Robust location measures for imprecise-valued random elements

Beatriz Sinova Fernández

Universidad de Oviedo

Doctoral Programme
in Mathematics and Statistics

PhD Dissertation

Robust location measures for imprecise-valued random elements

Beatriz Sinova Fernández

Universidad de Oviedo

Doctoral Programme
in Mathematics and Statistics

UNIVERSITEIT GENT

Doctorate and Doctoral
Training Programme
in Sciences: Mathematics

PhD Dissertation

supervised by

María Ángeles Gil, Gil González Rodríguez
University of Oviedo

and Stefan Van Aelst
University of Ghent

A mis padres y a Manu, mi hermano

Acknowledgements

"I deem that there are lyric days
so ripe with radiance and cheer,
so rich with gratitude and praise
that they enrapture all the year."
(Robert William Service)

I would not like to begin this work without the explicit acknowledgements to all the people who have contributed, in different ways, to make it possible. First, I would like to thank the dedication of my three supervisors, Ángeles, Gil and Stefan. Ángeles' constant and deep guidance has been as crucial as her unconditional human support and there are no words, in any of the languages used in this Doctoral Thesis, to say how much this help has meant to me. My love of programming was combined with Gil's valuable ideas and simulation techniques in my first contact with empirical analysis of this kind. Those first codes suggested by Gil gave me the chance to grow and acquire independence in simulation tasks. However, this definitely would not have been enough help without Stefan's knowledge of Robust Statistics, whose strong direction and useful proposals have been also decisive for me to take my first steps in this field.

It is my pleasant duty to name other people who have taken part in the work developed during the last three years and a half. Let me express my gratitude to Ana Colubi, for her helpful supervision during the first year of this project and the afterwards collaborations, to Carlos Carleos, for his generous spending of time designing the computerized format of the questionnaire conducted on students and included in this memory, to Norberto, for his inspiring idea of adapting the TIMSS/PIRLS questionnaries to the fuzzy rating scale, and to Asun, for her meticulous reading of this work. Of course, the fuzzy rating survey on the fourth grade students would have been impossible to realize without the solicitous attention of the Colegio San Ignacio in Oviedo.

During all my doctoral training I have been a granted student, thanks to the Ministerio de Educación, Cultura y Deporte, in the Departamento de Estadística e I.O. y D.M. of the Universidad de Oviedo. I am very grateful to the kind daily treatment that I have always received from all its members, starting from Manolo, the actual head of the department, who has always offered his hand to help me, showing a great confidence on me. The last two years I have also been part of the Vakgroep Toegepaste Wiskunde, Informatica en Statistiek from the Universiteit Gent by means of a Joint PhD program. Let me express my gratitude to the head of the department, Prof. Dr. Willy Govaerts, in representation of all its members, for the warm welcome every time I enjoyed a short stay in the Universiteit Gent. My thankful feeling of being part of these groups is extremely related to the friendly atmosphere that surrounds me in the SMIRE research group, with whom I have experimented, apart from their appreciated scientific comments, the luck of sharing a lot of unforgettable moments in different conferences, and in the group we founded in Gent during the lunch breaks and that leaded to some lasting friendships. All of them have made my work easier, or, at least, tolerable. Let me name explicitly Ana, María Brígida, Ángela, Marta, Sara, Luis, Tim, Johan, Bashir, Holger, Sjouke, Peter, Gustavo and Xianming. My colleague of office, Davide, and Itzíar and Adri have of course contributed to this nice environment too along these years.

Finally, but no less important, my complete gratitude to all my family, and in special to my parents and my brother, for their constant support during all my life, the values of effort and hard work, among others, that they have instilled in me since I was a child and their patient attention even when they did not understand the mathematical topics I was talking about. My appreciation is also extended to Lore, Andre, Ali, Sara, Julia and Diego for their blind confidence on me and all their affection, thanks for being always there for me. I am especially grateful to Diego for his infinite patience, understanding and help, also to his computer, these busy last months of work.

This work has been financially supported by/benefited from the grant Ayuda del Programa de FPU AP2009-1197 from the Spanish Ministerio de Educación, Cultura y Deporte, one of its Ayudas para las Estancias Breves, the projects SV-PA-13-ECOEMP-66 from the Principado de Asturias and MTM2009-09440-C02-01 and MTM2006-07501 from the Ministerio de Ciencia e Innovación, an Ayuda a la Investigación de la Fundación Banco Herrero (from April 2011 to April 2012) and three grants for short stays from the COST Action IC0702.

Oviedo-Gent, 22nd April, 2014

Prologue

"Paradoxically, one of the principal contributions of fuzzy logic is its high power of precisiation of what is imprecise."
(Zadeh, 2008)

In the current information age, new types of data are emerging in the everyday life and we should face the new challenging problems associated with the analysis of these data. In consequence, to tackle these problems, new statistical ideas and developments are needed.

Among these new types of data, one can think about the (crisp or fuzzy) set-valued ones, which have often been referred to as imprecise data for the last years. These data come from two different situations in real-life: the imprecise observation/perception of the values of a random variable/vector and the expression of the values of intrinsically imprecise attributes.

The work developed for this dissertation deals with imprecise data modeled as either nonempty compact convex sets or fuzzy sets of finite-dimensional Euclidean spaces. The interest is focused on the robust analysis of the location or central tendency of the random mechanism generating such imprecise data. More concretely, imprecise data are supposed to come from the performance of an imprecise-valued random element (that is, a random set or random fuzzy set) and the statistical analysis concerns the location of this random element.

A well-known location measure of the distribution of a random (fuzzy) set is the associated Aumann-type mean, which satisfies that:

- it is well-defined under quite general conditions,
- it extends the mean value/vector of a random variable/vector, respectively,
- and it preserves the main valuable properties of the real/vectorial-valued means.

Nevertheless, the Aumann-type mean also inherits a rather negative feature from the mean of a random variable/vector: its high sensitivity to either the existence of outliers or data changes.

In looking for a more robust location measure, it seems convenient to follow some of the most successful approaches in dealing with other types of data. In this respect, the work is to be centered on the trimmed means and M-estimators approaches.

To develop or adapt these approaches when the available data are imprecise, two methodologies will be considered, namely:

- On one hand, imprecise data can be represented as functional data from (a convex cone within) certain Hilbert space-valued random elements. Consequently, we may particularize results and methods from Functional Data Analysis which have already been stated, like those regarding the trimmed means and M-estimators to the imprecise-valued case. This can be properly made whenever one guarantees that the particularization does not move out of the cone of the imprecise data.
- On the other hand, when either the preceding methodology fails or the design techniques are exact or give a better approximation, we may develop *ad hoc* concepts and methods by combining notions and results from both (Fuzzy) Set-Valued Analysis and Large Sample/Resampling Statistics.

With this general goal, the work for this dissertation is structured as follows:

Chapter 1 gathers all the main preliminary and supporting tools that will be used in developing the two mentioned approaches. Some concepts and results had already been introduced in the literature in connection with problems different from the ones we have focused on. Nevertheless, a substantial part of the chapter regards notions and results which have been expressly introduced during the course of this thesis.

Firstly, the types of data to be dealt with are presented by recalling their definition and also by showing some useful representations that characterize them. The usual arithmetics between imprecise data are later presented in terms of both their definitions and their alternative representations. Remarks in connection with some distinctive structural characteristics will be stated. Metrics have already been a valuable tool for statistics with precise data, but the distinctive structural characteristics we have just mentioned make the role of metrics even more outstanding and crucial for statistics with imprecise data, especially in the location approaches that will be considered in this work. For this purpose, some suitable distances between imprecise data will be recalled and new metrics, based on the characterizing representations of these data, will be also introduced. The random mechanisms that model the generation of imprecise data will be also presented within the chapter, along with the Aumann/Aumann-type means of these models and some of its main properties. To motivate the main body of the work, contained in Chapters 2, 3 and 4, some simulation studies are carried out to corroborate empirically that the Aumann and Aumann-type means are highly sensitive to outliers and data changes in the imprecise-valued case.

Therefore, one is moved to look for robust location measures for these random elements. Nowadays, in dealing with real-valued data the most widely used robust estimators of central tendency are trimmed means and M-estimators of location, both including the median as a special case. Depending on the distribution of the considered random variable, M-estimators outperform trimmed means in terms of efficiency and robustness or *vice versa*. For this reason, the work has been centered on studying the extension of these two notions to deal with imprecise-valued data.

Chapter 2 is devoted to the extension of trimmed means to deal with imprecise-valued data. At this point, we could particularize to some extent some published concepts and results from Functional Data Analysis. More concretely, the population trimmed mean has been previously introduced in the literature and its existence and uniqueness (under ideal assumptions) have been also proved. On the other hand, the empirical trimmed mean has been approximated by choosing algorithmically a convenient point from the sample and the consistency of this approximation has been demonstrated. In this chapter, after remarking that the required ideal assumptions to guarantee the uniqueness of the population trimmed means do not cover some realistic situations, a new algorithm to compute the empirical trimmed mean and not only an approximation is introduced. The new algorithm is examined and some results are obtained. The empirical trimmed mean is proved to be consistent and its finite sample breakdown point is also computed as an objective tool measuring its robustness.

A theoretical comparative analysis of the complexity of the new and the previous algorithms for the trimmed mean is carried out. Furthermore, a comparative study of the efficiency of the empirical trimmed mean computed using the new algorithm in contrast to the mean, the trimmed mean based on depths and the approximation of the empirical trimmed mean is developed through simulations of functional data.

The ideas and results in this chapter are finally particularized to set- and fuzzy set-valued data by means of two illustrative real-life examples and some comparative simulations in connection with the Aumann/Aumann-type means and analogue to the ones presented for functional data.

Chapter 3 deals with the extension of M-estimates of location for imprecise-valued data. As for the trimmed means, the first attempt to approach the problem has led to the search of M-estimates in the realm of Functional Data Analysis. Some recent studies in the context of robust nonparametric density estimation combine ideas from both the traditional kernel density estimation and the classical M-estimation. These studies have been first adapted to Hilbert space-valued random elements in such a way that necessary and sufficient conditions for the existence of M-estimates in this setting have been stated. Moreover, these conditions allow us to ensure that each of these estimates can be expressed as a convex linear combination or weighted mean of the sample elements the estimate is based on. This special feature is crucial for the particularization to imprecise-valued data, in order to guarantee that the M-estimates remain in the space of values the location measure should belong to. Furthermore, an iterative algorithm, extension of the iteratively re-weighted least squares algorithm used in classical M-estimation, is also adapted to approximate the sample location M-estimate for Hilbert space-valued data under the above-referred conditions. Finally, the consistency is proved and the finite sample breakdown point of the M-estimates is calculated, both results under some other assumptions.

Although the conditions for the existence of M-estimates cover the extension of several interesting and valuable classical M-estimates, other valuable and natural ones do not fulfill them. At this point, the second attempt to approach the problem has consisted of developing *ad hoc* procedures for imprecise-valued random elements related to the 1-dimensional Euclidean space. Two of these procedures are based on convenient L^1 metrics between imprecise values (which have been either introduced or recalled in Chapter 1) and the other one makes use of an L^2 metric, being inspired by the spatial approach to the median.

The ideas and results in this chapter are finally illustrated by means of two real-life examples and some comparative simulations in connection with the Aumann-type mean are carried out.

Chapter 4 aims to compare the robust behaviour of the different location measures introduced in this work from an empirical point of view. While the aim of the simulations developed in Chapters 2 and 3 was to compare the robustness of the new location measures in contrast to the sensitivity of the Aumann/Aumann-type means, at this chapter the comparison will be stated among all the approaches which have been suggested in the course of this work. A summarized discussion is to be presented after them.

Each of the chapters in the work will end with some common types of remarks, namely,

- those clearly highlighting the main contributions in the chapter;
- those pointing out the ideas and results which, having been developed for the chapter, have already been published, accepted or submitted for publication.

The work for this dissertation ends with some comments and suggestions concerning open problems and future directions in connection with the topic covered in it.

Prólogo

"Paradoxically, one of the principal contributions of fuzzy logic is its high power of precisiation of what is imprecise."
(Zadeh, 2008)

En la era de la información en la que nos encontramos inmersos, continuamente surgen nuevos tipos de datos procedentes de la vida cotidiana. Los problemas asociados al análisis de dichos datos suponen un desafío al que hay que enfrentarse y abordar a través de desarrollos estadísticos novedosos.

Entre los tipos de datos que están emergiendo, se encuentran aquellos que toman valores de conjunto (clásico o fuzzy) y que en los últimos años suelen denominarse datos imprecisos. En la vida real, estos datos pueden provenir de dos situaciones completamente diferentes: por un lado, de la observación o percepción imprecisa de los valores de una variable o de un vector aleatorios y, por otro, de la expresión de los valores de atributos que son intrínsecamente imprecisos.

El trabajo desarrollado para esta tesis se ocupa de los datos imprecisos que se modelizan, o bien mediante conjuntos no vacíos, compactos y convexos, o bien mediante conjuntos fuzzy de espacios euclídeos de dimensión finita. El interés está focalizado en el análisis robusto de la localización o tendencia central del mecanismo aleatorio que genera tales datos. Más concretamente, se supone que los datos que manejamos son resultado de la realización de un elemento aleatorio con valores imprecisos (es decir, de un conjunto aleatorio o de un conjunto fuzzy aleatorio) y el análisis estadístico considerado hace referencia a la tendencia central de la distribución de ese elemento aleatorio.

Una medida de localización muy conocida y habitual en el contexto de los conjuntos aleatorios (clásicos o fuzzy) es la media (de Aumann o tipo Aumann, respectivamente) asociada a los mismos, la cual satisface que:

- está bien definida en condiciones bastante generales,
- extiende el valor/vector media de las variables/vectores aleatorios,

- y conserva las propiedades más relevantes de las medias del caso real y vectorial.

No obstante, esta generalización de la media comparte con los casos real y vectorial un rasgo muy negativo: su elevada sensibilidad a la presencia de outliers o al cambio de datos.

Es deseable, por lo tanto, la búsqueda de una medida de localización con un comportamiento más robusto. Para ello, parece conveniente adoptar algunos de los enfoques más exitosos en el análisis de otros tipos de datos. Ese es el motivo por el que este trabajo va a centrarse en las medias recortadas y en los M-estimadores.

Para desarrollar o adaptar estos enfoques cuando los datos disponibles son imprecisos, van a considerarse dos metodologías:

- Por un lado, los datos imprecisos pueden representarse como datos funcionales provenientes de elementos aleatorios con valores en (un cono convexo dentro de) cierto espacio de Hilbert. Por lo tanto, podrían particularizarse los resultados y métodos ya establecidos para el Análisis de Datos Funcionales, como los relativos a las medias recortadas y a los M-estimadores, al caso de datos imprecisos. Esta aproximación será válida siempre y cuando pueda garantizarse que la particularización se lleva a cabo dentro del cono correspondiente a los datos imprecisos.
- Por otro lado, cuando la metodología anterior falla o es posible establecer técnicas exactas o que proporcionen una aproximación más adecuada, pueden desarrollarse conceptos y métodos *ad hoc* combinando nociones y resultados del Análisis de Valores de Conjunto (o Conjunto Fuzzy) y las Estadísticas de Grandes Muestras y de Remuestreo.

Con este objetivo general, el trabajo de la tesis se ha estructurado como sigue:

El **Capítulo 1** recoge las herramientas preliminares y de apoyo que van a usarse en el desarrollo de los dos enfoques mencionados. Algunos conceptos y propiedades se han introducido previamente en la literatura en relación con otros problemas. Sin embargo, una parte importante del capítulo se refiere a nociones y resultados que se han introducido de forma expresa en el curso de esta tesis.

En primer término se presentan los tipos de datos que van a tratarse, recordando sus definiciones y mostrando algunas representaciones útiles que los caracterizan. A continuación, se exponen las aritméticas usuales entre datos imprecisos (tanto en términos de sus definiciones originales como de las alternativas en función de sus representaciones caracterizadoras), señalando algunas características estructurales distintivas inherentes a su empleo. Hay que tener en cuenta que, si bien las métricas entre datos son una herramienta muy notable y valiosa en la Estadística clásica, las peculiaridades estructurales a las que acabamos de referirnos hacen que su papel sea aún más destacado si cabe en el caso de datos imprecisos, especialmente en los enfoques del estudio de la localización que van a considerarse a lo largo de esta tesis. Con esta finalidad, se recopilan algunas de las distancias más adecuadas entre datos imprecisos y se introducen nuevas métricas basadas en representaciones caracterizadoras de dichos datos. También se incluyen en este capítulo los mecanismos aleatorios que modelizan la generación de datos imprecisos, junto con las medias de Aumann y tipo Aumann de esos modelos, así como algunas de sus propiedades más notables. Para motivar la contribución principal de esta memoria, recogida en los Capítulos 2, 3 y 4, se llevan a cabo algunos estudios de simulación que corroboran empíricamente que las medias de Aumann y tipo Aumann son muy sensibles ante outliers y cambios de datos.

Todo ello justifica la propuesta de medidas de localización robustas para estos elementos aleatorios. En la actualidad, las medias recortadas y los M-estimadores de localización son los estimadores robustos de tendencia central más utilizados al trabajar con valores reales, siendo la mediana un caso especial de los anteriores. Dependiendo de las distribuciones de las variables consideradas, los M-estimadores resultan más eficientes y robustos que las medias recortadas o viceversa. Por esta razón, la tesis se ha centrado en el estudio de la generalización de ambas nociones para el tratamiento de datos imprecisos.

El **Capítulo 2** está dedicado a la extensión de las medias recortadas para datos con valores imprecisos. En este punto, pueden particularizarse algunos conceptos y resultados del Análisis de Datos Funcionales. Más concretamente, las medias recortadas poblacionales ya han sido introducidas previamente por otros autores y se ha demostrado su existencia y unicidad bajo ciertas condiciones ideales. Por otra parte, la media recortada empírica se ha aproximado eligiendo algorítmicamente un punto apropiado a partir de la muestra y su consistencia ha sido demostrada. En este capítulo, tras observar que las suposiciones requeridas para garantizar la unicidad de las medias recortadas poblacionales no se ajustan a las situaciones más realistas, se introduce un nuevo algoritmo para calcular la media recortada empírica de forma exacta, y no simplemente una aproximación de la misma. Después de examinar el algoritmo propuesto y obtener algunos resultados teóricos, se prueba que la media recortada empírica es consistente y se determina su punto de ruptura muestral finito como una herramienta objetiva de su robustez.

La complejidad de ambos algoritmos se va a comparar a nivel teórico y también a través de un estudio de simulación con datos funcionales, contrastando la eficiencia de la media recortada empírica calculada frente a su aproximación, la media y la media recortada basada en profundidades.

Finalmente, se particularizan las ideas y resultados de este capítulo a datos con valores de conjunto y de conjunto fuzzy mediante dos ejemplos ilustrativos del mundo real y algunas simulaciones comparativas en relación con las medias de Aumann y tipo Aumann siguiendo pautas análogas a las de las comparaciones para datos funcionales.

El **Capítulo 3** se ocupa de la extensión de las M-estimaciones de localización para datos con valores imprecisos. Al igual que ocurría con las medias recortadas, el primer acercamiento al problema ha conducido a la búsqueda de M-estimaciones en el marco del Análisis de Datos Funcionales. Algunos estudios recientes, en el contexto de la estimación de densidad no paramétrica robusta, combinan ideas de la estimación de densidad núcleo tradicional y la M-estimación clásica. Estos estudios se han adaptado en primer lugar a elementos aleatorios con valores en un espacio de Hilbert, estableciendo condiciones necesarias y suficientes para la existencia de M-estimaciones en esa situación. Gracias a esas condiciones es posible asegurar que las estimaciones pueden expresarse como una combinación lineal convexa o media ponderada de los elementos muestrales a partir de los cuales se calcula la estimación. Esta propiedad resulta crucial para la particularización al caso de datos imprecisos, puesto que permite garantizar que las M-estimaciones no toman valores fuera del espacio al que debe pertenecer la medida de localización. Además, para aproximar la M-estimación de localización muestral para datos con valores en espacios de Hilbert en las condiciones indicadas, se ha adaptado un procedimiento iterativo que extiende el algoritmo de mínimos cuadrados reponderado iterativamente y empleado habitualmente en la M-estimación clásica. Por último, y bajo otras condiciones, se prueba la consistencia y se calcula el punto de ruptura muestral finito de las M-estimaciones.

Aunque las condiciones para la existencia de las M-estimaciones son válidas para muchas de las funciones de pérdida clásicas más interesantes, otras relevantes y naturales no las satisfacen. En este punto, la segunda forma de acercarse al problema ha consistido en desarrollar procedimientos *ad hoc* para elementos aleatorios con valores imprecisos, limitándonos al caso unidimensional (números fuzzy aleatorios e intervalos aleatorios). Dos de esos procedimientos están basados en apropiadas métricas (entre valores imprecisos) de tipo L^1, las introducidas o rememoradas en el Capítulo 1, y el tercero recurre a una métrica de tipo L^2, inspirándose en el enfoque de la mediana espacial.

Para concluir, las ideas y resultados de este capítulo se ilustran a través de dos ejemplos de la vida real y algunas simulaciones comparativas en conexión con la media de Aumann o tipo Aumann.

El **Capítulo 4** compara el comportamiento robusto de las distintas medidas/estimaciones de localización introducidas en la tesis, desde una perspectiva empírica. Mientras que el objetivo de las simulaciones recogidas en los Capítulos 2 y 3 era contrastar la robustez de las nuevas medidas frente a la sensibilidad de las medias de Aumann o tipo Aumann, en este capítulo la comparación concierne a todas las propuestas presentadas a lo largo de esta memoria. Tras las simulaciones, se presenta una discusión que resume las conclusiones extraídas.

Cada uno de los capítulos de la tesis finaliza con una serie de observaciones comunes:

- las que subrayan las contribuciones más destacables del capítulo;
- las que indican las ideas y resutados que se han desarrollado expresamente para dicho capítulo y ya se han publicado, aceptado para publicación o enviado para su consideración en revista.

La memoria concluye con algunos comentarios y sugerencias de problemas abiertos y líneas futuras de investigación estrechamente conectadas con el tema de la tesis.

Proloog

"Paradoxically, one of the principal contributions of fuzzy logic is its high power of precisiation of what is imprecise."

(Zadeh, 2008)

In het huidige informatie tijdperk ontstaan nieuwe types gegevens in het dagelijkse leven en we moeten de uitdagende problemen die gepaard gaan met de analyse van deze nieuwe soorten gegevens aanpakken. Om deze problemen op te lossen zijn nieuwe statistische ideeën en ontwikkelingen noodzakelijk.

Eén van de nieuwe types data waar men kan aan denken zijn de (crisp of fuzzy) verzamelingwaardige gegevens die in de laatste jaren dikwijls niet-precieze gegevens genoemd worden. Deze gegevens zijn afkomstig van twee verschillende situaties in de praktijk: De niet-precieze observatie/perceptie van de waarde van een stochastische veranderlijke/vector en de expressie van de waarde van intrinsiek niet-precieze attributen.

Het werk dat in deze thesis ontwikkeld werd behandelt niet-precieze gegevens die gemodelleerd worden als ofwel niet-lege compacte convexe verzamelingen ofwel als vaagverzamelingen van eindige Euclidische ruimten. De hoofdinteresse ligt bij de robuuste analyse van de locatie of centrale tendens van het toevalsafhankelijke mechanisme that zulke niet-precieze gegevens genereert. Meer concreet, niet-precieze gegevens worden verondersteld afkomstig te zijn van een niet-precieswaardig stochastisch element (d.w.z. een stochastische verzameling of een stochastische vaagverzameling) en de statistische analyse heeft betrekking op de locatie van dit stochastisch element.

Een welgekende maat voor de locatie van de verdeling van een stochastische (vaag)verzameling is het Aumann-type gemiddelde, die voldoet aan:

- Het is goed-gedefinieerd onder vrij algemene voorwaarden,
- Het is een uitbreiding van de gemiddelde waarde/vector van een stochastische veranderlijke/vector, respectievelijk,
- en het behoudt de belangrijkste waardevolle eigenschappen van de reëel/vector-waardige gemiddelden.

Desondanks erft het Aumann-type gemiddelde ook een negatieve eigenschap van het gemiddelde van een stochastische veranderlijke/vector: zijn extreem hoge gevoeligheid aan de aanwezigheid van uitschieters of (kleine) veranderingen in de gegevens.

In de zoektocht naar meer robuuste locatiematen, is het voor de hand liggend om de meest succesvolle procedures bij andere types gegevens te volgen. Dit werk concentreert zich dan ook op het getrimde gemiddelde en de M-schatter aanpak.

Om deze technieken te ontwikkelen of aan te passen voor niet-precieze gegevens, worden twee methoden beschouwd, namelijk

- Enerzijds kunnen niet-precieze gegevens voorgesteld worden als functionele gegevens afkomstig van (een convexe kegel in) zekere Hilbertruimte-waardige stochastische elementen. Bijgevolg kunnen we reeds bestaande resultaten en methoden van functionele data analyse, zoals deze voor getrimde gemiddelden en M-schatters, toepassen op het speciale geval van niet-precieswaardige gegevens. Deze aanpak kan volledig uitgewerkt worden als gegarandeerd kan worden dat de toepassing van deze procedures op niet-precieswaardige gegevens niet buiten de kegel van de niet-precieze gegevens kan terechtkomen.
- Anderzijds, als de voorgaande methodologie faalt of als alternatieve design-technieken exacte resultaten of betere benaderingen opleveren, dan kunnen we *ad hoc* concepten en methoden ontwikkelen door noties en resultaten van (vaag)verzamelingwaardige analyse te combineren met grote steekproeven / resampling statistiek.

Met dit algemene doel is het werk voor deze verhandeling gestructureerd als volgt:

Hoofdstuk 1 verzamelt alle belangrijke voorafgaande en ondersteunende instrumenten die gebruikt zullen worden in de ontwikkeling van de twee hierboven vermelde aanpakken. Enkele concepten en resultaten werden eerder al in de literatuur geïntroduceerd in verband met andere problemen dan degenen die hier bestudeerd worden. Een substantieel deel van het hoofdstuk behandelt echter begrippen en resultaten die nieuw geïntroduceerd werden, specifiek voor het onderzoek in deze thesis.

Eerst worden de types gegevens die behandeld worden voorgesteld door hun definitie te herhalen en ook door nuttige representaties die de gegevens karakteriseren aan te tonen. De standaard rekenregels voor niet-precieze gegevens worden dan voorgesteld, zowel in termen van hun definitie als in termen van hun alternatieve representaties. Commentaren in verband met verschillende structurele karakterisaties worden ook toegevoegd. Metrieken zijn al een heel nuttig instrument gebleken voor de statistische analyse van niet-precieze gegevens, maar de nieuwe representaties zorgen ervoor dat de rol van metrieken nog crucialer wordt voor de statistische analyse van niet-precieze gegevens, in het bijzonder voor de locatieprocedures die in dit werk uiteengezet worden. Met dit doel worden handige afstanden tussen niet-precieze gegevens herhaald, maar ook worden nieuwe metrieken gebaseerd op de nieuwe representaties voor deze gegevens geïntroduceerd. De stochastische mechanismen die het genereren van niet-precieze gegevens modelleren worden ook in dit hoofdstuk voorgesteld, samen met de Aumann-type gemiddelden voor deze modellen en hun belangrijkste eigenschappen. Om het werk in hoofdstukken 2, 3 en 4, die de hoofdzaak van deze thesis vormen, te motiveren worden simulatieresultaten getoond die empirisch bevestigen dat de Aumann-type gemiddelden extreem gevoelig zijn aan uitschieters en veranderingen in de niet-precieswaardige gegevens.

Daarom wordt men aangezet om op zoek te gaan naar robuuste locatiematen voor deze stochastische elementen. De huidige meest gebruikte robuuste schatters voor het centrum van reëelwaardige gegevens zijn de getrimde gemiddelden en M-schatters, die beide de mediaan bevatten als speciaal geval. Afhankelijk van de verdeling van de beschouwde stochastische veranderlijke kunnen M-schatters de getrimde gemiddelden overtreffen in termen van efficiëntie en robuustheid of *vice versa*. Voor deze reden richt dit werk zich op het uitbreiden van deze twee noties voor de behandeling van niet-precieswaardige gegevens.

Hoofdstuk 2 is gewijd aan de uitbreiding van getrimde gemiddelden voor de behandeling van niet-precieswaardige gegevens. Hiervoor konden in zekere mate gepubliceerde begrippen en resultaten van functionele data analyse gebruikt worden. Meer concreet, het populatie getrimde gemiddelde werd al in de literatuur geïntroduceerd en zijn bestaan en uniciteit (onder ideale voorwaarden) werden reeds bewezen. Anderzijds werd het steekproef getrimde gemiddelde benaderd door op een algoritmische wijze een geschikte observatie van de steekproef te kiezen en de consistentie van deze schatter werd aangetoond. In dit hoofdstuk wordt eerst opgemerkt dat de ideale voorwaarden om uniciteit van het populatie getrimde gemiddelde te garanderen niet voldaan zijn in een aantal realistisch situaties en wordt daarna een nieuw algoritme voor de berekening van het steekproef getrimde gemiddelde, en niet slechts een benadering, voorgesteld. De eigenschappen van het nieuwe algoritme worden onderzocht. Er wordt aangetoond dat het steekproef getrimde gemiddelde consistent is en het finite-sample breekpunt van de schatter wordt ook berekend als een objectieve manier om zijn robuustheid te meten.

Een theoretische vergelijking van de complexiteit van het nieuwe en bestaande algoritme voor het steekproef getrimde gemiddelde wordt uitgevoerd. Daarenboven wordt een vergelijkende studie opgezet om de efficiëntie van het steekproef getrimde gemiddelde, berekend aan de hand van het nieuwe algoritme, te vergelijken met het gemiddelde, het getrimde gemiddelde gebaseerd op dieptefuncties en de eerdere benadering van het steekproef getrimde gemiddelde aan de hand van simulaties met functionele gegevens.

De ideeën en resultaten in dit hoofdstuk worden tenslotte toegepast op het specifieke geval van verzameling- en vaagverzamelingwaardige gegevens door middel van twee illustraties met reële data voorbeelden en enkele vergelijkende simulaties ten opzichte van de Aumann-type gemiddelden, analoog aan de simulaties met functionele gegevens.

Hoofdstuk 3 behandelt de uitbreiding van M-schatters voor locatie naar niet-precieswaardige gegevens. Zoals bij getrimde gemiddelden was de eerste poging om het probleem aan te pakken om op zoek te gaan naar M-schatters in de context van functionele data analyse, Enkele recente studies in de context van robuuste niet-parametrische dichtheidsschatting combineren ideeën van traditionele kernel dichtheidsschatting met M-schatters. Deze studies werden eerst aangepast naar Hilbertruimte-waardige stochastische elementen en wel zodanig dat voldoende en nodige voorwaarden voor het bestaan van M-schatters in deze setting opgesteld werden. Deze voorwaarden stellen ons daarenboven in staat om te garanderen dat deze schatters kunnen uitgedrukt worden als een convexe lineaire combinatie of een gewogen gemiddelde van de steekproefelementen waarop de schatting gebaseerd is. Deze bijzondere eigenschap is cruciaal voor de toepassing op het speciale geval van niet-precieswaardige gegevens om te kunnen garanderen dat de M-schatter in de ruimte blijft waartoe de locatieschatter moet behoren. Door het iteratieve herwogen kleinste kwadratenalgoritme voor klassieke M-schatters aan te passen, wordt bovendien een iteratief algoritme uitgewerkt om de steekproef locatie M-schatting voor Hilbertruimte-waardige gegevens te benaderen onder de reeds aangehaalde voorwaarden. Ten slotte wordt de consistentie aangetoond en ook het finite-sample breekpunt van de M-schatters berekend onder geschikte voorwaarden.

Alhoewel de voorwaarden voor het bestaan van M-schatters de uitbreiding van meerdere interessante en belangrijke M-schatters omvatten, sluiten de voorwaarden ook andere natuurlijke, belangrijke M-schatters uit. Daarom heeft de tweede aanpak voor het probleem erin bestaan om *ad hoc* procedures te ontwikkelen voor niet-precieswaardige stochastische elementen gerelateerd aan de ééndimensionele Euclidische ruimte. Twee van deze procedures zijn gebaseerd op geschikte L^1 metrieken tussen niet-precieze gegevens (die herhaald of geïntroduceerd werden in Hoofdstuk 1) en de andere procedure gebruikt een L^2 metriek, geïnspireerd door de spatiale aanpak voor de mediaan.

De ideeën en resultaten in dit hoofdstuk worden tenslotte geïllustreerd door middel van twee voorbeelden met reële data en enkele vergelijkende simulaties ten opzichte van het Aumann-type gemiddelde worden ook uitgevoerd.

Hoofdstuk 4 beoogt om de robuustheid van de verschillende locatiematen die in dit werk geïntroduceerd werden te vergelijken vanuit een empirisch standpunt. Daar waar de simulaties in de hoofdstukken 2 en 3 bedoeld waren om de robuustheid van de nieuwe locatiematen te vergelijken ten opzichte van de gevoeligheid van de Aumann-type gemiddelden, willen we in dit hoofdstuk een vergelijking maken tussen alle procedures die tijdens dit werk voorgesteld werden. Een samenvattend overzicht wordt op het einde weergegeven.

Elk van de hoofdstukken in dit werk eindigt met eenzelfde soort van opmerkingen, namelijk

- opmerkingen die duidelijk de belangrijkste bijdragen van het hoofdstuk in de verf zetten.

- opmerkingen die duidelijk aangeven welke ideeën en resultaten die in het hoofdstuk uitgewerkt werden al gepubliceerd werden, aanvaard werden voor publicatie of ter publicatie aangeboden werden.

Het werk voor deze verhandeling eindigt met enkele commentaren en suggesties met betrekking tot open problemen en toekomstige richtingen voor onderzoek in verband met de onderwerpen die hier aan bod gekomen zijn.

Contents

List of Figures

List of Tables

Chapter 1

Preliminary tools and supporting results

Nowadays, data is the raw material and the essence of information and knowledge. Statistical data analysis methodology is permanently evolving due to either the emergence of new types of data in real-life studies or data analysts being able to develop tools to deal with more or less complex data. In this respect, the first statistical techniques were simply designed to manage either quantitative or qualitative data. At present, we can find numerous statistical procedures to handle functional data (see, for instance, some recent studies about by Febrero-Bande and González-Manteiga [68], Arribas-Gil and Müller [4], Jacques and Preda [110]), incomplete/missing data (see, for instance, Bianco *et al.* [17], Ferraty *et al.* [74], Zhao *et al.* [223], Lin [123], Sinha *et al.* [170]), and several other types of data.

Among the new types of data that can be frequently found in the everyday life, one can consider imprecise data. As a quite natural, inherent and almost inevitable fact, the meaning of "imprecise" is itself rather imprecise and there is no unanimous agreement on it.

One of the most commonly used meaning for the imprecise data has been that corresponding to mixtures of interval-valued and ordinal data (see, for instance, Cooper *et al.* [36]).

Another meaning, the one considered along this work and that is prevailing in the last years (see, for instance, Petit-Renaud and Denoeux [149], Coppi *et al.* [41], Gil *et al.* [85], Hsu *et al.* [105], Zerafat *et al.* [222], Blanco-Fernández *et al.* [18], Ferraro and Giordani [73]), is that imprecise data stand for set or fuzzy set-valued data. These data are assumed to derive from either the imprecise observation/description of a real/vectorial data or the valuation of an intrinsically imprecise-valued data. It should be clearly stated that the statistics in this work will be supposed to refer to either these data or the random mechanism generating them, independently of the origin of the imprecise data.

In this chapter, several models, concepts and results regarding imprecise data will be exposed. Section 1.1 concerns the mathematical modeling for imprecise data that is to be adopted in all the studies for this dissertation. Section 1.2 refers to the arithmetic between these data. Section 1.3 presents some useful distances between imprecise data which will be employed in Chapters 2 and 3. Section 1.4 recalls suitable models for the imprecise-valued random elements generating imprecise data. Section 1.5 includes empirical results showing how the Aumann-type mean of such random elements is influenced by data contamination, what motivates the interest of the studies in subsequent chapters. The chapter ends with a summary of the novelties and the related publications derived from it.

1.1 Types of imprecise data and some valuable representations

Imprecise data, in accordance with the considered meaning, model many complex objects or appear in different domains.

In this way, set-valued data (and, in particular, interval-valued data) are a type of data that often arise in econometric and financial applications. For example, as indicated by Molchanov and others (see, for instance, Haval and Molchanov [111], Cascos and Molchanov [26], Molchanov [139]): earners may report a salary bracket instead of the exact salary or the profit of a firm may be intentionally converted to an interval to ensure anonymity, it is possible to be interested in the range of prices (which are always non-unique in case of transaction costs) and in case of several financial assets this range leads to a parallelepiped (or to a more general convex set if simultaneous transactions on several assets attract an extra discount). Also, set-valued data are present in many medical and image databases recording, for instance: the set of alleles present at a particular genomic location for each person in a group; a particle; the set of symptoms/illnesses or the set of objects appearing in a picture; and so on (see, for instance, Zhang *et al.* [224], Díaz *et al.* [58], and others).

Similarly, the kind of data called fuzzy set-valued data (and, in particular, fuzzy number-valued data) describe ratings, opinions, judgements, perceptions and other data often in connection with human valuations in a natural and very expressive way. More concretely, when conducting quality ratings, satisfaction valuations and many other surveys, responses cannot usually be expected to be expressible in terms of values in a precise scale, since they are essentially imprecise (see De la Rosa de Sáa *et al.* [53] for a recent detailed discussion about this point). Examples of fuzzy set-valued data with higher dimension are scarcely found in the literature of applications, but some references appear to conical fuzzy data in an approach to the tone and color triangle designs (see Sugano [192] and Sugano *et al.* [193]), in analyzing the perforation of a plate by projectiles (see Celmiņš [28]) or in clusterizing fuzzy ecological data (see Salski [165]), among others.

The formalization of set- and fuzzy set-valued data is now presented in the next two subsections.

1.1.1 Set-valued data and helpful representations

Let $(\mathbb{R}^p, \|\cdot\|)$ be the p-dimensional Euclidean space with the associated norm, where $p \in \mathbb{N}$.

Definition 1.1.1. $\mathcal{K}_c(\mathbb{R}^p)$ *is the space of nonempty compact convex subsets of* $\mathbb{R}^p$*. Consequently,* $\mathcal{K}_c(\mathbb{R})$ *denotes the space of closed and bounded nonempty intervals.*

Along this work, when we refer to ***set-valued data***, we will be concerned with elements in $\mathcal{K}_c(\mathbb{R}^p)$ with $p > 1$, whereas when we refer to ***interval-valued data***, we will be concerned with elements in $\mathcal{K}_c(\mathbb{R})$.

Convex bodies of $\mathbb{R}^p$, i.e., the elements in $\mathcal{K}_c(\mathbb{R}^p)$, can be represented in terms of the support function introduced by Minkowski [134] (see, for instance, Castaing and Valadier [27], Schneider [167] and Ghosh and Kumar [82] for more details).

Definition 1.1.2. *The* ***support function*** *of* $K \in \mathcal{K}_c(\mathbb{R}^p)$ *is the mapping* $s_K : \mathbb{S}^{p-1} \to \mathbb{R}$ *(where* $\mathbb{S}^{p-1}$ *denotes the unit sphere of* $\mathbb{R}^p$*, that is,* $\mathbb{S}^{p-1} = \{\mathbf{u} \in \mathbb{R}^p : \|\mathbf{u}\| = 1\}$*) given by*

$$s_K(\mathbf{u}) = \sup_{\mathbf{v} \in K} \langle \mathbf{u}, \mathbf{v} \rangle$$

for all $\mathbf{u} \in \mathbb{S}^{p-1}$*, with* $\langle \cdot, \cdot \rangle$ *denoting the inner product on* $\mathbb{R}^p$*.*

The value $s_K(\mathbf{u})$ represents the signed (i.e., oriented) distance from $\mathbf{0} \in \mathbb{R}^p$ to the supporting hyperplane of K which is orthogonal to $\mathbf{u}$. Figure 1.1 shows the graphical interpretation of the support function of a nonempty compact convex set in dimensions $p = 1$ and $p = 2$. Ghosh and Kumar [82] have also graphically illustrated the computation of the support function for different situations in case $p = 2$.

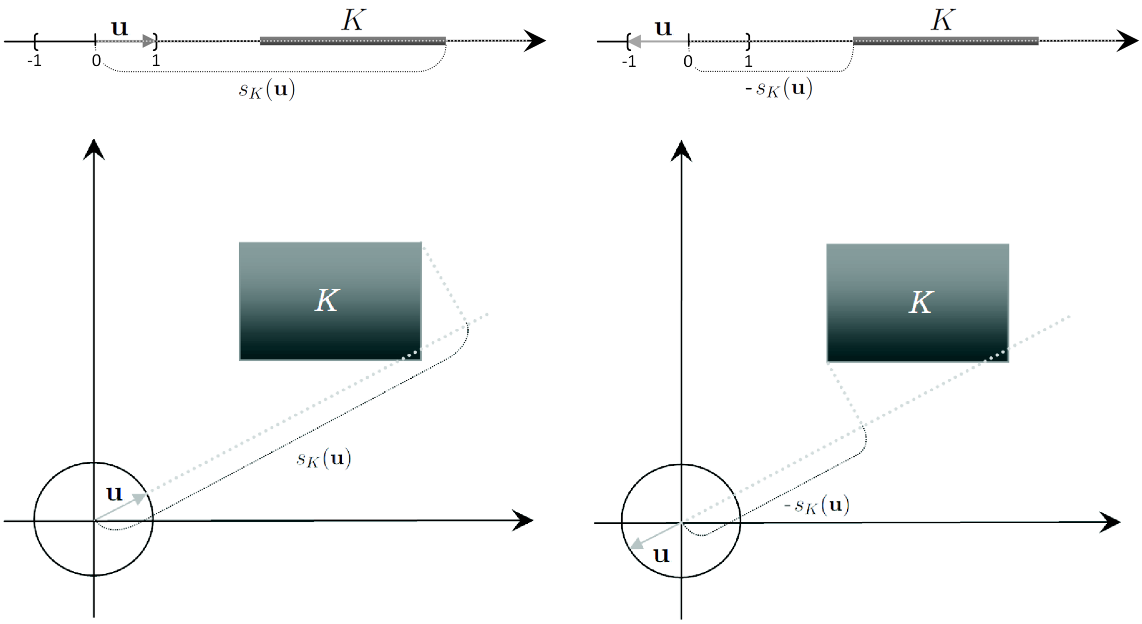


Figure 1.1: Support function of a nonempty compact convex set. On the top, case $p = 1$, $\mathbb{S}^0 = \{-1, 1\}$. On the bottom, case $p = 2$, $\mathbb{S}^1 =$ circumference (center $= (0, 0)$, radius $= 1$)

The support function was originally defined on the space $\mathbb{R}^p$, but because of the support function being positive homogeneous (i.e., $s_K(\gamma \cdot \mathbf{v}) = \gamma \cdot s_K(\mathbf{v})$ for all $\mathbf{v} \in \mathbb{R}^p, \gamma > 0$), then the definition can be constrained to the unit sphere.

It should be emphasized that a set in $\mathcal{K}_c(\mathbb{R}^p)$ is uniquely determined by its support function. Actually, necessary and sufficient conditions are known to ensure such a uniqueness (see, for instance, Schneider [167], Theorem 1.7.1, p. 38):

Proposition 1.1.1. [167] *If $s : \mathbb{S}^{p-1} \to \mathbb{R}$ is a subadditive function, that is,*

$$\|\mathbf{u} + \mathbf{v}\| \cdot s\left(\frac{\mathbf{u} + \mathbf{v}}{\|\mathbf{u} + \mathbf{v}\|}\right) \leq s(\mathbf{u}) + s(\mathbf{v})$$

for all $\mathbf{u}, \mathbf{v} \in \mathbb{S}^{p-1}$ (and, hence, one can trivially extend it to a sublinear –i.e., subadditive and positive homogeneous– function on $\mathbb{R}^p$), then there exists a unique $K \in \mathcal{K}_c(\mathbb{R}^p)$ such that $s = s_K$. More concretely,

$$K = \left\{\mathbf{v} \in \mathbb{R}^p \,:\, \langle \mathbf{v}, \mathbf{u} \rangle \leq s(\mathbf{u}) \text{ for all } \mathbf{u} \in \mathbb{S}^{p-1}\right\}.$$

In the particular case $p = 1$, that is, in dealing with interval-valued data, the two following characterizing representations are intuitive and easy-to-use:

Definition 1.1.3. *The* ***inf/sup representation*** *of the interval* $K \in \mathcal{K}_c(\mathbb{R})$ *is the vector* $\boldsymbol{\iota}_K = (\inf K, \sup K) = (-s_K(-1), s_K(1)) \in \{(x,y) \in \mathbb{R}^2 : x \leq y\}$. *In fact, for any vector* $\mathbf{w} = (w_1, w_2) \in \mathbb{R}^2$ *with* $w_1 \leq w_2$, *there exists a unique interval* $K = [w_1, w_2] \in \mathcal{K}_c(\mathbb{R})$ *such that* $\boldsymbol{\iota}_K = \mathbf{w}$.

