.2	10	0.7946	0.8369	0.81	0.4599	0.5603
	30	0.7905	0.833	0.8054	0.4594	0.5631
	50	0.7944	0.8364	0.8121	0.4587	0.5649
	100	0.798	0.8421	0.8129	0.4585	0.5649
	300	0.7943	0.838	0.8092	0.4586	0.5635
0.25	10	0.7455	0.8365	0.8091	0.4619	0.5514
	30	0.7417	0.8325	0.8044	0.4622	0.5541
	50	0.7454	0.836	0.8111	0.4605	0.5556
	100	0.7487	0.8417	0.812	0.4601	0.5557
	300	0.7452	0.8375	0.8083	0.4606	0.5541
0.5	10	0.5001	0.8327	0.8016	0.4768	0.488
	30	0.4976	0.829	0.7973	0.4763	0.4914
	50	0.5	0.8323	0.8036	0.4755	0.4903
	100	0.5023	0.8381	0.8046	0.4758	0.4896
	300	0.5	0.834	0.801	0.4759	0.4895

Tabla 2.7: Simulaciones de los mid_1 a partir de $\beta(4, 2)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	75.4	85.5	75.2	99.9	97.8
	30	85.5	96.4	88.4	100	100
	50	91.9	98.8	94.7	100	100
	100	97.5	99.9	98.5	100	100
	300	100	100	100	100	100
0.0909	10	78.4	85.6	75.4	99.9	98
	30	90.1	96.5	88.7	100	100
	50	95.6	98.8	94.9	100	100
	100	99.1	99.9	98.5	100	100
	300	100	100	100	100	100
0.2	10	86.5	85.6	76.2	99.9	98.4
	30	97.7	96.6	90.1	100	100
	50	99.3	98.8	96	100	100
	100	100	99.9	98.8	100	100
	300	100	100	100	100	100
0.25	10	89.6	85.6	76.5	99.9	98.7
	30	98.5	96.6	90.5	100	100
	50	99.6	98.8	96.4	100	100
	100	100	99.9	98.8	100	100
	300	100	100	100	100	100
0.5	10	97.7	85.6	79.8	99.9	99.6
	30	100	96.7	94	100	100
	50	100	98.8	97.9	100	100
	100	100	99.9	99.8	100	100
	300	100	100	100	100	100

Tabla 2.8: Simulaciones de los mid_1 a partir de $\beta(4, 2)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.8226	0.6942	0.8132
30	0.8342		0.7061	0.8191	0.4402	0.5502
50	0.8375		0.7095	0.8218	0.4428	0.5529
100	0.8366		0.7075	0.8231	0.4434	0.5541
300	0.8332		0.7047	0.8198	0.4425	0.5535
0.0909	10	0.7857	0.6938	0.8109	0.4385	0.541
	30	0.7968	0.7057	0.8169	0.4409	0.5426
	50	0.8	0.7091	0.8197	0.4436	0.5465
	100	0.7992	0.7071	0.8209	0.4442	0.5469
	300	0.7959	0.7042	0.8175	0.4432	0.5469
0.2	10	0.693	0.6926	0.8043	0.4416	0.5225
	30	0.7028	0.7045	0.8104	0.4431	0.5236
	50	0.7056	0.7079	0.8132	0.4457	0.528
	100	0.7048	0.7059	0.8143	0.4465	0.5282
	300	0.7019	0.703	0.811	0.4453	0.5283
0.25	10	0.6505	0.6919	0.8007	0.4426	0.5131
	30	0.6597	0.7038	0.8069	0.4441	0.5127
	50	0.6623	0.7072	0.8097	0.447	0.5179
	100	0.6616	0.7052	0.8107	0.448	0.518
	300	0.6589	0.7023	0.8074	0.4464	0.5179
0.5	10	0.4379	0.6866	0.7733	0.4518	0.4513
	30	0.4441	0.6986	0.78	0.4516	0.4482
	50	0.4458	0.702	0.7827	0.4558	0.4494
	100	0.4453	0.6998	0.7832	0.4565	0.4518
	300	0.4435	0.697	0.7801	0.4548	0.4509

Tabla 2.9: Simulaciones de los mid_1 a partir de $\beta(6, 1)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	91.5	97.5	68.4	69.4	65.6
	30	99.4	100	79	90.1	78.7
	50	99.9	100	84.3	95.6	89.5
	100	100	100	92.3	99	95.4
	300	100	100	99.6	100	99.9
0.0909	10	93.1	97.5	69.1	69.6	64.9
	30	99.7	100	79.7	90.5	77.9
	50	100	100	85.6	95.7	89.4
	100	100	100	93.1	99.3	95
	300	100	100	99.6	100	99.8
0.2	10	96.2	97.5	71.2	70.7	63.7
	30	100	100	82.9	91.7	76.1
	50	100	100	88.3	96	86.9
	100	100	100	95.2	99.3	95.2
	300	100	100	99.9	100	99.6
0.25	10	96.9	97.5	72.6	70.9	62.1
	30	100	100	84.4	92	76.3
	50	100	100	89.8	96.2	86.2
	100	100	100	96.3	99.2	94.1
	300	100	100	100	100	99.6
0.5	10	98.3	97.7	80.5	74.8	48.8
	30	100	100	93.6	94.2	66.6
	50	100	100	97.1	97.6	78.4
	100	100	100	99.8	99.6	87.6
	300	100	100	100	100	98.3

Tabla 2.10: Simulaciones de los mid_1 a partir de $\beta(6, 1)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.6047	0.4352	0.8296
	30	0.61	0.4394	0.8357	0.6802	0.767
	50	0.6111	0.4406	0.8353	0.6841	0.7599
	100	0.6078	0.4385	0.83	0.6797	0.7601
	300	0.6077	0.4378	0.8319	0.681	0.7626
0.0909	10	0.5779	0.4347	0.8223	0.6923	0.7661
	30	0.583	0.4389	0.8284	0.678	0.7678
	50	0.584	0.4402	0.828	0.6816	0.7604
	100	0.5809	0.438	0.8228	0.6773	0.7603
	300	0.5808	0.4374	0.8247	0.6785	0.7634
0.2	10	0.5105	0.4333	0.8013	0.6829	0.7671
	30	0.515	0.4376	0.8074	0.6691	0.7689
	50	0.5159	0.4388	0.8071	0.6736	0.7634
	100	0.5132	0.4367	0.8021	0.6699	0.7618
	300	0.513	0.436	0.8038	0.6711	0.766
0.25	10	0.4797	0.4326	0.7902	0.6802	0.7719
	30	0.4839	0.4368	0.7962	0.6648	0.7704
	50	0.4847	0.438	0.796	0.6696	0.7666
	100	0.4821	0.4359	0.7911	0.667	0.7641
	300	0.482	0.4352	0.7927	0.6669	0.7682
0.5	10	0.3252	0.4268	0.7126	0.6473	0.799
	30	0.328	0.431	0.7182	0.634	0.7962
	50	0.3285	0.4321	0.7183	0.6404	0.7879
	100	0.3268	0.43	0.7141	0.6383	0.7888
	300	0.3268	0.4294	0.7151	0.6373	0.792

Tabla 2.11: Simulaciones de los mid_1 a partir de $\beta(6, 10)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	58.3	55.5	76.2	98.2	92.4
	30	63.5	64.2	90.2	100	100
	50	70.4	71.4	95.2	100	100
	100	80.5	81.8	98.5	100	100
	300	92.8	94	100	100	100
0.0909	10	61	55.4	76.2	98.1	92.8
	30	69.1	64.2	90.2	100	100
	50	77.1	71.4	95.2	100	100
	100	86.2	81.7	98.5	100	100
	300	97.4	93.9	100	100	100
0.2	10	69.6	54.8	76	98	93.9
	30	82.3	63.9	90.1	100	100
	50	89.5	70.7	94.9	100	100
	100	96.1	81.2	98.5	100	100
	300	99.9	93.2	100	100	100
0.25	10	73.4	54.6	75.9	98	94.7
	30	86.3	63.7	90	100	100
	50	92.8	70.4	94.9	100	100
	100	97.8	80.6	98.4	100	100
	300	100	93	100	100	100
0.5	10	85.8	53.2	75.3	96.2	97.5
	30	96.8	61.6	89.2	100	100
	50	99.3	68.1	94.2	100	100
	100	99.9	78.2	98.1	100	100
	300	100	89.9	100	100	100

Tabla 2.12: Simulaciones de los mid_1 a partir de $\beta(6, 10)$

τ	n	$\frac{\text{ECMF}}{\text{ECML}}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{F})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{F})}$
		0.0476	10	0.8688	0.8453	0.6502
	30	0.8972	0.8746	0.6705	0.345	0.4779
	50	0.8807	0.8563	0.6592	0.3446	0.4842
	100	0.8794	0.8545	0.6585	0.3455	0.4843
	300	0.8836	0.861	0.6606	0.3467	0.4845
0.0909	10	0.8308	0.8462	0.6511	0.3469	0.4822
	30	0.858	0.8756	0.6715	0.3467	0.474
	50	0.8421	0.8572	0.6602	0.3461	0.48
	100	0.8409	0.8554	0.6594	0.347	0.4803
	300	0.845	0.8619	0.6616	0.3484	0.4803
0.2	10	0.735	0.8488	0.6538	0.3528	0.4696
	30	0.7591	0.8785	0.6744	0.3519	0.462
	50	0.745	0.8599	0.6629	0.3505	0.4678
	100	0.7439	0.858	0.6622	0.352	0.4682
	300	0.7475	0.8647	0.6643	0.3534	0.4681
0.25	10	0.691	0.8503	0.6553	0.3554	0.4626
	30	0.7138	0.8801	0.6759	0.355	0.455
	50	0.7005	0.8614	0.6644	0.3535	0.4602
	100	0.6994	0.8595	0.6637	0.3549	0.4618
	300	0.7029	0.8662	0.6659	0.356	0.4614
0.5	10	0.4714	0.8621	0.6672	0.3791	0.409
	30	0.4872	0.8927	0.6886	0.3769	0.4099
	50	0.478	0.8734	0.6766	0.3748	0.411
	100	0.4772	0.8713	0.6757	0.3761	0.4135
	300	0.4796	0.8783	0.6781	0.3766	0.4132