Definition 1.1.4. *The* ***mid/spr representation*** *of the interval* $K \in \mathcal{K}_c(\mathbb{R})$ *is the vector* $\boldsymbol{\eta}_K = (\text{mid}\, K, \text{spr}\ K) = ((\inf K + \sup K)/2, (\sup K - \inf K)/2) \in \mathbb{R} \times [0, \infty)$ *(i.e.,* $\text{mid}\, K$ *is the mid-point/centre of* K *and* $\text{spr}\, K$ *denotes the spread/radius of* K*). In fact, for any vector* $\mathbf{w} = (w_1, w_2) \in \mathbb{R} \times [0, \infty)$ *there exists a unique interval* $K = [w_1 - w_2, w_1 + w_2] \in \mathcal{K}_c(\mathbb{R})$ *such that* $\boldsymbol{\eta}_K = \mathbf{w}$.

The use of the support function representation is quite interesting from a theoretical viewpoint, as we will later see, but in many cases it is not easy-to-specify, to-handle and to-interpret. In contrast, both the inf/sup and the mid/spr are always easy-to-compute and to-use, although they are limited to the one-dimensional case.

1.1.2 Fuzzy set-valued data and helpful representations

Fuzzy set-valued data are a 'level-wise' extension of set-valued ones, in which the levels add a certain gradualness to the imprecision of set-valued data.

Definition 1.1.5. $\mathcal{F}_c(\mathbb{R}^p)$ *is the space of the fuzzy subsets of* $\mathbb{R}^p$ *(that is, the mappings* $\widetilde{U} : \mathbb{R}^p \to [0,1]$*), such that their* α*-levels*

$$\widetilde{U}_\alpha = \begin{cases} \{\mathbf{x} \in \mathbb{R}^p : \widetilde{U}(\mathbf{x}) \geq \alpha\} & \text{if } \alpha \in (0,1] \\ \text{cl}\{\mathbf{x} \in \mathbb{R}^p : \widetilde{U}(\mathbf{x}) > 0\} & \text{if } \alpha = 0, \end{cases}$$

belong to $\mathcal{K}_c(\mathbb{R}^p)$, *where* $\widetilde{U}(\mathbf{x})$ *means the 'degree of compatibility of* $\mathbf{x}$ *with* $\widetilde{U}$*' (or 'degree of truth of the assertion «*$\mathbf{x}$ *is* $\widetilde{U}$*»').*

Equivalently, a fuzzy set value can be defined as a normal (i.e., having nonempty 1-level) upper semi-continuous element of $[0,1]^{\mathbb{R}^p}$ with bounded 0-level.

Along this work, when we refer to ***fuzzy set-valued data***, we will be concerned with elements in $\mathcal{F}_c(\mathbb{R}^p)$ with $p \in \mathbb{N}$. We will distinguish two situations: when $p > 1$, we will call them ***fuzzy vectors***, while when $p = 1$, they will be referred to as ***fuzzy number-valued data*** or simply ***fuzzy numbers***.

As for the set-valued case, elements in $\mathcal{F}_c(\mathbb{R}^p)$ can be represented in terms of the (extended) support function introduced by Puri and Ralescu [156] (see, for instance, Diamond and Kloeden [56] and Liang *et al.* [122] for more details).

Definition 1.1.6. *The **support function** of $\widetilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$ is given by the mapping $s_{\widetilde{U}} : [0,1] \times \mathbb{S}^{p-1} \to \mathbb{R}$ defined so that*

$$s_{\widetilde{U}}(\alpha, \mathbf{u}) = s_{\widetilde{U}_\alpha}(\mathbf{u}) = \sup_{\mathbf{x} \in \widetilde{U}_\alpha} \langle \mathbf{u}, \mathbf{x} \rangle$$

for all $\mathbf{u} \in \mathbb{S}^{p-1}, \alpha \in [0,1]$.

It should be emphasized that a fuzzy set in $\mathcal{F}_c(\mathbb{R}^p)$ is uniquely determined by its support function. Actually, necessary and sufficient conditions are known to ensure such a uniqueness (see, for instance, Butnariu *et al.* [22], Theorem 1, p. 24):

Proposition 1.1.2. [22] *If $s : [0,1] \times \mathbb{S}^{p-1} \to \mathbb{R}$ is a function such that*

s.i) is subadditive in $\mathbb{S}^{p-1}$, that is, for all $\alpha \in [0,1]$, $\mathbf{u}, \mathbf{v} \in \mathbb{S}^{p-1}$,

$$\|\mathbf{u} + \mathbf{v}\| \cdot s\left(\alpha, \frac{\mathbf{u} + \mathbf{v}}{\|\mathbf{u} + \mathbf{v}\|}\right) \leq s(\alpha, \mathbf{u}) + s(\alpha, \mathbf{v})$$

and

s.ii) for all $\mathbf{u} \in \mathbb{S}^{p-1}$, the function $s(\cdot, \mathbf{u})$ is non-increasing, left-continuous on $(0,1]$ and right-continuous at 0,

then, there exists a unique $\widetilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$ such that $s = s_{\widetilde{U}}$. Actually, $\widetilde{U}$ is the fuzzy set such that for all $\alpha \in [0,1]$:

$$\widetilde{U}_\alpha = \{\mathbf{x} \in \mathbb{R}^p \,:\, \langle \mathbf{u}, \mathbf{x} \rangle \leq s(\alpha, \mathbf{u}) \text{ for all } \mathbf{u} \in \mathbb{S}^{p-1}\}.$$

In the particular case $p = 1$, i.e., in dealing with fuzzy number-valued data, two characterizing representations are intuitive and easy-to-use.

Definition 1.1.7. *The **inf/sup representation** of the fuzzy number $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ is the vector-valued function $\boldsymbol{\iota}_{\widetilde{U}} = (\iota^l_{\widetilde{U}}, \iota^r_{\widetilde{U}}) : [0,1] \to \{(x,y) \in \mathbb{R}^2 \,:\, x \leq y\}$ such that $\boldsymbol{\iota}_{\widetilde{U}}(\alpha) = \boldsymbol{\iota}_{\widetilde{U}_\alpha}$, that is, $\iota^l_{\widetilde{U}}(\alpha) = \inf \widetilde{U}_\alpha$, $\iota^r_{\widetilde{U}}(\alpha) = \sup \widetilde{U}_\alpha$.*

It should be emphasized that a fuzzy set in $\mathcal{F}_c(\mathbb{R})$ is uniquely determined by its inf/sup representation, as the following result states (see, for instance, Goetschel and Voxman [87] and Ming [133]-Theorem 3.1, pp. 187-188):

Proposition 1.1.3. [87] *Given a fuzzy number $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$, there exist two functions $l : [0,1] \to \mathbb{R}$ and $r : [0,1] \to \mathbb{R}$ satisfying that*

i) l and r are
  - *left-continuous on $(0,1]$,*
  - *right-continuous at 0,*
  - *non-increasing on $[0,1]$,*

ii) $-l(1) \leq r(1)$,

such that

$$\boldsymbol{\iota}_{\widetilde{U}}(\alpha) = (-l(\alpha), r(\alpha)) \text{ for all } \alpha \in [0,1].$$

Conversely, let $l : [0,1] \to \mathbb{R}$ *and* $r : [0,1] \to \mathbb{R}$ *be two functions satisfying Conditions i) and ii). Then, there exists a unique* $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ *such that the vector-valued function* $(-l, r)$ *is its inf/sup representation.*

Another representation that has been used in providing an alternative interpretation and expression for the metric by Bertoluzza *et al.* [16] (see Trutschnig *et al.* [201], Casals *et al.* [25]) and in formalizing some statistical developments with fuzzy number-valued data is the *mid/spr representation*. This representation associates each fuzzy number $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ with the vector-valued function $\boldsymbol{\eta}_{\widetilde{U}} = (\eta^m_{\widetilde{U}}, \eta^s_{\widetilde{U}}) : [0,1] \to \mathbb{R} \times [0,\infty)$ such that $\boldsymbol{\eta}_{\widetilde{U}}(\alpha) = \boldsymbol{\eta}_{\widetilde{U}_\alpha}$, that is, $\eta^m_{\widetilde{U}}(\alpha) = \text{mid}\, \widetilde{U}_\alpha$, $\eta^s_{\widetilde{U}}(\alpha) = \text{spr}\, \widetilde{U}_\alpha$. Components in this representation satisfy that $\eta^s_{\widetilde{U}}$ is a left-continuous non-increasing and nonnegative function on $(0,1]$ and right-continuous at 0, whereas $\eta^m_{\widetilde{U}}$ is also left-continuous on $(0,1]$ and right-continuous at 0, but nothing can be ensured in general in connection with its monotonicity. In fact, in contrast to the inf/sup representation, one cannot state a set of necessary and sufficient conditions for the functions involved in the mid/spr representation to characterize a fuzzy number.

Alternatively, and aiming to extend the mid/spr representation of the interval-valued case, one can introduce a new representation. This representation also takes into account the center and the shape, although in a slightly different way, so it is possible to establish necessary and sufficient conditions to determine a fuzzy number. Since this representation is a novelty, its construction will be explained in detail.

The new representation is based on considering an alternative indicator of the 'center' (instead of considering the mid function) along with an indicator of the 'shape' quantifying the deviation with respect to the center (instead of considering the spr function). A suitable indicator of the 'center' of a fuzzy number is the one given by Yager [218] and later extended by De Campos and González [52] (as the .5-average index) and by Nasibov [142] (as the weighted averaging based on levels -see also Nasibov *et al.* [143]). For any $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$, the ***weighted averaging based on levels*** is defined as the real number in the interior set $\text{int}(\widetilde{U}_0)$ such that

$$\text{wabl}^\varphi(\widetilde{U}) = \int_{[0,1]} \text{mid}\, \widetilde{U}_\alpha \, d\varphi(\alpha),$$

where φ is a weighting measure on the measurable space $([0,1], \mathcal{B}_{[0,1]})$ that can be formalized by means of an absolutely continuous probability measure with positive mass function on $(0,1)$.

It should be pointed out that no stochastic meaning is actually associated with φ, but it allows us to weight the 'degrees of compatibility' given by the α-levels.

The wabl^φ is often used as a defuzzification function to rank fuzzy numbers and it coincides with the well-known generalized Steiner point (or centroid) of a fuzzy number (see, for instance, Diamond and Kloeden [55, 56], Körner [117], Diamond and Körner [57], Butnariu *et al.* [22], Vetterlein and Navara [209, 210], and Liang *et al.* [122]) by extending level-wise the Steiner points for convex sets (see Schneider [166, 167]).

The wabl^φ is one of the three components of the new representation of fuzzy numbers. The other two components are level-wise indicators of the shape of a fuzzy number with respect to the considered center. They can be formalized as the following functions:

$$\text{ldev}^\varphi_{\widetilde{U}} : [0,1] \to \mathbb{R}, \quad \alpha \mapsto \text{ldev}^\varphi_{\widetilde{U}}(\alpha) = \text{wabl}^\varphi(\widetilde{U}) - \inf \widetilde{U}_\alpha,$$
$$\text{rdev}^\varphi_{\widetilde{U}} : [0,1] \to \mathbb{R}, \quad \alpha \mapsto \text{rdev}^\varphi_{\widetilde{U}}(\alpha) = \sup \widetilde{U}_\alpha - \text{wabl}^\varphi(\widetilde{U}).$$

On the basis of these three components, we obtain a representation of fuzzy numbers.

Definition 1.1.8. *Let φ be an absolutely continuous probability measure associated with the measurable space $([0,1], \mathcal{B}_{[0,1]})$ and having positive mass function on $(0,1)$. The **φ-wabl/ldev/rdev representation** of the fuzzy number $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ is the vector-valued function $\boldsymbol{\upsilon}^\varphi_{\widetilde{U}} = (\upsilon^w_{\widetilde{U}}, \upsilon^l_{\widetilde{U}}, \upsilon^r_{\widetilde{U}}) : [0,1] \to \mathbb{R}^3$ such that $\upsilon^w_{\widetilde{U}}$ is constantly equal to $\text{wabl}^\varphi(\widetilde{U})$, $\upsilon^l_{\widetilde{U}}(\alpha) = \text{ldev}_{\widetilde{U}}(\alpha)$ and $\upsilon^r_{\widetilde{U}}(\alpha) = \text{rdev}^\varphi_{\widetilde{U}}(\alpha)$.*

For symmetric fuzzy number-valued data, the φ-wabl/ldev/rdev representation coincides with the mid/spr one, irrespective of φ. Consequently, it is indeed an extension of the mid/spr representation for interval-valued data.

As for the inf/sup representation, one can state necessary and sufficient conditions characterizing fuzzy numbers by their φ-wabl/ldev/rdev representation. Thus (see Appendix for the proof),

Proposition 1.1.4. *Given a fuzzy number $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ there exist a value $m \in \mathbb{R}$ and two functions $l^* : [0,1] \to \mathbb{R}$ and $r^* : [0,1] \to \mathbb{R}$ satisfying that*

i) l^* and r^* are

- *left-continuous functions at any $\alpha \in (0,1]$,*
- *right-continuous at 0,*
- *and non-increasing on $[0,1]$,*

with

ii) $-l^*(1) \leq r^*(1)$,

and such that for all $\alpha \in [0,1]$,

$$\widetilde{U}_\alpha = [m - l^*(\alpha), m + r^*(\alpha)] .$$

Conversely, let $m \in \mathbb{R}$ *and let* $l^* : [0,1] \to \mathbb{R}$ *and* $r^* : [0,1] \to \mathbb{R}$ *be functions satisfying Conditions i) and ii). Then there exists a unique* $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ *such that for all* $\alpha \in [0,1]$

$$\widetilde{U}_\alpha = [m - l^*(\alpha), m + r^*(\alpha)] .$$

Furthermore, if there is an absolutely continuous probability measure φ *on* $([0,1], \mathcal{B}_{[0,1]})$ *with positive mass function on* $(0,1)$ *and such that*

iii) $\displaystyle\int_{[0,1]} l^*(\alpha)\, d\varphi(\alpha) = \int_{[0,1]} r^*(\alpha)\, d\varphi(\alpha)$,

then, (m, l^*, r^*) *is the* φ-wabl/ldev/rdev *representation of* $\widetilde{U}$.

This result will be illustrated by means of an example.

Example 1.1.1. Let $m = 8$, $l^*(\alpha) = 5 - 3\alpha^2$ and $r^*(\alpha) = 7 - 6\alpha$ for $\alpha \in [0,1]$. Since functions l^* and r^* satisfy Conditions *i)* and *ii)* in Proposition 1.1.4, there exists a unique bounded fuzzy number $\widetilde{U}$ such that

$$\widetilde{U}_\alpha = [m - l^*(\alpha), m + r^*(\alpha)] = \left[3 + 3\alpha^2, 15 - 6\alpha\right]$$

for every $\alpha \in [0,1]$. This fuzzy number $\widetilde{U}$ is shown in Figure 1.2 and is given by

$$\widetilde{U}(x) = \begin{cases} \sqrt{(x-3)/3} & \text{if } x \in [3,6) \\ 1 & \text{if } x \in [6,9) \\ (15-x)/6 & \text{if } x \in [9,15] \\ 0 & \text{otherwise.} \end{cases}$$

On the other hand, the equality

$$\int_{[0,1]} (5 - 3\alpha^2)\, d\alpha = 4 = \int_{[0,1]} (7 - 6\alpha)\, d\alpha$$

implies that for $\varphi \equiv \ell \equiv$ Lebesgue measure in $[0,1]$ we have that

$$\text{wabl}^\ell(\widetilde{U}) = 8, \ \text{ldev}^\ell_{\widetilde{U}}(\alpha) = 5 - 3\alpha^2, \ \text{rdev}^\ell_{\widetilde{U}}(\alpha) = 7 - 6\alpha.$$

Some considerations should be made in connection with the last representation.

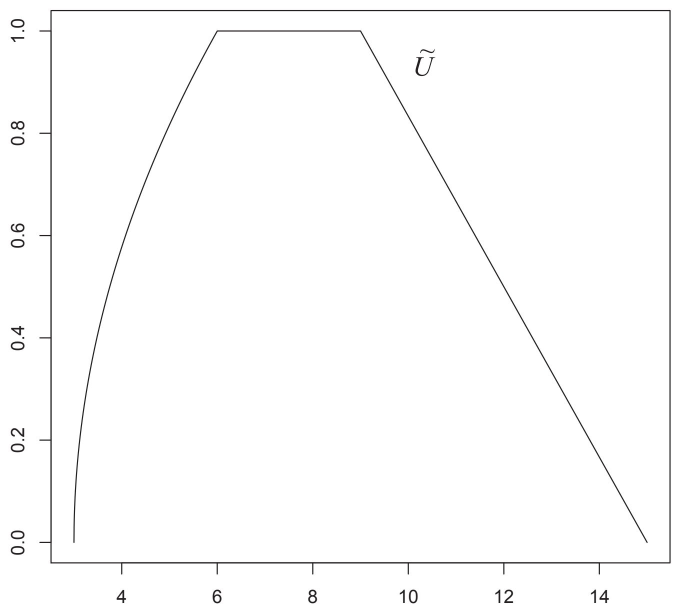


Figure 1.2: Fuzzy number for the given ℓ-wabl/ldev/rdev representation in Example 1.1.1

Remark 1.1.1. The wabl/ldev/rdev representation for this fuzzy number is not unique, since it is indeed possible to find a wabl/ldev/rdev representation for every choice of φ. For instance, if one chooses $\varphi \equiv \text{Beta}(5,1)$ (i.e., the probability measure associated with the Beta$(5,1)$ distribution), then

$$\text{wabl}^{\varphi}(\widetilde{U}) = 53/7,\ \text{ldev}^{\varphi}_{\widetilde{U}}(\alpha) = (32 - 21\alpha^2)/7,\ \text{rdev}^{\varphi}_{\widetilde{U}}(\alpha) = (52 - 42\alpha)/7.$$

Note that for the Beta$(5,1)$ distribution, the larger the α-level of a set (that is to say, the greater the degree of compatibility with $\widetilde{U}$), the larger its weight in the corresponding wabl/ldev/rdev representation.

Remark 1.1.2. It should also be emphasized that, on the basis of the Weighted Mean Value Theorem for integrals, whenever $\text{mid}\,\widetilde{U}_\alpha$ is a continuous function of α in $[0,1]$, then for each φ there exists at least one $\beta_\varphi \in [0,1]$ such that

$$\text{mid}\,\widetilde{U}_{\beta_\varphi} = \int_{[0,1]} \text{mid}\,\widetilde{U}_\alpha \, d\varphi(\alpha) = \text{wabl}^{\varphi}(\widetilde{U}).$$

The wabl/ldev/rdev representation can be extended to deal with fuzzy vector-valued data as follows:

Definition 1.1.9. *Given an absolutely continuous probability measure φ on the measurable space $([0,1], \mathcal{B}_{[0,1]})$, the **$\varphi$-support/Steiner representation** of the fuzzy vector $\widetilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$ is given by the vector-valued function:*

$$\boldsymbol{\tau}^{\varphi}_{\widetilde{U}} : [0,1] \times \mathbb{S}^{p-1} \to \mathbb{R}^p \times \mathbb{R}, \quad (\alpha, \mathbf{u}) \mapsto \boldsymbol{\tau}^{\varphi}_{\widetilde{U}}(\alpha, \mathbf{u}) = (\mathbf{S}^{\varphi}(\widetilde{U}), s_{\widetilde{U}}(\alpha, \mathbf{u})),$$

*where $\mathbf{S}^{\varphi}(\widetilde{U})$ stands for the **φ-Steiner point** of $\widetilde{U}$, which, if it exists, corresponds to the vector value:*

$$\mathbf{S}^{\varphi}(\widetilde{U}) = \int_{[0,1]\times\mathbb{S}^{p-1}} \mathbf{u} \cdot s_{\widetilde{U}}(\alpha, \mathbf{u})\, d\lambda_p(\mathbf{u})\, d\varphi(\alpha),$$

with λ_p denoting the normalized Lebesgue measure on $\mathbb{S}^{p-1}$.

The following result establishes necessary and sufficient conditions to characterize each fuzzy vector by a support/Steiner representation.

Proposition 1.1.5. *If there exist a function $s : [0,1] \times \mathbb{S}^{p-1} \to \mathbb{R}$ fulfilling conditions s.i) and s.ii) in Proposition 1.1.2 and an absolutely continuous probability measure φ on $([0,1], \mathcal{B}_{[0,1]})$, then, there exists a unique fuzzy vector $\widetilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$ such that*

$$\boldsymbol{\tau}^{\varphi}(\widetilde{U}) = (\mathbf{m}, s),$$

with $\mathbf{m} = \displaystyle\int_{[0,1]\times\mathbb{S}^{p-1}} \mathbf{u} \cdot s(\alpha, \mathbf{u})\, d\lambda_p(\mathbf{u})\, d\varphi(\alpha)$.

1.2 Arithmetics with imprecise data

In performing statistics with imprecise data, one of the key tools is given by the arithmetic to operate with these data. More concretely, the elementary operations to be specified are the sum of imprecise data and the multiplication of scalars by imprecise data.

Although there is no full agreement about how these operations should be formalized, most of the theoretical and practical studies with imprecise data consider the usual and natural approaches, which will be recalled in next subsections.

1.2.1 Arithmetic with set-valued data

In extending the sum and the product by a scalar from the Euclidean space $\mathbb{R}^p$ to $\mathcal{K}_c(\mathbb{R}^p)$, a natural way to proceed consists of defining these operations as the image sets of the involved set-valued data through the function associated with the corresponding operation (see Minkowski [134]). Thus,

Definition 1.2.1. *Let $K, K' \in \mathcal{K}_c(\mathbb{R}^p)$. The* ***Minkowski sum*** *(or* ***Minkowski addition****) of K and K' is defined as the set value $K + K' \in \mathcal{K}_c(\mathbb{R}^p)$ given by*

$$K + K' = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in K, \mathbf{y} \in K'\} = \bigcup_{\mathbf{y} \in K'} (K + \mathbf{y}) = \bigcup_{\mathbf{x} \in K} (\mathbf{x} + K'),$$

where $K + \mathbf{y} = K + \{\mathbf{y}\} = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in K\} = \mathbf{y} + K = \{\mathbf{y}\} + K$.

In particular, if $p = 1$, $K = [a, b]$ and $K' = [a', b']$, then $K + K' = [a + a', b + b']$.

As already indicated by Schneider [167], the Minkowski sum $K + K'$ can be 'kinematically' interpreted as the set that is covered if K undergoes all translations by vector values in K'.

Figure 1.3 graphically illustrates the sum of two set-valued data in case $p = 2$.

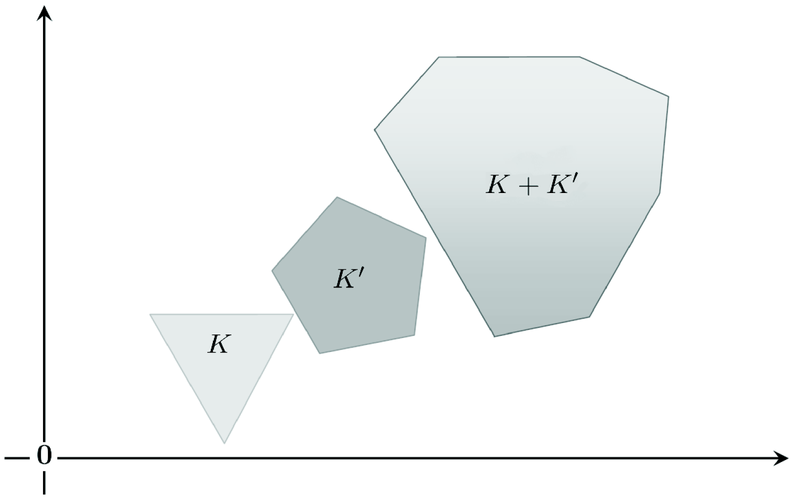


Figure 1.3: Example of the Minkowski sum of two set-valued data

On the other hand,

Definition 1.2.2. *Let $K \in \mathcal{K}_c(\mathbb{R}^p)$ and $\gamma \in \mathbb{R}$. The* ***product*** *of K* ***by the scalar*** *γ is defined as the set value $\gamma \cdot K \in \mathcal{K}_c(\mathbb{R}^p)$ given by*

$$\gamma \cdot K = \{\gamma \mathbf{x} : \mathbf{x} \in K\}.$$

In particular, if $p = 1$ and $K = [a, b]$, then

$$\gamma \cdot K = \begin{cases} [\gamma\, a, \gamma\, b] & \text{if } \gamma \geq 0 \\ \\ [\gamma\, b, \gamma\, a] & \text{otherwise.} \end{cases}$$

Figure 1.4 graphically illustrates the product of a set-valued data by a scalar in case $p = 2$.

These operations satisfy the following properties (see, for instance, Schneider [167]):

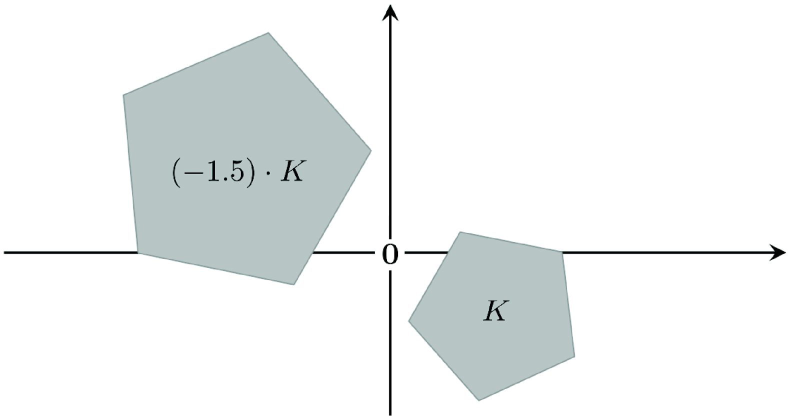


Figure 1.4: Example of the product of a set-valued datum by a scalar

Proposition 1.2.1. [167] *The Minkowski sum and the product by a scalar on $\mathcal{K}_c(\mathbb{R}^p)$ satisfy that for all $K, K', K'' \in \mathcal{K}_c(\mathbb{R}^p)$ and $\gamma, \varrho \in \mathbb{R}$*

- $K + K' = K' + K$ *(commutativity of the Minkowski sum);*
- $K + (K' + K'') = (K + K') + K''$ *(associativity of the Minkowski sum);*
- $K + \{\mathbf{0}\} = K$ *(the neutral element of the Minkowski sum is $\{\mathbf{0}\}$);*
- $\gamma \cdot (\varrho \cdot K) = (\gamma \cdot \varrho) \cdot K$ *(associativity of the product by a scalar);*
- $1 \cdot K = K$ *(the neutral scalar element of the product by a scalar is 1);*
- $\gamma \cdot (K + K') = \gamma \cdot K + \gamma \cdot K'$ *whenever $\gamma, \varrho \in [0, \infty)$ (distributivity of the product by a scalar w.r.t. the Minkowski sum);*
- $(\gamma + \varrho) \cdot K = \gamma \cdot K + \varrho \cdot K$ *(distributivity of the Minkowski sum w.r.t. the product by nonnegative scalars);*
- $K + K' = K'' + K'$ *implies that* $K = K''$ *(cancellation law of the Minkowski sum).*

The point-wise 'opposite' and 'difference' of set-valued data could be defined as

$$-K = \{-\mathbf{x} : \mathbf{x} \in K\}, \quad K - K' = \{\mathbf{x} - \mathbf{y} : \mathbf{x} \in K, \mathbf{y} \in K'\},$$

whence $-K$ is the image of K under reflection w.r.t. $\{\mathbf{0}\}$, and $K - K'$ is the set that is covered if K undergoes all reflected w.r.t. $\{\mathbf{0}\}$ translations by vector values in K'.

It can be straightforwardly proved that these point-wisely defined terms can also be formalized through the two elementary algebraic operations on $\mathcal{K}_c(\mathbb{R}^p)$ introduced in Definitions 1.2.1 and 1.2.2, so that

$$-K = (-1) \cdot K, \quad K - K' = K + (-1) \cdot K'.$$

Unlike what happens in the vector-valued case, a remarkable differential aspect of this arithmetic is that $K-K$ does not coincide in general with the neutral element $\{\mathbf{0}\}$, and hence, $(K-K')+K'$ and $(K+K')-K'$ do not coincide with K. Actually (see, for instance, Schneider [167]),

Proposition 1.2.2. [167] *Whatever $K, K' \in \mathcal{K}_c(\mathbb{R}^p)$ may be*

- $\{\mathbf{0}\} \subset K-K$, *with*
  - •• $\{\mathbf{0}\} = K-K$ *if and only if K reduces to a singleton in $\mathbb{R}^p$,*
  - •• $K-K$ *being centrally symmetric w.r.t.* $\mathbf{0}$ *(i.e., $\mathbf{x} \in K-K$ if and only if $-\mathbf{x} \in K-K$);*
- $K \subset (K+K')-K'$ *and* $K \supset (K-K')+K'$.

As a consequence from the last preceding results one can conclude that

Proposition 1.2.3. [167] *The space $\mathcal{K}_c(\mathbb{R}^p)$ satisfies that*

- *with the Minkowski sum is a commutative semigroup, although not a group;*
- *with the Minkowski sum and the product by a scalar is a semilinear space (in fact, a convex cone), but not a linear (vector) space.*

The operations in Definitions 1.2.1 and 1.2.2 could be alternatively defined in terms of their characterizing representations in Subsection 1.1.1. Thus, it is well-known that by considering the functional arithmetic with support functions

Proposition 1.2.4. [134] *Whatever $K, K' \in \mathcal{K}_c(\mathbb{R}^p)$ and $\gamma \in [0,\infty)$ may be*

- $s_{K+K'}(\mathbf{u}) = s_K(\mathbf{u}) + s_{K'}(\mathbf{u})$ *for all* $\mathbf{u} \in \mathbb{S}^{p-1}$*;*
- $s_{\gamma\cdot K}(\mathbf{u}) = \gamma \cdot s_K(\mathbf{u})$ *for all* $\mathbf{u} \in \mathbb{S}^{p-1}$.

In case $p=1$, it is trivial that

$$\inf(K+K') = \inf K + \inf K', \quad \inf(\gamma\cdot K) = \begin{cases} \gamma\cdot\inf K & \text{if } \gamma \geq 0 \\ \gamma\cdot\sup K & \text{otherwise} \end{cases}$$

$$\sup(K+K') = \sup K + \sup K', \quad \sup(\gamma\cdot K) = \begin{cases} \gamma\cdot\sup K & \text{if } \gamma \geq 0 \\ \gamma\cdot\inf K & \text{otherwise} \end{cases}$$

and

$$\operatorname{mid}(K+K') = \operatorname{mid} K + \operatorname{mid} K', \quad \operatorname{mid}(\gamma\cdot K) = \gamma\cdot\operatorname{mid} K,$$

$$\operatorname{spr}(K+K') = \operatorname{spr} K + \operatorname{spr} K', \quad \operatorname{spr}(\gamma\cdot K) = |\gamma|\cdot\operatorname{mid} K,$$

whence

Proposition 1.2.5. *Whatever $K, K' \in \mathcal{K}_c(\mathbb{R})$ and $\gamma \in [0,\infty)$ may be*

$$\boldsymbol{\iota}_{K+K'} = \boldsymbol{\iota}_K + \boldsymbol{\iota}_{K'}, \quad \boldsymbol{\iota}_{\gamma\cdot K} = \gamma\cdot\boldsymbol{\iota}_K,$$
$$\boldsymbol{\eta}_{K+K'} = \boldsymbol{\eta}_K + \boldsymbol{\eta}_{K'}, \quad \boldsymbol{\eta}_{\gamma\cdot K} = \gamma\cdot\boldsymbol{\eta}_K.$$

1.2.2 Arithmetic with fuzzy set-valued data

The common way to extend the sum and the product by a scalar from the Euclidean space $\mathbb{R}^p$ to $\mathcal{F}_c(\mathbb{R}^p)$ is to use *Zadeh's extension principle* (Zadeh [219]). It provides a general method for extending nonfuzzy mathematical concepts in order to deal with fuzzy set-valued data. In particular, it has been systematically applied to the algebra of real and vector values from which operations on fuzzy set-values have been extensively developed.

These operations generalize the usual set-valued arithmetic. Furthermore, because of the assumed compactness of the level sets and the consequent upper semicontinuity of the fuzzy set-valued data, the results in Nguyen [146] guarantee the equivalence of Zadeh's principle with its *level set form*. This equivalence indicates that the sum and the product by a scalar of fuzzy set-valued data based on Zadeh's extension principle are equivalent level-wise to the corresponding set-valued operations. Since this set-valued approach is in general much simpler than the one based on functions, it will be more convenient to use it for most of the theoretical and practical developments.

Now, the two elementary operations are to be recalled following the two equivalent approaches on $\mathcal{F}_c(\mathbb{R}^p)$.

Definition 1.2.3. *Let $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$. The* ***sum*** *of $\widetilde{U}$ and $\widetilde{V}$ is defined as the fuzzy set value $\widetilde{U} + \widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$ given by*

$$(\widetilde{U} + \widetilde{V})(\mathbf{t}) = \sup_{\mathbf{y},\mathbf{z}\in\mathbb{R}^p\,:\,\mathbf{y}+\mathbf{z}=\mathbf{t}} \min\big\{\widetilde{U}(\mathbf{y}), \widetilde{V}(\mathbf{z})\big\}.$$

Equivalently, for each $\alpha \in [0,1]$

$$(\widetilde{U} + \widetilde{V})_\alpha = \textit{Minkowski sum of } \widetilde{U}_\alpha \textit{ and } \widetilde{V}_\alpha = \big\{\mathbf{y} + \mathbf{z} \,:\, \mathbf{y} \in \widetilde{U}_\alpha, \mathbf{z} \in \widetilde{V}_\alpha\big\}.$$

Figure 1.5 illustrates the sum of two fuzzy set-valued data in case $p = 1$ graphically. It has been obtained by applying the R package SAFD (Statistical Analysis of Fuzzy Data), which has been developed (see Trutschnig and Lubiano [202] and Trutschnig *et al.* [203]), among other purposes, to ease the performance of the standard operations on the class of fuzzy numbers. On the other hand,

Definition 1.2.4. *Let $\widetilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$ and $\gamma \in \mathbb{R}$. The* ***product*** *of $\widetilde{U}$* ***by the scalar*** *γ is defined as the fuzzy set value $\gamma \cdot \widetilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$ given by*

$$(\gamma \cdot \widetilde{U})(\mathbf{t}) = \sup_{\mathbf{y}\in\mathbb{R}^p\,:\,\mathbf{y}=\gamma\mathbf{t}} \widetilde{U}(\mathbf{y}) = \begin{cases} \widetilde{U}(\mathbf{t}/\gamma) & \text{if } \gamma \neq 0 \\ \mathbb{1}_{\{0\}}(\mathbf{t}) & \text{if } \gamma = 0. \end{cases}$$

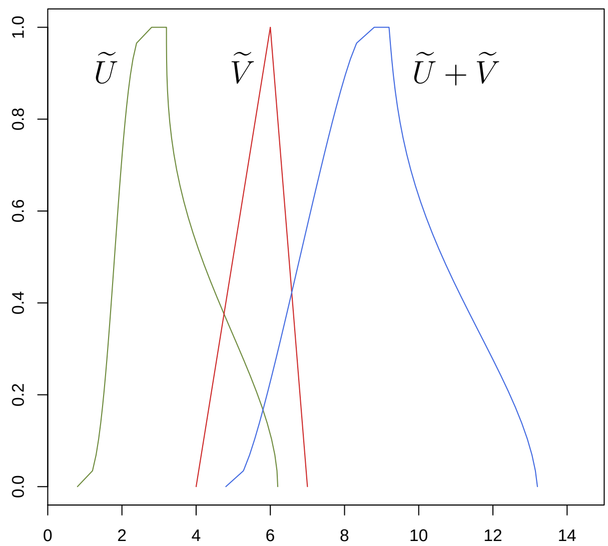


Figure 1.5: Example of the sum of two fuzzy set-valued data

Equivalently, for each $\alpha \in [0,1]$

$$(\gamma \cdot \widetilde{U})_\alpha = \gamma \cdot \widetilde{U}_\alpha = \{\gamma \cdot \mathbf{y} \,:\, \mathbf{y} \in \widetilde{U}_\alpha\}.$$

Figure 1.6 graphically illustrates the product of a scalar by a fuzzy set-valued datum in case $p = 1$, and it has been obtained by applying SAFD.

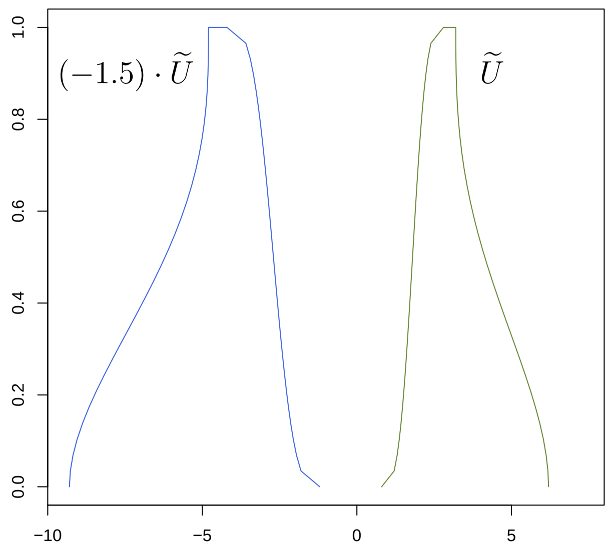


Figure 1.6: Example of the product of a fuzzy set-valued datum by a scalar

Based on the properties in the set-valued case (Proposition 1.2.1) and the level set form of the extension principle, one can trivially conclude that the operations above satisfy the following properties:

Proposition 1.2.6. *The sum and the product by a scalar on $\mathcal{F}_c(\mathbb{R}^p)$ based on Zadeh's extension principle satisfy that for all $\widetilde{U}, \widetilde{V}, \widetilde{W} \in \mathcal{F}_c(\mathbb{R}^p)$ and $\gamma, \varrho \in \mathbb{R}$*

- $\widetilde{U} + \widetilde{V} = \widetilde{V} + \widetilde{U}$ *(commutativity of the fuzzy sum);*
- $\widetilde{U} + (\widetilde{V} + \widetilde{W}) = (\widetilde{U} + \widetilde{V}) + \widetilde{W}$ *(associativity of the fuzzy sum);*
- $\widetilde{U} + \mathbb{1}_{\{\mathbf{0}\}} = \widetilde{U}$ *(the neutral element of the fuzzy sum is the indicator $\mathbb{1}_{\{0\}}$);*
- $\gamma \cdot (\varrho \cdot \widetilde{U}) = (\gamma \cdot \varrho) \cdot \widetilde{U}$ *(associativity of the fuzzy product by a scalar);*
- $1 \cdot \widetilde{U} = \widetilde{U}$ *(the neutral scalar element of the fuzzy product by a scalar is 1);*
- $\gamma \cdot (\widetilde{U} + \widetilde{V}) = \gamma \cdot \widetilde{U} + \gamma \cdot \widetilde{V}$ *(distributivity of the fuzzy product by a scalar w.r.t. the fuzzy sum);*
- $(\gamma + \varrho) \cdot \widetilde{U} = \gamma \cdot \widetilde{U} + \varrho \cdot \widetilde{U}$ *whenever $\gamma, \varrho \in [0, \infty)$ (distributivity of the fuzzy sum w.r.t. the fuzzy product by nonnegative scalars);*
- $\widetilde{U} + \widetilde{V} = \widetilde{W} + \widetilde{V}$ *implies that $\widetilde{U} = \widetilde{W}$ (cancellation law of the fuzzy sum).*

Due to the equivalence between the set-valued operations and their point-wise approach, the point- and level-wise 'opposite' and 'difference' of fuzzy set-valued data could be immediately defined as

$$-\widetilde{U} = (-1) \cdot \widetilde{U}, \quad \widetilde{U} - \widetilde{V} = \widetilde{U} + (-1) \cdot \widetilde{V}.$$

As for the set-valued case, a remarkable differential aspect of this arithmetic with respect to the real and vectorial ones is that $\widetilde{U} - \widetilde{U}$ does not coincide in general with the neutral element $\mathbb{1}_{\{\mathbf{0}\}}$, and hence, $(\widetilde{U} - \widetilde{V}) + \widetilde{V}$ and $(\widetilde{U} + \widetilde{V}) - \widetilde{V}$ do not coincide with $\widetilde{U}$. Actually, and based on Proposition 1.2.2, one can trivially conclude that

Proposition 1.2.7. *Whatever $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$ may be*

- $\{\mathbf{0}\} \subset (\widetilde{U} - \widetilde{U})_\alpha$ *for all $\alpha \in [0, 1]$, with*
  - $\mathbb{1}_{\{\mathbf{0}\}} = \widetilde{U} - \widetilde{U}$ *if and only if $\widetilde{U}$ reduces to the indicator function of a singleton in $\mathbb{R}^p$,*
  - $(\widetilde{U} - \widetilde{U})_\alpha$ *being centrally symmetric w.r.t. $\mathbf{0}$ for all α;*
- $\widetilde{U} \subset (\widetilde{U} + \widetilde{V}) - \widetilde{V}$ *and* $\widetilde{U} \supset (\widetilde{U} - \widetilde{V}) + \widetilde{V}$.

As a consequence one can derive that

Proposition 1.2.8. *The space $\mathcal{F}_c(\mathbb{R}^p)$ satisfies that*

- *with the fuzzy sum is a commutative semigroup, although not a group;*

- *with the fuzzy sum and the product by a scalar is a semilinear space (in fact, a convex cone), but not a linear (vector) space.*

The operations in Definitions 1.2.3 and 1.2.4 could be alternatively defined in terms of their characterizing representations in Subsection 1.1.2. Thus, it can be trivially stated (see, for instance, Puri and Ralescu [156]) that by considering the functional arithmetic with support functions

Proposition 1.2.9. [156] *Whatever $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$ and $\gamma \in [0,\infty)$ may be*

- $s_{\widetilde{U}+\widetilde{V}}(\alpha, \mathbf{u}) = s_{\widetilde{U}}(\alpha, \mathbf{u}) + s_{\widetilde{V}}(\alpha, \mathbf{u})$ *for all* $\mathbf{u} \in \mathbb{S}^{p-1}$ *and* $\alpha \in [0,1]$*;*
- $s_{\gamma\cdot\widetilde{U}}(\alpha, \mathbf{u}) = \gamma \cdot s_{\widetilde{U}}(\alpha, \mathbf{u})$ *for all* $\mathbf{u} \in \mathbb{S}^{p-1}$ *and* $\alpha \in [0,1]$.

In case $p = 1$, it is trivial that for all $\alpha \in [0,1]$

$$\inf(\widetilde{U}+\widetilde{V})_\alpha = \inf \widetilde{U}_\alpha + \inf \widetilde{V}_\alpha, \quad \inf(\gamma \cdot \widetilde{U})_\alpha = \begin{cases} \gamma \cdot \inf \widetilde{U}_\alpha & \text{if } \gamma \geq 0 \\ \gamma \cdot \sup \widetilde{U}_\alpha & \text{otherwise} \end{cases}$$

$$\sup(\widetilde{U}+\widetilde{V})_\alpha = \sup \widetilde{U}_\alpha + \sup \widetilde{V}_\alpha, \quad \sup(\gamma \cdot \widetilde{U})_\alpha = \begin{cases} \gamma \cdot \sup \widetilde{U}_\alpha & \text{if } \gamma \geq 0 \\ \gamma \cdot \inf \widetilde{U}_\alpha & \text{otherwise} \end{cases}$$

and

$$\text{wabl}^\varphi(\widetilde{U}+\widetilde{V}) = \text{wabl}^\varphi(\widetilde{U}) + \text{wabl}^\varphi(\widetilde{V}), \quad \text{wabl}^\varphi(\gamma \cdot \widetilde{U}) = \gamma \cdot \text{wabl}^\varphi(\widetilde{U}),$$

$$\text{ldev}^\varphi_{\widetilde{U}+\widetilde{V}}(\alpha) = \text{ldev}^\varphi_{\widetilde{U}}(\alpha) + \text{ldev}^\varphi_{\widetilde{V}}(\alpha), \quad \text{ldev}^\varphi_{\gamma\cdot\widetilde{U}}(\alpha) = \begin{cases} \gamma \cdot \text{ldev}^\varphi_{\widetilde{U}}(\alpha) & \text{if } \gamma \geq 0 \\ -\gamma \cdot \text{rdev}^\varphi_{\widetilde{U}}(\alpha) & \text{otherwise} \end{cases}$$

$$\text{rdev}^\varphi_{\widetilde{U}+\widetilde{V}}(\alpha) = \text{rdev}^\varphi_{\widetilde{U}}(\alpha) + \text{rdev}^\varphi_{\widetilde{V}}(\alpha), \quad \text{rdev}^\varphi_{\gamma\cdot\widetilde{U}}(\alpha) = \begin{cases} \gamma \cdot \text{rdev}^\varphi_{\widetilde{U}}(\alpha) & \text{if } \gamma \geq 0 \\ -\gamma \cdot \text{ldev}^\varphi_{\widetilde{U}}(\alpha) & \text{otherwise} \end{cases}$$

whence

Proposition 1.2.10. *Whatever $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R})$ and $\gamma \in [0,\infty)$ may be, for all $\alpha \in [0,1]$ we have that*

$$\boldsymbol{\iota}_{\widetilde{U}+\widetilde{V}}(\alpha) = \boldsymbol{\iota}_{\widetilde{U}}(\alpha) + \boldsymbol{\iota}_{\widetilde{V}}(\alpha), \quad \boldsymbol{\iota}_{\gamma\cdot\widetilde{U}}(\alpha) = \gamma \cdot \boldsymbol{\iota}_{\widetilde{U}}(\alpha),$$

$$\boldsymbol{\upsilon}^\varphi_{\widetilde{U}+\widetilde{V}}(\alpha) = \boldsymbol{\upsilon}^\varphi_{\widetilde{U}}(\alpha) + \boldsymbol{\upsilon}^\varphi_{\widetilde{V}}(\alpha), \quad \boldsymbol{\upsilon}^\varphi_{\gamma\cdot\widetilde{U}}(\alpha) = \gamma \cdot \boldsymbol{\upsilon}^\varphi_{\widetilde{U}}(\alpha).$$

The last conclusion can be extended to $\mathcal{F}_c(\mathbb{R}^p)$ through the support/Steiner representation, so that

Proposition 1.2.11. *Whatever $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$ and $\gamma \in [0,\infty)$ may be, for all $\alpha \in [0,1]$ we have that*

$$\boldsymbol{\tau}^\varphi_{\widetilde{U}+\widetilde{V}}(\alpha) = \boldsymbol{\tau}^\varphi_{\widetilde{U}}(\alpha) + \boldsymbol{\tau}^\varphi_{\widetilde{V}}(\alpha), \quad \boldsymbol{\tau}^\varphi_{\gamma\cdot\widetilde{U}}(\alpha) = \gamma \cdot \boldsymbol{\tau}^\varphi_{\widetilde{U}}(\alpha).$$

1.3 Metrics between imprecise data

A relevant consequence from the nonlinearity of the spaces of set and fuzzy set values is that *there is no 'difference operation'* between these values that is simultaneously well-defined and preserves the main properties of the difference between real/vectorial values in connection with the sum. In fact, there exists a difference notion (Hukuhara's one) satisfying the last condition, but it cannot be defined for many set and fuzzy set values.

Moreover, it should be pointed out that, although fuzzy set values are formalized as $[0,1]$-valued functions and set values can be trivially identified with $\{0,1\}$-valued functions, one cannot treat directly imprecise data as if they were functional, in the way they are usually handled in Functional Data Analysis. This is due to the fact that none of the above-presented arithmetics coincide with the usual arithmetic with functions, so when we apply the functional arithmetic on either $\mathcal{K}_c(\mathbb{R}^p)$ (more concretely, the corresponding indicator functions) or $\mathcal{F}_c(\mathbb{R}^p)$, outputs are quite often out of this space and their meaning is lost.

These concerns have been substantially overcome in developing statistics with imprecise data by incorporating suitable distances between them. On one hand, distances will allow to 'translate' the equality of set/fuzzy set values into the distance between these values being equal to 0, as in the case of real values. On the other hand, as it will be shown later, appropriate distances also allow us to 'identify' set-valued and fuzzy set-valued data with functional ones through the support function.

Distances between imprecise data are a topic that has received a deep attention in the literature. In this way, it has often been considered in connection with studies of similarity between sets/fuzzy sets, as well as for statistical purposes such as classification of set/fuzzy set-valued elements or inferential statistics with set/fuzzy set-valued random elements. Regarding the last target, metrics have been used, among other applications,

- to obtain some limit and probabilistic results for set- and fuzzy set-valued random elements (see, for instance, Artstein and Vitale [5], Lyasenkho [129], Cuesta and Matrán [48], Molchanov [135, 137], Hess [103], Colubi *et al.* [31, 35], Proske and Puri [153], Krätschmer [120], Terán and Molchanov [200], Terán [197, 199], Quang and Thuan [147], and Aletti and Bongiorno [3]),
- in optimization problems, image analysis, signal theory, etc. (see, for instance, Huttenlocher *et al.* [109], Abbasbandy and Asady [2], Abbasbandy and Amirfakhrian [1], Ayala and López-Díaz [7], Ayala *et al.* [8], Báez-Sánchez *et al.* [10], Ban and Coroianu [11], Bera *et al.* [14], Coroianu [42], and Coroianu *et al.* [43]),

- and especially in performing many statistical analyses, like classifying imprecise data, testing hypotheses about the mean(s) or variance(s) of set/fuzzy set-valued random elements, regression analysis, fuzzy clustering, fuzzy decision tree, and so on (see, for instance, Cressie and Laslett [45], Näther [144, 145], Körner and Näther [119], Körner [118], García *et al.* [79], Montenegro *et al.* [140, 141], D'Urso [61], Coppi and D'Urso [37, 38], Ayala *et al.* [9], Gil *et al.* [85], Coppi *et al.* [39, 40], D'Urso and Giordani [63], D'Urso and Santoro [66, 67], Beresteanu and Molinari [15], González-Rodríguez *et al.* [88, 90, 91], D'Urso [62], D'Urso *et al.* [65, 64], Ferraro *et al.* [72], García-García and Santos-Rodríguez [80], García-García *et al.* [81], Ramos-Guajardo and Lubiano [161], Cappelli *et al.* [24], Ferraro and Giordani [73] and Guillaume *et al.* [98]).

In connection with set-valued data, the best known and most used metric is likely Hausdorff's one [101]. This metric (which is defined on a general metric space) is particularized on $\mathcal{K}_c(\mathbb{R}^p)$ as follows:

Definition 1.3.1. *Let $K, K' \in \mathcal{K}_c(\mathbb{R}^p)$. The* ***Hausdorff distance*** *between K and K' is given by*

$$d_H(K, K') = \delta_\infty(K, K') = \max\left\{\sup_{\mathbf{x}\in K}\inf_{\mathbf{y}\in K'}\|\mathbf{x}-\mathbf{y}\|, \sup_{\mathbf{y}\in K'}\inf_{\mathbf{x}\in K}\|\mathbf{x}-\mathbf{y}\|\right\}.$$

Hausdorff's distance is also denoted as δ_∞ because it is an L_∞ distance between the corresponding support functions, that is,

$$d_H(K, K') = \delta_\infty(K, K') = \sup_{\mathbf{u}\in\mathbb{S}^{p-1}}\|s_K(\mathbf{u}) - s_{K'}(\mathbf{u})\|$$

(see, for instance, Weil [214] and McClure and Vitale [131]).

The Hausdorff metric is sometimes referred to in the literature as the Pompeiu-Hausdorff metric. In fact, as Bârsan and Tiba have clearly explained [13], Pompeiu first defined in his PhD Thesis [151] a distance between sets by considering the so-called *écart mutuel*, which is given by $\sup_{\mathbf{x}\in K}\inf_{\mathbf{y}\in K'}\|\mathbf{x}-\mathbf{y}\|+\sup_{\mathbf{y}\in K'}\inf_{\mathbf{x}\in K}\|\mathbf{x}-\mathbf{y}\|$.

Anyway, in spite of the many advantages of Hausdorff's metric, for the purposes of this dissertation we will mainly consider L^2 and L^1 metrics. In this way, this section aims to either recall or introduce some suitable L^2 and L^1 metrics which are on the basis of the approaches to the robust location measures for imprecise data presented in Chapters 2 and 3 of this work.

1.3.1 L^2 metrics for imprecise data

Since set- and fuzzy set-valued data are characterized by their respective support functions, some rather immediate L^2 metrics to consider are those defined by Vitale [211] for set values and extended to fuzzy set values by Diamond and Kloeden [54].

Definition 1.3.2. [211] *Let $K, K' \in \mathcal{K}_c(\mathbb{R}^p)$. The **$\delta_2$ distance** between K and K' is given by*

$$\delta_2(K, K') = \sqrt{\int_{\mathbb{S}^{p-1}} \|s_K(\mathbf{u}) - s_{K'}(\mathbf{u})\|^2 \, d\lambda_p(\mathbf{u})}.$$

If $p = 1$, then $\delta_2(K, K') = \sqrt{\dfrac{[\inf K - \inf K']^2}{2} + \dfrac{[\sup K - \sup K']^2}{2}}$.

Vitale [211] proved that δ_2 and d_H induce the same topology on $\mathcal{K}_c(\mathbb{R}^p)$ and yield separable metric spaces.

Definition 1.3.3. [54] *Let $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$. The **$\boldsymbol{\rho_2}$ distance** between $\widetilde{U}$ and $\widetilde{V}$ is given by*

$$\boldsymbol{\rho_2}(\widetilde{U}, \widetilde{V}) = \sqrt{\int_{[0,1]\times\mathbb{S}^{p-1}} \|s_{\widetilde{U}}(\alpha, \mathbf{u}) - s_{\widetilde{V}}(\alpha, \mathbf{u})\|^2 \, d\lambda_p(\mathbf{u}) \, d\ell(\alpha)}.$$

The metric $\boldsymbol{\rho_2}$ can be extended by weighting the relevance of different levels. Let φ be an absolutely continuous probability measure on $([0,1], \mathcal{B}_{[0,1]})$ with the mass function being positive in $(0,1)$. The $\boldsymbol{\rho_2^\varphi}$ distance between $\widetilde{U}$ and $\widetilde{V}$ is given by

$$\boldsymbol{\rho_2^\varphi}(\widetilde{U}, \widetilde{V}) = \sqrt{\int_{[0,1]\times\mathbb{S}^{p-1}} \|s_{\widetilde{U}}(\alpha, \mathbf{u}) - s_{\widetilde{V}}(\alpha, \mathbf{u})\|^2 \, d\lambda_p(\mathbf{u}) \, d\varphi(\alpha)}.$$

Diamond and Kloeden [54] proved that $\boldsymbol{\rho_2}$ is topologically equivalent on $\mathcal{F}_c(\mathbb{R}^p)$ to $\boldsymbol{d_2}$, which is given (see Klement *et al.* [116]) by

$$\boldsymbol{d_2}(\widetilde{U}, \widetilde{V}) = \sqrt{\int_{[0,1]} \left[d_H(\widetilde{U}_\alpha, \widetilde{U}_\alpha)\right]^2 \, d\ell(\alpha)}$$

and both yield separable metric spaces. Analogously, one can easily prove that $\boldsymbol{\rho_2^\varphi}$ is topologically equivalent to $\boldsymbol{d_2^\varphi}(\widetilde{U}, \widetilde{V}) = \sqrt{\int_{[0,1]} \left[d_H(\widetilde{U}_\alpha, \widetilde{U}_\alpha)\right]^2 \, d\varphi(\alpha)}$, and the separability is also straightforwardly concluded.

By following the ideas in Rådström [159] on $\mathcal{K}_c(\mathbb{R}^p)$ and in Puri and Ralescu [154] on $\mathcal{F}_c(\mathbb{R}^p)$, the support functions of elements in these spaces allow us to embed isometrically each of these spaces into a convex cone of a Hilbert space of functions (more concretely, the space of the L^2-type real-valued functions on $\mathbb{S}^{p-1}$ and $[0,1] \times \mathbb{S}^{p-1}$, respectively, with the metrics λ_p and $\lambda_p \otimes \ell$). As it will be remarked later, this embedding allows us to identify set- and fuzzy set-valued data with functional data.

The metric $\boldsymbol{\rho_2^\varphi}$ can be extended by considering two families of L^2 metrics which pay attention to the 'center' of the involved values as well as to their 'shape' separately. One of these families has been introduced in previous papers and will be recalled here with a certain detail, while the other one is to be introduced for this work. Since set values are a special type of fuzzy set values, the metrics are first presented on $\mathcal{F}_c(\mathbb{R}^p)$ and will be particularized later to the case in which $p = 1$ and $\mathcal{K}_c(\mathbb{R}^p)$.

The first family of distances extending $\boldsymbol{\rho_2^\varphi}$ is that one introduced for fuzzy vector-valued data by Trutschnig *et al.* [201] as an extension of the metric for fuzzy number-valued data given by Bertoluzza *et al.* [16] (see Casals *et al.* [25] for a recent review about). This family is formalized as follows:

Definition 1.3.4. [201] *Let $\theta \in (0, +\infty)$ and let φ be an absolutely continuous probability measure on $([0,1], \mathcal{B}_{[0,1]})$ with the mass function being positive in $(0,1)$. Then, the **mid/spr-based** L^2 **distance** is defined as the mapping $D_\theta^\varphi : \mathcal{F}_c(\mathbb{R}^p) \times \mathcal{F}_c(\mathbb{R}^p) \to [0, +\infty)$ such that it associates each pair of elements of $\mathcal{F}_c(\mathbb{R}^p)$, $\widetilde{U}$ and $\widetilde{V}$, with the value*

$$D_\theta^\varphi(\widetilde{U}, \widetilde{V}) = \left[\int_{[0,1]\times\mathbb{S}^{p-1}} \left[\operatorname{mid} \Pi\, \widetilde{U}_\alpha(\mathbf{u}) - \operatorname{mid} \Pi\, \widetilde{V}_\alpha(\mathbf{u}) \right]^2 d\lambda_p(\mathbf{u})\, d\varphi(\alpha) \right.$$
$$\left. + \theta \int_{[0,1]\times\mathbb{S}^{p-1}} \left[\operatorname{spr} \Pi\, \widetilde{U}_\alpha(\mathbf{u}) - \operatorname{spr} \Pi\, \widetilde{V}_\alpha(\mathbf{u}) \right]^2 d\lambda_p(\mathbf{u})\, d\varphi(\alpha) \right]^{1/2},$$

where $\Pi\, \widetilde{U}_\alpha(\mathbf{u})$ denotes the projection of $\widetilde{U}_\alpha$ over the direction $\mathbf{u} \in \mathbb{S}^{p-1}$.

The mid/spr-based L^2 distance between elements in $\mathcal{K}_c(\mathbb{R}^p)$ is denoted by d_θ.

Notice that $\Pi\, \widetilde{U}_\alpha(\mathbf{u}) = [-s_{\widetilde{U}}(\alpha, -\mathbf{u}), s_{\widetilde{U}}(\alpha, \mathbf{u})]$.

Remark 1.3.1. Due to the meaning of $\operatorname{mid} \Pi\, \widetilde{U}_\alpha(\mathbf{u})$ and $\operatorname{spr} \Pi\, \widetilde{U}_\alpha(\mathbf{u})$, the choice of θ allows us to weight the effect of the deviation between spreads (which could be intuitively translated into the difference in 'shape') in contrast to the effect of the deviation between mid's (which can be intuitively translated into the difference in 'center') for each level.

On the other hand, the choice of φ enables to weight the relevance of different levels (i.e., the degree of 'imprecision'). Although it has been formalized as a probability measure, its role is not stochastic, but weighting.

Since $s(\alpha, \mathbf{u}) = \operatorname{mid} \Pi\, \widetilde{U}_\alpha(\mathbf{u}) + \operatorname{spr} \Pi\, \widetilde{U}_\alpha(\mathbf{u})$, being this decomposition orthogonal, one can trivially prove that $D_1^\varphi(\widetilde{U}, \widetilde{V}) = \boldsymbol{\rho_2^\varphi}(\widetilde{U}, \widetilde{V})$ for all $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$.

Furthermore, following ideas similar to those in [201], it can be proved that, in case $\theta \in (0,1]$, there exists a probability measure W_θ on $([0,1], \mathcal{B}_{[0,1]})$ such that

$$D_\theta^\varphi(\widetilde{U}, \widetilde{V}) = \sqrt{\int_{[0,1]\times[0,1]\times\mathbb{S}^{p-1}} \left[(\Pi\, \widetilde{U}_\alpha(\mathbf{u}))^{[\xi]} - (\Pi\, \widetilde{V}_\alpha(\mathbf{u}))^{[\xi]}\right]^2 d\lambda_p(\mathbf{u})\, d\varphi(\alpha)\, dW_\theta(\xi)}$$

where $(\Pi\, \widetilde{U}_\alpha(\mathbf{u}))^{[\xi]} = \xi \cdot \sup \Pi\, \widetilde{U}_\alpha(\mathbf{u}) + (1-\xi) \cdot \inf \Pi\, \widetilde{U}_\alpha(\mathbf{u})$ for all $\xi \in [0,1]$.

This leads to an alternative *interpretation of the weighting parameter* θ in terms of certain weighted averages of the squared Euclidean distances between the same convex linear combinations of the extreme values of $\Pi\, \widetilde{U}_\alpha(\mathbf{u})$ and $\Pi\, \widetilde{V}_\alpha(\mathbf{u})$ (i.e., for the same ξ) and can be helpful also in choosing θ.

As some interesting examples we find the following:

- D_1^φ can be equivalently expressed as

$$D_1^\varphi(\widetilde{U}, \widetilde{V}) = \left[\frac{1}{2}\int_{[0,1]\times\mathbb{S}^{p-1}} \left[\inf \Pi\, \widetilde{U}_\alpha(\mathbf{u}) - \inf \Pi\, \widetilde{V}_\alpha(\mathbf{u})\right]^2 d\lambda_p(\mathbf{u})\, d\varphi(\alpha) + \frac{1}{2}\int_{[0,1]\times\mathbb{S}^{p-1}} \left[\sup \Pi\, \widetilde{U}_\alpha(\mathbf{u}) - \sup \Pi\, \widetilde{V}_\alpha(\mathbf{u})\right]^2 d\lambda_p(\mathbf{u})\, d\varphi(\alpha)\right]^{1/2},$$

  that is, D_1^φ corresponds to the probability measure W_θ associated with the uniform distribution on $\{0,1\}$ (and hence taking only into account the infima and suprema of the projection intervals);

- $D_{1/3}^\varphi$ can be equivalently expressed as

$$D_{1/3}^\varphi(\widetilde{U}, \widetilde{V}) = \sqrt{\int_{[0,1]}\int_{[0,1]\times\mathbb{S}^{p-1}} \left[(\Pi\, \widetilde{U}_\alpha(\mathbf{u}))^{[\xi]} - (\Pi\, \widetilde{V}_\alpha(\mathbf{u}))^{[\xi]}\right]^2 d\lambda_p(\mathbf{u})\, d\varphi(\alpha)\, d\ell(\xi)},$$

  that is, $D_{1/3}^\varphi$ corresponds to the probability measure $W_\theta = \ell$ associated with the uniform distribution on $[0,1]$ (and hence taking into account and equally weighting all the points of the projection intervals).

The L^2 mid/spr-based metric fulfills several valuable metric properties, and it allows us to establish a Rådstrom-type isometry enabling us to identify each fuzzy set-valued datum with a functional data and to connect one-to-one the corresponding arithmetics and metrics. These properties can be found in detail in Trutschnig et al. [201] and González-Rodríguez *et al.* [90] (see also Blanco-Fernández *et al.* [19] and Gil *et al.* [83]).

Proposition 1.3.1. *Let $\theta \in (0, +\infty)$ and let φ be an absolutely continuous probability measure on $([0,1], \mathcal{B}_{[0,1]})$ with the mass function being positive in $(0,1)$. Let $\mathbb{H}_2 = \{L^2$-type real-valued functions defined on $[0,1] \times \mathbb{S}^{p-1}$ w.r.t. $\ell \otimes \lambda_p\}$. Then, the L^2 mid/spr-based metric satisfies that*

i) D_θ^φ *is an L^2-type metric on $\mathcal{F}_c(\mathbb{R}^p)$.*

ii) D_θ^φ *is translational and rotational invariant, i.e.,* $D_\theta^\varphi(\widetilde{U} + \widetilde{W}, \widetilde{V} + \widetilde{W}) = D_\theta^\varphi(\widetilde{U}, \widetilde{V})$ *and* $D_\theta^\varphi((-1) \cdot \widetilde{U}, (-1) \cdot \widetilde{V}) = D_\theta^\varphi(\widetilde{U}, \widetilde{V})$.

iii) D_θ^φ *is topologically equivalent to* $\boldsymbol{\rho}_2^\varphi$ *(and, hence, to* $\mathbf{d}_2^\varphi$*).*

iv) $(\mathcal{F}_c(\mathbb{R}^p), D_\theta^\varphi)$ *is a separable metric space.*

v) The support function $s : \mathcal{F}_c(\mathbb{R}^p) \to \mathbb{H}_2$ *(with* $s(\widetilde{U}) = s_{\widetilde{U}}$*) states an isometric embedding of* $\mathcal{F}_c(\mathbb{R}^p)$ *with the fuzzy arithmetic and* D_θ^φ *onto a convex cone of the Hilbert space* $\mathbb{H}_2$ *with the functional arithmetic and the distance induced by the norm*

$$\|h - h'\|_\theta^\varphi = \sqrt{\langle h - h', h - h' \rangle_\theta^\varphi},$$

with

$$\langle f, g \rangle_\theta^\varphi = \int_{[0,1]\times\mathbb{S}^{p-1}} \operatorname{mid} f(\alpha, \mathbf{u}) \cdot \operatorname{mid} g(\alpha, \mathbf{u}) \, d\lambda_p(\mathbf{u}) \, d\varphi(\alpha)$$
$$+\theta \int_{[0,1]\times\mathbb{S}^{p-1}} \operatorname{spr} f(\alpha, \mathbf{u}) \cdot \operatorname{spr} g(\alpha, \mathbf{u}) \, d\lambda_p(\mathbf{u}) \, d\varphi(\alpha)$$

and

$$\operatorname{mid} f(\alpha, \mathbf{u}) = \frac{f(\alpha, \mathbf{u}) - f(\alpha, -\mathbf{u})}{2}, \quad \operatorname{spr} f(\alpha, \mathbf{u}) = \frac{f(\alpha, \mathbf{u}) + f(\alpha, -\mathbf{u})}{2}.$$

Remark 1.3.2. An immediate and crucial implication from Proposition 1.3.1.*v*) is that any fuzzy set value $\widetilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$ can be identified with the corresponding function $s_{\widetilde{U}}$ and this identification is accompanied by the correspondences between the usual arithmetics and L^2 metrics. Consequently, data in the setting of fuzzy set-valued data with the fuzzy arithmetic and the metric D_θ^φ can be systematically translated into data in the setting of functional data with the functional arithmetic and the metric based on the associated norm. In this way, despite the fact that fuzzy data should not be treated directly as functional data, they can be treated as functional data by considering the identification *via* the support function.

Then, we can now assert formally as a relevant implication for statistical purposes that several developments in Functional Data Analysis could be particularized to fuzzy set-valued data by using the adequate identifications and correspondences. However, it should be guaranteed that the resulting elements/outputs remain in the cone $s\left(\mathcal{F}_c(\mathbb{R}^p)\right)$. In case either the functional developments become very complex or the resulting elements/outputs are out of $s\left(\mathcal{F}_c(\mathbb{R}^p)\right)$, *ad hoc* techniques should be developed, as we will show in the next chapters.

The mid/spr-based L^2 metric has been shown to be very suitable in the development of statistical methodology for experimental fuzzy set-valued data. For instance, González-Rodríguez *et al.* [90] provides a detailed explanation of an approach to the ANOVA with fuzzy data based on the functional data identification. The recent reviews of Blanco-Fernández *et al.* [18, 19, 20] and Gil *et al.* [83] summarize most of these statistical methods.

The particularization of the mid/spr-based L^2 metric to fuzzy numbers and to set and interval values is immediate and will be employed in some developments in Chapters 2 and 3.

Now we will introduce the second family extending $\boldsymbol{\rho}_{\mathbf{2}}^{\varphi}$ for fuzzy vector-valued data with a double purpose, namely, taking into account the influence of the deviation in 'center' and the influence of the deviation in 'shape' separately and being based on a representation for which there exist sufficient conditions characterizing fuzzy set values.

For the first family of distances, the center and shape have been substantiated through the mid and spr functions, respectively. As it has been already pointed out in Section 1.1.2, for the mid/spr representation there is not a set of sufficient conditions characterizing fuzzy set-valued data.

In this new family, we consider the Steiner point (and, consequently, the weighted averaging based on levels in case $p = 1$) as indicator of the center and the level-wise deviations of the levels w.r.t. the center (left and right deviations of the extreme points w.r.t. the wabl in case $p = 1$) as indicators of the shape. Therefore, the generalized metric will be based on the support/Steiner representation for which sufficient conditions have been stated (Proposition 1.1.5) to characterize fuzzy set-valued data. Thus,

Definition 1.3.5. *Given an absolutely continuous probability measure φ on the measurable space $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function on $(0,1)$ and a parameter $\theta \in (0,1]$, the* ***support/Steiner-based*** *L^2* ***metric*** *is the mapping $\mathfrak{D}^\varphi_\theta : \mathcal{F}_c(\mathbb{R}^p) \times \mathcal{F}_c(\mathbb{R}^p) \to [0,+\infty)$ such that for $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$:*

$$\mathfrak{D}^\varphi_\theta(\widetilde{U},\widetilde{V}) = \Bigg[(1-\theta)\|\mathbf{S}^\varphi(\widetilde{U}) - \mathbf{S}^\varphi(\widetilde{V})\|^2 + \theta \int_{[0,1]\times\mathbb{S}^{p-1}} \left[s_{\widetilde{U}}(\alpha,\mathbf{u}) - s_{\widetilde{V}}(\alpha,\mathbf{u})\right]^2 d\lambda_p(\mathbf{u})\, d\varphi(\alpha)\Bigg]^{1/2}$$

$$= \sqrt{\|\mathbf{S}^\varphi(\widetilde{U}) - \mathbf{S}^\varphi(\widetilde{V})\|^2 + \theta \int_{[0,1]\times\mathbb{S}^{p-1}} \|\mathbf{dev}^\varphi_{\widetilde{U}}(\alpha,\mathbf{u}) - \mathbf{dev}^\varphi_{\widetilde{V}}(\alpha,\mathbf{u})\|^2\, d\lambda_p(\mathbf{u})\, d\varphi(\alpha)},$$

where $\mathbf{dev}^\varphi_{\widetilde{U}}(\alpha,\mathbf{u}) = \mathbf{u}\cdot s_{\widetilde{U}}(\alpha,\mathbf{u}) - \mathbf{S}^\varphi(\widetilde{U})$.

The support/Steiner-based L^2 distance between elements in $\mathcal{K}_c(\mathbb{R}^p)$ will be denoted by $\mathfrak{d}_\theta$.

The support/Steiner-based L^2 metric can be also expressed in terms of the $\boldsymbol{\rho}^\varphi_{\mathbf{2}}$. More concretely,

$$\mathfrak{D}^\varphi_\theta(\widetilde{U},\widetilde{V}) = \sqrt{(1-\theta)\|\mathbf{S}^\varphi(\widetilde{U}) - \mathbf{S}^\varphi(\widetilde{V})\|^2 + \theta\left[\boldsymbol{\rho}^\varphi_{\mathbf{2}}(\widetilde{U},\widetilde{V})\right]^2}.$$

The mapping $\mathfrak{D}^\varphi_\theta$ is a distance between fuzzy vectors. In consequence (see Appendix for the proof),

Proposition 1.3.2. *$\left(\mathcal{F}_c(\mathbb{R}^p), \mathfrak{D}^\varphi_\theta\right)$ is a metric space.*

As for the mid/spr-based L^2 metric, the parameter θ and the measure φ do not have a stochastic meaning in the support/Steiner-based distance. The parameter θ weighs the influence of the 'deviation in shape' between the fuzzy vectors (quantified through $\mathbf{dev}^\varphi$) with respect to the influence of their 'deviation in center' (quantified through the generalized Steiner point $\mathbf{S}^\varphi$), while the choice of φ allows us to weigh the influence of each α-level (i.e., the different degrees of 'compatibility').

The support/Steiner-based metric is in fact an L^2 metric that allows us to embed the space of fuzzy set values into a convex cone of a Hilbert space through the support/Steiner representation. In this way, an inner product in $\mathbb{R}^p \times \mathbb{H}_2$ can be defined as follows:

Let $\theta \in (0,1]$ and let φ be a weighting measure formalized as an absolutely continuous probability measure on $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function in $(0,1)$. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$ and $f, g \in \mathbb{H}_2$, consider the inner product

$$\big\langle\big\langle(\mathbf{x},f),(\mathbf{y},g)\big\rangle\big\rangle^\varphi_\theta = \int_{[0,1]\times\mathbb{S}^{p-1}} \langle(\mathbf{x}, f(\alpha,\mathbf{u})),(\mathbf{y}, g(\alpha,\mathbf{u}))\rangle_\theta\, d\lambda_p(\mathbf{u})\, d\varphi(\alpha),$$

where the Euclidean inner product $\langle\cdot,\cdot\rangle_\theta$ on $\mathbb{R}^p \times \mathbb{R}$ is based on the weighted dot product given by

$$\langle(\mathbf{x}_1, x_2), (\mathbf{y}_1, y_2)\rangle_\theta = (1-\theta)\,\mathbf{x}_1 \cdot \mathbf{y}_1 + \theta\, x_2\, y_2.$$

Obviously, if $||\cdot||_\theta^\varphi$ denotes the norm associated with the inner product $\langle\langle\cdot,\cdot\rangle\rangle_\theta^\varphi$, we have that

$$\mathfrak{D}_\theta^\varphi(\widetilde{U},\widetilde{V}) = \left|\left|\boldsymbol{\tau}^\varphi(\widetilde{U}) - \boldsymbol{\tau}^\varphi(\widetilde{V})\right|\right|_\theta^\varphi = \sqrt{\langle\langle\boldsymbol{\tau}^\varphi(\widetilde{U}) - \boldsymbol{\tau}^\varphi(\widetilde{V}), \boldsymbol{\tau}^\varphi(\widetilde{U}) - \boldsymbol{\tau}^\varphi(\widetilde{V})\rangle\rangle_\theta^\varphi}.$$

Then, due to the properties of the generalized Steiner points and the support functions (see, for instance, Butnariu *et al.* [22], Vetterlein and Navara [209, 210] and Liang *et al.* [122]), the following result holds:

Proposition 1.3.3. *Let φ be an arbitrarily fixed absolutely continuous probability measure on $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function on $(0,1)$ and $\theta \in (0,1]$ be a weighting parameter. Then,*

i) $\langle\langle\cdot,\cdot\rangle\rangle_\theta^\varphi$ *is an inner product in* $\mathbb{R}^p \times \mathbb{H}_2$.

ii) $\left(\mathbb{R}^p \times \mathbb{H}_2, \langle\langle\cdot,\cdot\rangle\rangle_\theta^\varphi\right)$ *is a Hilbert space.*

iii) $\mathfrak{D}_\theta^\varphi$ *is an L^2-type metric and it is translational and rotational invariant.*

iv) For a fixed φ, the function $\boldsymbol{\tau}^\varphi : \mathcal{F}_c(\mathbb{R}^p) \to \mathbb{R}^p \times \mathbb{H}_2$ with $\boldsymbol{\tau}^\varphi(\widetilde{U}) = \boldsymbol{\tau}_{\widetilde{U}}^\varphi$ for all $\widetilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$ satisfies that

- $\boldsymbol{\tau}^\varphi$ *is an isometry from* $(\mathcal{F}_c(\mathbb{R}^p), \mathfrak{D}_\theta^\varphi)$ *into* $\left(\mathbb{R}^p \times \mathbb{H}_2, \langle\langle\cdot,\cdot\rangle\rangle_\theta^\varphi\right)$,
- $\boldsymbol{\tau}^\varphi(\widetilde{U} + \widetilde{V}) = \boldsymbol{\tau}^\varphi(\widetilde{U}) + \boldsymbol{\tau}^\varphi(\widetilde{V})$ *for all* $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$,
- $\boldsymbol{\tau}^\varphi(\gamma \cdot \widetilde{U}) = \gamma \cdot \boldsymbol{\tau}^\varphi(\widetilde{U})$ *for all* $\widetilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$ *and* $\gamma > 0$.

Consequently, the $\boldsymbol{\tau}^\varphi$ function preserves the semilinearity of $\mathcal{F}_c(\mathbb{R}^p)$ and relates the fuzzy arithmetic to the vectorial-valued functional arithmetic, which implies that $\mathcal{F}_c(\mathbb{R}^p)$ can be isometrically embedded into a convex cone of the Hilbert space $\left(\mathbb{R}^p \times \mathbb{H}_2, \langle\langle\cdot,\cdot\rangle\rangle_\theta^\varphi\right)$.

The metric space $(\mathcal{F}_c(\mathbb{R}^p), \mathfrak{D}_\theta^\varphi)$ is separable. This assertion is justified by the fact that $\mathfrak{D}_\theta^\varphi$ is topologically equivalent (in fact, strongly equivalent) to the metric $\boldsymbol{\rho}_2^\varphi$ and, hence, to $\boldsymbol{d}_2^\varphi$. Specifically (see Appendix for the proof),

Proposition 1.3.4. *Let $\theta \in (0,1]$ be a weight parameter and let φ be an arbitrarily fixed absolutely continuous probability measure on $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function in $(0,1)$. The metric $\mathfrak{D}_\theta^\varphi$ is topologically equivalent to the metric $\boldsymbol{\rho}_2^\varphi$ on $\mathcal{F}_c(\mathbb{R}^p)$. More precisely,*

$$\sqrt{\theta} \cdot \boldsymbol{\rho}_2^\varphi(\widetilde{U},\widetilde{V}) \le \mathfrak{D}_\theta^\varphi(\widetilde{U},\widetilde{V}) \le \boldsymbol{\rho}_2^\varphi(\widetilde{U},\widetilde{V})$$

for all $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$.

The practical computation of $\mathfrak{D}_\theta^\varphi$ in the case $p = 1$ is quite simple and easy to implement in any programming language. This computation will be illustrated later by means of some examples and simulations.

Unlike the one-dimensional case, when $p > 1$ the situation usually becomes much more complex. This is a consequence of the computational difficulties involved in determining most of the support functions of fuzzy vectors in practice (see Ghosh and Kumar [82] for some details about the set-valued case).

As an example illustrating the use of $\mathfrak{D}_\theta^\varphi$ in one of the simplest fuzzy vector-valued situations, we are now going to compute the distance between two conical fuzzy vectors arisen in an approach to the tone and color triangle designs (see Sugano [191, 192, 193] and Sugano *et al.* [194]).

Example 1.3.1. Sugano and collaborators have developed studies on a system of the three primary colors RGB presented on a color triangle. The usual triangle involves sixty-six fuzzy inputs (fundamental type) on parts of the tone triangle designated as darkness-blackness, lightness-whiteness and chromaticness.

The main (fundamental type) color names and modifiers are No. 1: darkest, No. 11: lightest, and No. 66: maximum chromaticness in the tone triangle and No. 1: blue, No. 11: green, and No. 66: red in the color triangle.

The fuzzy inputs are formed by right circular conical fuzzy vectors and they can mutually overlap (actually, any fuzzy input overlaps with some other ones). Detailed arguments leading to the conical fuzzy inputs can be found in, among others, Sugano [191, 192, 193] and Sugano *et al.* [194].

Figure 1.7 displays the general scheme, but only three fuzzy inputs have been fully represented, namely:

- No. 1, corresponding to the right cone $\widetilde{C}^1$ such that for each $\alpha \in [0, 1]$

$$\widetilde{C}^1_\alpha = \text{circle with centre } (0,0) \text{ and radius } 10(1-\alpha),$$

- No. 36, corresponding to the right cone $\widetilde{C}^{36}$ such that for each $\alpha \in [0, 1]$

$$\widetilde{C}^{36}_\alpha = \text{circle with centre } (15\sqrt{3}, 65) \text{ and radius } 10(1-\alpha),$$

- No. 58, corresponding to the right cone $\widetilde{C}^{58}$ such that for each $\alpha \in [0, 1]$

$$\widetilde{C}^{58}_\alpha = \text{circle with centre } (35\sqrt{3}, 45) \text{ and radius } 10(1-\alpha).$$

For the rest of the inputs only their 0.5-levels have been displayed graphically.

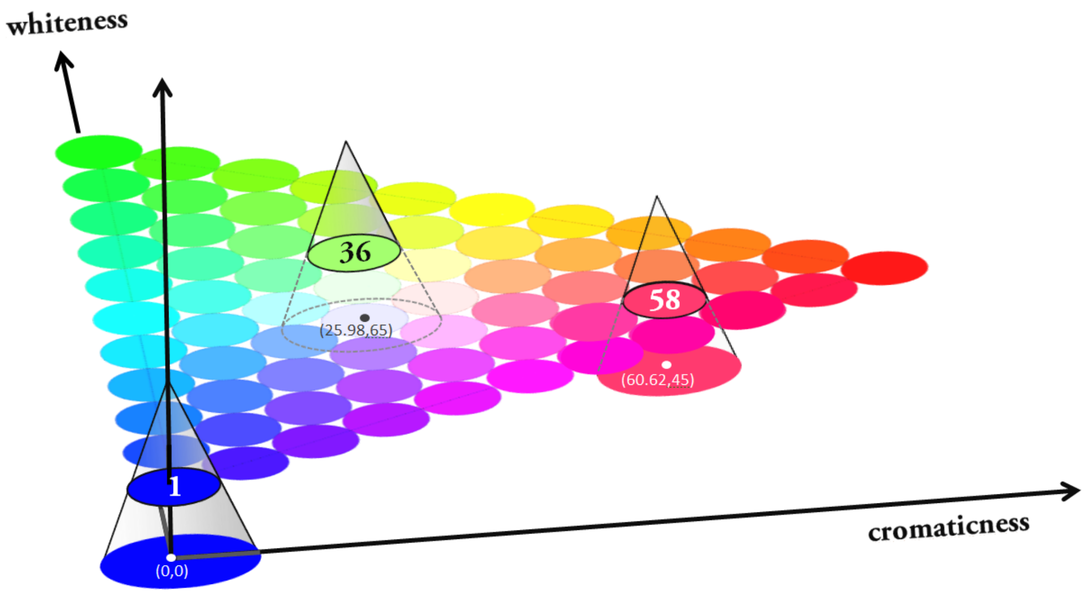


Figure 1.7: Conical fuzzy vectors in the usual color triangle (fundamental type)

The support function associated with a right circular conical fuzzy vector $\widetilde{C}$ such that, for each $\alpha \in [0,1]$, the α-level corresponds to $\widetilde{C}_\alpha =$ circle with centre (x,y) and radius $r(1-\alpha)$ can be calculated using some trigonometric results and is given for $\mathbf{u} = (\cos\beta_{\mathbf{u}}, \sin\beta_{\mathbf{u}}) \in \mathbb{S}^1 =$ circumference with centre $(0,0)$ and radius 1 by

$$s_{\widetilde{C}}(\alpha,\mathbf{u}) = x\cdot\cos\beta_{\mathbf{u}} + y\cdot\sin\beta_{\mathbf{u}} + r(1-\alpha).$$

Consequently, one can obtain the following distances using the already seen expression

$$\left[\mathfrak{D}_\theta^\varphi(\widetilde{U},\widetilde{V})\right]^2 = (1-\theta)\|\mathbf{S}^\varphi(\widetilde{U}) - \mathbf{S}^\varphi(\widetilde{V})\|^2 + \theta\left[\boldsymbol{\rho}_{\mathbf{2}}^\varphi(\widetilde{U},\widetilde{V})\right]^2.$$

Note that in this case, because of the 'shape' of the α-levels of the three conical fuzzy vectors coinciding, they are irrespective of the chosen φ:

$$\left[\mathfrak{D}_\theta^\varphi(\widetilde{C}^1,\widetilde{C}^{36})\right]^2 = (1-\theta)\,\|(0,0) - (15\sqrt{3},65)\|^2$$

$$+\frac{\theta}{2\pi}\int_{[0,1]\times[0,2\pi)}[10(1-\alpha) - 15\sqrt{3}\cos\beta - 65\sin\beta - 10(1-\alpha)]^2\,d\beta\,d\varphi(\alpha)$$

$$= 4900(1-\theta),$$

$$\left[\mathfrak{D}_\theta^\varphi(\widetilde{C}^1,\widetilde{C}^{58})\right]^2 = (1-\theta)\,\|(0,0) - (35\sqrt{3},45)\|^2$$

$$+\frac{\theta}{2\pi}\int_{[0,1]\times[0,2\pi)}[10(1-\alpha) - 35\sqrt{3}\cos\beta - 45\sin\beta - 10(1-\alpha)]^2\,d\beta\,d\varphi(\alpha)$$

$$= 5700(1-\theta),$$

$$\left[\mathfrak{D}_\theta^\varphi(\widetilde{C}^{36},\widetilde{C}^{58})\right]^2 = (1-\theta)\,\|(15\sqrt{3},65)-(35\sqrt{3},45)\|^2$$
$$+\frac{\theta}{2\pi}\int_{[0,1]\times[0,2\pi)}[15\sqrt{3}\cos\beta+65\sin\beta+10(1-\alpha)$$
$$-35\sqrt{3}\cos\beta-45\sin\beta-10(1-\alpha)]^2\,d\beta\,d\varphi(\alpha)=1600(1-\theta).$$

Therefore, and whatever the weight θ may be, we can conclude that the fuzzy inputs No. 36 and No. 58 are closer than No. 1 and No. 36 and than No. 1 and No. 58.

In summary, the support/Steiner-based L^2 metrics represent a parameterized family of topologically equivalent metrics to the distance $\boldsymbol{\rho}_{\mathbf{2}}^{\varphi}$ (based on the 2-norm and making use of the support function representation of the fuzzy value). Consequently, they share all the topological advantages of $\boldsymbol{\rho}_{\mathbf{2}}^{\varphi}$, but they also allow us to control the relative influence of the center and the shape of the fuzzy values, whereas $\boldsymbol{\rho}_{\mathbf{2}}^{\varphi}$ does not. This double control was also allowed by Trutschnig *et al.*'s metric, but, unfortunately, there is not a set of sufficient conditions for the mid/spr representation to characterize fuzzy set-valued data, what becomes a rather serious drawback in many optimization problems.

A key question that can arise when employing the distance $\mathfrak{D}_\theta^\varphi$ in real-life problems involving fuzzy data is the selection of a particular element of the family of metrics. For this purpose, we will interpret the roles played by φ and θ in the metric in a deeper way.

The probability measure φ can be formally identified with a measure weighting the 'importance' given to the different α-levels of the fuzzy set-valued data. For instance,

- the choice $\varphi \equiv \ell$ indicates that one gives the same relevance to all levels in quantifying the distance between fuzzy set values;
- choosing φ such that the greater the value of α, the greater its weight (e.g. $\varphi \equiv \text{Beta}(p,1)$ with $p >> 1$) indicates that one gives higher relevance to high levels in quantifying the distance between fuzzy values, that is, one mainly focuses on the levels with high degree of compatibility;
- choosing φ such that the greater the value of α, the lower its weight (e.g. $\varphi \equiv \text{Beta}(1,p)$ with $p >> 1$) indicates that one gives higher relevance to low levels in quantifying the distance between fuzzy values, that is, one mainly focuses on the levels with low degree of compatibility.