2.2.2. Tablas para el caso $k = 5$ Tabla 2.13: Simulaciones de los mid_1 a partir de $\beta(1, 1)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	63.1	95.7	96.5	84.3	98.7
	30	75.2	100	100	98.6	100
	50	82.4	100	100	99.6	100
	100	89	100	100	100	100
	300	98.7	100	100	100	100
0.0909	10	72.2	95.8	96.7	85.3	98.7
	30	88.6	100	100	98.8	100
	50	94.1	100	100	99.6	100
	100	98.8	100	100	100	100
	300	100	100	100	100	100
0.2	10	90.6	96.1	96.8	89.3	99
	30	99.4	100	100	99.5	100
	50	99.9	100	100	99.8	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.25	10	95	96.2	96.8	91	99.1
	30	99.8	100	100	99.6	100
	50	99.9	100	100	99.9	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	99.9	97	97.3	97.8	99.6
	30	100	100	100	100	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.14: Simulaciones de los mid_1 a partir de $\beta(1, 1)$

τ	n	$\frac{\text{ECMF}}{\text{ECML}}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{F})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{F})}$
		0.0476	10	0.9483	0.7973	0.7483
	30	0.9454	0.7969	0.7479	0.8626	0.7375
	50	0.942	0.7963	0.7465	0.8622	0.7386
	100	0.9451	0.7974	0.7482	0.8632	0.7391
	300	0.9444	0.7977	0.748	0.8629	0.7385
0.0909	10	0.9054	0.7963	0.7477	0.8627	0.7325
	30	0.9027	0.7959	0.7473	0.8571	0.7333
	50	0.8995	0.7953	0.7459	0.8565	0.7346
	100	0.9024	0.7965	0.7476	0.8581	0.7351
	300	0.9018	0.7967	0.7474	0.8575	0.7345
0.2	10	0.7974	0.7934	0.7459	0.8444	0.7208
	30	0.795	0.7931	0.7456	0.8409	0.722
	50	0.7922	0.7924	0.7441	0.8402	0.7237
	100	0.7947	0.7936	0.7459	0.8417	0.7234
	300	0.7942	0.7939	0.7456	0.841	0.7231
0.25	10	0.7479	0.7918	0.7449	0.8358	0.7151
	30	0.7457	0.7915	0.7446	0.8334	0.7165
	50	0.743	0.7908	0.7431	0.8311	0.7173
	100	0.7454	0.7921	0.7449	0.8331	0.7169
	300	0.7449	0.7923	0.7446	0.8323	0.7167
0.5	10	0.5003	0.7793	0.7371	0.7714	0.6683
	30	0.4989	0.7791	0.7369	0.7679	0.6703
	50	0.4971	0.7784	0.7355	0.7696	0.6726
	100	0.4988	0.7797	0.7373	0.7694	0.6707
	300	0.4984	0.7799	0.737	0.7687	0.6713

Tabla 2.15: Simulaciones de los mid_1 a partir de $\beta(0,75; 0,75)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	68.8	97	97.8	89.7	99.6
	30	81.6	100	100	98.9	100
	50	87.2	100	100	99.6	100
	100	95.3	100	100	100	100
	300	99.9	100	100	100	100
0.0909	10	78.5	97	97.8	90.7	99.7
	30	93.1	100	100	99.2	100
	50	96.9	100	100	99.8	100
	100	99.7	100	100	100	100
	300	100	100	100	100	100
0.2	10	94.4	97.4	97.9	93.3	99.8
	30	99.7	100	100	99.7	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.25	10	97.3	97.4	97.9	94.5	99.9
	30	100	100	100	99.8	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	99.9	97.9	98.1	98.5	99.8
	30	100	100	100	100	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.16: Simulaciones de los mid_1 a partir de $\beta(0,75; 0,75)$

τ	n	$\frac{\text{ECMF}}{\text{ECML}}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{F})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{F})}$
		0.0476	10	0.9406	0.8234	0.7678
	30	0.9423	0.8266	0.7711	0.8782	0.7599
	50	0.941	0.827	0.7712	0.8806	0.7627
	100	0.9405	0.8264	0.7709	0.8807	0.7618
	300	0.9408	0.8257	0.7701	0.8793	0.7597
0.0909	10	0.8981	0.8223	0.7672	0.8716	0.7521
	30	0.8997	0.8256	0.7705	0.8732	0.7569
	50	0.8984	0.8259	0.7706	0.8759	0.7593
	100	0.8979	0.8253	0.7703	0.8755	0.7586
	300	0.8982	0.8246	0.7695	0.8741	0.7563
0.2	10	0.7908	0.8192	0.7654	0.8564	0.7399
	30	0.7922	0.8224	0.7687	0.8586	0.7463
	50	0.7911	0.8227	0.7688	0.8613	0.7494
	100	0.7907	0.8221	0.7685	0.8599	0.7487
	300	0.7909	0.8214	0.7676	0.8587	0.746
0.25	10	0.7416	0.8174	0.7644	0.8493	0.7346
	30	0.743	0.8206	0.7677	0.8507	0.7409
	50	0.742	0.821	0.7678	0.852	0.7437
	100	0.7416	0.8204	0.7675	0.852	0.7431
	300	0.7418	0.8197	0.7666	0.8507	0.7406
0.5	10	0.4959	0.8037	0.7564	0.7889	0.6963
	30	0.4968	0.8068	0.7597	0.7916	0.7004
	50	0.4961	0.8072	0.7599	0.795	0.7026
	100	0.4958	0.8067	0.7596	0.7935	0.7025
	300	0.496	0.8059	0.7587	0.792	0.7003

Tabla 2.17: Simulaciones de los mid_1 a partir de $\beta(2, 2)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	59.5	96.5	96.6	86.7	96.7
	30	59.5	99.9	99.9	97.1	100
	50	62.1	100	100	99.7	100
	100	69.4	100	100	100	100
	300	80.1	100	100	100	100
0.0909	10	65.9	96.6	96.6	87.2	96.8
	30	74.5	99.9	99.9	97.6	100
	50	79.4	100	100	99.8	100
	100	89	100	100	100	100
	300	97.7	100	100	100	100
0.2	10	82.3	96.6	96.6	89.1	97.5
	30	94.8	99.9	99.9	98.5	100
	50	98.4	100	100	99.9	100
	100	99.8	100	100	100	100
	300	100	100	100	100	100
0.25	10	87.5	96.7	96.7	90.6	97.7
	30	97.9	99.9	99.9	98.8	100
	50	99.7	100	100	99.9	100
	100	99.9	100	100	100	100
	300	100	100	100	100	100
0.5	10	99.8	97.1	96.9	96.4	99.3
	30	100	99.9	99.9	99.8	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.18: Simulaciones de los mid_1 a partir de $\beta(2, 2)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.9538	0.7109	0.6776
	30	0.9712	0.7231	0.6914	0.8236	0.6914
	50	0.9712	0.7244	0.6905	0.8214	0.6936
	100	0.9693	0.7192	0.6866	0.8171	0.6897
	300	0.9693	0.72	0.6875	0.818	0.6901
0.0909	10	0.9109	0.7106	0.6773	0.8071	0.6831
	30	0.9275	0.7227	0.691	0.8193	0.6865
	50	0.9274	0.724	0.6902	0.8172	0.6892
	100	0.9257	0.7188	0.6863	0.8126	0.6852
	300	0.9257	0.7197	0.6872	0.8137	0.6856
0.2	10	0.8027	0.7095	0.6763	0.7901	0.669
	30	0.8173	0.7216	0.69	0.8057	0.6733
	50	0.8173	0.7229	0.6891	0.803	0.6754
	100	0.8158	0.7178	0.6853	0.7987	0.6718
	300	0.8157	0.7187	0.6862	0.8	0.6721
0.25	10	0.7531	0.7089	0.6757	0.7803	0.662
	30	0.7668	0.7211	0.6894	0.7985	0.6661
	50	0.7668	0.7223	0.6886	0.7954	0.6682
	100	0.7654	0.7173	0.6848	0.7916	0.6648
	300	0.7654	0.7181	0.6856	0.7928	0.6649
0.5	10	0.5052	0.7042	0.6713	0.7206	0.6075
	30	0.5145	0.7164	0.685	0.734	0.6109
	50	0.5144	0.7175	0.6841	0.7328	0.6139
	100	0.5135	0.7128	0.6804	0.7304	0.6103
	300	0.5135	0.7137	0.6813	0.7307	0.6109

Tabla 2.19: Simulaciones de los mid_1 a partir de $\beta(4, 2)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	66.7	91.2	98.3	83.1	95.1
	30	75.5	99.3	100	96.7	99.9
	50	84.1	100	100	98.8	100
	100	92.5	100	100	100	100
	300	99.2	100	100	100	100
0.0909	10	72.4	91.2	98.3	83.5	95.7
	30	84.1	99.3	100	97.2	99.9
	50	92.2	100	100	99.1	100
	100	98	100	100	100	100
	300	99.9	100	100	100	100
0.2	10	84.8	91.2	98.5	85.2	96.7
	30	96.4	99.3	100	98.4	100
	50	98.8	100	100	99.5	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.25	10	88.9	91.5	98.5	86.7	97
	30	98.4	99.2	100	98.7	100
	50	99.8	100	100	99.7	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	99.3	92.8	98.6	94	99.3
	30	100	99.4	100	99.7	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.20: Simulaciones de los mid_1 a partir de $\beta(4, 2)$