We illustrate these assertions with the following example in case $p = 1$.

Example 1.3.2. Consider the four couples of triangular fuzzy numbers in Figure 1.8. Note that:

- $\widetilde{U}$ and $\widetilde{V}$ share the 0-level -the interval $[0, 2]$-, but differ strongly in the shape, being 'closer' at lower levels;
- $\widetilde{U}'$ and $\widetilde{V}'$ have 0-levels -the intervals $[0, 1]$ and $[1, 2]$- which only overlap in a singleton and also show strongly different shapes, being also 'closer' at lower levels;
- $\widetilde{U}''$ and $\widetilde{V}''$ have 0-levels -the intervals $[0, 1]$ and $[1, 2]$- which only overlap in a singleton, but show the same shape and they are both symmetric;
- $\widetilde{U}'''$ and $\widetilde{V}'''$ have 0-levels -the intervals $[-2, 0]$ and $[0, 2]$- which only overlap in a singleton, differ very strongly in the shape and they are 'closer' at higher levels.

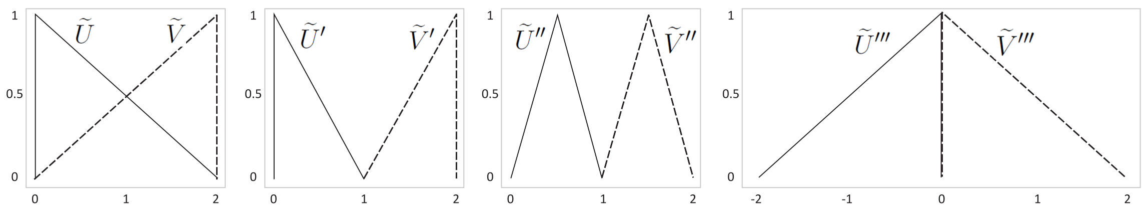


Figure 1.8: Different couples of fuzzy numbers

Figure 1.9 displays graphically the distance $\mathfrak{D}^{\varphi}_{1/3}$ between the fuzzy numbers $\widetilde{U}$ and $\widetilde{V}$, $\widetilde{U}'$ and $\widetilde{V}'$, $\widetilde{U}''$ and $\widetilde{V}''$ and $\widetilde{U}'''$ and $\widetilde{V}'''$ as a function of $p \in (0, \infty)$ when φ is taken as $\beta(p, 1)$ and $\beta(1, p)$, respectively (note that the situation $p = 1$ corresponds to $\varphi \equiv \ell$).

The conclusions from Figure 1.9 are not unexpected looking at the center and shape of the considered fuzzy numbers in Figure 1.8. First note that, due to the symmetry of the corresponding fuzzy numbers, $\mathfrak{D}^{\varphi}_{1/3}(\widetilde{U}'', \widetilde{V}'')$ is constantly equal to 1, independently from the choice of φ. On the other hand, $\mathfrak{D}^{\varphi}_{1/3}(\widetilde{U}, \widetilde{V})$ and $\mathfrak{D}^{\varphi}_{1/3}(\widetilde{U}', \widetilde{V}')$ increase as the weight of the high levels increases, and decreases otherwise, whereas $\mathfrak{D}^{\varphi}_{1/3}(\widetilde{U}''', \widetilde{V}''')$ decreases as the weight to the low levels increases, and increases otherwise.

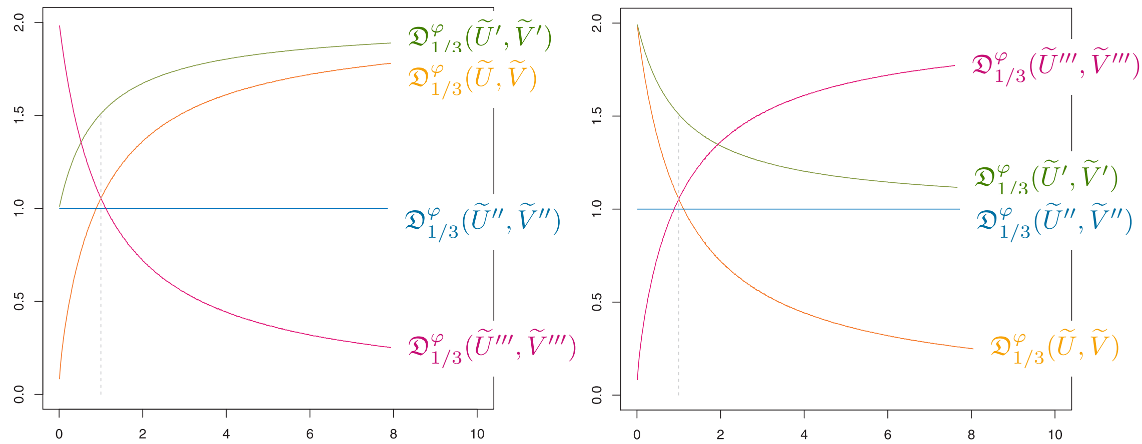


Figure 1.9: Distance $\mathfrak{D}^{\varphi}_{1/3}$ between the couples of fuzzy numbers in Fig. 1.8 for $\varphi \equiv \beta(p, 1)$ (left) and $\beta(1, p)$ (right)

As explained in introducing the support/Steiner-based L^2 metric, the general role of the parameter θ is to weigh the influence of the squared deviation in center (measured by $\mathbf{S}^{\varphi}$ or wabl$^{\varphi}$) of the fuzzy values in contrast to the influence of their squared deviation in shape (measured level-wise by $\mathbf{dev}^{\varphi}$ or ldev$^{\varphi}$ and rdev$^{\varphi}$ functions). However, although the exact meaning of the value θ takes is not clear yet, it can be identified intuitively in the situation $p = 1$. First, we will particularize this metric to the fuzzy number-valued case.

When $p = 1$, the metric $\mathfrak{D}^{\varphi}_{\theta}$ reduces to a distance based on the wabl/ldev/rdev representation of the involved fuzzy numbers. In this way (see Appendix for the proof),

Proposition 1.3.5. *If* $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R})$, *then*

$$\mathfrak{D}^{\varphi}_{\theta}(\widetilde{U}, \widetilde{V}) = \Bigg[\Big(\text{wabl}^{\varphi}(\widetilde{U}) - \text{wabl}^{\varphi}(\widetilde{V})\Big)^2$$

$$+ \frac{\theta}{2} \int_{[0,1]} \Big[\text{ldev}^{\varphi}_{\widetilde{U}}(\alpha) - \text{ldev}^{\varphi}_{\widetilde{V}}(\alpha)\Big]^2 d\varphi(\alpha) + \frac{\theta}{2} \int_{[0,1]} \Big[\text{rdev}^{\varphi}_{\widetilde{U}}(\alpha) - \text{rdev}^{\varphi}_{\widetilde{V}}(\alpha)\Big]^2 d\varphi(\alpha)\Bigg]^{1/2}.$$

Remark 1.3.3. It should be noticed that if the involved fuzzy numbers are symmetric (and, in particular, interval-valued), the distances $\mathfrak{D}^{\varphi}_{\theta}$ and D^{φ}_{θ} between them coincide.

A result can be established now to allow us to interpret better the role of the value of θ when $p = 1$. For this purpose, for an arbitrarily given fuzzy number $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ and level $\alpha \in [0, 1]$, each real value

$$x \in \left[\min\{\text{wabl}^\varphi(\widetilde{U}), \inf \widetilde{U}_\alpha\}, \max\{\text{wabl}^\varphi(\widetilde{U}), \sup \widetilde{U}_\alpha\}\right]$$

can be written as a particular linear combination of the components of the wabl/ ldev/rdev representation. More precisely, for any of these x's there exists a $\xi \in [-1, 1]$ such that

$$x = f^\varphi_{\widetilde{U}}(\alpha, \xi) = \text{wabl}^\varphi(\widetilde{U}) - M_0(-\xi) \cdot \text{ldev}^\varphi_{\widetilde{U}}(\alpha) + M_0(\xi) \cdot \text{rdev}^\varphi_{\widetilde{U}}(\alpha),$$

with $M_0(\xi) = \max\{0, \xi\}$.

Note that $\left[\min\{\text{wabl}^\varphi(\widetilde{U}), \inf \widetilde{U}_\alpha\}, \max\{\text{wabl}^\varphi(\widetilde{U}), \sup \widetilde{U}_\alpha\}\right]$ does not always represent the level $\widetilde{U}_\alpha$, but in case $\text{wabl}^\varphi(\widetilde{U}) \notin \widetilde{U}_\alpha$, the α-level interval is 'enlarged' to include $\text{wabl}^\varphi(\widetilde{U})$.

Based on the above expression, for any two fuzzy numbers $\widetilde{U}$ and $\widetilde{V}$, a one-to-one correspondence between them can be stated by considering the functions $f^\varphi_{\widetilde{U}}(\alpha, \xi)$ and $f^\varphi_{\widetilde{V}}(\alpha, \xi)$, so that it seems plausible to consider the distance between $\widetilde{U}$ and $\widetilde{V}$ as given by

$$\mathcal{D}^\varphi_\vartheta(\widetilde{U}, \widetilde{V}) = \sqrt{\int_{[0,1]} \int_{[-1,1]} \left[f^\varphi_{\widetilde{U}}(\alpha, \xi) - f^\varphi_{\widetilde{V}}(\alpha, \xi)\right]^2 d\vartheta(\xi)\, d\varphi(\alpha)},$$

where ϑ is a measure which can be identified formally with a symmetric and non-degenerate probability measure on $\big([-1, 1], \mathcal{B}_{[-1,1]}\big)$.

Since ϑ is assumed to be symmetric on $[-1, 1]$, it can be expressed as a finite mixture $\vartheta = .5 \cdot \zeta + .5 \cdot \nu$, where ν is a (non-degenerate at 0) probability measure on $[0, 1]$ and $\zeta(\xi) = \nu(-\xi)$. Therefore, the distance $\mathcal{D}^\varphi_\vartheta(\widetilde{U}, \widetilde{V})$ can be rewritten by taking into account that

$$\int_{[-1,1]} \left[f^\varphi_{\widetilde{U}}(\alpha, \xi) - f^\varphi_{\widetilde{V}}(\alpha, \xi)\right]^2 d\vartheta(\xi)$$

$$= \frac{1}{2} \int_{[0,1]} \left[\Big(\text{wabl}^\varphi(\widetilde{U}) - \xi \cdot \text{ldev}^\varphi_{\widetilde{U}}(\alpha)\Big) - \Big(\text{wabl}^\varphi(\widetilde{V}) - \xi \cdot \text{ldev}^\varphi_{\widetilde{V}}(\alpha)\Big)\right]^2 d\nu(\xi)$$

$$+ \frac{1}{2} \int_{[0,1]} \left[\Big(\text{wabl}^\varphi(\widetilde{U}) + \xi \cdot \text{rdev}^\varphi_{\widetilde{U}}(\alpha)\Big) - \Big(\text{wabl}^\varphi(\widetilde{V}) + \xi \cdot \text{rdev}^\varphi_{\widetilde{V}}(\alpha)\Big)\right]^2 d\nu(\xi).$$

The next result shows that this distance is an equivalent definition for $\mathfrak{D}^\varphi_\theta$. Based on this equivalence, the role of the parameter θ becomes easier to interpret (see Appendix for the proof).

Proposition 1.3.6. *The family of metrics* $\mathcal{D}^\varphi_\vartheta$ *is equivalent to the family of metrics* $\mathfrak{D}^\varphi_\theta$.

As an immediate implication from the preceding proposition we can interpret some choices of the value of θ that will be useful for practical purposes. Among the most relevant metrics, we can highlight:

- the choice of $\theta = 1/3$ (considered in Example 1.3.2) corresponds to choosing ν as the Lebesgue measure ℓ on $[0, 1]$ (i.e. the points in each 'enlarged' level being weighted uniformly);
- the choice of $\theta = 1$ corresponds, among others, to choosing ν as the indicator function of $\{1\}$, so that as already commented $\mathfrak{D}_1^{\varphi} = \boldsymbol{\rho}_{\mathbf{2}}^{\varphi}$.

In summary, another advantage of this family of metrics is that the choice of the parameter θ can be interpreted in a similar way to that of Bertoluzza *et al.*'s metric, what can make its choice easier in practical problems. Nevertheless, the use of the parameterized version simplifies many practical and theoretical developments.

To illustrate the computation of the distance $\mathfrak{D}_\theta^{\varphi}$ when $p = 1$ and interpret the role of θ, we make use of another example.

Example 1.3.3. Setnes *et al.* [169] considered the modeling of a real-world system, a pressure-controlled fermenter tank, by means of the FAIR (Fuzzy Arithmetic-based Interpolative Reasoning) method. A model with fuzzy outputs was identified by least squares approximation, where the identification data consist of fuzzy observations.

One of the variables that should be carefully controlled during a fermentation process is the pressure in the fermenter tank. The inference in FAIR consists of a set of fuzzy rules, each of which describes a local relation between the inputs (or antecedents) and the outputs (or consequents). By using previous knowledge about the process, the antecedents are partitioned into four 'fuzzy regions', associated with the triangular values $\widetilde{U}_1 = \text{Tri}(1, 1, 1.2)$, $\widetilde{U}_2 = \text{Tri}(1, 1.2, 1.65)$, $\widetilde{U}_3 = \text{Tri}(1.2, 1.65, 2.2)$ and $\widetilde{U}_4 = \text{Tri}(1.65, 2.2, 2.2)$.

In case there is no knowledge the researchers can use to determine the antecedent fuzzy partition, data-driven approaches, like fuzzy clustering, can be applied. Given the antecedent partition above, the fuzzy numbers in the consequents are estimated from the identification data (see Setnes *et al.* [169] for more details). One of the fuzzy consequents obtained following this process corresponds to $\widetilde{V} = \text{Tri}(1.27, 1.476, 1.673)$, which is definitely 'in between' $\widetilde{U}_2$ and $\widetilde{U}_3$.

Figure 1.10 displays the fuzzy partition of the antecedents (in grey lines) and the examined consequent (in pink line).

Assume that, for purposes of classification or others, we want to look for the closest antecedent to $\widetilde{V}$ in the considered particion.

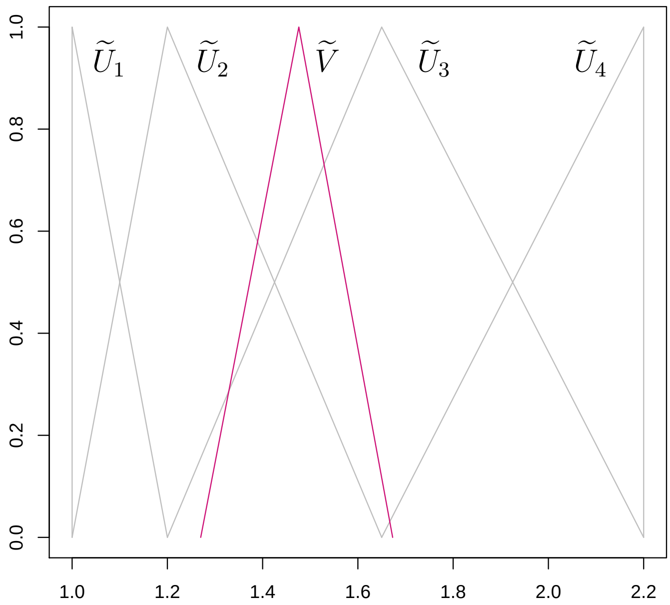


Figure 1.10: Fuzzy partition of the antecedents and consequent $\widetilde{V}$ in Example 1.3.3

Then, by considering the metric $\mathfrak{D}_\theta^\ell$ we obtain that:

$$\mathfrak{D}_\theta^\ell(\widetilde{U}_2, \widetilde{V}) = \sqrt{0.0446 + 0.0064\,\theta}, \quad \mathfrak{D}_\theta^\ell(\widetilde{U}_3, \widetilde{V}) = \sqrt{0.0405 + 0.0299\,\theta}.$$

This situation shows that depending on the choice of θ we will classify $\widetilde{U}_2$ as closer to $\widetilde{V}$ than $\widetilde{U}_3$. More concretely, if $\theta \leq 0.1744$ then $\widetilde{U}_3$ is closer to $\widetilde{V}$ than $\widetilde{U}_2$, otherwise the situation is the contrary (see Figure 1.11).

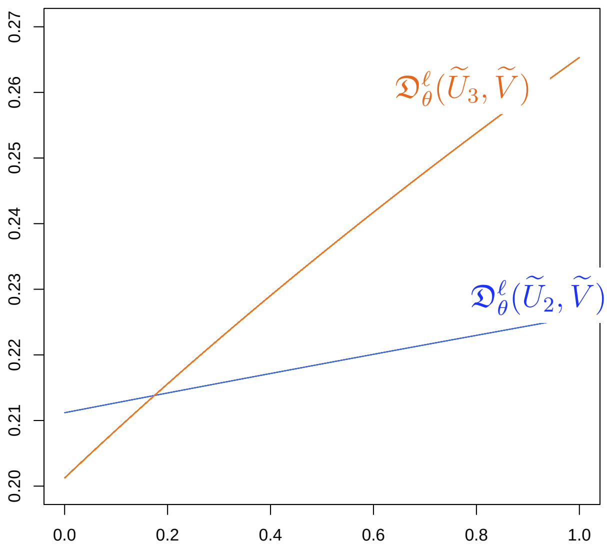


Figure 1.11: Distance $\mathfrak{D}_\theta^\ell$ between the couples of fuzzy numbers in Fig. 1.10, $(\widetilde{U}_2, \widetilde{V})$ and $(\widetilde{U}_3, \widetilde{V})$, as a function of θ

In short, when the metric assesses much higher influence to the 'center' than to the 'shape' (i.e., θ takes on very low values) $\widetilde{V}$ is slightly closer to $\widetilde{U}_3$, but the smaller the relative importance of the center in contrast to that of the shape, the closer $\widetilde{V}$ to $\widetilde{U}_2$.

In Section 1.5 it will be shown that metrics D_θ^φ and $\mathfrak{D}_\theta^\varphi$ lead to close outputs for many interesting cases, so the main advantage of the support/Steiner-based L^2 metric in contrast to the mid/spr-based distance is the highlighted fact that a set of sufficient conditions exist for the second one to characterize imprecise-valued data.

1.3.2 L^1 metrics for imprecise data

As for the L^2-type metrics, some rather immediate L^1 metrics to consider for imprecise values are the ones defined by Vitale [211] for set values and their level-wise extension for fuzzy set values by Diamond and Kloeden [54].

Definition 1.3.6. [211] *Let $K, K' \in \mathcal{K}_c(\mathbb{R}^p)$. The* $\boldsymbol{\delta_1}$ ***distance*** *between K and K' is given by*

$$\delta_1(K, K') = \int_{\mathbb{S}^{p-1}} \|s_K(\mathbf{u}) - s_{K'}(\mathbf{u})\| \, d\lambda_p(\mathbf{u}).$$

If $p = 1$, then $\delta_1(K, K') = |\inf K - \inf K'|/2 + |\sup K - \sup K'|/2$.

δ_1 and d_H induce the same topology on $\mathcal{K}_c(\mathbb{R}^p)$ and yield separable metric spaces, as proven in Vitale [211].

Definition 1.3.7. [54] *Let $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$. The* $\boldsymbol{\rho_1}$ ***distance*** *between $\widetilde{U}$ and $\widetilde{V}$ is given by*

$$\boldsymbol{\rho_1}(\widetilde{U}, \widetilde{V}) = \int_{[0,1]\times\mathbb{S}^{p-1}} \|s_{\widetilde{U}}(\alpha, \mathbf{u}) - s_{\widetilde{V}}(\alpha, \mathbf{u})\| \, d\lambda_p(\mathbf{u}) \, d\ell(\alpha).$$

The metric $\boldsymbol{\rho_1}$ can be 'corrected' by weighting the relevance of the different levels. Let φ be an absolutely continuous probability measure on $([0,1], \mathcal{B}_{[0,1]})$ with the mass function being positive in $(0,1)$. The $\boldsymbol{\rho_1^\varphi}$ distance between $\widetilde{U}$ and $\widetilde{V}$ is given by

$$\boldsymbol{\rho_1^\varphi}(\widetilde{U}, \widetilde{V}) = \int_{[0,1]\times\mathbb{S}^{p-1}} \|s_{\widetilde{U}}(\alpha, \mathbf{u}) - s_{\widetilde{V}}(\alpha, \mathbf{u})\| \, d\lambda_p(\mathbf{u}) \, d\varphi(\alpha).$$

Diamond and Kloeden [54] proved that $\boldsymbol{\rho_1}$ is topologically equivalent on $\mathcal{F}_c(\mathbb{R}^p)$ to $\boldsymbol{d_1}$, which is given (see Klement *et al.* [116]) by

$$\boldsymbol{d_1}(\widetilde{U}, \widetilde{V}) = \int_{[0,1]} d_H(\widetilde{U}_\alpha, \widetilde{V}_\alpha) \, d\ell(\alpha)$$

and both yield separable metric spaces. Analogously, one can easily prove that $\boldsymbol{\rho}_{\mathbf{1}}^{\varphi}$ is topologically equivalent to $\boldsymbol{d}_{\mathbf{1}}^{\varphi}(\widetilde{U},\widetilde{V}) = \int_{[0,1]} d_H(\widetilde{U}_\alpha, \widetilde{U}_\alpha)\, d\varphi(\alpha)$ and that both yield separable metric spaces.

By following the ideas in Rådström [159] on $\mathcal{K}_c(\mathbb{R}^p)$ and in Puri and Ralescu [154] on $\mathcal{F}_c(\mathbb{R}^p)$, the support functions of elements in these spaces allow us to embed isometrically each of these spaces into a convex cone of a Banach space of functions (more concretely, the space of the L^1-type real-valued functions on $\mathbb{S}^{p-1}$ and $[0,1] \times \mathbb{S}^{p-1}$, respectively, with the metrics λ_p and $\lambda_p \otimes \ell$).

In particular, when $p = 1$ we have that

Proposition 1.3.7. *Let φ be an arbitrarily fixed absolutely continuous probability measure on $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function on $(0,1)$. Then,*

i) *$\boldsymbol{\rho}_{\mathbf{1}}^{\varphi}$ is an L^1 metric on $\mathcal{F}_c(\mathbb{R})$ and it is translational and rotational invariant.*

ii) *For a fixed φ, the function $\boldsymbol{\iota}^{\varphi} : \mathcal{F}_c(\mathbb{R}) \to \mathbb{H}_1 = \{L^1$-type 2-dimensional vector-valued functions defined on $[0,1]\}$ satisfies that*

- *$\boldsymbol{\iota}^{\varphi}$ is an isometry from $(\mathcal{F}_c(\mathbb{R}), \boldsymbol{\rho}_{\mathbf{1}}^{\varphi})$ into $\mathbb{H}_1$,*
- *$\boldsymbol{\iota}^{\varphi}(\widetilde{U} + \widetilde{V}) = \boldsymbol{\iota}^{\varphi}(\widetilde{U}) + \boldsymbol{\iota}^{\varphi}(\widetilde{V})$ for all $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R})$,*
- *$\boldsymbol{\iota}^{\varphi}(\gamma \cdot \widetilde{U}) = \gamma \cdot \boldsymbol{\iota}^{\varphi}(\widetilde{U})$ for all $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ and $\gamma > 0$;*

Consequently, the $\boldsymbol{\iota}^{\varphi}$ function preserves the semilinearity of $\mathcal{F}_c(\mathbb{R})$ and relates the fuzzy arithmetic to the functional arithmetic, what implies that $\mathcal{F}_c(\mathbb{R})$ can be isometrically embedded into a convex cone of the Banach space $\left(\mathbb{H}_1, \|\cdot\|_1^{\varphi}\right)$ with

$$\|f - g\|_1^{\varphi} = \int_{[0,1]} \left(\frac{1}{2}|f_1(\alpha) - g_1(\alpha)| + \frac{1}{2}|f_2(\alpha) - g_2(\alpha)| \right) d\varphi(\alpha)$$

for $f = (f_1, f_2), g = (g_1, g_2) \in \mathbb{H}_1$.

The metrics δ_1^{φ} and $\boldsymbol{\rho}_{\mathbf{1}}^{\varphi}$ can be complemented (in this case, it does not mean an extension) by considering families of L^1 metrics paying separate attention to the 'center' and the 'shape' of the involved values.

It should be emphasized that the L^1 metrics between imprecise values will be involved in this work to develop *ad hoc* approaches in extending the median of imprecise-valued random elements.

At this point, on one hand, since under some usual conventions the median of a real-valued random variable preserves monotonicity and continuity and is equivariant under the product by a scalar, but nothing can be said in general in connection with additivity/subadditivity, we will constrain the use of the L^1 metrics in this setting to the case $p = 1$. On the other hand, the fact of having a set of sufficient conditions to characterize imprecise values becomes crucial, so the family of metrics based on the wabl/ldev/rdev representation will be considered in this respect.

Definition 1.3.8. *Given an absolutely continuous probability measure φ on the measurable space $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function on $(0,1)$ and a parameter $\theta \in (0,1]$, the* ***wabl/ldev/rdev-based*** L^1 ***metric*** *is the mapping $\mathscr{D}_\theta^\varphi : \mathcal{F}_c(\mathbb{R}) \times \mathcal{F}_c(\mathbb{R}) \to [0,+\infty)$ such that for $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R})$:*

$$\mathscr{D}_\theta^\varphi(\widetilde{U},\widetilde{V}) = |\text{wabl}^\varphi(\widetilde{U}) - \text{wabl}^\varphi(\widetilde{V})|$$
$$+\frac{\theta}{2}\int_{[0,1]} |\text{ldev}_{\widetilde{U}}^\varphi(\alpha) - \text{ldev}_{\widetilde{V}}^\varphi(\alpha)|\, d\varphi(\alpha) + \frac{\theta}{2}\int_{[0,1]} |\text{rdev}_{\widetilde{U}}^\varphi(\alpha) - \text{rdev}_{\widetilde{V}}^\varphi(\alpha)|\, d\varphi(\alpha).$$

The wabl/ldev/rdev-based L^1 distance between elements in $\mathcal{K}_c(\mathbb{R})$ is denoted by $\mathscr{d}_\theta$ (with $\mathscr{d}_1 = d_H$ on $\mathcal{K}_c(\mathbb{R})$).

The metric $\mathscr{D}_\theta^\varphi$ can be also expressed as follows:

$$\mathscr{D}_\theta^\varphi(\widetilde{U},\widetilde{V}) = \int_{[0,1]} |\boldsymbol{v}_{\widetilde{U}}^\varphi(\alpha) - \boldsymbol{v}_{\widetilde{V}}^\varphi(\alpha)|_\theta^1\, d\varphi(\alpha),$$

where $|\cdot|_\theta^1$ is the L^1 norm in $\mathbb{R}^3$ given for $\mathbf{x} = (x_1, x_2, x_3), \mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ by

$$|\mathbf{x} - \mathbf{y}|_\theta^1 = |x_1 - y_1| + \frac{\theta}{2} \cdot |x_2 - y_2| + \frac{\theta}{2} \cdot |x_3 - y_3|.$$

By following arguments similar to those in Propositions 1.3.2 and 1.3.3, the mapping $\mathscr{D}_\theta^\varphi$ is a distance between fuzzy numbers. Thus,

Proposition 1.3.8. *Let φ be an arbitrarily fixed absolutely continuous probability measure on $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function on $(0,1)$, and $\theta \in (0,1]$ be a weighting parameter. Then,*

i) *$\mathscr{D}_\theta^\varphi$ is an L^1 metric on $\mathcal{F}_c(\mathbb{R})$, both translational and rotational invariant.*

ii) *For a fixed φ, the function $\boldsymbol{v}^\varphi : \mathcal{F}_c(\mathbb{R}) \to \mathbb{H}_1^\star = \{L^1\text{-type 3-dimensional vector-valued functions defined on } [0,1]\}$ satisfies that*

- *$\boldsymbol{v}^\varphi$ is an isometry from $(\mathcal{F}_c(\mathbb{R}), \mathscr{D}_\theta^\varphi)$ into $\mathbb{H}_1^\star$,*
- *$\boldsymbol{v}^\varphi(\widetilde{U} + \widetilde{V}) = \boldsymbol{v}^\varphi(\widetilde{U}) + \boldsymbol{v}^\varphi(\widetilde{V})$ for all $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R})$,*
- *$\boldsymbol{v}^\varphi(\gamma \cdot \widetilde{U}) = \gamma \cdot \boldsymbol{v}^\varphi(\widetilde{U})$ for all $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ and $\gamma > 0$;*

Consequently, the $\boldsymbol{v}^\varphi$ function preserves the semilinearity of $\mathcal{F}_c(\mathbb{R})$ and relates the fuzzy arithmetic to the functional arithmetic, what implies that $\mathcal{F}_c(\mathbb{R})$ can be isometrically embedded into a convex cone of the Banach space $\left(\mathbb{H}_1^\star, \|\cdot\|_\theta^{\varphi\star}\right)$ with $\|f - g\|_\theta^{\varphi\star} = \int_{[0,1]} |f(\alpha) - g(\alpha)|_\theta^1\, d\varphi(\alpha)$.

The metric space $(\mathcal{F}_c(\mathbb{R}), \mathscr{D}_\theta^\varphi)$ is separable. This assertion is justified by the fact that $\mathscr{D}_\theta^\varphi$ is topologically equivalent (in fact, strongly equivalent) to the metric $\boldsymbol{\rho}_\mathbf{1}^\varphi$ and, hence, to $\boldsymbol{d}_\mathbf{1}^\varphi$. Specifically (see Appendix for the proof),

Proposition 1.3.9. *Let $\theta \in (0,1]$ be a weight parameter and let φ be an arbitrarily fixed absolutely continuous probability measure on $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function in $(0,1)$. The metric $\mathscr{D}_\theta^\varphi$ is topologically equivalent to the metric $\boldsymbol{\rho}_\mathbf{1}^\varphi$ on $\mathcal{F}_c(\mathbb{R})$. More precisely,*

$$\theta \cdot \boldsymbol{\rho}_\mathbf{1}^\varphi(\widetilde{U}, \widetilde{V}) \leq \mathscr{D}_\theta^\varphi(\widetilde{U}, \widetilde{V}) \leq (2 + 3\,\theta) \cdot \boldsymbol{\rho}_\mathbf{1}^\varphi(\widetilde{U}, \widetilde{V})$$

for all $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R})$.

Remark 1.3.4. As pointed out in Remark 1.3.2 and as a consequence from the embeddings in Propositions 1.3.7 and 1.3.8, at first glance one might be tempted to consider the following reasoning: in extending statistical methods dealing with real-valued data to either interval- or fuzzy number-valued data, one can take into account that any interval value can be identified with a certain two-dimensional vector (by considering either its infimum/supremum or its mid-point/spread representation) and many fuzzy numbers (those being characterized by some few points, like trapezoidals) can be identified with finite-dimensional vectors, so we can immediately particularize the well-known multivariate procedures to analyze these data.

Unfortunately, the last reasoning fails due to the nonlinearities of the space of interval values and the space of fuzzy numbers: the vectors associated with interval and fuzzy number values through the isometries should fulfill some constraints, since not all the elements in $\mathbb{R}^p$ determine a compact interval (in case $p = 2$) or a fuzzy number of the concrete referred shapes. Consequently, if one tries to particularize directly well-known multivariate procedures to analyze either interval- or fuzzy number-valued data, one should never forget that the mathematical and probabilistic results and developments supporting these procedures neglect such constraints, so the particularization is wrong in general. Even when some procedures could be 'physically' applied, conclusions from them could be neither reliable nor well-supported.

1.4 Imprecise-valued random elements

Mathematical modeling is essential in developing data analysis. With this in mind, appropriate models have been recalled in Section 1.1 for set- and fuzzy set-valued data, as well as some useful representations for them. The key operations for statistical purposes and metrics between these data have been presented in Sections 1.2 and 1.3, respectively.

In Statistics data are usually a sample of observations/perceptions from a population of interest. In other words, data are assumed to come from the repeated performance of a random mechanism.

Fréchet [77, 78] anticipated that future mathematics would have to incorporate new and unexpected sorts of objects quite beyond numbers and vectors. In response to this need he introduced random elements taking on values in metric spaces and he also pointed out the valuable implications associated with the introduction of a distance between elements in the considered space. In accordance with their current usage, a random element is defined to be a measurable function between a sample space and a metric space equipped with its Borel σ-algebra.

Random sets and random fuzzy sets determine a well-stated and supported model for the random mechanisms generating set-valued and fuzzy set-valued data within the probabilistic setting, respectively. They integrate both randomness and imprecision, so that the first one affects the generation of experimental data, whereas the second one has effect on the nature of the experimental data, which are assumed to be intrinsically imprecise.

1.4.1 Random compact convex sets

Random sets were rigorously formalized as random elements of a space of sets by Matheron [130], although Robbins [163, 164] had already presented some rather informal instances of the notion several decades before.

In this way, Matheron stated the fundamentals of the theory of random closed sets, as well as the appropriate model and basic tools within the probabilistic setting. In Molchanov [138], one can find a wide and quite updated monograph on random sets.

In Hiai and Umegaki [104], several equivalent definitions for the notion of measurable compact set-valued random elements were established. When the set values belong to $\mathcal{K}_c(\mathbb{R}^p)$, one of the most suitable definitions among them is the following one:

Definition 1.4.1. *Given a probability space* $(\Omega, \mathcal{A}, P)$*, a mapping* $\mathsf{X} : \Omega \to \mathcal{K}_c(\mathbb{R}^p)$ *is said to be a* ***random compact convex set*** *associated with it if* X *is measurable with respect to the Borel* σ*-algebra generated by the topology induced by the Hausdorff metric on* $\mathcal{K}_c(\mathbb{R}^p)$*.*

In case $p = 1$*, we will refer to a random compact convex set as a* ***random interval****.*

Because of the topological equivalences by Vitale [211], one can trivially conclude that:

Proposition 1.4.1. *Given a probability space* $(\Omega, \mathcal{A}, P)$*, a mapping* $\mathsf{X} : \Omega \to \mathcal{K}_c(\mathbb{R}^p)$ *is said to be a* ***random compact convex set*** *associated with it if* X *is measurable with respect to the Borel* σ*-algebra generated by the topology induced by the metrics* δ_1 *or* δ_2 *on* $\mathcal{K}_c(\mathbb{R}^p)$*.*

1.4.2 Random fuzzy sets

On the other hand, random fuzzy sets were initially introduced by Féron [69, 70, 71] in a double way: as a Borel-measurable function (i.e., following Fréchet's approach), and as a level-wise extension of random sets. Féron sketched the guiding idea in the notion, but without specifying some key terms like the involved metrics.

Concerning the level-wise extension of random sets, Puri and Ralescu [157] have made this specification as follows:

Definition 1.4.2. *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space. A mapping* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ *is said to be a* ***random fuzzy set*** *(or* ***fuzzy random variable in Puri and Ralescu's sense****) if for each* $\alpha \in [0,1]$ *the set-valued mapping* $\mathcal{X}_\alpha : \Omega \to \mathcal{K}_c(\mathbb{R}^p)$ *(with* $\mathcal{X}_\alpha(\omega) = \big(\mathcal{X}(\omega)\big)_\alpha$ *for all* $\omega \in \Omega$*) is a random compact convex set.*

When $p = 1$*, we will refer to a random fuzzy set as a* ***random fuzzy number*** *and, otherwise, as a* ***random fuzzy vector****.*

In addition to the level-wise measurability in Definition 1.4.2, Puri and Ralescu also considered to define random fuzzy sets as Borel-measurable functions, as suggested by Féron and in agreement with Fréchet's approach. Thereby, notions like the induced distribution or the independence would be inherited from those in the sample space. In this respect, Puri and Ralescu [156] (see, also, Klement *et al.* [116]) defined alternatively, but not equivalently, a fuzzy random variable associated with a probability space as a Borel-measurable mapping in connection with the metric $\boldsymbol{d_\infty}(\widetilde{U}, \widetilde{V}) = \sup_{\alpha \in [0,1]} d_H(\widetilde{U}_\alpha, \widetilde{V}_\alpha)$, which leads to a non-separable metric space.

Anyway, this $\boldsymbol{d_\infty}$-Borel measurability is stronger than the level-wise one in Definition 1.4.2 (see Colubi *et al.* [32, 33], Kim [115] and Terán [196]).

On $\mathcal{F}_c(\mathbb{R}^p)$, several equivalences to the level-wise measurability can be stated in terms of metrics (see González-Rodríguez *et al.* [90] and Gil *et al.* [83] for a detailed study on one of these equivalences and some interesting implications).

Proposition 1.4.2. *Let $(\Omega, \mathcal{A}, P)$ be a probability space. A mapping $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ is a random fuzzy set if and only if*

i) it is measurable with respect to the Borel σ-algebra generated by the topology induced by the metric D_θ^φ on $\mathcal{F}_c(\mathbb{R}^p)$ (or $\mathfrak{D}_\theta^\varphi$, $\boldsymbol{d_1^\varphi}$ and $\boldsymbol{d_2^\varphi}$), whatever the probability measure φ and the parameter θ may be;

ii) the mapping $s_{\mathcal{X}} = s \circ \mathcal{X} : \Omega \to \mathbb{H}_2$ is a random element of $\mathbb{H}_2$ (the Borel measurability of this random element being associated with the corresponding distance in the isometric embedding within this Hilbert space);

iii) for each $\alpha \in [0,1]$ and $\mathbf{u} \in \mathbb{S}^{p-1}$, the real-valued function $s_{\mathcal{X}}(u, \alpha)$ is a real-valued random variable;

iv) for each $\alpha \in [0,1]$ and $\mathbf{u} \in \mathbb{S}^{p-1}$, the real-valued functions $\mathrm{mid}\, s_{\mathcal{X}}(u, \alpha)$ and $\mathrm{spr}\, s_{\mathcal{X}}(u, \alpha)$ are real-valued random variables, the second one being always non-negative.

It should be remarked again that the Borel measurability of random compact convex sets and random fuzzy sets ensures that one can properly refer to the *induced distribution* of these random elements, the *independence* of two of them, and so on, without needing to state these notions.

1.4.3 Aumann and Aumann-type mean of random sets and random fuzzy sets

Location measures are the first natural attempt when we want to summarize the distribution of either a random compact convex set or a random fuzzy set. The best known location measures are the so-called Aumann mean for random compact convex sets (see Aumann [6], and Artstein and Vitale [5]) and Aumann-type mean value for random fuzzy sets (see Puri and Ralescu [157]), which are formalized as follows:

Definition 1.4.3. [6] *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space and* $\mathsf{X} : \Omega \to \mathcal{K}_c(\mathbb{R}^p)$ *be an associated random compact convex set which is integrably bounded (i.e.,* $E(\|\mathsf{X}\|) < \infty$*, where* $\|\mathsf{X}\|$ *is the random vector* $\sup\{\|\mathbf{x}\|, : \mathbf{x} \in \mathsf{X}\}$*). The* ***(population) Aumann mean*** *or* ***expected value*** *of* X *is the set value* $\mathsf{E}[\mathsf{X}] \in \mathcal{K}_c(\mathbb{R}^p)$ *such that*

$$\mathsf{E}[\mathsf{X}] = \left\{ E(f) \, / \, f : \Omega \to \mathbb{R}^p, \; f \in L^1(\Omega, \mathcal{A}, P), \; f \in \mathsf{X} \text{ a.s. } [P] \right\},$$

that is, it is the set of the mean values of all the 'selections' of X*.*

The Aumann mean of a random compact convex set can be equivalently defined (see, for instance, Vitale [212]) as

Proposition 1.4.3. [212] *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space and* $\mathsf{X} : \Omega \to \mathcal{K}_c(\mathbb{R}^p)$ *be an associated integrably bounded random compact convex set. Then,* $\mathsf{E}[\mathsf{X}]$ *satisfies that*

$$s_{\mathsf{E}[\mathsf{X}]} = E(s_{\mathsf{X}}),$$

where the right-hand side is a Bochner integral.

In the fuzzy set-valued case

Definition 1.4.4. [157] *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space and* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ *be an associated random fuzzy set which is integrably bounded (i.e.,* $E(\|\mathcal{X}_0\|) < \infty$*). The* ***(population) Aumann-type mean*** *or* ***expected value*** *of* $\mathcal{X}$ *is the fuzzy value* $\widetilde{E}(\mathcal{X}) \in \mathcal{F}_c(\mathbb{R}^p)$ *such that for all* $\alpha \in [0,1]$

$$\left(\widetilde{E}(\mathcal{X})\right)_\alpha = \mathsf{E}[\mathcal{X}_\alpha].$$

The Aumann-type mean of a random fuzzy set can be equivalently defined (see Puri and Ralescu [156]) as

Proposition 1.4.4. [156] *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space and* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ *be an associated integrably bounded random fuzzy set. Then,* $\widetilde{E}(\mathcal{X})$ *satisfies that*

$$s_{\widetilde{E}(\mathcal{X})} = E(s_{\mathcal{X}}),$$

where the right-hand side is a Bochner integral.

The preceding imprecise-valued means satisfy several valuable properties similar to those in the classical case. In this way (see for reviews Molchanov [138] and González-Rodríguez *et al.* [90]):

Proposition 1.4.5. E *and* $\widetilde{E}$ *are equivariant under affine transformations on* $\mathcal{K}_c(\mathbb{R}^p)$ *and* $\mathcal{F}_c(\mathbb{R}^p)$*, respectively, that is, if* $\gamma \in \mathbb{R}$*,* $K \in \mathcal{K}_c(\mathbb{R}^p)$*,* $\widetilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$ *and* X *and* $\mathcal{X}$ *are integrably bounded random compact convex set and random fuzzy set, respectively, then*

$$\mathsf{E}[\gamma \cdot \mathsf{X} + K] = \gamma \cdot \mathsf{E}[\mathsf{X}] + K, \quad \widetilde{E}(\gamma \cdot \mathcal{X} + \widetilde{U}) = \gamma \cdot \widetilde{E}(\mathcal{X}) + \widetilde{U}.$$

Proposition 1.4.6. E *and* $\widetilde{E}$ *are additive, that is, if* X *and* Y *are integrably bounded random compact convex sets associated with the same probability space and* $\mathcal{X}$ *and* $\mathcal{Y}$ *are integrably bounded random fuzzy sets associated with the same probability space, then*

$$\mathsf{E}[\mathsf{X} + \mathsf{Y}] = \mathsf{E}[\mathsf{X}] + \mathsf{E}[\mathsf{Y}], \quad \widetilde{E}(\mathcal{X} + \mathcal{Y}) = \widetilde{E}(\mathcal{X}) + \widetilde{E}(\mathcal{Y}).$$

Proposition 1.4.7. E *and* $\widetilde{E}$ *are coherent with the usual set- and fuzzy set-valued arithmetics, respectively, so that if* X *and* $\mathcal{X}$ *are finite-valued random compact convex set and random fuzzy set, respectively, that is,* $\mathsf{X}(\Omega) = \{K_1, \ldots, K_m\} \subset \mathcal{K}_c(\mathbb{R}^p)$ *and* $\mathcal{X}(\Omega) = \{\widetilde{U}_1, \ldots, \widetilde{U}_m\} \subset \mathcal{F}_c(\mathbb{R}^p)$*, then if* $p_i = P\big(\{\omega \in \Omega : \mathsf{X}(\omega) = K_i\}\big)$ *or* $p_i = P\big(\{\omega \in \Omega : \mathcal{X}(\omega) = \widetilde{U}_i\}\big)$*, we have that*

$$\mathsf{E}[\mathsf{X}] = p_1 \cdot K_1 + \ldots + p_m \cdot K_m, \quad \widetilde{E}(\mathcal{X}) = p_1 \cdot \widetilde{U}_1 + \ldots + p_m \cdot \widetilde{U}_m.$$

In fact, the ***sample Aumann mean*** associated with a simple random sample $(\mathsf{X}_1, \ldots, \mathsf{X}_n)$ from a random compact convex set X (i.e., $\mathsf{X}_1, \ldots, \mathsf{X}_n$ are independent random compact convex sets which are identically distributed as X) is given by

$$\overline{\mathsf{X}_n} = \frac{1}{n} \cdot \big[\mathsf{X}_1 + \ldots + \mathsf{X}_n\big].$$

Analogously, the ***sample Aumann-type mean*** associated with a simple random sample $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ from a random fuzzy set $\mathcal{X}$ (i.e., $\mathcal{X}_1, \ldots, \mathcal{X}_n$ are independent random fuzzy sets which are identically distributed as $\mathcal{X}$) is given by

$$\overline{\mathcal{X}_n} = \frac{1}{n} \cdot \big[\mathcal{X}_1 + \ldots + \mathcal{X}_n\big].$$

The computation of the Aumann-type mean is now illustrated by considering the situation in Example 1.3.1.

Example 1.4.1. We consider the 'sample' of the 66 conical fuzzy vectors representing the main fundamental type colors and modifiers shown in Figure 1.7.

These 66 fuzzy vectors are such that No. i corresponds to the right cone $\widetilde{C}^i$ such that for each $\alpha \in [0, 1]$

$$\widetilde{C}^i_\alpha = \text{circle with centre } (x_i, y_i) \text{ and radius } 10(1 - \alpha).$$

All the centres have been detailed in Table 1.1.

$\widetilde{C}^\bullet$	centre	$\widetilde{C}^\bullet$	centre	$\widetilde{C}^\bullet$	centre	$\widetilde{C}^\bullet$	centre	$\widetilde{C}^\bullet$	centre	$\widetilde{C}^\bullet$	centre
1	$(0,0)$	**12**	$(5\sqrt{3},5)$	**23**	$(10\sqrt{3},20)$	**34**	$(15\sqrt{3},45)$	**45**	$(20\sqrt{3},80)$	**56**	$(30\sqrt{3},70)$
2	$(0,10)$	**13**	$(5\sqrt{3},15)$	**24**	$(10\sqrt{3},30)$	**35**	$(15\sqrt{3},55)$	**46**	$(25\sqrt{3},25)$	**57**	$(35\sqrt{3},35)$
3	$(0,20)$	**14**	$(5\sqrt{3},25)$	**25**	$(10\sqrt{3},40)$	**36**	$(15\sqrt{3},65)$	**47**	$(25\sqrt{3},35)$	**58**	$(35\sqrt{3},45)$
4	$(0,30)$	**15**	$(5\sqrt{3},35)$	**26**	$(10\sqrt{3},50)$	**37**	$(15\sqrt{3},75)$	**48**	$(25\sqrt{3},45)$	**59**	$(35\sqrt{3},55)$
5	$(0,40)$	**16**	$(5\sqrt{3},45)$	**27**	$(10\sqrt{3},60)$	**38**	$(15\sqrt{3},85)$	**49**	$(25\sqrt{3},55)$	**60**	$(35\sqrt{3},65)$
6	$(0,50)$	**17**	$(5\sqrt{3},55)$	**28**	$(10\sqrt{3},70)$	**39**	$(20\sqrt{3},20)$	**50**	$(25\sqrt{3},65)$	**61**	$(40\sqrt{3},40)$
7	$(0,60)$	**18**	$(5\sqrt{3},65)$	**29**	$(10\sqrt{3},80)$	**40**	$(20\sqrt{3},30)$	**51**	$(25\sqrt{3},75)$	**62**	$(40\sqrt{3},50)$
8	$(0,70)$	**19**	$(5\sqrt{3},75)$	**30**	$(10\sqrt{3},90)$	**41**	$(20\sqrt{3},40)$	**52**	$(30\sqrt{3},30)$	**63**	$(40\sqrt{3},60)$
9	$(0,80)$	**20**	$(5\sqrt{3},85)$	**31**	$(15\sqrt{3},15)$	**42**	$(20\sqrt{3},50)$	**53**	$(30\sqrt{3},40)$	**64**	$(45\sqrt{3},45)$
10	$(0,90)$	**21**	$(5\sqrt{3},95)$	**32**	$(15\sqrt{3},25)$	**43**	$(20\sqrt{3},60)$	**54**	$(30\sqrt{3},50)$	**65**	$(45\sqrt{3},55)$
11	$(0,100)$	**22**	$(10\sqrt{3},10)$	**33**	$(15\sqrt{3},35)$	**44**	$(20\sqrt{3},70)$	**55**	$(30\sqrt{3},60)$	**66**	$(50\sqrt{3},50)$

Table 1.1: Centres (x_i, y_i) of the 66 fundamental type colors in Figure 1.7 (Example 1.3.1)

By using Proposition 1.4.4, the sample Aumann-type mean,

$$\frac{1}{66}\cdot\left[\widetilde{C}^1+\ldots+\widetilde{C}^{66}\right],$$

satisfies that for all $\alpha \in [0,1]$ and $\mathbf{u} \in \mathbb{S}^1$

$$s_{\frac{1}{66}\cdot[\widetilde{C}^1+\ldots+\widetilde{C}^{66}]}(\alpha,\mathbf{u}) = \frac{1}{66}\sum_{i=1}^{66} s_{\widetilde{C}^i}(\alpha,\mathbf{u}).$$

In accordance with the expression given for $s_{\widetilde{C}^i}(\alpha,\mathbf{u})$ in Example 1.3.1, we can conclude that

$$\begin{aligned} s_{\frac{1}{66}\cdot[\widetilde{C}^1+\ldots+\widetilde{C}^{66}]}(\alpha,\mathbf{u}) &= \cos\beta_{\mathbf{u}}\cdot\frac{1}{66}\sum_{i=1}^{66} x_i + \sin\beta_{\mathbf{u}}\cdot\frac{1}{66}\sum_{i=1}^{66} y_i + 10(1-\alpha) \\ &= \frac{50\sqrt{3}}{3}\cdot\cos\beta_{\mathbf{u}} + 50\cdot\sin\beta_{\mathbf{u}} + 10(1-\alpha), \end{aligned}$$

which is the support function associated with the right cone $\widetilde{C}^*$ such that for each $\alpha \in [0,1]$

$$\widetilde{C}^*_\alpha = \text{circle with centre } (50\sqrt{3}/3, 50) \text{ and radius } 10(1-\alpha).$$

Obviously, this mean value does not coincide with any of the 66 fundamental type colors, but it is surrounded by $\widetilde{C}_{34}$, $\widetilde{C}_{35}$ and $\widetilde{C}_{42}$.

If we consider the detail type color triangle delimited by these three main colors (see Figure 1.12 on the left), we obtain that the centre corresponds to white (No. 104), which is surrounded by six neighboring colors (No. 101 − No. 107), all of them surrounded by No. 34, No. 35 and No. 42. Therefore, $\widetilde{C}^*$ corresponds to the right cone associated with white and being displayed in Figure 1.12 (on the right).

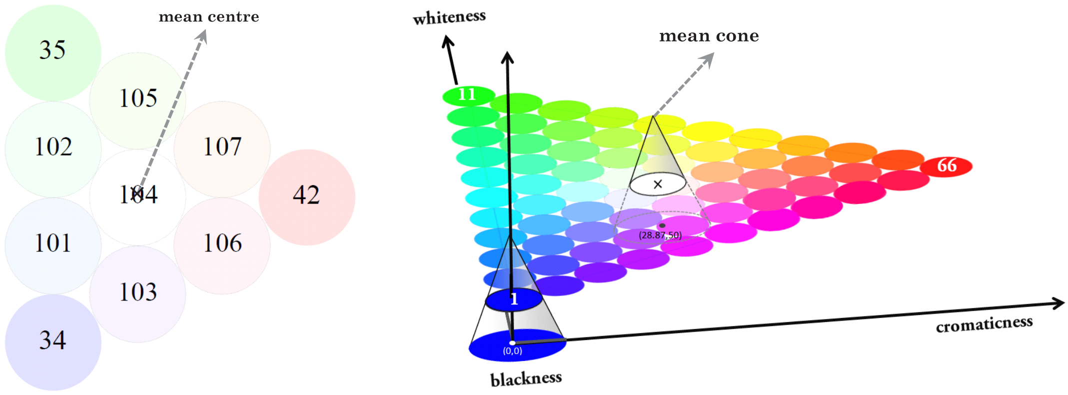


Figure 1.12: Conical fuzzy vectors in the usual color triangle

Although other possible definitions for the mean of random imprecise-valued elements have been suggested in the literature, it should be pointed out that the Aumann and Aumann-type concepts are supported by *Strong Laws of Large Numbers* for random sets (see, for instance, Artstein and Vitale [5], Giné *et al.* [86], Puri and Ralescu [155], Molchanov [136] and Terán [198]) and for random fuzzy sets (see, for instance, Klement *et al.* [116], Colubi *et al.* [35], Molchanov [137], Proske and Puri [153] and Terán [195]). In accordance with these results, the population Aumman and Aumman-type means are the almost sure limit (in different metrics' sense) of the sample Aumann and Aumann-type means, respectively.

Furthermore, the Aumann and Aumann-type means are the Fréchet expectations for all the L^2 metrics considered in this work. It is already known (see, for instance, González-Rodríguez et al. [90]) that

$$\mathsf{E}[\mathsf{X}] = \arg \min_{K \in \mathcal{K}_c(\mathbb{R}^p)} E\left(\left[d_\theta^\varphi(\mathrm{X}, K)\right]^2\right) \text{ and } \widetilde{E}(\mathcal{X}) = \arg \min_{\widetilde{U} \in \mathcal{F}_c(\mathbb{R}^p)} E\left(\left[D_\theta^\varphi(\mathcal{X}, \widetilde{U})\right]^2\right).$$

This is also true under quite general conditions when the mid/spr-based L^2 metric is replaced by the support/Steiner-based L^2 one. Thus,

Proposition 1.4.8. *Let φ be an absolutely continuous probability measure on the measurable space $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function on $(0,1)$, $\theta \in (0,1]$ be a parameter and $\mathcal{X}$ be a random fuzzy set associated with the probability space $(\Omega, \mathcal{A}, P)$. Assume that $s_{\mathcal{X}} \in L^2(\Omega, \mathcal{A}, P)$ and that $\mathbf{u} \cdot s_{\mathcal{X}}(\alpha, \mathbf{u})$ fulfills sufficient conditions to ensure the exchange of the iterated integrals of this vector-valued function on Ω and $[0,1] \times \mathbb{S}^{p-1}$, then*

$$\widetilde{E}(\mathcal{X}) = \arg \min_{\widetilde{U} \in \mathcal{F}_c(\mathbb{R}^p)} E\left(\left[\mathfrak{D}_\theta^\varphi(\mathcal{X}, \widetilde{U})\right]^2\right).$$

1.4.4 Symmetric random fuzzy numbers/intervals

The behaviour of location measures when the distribution is symmetric is especially interesting. The notion of symmetry of a distribution can be extended from the real- to the interval- and fuzzy number-valued cases in a natural way. To formalize this extension one can state the following:

Definition 1.4.5. *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space and let* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ *be a* ***random fuzzy number*** *associated with* $(\Omega, \mathcal{A}, P)$. $\mathcal{X}$ *is said to be* ***symmetric about*** $c \in \mathbb{R}$ *if, and only if,* $\mathcal{X} - c \stackrel{d}{=} c - \mathcal{X}$ *or, equivalently,* $\mathcal{X} \stackrel{d}{=} 2c - \mathcal{X}$, *where* $\stackrel{d}{=}$ *denotes the identity in distribution.*

Obviously, $\mathcal{X}$ is a symmetric random fuzzy number about c if, and only if, $\mathcal{X} - c$ is a symmetric random fuzzy number about 0. Then, if $\mathcal{X}$ is a symmetric random fuzzy number about c, it can be rewritten as $\mathcal{X} = \mathcal{E} + c$, where $\mathcal{E}$ is a symmetric random fuzzy number about 0.

To interpret this notion, we can consider the following three examples. The first one, really simple, indicates a key divergence with respect to the real-valued settings: the fact that the space is nonlinear.

Example 1.4.2. Let $\mathcal{X}$ be a random fuzzy number associated with a probability space $(\Omega, \mathcal{A}, P)$ and let $\mathcal{O}_{\mathcal{X}} = \mathcal{X} - \mathcal{X}$. For any $\omega \in \Omega$, we have that $\mathcal{O}_{\mathcal{X}}(\omega)$ is a symmetric fuzzy number about 0.

The second example is based on the one supplied by Chou [30].

Example 1.4.3. In many social surveys, respondents are customarily asked by means of a questionnaire to indicate their choices from a set of prefixed Likert-type items. Many researchers consider Likert-type labels as fuzzy number-valued ones, by identifying the generic response to a question with a fuzzy linguistic variable (see, for instance, Serrano-Guerrero *et al.* [168] or Porcel *et al.* [152] for recent studies about). As indicated by Chou [30], "often the wording of response levels clearly implies a symmetry of response levels about a middle category; at the very least, such an item would fall between ordinal-level and interval-level measurement... The use of fuzzy sets is central in computing with words or labels as they provide a means of modeling the vagueness underlying most natural linguistic terms (see, for instance, Zadeh [220]). The semantic elements of the term set are given by fuzzy numbers defined on a bounded interval (say $[0, 1]$).

In practice, triangular fuzzy numbers are a uniformly distributed ordered set of linguistic terms, so they provide a relatively simple way to capture the vagueness of linguistic assessments,...", like the ones graphically displayed in Figure 1.13, where VD = STRONGLY DISAGREE, D = DISAGREE, SD = SOMEWHAT DISAGREE, N = NEITHER AGREE NOR DISAGREE, SA = SOMEWHAT AGREE, A = AGREE and VA = STRONGLY AGREE.

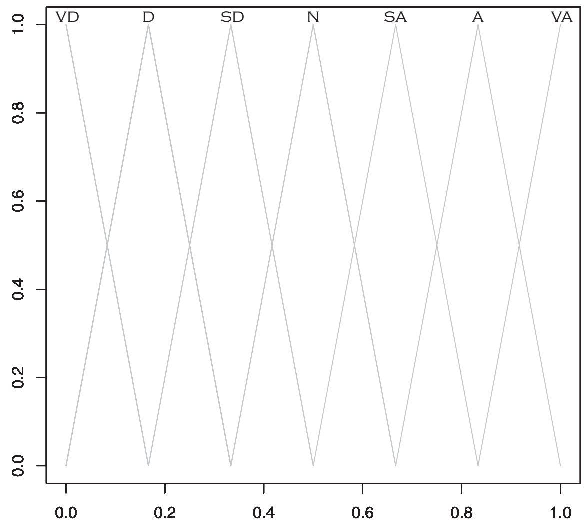


Figure 1.13: Semantic elements of a term set given by 7 fuzzy triangular numbers

Assume that the survey has been performed in a population/sample $\{\omega_1, \dots, \omega_{953}\}$ of 953 respondents and that the distribution of the random response $\mathcal{X}$ is the following:

label	VD	D	SD	N	SA	A	VA
absol. freq.	38	143	207	177	207	143	38

Since

$$\text{VD} = 1 - \text{VA},\ \text{D} = 1 - \text{A},\ \text{SD} = 1 - \text{SA},$$

and

$$\#\{\omega_j : \mathcal{X}(\omega_j) = \text{VD}\} = 38 = \#\{\omega_j : \mathcal{X}(\omega_j) = \text{VA}\},$$

$$\#\{\omega_j : \mathcal{X}(\omega_j) = \text{D}\} = 143 = \#\{\omega_j : \mathcal{X}(\omega_j) = \text{A}\},$$

$$\#\{\omega_j : \mathcal{X}(\omega_j) = \text{SD}\} = 207 = \#\{\omega_j : \mathcal{X}(\omega_j) = \text{SA}\},$$

then the random fuzzy set $\mathcal{X}$ is symmetric about $c = 0.5$.

The third example has been drawn from the so-called 'characterizing fuzzy representation' of real-valued random variables (see González-Rodríguez *et al.* [89]).

Example 1.4.4. Let X be a random variable associated with a probability space and assume that X has a Binomial distribution $\mathrm{Bin}(4, 0.5)$. González-Rodríguez *et al.* [89] have introduced a generalized fuzzy representation that characterizes the distribution of a real-valued random variable by means of the Aumann-type expected value of the random fuzzy set corresponding to the composition of the fuzzy representation and the random variable.

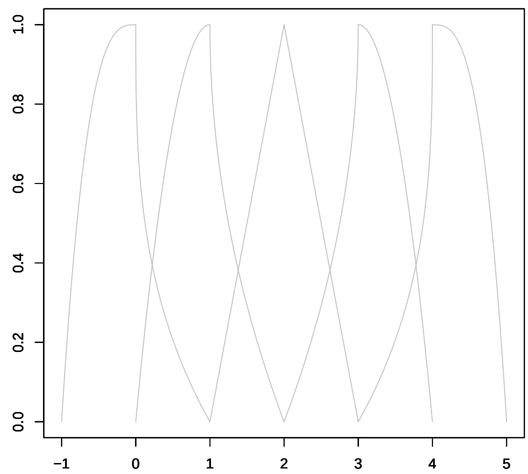


Figure 1.14: Values of a characterizing fuzzy representation of an RV taking on values 0, 1, 2, 3 and 4

In this way, if $X \sim \mathrm{Bin}(4, 0.5)$, the random fuzzy number $\mathcal{X} = \gamma_{(2)} \circ X$ such that for all $\alpha \in [0, 1]$

$$\left(\gamma_{(2)}(x)\right)_\alpha = \begin{cases} \left[x - (1-\alpha)^{1/(3-x)}, x + (1-\alpha)^{3-x}\right] & \text{if } x = 0, 1 \\ \\ \left[x - (1-\alpha)^{x-1}, x + (1-\alpha)^{1/(x-1)}\right] & \text{if } x = 2, 3, 4 \end{cases}$$

is a symmetric random fuzzy number about $c = 2$. The five different fuzzy values it takes on have been graphically displayed on Figure 1.14 and their corresponding probabilities are those associated with the five values of the Binomial.

Now it will be proven that the symmetry of a random fuzzy number entails the symmetry of all its α-level random intervals:

Proposition 1.4.9. *Let $\mathcal{X}$ be a random fuzzy number associated with a probability space $(\Omega, \mathcal{A}, P)$. If $\mathcal{X}$ is symmetric about $c \in \mathbb{R}$, then, for all $\alpha \in [0, 1]$, the α-level random interval $\mathcal{X}_\alpha$ is symmetric about c.*

The converse assertion is not true, that is, if for each $\alpha \in [0,1]$ the random interval $\mathcal{X}_\alpha$ is symmetric about c, the random fuzzy number $\mathcal{X}$ is not necessarily symmetric about c. As a counterexample we can consider the following:

Example 1.4.5. Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $\mathcal{A} = \mathcal{P}(\Omega)$ and P being associated with a uniform distribution on Ω. Let $\mathcal{X}$ be the random fuzzy number such that

$$\mathcal{X}(\omega_1)(x) = \begin{cases} x^2/2 & \text{if } x \in [0,1] \\ -(x^2 - 4x + 2)/2 & \text{if } x \in (1,3] \\ (x-4)^2/2 & \text{si } x \in (3,4] \\ 0 & \text{otherwise} \end{cases}$$

$$\mathcal{X}(\omega_2)(x) = \text{Tri}(0,2,4),$$

where $\text{Tri}(a,b,c)$ denotes the triangular fuzzy number such that $\text{Tri}(a,b,c)_0 = [a,c]$, $\text{Tri}(a,b,c)_1 = \{b\}$,

$$\mathcal{X}(\omega_3)(x) = \begin{cases} -x/2 & \text{if } x \in (-1,0] \\ -(x^2 + 4x + 2)/2 & \text{if } x \in (-3,-1] \\ (x+4)/2 & \text{si } x \in [-4,-3] \\ 0 & \text{otherwise} \end{cases}$$

$$\mathcal{X}(\omega_4)(x) = \begin{cases} x^2/2 & \text{if } x \in (-1,0] \\ -x/2 & \text{if } x \in (-2,-1] \\ (x+4)/2 & \text{if } x \in (-3,-2] \\ (x+4)^2/2 & \text{si } x \in [-4,-3] \\ 0 & \text{otherwise} \end{cases}$$

One can easily prove that, for all $\alpha \in [0,1]$, the random interval $\mathcal{X}_\alpha$ is symmetric about 0:

- on one hand, for all $\alpha \in [0, 0.5]$

  $$\mathcal{X}_\alpha(\omega_1) = -\mathcal{X}_\alpha(\omega_4),\ \mathcal{X}_\alpha(\omega_2) = -\mathcal{X}_\alpha(\omega_3),$$

  whence, because of P being associated with the uniform distribution on Ω, $\mathcal{X}_\alpha$ and $-\mathcal{X}_\alpha$ are identically distributed, i.e., $\mathcal{X}_\alpha$ is symmetric about 0;
- on the other hand, for all $\alpha \in (0.5, 1]$

  $$\mathcal{X}_\alpha(\omega_1) = -\mathcal{X}_\alpha(\omega_3),\ \mathcal{X}_\alpha(\omega_2) = -\mathcal{X}_\alpha(\omega_4),$$

  whence, because of P being associated with the uniform distribution on Ω, $\mathcal{X}_\alpha$ and $-\mathcal{X}_\alpha$ are identically distributed, i.e., $\mathcal{X}_\alpha$ is symmetric about 0.

However, each of the four distinct values of $\mathcal{X}$ differ from the four distinct values of $-\mathcal{X}$, so $\mathcal{X}$ is not symmetric about 0.

The Aumann-type mean value of a symmetric random fuzzy number satisfies the following property:

Proposition 1.4.10. *Let $(\Omega, \mathcal{A}, P)$ be a probability space and let $\mathcal{X}$ be an integrably bounded symmetric random fuzzy number about $c \in \mathbb{R}$. Then, the Aumann-type mean value of $\mathcal{X}$ is a symmetric fuzzy number about c.*

The result in Proposition 1.4.10 is now illustrated by computing the mean values of the symmetric random fuzzy numbers in Examples 1.4.3 and 1.4.4.

Example 1.4.6. The Aumann-type mean value of the symmetric random fuzzy number about 0.5 in Example 1.4.3 is graphically displayed in Figure 1.15:

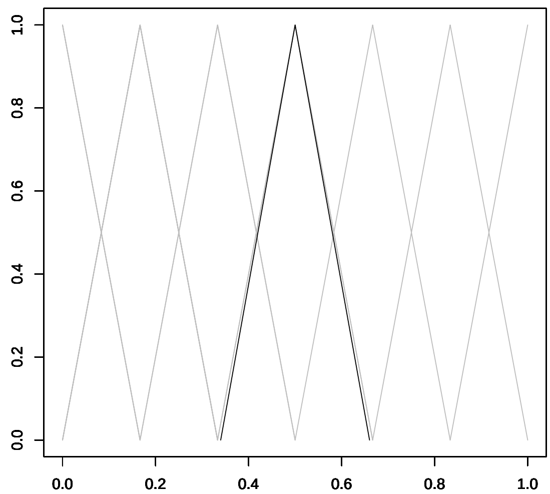


Figure 1.15: Aumann-type mean (in black) of the 953 responses in Example 1.4.3

The Aumann-type mean expected value of the symmetric random fuzzy number about 2 in Example 1.4.4 is graphically displayed in Figure 1.16.

1.5 Motivating and clarifying simulation studies

This section aims to illustrate empirically (and sometimes simultaneously) several facts. Although similar conclusions could be drawn for both set- and fuzzy set-valued data, the empirical analysis will be focused on the more general case of fuzzy

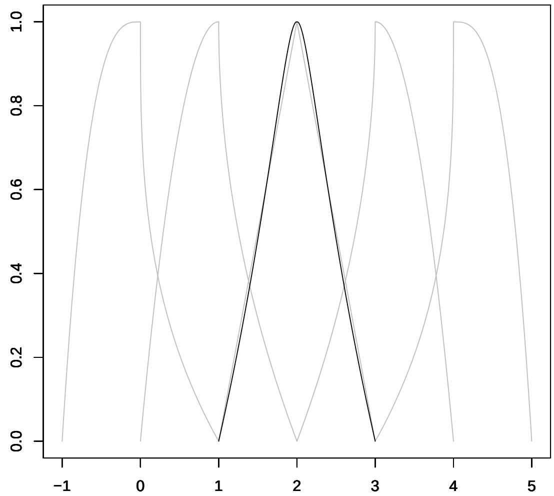


Figure 1.16: Aumann-type mean (in black) of the characterizing fuzzy representation of the Bin$(4, 0.5)$ in Example 1.4.4

set-valued data. Actually, the simulations involve trapezoidal fuzzy numbers and interval-valued data are indeed special trapezoidal fuzzy numbers.

The most relevant fact to be illustrated refers to the evidence that, although the Aumann and Aumann-type means are the most common candidates to get some idea about the central tendency of a sample or population of set- or fuzzy set-valued data, they not only inherit from the real-valued case the properties indicated at the end of last section, but also the sensitivity of the mean to either data perturbations or the existence of extreme values (outliers).

The second aspect we will focus on is related to the comparison of the behaviour of the mid/spr and the support/Steiner-based L^2 metrics, as well as the effect the choice of the weight parameter θ has on them. The third point is analogous to the second one, but replacing L^2 by L^1 metrics.

The **first simulations** are addressed to the first target. They have been performed as follows:

Step 1. A sample of size $n = 100000$ of trapezoidal fuzzy numbers has been simulated for each of several different situations in such a way that

- to generate the trapezoidal fuzzy data, we have considered four real-valued random variables as follows: $\mathcal{X} = \mathrm{Tra}(X_1 - X_2 - X_3, X_1 - X_2, X_1 + X_2, X_1 + X_2 + X_4)$, with $X_1 = \mathrm{mid}\,\mathcal{X}_1$, $X_2 = \mathrm{spr}\,\mathcal{X}_1$, $X_3 = \inf \mathcal{X}_1 - \inf \mathcal{X}_0$, $X_4 = \sup \mathcal{X}_0 - \sup \mathcal{X}_1$;
- each sample is assumed to be split into a subsample of size $n(1-c_p)$ (where c_p denotes the proportion of contamination and is supposed to range in

$\{0, 0.1, 0.2, 0.4\}$) associated with a non-contaminated distribution and a subsample of size $n \cdot c_p$ associated with a contaminated one. C_D plays an additional contamination role, measuring how far the distribution of the contaminated subsample is from the distribution of the non-contaminated one (and ranges in $\{0, 1, 5, 10, 100\}$);

- 16 situations for different values of c_p and C_D have been considered and for each of them two cases have been selected, namely, one in which the random variables X_i are independent (CASE 1) and another one in which they are dependent (CASE 2). More specifically, CASE 1 assumes that:
  - •• $X_1 \sim \mathcal{N}(0, 1)$ and $X_2, X_3, X_4 \sim \chi_1^2$ for the non-contaminated subsample,
  - •• $X_1 \sim \mathcal{N}(0, 3) + C_D$ and $X_2, X_3, X_4 \sim \chi_4^2 + C_D$ for the contaminated subsample,

  whereas CASE 2 assumes that:
  - •• $X_1 \sim \mathcal{N}(0, 1)$ and $X_2, X_3, X_4 \sim 1/(X_1^2 + 1)^2 + 0.1 \cdot \chi_1^2$ for the non-contaminated subsample (with χ_1^2 independent of X_1),
  - •• $X_1 \sim \mathcal{N}(0, 3) + C_D$ and $X_2, X_3, X_4 \sim 1/(X_1^2 + 1)^2 + 0.1 \cdot \chi_1^2 + C_D$ for the contaminated subsample (with χ_1^2 independent of X_1),

  obviously, being χ_1^2 independent from $1/(X_1^2 + 1)^2$.

Step 2. $N = 1000$ replications of *Step 1* have been considered, so that for each of the 16 situations concerning c_p and c_D there are 1000 available samples of size $n = 100000$.

Step 3. The population Aumann-type mean for the non-contaminated distribution has been approximated by a Monte Carlo approach on the basis of the 1000×100000 data from the first situation ($c_p = C_D = 0$). For each of the 16 situations, the 1000 sample Auman-type means are computed as estimates of the population Aumann-type mean for the non-contaminated distribution and

- the mean distance between the non-contaminated distribution to each sample mixed Aumann-type mean is computed and finally the mean over the 1000 samples (MD) is obtained;
- the mean of the 1000 distances (DM) between the mixed and non-contaminated Aumann-type means is obtained.

		CASE 1		CASE 2	
c_p	c_D	MD	DM	MD	DM
0	0	1.590684	0.005446	1.0027501	0.003035
0.1	0	1.685077	0.340882	1.004752	0.012257
0.1	1	1.727329	0.457775	0.990237	0.120268
0.1	5	1.958604	0.965272	1.085708	0.641559
0.1	10	2.355203	1.614466	1.568303	1.306176
0.1	100	13.552122	13.479923	13.187656	13.156037
0.2	0	1.825401	0.680722	1.007136	0.023612
0.2	1	1.946608	0.924358	0.988365	0.241804
0.2	5	2.601743	1.952914	1.548887	1.283131
0.2	10	3.658885	3.261031	2.728677	2.5792370
0.2	100	26.827333	26.793580	26.308354	26.292482
0.4	0	2.212263	1.364644	1.012141	0.046735
0.4	1	2.558034	1.874279	1.025340	0.485976
0.4	5	4.194068	3.860446	2.701750	2.551476
0.4	10	6.646680	6.470993	5.211615	5.133452
0.4	100	54.186654	54.170428	51.864586	51.856741

Table 1.2: Mean distances of the mixed Aumann-type mean to the non-contaminated distribution (MD), and distances between the sample mixed and non-contaminated Aumann-type means (DM)

Distances have been computed by considering the L^2 metric $\boldsymbol{\rho_2} = D_1^\ell = \mathfrak{D}_1^\ell$. The outputs of this simulation study have been collected in Table 1.2.

On the basis of the outputs in Table 1.2, one can conclude from an empirical point of view that in CASE 1 the higher the perturbation (especially, the one through c_D), the worse the sample mean summarizes the non-contaminated distribution. More concretely, for a fixed level of contamination c_p,

- the farther the contaminated distribution from the non-contaminated one, the substantially greater mean distance between the approximated mean and the non-contaminated distribution;
- the farther the contaminated distribution from the non-contaminated one, the substantially greater distance between the contaminated and the non-contaminated means.

The behaviour is quite close, although slightly less monotonic in CASE 2. Therefore, this first empirical study greatly motivates the developments in Chapters 2 and 3 of this work.

The **second simulations** will be addressed to the two first targets and they have been performed as follows:

Step 1 and *Step 2* are similar to those for the first simulations.

Step 3'. The population Aumann-type mean for the non-contaminated distribution has been approximated by a Monte Carlo approach on the basis of the 1000×100000 data from from the first situation ($c_p = C_D = 0$). For each of the 16 situations, the 1000 sample Auman-type means are computed as estimates of the population Aumann-type mean for the non-contaminated distribution. The mean squared error (MSE) of the estimator has been approximated by the mean squared distance between the sample mixed and the approximate population mean for the non-contaminated distribution.

Distances have been computed by considering the L^2 metrics D_θ^ℓ and $\mathfrak{D}_\theta^\ell$. The outputs for this simulation study have been collected in Table 1.3 for D_θ^ℓ, and those for $\mathfrak{D}_\theta^\ell$ are in Table 1.4. It is convenient to realize that

$$\widehat{MSE}(\mathcal{X}, \widetilde{E}(\mathcal{X})) = E\left(\left[D_\theta^\varphi(\mathcal{X}, \widetilde{E}(\mathcal{X}))\right]^2\right) = V_1 + \theta \cdot V_2,$$

$$\widehat{\mathfrak{MSE}}(\mathcal{X}, \widetilde{E}(\mathcal{X})) = E\left(\left[\mathfrak{D}_\theta^\varphi(\mathcal{X}, \widetilde{E}(\mathcal{X}))\right]^2\right) = \mathfrak{V}_1 + \theta \cdot \mathfrak{V}_2,$$

with

$$V_1 = \int_{[0,1]} \mathrm{Var}(\mathrm{mid}\, \mathcal{X}_\alpha)\, d\varphi(\alpha), \quad V_2 = \int_{[0,1]} \mathrm{Var}(\mathrm{spr}\, \mathcal{X}_\alpha)\, d\varphi(\alpha),$$

		CASE 1			**CASE 2**		
c_p	c_D	V_1	V_2	$\widehat{MSE}$	V_1	V_2	$\widehat{MSE}$
0	0	0.000014	0.000023	0.000014 + 0.000023 θ	0.000010	0.000003	0.000010 +0.000003 θ
0.1	0	0.000019	0.127470	0.000019 + 0.127470 θ	0.000015	0.000243	0.000015 +0.000243 θ
0.1	1	0.003466	0.221471	0.003466 + 0.221471 θ	0.003402	0.011613	0.003402 + 0.011613 θ
0.1	5	0.079311	0.923035	0.079311 + 0.923035 θ	0.086931	0.336959	0.086931 +0.336959 θ
0.1	10	0.334994	2.401428	0.334994 + 2.401428 θ	0.336682	1.375143	0.336682 + 1.375143 θ
0.1	100	32.67661	153.0056	32.67661 + 153.0056 θ	32.81563	141.5788	32.81563+141.5788 θ
0.2	0	0.000024	0.515910	0.000024 + 0.515910 θ	0.000017	0.000874	0.000017 + 0.000874 θ
0.2	1	0.013199	0.927162	0.013199 + 0.927162 θ	0.013452	0.047189	0.013452 + 0.047189 θ
0.2	5	0.337335	3.648890	0.337335 + 3.648890 θ	0.334835	1.341384	0.334835 +1.341384 θ
0.2	10	1.356971	9.576707	1.356971 + 9.576707 θ	1.298018	5.573106	1.298018 +5.573106 θ
0.2	100	127.7962	616.7393	127.7962 + 616.7393 θ	134.6084	552.2032	134.6084+552.2032 θ
0.4	0	0.000033	2.092744	0.000033 + 2.092744 θ	0.000025	0.003553	0.000025 + 0.003553 θ
0.4	1	0.052931	3.657312	0.052931 + 3.657312 θ	0.053656	0.188999	0.053656 + 0.188999 θ
0.4	5	1.330965	14.69655	1.330965 + 14.69655 θ	1.358108	5.459095	1.358108 + 5.459095 θ
0.4	10	5.257283	39.60472	5.257283 + 39.60472 θ	5.622794	21.18859	5.622794+ 21.18859 θ
0.4	100	542.4229	2425.985	542.4229 + 2425.985 θ	542.4764	2244.256	542.4764+2244.256 θ

Table 1.3: Approximate D_θ^ℓ-based mean squared error of the Aumann-type sample-mean estimator based on mixed samples

$$\mathfrak{V}_1 = \mathrm{Var}(\mathrm{wabl}^\varphi(\mathcal{X})),$$

		CASE 1			**CASE 2**		
c_p	c_D	$\mathfrak{V}_1$	$\mathfrak{V}_2$	$\widehat{\mathfrak{MSE}}$	$\mathfrak{V}_1$	$\mathfrak{V}_2$	$\widehat{\mathfrak{MSE}}$
0	0	0.000013	0.000024	$0.000013 + 0.000024\,\theta$	0.000010	0.000003	$0.000010 + 0.000003\,\theta$
0.1	0	0.000018	0.127471	$0.000018 + 0.127471\,\theta$	0.000015	0.000243	$0.000015 + 0.000243\,\theta$
0.1	1	0.003465	0.221472	$0.003465 + 0.221472\,\theta$	0.003402	0.011613	$0.003402 + 0.011613\,\theta$
0.1	5	0.079310	0.923036	$0.079310 + 0.923036\,\theta$	0.086931	0.336959	$0.086931 + 0.336959\,\theta$
0.1	10	0.334993	2.401429	$0.334993 + 2.401429\,\theta$	0.336682	1.375143	$0.336682 + 1.375143\,\theta$
0.1	100	32.67661	153.0056	$32.67661 + 153.0056\,\theta$	32.81563	141.5788	$32.81563+141.5788\ \theta$
0.2	0	0.000022	0.515911	$0.000022 + 0.515911\,\theta$	0.000017	0.000874	$0.000017 + 0.000874\,\theta$
0.2	1	0.013198	0.927163	$0.013198 + 0.927163\,\theta$	0.013452	0.047189	$0.013452 + 0.047189\,\theta$
0.2	5	0.337334	3.648892	$0.337334 + 3.648892\,\theta$	0.334835	1.341384	$0.334835 + 1.341384\,\theta$
0.2	10	1.356969	9.576708	$1.356969 + 9.576708\,\theta$	1.298018	5.573106	$1.298018 + 5.573106\,\theta$
0.2	100	127.7962	616.7393	$127.7962 + 616.7393\,\theta$	134.6084	552.2032	$134.6084+552.2032\ \theta$
0.4	0	0.000031	2.092746	$0.000031 + 2.092746\,\theta$	0.000025	0.003553	$0.000025 +0.003553\ \theta$
0.4	1	0.052929	3.657314	$0.052929 + 3.657314\,\theta$	0.53656	0.188999	$0.53656 +0.188999\ \theta$
0.4	5	1.330964	14.696555	$1.330964 + 14.69655\,\theta$	1.358108	5.459095	$1.358108 + 5.459095\,\theta$
0.4	10	5.257282	39.604723	$5.257282 + 39.60472\,\theta$	5.622794	21.18859	$5.622794 + 21.18859\,\theta$
0.4	100	542.4228	2425.9857	$542.4228 + 2425.985\,\theta$	542.4764	2244.256	$542.4764+ 2244.256\ \theta$

Table 1.4: Approximate $\mathfrak{D}^{\ell}_{\theta}$-based mean squared error of the Aumann-type sample-mean estimator based on mixed samples

$$\mathfrak{V}_2 = \frac{1}{2}\int_{[0,1]} \mathrm{Var}^{\varphi}(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha))\, d\varphi(\alpha) + \frac{1}{2}\int_{[0,1]} \mathrm{Var}^{\varphi}(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha))\, d\varphi(\alpha).$$

On the basis of the last two tables, one can conclude empirically in connection with the first target that the sample mean is quite sensitive. More specifically, the approximated mean squared error substantially increases with the contamination, especially when the distribution of the contaminated subsample is very far from the distribution of the non-contaminated one (i.e., with large c_D). The sensitivity also increases with the value of the parameter θ, which weighs the effect of the deviation in 'shape' with respect to the deviations in the center.

It is also interesting to note that using $\mathfrak{D}^{\varphi}_{\theta}$ instead of D^{φ}_{θ} scarcely affects the conclusions, regarding the second target.

The **third simulations** will be addressed to the first and third targets. The simulations have been performed as the second ones, but computing the distances by means of the L^1 metric $\mathrm{D}^{\ell}_{\theta}$, where

$$\mathrm{D}^{\ell}_{\theta}(\widetilde{U},\widetilde{V}) = \int_{[0,1]} \left[|\mathrm{mid}\,\widetilde{U}_{\alpha} - \mathrm{mid}\,\widetilde{V}_{\alpha}| + \theta\cdot|\mathrm{spr}\,\widetilde{U}_{\alpha} - \mathrm{spr}\,\widetilde{V}_{\alpha}|\right] d\ell(\alpha).$$

Note that due to the fact that the mid/spr representation of fuzzy numbers does not have associated sufficient conditions to characterize fuzzy numbers, it will not be used hereinafter. When $\theta = 1$, the $\mathrm{D}^{\ell}_{\theta}$ metric reduces to $\boldsymbol{d_1}$ by Klement *et al.* [116].

Secondly, the distance $\mathscr{D}^{\ell}_{\theta}$ will be considered.

		CASE 1			CASE 2		
c_p	c_D	$\widehat{\mathrm{ME}}\,(\theta=1/3)$	$\widehat{\mathrm{ME}}\,(\theta=2/3)$	$\widehat{\mathrm{ME}}\,(\theta=1)$	$\widehat{\mathrm{ME}}\,(\theta=1/3)$	$\widehat{\mathrm{ME}}\,(\theta=2/3)$	$\widehat{\mathrm{ME}}\,(\theta=1)$
0	0	0.000022	0.000036	0.000054	0.000012	0.000015	0.000018
0.1	0	0.014645	0.056992	0.127061	0.000060	0.000153	0.000294
0.1	1	0.040604	0.127047	0.262563	0.007495	0.014028	0.23038
0.1	5	0.311624	0.728550	1.335324	0.203591	0.391977	0.651849
0.1	10	1.039781	2.252823	3.981161	0.794486	1.552579	2.604948
0.1	100	83.41995	167.5429	284.8067	82.74281	161.1959	269.7601
0.2	0	0.055559	0.218843	0.489873	0.000171	0.000520	0.001064
0.2	1	0.165305	0.513345	1.057647	0.029820	0.056282	0.092764
0.2	5	1.250706	2.925262	5.365626	0.791848	1.542378	2.583662
0.2	10	4.177762	8.946755	15.71538	3.193542	6.210516	10.38316
0.2	100	325.989	655.7083	1115.892	333.0015	648.4194	1084.466
0.4	0	0.224389	0.889759	1.996140	0.000552	0.001872	0.003986
0.4	1	0.658307	2.051380	4.232302	0.117735	0.221450	0.364607
0.4	5	5.083451	11.96329	21.99130	3.133536	6.135763	10.30930
0.4	10	16.24498	35.48329	62.95539	13.09717	25.45884	42.52502
0.4	100	1334.663	2637.734	4448.512	1309.402	2573.645	4322.966

Table 1.5: Approximate $\mathrm{D}^{\ell}_{\theta}$-based mean squared error of the Aumann-type sample-mean estimator based on mixed samples

		CASE 1			CASE 2		
c_p	c_D	$\widehat{\mathbf{ME}}\,(\theta=1/3)$	$\widehat{\mathbf{ME}}\,(\theta=2/3)$	$\widehat{\mathbf{ME}}\,(\theta=1)$	$\widehat{\mathbf{ME}}\,(\theta=1/3)$	$\widehat{\mathbf{ME}}\,(\theta=2/3)$	$\widehat{\mathbf{ME}}\,(\theta=1)$
0	0	0.000021	0.000035	0.000054	0.000012	0.000015	0.000018
0.1	0	0.014628	0.056960	0.127013	0.000060	0.000153	0.000294
0.1	1	0.040603	0.127044	0.262560	0.007495	0.010285	0.023038
0.1	5	0.311623	0.728550	1.335323	0.203591	0.391977	0.651849
0.1	10	1.039781	2.252823	3.981160	0.794486	1.552579	2.604948
0.1	100	83.41995	167.5429	284.8067	82.74281	161.1959	269.7601
0.2	0	0.055519	0.218763	0.489753	0.000171	0.000520	0.001064
0.2	1	0.165304	0.513343	1.057644	0.029820	0.056282	0.092763
0.2	5	1.250702	2.925254	5.365614	0.791848	1.542378	2.583662
0.2	10	4.177758	8.946746	15.71536	3.193542	6.210516	10.38316
0.2	100	325.9893	655.7083	1115.892	333.0015	648.4194	1084.466
0.4	0	0.224304	0.889589	1.995886	0.000552	0.001872	0.003986
0.4	1	0.658305	2.051376	4.232296	0.117735	0.221450	0.364607
0.4	5	5.083451	11.96329	21.99130	3.133536	6.135763	10.30930
0.4	10	16.24498	35.48329	62.95539	13.09717	25.45884	42.52502
0.4	100	1334.663	2637.734	4448.512	1309.402	2573.645	4322.966

Table 1.6: Approximate $\mathscr{D}^{\ell}_{\theta}$-based mean squared error of the Aumann-type sample-mean estimator based on mixed samples

The outputs for this simulation study have been collected in Tables 1.5 (for $\mathrm{D}^{\ell}_{\theta}$) and 1.6 (for $\mathscr{D}^{\ell}_{\theta}$). Since one cannot get a general expression of the mean squared errors as a function of θ when the distance is L^1-type, we now present the outputs for three choices of θ (actually, $\theta = 1/3, 2/3, 1$).

Analyzing the last two tables, the empirical conclusion in connection with the first target in this section is that the sample mean is quite sensitive, especially with respect to the changes in c_D. The sensitivity also increases with the value of the parameter θ.

On the other hand, regarding the third target, it is also interesting to note that using $\mathscr{D}_\theta^\varphi$ instead of $\mathrm{D}_\theta^\varphi$ scarcely affects the conclusions (indeed, the difference is so tiny that more digits would be needed to show it).