τ	n	$\frac{\text{ECMF}}{\text{ECML}}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{F})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{F})}$
		0.0476	10	0.908	0.7174	0.6085
	30	0.9043	0.7106	0.6033	0.7758	0.6433
	50	0.9014	0.7091	0.6021	0.7746	0.6415
	100	0.8988	0.7086	0.6015	0.7726	0.641
	300	0.902	0.7111	0.603	0.7757	0.6419
0.0909	10	0.8673	0.7165	0.608	0.7727	0.6378
	30	0.8638	0.7097	0.6029	0.771	0.6381
	50	0.861	0.7083	0.6017	0.7691	0.6363
	100	0.8586	0.7077	0.601	0.7674	0.6359
	300	0.8616	0.7102	0.6025	0.7704	0.6365
0.2	10	0.7649	0.7137	0.6066	0.7576	0.6204
	30	0.7619	0.7071	0.6015	0.7556	0.6211
	50	0.7594	0.7057	0.6003	0.7538	0.6203
	100	0.7572	0.705	0.5996	0.7523	0.6194
	300	0.7599	0.7075	0.6011	0.7547	0.6198
0.25	10	0.718	0.7122	0.6058	0.7467	0.6128
	30	0.7151	0.7057	0.6008	0.7469	0.6117
	50	0.7128	0.7042	0.5996	0.7453	0.6114
	100	0.7108	0.7035	0.5989	0.7436	0.6108
	300	0.7133	0.706	0.6004	0.7461	0.6107
0.5	10	0.4833	0.7003	0.5997	0.6751	0.548
	30	0.4814	0.6946	0.5951	0.6766	0.5458
	50	0.4798	0.6931	0.5939	0.6767	0.5481
	100	0.4784	0.6921	0.593	0.6743	0.5481
	300	0.4802	0.6946	0.5945	0.6762	0.5471

Tabla 2.21: Simulaciones de los mid_1 a partir de $\beta(6, 1)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	92.2	66.1	92.2	75	82.1
	30	99.5	77.2	99.3	90.2	95.6
	50	99.9	86.1	99.9	95.6	99.3
	100	100	94.4	100	99.5	100
	300	100	99.6	100	100	100
0.0909	10	93.8	67	92.6	75.7	82.4
	30	99.8	79.4	99.3	91.4	95.4
	50	100	87.5	99.9	96.3	99
	100	100	95.2	100	99.6	100
	300	100	99.7	100	100	100
0.2	10	97.6	70.5	93.6	78.2	82.9
	30	100	84.5	99.5	93.6	95.4
	50	100	91.3	99.9	98	99.4
	100	100	97.3	100	99.9	100
	300	100	100	100	100	100
0.25	10	98.6	72.8	94	80.5	82.9
	30	100	86.9	99.5	94.8	95.9
	50	100	93.1	99.9	98.6	99.6
	100	100	98	100	99.9	100
	300	100	100	100	100	100
0.5	10	99.7	85.9	97.2	88.6	83.6
	30	100	97.9	100	99.2	96.5
	50	100	99.8	100	100	99.3
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.22: Simulaciones de los mid_1 a partir de $\beta(6, 1)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.6583	0.8275	0.5879
	30	0.6658	0.8332	0.5922	0.748	0.6785
	50	0.6645	0.8301	0.59	0.758	0.6736
	100	0.6649	0.8322	0.5915	0.7615	0.6758
	300	0.6639	0.8321	0.591	0.758	0.677
0.0909	10	0.6292	0.8193	0.5847	0.761	0.6754
	30	0.6364	0.825	0.589	0.7396	0.6758
	50	0.6351	0.822	0.5869	0.75	0.6704
	100	0.6355	0.824	0.5884	0.7532	0.6725
	300	0.6346	0.8239	0.5879	0.7498	0.6739
0.2	10	0.5558	0.7955	0.5755	0.7367	0.6666
	30	0.5622	0.8014	0.5799	0.7178	0.6677
	50	0.561	0.7987	0.5778	0.727	0.662
	100	0.5614	0.8005	0.5792	0.7302	0.6639
	300	0.5606	0.8003	0.5787	0.727	0.6652
0.25	10	0.5222	0.783	0.5705	0.7245	0.663
	30	0.5281	0.789	0.5749	0.7064	0.6626
	50	0.527	0.7863	0.5729	0.7155	0.6572
	100	0.5274	0.788	0.5743	0.7187	0.6597
	300	0.5266	0.7878	0.5737	0.7154	0.6608
0.5	10	0.3541	0.6968	0.5342	0.6521	0.6417
	30	0.3581	0.7033	0.5388	0.6379	0.6367
	50	0.3573	0.7014	0.5371	0.6458	0.6329
	100	0.3575	0.7023	0.5381	0.6485	0.6343
	300	0.357	0.702	0.5376	0.6454	0.6354

Tabla 2.23: Simulaciones de los mid_1 a partir de $\beta(6, 10)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	76.3	97.6	92.6	95	91.9
	30	88	99.9	99.7	99.9	99.5
	50	93	100	99.9	100	100
	100	99.2	100	100	100	100
	300	100	100	100	100	100
0.0909	10	80.1	97.6	92.7	95.2	92.1
	30	91.9	99.9	99.7	99.9	99.6
	50	95.7	100	99.9	100	100
	100	99.9	100	100	100	100
	300	100	100	100	100	100
0.2	10	87.3	97.7	92.7	95.2	93.1
	30	98.2	99.9	99.8	99.9	99.8
	50	99.6	100	99.9	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.25	10	90.1	97.7	92.7	95.1	93.5
	30	99	99.9	99.8	99.9	99.9
	50	99.9	100	99.9	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	99.4	97.6	93.2	96	94.9
	30	100	99.9	99.7	100	100
	50	100	100	99.9	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.24: Simulaciones de los mid_1 a partir de $\beta(6, 10)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.7902	0.5162	0.5848
	30	0.8068	0.5241	0.5925	0.5716	0.5838
	50	0.7954	0.5177	0.5864	0.5639	0.5826
	100	0.797	0.5195	0.5876	0.5685	0.5845
	300	0.7972	0.5189	0.5862	0.5684	0.5843
0.0909	10	0.7556	0.5169	0.5848	0.561	0.5796
	30	0.7715	0.5249	0.5926	0.5721	0.5814
	50	0.7606	0.5185	0.5864	0.5643	0.5807
	100	0.7621	0.5203	0.5877	0.5686	0.5822
	300	0.7623	0.5197	0.5863	0.5684	0.5821
0.2	10	0.6684	0.5192	0.5849	0.5611	0.5718
	30	0.6826	0.5274	0.5928	0.5719	0.5745
	50	0.6729	0.5209	0.5865	0.564	0.5747
	100	0.6742	0.5226	0.5877	0.5679	0.5757
	300	0.6744	0.522	0.5864	0.568	0.5759
0.25	10	0.6285	0.5205	0.5849	0.5601	0.5685
	30	0.6418	0.5288	0.5929	0.5715	0.5713
	50	0.6327	0.5222	0.5865	0.5637	0.5713
	100	0.6339	0.5239	0.5878	0.5674	0.5725
	300	0.6341	0.5233	0.5864	0.5676	0.5724
0.5	10	0.4288	0.5306	0.5851	0.5496	0.5466
	30	0.4381	0.5397	0.5936	0.5577	0.5453
	50	0.4317	0.5326	0.5869	0.5505	0.545
	100	0.4324	0.5342	0.5881	0.5539	0.5461
	300	0.4326	0.5338	0.587	0.5547	0.5462

2.2.3. Tablas para el caso $k = 6$ Tabla 2.25: Simulaciones de los mid_1 a partir de $\beta(1, 1)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	68.3	92.8	100	98.1	100
	30	76.7	99.5	100	100	100
	50	82.4	100	100	100	100
	100	90.6	100	100	100	100
	300	98.9	100	100	100	100
0.0909	10	79.1	92.9	100	98.1	100
	30	91.6	99.4	100	100	100
	50	95.5	100	100	100	100
	100	98.8	100	100	100	100
	300	100	100	100	100	100
0.2	10	94.2	92.9	100	98.5	100
	30	99.6	99.4	100	100	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.25	10	97.3	92.9	100	98.6	100
	30	100	99.4	100	100	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	100	93.1	100	99.1	100
	30	100	99.4	100	100	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.26: Simulaciones de los mid_1 a partir de $\beta(1, 1)$

τ	n	$\frac{\text{ECMF}}{\text{ECML}}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{F})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{F})}$
		0.0476	10	0.9468	0.8501	0.4413
	30	0.9505	0.8538	0.4441	0.8017	0.4415
	50	0.951	0.8546	0.4452	0.7994	0.443
	100	0.9523	0.8555	0.445	0.8014	0.4415
	300	0.9505	0.8544	0.445	0.8014	0.4421
0.0909	10	0.904	0.8501	0.441	0.7998	0.4409
	30	0.9075	0.8538	0.4437	0.8	0.4417
	50	0.908	0.8545	0.4448	0.7975	0.4431
	100	0.9093	0.8554	0.4446	0.7996	0.4416
	300	0.9076	0.8544	0.4447	0.7995	0.4422
0.2	10	0.7962	0.8499	0.44	0.7934	0.4415
	30	0.7993	0.8536	0.4427	0.794	0.442
	50	0.7997	0.8543	0.4438	0.7919	0.4433
	100	0.8008	0.8552	0.4436	0.7937	0.4419
	300	0.7993	0.8542	0.4436	0.7937	0.4425
0.25	10	0.7467	0.8498	0.4394	0.7901	0.4418
	30	0.7497	0.8535	0.4421	0.7907	0.4422
	50	0.75	0.8542	0.4432	0.7888	0.4434
	100	0.7511	0.8552	0.443	0.7905	0.4421
	300	0.7497	0.8541	0.4431	0.7905	0.4427
0.5	10	0.4996	0.8489	0.4349	0.7665	0.4429
	30	0.5016	0.8528	0.4377	0.7676	0.4432
	50	0.5018	0.8535	0.4387	0.7658	0.4441
	100	0.5026	0.8545	0.4386	0.7662	0.4429
	300	0.5016	0.8534	0.4386	0.7665	0.4435