1.6 Concluding remarks of this chapter

This chapter has presented the preliminary and supporting concepts and results corresponding to the type of data to be dealt with, as well as the basic models and tools to formalize and handle them in a probabilistic/statistical setting. Nevertheless, the statistical literature on robust location measures that have inspired or have been extended/adapted in Chapters 2 and 3 has not been reviewed due to the presupposed background and expertise of the potential readers of this work.

It should be noted that the set- and fuzzy set-valued data as used here could be assumed to be more general, say: unbounded (or having unbounded 0-level in the fuzzy case); non-convex; subsets or fuzzy subsets of infinite dimensional or non-Euclidean spaces, etc. However, for practical purposes, the spaces $\mathcal{K}_c(\mathbb{R}^p)$ and $\mathcal{F}_c(\mathbb{R}^p)$ fit and cover most of the real-life and realistic imprecise data and their mathematical management and analysis are pretty simplified.

It should also be clarified that, although several concepts and results in this chapter are well-known ones from the existing literature, some other ones have been expressly conceived for this work. In this regard, the main contribution in the chapter refers to:

- the wabl/ldev/rdev representation of fuzzy numbers in Definition 1.1.8, along with the associated L^2 and L^1 metrics (Proposition 1.3.5 and Definition 1.3.8, respectively) and properties for all in Propositions 1.1.4, 1.2.10, 1.3.6, 1.3.8 and 1.3.9,

- the support/Steiner representation of fuzzy vectors in Definition 1.1.9 along with the associated L^2 metric in Definition 1.3.5 and properties for both in Propositions 1.1.5, 1.2.11, 1.3.2, 1.3.3, 1.3.4 and 1.4.8,

- the concept of symmetry of a random fuzzy number in Definition 1.4.5 and the related results in Propositions 1.4.9 and 1.4.10,
- and the simulation developments in Section 1.5, which provide valuable arguments to motivate the study in the two other chapters and support the choices of some metrics in them.

The new ideas and results in this chapter have been gathered in several published, accepted or submitted manuscripts, namely, Sinova *et al.* [173, 176, 179, 180, 182, 185].

Chapter 2

Trimmed means for imprecise-valued data

In dealing with real-valued data, an appealing robust measure of location is the class of trimmed means, introduced by Tukey [204] which includes the median as a very special element. In the real-valued settings, trimming entails removing a given percentage of data from the tails of the distribution and computing the mean of the remaining central observations. The median would correspond to the special case in which all, except one or two data, are removed from the tails.

Although one can properly refer to the distribution of the random element generating imprecise-valued data, one cannot properly talk about its 'tails'. As it has already been remarked,

- there is no universally accepted ranking of elements in either $\mathcal{K}_c(\mathbb{R}^p)$ or $\mathcal{F}_c(\mathbb{R}^p)$, so how could we model formally the tails of such a distribution?;
- there are not easy-to-use and realistic models for such a distribution yet.

Fortunately, despite the last drawbacks, trimmed means in Hilbert spaces have already been defined and studied in the literature under some ideal conditions. Furthermore, an algorithm to calculate the trimmed mean estimate of the center of a functional distribution has been proposed.

Cuesta-Albertos and Fraiman [46] have introduced the trimmed mean in the vector space $L^2[0,1] = \{f : [0,1] \to \mathbb{R} : f \text{ is square integrable}\}$, although most of the results apply more generally to uniformly convex Banach spaces. Theoretical properties of the trimmed mean, like the existence, consistency of the corresponding estimator and qualitative robustness under some conditions, have been examined and discussed. Moreover, Cuesta-Albertos and Fraiman [46] have proposed an algorithm that approximates the empirical trimmed mean associated with a sample by obtaining an element of the sample that converges to the trimmed mean whenever this value is in the support of the considered distribution. However, this algorithm is not applicable for large data sets given its computational complexity and it is clearly theoretically oriented, while our interest is not only theoretical, but also practical.

This chapter first focuses on the trimmed means in separable Hilbert spaces. As a consequence, these results and conclusions will be applicable not only to functional data, but also to set- and fuzzy set-valued ones.

The performance of trimmed means will be discussed paying attention to contamination models. The problem of finding the empirical trimmed mean is discussed and a new algorithm is proposed to find the solution in a more efficient way. This algorithm is inspired by the FAST-LTS (see Rousseeuw and Van Driessen [158]). The effectiveness and efficiency of this algorithm will be shown to be higher than those for the Cuesta-Albertos and Fraiman algorithm theoretically and empirically. One of the most important measures of robustness, the finite sample breakdown point, will be considered in order to analyze the behaviour of the new estimator of the trimmed mean.

Since there are other robust location measures in functional Hilbert spaces, such as the usual functional median or the trimmed mean based on depth functions (see Cuevas and Fraiman [51] and Cuesta-Albertos and Nieto-Reyes [49]), a comparison with representative estimators of this type will be also shown. Despite the fact that both the sample mean and the trimmed means will be compared in all the examples and applications to unify the studies, it is important to note that, although the trimmed means are estimators of the population mean in the location-scale model, this is not true in a general Hilbert space because it lacks symmetry, which is necessary for the population trimmed means to coincide with the population mean.

In Section 2.1, the population trimmed mean for a Hilbert space-valued random element is presented, and the already known theoretical results related to its existence and uniqueness are recalled. Section 2.2 is devoted to the empirical trimmed mean and it introduces the proposed algorithm for its computation and discusses some of its main properties, such as the consistency and the finite sample breakdown point. The suggested algorithm will be compared with the existing one by Cuesta-Albertos and Fraiman, both theoretically and empirically, in Section 2.3. The performance against estimators of this type will be analyzed in Section 2.4. Applications to fuzzy and set-valued data are shown in Section 2.5. The chapter ends with a Section 2.6 containing the summary of the novelties and the related publications derived from it.

2.1 Population trimmed means for Hilbert space-valued random elements

In this section, the population trimmed mean of a Hilbert space-valued random element, along with some remarks about its existence and uniqueness, is recalled (for details, see Cuesta-Albertos and Fraiman [46]).

Definition 2.1.1. [46] *Consider a random element $X : \Omega \to \mathbb{H}$, where $(\Omega, \mathcal{A}, P)$ is a probability space, $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$ is a separable Hilbert space and P_X is the induced probability distribution on the Borel σ-algebra on $\mathbb{H}$. For any $\beta \in (0,1)$, the corresponding* ***population trimmed mean*** *is any $h_{P_X} \in \mathbb{H}$ such that there exists a* ***trimming function*** *$\tau_{P_X} \in \mathcal{P}_\beta$, where*

$$\mathcal{P}_\beta = \left\{ \tau : \mathbb{H} \to [0,1] \, : \, \tau \,\text{measurable}, \int_{\mathbb{H}} \tau(x)\, dP_X(x) \geq 1 - \beta \right\},$$

satisfying that

$$\int_{\mathbb{H}} \|x - h_{P_X}\|^2_{\mathbb{H}} \tau_{P_X}(x)\, dP_X(x) = \inf_{h\in \mathbb{H},\, \tau\in \mathcal{P}_\beta} \int_{\mathbb{H}} \|x - h\|^2_{\mathbb{H}} \tau(x)\, dP_X(x).$$

Note that this definition generalizes the notion of trimmed means based on trimming regions, since that situation is equivalent to using as trimming function only indicator functions of sets, i.e.,

$$\min_{A\in \mathcal{E}} \int_A \|x - E(X|A)\|^2_{\mathbb{H}}\, dP_X(x) = \min_{A\in \mathcal{E},\, h\in \mathbb{H}} \int_A \|x - h\|^2_{\mathbb{H}}\, dP_X(x),$$

where $\mathcal{E} = \{A \subset \mathbb{H} \, : \, P_X(A) = 1 - \beta\}$.

However, it turns out that the best trimming function essentially coincides with the indicator function of a set (as proven in Cuesta-Albertos *et al.* [47] for random vectors and in Cuesta-Albertos and Fraiman [46] for Hilbert space-valued random elements), so we can restrict ourselves to this case.

Remark 2.1.1. The existence of the trimmed mean and sufficient conditions for its uniqueness have already been established in Cuesta-Albertos and Fraiman [46]. These ideal conditions do not cover contamination models and general sample distributions in robust statistics.

To illustrate this assertion one can consider the following example. The ideal conditions guaranteeing uniqueness in [46] require the existence of a point m_0 in the functional space $L^2[0,1]$ such that $P(\overline{B}(m_0, r)) > P(\overline{B}(m, r))$ for all m in $L^2([0,1])$ and all nonnegative real values r.

Let X be a continuous $L^2[0,1]$-valued random element satisfying this assumption. Choose ϵ and $r > 0$ such that $P(\overline{B}(m_0, r)) < \epsilon$. If a fraction ϵ of outliers is now inserted at $\overline{B}(m_1, r)$ with $m_1 \neq m_0$, then $P(\overline{B}(m_1, r)) \geq \epsilon$ and thus the uniqueness condition is not satisfied anymore.

2.2 Empirical trimmed means for Hilbert space-valued data

This section aims to analyze the consistency and robustness of a new trimmed mean estimator.

Let $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$ be a separable Hilbert space, $(\Omega, \mathcal{A}, P)$ be a probability space, $X : \Omega \to \mathbb{H}$ be a continuous random element and $\mathbf{h_n} = (h_1, \ldots, h_n)$ be a sample of independent observations.

For each $\beta \in (0,1)$, we consider the problem of finding the ***β-trimmed mean*** (or simply trimmed mean) and the ***trimming region*** $\widehat{E}$, given by

$$\widehat{E} = \arg\min_{E \in \mathcal{E}} \frac{1}{n_\beta} \sum_{i \in E} \left\| h_i - \frac{1}{n_\beta} \sum_{j \in E} h_j \right\|_{\mathbb{H}}^2 = \arg\min_{E \in \mathcal{E}} \operatorname{Var}(\mathbf{h_n} | E)$$

where $n_\beta = n - [n\beta]$ is the cardinal of the trimming region $\widehat{E}$ and the set

$$\mathcal{E} = \{E \subset \{1, \ldots, n\} \, : \, \#E = n_\beta\}$$

consists of all the subsets of n_β different natural numbers which are up to the sample size. Therefore, the trimming region, for a fixed proportion of trimming, can be seen as the set containing the remaining proportion of sample data with minimum variance.

This problem has at least one solution because it is finite combinatorial. Once a trimming region $\widehat{E}$ is determined, the associated trimmed mean and variance are simply the mean and variance of the sample conditioned to $\widehat{E}$.

The proposed algorithm to find one of the solutions for the trimming problem is to be inspired on the common strategy in real-valued settings. This strategy makes an initial choice of the trimming region (the simplest way would be a random choice); then, this region is adjusted with a method to construct new trimming regions which are more concentrated than the previous ones (see Rousseeuw and Van Driessen [158]). Following a similar reasoning for the Hilbert case, it is easy to prove that the variance decreases as the set of the n_β closest observations to the mean of the original region is considered as the new region. The theorem below formalizes this theoretical conclusion:

Theorem 2.2.1. *Let $E_1 \in \mathcal{E}$ and denote by $E_2 \in \mathcal{E}$ a set of indices corresponding to the n_β observations with smallest distances to $\sum_{i\in E_1} h_i/n_\beta$. Then,*

$$\mathrm{Var}(\mathbf{h_n}|E_2) \leq \mathrm{Var}(\mathbf{h_n}|E_1),$$

where the equality holds if and only if $E_1 = E_2$.

Proof. Since E_2 corresponds to the indices of the n_β closest observations to $\frac{1}{n_\beta}\sum_{j\in E_1} h_j$, we have that:

$$\frac{1}{n_\beta}\sum_{i\in E_2} \|h_i - \frac{1}{n_\beta}\sum_{j\in E_1} h_j\|^2_{\mathbb{H}} \leq \frac{1}{n_\beta}\sum_{i\in E_1} \|h_i - \frac{1}{n_\beta}\sum_{j\in E_1} h_j\|^2_{\mathbb{H}} = \mathrm{Var}(\mathbf{h_n}|E_1).$$

It is well-known that the mean minimizes the sum of the squared distances to all the considered observations, so:

$$\mathrm{Var}(\mathbf{h_n}|E_2) = \frac{1}{n_\beta}\sum_{i\in E_2} \|h_i - \frac{1}{n_\beta}\sum_{j\in E_2} h_j\|^2_{\mathbb{H}} \leq \frac{1}{n_\beta}\sum_{i\in E_2} \|h_i - \frac{1}{n_\beta}\sum_{j\in E_1} h_j\|^2_{\mathbb{H}}$$

and the inequality follows. □

As a consequence, to compute the trimmed mean of a sample we can consider the following strategy:

Algorithm to compute the trimmed mean of a sample (ETMA)

Step 1. Set $n_\beta = n - \lfloor n\beta \rfloor \in \{1, \ldots, n\}$ ($\lfloor\cdot\rfloor$ = floor function), the size of the trimming region, fix **NS** the number of starting points and **nbest** the number of best trimming regions selected after **nrep** initial steps. Initialize $MSE = \infty$;

Step 2. Choose at random either three data from the sample (considered a seed) and compute its mean, f, or an observation considered as an initial mean, f. It is then possible to build a first region of size n_β centered around f. Since the three data or the observation are chosen randomly, this first region will be also random;

Step 3. Select the n_β closest data to the mean f (where closeness refers to the distance corresponding to the norm associated with the inner product of the Hilbert space):

$$\{h_{k_1}, \ldots, h_{k_{n_\beta}}\};$$

Step 4. Compute the mean f^* of the n_β observations in *Step 3*, and then calculate the corresponding mean squared error given by

$$f^* = \sum_{i=1}^{n_\beta} h_{k_i}/n_\beta, \ MSE^* = \sum_{i=1}^{n_\beta} \|h_{k_i} - f^*\|^2_{\mathbb{H}}/n_\beta;$$

Update the value of the upper bound, that is,

$$MSE = MSE^*;$$

Step 5. *Steps 3* and *4* are repeated **nrep** times;

Step 6. Repeat *Steps 2-5* **NS** times and choose the **nbest** trimming regions with lowest associated MSE^*;

Step 7. For each of the **nbest** trimming regions, repeat *Steps 3-4* until convergence;

Step 8. The estimate associated with the smallest MSE^* will be the final estimate of the trimmed mean. Moreover, the corresponding value MSE^* will be the trimmed Mean Squared Error associated with it.

According to Theorem 2.2.1, one can easily conclude that the convergence of this iteration process is always reached after a finite number of iterations. Furthermore, due to *Step 2*, this new algorithm always performs at least as well as the Cuesta-Albertos and Fraiman alternative.

2.2.1 Consistency and robustness of the ETMAs

Some properties of the empirical trimmed mean as estimator of its population value can be analyzed now. The following result is similar as in [46] and shows the strong consistency of the sample trimmed mean as estimator of the population trimmed mean under uniqueness conditions.

Theorem 2.2.2. *Let $\beta \in (0,1)$ and assume that the probability measure P_X has a unique trimmed mean parameter. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of i.i.d. random elements with distribution P_X, and denote by P_n the empirical probability measure, that is,*

$$P_n(\omega) := \frac{1}{n}\sum_{i\leq n}\delta_{X_i(\omega)} \text{ for all } n\in\mathbb{N},\ \omega\in\Omega,$$

with δ_x being the Dirac delta measure on x, and h_n^ω being any of its corresponding empirical trimmed means. Then,

$$\lim_{n\to\infty}\|h_n^\omega - h_{P_X}\|_{\mathbb{H}} = 0 \ a.s.\,[P].$$

It should be emphasized that the difference between the consistency for the estimator from the new algorithm and the one from [46] lies in the fact that the latter does not search for the point h_n^ω, but for the (also consistent) estimate $\widehat{m}_{k_n}^\omega$, which is the result of constraining to the support of P_n the search of the minimum in computing the trimmed mean.

The breakdown point of h_n^ω as estimator of the empirical trimmed mean is derived now. As indicated by Cuevas *et al.* [50] this notion, originally introduced by

Hampel [99] and formalized later in the current way by Donoho and Huber [59], can be adapted to estimators taking values in general metric spaces. Following Donoho and Huber, the finite sample breakdown point (denoted fsbp) of the sample trimmed mean corresponding to a sample of size n from a random element $X : \Omega \to \mathbb{H}$, where $(\Omega, \mathcal{A}, P)$ is a probability space and $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$ is a separable Hilbert space, is given by

$$\text{fsbp}(\widehat{(h_{P_X})}_n, \mathbf{h_n}, d) = \frac{1}{n} \min \left\{ k \in \{1, \ldots, n\} \, : \, \sup_{Q_{n,k}} d(h_{P_n}, h_{Q_{n,k}}) = \infty \right\},$$

where $\mathbf{h_n}$ denotes the considered sample of n observations from the metric space $(\mathbb{H}, d)$ (being d the distance associated with the inner product) in which $\sup_{h,h' \in \mathbb{H}} d(h, h') = \infty$, P_n is the empirical distribution of $\mathbf{h_n}$ and $Q_{n,k}$ is the empirical distribution of sample $\mathbf{y_{n,k}}$ obtained from the original sample $\mathbf{h_n}$ by perturbing up to k of its elements. Then, we have that

Theorem 2.2.3. *Consider a continuous random element $X : \Omega \to \mathbb{H}$ such that*

$$\sup_{h,h' \in \mathbb{H}} d(h, h') = \infty$$

(with d the distance associated with the inner product), where $(\Omega, \mathcal{A}, P)$ is a probability space, $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$ is a separable Hilbert space and P_X is a fixed probability distribution on the Borel σ-algebra on $\mathbb{H}$. For any $\beta \in (0,1)$, the finite sample breakdown point of the corresponding sample trimmed mean equals

$$\text{fsbp}\big(\widehat{(h_{P_X})}_n, \mathbf{h_n}, d\big) = \frac{\lfloor n \cdot \beta \rfloor + 1}{n}.$$

Proof. First note that

$$\min \left\{ k \in \{1, \ldots, n\} \, : \, \sup_{Q_{n,k}} d(h_{P_n}, h_{Q_{n,k}}) = \infty \right\} \leq \lfloor n \cdot \beta \rfloor + 1.$$

Indeed, if there are at least $\lfloor n \cdot \beta \rfloor + 1$ perturbed points ($k \geq \lfloor n \cdot \beta \rfloor + 1$ fixed) which lie arbitrarily far from the original observations, then any possible trimming region of size $n_\beta = n - \lfloor n \cdot \beta \rfloor$ must contain at least one of these 'contaminated' points and the estimate of the trimmed mean (the mean over the trimming region), $h_{Q_{n,k}}$, will be arbitrarily far from the original observations too, so that $\sup_{Q_{n,k}} d(h_{P_n}, h_{Q_{n,k}}) = \infty$.

Now, it is proved that

$$\sup_{Q_{n,k}} d(h_{P_n}, h_{Q_{n,k}}) < \infty$$

for any $k \leq \lfloor n \cdot \beta \rfloor$, so that the other inequality also holds. For this purpose two type of regions can be distinguished, namely,

- The regions of size n_β with no contaminated points and, therefore, a finite trimmed mean and variance.
- The regions of size n_β containing at least one contaminated point, with trimmed mean arbitrarily far away and arbitrarily big variance.

Therefore, the trimming region minimizing the corresponding trimmed mean squared error will correspond to the first type: it will not contain any contaminated points and its trimmed mean will be finite, so

$$\sup_{Q_{n,\lfloor n\cdot\beta\rfloor}} d(h_{P_n}, h_{Q_{n,\lfloor n\cdot\beta\rfloor}}) < \infty. \qquad \square$$

Clearly, the finite sample breakdown point tends to β as n tends to ∞, so the asymptotic breakdown point is β.

2.3 Comparative study between the ETMA and the Cuesta-Albertos and Fraiman algorithm

In this section the algorithm introduced in Section 2.2 is compared to the algorithm of Cuesta-Albertos and Fraiman. As it has already been pointed out, whereas the new algorithm searches for the empirical trimmed mean, the algorithm in [46] selects an element of the sample converging to the trimmed mean whenever it belongs to the support of the distribution.

Table 2.1 compares the theoretical complexity of both algorithms by distinguishing all the involved steps. Steps of the Empirical Trimmed Mean Algorithm (ETMA) have been detailed in Section 2.2. The *steps for the algorithm by Cuesta-Albertos and Fraiman* (C&F) can be summarized as follows:

Step 1. Fix $0 < n_\beta \leq n$;

Step 2. Compute the squared distances between all the observations;

Step 3. Compute the mean and MSE corresponding to the n_β closest elements of each observation;

Step 4. The outcome is the observation with minimal MSE of the associated region (consisting of the n_β closest points to this observation).

Denote by $\mathbf{Q}$ and $\mathbf{q}$ the costs of computing the squared norms and the means, respectively, and by lt the number of evaluations $MSE^* < MSE$, then the complexity of the algorithms is gathered in Table 2.1.

Step	ETMA	C&F
1	-	-
2	-	$\approx \mathbf{Q} \times \mathbf{n^2}$
3	$\mathbf{NS} \times \mathbf{n} \log \mathbf{n}$	$\approx \mathbf{n^2} \log \mathbf{n}$
4	$\approx \mathbf{NS} \times (\mathbf{Q} + \mathbf{q}) \times \mathbf{n}$	-
5	$\approx \mathbf{NS} \times ((\mathbf{Q} + \mathbf{q}) \times \mathbf{n}$ $+\mathbf{n} \log \mathbf{n}) \times \mathbf{nrep}$	It does not apply
6	$\mathbf{NS} \log \mathbf{NS}$	It does not apply
7	$\approx \mathbf{nbest} \times ((\mathbf{Q} + \mathbf{q}) \times \mathbf{n}$ $+\mathbf{n} \log \mathbf{n}) \times \mathbf{nrep}$	It does not apply
8	$\mathbf{nbest} \log \mathbf{nbest}$	It does not apply
Total	$\mathrm{O}(\mathbf{n} \log \mathbf{n})$	$\mathrm{O}(\mathbf{n^2} \log \mathbf{n})$

Table 2.1: Complexity of the structure of the two algorithms, ETMA and C&F

The cost $\mathbf{Q}$ of computing the squared norms is obviously much bigger than $\mathbf{q}$, the cost of computing the means only using additions and division. For that reason, when the sample size n is small, the influence of $\mathbf{Q}$ in this calculus is enormous. Because of this, computing the distances between all the observations in Cuesta-Albertos and Fraiman's algorithm is very expensive from the computational point of view. On the other hand, these squared distances are only computed between each observation and a reference element, the mean, in the proposed algorithm (and the same happens with the quicksort function). The new algorithm is also more efficient, unless the sample size is smaller than the constant $\mathbf{NS}$, since its computational complexity ($\mathrm{O}(\mathbf{n} \log \mathbf{n})$) is smaller than the one of the existing algorithm ($\mathrm{O}(\mathbf{n}^2 \log \mathbf{n})$). Otherwise, the Empirical Trimmed Mean Algorithm is clearly not faster than the Cuesta-Albertos and Fraiman alternative.

To illustrate this assertion, and support the effectiveness of the ETMA by reaching a more desirable solution, we consider an example.

Example 2.3.1. The considered database, which can be found in the R package fda, is 'Canadian average annual weather cycle'.

It collects the daily temperature and precipitation at 35 different locations in Canada averaged over 1960 to 1994 (see Ramsay and Silverman [162] and Figure 2.1 on the left).

The trimmed mean is not computed for these functions, but for their smoothed version estimated using the 65-element Fourier basis (see Figure 2.1 on the right), which is common for functional data.

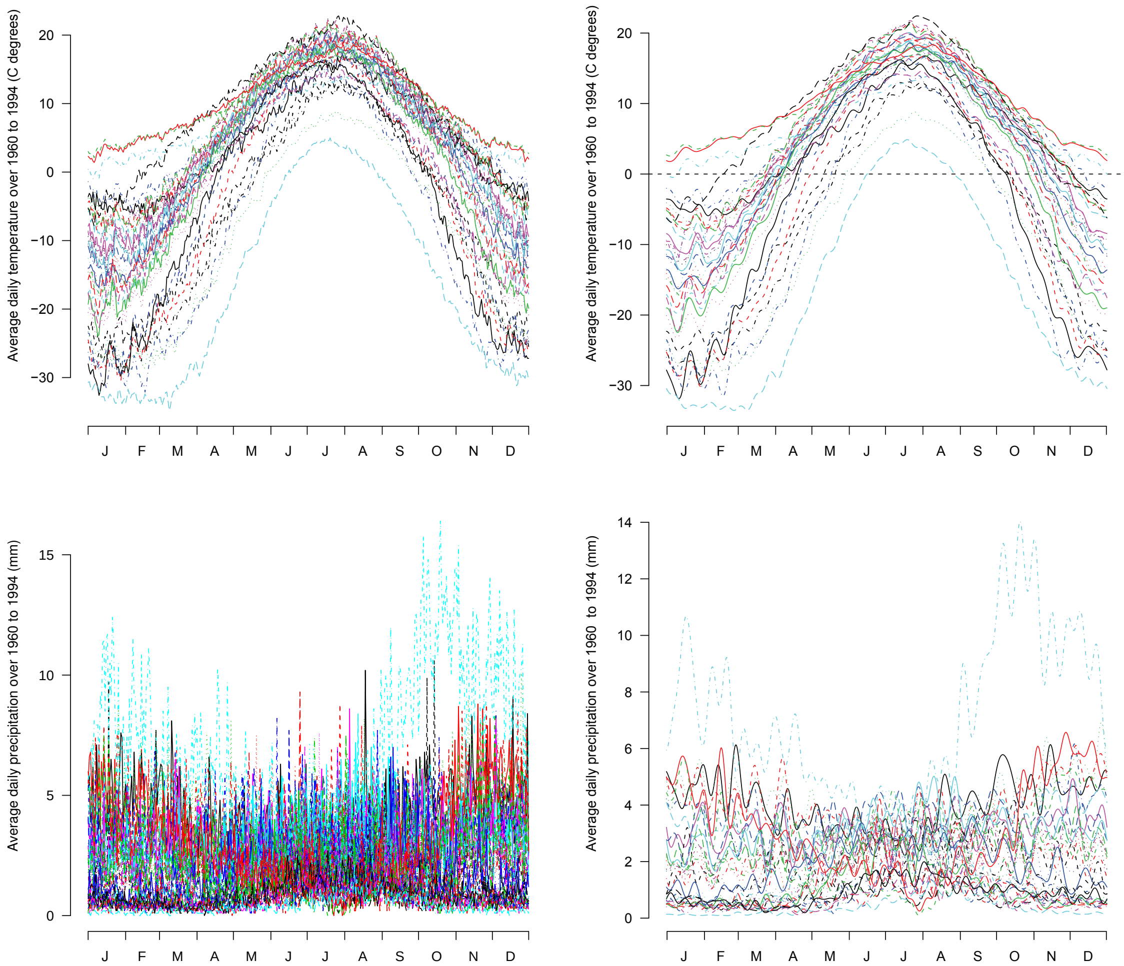


Figure 2.1: On the left, average daily temperature and precipitation in 35 different Canadian locations over 1960 to 1994; on the right, the smoothed functions for the average daily temperature and precipitation over 1960 to 1994

The 10%, 20% and 30% trimmed means for this smoothed database are now computed by using the Empirical Trimmed Mean Algorithm 400 times and the Cuesta-Albertos and Fraiman just once. For each of them, a graph of the smoothed sample observations and both estimates is displayed in Figure 2.2 for the average daily temperature and in Figure 2.3 for the average daily precipitation.

Cuesta-Albertos and Fraiman's estimate always refers to the same optimal sample element after having compared all the possibilities, whereas the estimate given by the ETMA is the average of the corresponding fraction of most concentrated data, reached at different times depending on the starting point.

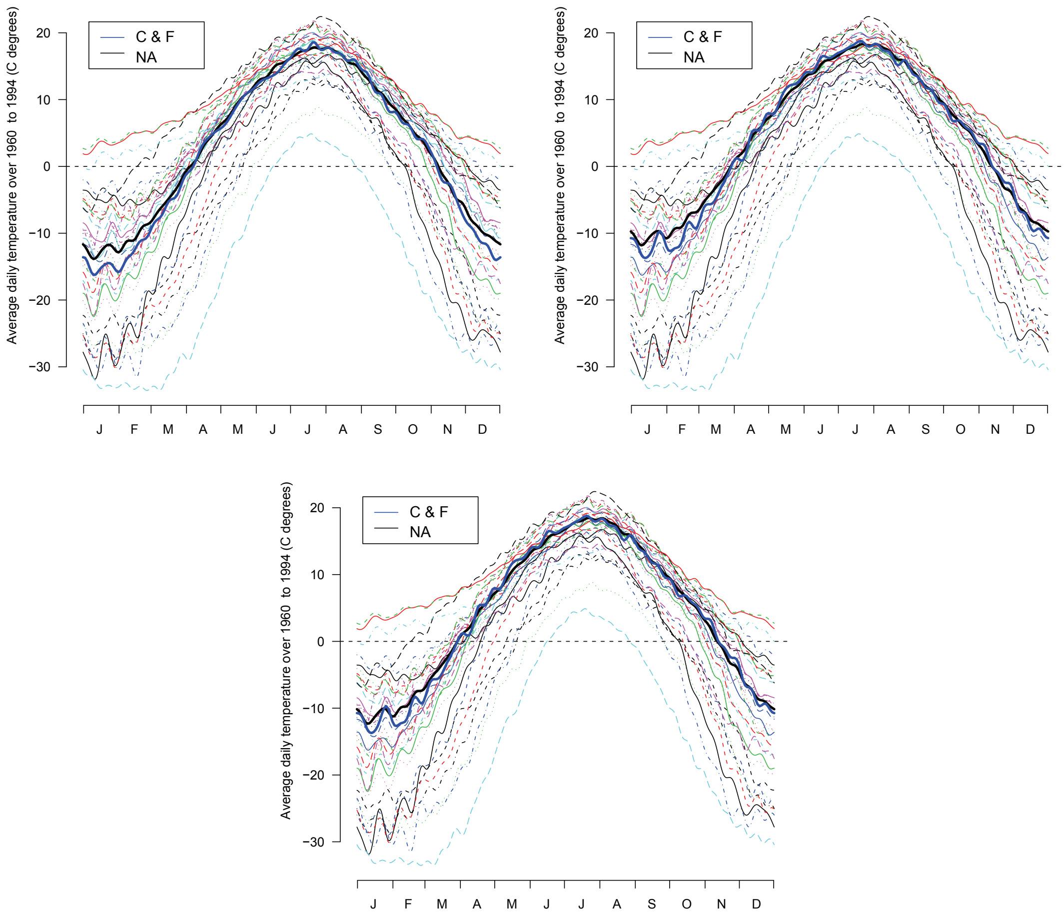


Figure 2.2: The smoothed database for the average daily temperature and the 10% (top left), 20% (top right) and 30% (bottom) trimmed means computed using the ETMA and the C&F algorithm

Moreover, this can be seen in Tables 2.2 and 2.3 where the computation time, the mean squared error (MSE) in the trimming region and the number of distances computed in both cases are gathered.

As it has been just shown, the algorithm in Section 2.2 reaches its solution faster (in mean) than Cuesta-Albertos and Fraiman's. This relates to the number of distances to be computed, which is smaller for the new algorithm even in the worst case in which the initial seed of the algorithm lies far from the true value of the trimmed mean.

Furthermore, in this example it turns out that for all random starts the new algorithm converged to a better solution since its corresponding trimmed mean squared error is smaller than for Cuesta and Fraiman's solution (actually, C&F algorithm does not search for the sample trimmed mean, but for the element in the sample minimizing the trimmed mean squared error).

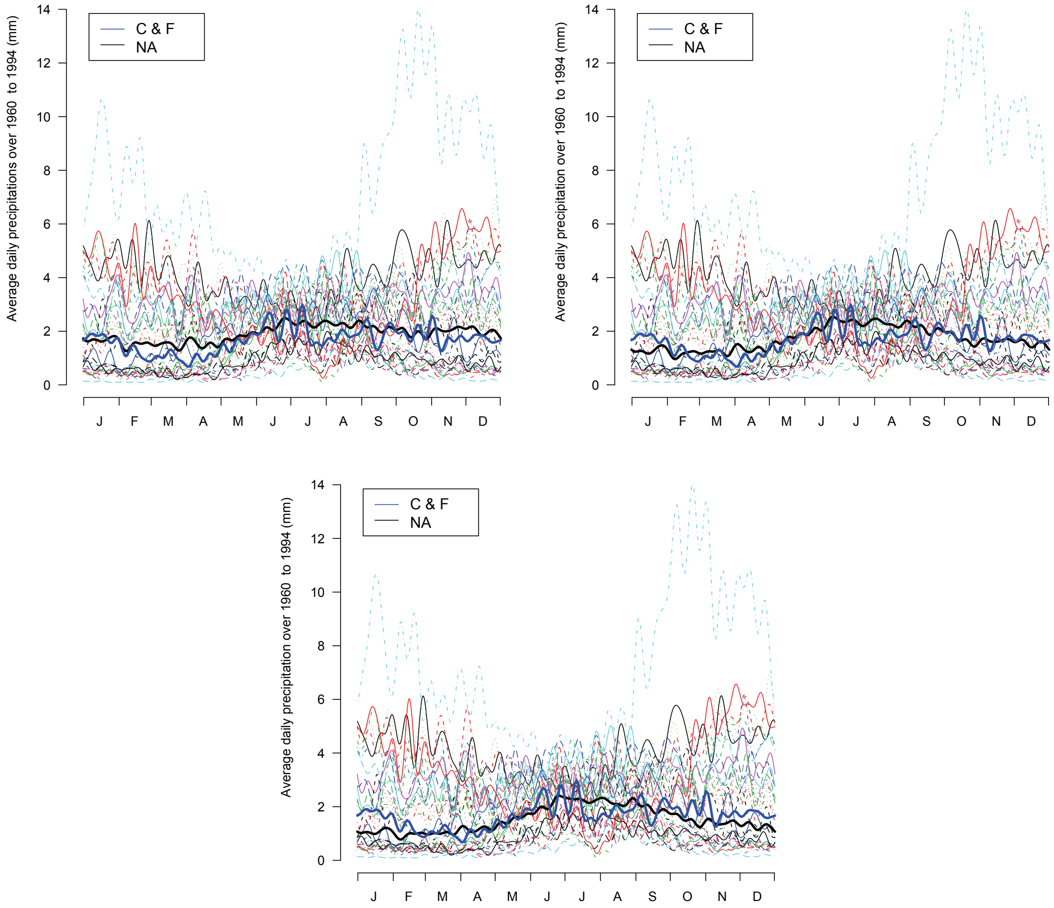


Figure 2.3: The smoothed database for the average daily precipitation and the 10% (top left), 20% (top right) and 30% (bottom) trimmed mean computed using the ETMA and the C&F algorithm

Algorithm	ETMA			C&F		
Trimming	10%	20%	30%	10%	20%	30%
Time	1.56 s	2.15 s	2.12 s	8.01 s	7.97 s	8.04 s
MSE	10797	6490	4267	11128	6852	4551
Distances	105 (66.75%) 140 (33.25%)	105 (37.75%) 140 (9.50%) 175 (18.25%) 210 (32.25%) 245 (4.25%)	105 (24.50%) 140 (21.75%) 175 (38.50%) 210 (15.25%)	630	630	630

Table 2.2: Computation time, mean squared error (MSE) and number of distances (in parentheses, the percentage over the 400 times this number has been obtained) computed for the ETMA and the C&F algorithms when estimating the 10%, 20%, and 30% trimmed means for the average daily temperature

Algorithm	ETMA			C&F		
Trimming	10%	20%	30%	10%	20%	30%
Time	2.33 s	2.29 s	2.65 s	10.09 s	10.01 s	10.04 s
MSE	537	369	280	593	422	342
Distances	105 (31.25%) 140 (28.25%) 175 (18.25%) 210 (40.50%)	105 (28.25%) 140 (40.75%) 175 (28.00%) 210 (3.00%)	105 (31.00%) 140 (18.50%) 175 (18.25%) 175 (30.25%) 245 (2.00%)	630	630	630

Table 2.3: Computation time, mean squared error (MSE) and number of distances (in parentheses, the percentage over the 400 times this number has been obtained) computed for the ETMA and the C&F algorithms when estimating the 10%, 20%, and 30% trimmed means for the average daily precipitation

Similar behaviour is observed for the case of the average daily precipitation when computing the 10%, 20% and 30% trimmed means (see Figure 2.3 and Table 2.3).

Finally, the last graph for the daily temperature in the Canadian locations shows that the results are almost the same when some outliers occur in the sample. Figure 2.4 shows the estimates when the three first coefficients (the most important ones) of the seven first functions (20% of the sample) have been reduced to half of its original value. It can be seen that, although some outliers are present, the estimate does not change much because of its robustness.

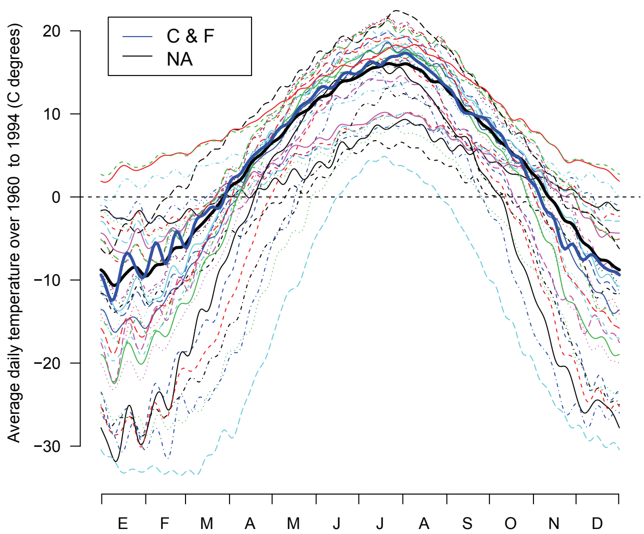


Figure 2.4: The smoothed database for the contaminated average daily temperature and the 20% trimmed mean computed using the ETMA and the C&F algorithm

2.4 Simulation studies about trimmed means for Hilbert space-valued data

This section develops a simulation study for functional data following the scheme of this kind of comparisons in the literature (see, for instance, Fraiman and Muniz [75] and López-Pintado and Romo [125]):

- $N = 500$ replications in order to compute the mean squared error;
- $n = 50, 80$ sample size considered, that is to say, number of curves generated;
- $q = .05, .1$ contamination proportion;
- $M = 5, 25$ contamination magnitude;
- $\beta = .2$ trimming proportion;
- $T = 30$ number of equidistant points chosen as a partition of [0.1].

The considered models analyze different types of outliers. Since a curve is an outlier if it is generated in a different way than the other curves in the sample, this could happen due to many reasons. In *Model 1* there is no contamination. *Models 2-5* contain magnitude outliers, that is, curves that are really far from the mean value. *Models 6-9* contain shape outliers that are not necessarily far from the mean, but present a different shape or pattern. Only curves defined on the domain $[0, 1]$ are considered. More concretely, concerning *Models 1-5*, the situations are the following:

- *Model 1:* In this model, the curves, represented by $X_i(t)$ for $i = 1, \ldots, n$, are generated following the distribution:

$$X_i(t) = 4t + e_i(t),$$

  where $e_i(t)$ is a Gaussian stochastic process with mean 0 and covariance function $\gamma(s, t) = e^{-|t-s|}$.

- *Model 2:* Symmetric contamination obtained by generating the curves:

$$Y_i(t) = X_i(t) + \varepsilon_i \sigma_i M, \quad 1 \leq i \leq n$$

  where $\{\varepsilon_i\}_{i=1}^n$ and $\{\sigma_i\}_{i=1}^n$ are independent sequences of random variables following a Bernoulli distribution $B(q)$ and a discrete uniform on $\{-1, 1\}$ distributions, respectively.

- *Model 3:* Asymmetric contamination given by the curves:

$$Y_i(t) = X_i(t) + \varepsilon_i M, \quad 1 \leq i \leq n.$$

- *Model 4:* Partial (trajectories) contamination represented by the curves:

$$Y_i(t) = \begin{cases} X_i(t) + \varepsilon_i \sigma_i M, & \text{if } t \geq T_i \\ X_i(t), & \text{if } t < T_i \end{cases},$$

  where $1 \leq i \leq n$ and the corresponding T_i is a random number generated from an $\mathcal{U}(0, 1)$ distribution.

- *Model 5:* Peaks contamination introduced by the expression:

$$Y_i(t) = \begin{cases} X_i(t) + \varepsilon_i \sigma_i M, & \text{if } T_i \leq t \leq T_i + l \\ X_i(t), & \text{if } t \notin [T_i, T_i + l] \end{cases},$$

  for $1 \leq i \leq n$, $l = 2/30$ and the corresponding T_i is a random number generated from an $\mathcal{U}(0, 1 - l)$ distribution.

For the remaining situations (containing shape outliers), a basic model has been considered, namely,

$$X_i(t) = g(t) + e_{1i}(t), \quad 1 \leq i \leq n,$$

where the function $g(t)$ is either $g(t) = 4t$ or $g(t) = 4t^2$ and $e_{1i}(t)$ is the Gaussian process with mean 0 and covariance function $\gamma_1(s, t) = e^{-|t-s|^2}$.

The contamination is introduced by mixing this basic model with:

$$Y_i(t) = g(t) + e_{2i}(t), \quad 1 \leq i \leq n,$$

where now the Gaussian process $e_{2i}(t)$ still has mean 0, but its covariance function is $\gamma_2(s, t) = k \cdot e^{-c|t-s|^\mu}$, with nonnegative parameters k, c and μ (for more details about this family of models, see Wood and Chan [217]). The role of these parameters is to control the shape of the curves: when increasing μ and k the generated functions are softer, whereas when increasing c, they become more irregular. The contaminated models then are

$$Z_i(t) = (1 - \varepsilon) X_i(t) + \varepsilon Y_i(t), \quad 1 \leq i \leq n,$$

where ε follows a Bernoulli distribution $B(q)$. The different choices for the parameters are the following:

- *Model 6:* $k = 1$, $c = 1$, $\mu = .2$ and $g(t) = 4t$.
- *Model 7:* $k = 1$, $c = 1$, $\mu = .1$ and $g(t) = 4t$.
- *Model 8:* $k = 1$, $c = 1$, $\mu = .2$ and $g(t) = 4t^2$.
- *Model 9:* $k = 1$, $c = 1$, $\mu = .1$ and $g(t) = 4t^2$.

For any of the 500 iterations, the integrated squared error has been calculated in *Models 1-5* as follows:

$$E(j) = \frac{1}{T}\sum_{k=1}^{T}\left[\hat{g}_n\left(\frac{k}{T}\right) - g\left(\frac{k}{T}\right)\right]^2,$$

where $\hat{g}_n$ represents any of the proposed estimators (mean, trimmed mean, trimmed mean defined through a depth or trimmed mean computed using the Cuesta-Albertos and Fraiman's algorithm).

In *Models 6-9*, instead of using the Euclidean distance to compute both the trimmed means and the integrated error, the distance defined by means of the norm on the Sobolev space $W^{1,2}((0,1))$ is chosen. The reason is that shape outliers are not properly identified when computing the Euclidean distances because they are quite close to the mean value. However, the distance on the Sobolev space is defined not only involving the Euclidean distance between the functions, but also the Euclidean distance between their corresponding derivatives, so it is more practical for finding differences due to the shape of the functions.

Finally, the sample depth considered to sort all the curves and choose the deepest ones is:

$$D(X_{i_0}(t))_n = 1 - \frac{1}{T}\sum_{t=0}^{T}\left|\frac{1}{2} - \frac{1}{n}\sum_{i=1}^{n} I_{(-\infty, X_{i_0}(t)]}(X_i(t))\right|,$$

for any fixed $i_0 \in \{1,\dots,n\}$ and $I_{(-\infty,X_{i_0}(t)]}$ denoting the indicator function of the set $(-\infty, X_{i_0}(t)]$.

Therefore, after sorting (from deeper to less deep) the curves, the sample depth trimmed mean is computed as:

$$\hat{\mu}^D_{n,\beta}(t) = \frac{\sum_{i=1}^{n-\lfloor n\beta\rfloor} X^{(i)}(t)}{n - \lfloor n\beta\rfloor}.$$

Obviously, the sample mean is $\hat{\mu}_n(t) = \sum_{i=1}^{n} X_i(t)/n$. The remaining estimators have previously been introduced.