Tabla 2.27: Simulaciones de los mid_1 a partir de $\beta(0,75; 0,75)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	70.9	92.5	100	97.8	100
	30	83.1	99.8	100	100	100
	50	89.8	100	100	100	100
	100	96.8	100	100	100	100
	300	99.9	100	100	100	100
0.0909	10	81	92.5	100	97.9	100
	30	94.3	99.8	100	100	100
	50	98.5	100	100	100	100
	100	99.9	100	100	100	100
	300	100	100	100	100	100
0.2	10	96.5	92.5	100	97.9	100
	30	100	99.8	100	100	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.25	10	99	92.6	100	98.2	100
	30	100	99.8	100	100	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	100	92.9	100	99.6	100
	30	100	99.8	100	100	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.28: Simulaciones de los mid_1 a partir de $\beta(0,75; 0,75)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.9414	0.8626	0.4669
	30	0.9428	0.8639	0.4677	0.8147	0.463
	50	0.944	0.8658	0.4687	0.817	0.4642
	100	0.9427	0.8645	0.4678	0.8155	0.4631
	300	0.9421	0.8637	0.4677	0.8154	0.4633
0.0909	10	0.8988	0.8625	0.4665	0.8137	0.4636
	30	0.9001	0.8637	0.4674	0.8133	0.4631
	50	0.9013	0.8657	0.4683	0.8149	0.4642
	100	0.9001	0.8643	0.4675	0.8138	0.4631
	300	0.8995	0.8636	0.4673	0.8138	0.4634
0.2	10	0.7914	0.8621	0.4655	0.8087	0.4637
	30	0.7926	0.8634	0.4663	0.8086	0.4635
	50	0.7937	0.8653	0.4673	0.8096	0.4638
	100	0.7926	0.864	0.4664	0.8086	0.4633
	300	0.7921	0.8632	0.4663	0.8084	0.4635
0.25	10	0.7422	0.8619	0.4649	0.806	0.464
	30	0.7434	0.8632	0.4657	0.8055	0.4634
	50	0.7443	0.8651	0.4667	0.8069	0.4638
	100	0.7433	0.8638	0.4659	0.8056	0.4635
	300	0.7429	0.863	0.4657	0.8055	0.4636
0.5	10	0.4963	0.8603	0.4603	0.7833	0.4654
	30	0.497	0.8616	0.4612	0.7831	0.4637
	50	0.4977	0.8635	0.4622	0.786	0.4639
	100	0.497	0.8622	0.4614	0.7838	0.4635
	300	0.4967	0.8615	0.4612	0.7837	0.4637

Tabla 2.29: Simulaciones de los mid_1 a partir de $\beta(2, 2)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	52.8	88.6	100	94.4	100
	30	61.4	98.7	100	99.8	100
	50	66	99.9	100	100	100
	100	68.7	100	100	100	100
	300	79.9	100	100	100	100
0.0909	10	62.5	88.5	100	94.4	100
	30	77.6	98.7	100	99.8	100
	50	85.4	99.9	100	100	100
	100	91.7	100	100	100	100
	300	98.8	100	100	100	100
0.2	10	84.3	88.1	100	94.8	100
	30	97.7	98.6	100	99.9	100
	50	99.2	99.9	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.25	10	91	88	100	94.8	100
	30	99.4	98.6	100	100	100
	50	100	99.8	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	100	87.3	100	96.1	100
	30	100	98.3	100	100	100
	50	100	99.7	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.30: Simulaciones de los mid_1 a partir de $\beta(2, 2)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.9798	0.8301	0.3935
	30	0.9723	0.8248	0.3902	0.7766	0.3968
	50	0.9702	0.8233	0.3904	0.7741	0.3982
	100	0.9751	0.8256	0.3888	0.7762	0.3967
	300	0.9749	0.8258	0.3897	0.7763	0.3974
0.0909	10	0.9357	0.8303	0.3931	0.7733	0.3986
	30	0.9285	0.825	0.3899	0.7746	0.3972
	50	0.9265	0.8235	0.3901	0.7723	0.3983
	100	0.9313	0.8258	0.3885	0.7745	0.397
	300	0.9311	0.826	0.3894	0.7747	0.3976
0.2	10	0.8246	0.831	0.3922	0.7676	0.3995
	30	0.8183	0.8258	0.389	0.7705	0.3982
	50	0.8165	0.8242	0.3892	0.7677	0.399
	100	0.8207	0.8266	0.3876	0.7695	0.3976
	300	0.8205	0.8267	0.3885	0.7694	0.3984
0.25	10	0.7736	0.8314	0.3917	0.7655	0.3997
	30	0.7677	0.8262	0.3885	0.7674	0.3991
	50	0.7661	0.8246	0.3887	0.7652	0.3994
	100	0.77	0.827	0.3871	0.7667	0.3979
	300	0.7698	0.8272	0.388	0.7667	0.3988
0.5	10	0.519	0.8345	0.3877	0.7413	0.4015
	30	0.5151	0.8293	0.3847	0.7434	0.4021
	50	0.5139	0.8277	0.3847	0.7419	0.4026
	100	0.5165	0.8302	0.3832	0.7435	0.4012
	300	0.5165	0.8304	0.3841	0.7432	0.4019

Tabla 2.31: Simulaciones de los mid_1 a partir de $\beta(4, 2)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	65.6	86.8	99.9	93.4	99.9
	30	72.7	97.7	100	99.6	100
	50	79.7	99.4	100	100	100
	100	87	99.9	100	100	100
	300	98.1	100	100	100	100
0.0909	10	73.2	86.6	99.9	93.6	99.9
	30	82.3	97.7	100	99.8	100
	50	90.3	99.4	100	100	100
	100	97.5	99.9	100	100	100
	300	99.9	100	100	100	100
0.2	10	87.3	86.5	99.9	94.4	99.9
	30	97.7	97.7	100	99.9	100
	50	99.6	99.4	100	100	100
	100	100	99.9	100	100	100
	300	100	100	100	100	100
0.25	10	91.8	86.4	99.9	94.4	99.9
	30	99.4	97.6	100	100	100
	50	100	99.4	100	100	100
	100	100	99.9	100	100	100
	300	100	100	100	100	100
0.5	10	99.6	86.3	99.9	95.6	99.9
	30	100	97.1	100	100	100
	50	100	99.3	100	100	100
	100	100	99.9	100	100	100
	300	100	100	100	100	100

Tabla 2.32: Simulaciones de los mid_1 a partir de $\beta(4, 2)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.9203	0.7993	0.3901
	30	0.9331	0.8074	0.39	0.752	0.3808
	50	0.9303	0.8054	0.3899	0.749	0.3839
	100	0.9309	0.8071	0.3922	0.751	0.3838
	300	0.9297	0.8058	0.3909	0.7502	0.3825
0.0909	10	0.8791	0.7995	0.3897	0.7454	0.3851
	30	0.8914	0.8076	0.3896	0.7498	0.3802
	50	0.8887	0.8057	0.3895	0.7473	0.3833
	100	0.8892	0.8074	0.3918	0.7491	0.3833
	300	0.8881	0.8061	0.3905	0.7483	0.3821
0.2	10	0.7753	0.8003	0.3885	0.7389	0.3819
	30	0.7862	0.8085	0.3884	0.7445	0.3783
	50	0.7838	0.8065	0.3883	0.7418	0.3818
	100	0.7842	0.8082	0.3906	0.743	0.3819
	300	0.7833	0.8069	0.3893	0.7426	0.3805
0.25	10	0.7278	0.8008	0.3879	0.7354	0.3805
	30	0.7379	0.809	0.3877	0.7415	0.3773
	50	0.7357	0.807	0.3876	0.7386	0.381
	100	0.7361	0.8086	0.3899	0.7399	0.3811
	300	0.7352	0.8073	0.3886	0.7392	0.3797
0.5	10	0.4899	0.8043	0.3828	0.7108	0.3751
	30	0.4968	0.8127	0.3825	0.7162	0.3706
	50	0.4953	0.8107	0.3825	0.7131	0.3741
	100	0.4955	0.8121	0.3847	0.7144	0.3737
	300	0.4949	0.811	0.3834	0.7142	0.3726

Tabla 2.33: Simulaciones de los mid_1 a partir de $\beta(6, 1)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	92.2	82.8	63.6	84.6	74.3
	30	99.5	96.5	83.9	97.9	94.7
	50	99.9	98.8	89	99.3	97.8
	100	100	99.9	97.2	99.9	99.8
	300	100	100	100	100	100
0.0909	10	94.1	83.1	63.6	84.7	74.4
	30	99.7	96.7	83.9	97.8	94.9
	50	100	98.8	89.2	99.5	97.9
	100	100	99.9	97.2	99.9	99.8
	300	100	100	100	100	100
0.2	10	98.5	83.5	63.7	85	75.7
	30	100	97.1	84.3	97.8	95.4
	50	100	99	89.4	99.6	98.4
	100	100	99.9	97.2	99.9	99.8
	300	100	100	100	100	100
0.25	10	99.4	83.7	63.7	85.2	75.9
	30	100	97.4	84.4	98.1	95.7
	50	100	99.2	89.5	99.7	98.4
	100	100	99.9	97.2	99.9	99.8
	300	100	100	100	100	100
0.5	10	99.8	85.8	64.1	86.3	78.9
	30	100	98.8	85	99.2	97.1
	50	100	99.6	90.1	99.7	98.9
	100	100	100	97.4	100	99.9
	300	100	100	100	100	100

Tabla 2.34: Simulaciones de los mid_1 a partir de $\beta(6, 1)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.7167	0.7766	0.7181
	30	0.7151	0.7712	0.6982	0.7447	0.5874
	50	0.7214	0.7793	0.7127	0.7535	0.5923
	100	0.7178	0.7747	0.7058	0.7492	0.5891
	300	0.717	0.7743	0.7058	0.7482	0.5904
0.0909	10	0.685	0.7755	0.7174	0.7491	0.6023
	30	0.6834	0.7702	0.6975	0.743	0.5844
	50	0.6895	0.7782	0.712	0.7519	0.5885
	100	0.686	0.7736	0.7052	0.7476	0.586
	300	0.6853	0.7733	0.7051	0.7467	0.5873
0.2	10	0.6051	0.7723	0.7153	0.7443	0.5877
	30	0.6037	0.7671	0.6955	0.7388	0.5743
	50	0.6091	0.775	0.7098	0.7471	0.5786
	100	0.606	0.7705	0.7031	0.7432	0.5773
	300	0.6053	0.7701	0.703	0.7421	0.5784
0.25	10	0.5685	0.7705	0.7141	0.7423	0.5822
	30	0.5672	0.7655	0.6943	0.7364	0.5698
	50	0.5722	0.7733	0.7087	0.7448	0.5743
	100	0.5693	0.7688	0.7019	0.7408	0.5734
	300	0.5687	0.7684	0.7019	0.7397	0.5733
0.5	10	0.3854	0.7568	0.7051	0.7263	0.5431
	30	0.3846	0.7525	0.6855	0.7204	0.5344
	50	0.3879	0.7599	0.6996	0.7291	0.5382
	100	0.3859	0.7555	0.6931	0.7251	0.5386
	300	0.3856	0.7552	0.693	0.7239	0.537

Tabla 2.35: Simulaciones de los mid_1 a partir de $\beta(6, 10)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	74.6	90.2	99.3	92.1	99.9
	30	82.8	98.6	100	99.4	100
	50	92	99.9	100	100	100
	100	97.3	100	100	100	100
	300	100	100	100	100	100
0.0909	10	79	90.1	99.3	92.2	99.9
	30	88.4	98.6	100	99.4	100
	50	95.8	99.9	100	100	100
	100	99	100	100	100	100
	300	100	100	100	100	100
0.2	10	87.6	89.7	99.3	91.9	99.9
	30	96.8	98.5	100	99.3	100
	50	99.9	99.9	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.25	10	90.6	89.6	99.3	91.9	99.9
	30	98.7	98.5	100	99.3	100
	50	99.9	99.9	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	98.4	87.9	99.2	91.2	99.9
	30	100	98.3	100	99.5	100
	50	100	99.9	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.36: Simulaciones de los mid_1 a partir de $\beta(6, 10)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.8398	0.68	0.3136
	30	0.8543	0.6904	0.3174	0.6774	0.3783
	50	0.8474	0.6849	0.3151	0.6704	0.3754
	100	0.8512	0.6878	0.3165	0.6715	0.3755
	300	0.8519	0.6884	0.3171	0.6734	0.3757
0.0909	10	0.803	0.6811	0.3139	0.6655	0.3735
	30	0.817	0.6916	0.3177	0.6778	0.3798
	50	0.8103	0.6861	0.3153	0.6707	0.3763
	100	0.8139	0.6889	0.3167	0.6719	0.3771
	300	0.8146	0.6896	0.3173	0.6739	0.3772
0.2	10	0.7104	0.6847	0.3146	0.6661	0.3771
	30	0.7228	0.6953	0.3185	0.6785	0.3836
	50	0.7169	0.6897	0.3161	0.6714	0.38
	100	0.7201	0.6925	0.3175	0.6728	0.3811
	300	0.7206	0.6933	0.318	0.675	0.3816
0.25	10	0.668	0.6866	0.315	0.6668	0.3792
	30	0.6797	0.6974	0.3189	0.679	0.3863
	50	0.6741	0.6917	0.3165	0.6721	0.382
	100	0.677	0.6945	0.3179	0.6736	0.3834
	300	0.6776	0.6953	0.3185	0.6757	0.3839
0.5	10	0.4557	0.7023	0.3182	0.6698	0.3967
	30	0.4639	0.7137	0.3222	0.6787	0.404
	50	0.4599	0.7077	0.3197	0.6721	0.3997
	100	0.4619	0.7105	0.3211	0.6732	0.401
	300	0.4623	0.7113	0.3217	0.6761	0.4024

2.2.4. Tablas para el caso $k = 7$ Tabla 2.37: Simulaciones de los mid_1 a partir de $\beta(1, 1)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	71.7	85.4	74.8	100	100
	30	86.4	98.5	91.1	100	100
	50	89.4	99.3	95.4	100	100
	100	96.2	100	99.6	100	100
	300	99.9	100	100	100	100
0.0909	10	82.2	85.4	74.9	100	100
	30	95.9	98.6	91.2	100	100
	50	98.9	99.3	95.6	100	100
	100	100	100	99.6	100	100
	300	100	100	100	100	100
0.2	10	95.9	85.4	75.1	100	100
	30	100	98.6	91.2	100	100
	50	100	99.3	95.7	100	100
	100	100	100	99.6	100	100
	300	100	100	100	100	100
0.25	10	98.6	98.4	89.6	100	100
	30	100	100	99.3	100	100
	50	100	100	99.9	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	100	100	95.8	100	100
	30	100	100	100	100	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.38: Simulaciones de los mid_1 a partir de $\beta(1, 1)$

τ	n	$\frac{\text{ECMF}}{\text{ECML}}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{F})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{F})}$
		0.0476	10	0.9399	0.8436	0.9053
	30	0.9403	0.8465	0.9074	0.6114	0.7096
	50	0.943	0.8502	0.9126	0.6099	0.7097
	100	0.9447	0.8512	0.9128	0.6107	0.7099
	300	0.9451	0.8518	0.9132	0.6098	0.7092
0.0909	10	0.8974	0.8435	0.9052	0.6095	0.7094
	30	0.8978	0.8465	0.9073	0.6109	0.7092
	50	0.9004	0.8501	0.9124	0.6094	0.7096
	100	0.902	0.8512	0.9127	0.6103	0.7098
	300	0.9024	0.8517	0.9131	0.6094	0.709
0.2	10	0.7903	0.8434	0.9048	0.6082	0.709
	30	0.7907	0.8463	0.9069	0.6102	0.7088
	50	0.793	0.85	0.912	0.6084	0.7088
	100	0.7944	0.851	0.9123	0.6092	0.7093
	300	0.7947	0.8516	0.9127	0.6083	0.7085
0.25	10	0.7413	0.64	0.7779	0.6079	0.7085
	30	0.7416	0.6317	0.7745	0.6093	0.7081
	50	0.7438	0.6359	0.7812	0.6078	0.7085
	100	0.7451	0.6358	0.7816	0.6086	0.7089
	300	0.7454	0.6353	0.7817	0.6077	0.7081
0.5	10	0.4959	0.5107	0.6888	0.6047	0.7061
	30	0.4962	0.5068	0.682	0.6039	0.706
	50	0.4977	0.5058	0.6848	0.6027	0.7063
	100	0.4985	0.5078	0.6875	0.6036	0.7072
	300	0.4987	0.5076	0.6867	0.6024	0.7057

Tabla 2.39: Simulaciones de los mid_1 a partir de $\beta(0,75; 0,75)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	77.2	91.8	81.5	100	99.9
	30	90.1	99.2	94.6	100	100
	50	94.4	100	97.5	100	100
	100	99.3	100	99.8	100	100
	300	100	100	100	100	100
0.0909	10	87.5	91.8	81.5	100	99.9
	30	98.1	99.2	94.7	100	100
	50	99.6	100	97.5	100	100
	100	100	100	99.8	100	100
	300	100	100	100	100	100
0.2	10	98.6	91.7	81.5	100	99.9
	30	100	99.3	94.9	100	100
	50	100	100	97.5	100	100
	100	100	100	99.8	100	100
	300	100	100	100	100	100
0.25	10	99.4	99.2	92.5	100	99.9
	30	100	100	99.4	100	100
	50	100	100	99.9	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	100	100	97.5	100	100
	30	100	100	100	100	100
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.40: Simulaciones de los mid_1 a partir de $\beta(0,75; 0,75)$