To compare the behaviour of these estimators, the mean squared error (MSE) and its standard deviation (s) have been computed:

$$MSE = \frac{1}{N}\sum_{j=1}^{N} E(j), \quad s = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(E(j) - MSE)^2}.$$

Figure 2.5 shows an example of the sample of curves generated in one of the 500 replications considering each contamination model.

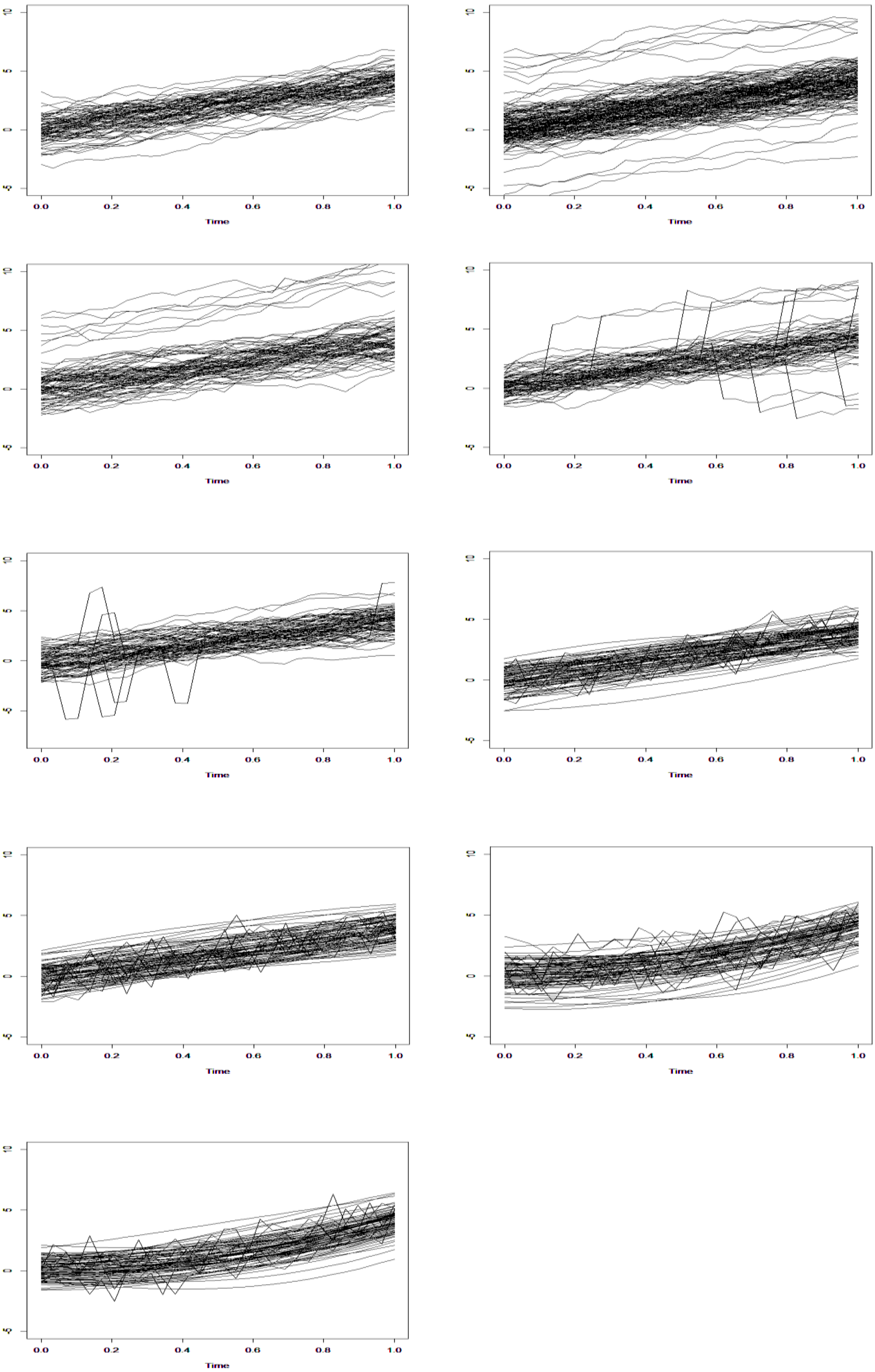


Figure 2.5: An example of the sample of curves generated in each of the Models 1-9 (starting on the top left corner to the right and down)

n	*q*	*M*	Estimator	*Model 1*	*Model 2*	*Model 3*	*Model 4*	*Model 5*
50	.05	5	Mean	**.019443** (.019470)	.045089 (.058401)	.100940 (.117672)	.033178 (.033396)	**.020686** (.018855)
			ETMA	.041967 (.044412)	.036431 (.042972)	.037849 (.041294)	.038451 (.046663)	.042351 (.047767)
			DTM	.027283 (.028692)	**.028227** (.030984)	**.027745** (.031989)	**.031958** (.033687)	.028919 (.027345)
			C&F	.235814 (0.219801)	.280490 (.285887)	.289691 (.313762)	.279875 (.266319)	.273206 (.275933)
80	.05	5	Mean	**.012736** (.014075)	.028164 (.037772)	.080471 (.030194)	.020019 (.025748)	**.013512** (.029245)
			ETMA	.028059 (.030684)	.023724 (.030090)	.019357 (.030194)	.019412 (.025748)	.017225 (.029245)
			DTM	.016765 (.017203)	**.016770** (.018414)	**.019357** (.020912)	**.019412** (.019376)	.017225 (.018445)
			C&F	.179295 (.120392)	.207068 (.182755)	.190714 (.124006)	.188384 (.109506)	.196137 (.196500)
50	.1	5	Mean	**.019443** (.019470)	.074806 (.093568)	.316388 (.268233)	.046796 (.052995)	**.022762** (.022055)
			ETMA	.041967 (.044412)	.037588 (.044766)	**.031499** (.035569)	**.034751** (.037954)	.037522 (.040879)
			DTM	.027283 (.028692)	**.026706** (.028457)	.044885 (.061108)	.038813 (.040508)	.030920 (.032898)
			C&F	.235814 (0.219801)	.404505 (.426930)	.344146 (.402058)	.295276 (.316343)	.284449 (.316469)
80	.1	5	Mean	**.012736** (.014075)	.046142 (.058836)	.288916 (.213707)	.026058 (.027837)	**.014167** (.012959)
			ETMA	.028059 (.030684)	.020693 (.023629)	**.0211832** (.023367)	.023479 (.027629)	.024345 (.024738)
			DTM	.016765 (.017203)	**.018577** (.021704)	.034285 (.040483)	**.023081** (.022272)	.018081 (.016679)
			C&F	.179295 (.120392)	.267904 (.313102)	.241447 (.204903)	.235759 (.200404)	.242181 (.326559)
50	.05	25	Mean	**.019443** (.019470)	.598505 (.877871)	2.00705 (2.18651)	.357199 (.452284)	.058920 (.037000)
			ETMA	.041967 (.044412)	.038324 (.045385)	.039744 (.047938)	**.039558** (.044655)	**.039143** (.042550)
			DTM	.027283 (.028692)	**.028224** (.034722)	**.028581** (.045668)	.142177 (.216643)	.073974 (.052580)
			C&F	.235814 (0.219801)	.293020 (.325771)	.306678 (.307732)	.315216 (.373383)	.316692 (.360712)
80	.05	25	Mean	**.012736** (.014075)	.380566 (.554507)	2.002653 (1.892887)	.214524 (.256844)	.037192 (.021657)
			ETMA	.028059 (.030684)	.028059 (.030684)	.026697 (.029234)	**.025204** (.026908)	**.025034** (.028828)
			DTM	.016765 (.017203)	**.019338** (.021305)	**.023875** (.057293)	.081431 (.110393)	.047306 (.029342)
			C&F	.179295 (.120392)	.202424 (.161379)	.199860 (.174267)	.200357 (.158083)	.198275 (.124842)
50	.1	25	Mean	**.019443** (.019470)	1.43842 (2.09170)	7.67767 (6.20711)	.650730 (.793654)	.104663 (.056480)
			ETMA	.041967 (.044412)	**.044679** (.103249)	**.040384** (.109995)	**.035046** (.039377)	**.035938** (.040379)
			DTM	.027283 (.028692)	.071070 (.359990)	.279066 (.907998)	.266246 (.325132)	.127604 (.072196)
			C&F	.235814 (0.219801)	.413356 (.491066)	.444635 (.638904)	.387787 (.460157)	.487706 (1.857527)
80	.1	25	Mean	**.012736** (.014075)	.721368 (.985035)	7.13882 (4.81952)	.393982 (.482538)	.058355 (.028707)
			ETMA	.028059 (.030684)	.020875 (.021987)	**.023616** (.042221)	**.022323** (.025796)	**.019207** (.019672)
			DTM	.016765 (.017203)	**.019215** (.022241)	.177194 (.766125)	.175461 (.216869)	.072071 (.036172)
			C&F	.179295 (.120392)	.294912 (.400697)	.258056 (.280400)	.283760 (.350824)	.244363 (.258518)

Table 2.4: Results of the simulations for functional data to compare the behaviour of the mean, the trimmed mean (ETMA), the depth trimmed mean (DTM) and the Cuesta-Albertos and Fraiman's trimmed mean (C&F) in each of the *Models 1-5*

n	q	Estimator	*Model 6*	*Model 7*	*Model 8*	*Model 9*
50	.05	Mean	.274843 (.274843)	.318066 (.201209)	.279586 (.175522)	.332575 (.221595)
		ETMA	**.102215** (.088751)	**.098173** (.087392)	**.102305** (.089215)	**.101730** (.090843)
		DTM	.387662 (.242542)	.452014 (.298218)	.392610 (.265891)	.478837 (.330253)
		C&F	.680324 (.690054)	.691095 (.774417)	.749653 (.793939)	.748472 (.930509)
80	.05	Mean	.177057 (.093181)	.202028 (.112033)	.170289 (.089113)	.197961 (.103885)
		ETMA	**.066588** (.059385)	**.064461** (.056836)	**.067025** (.059580)	**.064443** (.057663)
		DTM	.249272 (.135467)	.289686 (.163318)	.238519 (.125843)	.285359 (.154928)
		C&F	.401234 (.373435)	.405943 (.427162)	.441619 (.532938)	.410888 (.415632)
50	.1	Mean	.499788 (.254557)	.558465 (.295773)	.521980 (.272500)	.590299 (.312491)
		ETMA	**.098882** (.096812)	**.099748** (.092386)	**.101381** (.095071)	**.100544** (.094835)
		DTM	.716873 (.379071)	.809014 (.443945)	.746505 (.400653)	.855152 (.460487)
		C&F	1.09877 (1.27201)	1.37985 (7.06320)	1.14676 (1.44264)	1.22649 (1.55378)
80	.1	Mean	.309428 (.127167)	.357100 (.162768)	.304456 (.128494)	.363893 (.155640)
		ETMA	**.059315** (.051736)	**.063023** (.057198)	**.063683** (.060024)	**.059159** (.053112)
		DTM	.441946 (.189800)	.517605 (.244283)	.436498 (.195145)	.526328 (.227663)
		C&F	.685767 (.963506)	.594373 (.781261)	.674865 (.947538)	.579299 (.695001)

Table 2.5: Results of the simulations for functional data to compare the behaviour of the mean, the trimmed mean (ETMA), the depth trimmed mean (DTM) and the Cuesta-Albertos and Fraiman's trimmed mean (C&F) in each of the *Models 6-9*

The empirical results obtained can be seen in Tables 2.4 and 2.5. The numbers in parenthesis are the standard errors of the mean squared errors of the estimators and the bold number is the minimum mean squared error obtained in each situation and for each of the models.

For instance, independently of the case of study, the best estimator of the population mean is the sample mean in *Model 1*, what is logical because in that model there is no contamination at all.

In approximately half of the models with magnitude outliers the best choice is the trimmed mean based on the depth and in the other half, the trimmed mean proposed in this article, but one important remark is that in almost all the situations with more contamination (bigger values of q and M), the estimator chosen is this last one.

Moreover, the empirical trimmed mean is always the estimator with smallest mean squared error in all models with only shape outliers.

2.5 Applications to set- and fuzzy set-valued data

This section aims to particularize the ideas and developments about trimmed means for Hilbert space-valued data in previous sections to the special cases of set- and fuzzy set-valued data. On one hand, the particularization will be illustrated by means of two real-life examples and, on the other hand, simulations are carried out to compare in such cases the behaviour of the trimmed means in contrast to the alternative approaches introduced in Chapter 2.

2.5.1 Illustrative examples

This subsection illustrates the computation of the trimmed means for real-life applications of both fuzzy-valued data and set-valued data.

First of all, the computation of these location measures will be illustrated for a data set of fuzzy sets.

Example 2.5.1. The experiment called 'Perceptions' consists of measuring the subjective human perception about the relative length of different lines. Both the software developed and the whole data set (with its complete description) can be found on the web page http://bellman.ciencias.uniovi.es/SMIRE/Perceptions.html.

The experiment was carried out as follows: 9 men and 8 women had to compare the relative length of a line segment in contrast to the reference line segment showing the largest length considered. Their answer had to be given in two ways: using a linguistic descriptor (among 'VERY SMALL', 'SMALL', 'MEDIUM', 'LARGE' and 'VERY LARGE') and using a fuzzy set, as shown in Figure 2.6.

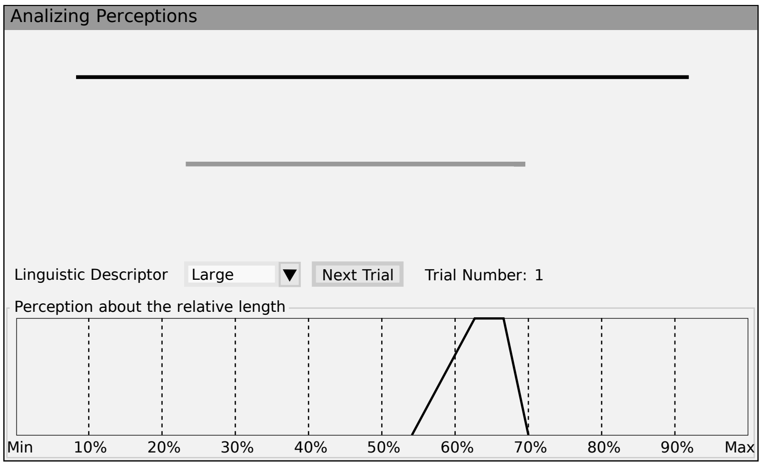


Figure 2.6: An example of an answer in the experiment 'Perceptions'

For a more detailed explanation, see Colubi *et al.* [34] and González-Rodríguez *et al.* [90].

The first 27 trials of each person correspond to random ordered relative sizes (of the line segment shown with respect to the reference one) in the fixed grid 3.7%, 15.2%, 26.8%, 38.4%, 50%, 61.5%, 73.1%, 84.7% and 96.3%. In the complete data set the answers of people who did not reach the 27 trials appear as well. In order to study the first 27 trials per person, the data set in González-Rodríguez *et al.* [90] will be used instead. Indeed, only the three trials referring to the line segment with length of about 84.7% per person will be considered.

In Figure 2.7 (left), the sample of fuzzy sets which is used in this example can be seen. The calculated location measures are given in Table 2.6 (their graphical display can be seen in Figure 2.7 (right)). Both in this example and in Example 2.5.2 (which deals with compact convex set-valued data), the weight parameter θ involved in the distance has been chosen to be equal to 1/3 and the trimming proportion to be equal to .2.

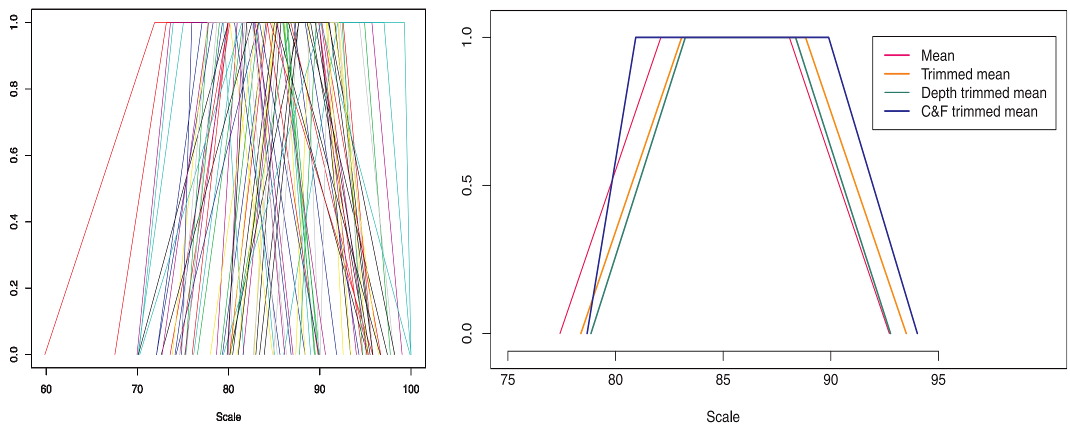


Figure 2.7: The sample considered in the experiment 'Perceptions' (left) and the graphical display of the estimates of the sample Aumann-type mean value, the empirical trimmed mean, the depth trimmed mean and the approximation of the empirical trimmed mean by Cuesta-Albertos and Fraiman (right)

For the computation of the depth trimmed mean, a partition of the domain determined by the 0-level of all the sample fuzzy sets had to be considered in order to express each fuzzy set as a function and apply the functional depth of Fraiman and Muniz [75].

It can be seen that all the estimates except for the approximation of the empirical trimmed mean by Cuesta-Albertos and Fraiman (which has to be an element from the sample) have the same shape. However, if they are seen as estimates of the central tendency, the sample Aumann-type mean value is observed to be located more to the left, influenced by the fuzzy number whose 0-level infimum is smaller (while the trimmed means, more robust, do not take that fuzzy number into account).

Estimator	inf 0-level	inf 1-level	sup 1-level	sup 0-level
Mean	77.42118	82.10667	88.07588	92.71765
ETMA	78.39098	83.07537	88.82756	93.50585
DTM	78.85927	83.24390	88.36415	92.77488
C&F	78.68	80.94	89.90	94.02

Table 2.6: Estimates of the sample Aumann-type mean value, the empirical trimmed mean (ETMA), the depth trimmed mean (DTM) and the approximation of the empirical trimmed mean by Cuesta-Albertos and Fraiman (C&F)

Now, an example of the computation of these location measures for a random (compact and convex) set is given.

Example 2.5.2. The temporary exhibition 'Spanish Drawings from the British Museum: Renaissance to Goya' was held in the Museo Nacional del Prado (Madrid, Spain) until 16th June 2013. Jointly organized with The British Museum, it consists of 71 works and two additional paintings from the collection of the Museo Nacional del Prado for which the corresponding drawings from the British Museum were preparatory. For more information, visit the web page:

http://www.museodelprado.es/en/exhibitions/exhibitions/at-the-museum/
el-trazo-espanol-en-el-british-museum-dibujos-del-renacimiento-a-goya/

The exhibit list, detailing the title, author, year and size of each drawing, can be also found on the web page:

http://bellman.ciencias.uniovi.es/SMIRE/Drawings.html

To organize the distribution of the drawings, it is useful to focus on their dimensions instead of their surface. The empirical trimmed mean and its approximation by Cuesta-Albertos and Fraiman were computed as well as the Aumann mean. As the support function of the Aumann mean is the mean of the support functions, it could be proven that indeed the sample Aumann mean is the rectangle with mean dimensions $\overline{a}_n \times \overline{b}_n$, being $a_i \times b_i$ $(i = 1, \ldots, n)$ the dimensions of the i-th drawing.

In Figure 2.8 (left) all the drawings are represented by a centered rectangle in order to compute the estimates. Note that there is a huge rectangle (Painting a)) that is a clear outlier. In Figure 2.8 (right), the sample is plotted again after removing that painting to be able to see the remaining observations more clearly.

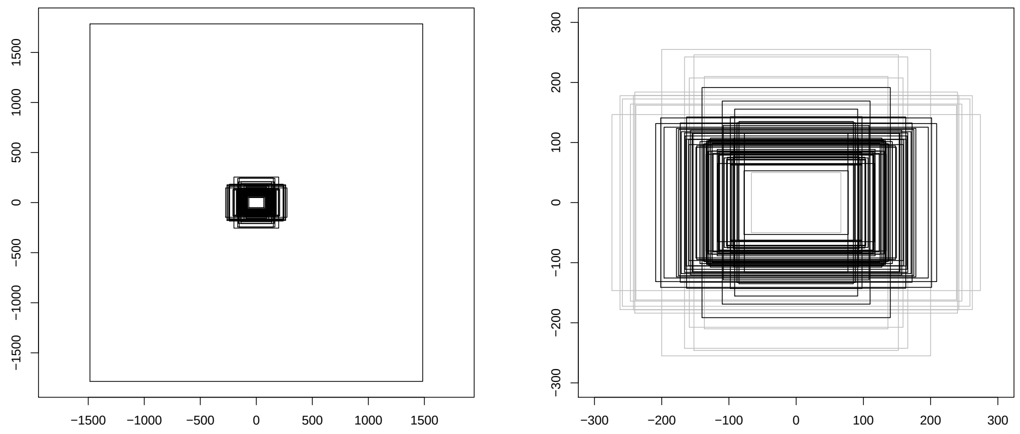


Figure 2.8: On the left, the sample of sets considered in this example; each drawing of dimensions $a \times b$ mm is represented by the centered rectangle $[-a/2, a/2] \times [-b/2, b/2]$. On the right, all the drawings (in grey) except from the outlier (Painting a); in black, the drawings that belong to the trimming region, used to compute the trimmed mean through the proposed algorithm

The values for the location measures are as follows:

- The sample Aumann-type mean value is the rectangle with width 323.4 mm and height 280.9 mm.
- The sample trimmed mean proposed (with trimming proportion .2) is the rectangle with dimensions 255.7 mm $\times$ 202.7 mm.
- The trimmed mean computed using the algorithm by Cuesta-Albertos and Fraiman (also with trimming proportion .2) is the rectangle with dimensions 255 mm $\times$ 215 mm.

All these estimates are plotted in Figure 2.9. As it can be seen, when estimating the central tendency, the mean is influenced by the outlier (Painting a)), while both trimmed means trim this outlier and their estimates have smaller dimensions, which correspond better to the majority of the sample.

In this case the depth trimmed mean does not appear because it is developed for functional data and no straightforward adaptation to compact and convex sets has been found. In Example 2.5.1, after taking a partition of the domain where all the elements of the sample are, the considered fuzzy numbers were expressed as functions and the depth could be computed. However, this approach is not valid for the rectangles, which cannot be treated as functions.

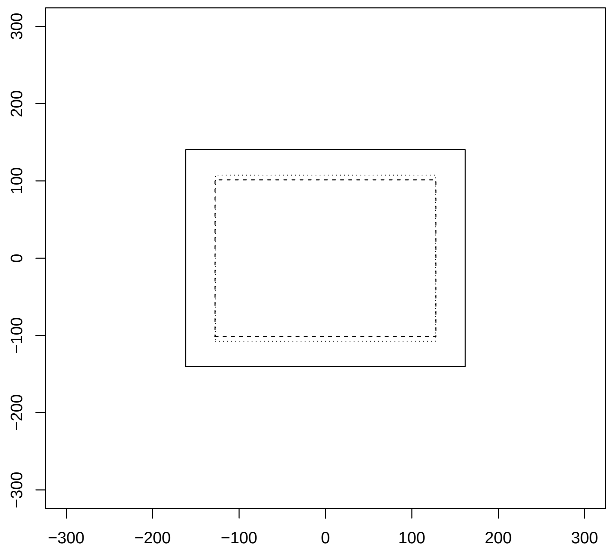


Figure 2.9: The rectangles corresponding to the estimates for the mean (continuous line), the trimmed mean (discontinuous line) and the Cuesta-Albertos and Fraiman's trimmed mean (dotted line)

2.5.2 Simulation-based comparative analysis

In this subsection some simulations with fuzzy number-valued data are shown to compare the ETMA, DTM and C&F procedures. More specifically, this empirical study joins the main design of the previous one with functional data in Section 2.4 and the simulation procedure in Section 1.5. The simulations in this subsection have been generated by dividing the generated sample of n fuzzy sets into a non-contaminated subsample, of size $n(1 - c_p)$, and a contaminated subsample, distinguishing CASE 1 and CASE 2 in Section 1.5.

The main change in the simulation procedure lies in *Steps 2* and *3*. Other changes affect the sample sizes and number of replications. The simulation steps are:

Step 1. A sample of size $n = 100$ of trapezoidal fuzzy numbers has been simulated for each of some different situations by following the scheme in Step 1 of the first simulations in p. 52.

Step 2. $NMC = 10000$ replications of *Step 1* have been considered for the situation $c_p = C_D = 0$, in order to approximate the population values of the Aumann-type mean, the trimmed mean, the depth trimmed mean and the Cuesta-Albertos and Fraiman's trimmed mean for the non-contaminated distribution by Monte Carlo. These three last measures have been approximated for both the considered trimming proportions (0.2 and 0.45).

That is to say, the sample Aumann-type mean, the ETMA, the DTM and the C&F estimates have been computed for each of the 10000 samples and, for each of the four measures, the mean of these 10000 estimates has been obtained.

Step 3. $N = 500$ replications of *Step 1* have been considered for all the 16 situations concerning c_p and C_D. For each situation and each of the 500 samples, the sample Aumann-type mean, the ETMA, the DTM and the C&F have been computed as estimates of their corresponding population values for the non-contaminated distribution. Afterwards, for each considered location measure, the mean squared distance between the 500 sample estimates and its population value approximated in *Step 2* (i.e., the mean squared error) is obtained. Finally, the standard deviation of the mean squared error is computed too.

The values chosen for other parameters are:

- $\theta = 1/3$ in the D_θ^φ distance between fuzzy sets;
- $n_{\mathcal{X}_0} = 500$ number of equidistant points chosen as a partition of the domain $[\min\{\inf \mathcal{X}_{i0}\}_{i=1}^n, \max\{\sup \mathcal{X}_{i0}\}_{i=1}^n]$ to represent the trapezoidal fuzzy sets as functions (in order to compute the depth, defined for functional data).

Taking into account that every fuzzy set with support function belonging to $\mathbb{H}_2$ can be 'identified' with an element of a Hilbert space (its own support function), for each $\beta \in (0,1)$ and a fixed probability distribution on the Borel σ algebra on $\mathbb{H}_2$, P_X, the corresponding trimmed mean of $\mathcal{X}$ is defined as $E(s_{\mathcal{X}}|A_{P_X})$, the trimming region being given by

$$\int_{A_{P_X}} \|s_{\tilde{x}} - E(s_{\mathcal{X}}|A_{P_X})\|_{\mathbb{H}_2}^2 dP_X(s_{\tilde{x}}) = \min_{\substack{A\subset\mathbb{H}_2 \\ P_X(A)\geq 1-\beta}} \int_A \|s_{\tilde{x}} - E(s_{\mathcal{X}}|A)\|_{\mathbb{H}_2}^2 dP_X(s_{\tilde{x}}).$$

Analogously, the sample trimming region is given by

$$\widehat{E} = \arg \min_{\substack{E\subset\{1,\ldots,n\} \\ \#E=h}} \frac{1}{h}\sum_{i\in E} \left\| s_{\tilde{x}_i} - \frac{1}{h}\sum_{j\in E} s_{\tilde{x}_j} \right\|_{\mathbb{H}_2}^2 = \arg\min_E \mathrm{Var}(\boldsymbol{s}_{\tilde{\mathbf{x}}_n}|E)$$

and the sample trimmed mean, $\frac{1}{h}\sum_{i\in\widehat{E}} s_{\tilde{x}_i}$ (where $\boldsymbol{s}_{\tilde{\mathbf{x}}_n} = (s_{\tilde{x}_1}, \ldots, s_{\tilde{x}_n})$ denotes the sample obtained from $\mathcal{X}$).

The results appear in Tables 2.7 and 2.8. The minimum mean squared error reached in each case is in bold letters, so it can be seen that the mean and the depth trimmed mean are the best estimators of the corresponding population value when there is no contamination.

c_p	C_D	Mean	ETMA ($\beta = .2$)	Depth TMean ($\beta = .2$)	C&F ($\beta = .2$)
.0	-	.020239 (.019538)	.030947 (.036265)	**.019270** (.019659)	.155957 (.161301)
.1	0	.071213 (.059865)	.033773 (.039581)	**.032877** (.034691)	.167204 (.160890)
.1	1	.107164 (.073137)	**.031587** (.036203)	.035388 (.037037)	.194660 (.184163)
.1	5	.413908 (.191479)	**.030264** (.029436)	.068187 (.081826)	.195983 (.186874)
.1	10	1.176410 (.481783)	**.027250** (.028361)	.192900 (.289635)	.207873 (.243604)
.1	100	86.942010 (33.71212)	**.031655** (.034465)	16.465730 (25.80862)	.246895 (.231829)
.2	0	.199625 (.128763)	**.056396** (.050108)	.065299 (.050222)	.233768 (.214643)
.2	1	.345419 (.190534)	**.059860** (.046930)	.089312 (.070942)	.263960 (.274717)
.2	5	1.547838 (.623109)	**.054135** (.043152)	.230970 (.253302)	.868819 (1.082253)
.2	10	4.620137 (1.608131)	**.069707** (.049117)	.681509 (.957643)	1.685437 (2.416464)
.2	100	328.1507 (111.1436)	**.070683** (.050285)	67.827430 (105.1396)	1.888415 (2.916228)
c_p	C_D	Mean	ETMA ($\beta = .45$)	Depth TMean ($\beta = .45$)	C&F ($\beta = .45$)
.4	0	.740745 (.373957)	.105939 (.086270)	**.093973** (.073935)	.215359 (.176504)
.4	1	1.285396 (.615249)	**.090338** (.076172)	.130757 (.142755)	.215994 (.201366)
.4	5	6.234710 (2.113790)	**.081873**(.066977)	.922707 (1.340607)	.340222 (.496529)
.4	10	18.20170 (5.113324)	**.084563** (.060761)	3.992455 (5.973074)	.412731 (.553588)
.4	100	1371.088 (411.6347)	**.081888** (.059663)	363.3129 (530.8564)	.451198 (.605826)

Table 2.7: CASE 1. Results of the simulations for fuzzy data to compare the behaviour of the mean, the trimmed mean (ETMA), the depth trimmed mean (Depth TMean) and the Cuesta-Albertos and Fraiman's trimmed mean (C&F) with trimming proportions .2 and .45

When the error proportion increases, the best choice is to use the trimmed mean defined in this section (with either .2 or .45 trimming proportion, depending on the amount of contamination in the considered sample).

The behaviour of the trimmed means is always more robust than for the Aumann mean, which has a finite sample breakdown point of only $1/n$. The adaption of the depth trimmed mean is not good enough because of the different 0-levels (support) of the fuzzy numbers generated, not having a common domain as in the functional case.

The population trimmed mean was approached by the Monte Carlo method in order to compute the squared error $E(j) = \left(D_\theta^\varphi(\hat{g}_n, g)\right)^2$, where $\hat{g}_n$ represents any of the proposed estimators, as seen before, and g represents the corresponding population value (see Figure 2.10).

The mean squared error (MSE) and its standard deviation (s) were computed through the same formulas as in the functional case.

c_p	C_D	Mean	ETMA ($\beta = .2$)	Depth TMean ($\beta = .2$)	C&F ($\beta = .2$)
.0	-	**.009870** (.011829)	.021766 (.030875)	.016095 (.019398)	.039595 (.056565)
.1	0	**.016444** (.022615)	.026524 (.034710)	.026389 (.036668)	.041681 (.052192)
.1	1	**.023079** (.024650)	.025725 (.030593)	.031291 (.038429)	.044325 (.057795)
.1	5	.220575 (.106323)	**.021713** (.025494)	.072698 (.114937)	.052428 (.135639)
.1	10	.884834 (.372503)	**.024308** (.029090)	.205296 (.307968)	.062374 (.177092)
.1	100	83.75032 (32.45057)	**.022880** (.029948)	16.01512 (25.96815)	.059158 (.200482)
.2	0	**.018996** (.025317)	.025711 (.034392)	.029631 (.038858)	.049089 (.071175)
.2	1	.053481 (.048911)	**.032670** (.030657)	.062243 (.076196)	.055639 (.119129)
.2	5	.833135 (.316060)	**.025657** (.023731)	.210033 (.329815)	.631538 (.980040)
.2	10	3.310474 (1.281787)	**.022801** (.020999)	.762788 (1.147577)	.959677 (1.194475)
.2	100	332.6713 (118.6947)	**.023279** (.019954)	63.21118 (100.7086)	1.167254 (1.718109)
c_p	C_D	Mean	ETMA ($\beta = .45$)	Depth TMean ($\beta = .45$)	C&F ($\beta = .45$)
.4	0	**.030258** (.043187)	.052824 (.077102)	.097097 (.132565)	.077202 (.112341)
.4	1	.158696 (.112643)	.091076 (.083862)	.204219 (.263046)	**.088566** (.164482)
.4	5	3.283897 (1.121385)	**.066523** (.045807)	1.295878 (2.255459)	.120824 (.289803)
.4	10	13.16870 (4.255767)	**.067270** (.046376)	5.044778 (8.545120)	.159640 (.423301)
.4	100	1309.785 (443.1194)	**.068764** (.045543)	401.3903 (715.6379)	.210200 (.577308)

Table 2.8: CASE 2. Results of the simulations for fuzzy data to compare the behaviour of the mean, the trimmed mean (ETMA), the depth trimmed mean (Depth TMean) and the Cuesta-Albertos and Fraiman's trimmed mean (C&F) with trimming proportions .2 and .45

2.6 Concluding remarks of this chapter

This chapter has been devoted to introduce a new algorithm to compute the sample trimmed mean in general Hilbert spaces. Its complexity and efficiency have been compared with the Cuesta-Albertos and Fraiman algorithm, although one should be aware that this last algorithm only attempts to approximate (not to compute) the empirical trimmed mean. Furthermore, the consistency has been established and the finite sample breakdown point of trimmed means has been derived to quantify its robustness.

The main contribution in the Chapter refers to

- the empirical studies focused on functional data, so fundamental in contemporary research, and on fuzzy (set-valued) data;
- since not only the proposed trimmed mean is applicable to those spaces, but also other trimmed means defined in different ways (through depths or other algorithms) and other location values, the simulation studies have considered some representative possibilities;

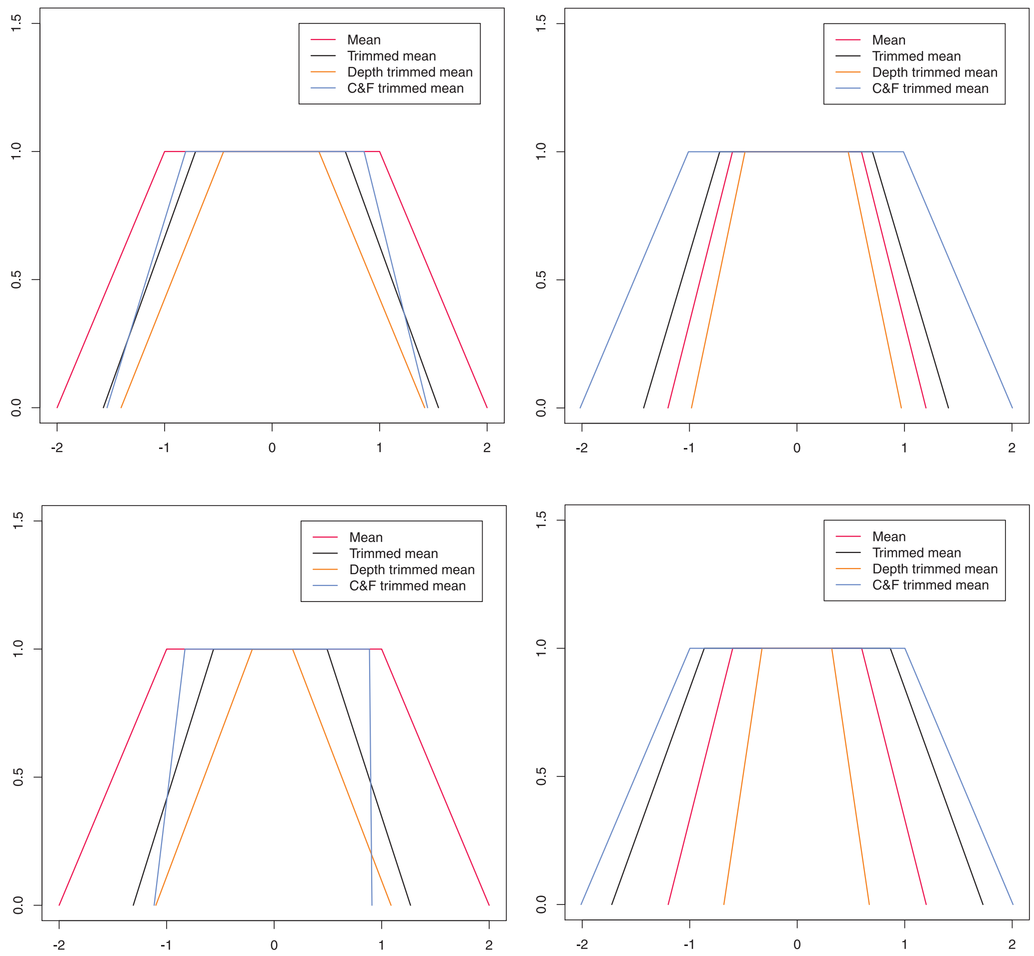


Figure 2.10: CASE 1: Population values of the mean and its Monte Carlo estimates for the trimmed mean, the depth trimmed mean and Cuesta-Albertos and Fraiman (C&F) trimmed mean when the trimming proportion is .2 (left above) and .45 (left below). CASE 2: Population values of the mean and its Monte Carlo estimates for the trimmed mean, the depth trimmed mean and Cuesta-Albertos and Fraiman (C&F) trimmed mean when the trimming proportion is .2 (right above) and .45 (right below)

- the robustness of the different estimators has been shown with the simulation studies, confirming the good behaviour of the new trimmed mean;
- the comparison has also been carried out for real-life examples involving fuzzy set-valued data;
- in the general application to set- and fuzzy set values of high dimension, the trimmed mean becomes quite useful.

The ideas and results in this chapter have been gathered in a submitted manuscript (González-Rodríguez *et al.* [92]) and two communications to conferences (Sinova *et al.* [177, 178]).

Chapter 3

Location M-estimates from imprecise-valued data

Another approach to estimating location that currently has considerable practical value consists of what are called M-estimators. M-estimators were introduced in estimating location from real-valued data by Huber [106] with the aim of limiting the influence of outliers in methods like the least squares one. For this purpose, the key idea consisted of replacing the square of the 'errors' with a (usually less rapidly increasing) loss function of the data and the parameter estimate. In this respect, M-estimators were presented as intermediaries between the sample mean and median.

In this chapter we are going to extend the notion of M-estimators to deal with imprecise-valued data. Since there exist some ideas and results in the literature which can be easily adapted to deal with Hilbert space-valued data, this adaptation is to be first carried out. Later, sufficient conditions of the loss function guaranteeing that the particularization leads to well-defined estimates are to be established. Nevertheless, some interesting and natural choices for the loss functions do not fulfill such conditions, so some *ad hoc* developments are to be considered in case of fuzzy number- and interval-valued data.

Following Huber ideas [106, 108], a location M-estimate of a Hilbert space-valued random element is formalized as follows: if $\{h_1, \ldots, h_n\}$ is a sample of independent values from a random element X taking on values on the Hilbert space $\mathbb{H}$, an estimate T_n defined by the minimization problem

$$\sum_{i=1}^{n} \rho(\| h_i - T_n \|_{\mathbb{H}}) = \min_{g\in\mathbb{H}} \sum_{i=1}^{n} \rho(\| h_i - g \|_{\mathbb{H}}),$$

where ρ is an arbitrary real-valued function (which will be referred to hereinafter as the *loss function*) is said to be a location M-estimate.

In Section 3.1 location M-estimates of a Hilbert-valued random element will be studied, and their expression will be discussed under some conditions over the loss function. Simulations for the functional-valued case are carried out in Section 3.1.5. In Section 3.2 the particularization of the M-estimates in Section 3.1 to the set- and fuzzy set-valued cases is examined.

In Sections 3.3, 3.4 and 3.5, solutions to some special choices of the loss function that are not covered by the preceding conditions will be achieved in the special case of random fuzzy numbers and random intervals through the use of *ad hoc* techniques. More concretely, the minimization problem above are to be solved

- when ρ is the absolute value function and two different norms associated with L^1 metrics on $\mathcal{F}_c(\mathbb{R})$ are considered, and
- for random intervals when ρ is the square root of the absolute value and an L^2 metric on $\mathcal{F}_c(\mathbb{R})$ is considered.

Given $(\widetilde{x}_1,\ldots,\widetilde{x}_n)$ a sample of independent values from a random fuzzy number $\mathcal{X}$, Section 3.3 deals with the minimization problem (see notations in Proposition 1.3.7, p. 37)

$$\sum_{i=1}^n \|\boldsymbol{\iota}_{\widetilde{x}_i}-\boldsymbol{\iota}_{\widetilde{T}_n}\|_1^{\varphi} = \min_{\widetilde{U}\in\mathcal{F}_c(\mathbb{R})}\sum_{i=1}^n \|\boldsymbol{\iota}_{\widetilde{x}_i}-\boldsymbol{\iota}_{\widetilde{U}}\|_1^{\varphi} = \min_{\widetilde{U}\in\mathcal{F}_c(\mathbb{R})}\sum_{i=1}^n \boldsymbol{\rho_1}(\widetilde{x}_i,\widetilde{U}),$$

whereas Section 3.4 deals with the minimization problem (see notations in Proposition 1.3.8, p. 38)

$$\sum_{i=1}^n \|\boldsymbol{\upsilon}^{\varphi}_{\widetilde{x}_i}-\boldsymbol{\upsilon}^{\varphi}_{\widetilde{T}_n}\|_\theta^{\varphi\star} = \min_{\widetilde{U}\in\mathcal{F}_c(\mathbb{R})}\sum_{i=1}^n \|\boldsymbol{\upsilon}^{\varphi}_{\widetilde{x}_i}-\boldsymbol{\upsilon}^{\varphi}_{\widetilde{U}}\|_\theta^{\varphi\star} = \min_{\widetilde{U}\in\mathcal{F}_c(\mathbb{R})}\sum_{i=1}^n \mathscr{D}^{\varphi}_\theta(\widetilde{x}_i,\widetilde{U}).$$

Given $(\mathsf{x}_1,\ldots,\mathsf{x}_n)$ a sample of independent values from a random interval X, Section 3.5 concerns the minimization problem (notations in Proposition 1.3.1, p. 24)

$$\sum_{i=1}^n \sqrt{\|\boldsymbol{\eta}_{\mathsf{x}_i}-\boldsymbol{\eta}_{\mathsf{T}_n}\|_\theta} = \min_{K\in\mathcal{K}_c(\mathbb{R})}\sum_{i=1}^n \sqrt{\|\boldsymbol{\eta}_{\mathsf{x}_i}-\boldsymbol{\eta}_{K}\|_\theta} = \min_{K\in\mathcal{K}_c(\mathbb{R})}\sum_{i=1}^n d_\theta(\mathsf{x}_i,K).$$

Section 3.6 is devoted to illustrate the main ideas in the chapter by means of some real-life examples. The chapter ends with a summary of the novelties and the related publications derived from it in Section 3.7.

3.1 Location M-estimates for Hilbert space-valued random elements

In searching for a robust nonparametric density estimator, Kim and Scott [114] (see also Kim [112], and Kim and Scott [113]) have combined a traditional kernel density estimator with ideas from classical M-estimation. They have interpreted the kernel density estimator based on a radial, positive semi-definite kernel as a sample mean in the associated reproducing kernel Hilbert space (see, for instance, Steinwart and Christmann [190] for details about). To lower the sensitivity of the sample mean to outliers, Kim and Scott suggest to estimate it robustly *via* M-estimates yielding a robust kernel density estimator.

Despite the fact that Kim and Scott ideas have been developed for reproducing kernel Hilbert spaces, they have also generalized the results to other Hilbert spaces, always within the setting of kernel density estimation. Therefore, and although the adaptation of Kim and Scott's results and corresponding proofs to Hilbert space-valued random elements is straightforward, it is to be detailed now in the context of M-estimates from Hilbert space-valued data.

Then, in this section we will present the adaptation of the concepts and results by Kim and Scott for general Hilbert-valued random elements.