τ	n	$\frac{\text{ECMF}}{\text{ECML}}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{F})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{F})}$
		0.0476	10	0.9333	0.8259	0.8996
	30	0.9362	0.8309	0.9028	0.6254	0.7287
	50	0.9388	0.8331	0.9073	0.6264	0.7291
	100	0.9357	0.8306	0.9043	0.6274	0.7302
	300	0.9369	0.8316	0.9047	0.627	0.73
0.0909	10	0.891	0.8258	0.8995	0.6299	0.7341
	30	0.8939	0.8308	0.9027	0.6251	0.7289
	50	0.8963	0.8331	0.9072	0.626	0.7291
	100	0.8934	0.8306	0.9042	0.6271	0.7301
	300	0.8945	0.8315	0.9046	0.6266	0.7298
0.2	10	0.7846	0.8258	0.8993	0.6288	0.733
	30	0.7871	0.8307	0.9024	0.6244	0.7289
	50	0.7893	0.833	0.907	0.6249	0.7288
	100	0.7867	0.8305	0.9039	0.626	0.7297
	300	0.7877	0.8315	0.9044	0.6256	0.7294
0.25	10	0.7359	0.6506	0.7942	0.628	0.7321
	30	0.7382	0.6521	0.7956	0.6237	0.7286
	50	0.7402	0.6547	0.7988	0.6243	0.7287
	100	0.7378	0.6542	0.7951	0.6255	0.7296
	300	0.7388	0.6537	0.7963	0.625	0.7291
0.5	10	0.492	0.5427	0.7141	0.6231	0.729
	30	0.4936	0.5383	0.7126	0.6202	0.7275
	50	0.4949	0.5399	0.7157	0.6196	0.7267
	100	0.4933	0.5378	0.7141	0.6212	0.7281
	300	0.494	0.5389	0.7151	0.6206	0.7274

Tabla 2.41: Simulaciones de los mid_1 a partir de $\beta(2, 2)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	61.8	76	68.9	99.9	99.5
	30	71.4	89.3	81	100	100
	50	77	95.3	87.2	100	100
	100	83.4	98.9	95.3	100	100
	300	97.3	100	99.7	100	100
0.0909	10	73.6	76.1	68.9	99.9	99.5
	30	86.9	89.3	81.2	100	100
	50	93	95.3	87.3	100	100
	100	98	98.9	95.3	100	100
	300	100	100	99.8	100	100
0.2	10	92.8	76.2	69.1	99.9	99.5
	30	99.8	89.7	81.5	100	100
	50	100	95.4	87.6	100	100
	100	100	98.9	95.5	100	100
	300	100	100	99.8	100	100
0.25	10	96	96	82.9	99.9	99.5
	30	99.9	100	97.8	100	100
	50	100	100	99.7	100	100
	100	100	100	99.9	100	100
	300	100	100	100	100	100
0.5	10	100	100	91.1	99.9	99.7
	30	100	100	99.9	100	100
	50	100	100	99.9	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.42: Simulaciones de los mid_1 a partir de $\beta(2, 2)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.9585	0.8849	0.9186
30	0.956		0.8845	0.9209	0.5866	0.6772
50	0.9564		0.8841	0.9192	0.5857	0.6764
100	0.9595		0.8867	0.9233	0.586	0.6764
300	0.9561		0.8831	0.9197	0.5861	0.6762
0.0909	10	0.9154	0.8847	0.9184	0.5845	0.6775
	30	0.9129	0.8844	0.9206	0.5864	0.6767
	50	0.9133	0.8839	0.9189	0.5853	0.6761
	100	0.9163	0.8866	0.923	0.5856	0.6759
	300	0.9131	0.8829	0.9194	0.5857	0.6759
0.2	10	0.8066	0.8842	0.9176	0.5822	0.676
	30	0.8045	0.8838	0.9198	0.5848	0.6758
	50	0.8049	0.8833	0.9181	0.5842	0.675
	100	0.8075	0.886	0.9222	0.5848	0.675
	300	0.8046	0.8824	0.9187	0.5847	0.6752
0.25	10	0.7568	0.5973	0.7402	0.5816	0.6753
	30	0.7548	0.5961	0.7428	0.5843	0.6756
	50	0.7552	0.5914	0.7396	0.5837	0.6745
	100	0.7577	0.5955	0.7431	0.5843	0.6745
	300	0.755	0.5951	0.7433	0.5841	0.6747
0.5	10	0.5076	0.452	0.6239	0.5762	0.6712
	30	0.5063	0.4481	0.6256	0.5797	0.6718
	50	0.5066	0.4484	0.6232	0.5788	0.6705
	100	0.5083	0.4503	0.6276	0.5797	0.6709
	300	0.5065	0.4495	0.6259	0.5795	0.6707

Tabla 2.43: Simulaciones de los mid_1 a partir de $\beta(4, 2)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	63.8	72.4	38.3	99.7	98.9
	30	76.1	89.3	30.8	100	100
	50	83.4	93.4	27.5	100	100
	100	94.8	99.2	20.5	100	100
	300	99	100	8.1	100	100
0.0909	10	72.4	72.6	38.6	99.7	98.9
	30	87.8	89.4	30.7	100	100
	50	93.7	93.4	27.6	100	100
	100	99.4	99.2	20.7	100	100
	300	100	100	8.2	100	100
0.2	10	89.1	72.9	38.8	99.7	98.6
	30	99	89.6	30.9	100	100
	50	99.9	93.5	28.2	100	100
	100	100	99.2	21.2	100	100
	300	100	100	8.5	100	100
0.25	10	93.8	97.4	40.7	99.7	98.9
	30	100	100	42.2	100	100
	50	100	100	48.1	100	100
	100	100	100	56.7	100	100
	300	100	100	65.6	100	100
0.5	10	100	100	44.3	100	98.1
	30	100	100	52.5	100	100
	50	100	100	62	100	100
	100	100	100	78.1	100	100
	300	100	100	89.6	100	100

Tabla 2.44: Simulaciones de los mid_1 a partir de $\beta(4, 2)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.935	0.8661	1.0429
	30	0.9334	0.8641	1.0457	0.5728	0.6792
	50	0.9325	0.8657	1.0473	0.5705	0.6778
	100	0.9286	0.8598	1.0413	0.5711	0.6783
	300	0.9309	0.8629	1.0445	0.5721	0.6775
0.0909	10	0.8931	0.8659	1.0425	0.5708	0.676
	30	0.8916	0.8639	1.0454	0.5723	0.6797
	50	0.8908	0.8655	1.047	0.5702	0.678
	100	0.887	0.8597	1.041	0.5706	0.6786
	300	0.8893	0.8627	1.0442	0.5716	0.6776
0.2	10	0.7876	0.8653	1.0414	0.5684	0.676
	30	0.7863	0.8634	1.0445	0.571	0.6809
	50	0.7856	0.8649	1.046	0.568	0.6788
	100	0.7823	0.8591	1.04	0.5694	0.6795
	300	0.7843	0.8621	1.0432	0.5699	0.6787
0.25	10	0.7392	0.5641	0.9696	0.5681	0.6762
	30	0.738	0.5663	0.9778	0.5704	0.6812
	50	0.7374	0.5651	0.9755	0.5666	0.6792
	100	0.7343	0.5639	0.9706	0.5684	0.6799
	300	0.7362	0.5633	0.9732	0.5692	0.6792
0.5	10	0.4974	0.4362	0.905	0.5606	0.6805
	30	0.4968	0.4362	0.9158	0.563	0.6847
	50	0.4964	0.4332	0.917	0.5586	0.6824
	100	0.4943	0.4342	0.9101	0.5613	0.6837
	300	0.4956	0.4341	0.9137	0.5616	0.6826

Tabla 2.45: Simulaciones de los mid_1 a partir de $\beta(6, 1)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	88.8	97.2	50	96.3	83.1
	30	99.2	100	43.7	100	95.1
	50	99.9	100	40.4	100	98.9
	100	100	100	39.8	100	99.9
	300	100	100	32.4	100	100
0.0909	10	91.9	97.2	50	96.3	82.9
	30	99.7	100	43.8	100	95
	50	100	100	40.4	100	98.8
	100	100	100	39.6	100	99.9
	300	100	100	32.1	100	100
0.2	10	96.9	97.2	49.9	96.2	82.6
	30	100	100	43.5	100	94.8
	50	100	100	40.5	100	98.5
	100	100	100	39.7	100	99.9
	300	100	100	31.8	100	100
0.25	10	98.5	98.2	49.7	95.9	82.2
	30	100	100	43.8	100	94.7
	50	100	100	40.7	100	98.5
	100	100	100	39.6	100	99.9
	300	100	100	35.1	100	100
0.5	10	99.7	82.7	50	95.2	79.9
	30	100	96.4	43	100	93.1
	50	100	99	40.4	100	97.8
	100	100	100	39.3	100	99.7
	300	100	100	36.2	100	100

Tabla 2.46: Simulaciones de los mid_1 a partir de $\beta(6, 1)$

τ	n	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECML}}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{L})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_1(\text{F})}}$	$\frac{\overline{\text{ECMF}}}{\overline{\text{ECMC}_2(\text{F})}}$
		0.0476	10	0.7812	0.6121	1.027
	30	0.7725	0.6032	1.0224	0.6119	0.7718
	50	0.7755	0.6063	1.0263	0.6183	0.776
	100	0.7749	0.605	1.0228	0.6126	0.7742
	300	0.7714	0.6028	1.0204	0.6114	0.7734
0.0909	10	0.7466	0.6123	1.0271	0.611	0.7647
	30	0.7384	0.6035	1.0225	0.6122	0.7723
	50	0.7412	0.6066	1.0263	0.6181	0.777
	100	0.7406	0.6053	1.0229	0.6124	0.7754
	300	0.7373	0.6031	1.0204	0.6114	0.7746
0.2	10	0.6595	0.6131	1.0273	0.6102	0.769
	30	0.6523	0.6043	1.0227	0.6117	0.7748
	50	0.6547	0.6073	1.0264	0.6178	0.7799
	100	0.6542	0.6061	1.023	0.6119	0.7779
	300	0.6514	0.6039	1.0205	0.6114	0.778
0.25	10	0.6196	0.5509	1.0251	0.6098	0.7712
	30	0.6129	0.5478	1.0195	0.6114	0.7759
	50	0.6151	0.5526	1.0231	0.618	0.7816
	100	0.6147	0.5517	1.0188	0.6122	0.7804
	300	0.612	0.5481	1.0167	0.6114	0.7797
0.5	10	0.42	0.5989	1.0239	0.606	0.7867
	30	0.4157	0.6009	1.0165	0.6102	0.7957
	50	0.417	0.6034	1.0181	0.6168	0.7946
	100	0.4167	0.6032	1.0173	0.6117	0.7968
	300	0.4149	0.5989	1.0143	0.6103	0.7934

Tabla 2.47: Simulaciones de los mid_1 a partir de $\beta(6, 10)$

τ	n	%ECMF <				
		ECML	ECMC ₁ (L)	ECMC ₂ (L)	ECMC ₁ (F)	ECMC ₂ (F)
0.0476	10	71.1	79.3	86.5	96.7	94.9
	30	85.9	94.2	98.7	100	99.9
	50	91.6	96.8	99.4	100	100
	100	97.6	99.8	100	100	100
	300	100	100	100	100	100
0.0909	10	76.2	79.3	86.7	96.7	95.2
	30	92.9	94.3	98.7	100	99.9
	50	95.8	96.8	99.4	100	100
	100	99.2	99.8	100	100	100
	300	100	100	100	100	100
0.2	10	87.2	80	87	96.7	95.7
	30	98.9	94.5	98.8	100	99.9
	50	99.6	96.9	99.5	100	100
	100	100	99.8	100	100	100
	300	100	100	100	100	100
0.25	10	90.8	92.4	90.8	96.6	95.1
	30	99.5	99.8	99.5	99.9	99.9
	50	99.7	100	99.9	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100
0.5	10	99.4	97.1	94.7	97.6	93.2
	30	100	100	99.9	100	99.9
	50	100	100	100	100	100
	100	100	100	100	100	100
	300	100	100	100	100	100

Tabla 2.48: Simulaciones de los mid_1 a partir de $\beta(6, 10)$

τ	n	$\frac{\text{ECMF}}{\text{ECML}}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{L})}$	$\frac{\text{ECMF}}{\text{ECMC}_1(\text{F})}$	$\frac{\text{ECMF}}{\text{ECMC}_2(\text{F})}$
		0.0476	10	0.8623	0.7919	0.7383
	30	0.8607	0.7944	0.737	0.563	0.6321
	50	0.8575	0.7905	0.7348	0.5625	0.6318
	100	0.8579	0.7898	0.7345	0.5632	0.6329
	300	0.8562	0.7894	0.7329	0.5656	0.6361
0.0909	10	0.8246	0.7914	0.7379	0.5564	0.6254
	30	0.8231	0.7939	0.7366	0.5633	0.6328
	50	0.82	0.7901	0.7344	0.5618	0.6327
	100	0.8203	0.7893	0.7341	0.5627	0.6332
	300	0.8187	0.7889	0.7326	0.5649	0.6367
0.2	10	0.7295	0.79	0.7368	0.5567	0.6268
	30	0.7282	0.7924	0.7356	0.5616	0.6357
	50	0.7256	0.7887	0.7334	0.5605	0.6352
	100	0.7258	0.7878	0.7331	0.5607	0.6357
	300	0.7243	0.7874	0.7315	0.5631	0.6394
0.25	10	0.6859	0.5041	0.4835	0.5569	0.6292
	30	0.6847	0.5001	0.4722	0.561	0.6372
	50	0.6823	0.4963	0.4723	0.5597	0.6375
	100	0.6824	0.4949	0.4708	0.5598	0.6371
	300	0.6811	0.4963	0.4715	0.5621	0.6407
0.5	10	0.468	0.3769	0.3635	0.5477	0.6393
	30	0.4673	0.3733	0.3593	0.5523	0.6472
	50	0.4658	0.3721	0.3585	0.5501	0.6477
	100	0.4657	0.3734	0.3596	0.5496	0.6454
	300	0.4647	0.3717	0.3575	0.5513	0.6495

2.3. Conclusiones empíricas obtenidas a partir de las simulaciones

Como ya hemos apuntado a lo largo del trabajo, el objetivo del mismo ha sido extraer algunas conclusiones que corroboren la percepción intuitiva de que con el empleo de los cuestionarios de respuesta libre basados en la escala de números fuzzy se alcanza una riqueza de información y una fidelidad y flexibilidad en la formalización de los datos y su tratamiento estadístico posterior, superiores a los que se tendría con el empleo de cuestionarios basados en escalas Likert o en la escala de números fuzzy con formato prefijado.

Por lo que concierne a la *diversidad* las conclusiones son contundentes:

Como probamos en la Proposición 2.1.1, el valor del índice de diversidad de Gini-Simpson es siempre superior en la escala libre de números fuzzy que en la escala Likert o en la escala de números fuzzy con formato prefijado. Una traducción plausible de este hecho es que las respuestas a los cuestionarios fuzzy de formato libre proporcionan información más rica, captan mejor la subjetividad propia de las respuestas y, desde un punto de vista estadístico, permiten una explotación más interesante de los datos resultantes.

En lo que respecta a la *representatividad de la medida de tendencia central* que hemos estudiado, la media (tipo Aumann), podemos afirmar que en una inmensa mayoría de casos es más representativa en el caso fuzzy libre que el caso Likert o en el caso fuzzy prefijado, como podemos comprobar en las tablas anteriores a través del ECM.

Son muchos los problemas que el estudio de este trabajo plantea de forma inmediata y que esperan investigarse en un futuro próximo.

Por un lado, analizar si la elección del valor del parámetro τ cuando se emplea la distancia D_τ influye en los resultados, y si es así, de qué manera.

Por otro lado, examinar si la elección de un determinado conjunto de números fuzzy prefijados con los que vamos a identificar los fuzzy libres hace variar los resultados que se obtendrían si eligiésemos otro conjunto de fuzzy prefijados.

Finalmente, en relación con las herramientas de comparación, sería conveniente realizar también una discusión comparativa recurriendo a criterios basados en contrastes de hipótesis (por ejemplo, comparaciones de las conclusiones y la potencias asociadas a la aplicación del test ANOVA para datos fuzzy de González-Rodríguez *et al.* (2011) frente al test de Kruskal-Wallis u otros para las categorías Likert). La clave estará en encontrar los procedimientos y conceptos en los que se fundamente la comparación.

Bibliografía

- [1] O. Aydin, F. Pakdil (2008). Fuzzy SERVQUAL Analysis in Airline Services. *Organizacija* **41(3)**, 108–115.
- [2] S. Bajpai, A. Sachdeva, J.P. Gupta (2010). Security risk assessment: Applying the concepts of fuzzy logic. *J. Hazard. Mater.* **173**, 258-264.
- [3] B. Bharadwaj (2007). Development of a fuzzy Likert scale for the WHO ICF to include categorical definitions on the basis of a continuum. *ETD Collection for Wayne State University. Paper AAI1442894*. (<http://digitalcommons.wayne.edu/dissertations/AAI1442894>).
- [4] C. Bertoluzza, N. Corral, A. Salas (1995). On a new class of distances between fuzzy numbers. *Math & Soft Comput* **2**, 71–84.
- [5] V.N. Bobylev (1985). Support function of a fuzzy set and its characteristic properties. *Mathematical Notes* **37**, 281–285.
- [6] F.A. Bocklisch, S.F. Bocklisch, J.F. Krems (2010). How to translate words into numbers? A fuzzy approach for the numerical translation of verbal probabilities. En: *Computational Intelligence for Knowledge-Based Systems Design* (E. Hüllermeier, R. Kruse, F. Hoffmann, Eds.) (LNAI 6178). Springer, Heidelberg. pp. 614–623.
- [7] F.A. Bocklisch (2011). The vagueness of verbal probability and frequency expressions. *Int. J. Adv. Comp. Sci.* **1(2)**, 52–57.
- [8] J. Cai, Q. Li, L. Li, H. Peng, Y. Yang (2012). A hybrid FCASO-SQP method for solving the economic dispatch problems with valve-point effects. *Energy* **38**, 346-353.
- [9] C. Castaing, M. Valadier (1977). Convex Analysis and Measurable Multifunctions. *Lec. Notes in Math.* 580. Springer-Verlag, Berlin.
- [10] C.C. Chou, L.J. Liu, S.F. Huang, J.M. Yih, T.C. Han (2011). An evaluation of airline service quality using the fuzzy weighted SERVQUAL method. *Appl. Soft. Comp.* **11**, 2117–2128.

- [11] S. de la Rosa de Saa (2011a). *Las escalas Likert y fuzzy: comparación empírica por medidas de error estadístico*. Trabajo Integrado en el Plan de Estudios y Académicamente Dirigido. Universidad de Oviedo.
- [12] S. de la Rosa de Saa (2011b). *Representatividad de la mediana para datos en escalas Likert y fuzzy: estudio empírico comparativo*. Tesina de Licenciatura en Matemáticas. Universidad de Oviedo.
- [13] P. Diamond, P., Kloeden (1999). Metric spaces of fuzzy sets. *Fuzzy Sets Syst.* **100**, 63–71.
- [14] D. Dubois, H. Prade (1978). Operations on fuzzy numbers. *Int. J. Syst. Sci.* **9**, 613–626.
- [15] M.B. Ferraro, R. Coppi, G. González-Rodríguez, A. Colubi (2010). A linear regression model for imprecise response. *Int. J. Appr. Reas.* **51**, 759–770.
- [16] M.A. Gil, G. González-Rodríguez (2012). Fuzzy vs Likert scale in Statistics. En: *Combining Experimentation and Theory (A Hommage to Abe Mamdani)* (E. Trillas, P. Bonissone, L. Magdalena, J. Kacprzyk, Eds.). Series Studies in Fuzziness and Soft Computing, 271. Springer, Heidelberg, pp. 407-420.
- [17] M.A. Gil, M.A. Lubiano, M. Montenegro, M.T. López-García (2002). Least squares fitting of an affine function and strength of association for interval-valued data. *Metrika* **56**, 97-111.
- [18] M.A. Gil, M. Montenegro, G. González-Rodríguez, A. Colubi, M.R. Casals (2006). Bootstrap approach to the multi-sample test of means with imprecise data. *Comp. Stat. Data Anal.* **51**, 148–162.
- [19] E. Giné, J. Zinn (1990). Bootstrapping general empirical measures. *Ann. Probab.* **18**, 851–869.
- [20] G. González-Rodríguez, M. Montenegro, A. Colubi, M.A. Gil (2006). Bootstrap techniques and fuzzy random variables: Synergy in hypothesis testing with fuzzy data. *Fuzzy Sets Syst.* **157**, 2608–2613.
- [21] G. González-Rodríguez, A. Colubi, M.A. Gil (2011). Fuzzy data treated as functional data. A one-way ANOVA test approach. *Comp. Stat. Data Anal.*, **56**(4), 943-955.
- [22] G. González-Rodríguez, A. Blanco, A. Colubi, M.A. Lubiano (2009). Estimation of a simple linear regression model for fuzzy random variables. *Fuzzy Sets Syst.* **160**, 357-370.

- [23] N. Grégoire, M. Bouillot (1998). Hausdorff distance between convex polygons. (<http://cgm.cs.mcgill.ca/godfried/teaching/cg-projects/98/normand/main.html>).
- [24] F. Herrera (2008). Genetic fuzzy systems: taxonomy, current research trends and prospects. *Evol Intel*, **1**(1), 27–46.
- [25] B. Hesketh, B. Griffin, V. Loh (2011). A future-oriented retirement transition adjustment framework. *J. Vocat. Behav.* **79**, 303–314.
- [26] B. Hesketh, T. Hesketh, J.-I.Hansen, D.Goranson (1995). Use of fuzzy variables in developing new scales from strong interest inventory. *J. Couns. Psych.* **42**, 85–99.
- [27] T. Hesketh, B. Hesketh (1994). Computerised fuzzy ratings: the concept of a fuzzy class. *Behav. Res. Meth., Inst. & Comp.* **26**, 272–277.
- [28] T. Hesketh, R.G.L. Pryor, B. Hesketh (1988). An application of a computerised fuzzy graphic rating scale to the psychological measurement of individual differences. *Int. J. Man. Mach. Stud.* **29**, 21–35.
- [29] H.-Y. Hu, Y.-C. Lee, T.-M. Yen (2010). Service quality gaps analysis based on Fuzzy linguistic SERVQUAL with a case study in hospital out-patient services. *The TQM Journal* **22**, 499–515.
- [30] P. Kambaki-Vougioukli, A. Karakos, N. Lygeros, T. Vougiouklis (2011). Fuzzy instead of discrete. *Ann. Fuzzy Math. Inform.* **2**(1), 81–89.
- [31] E.P. Klement, M.L. Puri, D.A. Ralescu (1986). Limit theorems for fuzzy random variables. *Proc. Roy. Soc. London A* **407**, 171–182.
- [32] R. Körner (1997). *Linear Models with Random Fuzzy Variables*. PhD Thesis. University of Freiberg. (<http://www.mathe.tu-freiberg.de/inst/stoch/Koerner/dissa.ps>).
- [33] M. Lalla, G. Facchinetti, G. Mastroleo (2004). Ordinal scales and fuzzy set systems to measure agreement: an application to the evaluation of teaching activity. *Quality & Quantity* **38**, 577–601.
- [34] M.A. Lazim, M.T.A. Osman (2009). Measuring teachers' beliefs about Mathematics: a fuzzy set approach. *Int. J. Soc. Sci.* **4**, 39–43.
- [35] C.-Q. Li (2010). *A New Likert Scale Based on Fuzzy Sets Theory*. PhD Thesis. University of Connecticut. (<http://digitalcommons.uconn.edu/dissertations/AAI3451393/>).
- [36] R. Likert (1932). A technique for the measurement of attitudes. *Arch. Psychol.* **140**, 1–55.

- [37] L.M. Lozano, E. García-Cueto, J. Muñiz (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology* **4**, 73–79.
- [38] M.A. Lubiano (1999). *Medidas de Variación de Elementos Aleatorios Imprecisos*. Tesis Doctoral. Universidad de Oviedo. (<http://www.tesisenred.net/TDR-0209110-122449/>).
- [39] M.A. Lubiano, M.A. Gil (1999). Estimating the expected value of fuzzy random variables in random samplings from finite populations. *Stat. Pap.* **40**, 277–295.
- [40] M. Montenegro (2003). *Estadística con Datos Imprecisos Basada en una Métrica Generalizada*. Tesis Doctoral. Universidad de Oviedo. (<http://www.tesisenred.net/TDR-0209110-120109/>).
- [41] M. Montenegro, M.R. Casals, M.A. Lubiano, M.A. Gil (2001). Two-sample hypothesis tests of means of a fuzzy random variable. *Inform. Sci* **133**, 89–100.
- [42] M. Montenegro, A. Colubi, M.R. Casals, M.A. Gil (2004). Asymptotic and Bootstrap techniques for testing the expected value of a fuzzy random variable. *Metrika* **59**, 31–49.
- [43] I.A. Motawa, C.J. Anumba, A. El-Hamalawi (2006). A fuzzy system for evaluating the risk of change in construction projects. *Adv. Eng. Softw.* **37**, 583–591.
- [44] S. Nadarajah (2006). Explicit expressions for moments of beta order statistics. *Metron* **LXIV**, 377–384.
- [45] C.V. Negoita, D.A. Ralescu (1975). *Applications of Fuzzy Sets to Systems Analysis*. Stuttgart and Halsted Press, New York.
- [46] E. Plebankiewicz (2012). A fuzzy sets based contractor prequalification procedure. *Automat. Constr.* **22**, 433–443.
- [47] M.L. Puri, D.A. Ralescu (1985). The concept of normality for fuzzy random variables. *Ann. Probab.* **11**, 1373–1379
- [48] M.L. Puri, D.A. Ralescu (1986). Fuzzy random variables. *J. Math. Anal. Appl.* **114**, 409–422
- [49] M. Saltan, S. Saltan, A. Sahiner (2007). Fuzzy logic modeling of deflection behavior against dynamic loading in flexible pavements. *Constr. Build. Mater.* **21**, 1406–1414.
- [50] B. Sinova, M.A. Gil, A. Colubi, S. Van Aelst (2012a). Median of a random fuzzy number. The 1-norm distance approach. *Fuzzy Sets Syst.* **200**, 99–115.

- [51] B. Sinova, S. de la Rosa de Súa, M.A. Gil, S. Van Aelst (2012b). The left/right spreads L^1 distance-based approach to the median of a random fuzzy number. (Enviado para publicación).
- [52] G. Tayfur, S. Ozdemir, V.P. Singh (2003). Fuzzy logic algorithm for runoff-induced sediment transport from bare soil surfaces. *Adv. Water Resour.* **26**, 1249-1256.
- [53] W. Trutschnig, G. González-Rodríguez, A. Colubi, M.A. Gil (2009). A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread. *Inform. Sci.* **179**, 3964–3972.
- [54] I.B. Turksen, I.A. Willson (1994). A fuzzy set preference model for consumer choice. *Fuzzy Sets Syst.* **68**, 253–266.
- [55] H. van Laerhoven, H.J. van der Zaag-Loonen, B.H.F. Derkx (2004). A comparison of Likert scale and visual analogue scales as response options in children’s questionnaires. *Acta Pædiatr.* **93**, 830-835.
- [56] C.H. Yeh, R.J. Willis, H. Deng, H. Pan (1999). Task oriented weighting in multi-criteria analysys. *Eur. J. Oper. Res.* **119**, 130-146.
- [57] A. Yücel, A.F. Güneri (2011). A weighted additive fuzzy programming approach for multi-criteria supplier selection. *Expert Syst. Appl.* **38**, 6281-6286.
- [58] L.A. Zadeh (1965). Fuzzy sets. *Inform. Contr.* **7**, 338–353.
- [59] L.A. Zadeh (1975). The concept of a linguistic variable and its application to approximate reasoning, Part 1. *Inform. Sci.* **8**, 199–249; Part 2. *Inform. Sci.* **8**, 301–353; Part 3. *Inform. Sci.* **9**, 43–80

Índice temático

- D_τ -error cuadrático medio muestral de un NFA respecto de su media, 38
- α -nivel de un número fuzzy, 10
- aritmética con números fuzzy, 14
- distancia entre números fuzzy
 - distancia generalizada tipo L^2 , D_τ , 20
- error cuadrático medio muestral de un NFA Likertizado, 39
- escala de números fuzzy, 9
 - cuestionarios basados en, 9
- escala Likert, 5
- escalas Likert
 - cuestionarios basados en, 5
- estructura semilineal de $(\mathcal{F}_c(\mathbb{R}), +, \cdot)$, 16
- índice de diversidad de Gini-Simpson muestral de un NFA, 38
- índice de diversidad de Gini-Simpson muestral de un NFA Likertizado, 38
- número fuzzy, 9
 - α -nivel de un, 10
 - función soporte de un, 17
- número fuzzy aleatorio, 22
 - D_τ -error cuadrático medio, 26
 - distribución inducida por un, 22
 - implicaciones de medibilidad, 22
 - independencia entre números fuzzy aleatorios, 22
 - índice de diversidad de Gini-Simpson, 25
 - media o esperanza tipo Aumann de un, 23
- números fuzzy
 - aritmética con, 14
 - distancia entre, 20
 - producto de un escalar por, 14
 - suma de, 14
- variables lingüísticas, 4