3.1.1 Basic concepts for location M-estimates of Hilbert space-valued random elements

The population and sample M-estimates of location for Hilbert space-valued random elements are defined as follows:

Definition 3.1.1. *Let $\mathbb{H}$ be a Hilbert space with associated norm $\|\cdot\|_{\mathbb{H}}$, $(\Omega, \mathcal{A}, P)$ be a probability space, $X : \Omega \to \mathbb{H}$ be an associated Hilbert-valued random element and ρ be an arbitrary loss function. The **population M-estimate of location** is the element $g_P^M \in \mathbb{H}$ minimizing*

$$J_P(g) = \int_\Omega \rho(\|X(\omega) - g\|_{\mathbb{H}})\, dP(\omega),$$

i.e.,

$$g_P^M = \arg\min_{g\in\mathbb{H}} E[\rho(\|X(\omega) - g\|_{\mathbb{H}})].$$

Definition 3.1.2. *Let $\mathbb{H}$ be a Hilbert space with associated norm $\|\cdot\|_{\mathbb{H}}$, $(\Omega, \mathcal{A}, P)$ be a probability space, $X : \Omega \to \mathbb{H}$ be an associated Hilbert-valued random element, $(X_1, \ldots, X_n)$ be a simple random sample from X and ρ be an arbitrary loss function. The **sample M-estimate of location** is the Hilbert-valued statistic $\widehat{g^M[X]}_n$ such that for each realization from the simple random sample, $\mathbf{h_n} = (h_1, \ldots, h_n)$, the Hilbert value(s) $\widehat{g^M[\mathbf{h_n}]}$ is (are) the solution(s) of the following optimization problem:*

$$\min_{g\in\mathbb{H}} J(g) = \min_{g\in\mathbb{H}} \frac{1}{n} \sum_{i=1}^{n} \rho(\|h_i - g\|_{\mathbb{H}}).$$

3.1.2 Representer theorem for the sample location M-estimates of Hilbert space-valued random elements

An interesting contribution of Kim and Scott's studies lies in their analysis of the conditions to ensure the existence of sample M-estimates of location as well as the conditions to express them as weighted linear combinations of the sample elements.

First, necessary conditions for $\widehat{g^M[\mathbf{h_n}]}$ to be a minimizer of J in Definition 3.1.2 will be found. In this case, as the function J is defined over a Hilbert space, the necessary conditions are formulated by using the Gâteaux differential. Recall that given a vector space (in particular, the Hilbert space $\mathbb{H}$) and a function $J : \mathbb{H} \to \mathbb{R}$, the Gâteaux differential of J at $g \in \mathbb{H}$ with incremental $h \in \mathbb{H}$ is defined as

$$\delta J(g;h) = \lim_{\varsigma \to 0} \frac{J(g+\varsigma h) - J(g)}{\varsigma}$$

$$= \lim_{\varsigma \to 0} \frac{\frac{1}{n}\sum_{i=1}^{n}\rho(\|h_i - (g+\varsigma h)\|_{\mathbb{H}}) - \frac{1}{n}\sum_{i=1}^{n}\rho(\|h_i - g\|_{\mathbb{H}})}{\varsigma}$$

$$= \lim_{\varsigma \to 0} \frac{1}{n\,\varsigma} \sum_{i=1}^{n} \left[\rho(\|h_i - (g+\varsigma h)\|_{\mathbb{H}}) - \rho(\|h_i - g\|_{\mathbb{H}})\right].$$

In case $\delta J(\widehat{g^M[\mathbf{h_n}]};h)$ is defined for all $h \in \mathbb{H}$, a necessary condition for $\widehat{g^M[\mathbf{h_n}]}$ to be a minimum of J is that $\delta J(\widehat{g^M[\mathbf{h_n}]};h) = 0$ for all $h \in \mathbb{H}$ (see Luenberger [128]). The following lemma stems from such a condition:

Lemma 3.1.1. *Let $\phi(x) = \rho'(x)/x$ and suppose the assumptions below are satisfied:*

- *ρ is non-decreasing, $\rho(0) = 0$ and $\lim_{x\to 0} \rho(x)/x = 0$,*
- *$\phi(0) \triangleq \lim_{x\to 0} \phi(x)$ exists and it is finite.*

Then, the Gâteaux differential of J at $g \in \mathbb{H}$ is $\delta J(g;h) = -\langle V(g), h\rangle_{\mathbb{H}}$, where

$$\begin{array}{rccl} V: & \mathbb{H} & \to & \mathbb{H} \\ & g & \mapsto & V(g) = \frac{1}{n}\sum_{i=1}^{n}\phi(\|h_i - g\|_{\mathbb{H}})\cdot(h_i - g) \end{array}$$

and a necessary condition for $\widehat{g^M[\mathbf{h_n}]}$ to be a minimizer of J is $V(\widehat{g^M[\mathbf{h_n}]}) = 0$.

Proof. To calculate the Gâteaux differential of J, two cases can be distinguished. Thus,

- If $h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h) \neq 0$, then

$$\frac{\partial}{\partial\varsigma}\rho(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}})$$

$$= \rho'(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}}) \cdot \frac{\partial}{\partial \varsigma} \|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}}$$

$$= \rho'(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}}) \cdot \frac{\partial}{\partial \varsigma} \sqrt{\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|^2_{\mathbb{H}}}$$

$$= \rho'(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}}) \cdot \frac{\frac{\partial}{\partial \varsigma} \|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|^2_{\mathbb{H}}}{2\sqrt{\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|^2_{\mathbb{H}}}}$$

$$= \frac{\rho'(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}})}{2\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}}}$$

$$\cdot \frac{\partial}{\partial \varsigma} \left(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|^2_{\mathbb{H}} + \varsigma^2 \|h\|^2_{\mathbb{H}} - 2\langle h_i - \widehat{g^M[\mathbf{h_n}]}, \varsigma h \rangle_{\mathbb{H}} \right)$$

$$= \frac{\rho'(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}})}{2\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}}} \cdot \left(2\varsigma \|h\|^2_{\mathbb{H}} - 2\langle h_i - \widehat{g^M[\mathbf{h_n}]}, h \rangle_{\mathbb{H}} \right)$$

$$= \frac{\rho'(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}})}{\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}}} \cdot \left(\varsigma \|h\|^2_{\mathbb{H}} - \langle h_i - \widehat{g^M[\mathbf{h_n}]}, h \rangle_{\mathbb{H}} \right)$$

$$= \phi(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}}) \cdot \left(-\langle h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h), h \rangle_{\mathbb{H}} \right).$$

- Otherwise (i.e., if $h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h) = 0$),

$$\frac{\partial}{\partial \varsigma} \rho(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}})$$

$$= \lim_{\kappa \to 0} \frac{\rho(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h + \kappa h)\|_{\mathbb{H}}) - \rho(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}})}{\kappa}.$$

Since $h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h) = 0$, then

$$\frac{\partial}{\partial \varsigma} \rho(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}}) = \lim_{\kappa \to 0} \frac{\rho(\|\kappa h\|_{\mathbb{H}}) - \rho(0)}{\kappa}$$

$$= \lim_{\kappa \to 0} \frac{\rho(\kappa \|h\|_{\mathbb{H}})}{\kappa} = \begin{cases} \lim_{\kappa \to 0} \frac{\rho(0)}{\kappa} & \text{if } h = 0 \\ \lim_{\kappa \to 0} \frac{\rho(\kappa \|h\|_{\mathbb{H}})}{\kappa \|h\|_{\mathbb{H}}} \|h\|_{\mathbb{H}} & \text{otherwise.} \end{cases}$$

As $\rho(0) = 0$ and $\lim_{x \to 0} \rho(x)/x = 0$, for any of the two last situations we have that the limit vanishes, and because $\phi(0)$ is well-defined under the assumptions in the lemma

$$\frac{\partial}{\partial \varsigma} \rho(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}}) = 0$$

$$= \phi(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}}) \left(-\langle h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h), h \rangle_{\mathbb{H}} \right).$$

Therefore, for any $\widehat{g^M[\mathbf{h_n}]}, h \in \mathbb{H}$ and $h_i \in \mathbb{H}$:

$$\frac{\partial}{\partial \varsigma}\rho(\|h_i-(\widehat{g^M[\mathbf{h_n}]}+\varsigma h)\|_{\mathbb{H}}) = \phi(\|h_i-(\widehat{g^M[\mathbf{h_n}]}+\varsigma h)\|_{\mathbb{H}}) \left(-\langle h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h), h\rangle_{\mathbb{H}}\right),$$

whence

$$\begin{aligned}
\delta J(\widehat{g^M[\mathbf{h_n}]}; h) &= \lim_{\kappa\to 0} \frac{J(\widehat{g^M[\mathbf{h_n}]} + \kappa h) - J(\widehat{g^M[\mathbf{h_n}]})}{\kappa} \\
&= \lim_{\kappa\to 0} \frac{J(\widehat{g^M[\mathbf{h_n}]} + (\varsigma+\kappa) h) - J(\widehat{g^M[\mathbf{h_n}]}+\varsigma h)}{\kappa}\Bigg|_{\varsigma=0} \\
&= \frac{\partial}{\partial\varsigma} J(\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\Big|_{\varsigma=0} = \frac{\partial}{\partial\varsigma}\left(\frac{1}{n}\sum_{i=1}^{n}\rho(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}})\right)\Bigg|_{\varsigma=0} \\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial\varsigma}\rho(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}})\Big|_{\varsigma=0} \\
&= \frac{1}{n}\sum_{i=1}^{n}\phi(\|h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h)\|_{\mathbb{H}})\cdot(-\langle h_i - (\widehat{g^M[\mathbf{h_n}]} + \varsigma h), h\rangle_{\mathbb{H}})\Big|_{\varsigma=0} \\
&= -\frac{1}{n}\sum_{i=1}^{n}\phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}})\cdot\langle h_i - \widehat{g^M[\mathbf{h_n}]}, h\rangle_{\mathbb{H}} \\
&= -\langle\frac{1}{n}\sum_{i=1}^{n}\phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}})\cdot(h_i - \widehat{g^M[\mathbf{h_n}]}), h\rangle_{\mathbb{H}} \\
&= -\langle V(\widehat{g^M[\mathbf{h_n}]}), h\rangle_{\mathbb{H}}.
\end{aligned}$$

Since a necessary condition for $\widehat{g^M[\mathbf{h_n}]}$ to be a minimizer of J is that $\delta J(\widehat{g^M[\mathbf{h_n}]}; h)$ $= 0$ for all $h \in \mathbb{H}$, the equivalent expression we have just found, allows us to conclude that a necessary condition to minimize J can be rewritten as $V(\widehat{g^M[\mathbf{h_n}]}) = 0$, i.e.,

$$\frac{1}{n}\sum_{i=1}^{n}\phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}})\cdot(h_i - \widehat{g^M[\mathbf{h_n}]}) = 0. \qquad \square$$

Lemma 3.1.1 is now considered to establish the representer theorem that expresses the sample location M-estimate as a convex linear combination of the sample components, $h_1, \ldots, h_n$. Thus,

Theorem 3.1.2. *Under the assumptions*

- ρ *is non-decreasing,* $\rho(0) = 0$ *and* $\lim_{x\to 0}\rho(x)/x = 0$,
- $\phi(0)$ *exists and it is finite, and*
- $\sum_{i=1}^{n}\phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}) > 0$,

the sample M-estimate of location exists and it can be expressed as

$$\widehat{g^M[\mathbf{h_n}]} = \sum_{i=1}^{n} w_i \cdot h_i,$$

where $w_i \geq 0$ *and* $\sum_{i=1}^{n} w_i = 1$*. Furthermore,* $w_i \propto \phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}})$*.*

Proof. By using Lemma 3.1.1, the considered assumptions guarantee that

$$\frac{1}{n}\sum_{i=1}^{n} \phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}) \cdot (h_i - \widehat{g^M[\mathbf{h_n}]}) = 0,$$

and solving this estimation equation for $\widehat{g^M[\mathbf{h_n}]}$ one gets that

$$\sum_{i=1}^{n} \phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}) \cdot h_i = \sum_{i=1}^{n} \phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}) \cdot \widehat{g^M[\mathbf{h_n}]}.$$

As $\widehat{g^M[\mathbf{h_n}]}$ does not depend on i,

$$\sum_{i=1}^{n} \phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}) \cdot h_i = \widehat{g^M[\mathbf{h_n}]} \sum_{i=1}^{n} \phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}).$$

Since $\sum_{i=1}^{n} \phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}) > 0$, then

$$\widehat{g^M[\mathbf{h_n}]} = \frac{\sum_{i=1}^{n} \phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}) \cdot h_i}{\sum_{j=1}^{n} \phi(\|h_j - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}})} = \sum_{i=1}^{n} \frac{\phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}})}{\sum_{j=1}^{n} \phi(\|h_j - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}})} \cdot h_i.$$

By denoting $w_i = \phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}) / \sum_{j=1}^{n} \phi(\|h_j - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}})$, one can express the M-estimate as $\widehat{g^M[\mathbf{h_n}]} = \sum_{i=1}^{n} w_i \cdot h_i$ and

$$\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} \left[\phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}) / \sum_{j=1}^{n} \phi(\|h_j - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}) \right] = 1,$$

$$w_i = \phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}) / \sum_{j=1}^{n} \phi(\|h_j - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}}) \geq 0$$

because $\phi(0) = 0$ and $\phi(x) = \rho'(x)/x \geq 0$ if $x > 0$ due to ρ having been assumed to be non-decreasing. □

The necessary condition for $\widehat{g^M[\mathbf{h_n}]} = \sum_{i=1}^{n} w_i \cdot h_i$ to be the minimizer of J proved in Theorem 3.1.2 is also sufficient (which guarantees then existence and uniqueness) by adding another assumption on J, as it can be seen in the following result:

Theorem 3.1.3. *Under the assumptions*

- ρ *is non-decreasing,* $\rho(0) = 0$ *and* $\lim_{x\to 0} \rho(x)/x = 0$,
- $\phi(0)$ *exists and it is finite, and*
- J *is strictly convex,*

the following conditions

i) $\widehat{g^M[\mathbf{h_n}]} = \sum_{i=1}^n w_i \cdot h_i$,

ii) $w_i \propto \phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}})$,

iii) $\sum_{i=1}^n w_i = 1$

are sufficient conditions for $\widehat{g^M[\mathbf{h_n}]}$ *to minimize* $J(g) = \frac{1}{n}\sum_{i=1}^n \rho(\|h_i - g\|_{\mathbb{H}})$.

Proof. If J is strictly convex, the global minimum should be unique. Furthermore, $\widehat{g^M[\mathbf{h_n}]} = \sum_{i=1}^n w_i \cdot h_i$ with $w_i \propto \phi(\|h_i - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}})$ and $\sum_{i=1}^n w_i = 1$ yields a minimum because $V(\widehat{g^M[\mathbf{h_n}]}) = 0$ whence (see Lemma 3.1.1) $\delta J(\widehat{g^M[\mathbf{h_n}]}; h) = 0$ for all $h \in \mathbb{H}$. □

Finally, some sufficient conditions for the function J to be strictly convex are given.

Proposition 3.1.4. *The function* J *is strictly convex provided either of the following conditions is fulfilled:*

i) ρ *is strictly convex and non-decreasing.*

ii) ρ *is convex, strictly increasing,* $n \geq 3$ *and* $\mathbf{A} = (\langle h_i, h_j\rangle_{\mathbb{H}})_{i,j=1}^n$ *is positive definite.*

Proof. Indeed, whatever $\gamma \in (0,1)$ and $g, h \in \mathbb{H}$ with $g \neq h$ may be

$$J(\gamma g + (1-\gamma)h) = \frac{1}{n}\sum_{i=1}^n \rho(\|h_i - \gamma g - (1-\gamma)h\|_{\mathbb{H}})$$
$$= \frac{1}{n}\sum_{i=1}^n \rho(\|\gamma(h_i - g) + (1-\gamma)(h_i - h)\|_{\mathbb{H}}).$$

The triangular inequality for $\|\cdot\|_{\mathbb{H}}$ along with the fact that ρ is non-decreasing ensures that

$$J(\gamma g + (1-\gamma)h) \leq \frac{1}{n}\sum_{i=1}^n \rho(\lambda\|h_i - g\|_{\mathbb{H}} + (1-\lambda)\|h_i - h\|_{\mathbb{H}}).$$

The last inequality is strict under condition $i)$, since the strict convexity of ρ implies that

$$\frac{1}{n}\sum_{i=1}^{n}\rho(\lambda\|h_i-g\|_{\mathbb{H}}+(1-\lambda)\|h_i-h\|_{\mathbb{H}})$$
$$<\frac{1}{n}\sum_{i=1}^{n}\lambda\rho(\|h_i-g\|_{\mathbb{H}})+(1-\lambda)\rho(\|h_i-h\|_{\mathbb{H}})=\lambda J(g)+(1-\lambda)J(h).$$

To prove this strict inequality under condition *ii*), suppose by *reductio ad absurdum* that this inequality holds with equality. Since ρ is strictly increasing, this can happen only if

$$\|\lambda(h_i-g)+(1-\lambda)(h_i-h)\|_{\mathbb{H}}=\lambda\|h_i-g\|_{\mathbb{H}}+(1-\lambda)\|h_i-h\|_{\mathbb{H}} \text{ for all } i\in\{1,\dots,n\}.$$

Equivalently, it can only happen if h_i-g and h_i-h are linearly dependent for all $i\in\{1,\dots,n\}$.

However, from $n\geq 3$ and due to the positive definiteness of $\mathbf{A}$, there exist three distinct h_i's (let's denote them by z_1, z_2 and z_3) with positive definite matrix $\mathbf{A}'=(\langle z_i,z_j\rangle_{\mathbb{H}})_{i,j=1}^{3}$. Should this be the case, one can prove that h_i-g and h_i-h are linearly independent for some $i\in\{1,2,3\}$, so the inequality is strict. First, we have that

- *Let $z_1,\dots,z_m$ be distinct elements in $\mathbb{H}$. If $\mathbf{A}=(\langle z_i,z_j\rangle_{\mathbb{H}})_{i,j=1}^{m}$ is positive definite, then $z_1,\dots,z_m$ are linearly independent.*

Furthermore, we can conclude that

- *If $\mathbb{H}$ denotes a Hilbert space, z_1,z_2,z_3 are distinct elements in $\mathbb{H}$ and the matrix $\mathbf{A}=(\langle z_i,z_j\rangle_{\mathbb{H}})_{i,j=1}^{3}$ is positive definite, then z_i-g and z_i-h are linearly independent for some $i\in\{1,2,3\}$ whatever $g,h\in\mathbb{H}$ with $g\neq h$ may be.*

  Indeed, by *reductio ad absurdum* assume that z_i-g and z_i-h are linearly dependent for all $i\in\{1,2,3\}$. Then, there exists $(\varsigma_i,\kappa_i)\neq(0,0)$ for all $i\in\{1,2,3\}$ such that

  $$\varsigma_1(z_1-g)+\kappa_1(z_1-h)=0,\ \varsigma_2(z_2-g)+\kappa_2(z_2-h)=0,\ \varsigma_3(z_3-g)+\kappa_3(z_3-h)=0.$$

  Since $g\neq h$, then $\varsigma_i+\kappa_i\neq 0$ for all $i\in\{1,2,3\}$, and two cases are to be discussed, namely,

  - Case $\varsigma_2=0$: In such a situation, $\kappa_2(z_2-h)=0$, so $h=z_2$. Furthermore, $\varsigma_1\neq 0$ (otherwise, $\kappa_1(z_1-h)=0$ and $h=z_1$, but $z_1\neq z_2$) and analogously $\varsigma_3\neq 0$. As a consequence, $\varsigma_1(z_1-g)+\kappa_1(z_1-h)=0$ and $\varsigma_3(z_3-g)$ $+\kappa_3(z_3-h)=0$ assure that

    $$\frac{\kappa_1}{\varsigma_1}(z_1-z_2)+z_1=g,\quad \frac{\kappa_3}{\varsigma_3}(z_3-z_2)+z_3=g,$$

and hence

$$g = \frac{\varsigma_3 + \kappa_3}{\varsigma_3} z_1 - \frac{\kappa_3}{\varsigma_3} z_2, \quad g = \frac{\varsigma_1 + \kappa_1}{\varsigma_1} z_1 - \frac{\kappa_1}{\varsigma_1} z_2$$

or, equivalently,

$$\frac{\varsigma_1 + \kappa_1}{\varsigma_1} z_1 + \left(\frac{\kappa_3}{\varsigma_3} - \frac{\kappa_1}{\varsigma_1} \right) z_2 - \frac{\varsigma_3 + \kappa_3}{\varsigma_3} z_3 = 0$$

with at least one of the coefficients being non-zero, what leads to a contradiction because z_1, z_2 and z_3 are linearly independent under the positive definiteness of $\mathbf{A}$.

– Case $\varsigma_2 \neq 0$: In this situation, $\varsigma_2 (\varsigma_1(z_1 - g) + \kappa_1(z_1 - h)) = 0$ and $\varsigma_1 (\varsigma_2(z_2 - g) + \kappa_2(z_2 - h)) = 0$, whence

$$(\varsigma_1\kappa_2 - \varsigma_2\kappa_1)h = -\varsigma_2(\varsigma_1 + \kappa_1)z_1 + \varsigma_1(\varsigma_2 + \kappa_2)z_2.$$

Note that $\varsigma_1\kappa_2 - \varsigma_2\kappa_1 \neq 0$ (otherwise, $-\varsigma_2(\varsigma_1 + \kappa_1)z_1 + \varsigma_1(\varsigma_2 + \kappa_2)z_2 = 0$ and, because of z_1 and z_2 being linearly independent, $\varsigma_2(\varsigma_1 + \kappa_1) = 0$ and hence $\varsigma_2 = 0$ which is not the case).

Therefore, h can be expressed as $h = \gamma_1 z_1 + \gamma_2 z_2$, with

$$\gamma_1 = \frac{-\varsigma_2(\varsigma_1 + \kappa_1)}{\varsigma_1\kappa_2 - \varsigma_2\kappa_1}, \quad \gamma_2 = \frac{\varsigma_1(\varsigma_2 + \kappa_2)}{\varsigma_1\kappa_2 - \varsigma_2\kappa_1}.$$

Similarly, one can argue that h can be expressed as $h = \gamma_3 z_2 + \gamma_4 z_3$, with

$$\gamma_3 = \frac{-\varsigma_3(\varsigma_2 + \kappa_2)}{\varsigma_2\kappa_3 - \varsigma_3\kappa_2}, \quad \gamma_4 = \frac{\varsigma_2(\varsigma_3 + \kappa_3)}{\varsigma_2\kappa_3 - \varsigma_3\kappa_2}.$$

Consequently, $\gamma_1 z_1 + \gamma_2 z_2 = \gamma_3 z_2 + \gamma_4 z_3$, so $\gamma_1 z_1 + (\gamma_2 - \gamma_3)z_2 - \gamma_4 z_3 = 0$. The linear independence of z_1, z_2 and z_3 entails that $\gamma_1 = 0$, $\gamma_2 = \gamma_3$ and $\gamma_4 = 0$, but $\gamma_1 = 0$ implies that $\varsigma_2 = 0$ (since $\varsigma_1 + \kappa_1 \neq 0$), which is not the case.

Then, we have that

$$J(\lambda g + (1-\lambda)h) < \frac{1}{n}\sum_{i=1}^{n} \rho(\lambda\|h_i - g\|_{\mathbb{H}} + (1-\lambda)\|h_i - h\|_{\mathbb{H}}),$$

and because of the convexity of ρ we have that

$$\frac{1}{n}\sum_{i=1}^{n} \rho(\lambda\|h_i - g\|_{\mathbb{H}} + (1-\lambda)\|h_i - h\|_{\mathbb{H}})$$

$$\leq \frac{1}{n}\sum_{i=1}^{n} \lambda\rho(\|h_i - g\|_{\mathbb{H}}) + (1-\lambda)\rho(\|h_i - h\|_{\mathbb{H}}) = \lambda J(g) + (1-\lambda)J(h). \qquad \square$$

Remark 3.1.1. The importance of Proposition 3.1.4 is that J can be strictly convex even for loss functions like the well-known Huber loss [108],

$$\rho_a(x) = \begin{cases} x^2/2 & \text{if } |x| \leq a \\ \\ a(|x| - a/2) & \text{otherwise,} \end{cases}$$

(with $a > 0$ the so-called tuning parameter) which is convex, but not strictly convex. Huber's loss mean a hybrid approach between squared and absolute error losses, so that it corresponds to a parabola in the vicinity of 0 and increases linearly at a given level a so that one can put appropriate emphasis on large and small errors.

Remark 3.1.2. It should be pointed out that, in the context of M-estimates, the fulfillment of the sufficient conditions for ρ in this section should be simply checked on the interval $[0, \infty)$, since ρ is to be applied on the considered norm.

3.1.3 An algorithm to compute the sample location M-estimates of Hilbert space-valued random elements

In Kim and Scott [114] an iterative algorithm is proposed for the computation of the sample location M-estimate in the considered setting, since the problem $\arg\min_{g \in \mathbb{H}} \sum_{i=1}^n \rho(\|h_i - g\|_{\mathbb{H}})$ does not have an explicit solution in general. The algorithm in this subsection, which is an extension of the iteratively re-weighted least squares algorithm used in classical M-estimation (see Huber [106]), will be adapted to cover the M-estimation in general Hilbert spaces.

Step 1. Take the initial weights $w_i^{(0)} \in \mathbb{R}$, for $i \in \{1, \ldots, n\}$, such that $w_i^{(0)} \geq 0$ and $\sum_{i=1}^n w_i^{(0)} = 1$. Fix a tolerance ε.

Step 2. Generate a sequence $\{g_{(k)}^M\}_{k \in \mathbb{N}}$ by iterating on the following procedure:

$$g_{(k)}^M = \sum_{i=1}^n w_i^{(k-1)} \cdot h_i, \quad w_i^{(k)} = \frac{\phi(\|h_i - g_{(k)}^M\|_{\mathbb{H}})}{\sum_{j=1}^n \phi(\|h_j - g_{(k)}^M\|_{\mathbb{H}})}.$$

Step 3. Terminate the algorithm when

$$\frac{|J(g_{(k+1)}^M) - J(g_{(k)}^M)|}{J(g_{(k)}^M)} < \varepsilon.$$

This procedure can be interpreted as looking for a fixed point of the function

$$f((w_1, \ldots, w_n)) = \frac{\phi(\|h_i - \sum_{i=1}^n w_i \cdot h_i\|_{\mathbb{H}})}{\sum_{j=1}^n \phi(\|h_j - \sum_{i=1}^n w_i \cdot h_i\|_{\mathbb{H}})}.$$

For the computation of $\|h_j - g^M_{(k)}\|_{\mathbb{H}}$ take into account that

$$\|h_j - g^M_{(k)}\|^2_{\mathbb{H}} = \langle h_j - g^M_{(k)}, h_j - g^M_{(k)}\rangle_{\mathbb{H}} = \langle h_j, h_j\rangle_{\mathbb{H}} - 2\langle h_j, g^M_{(k)}\rangle_{\mathbb{H}} + \langle g^M_{(k)}, g^M_{(k)}\rangle_{\mathbb{H}}$$

and, as $g^M_{(k)} = \sum_{i=1}^n w_i^{(k-1)} \cdot h_i$,

- $\langle h_j, g^M_{(k)}\rangle_{\mathbb{H}} = \sum_{i=1}^n w_i^{(k-1)} \langle h_j, h_i\rangle_{\mathbb{H}}$
- $\langle g^M_{(k)}, g^M_{(k)}\rangle_{\mathbb{H}} = \sum_{i=1}^n \sum_{l=1}^n w_i^{(k-1)} w_l^{(k-1)} \langle h_i, h_l\rangle_{\mathbb{H}}$,

and hence

$$\|h_j - g^M_{(k)}\|^2_{\mathbb{H}} = \langle h_j, h_j\rangle_{\mathbb{H}} - 2\sum_{i=1}^n w_i^{(k-1)} \langle h_j, h_i\rangle_{\mathbb{H}} + \sum_{i=1}^n \sum_{l=1}^n w_i^{(k-1)} w_l^{(k-1)} \langle h_i, h_l\rangle_{\mathbb{H}}.$$

Note that the computational complexity is $O(n^2)$ per iteration.

The convergence of the algorithm in terms of $\{J(g^M_{(k)})\}_{k=1}^\infty$ and $\{g^M_{(k)}\}_{k=1}^\infty$ is now to be characterized. Thus,

Theorem 3.1.5. *Under the assumptions:*

- *ρ is non-decreasing, $\rho(0) = 0$ and $\lim_{x\to 0} \rho(x)/x = 0$,*
- *$\phi(0)$ exists and it is finite,*
- *ρ, ρ' and ϕ are continuous,*
- *ϕ is non-increasing,*

the sequence $\{J(g^M_{(k)})\}_{k=1}^\infty$ monotonically decreases and converges. Also, the set $S = \{g \in \mathbb{H} : V(g) = 0\}$ is nonempty and

$$\|g^M_{(k)} - S\|_{\mathbb{H}} \triangleq \inf_{g\in S} \|g^M_{(k)} - g\|_{\mathbb{H}} \underset{k\to\infty}{\longrightarrow} 0,$$

where, as it has already been said, $\{g^M_{(k)}\}_{k=1}^\infty$ denotes the sequence produced by the algorithm.

Proof. First, the monotone decreasing property of $\{J(g^M_{(k)})\}_{k=1}^\infty$ is to be proved. Given $r \in \mathbb{R}$, define the function

$$u(x; r) = \rho(r) - \frac{1}{2} r\rho'(r) + \frac{1}{2}\phi(r)x^2.$$

Since ϕ is assumed to be non-increasing, then u is a surrogate function of ρ (i.e., it is a function that mimics most of the properties of ρ, but it is much simpler either analytically or computationally). In Huber [108] (pp. 184–186), the following properties are proved:

– $u(r;r) = \rho(r)$;

– $u(x;r) \geq \rho(x)$ for all x.

The first property is trivial. To prove the second one, consider the difference

$$z(x) = u(x;r) - \rho(x) = \rho(r) - \frac{1}{2} r\rho'(x) + \frac{1}{2}\phi(r)x^2 - \rho(x).$$

It satisfies

$$z(r) = z(-r) = 0, \quad z'(r) = z'(-r) = 0,$$

since $z'(x) = \rho'(r) \cdot x/r - \rho'(x)$ and $\rho'(-r) = -\rho'(r)$. Moreover, $\phi(x) = \rho'(x)/x$ is supposed to be decreasing for $x > 0$, so that

$$z'(x) = \begin{cases} \frac{\rho'(r)}{r} x - \rho'(x) \leq \frac{\rho'(x)}{x} x - \rho'(x) = 0 & \text{if} \quad 0 < x \leq r \\ \\ \frac{\rho'(r)}{r} x - \rho'(x) \geq \frac{\rho'(x)}{x} x - \rho'(x) = 0 & \text{if} \quad x \geq r. \end{cases}$$

Hence, r is a minimum and $z(x) \geq z(r) = 0$ for $x \geq 0$ (and, because of the symmetry, $z(x) \geq 0$ also for $x \leq 0$). Then, $u(x;r) \geq \rho(x)$ for all x.

Now, define

$$Q(g; g^M_{(k)}) = \frac{1}{n} \sum_{i=1}^{n} u(\|h_i - g\|_{\mathbb{H}}, \|h_i - g^M_{(k)}\|_{\mathbb{H}}).$$

Since ρ' and ϕ are assumed to be continuous, then $Q(\cdot;\cdot)$ is continuous in both arguments, that is,

– $Q(x;\cdot)$ is continuous because ρ, ρ', ϕ and $\|\cdot\|_{\mathbb{H}}$ are continuous and $Q(x;\cdot)$ corresponds to the sum, product and composition of continuous functions;

– $Q(\cdot;r)$ is obviously continuous because of being a quadratic form.

Then, by using that $u(r;r) = \rho(r)$,

$$\begin{aligned} Q(g^M_{(k)}; g^M_{(k)}) &= \frac{1}{n} \sum_{i=1}^{n} u(\|h_i - g^M_{(k)}\|_{\mathbb{H}}, \|h_i - g^M_{(k)}\|_{\mathbb{H}}) \\ &= \frac{1}{n} \sum_{i=1}^{n} \rho(\|h_i - g^M_{(k)}\|_{\mathbb{H}}) = J(g^M_{(k)}). \end{aligned} \tag{3.1}$$

By using that $u(x;r) \geq \rho(x)$, for all $g \in \mathbb{H}$

$$\begin{aligned} Q(g; g^M_{(k)}) &= \frac{1}{n} \sum_{i=1}^{n} u(\|h_i - g\|_{\mathbb{H}}, \|h_i - g^M_{(k)}\|_{\mathbb{H}}) \\ &\geq \frac{1}{n} \sum_{i=1}^{n} \rho(\|h_i - g\|_{\mathbb{H}}) = J(g). \end{aligned} \tag{3.2}$$

The next iterate $g^M_{(k+1)}$ is the minimizer of $Q(g; g^M_{(k)})$, since

$$g^M_{(k+1)} = \sum_{i=1}^n w_i^{(k)} h_i = \sum_{i=1}^n \frac{\phi(\|h_i - g^M_{(k)}\|_{\mathbb{H}})}{\sum_{j=1}^n \phi(\|h_j - g^M_{(k)}\|_{\mathbb{H}})} \cdot h_i$$

$$= \arg\min_{g\in\mathbb{H}} \sum_{i=1}^n \phi(\|h_i - g^M_{(k)}\|_{\mathbb{H}}) \cdot \|h_i - g\|^2_{\mathbb{H}}.$$

The last equality holds because of the weighted mean minimizing the weighted sum of squared norms $\sum_{i=1}^n \phi(\|h_i - g^M_{(k)}\|_{\mathbb{H}}) \cdot \|h_i - g\|^2_{\mathbb{H}}$. Therefore,

$$g^M_{(k+1)} = \arg\min_{g\in\mathbb{H}} \sum_{i=1}^n \rho(\|h_i - g^M_{(k)}\|_{\mathbb{H}})$$

$$-\frac{1}{2}\|h_i - g^M_{(k)}\|_{\mathbb{H}}\rho'(\|h_i - g^M_{(k)}\|_{\mathbb{H}}) + \frac{1}{2}\phi(\|h_i - g^M_{(k)}\|_{\mathbb{H}})$$

$$= \arg\min_{g\in\mathbb{H}} \frac{1}{n}\sum_{i=1}^n u(\|h_i - g\|_{\mathbb{H}}, \|h_i - g^M_{(k)}\|_{\mathbb{H}}) = \arg\min_{g\in\mathbb{H}} Q(g; g^M_{(k)}),$$

whence

$$g^M_{(k+1)} = \arg\min_{g\in\mathbb{H}} Q(g; g^M_{(k)}), \tag{3.3}$$

i.e., $Q(g^M_{(k+1)}; g^M_{(k)}) \leq Q(g; g^M_{(k)})$ for all $g \in \mathbb{H}$. So,

$$J(g^M_{(k)}) = Q(g^M_{(k)}; g^M_{(k)}) \geq Q(g^M_{(k+1)}; g^M_{(k+1)}) \geq J(g^M_{(k+1)}),$$

by sequentially applying Equations 3.1, 3.3 and 3.2.

Next, it is to be proved that every limit point m^* of the sequence $\{g^M_{(k)}\}_{k=1}^\infty$ belongs to $S = \{g \in \mathbb{H} : V(g) = 0\}$. For this purpose, it is first to be checked that

- *Given $h_1, \ldots, h_n$ elements of $\mathbb{H}$, let $\mathbb{G}_n \subset \mathbb{H}$ be defined as $\mathbb{G}_n = \{g \in \mathbb{H} : g = \sum_{i=1}^n w_i h_i,\ w_i \geq 0,\ \sum_{i=1}^n w_i = 1\}$. The set $\mathbb{G}_n$ is compact.*

  It is enough to notice that $\mathbb{G}_n = W(A)$, where $A = \{(w_1, \ldots, w_n) \in \mathbb{R}^n : w_i \geq 0,\ \sum_{i=1}^n w_i = 1\}$ is compact and

$$\begin{array}{llll} W: & A & \longrightarrow & \mathbb{H} \\ & (w_1, \ldots, w_n) & \longmapsto & W((w_1, \ldots, w_n)) = \sum_{i=1}^n w_i h_i \end{array}$$

  is continuous.

Theorem 3.1.2 ensures that $g^M_{(k)} = \sum_{i=1}^n w_i^{(k-1)} h_i$ with $w_i^{(k-1)} \geq 0$ and $\sum_{i=1}^n w_i^{(k-1)} = 1$, so $g^M_{(k)}$ belongs to $\mathbb{G}_n$ for all k. Then, $\{g^M_{(k)}\}_{k=1}^\infty$ has a convergent subsequence $\{g^M_{(k_l)}\}_{l=1}^\infty$. If m^* is the limit of $\{g^M_{(k_l)}\}_{l=1}^\infty$, then

$$Q(g^M_{(k_{l+1})}; g^M_{(k_{l+1})}) = J(g^M_{(k_{l+1})}) \leq J(g^M_{(k_l+1)}) \leq Q(g^M_{(k_l+1)}; g^M_{(k_l)}) \leq Q(g; g^M_{(k_l)}),$$

for all $g \in \mathbb{H}$ using Equation 3.1, the property of J of monotonically decreasing $(k_{l+1} \geq k_l + 1)$ and Equations 3.2 and 3.3 (in this order).

By taking the limit on both sides of the inequality,

$$Q(m^*; m^*) \leq Q(g; m^*) \text{ for all } g \in \mathbb{H},$$

since Q is continuous for both arguments. Therefore, by following the reasoning in Equation 3.3,

$$m^* = \arg\min_{g \in \mathbb{H}} Q(g; m^*) = \sum_{i=1}^{n} \frac{\phi(\|h_i - m^*\|_{\mathbb{H}})}{\sum_{j=1}^{n} \phi(\|h_j - m^*\|_{\mathbb{H}})} h_i,$$

and thus, either

$$\left(\sum_{i=1}^{n} \frac{\phi(\|h_i - m^*\|_{\mathbb{H}})}{\sum_{j=1}^{n} \phi(\|h_j - m^*\|_{\mathbb{H}})} h_i \right) - \frac{\sum_{j=1}^{n} \phi(\|h_j - m^*\|_{\mathbb{H}})}{\sum_{j=1}^{n} \phi(\|h_j - m^*\|_{\mathbb{H}})} m^* = 0$$

or

$$\sum_{i=1}^{n} \frac{\phi(\|h_i - m^*\|_{\mathbb{H}})}{\sum_{j=1}^{n} \phi(\|h_j - m^*\|_{\mathbb{H}})} \cdot (h_i - m^*) = 0.$$

As the denominator does not depend on index i, one gets that

$$\sum_{i=1}^{n} \phi(\|h_i - m^*\|_{\mathbb{H}}) \cdot (h_i - m^*) = 0.$$

So $m^* \in S$ by recalling that $V(g) = \frac{1}{n} \sum_{i=1}^{n} \phi(\|h_i - g\|_{\mathbb{H}}) \cdot (h_i - g)$.

Finally, by *reductio ad absurdum* one can verify that $\|g^M_{(k)} - S\|_{\mathbb{H}} \to 0$. Suppose that $\inf_{g \in S} \|g^M_{(k)} - g\|_{\mathbb{H}} \nrightarrow 0$. Then, there exists $\varepsilon > 0$ such that for all $k \in \mathbb{N}$, there exists $k_0 \geq k$ with $\inf_{g \in S} \|g^M_{(k_0)} - g\|_{\mathbb{H}} \geq \varepsilon$.

Thus, an increasing sequence of indices $\{k_l\}_{l=1}^{\infty}$ such that $\inf_{g \in S} \|g^M_{(k_l)} - g\|_{\mathbb{H}} \geq \varepsilon$ for all $l \in \mathbb{N}$ can be constructed.

Since $\{g^M_{(k_l)}\}_{l=1}^{\infty}$ lies in the compact set $\mathbb{G}_n$, there is a subsequence from it converging to some m_0. One can choose j such that $\|g^M_{(k_j)} - m_0\|_{\mathbb{H}} < \frac{\varepsilon}{2}$.

Moreover, m_0 is also a limit point of $\{g^M_{(k)}\}_{k=1}^{\infty}$, so $m_0 \in S$. But one gets the contradiction that

$$\varepsilon \leq \inf_{g \in S} \|g^M_{(k_j)} - g\|_{\mathbb{H}} \leq \|g^M_{(k_j)} - m_0\|_{\mathbb{H}} < \frac{\varepsilon}{2}. \qquad \square$$

Theorem 3.1.5 can be interpreted by saying that $g^M_{(k)}$ gets arbitrarily close to the set of stationary points of J (those $g \in \mathbb{H}$ for which $\delta J(g; h) = 0$ for all $h \in \mathbb{H}$) as the number of iterations grows.

Corollary 3.1.6. *If the strict convexity of J is added to the assumptions in Theorem 3.1.5, then $\{g^M_{(k)}\}_{k=1}^{\infty}$ converges to $\widehat{g^M[\mathbf{h_n}]}$ in the $\mathbb{H}$-norm.*

Proof. Under the strict convexity of J, there is a unique minimum and, therefore, only one element in S. We know that $\widehat{g^M[\mathbf{h_n}]} \in S$, so $S = \{\widehat{g^M[\mathbf{h_n}]}\}$ and

$$\|g^M_{(k)} - \widehat{g^M[\mathbf{h_n}]}\|_{\mathbb{H}} = \inf_{g \in S} \|g^M_{(k)} - g\|_{\mathbb{H}} \underset{k\to\infty}{\longrightarrow} 0,$$

using the convergence proved in Theorem 3.1.5. □

3.1.4 Consistency and robustness of the sample M-estimates of Hilbert space-valued random elements

In analyzing the inferential behaviour of the sample M-estimates from Hilbert space-valued random elements, we are first going to prove their strong consistency. Although Vandermeulen and Scott [208] have discussed such a problem within Kim and Scott's settings, the conditions assumed on the loss function are rather restrictive, and they do not cover some of the best known loss functions. Consequently, developments have not been adapted from it. On the other hand, under the assumptions for the representer theorem, one can attempt to prove consistency by applying limit theorems for randomly weighted means of Hilbert space-valued random elements, but in this case weights depend on the random elements in the sample, and the required results are beyond the scope of this work.

The approach that has been considered involve sufficient assumptions to ensure Huber's conditions for consistency in [107], that is, those in the following:

Lemma 3.1.7. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, the parameter set Λ be a a locally compact space with a countable base, and $q(\omega, g)$ be some real-valued function on $\Omega \times \Lambda$. Assume that $\omega_1, \omega_2, \ldots$ are independent random variables with values in Ω having the common probability distribution P. Let $T_n(\omega_1, \ldots, \omega_n)$ be any sequence of functions $T_n : \Omega^n \to \Lambda$, measurable or not, such that*

$$\lim_{n\to\infty} \left[\frac{1}{n} \sum_{i=1}^{n} q(\omega_i, T_n(\omega_1, \ldots, \omega_n)) - \inf_{g \in \Lambda} \frac{1}{n} \sum_{i=1}^{n} q(\omega_i, g) \right] = 0 \qquad a.s.\ [P].$$

Then, under Huber's sufficient conditions (see Huber [107]), namely,

- *Assumption (A-1): For each fixed $g_0 \in \Lambda$, the function*

$$q_0 : \Omega \longrightarrow \mathbb{R}, \quad \omega \longmapsto q(\omega, g_0)$$

is $\mathcal{A}$-measurable and separable in Doob's sense (i.e., there is a P-null set N and a countable subset $S \subset \Lambda$ such that for every open set $U \subset \Lambda$ and every closed interval A, the sets

$$V_1 = \{\omega : q(\omega, g) \in A \text{ for all } g \in U\},\ V_2 = \{\omega : q(\omega, g) \in A \text{ for all } g \in U \cap S\}$$

differ by at most a subset of N);

- *Assumption (A-2): The function q is a.s. lower semi-continuous in g_0, that is,*

$$\inf_{g \in U} q(\omega, g) \longrightarrow q(\omega, g_0),$$

as the neighborhood U of g_0 shrinks to $\{g_0\}$;

- *Assumption (A-3): There is a measurable function $a : \Omega \to \mathbb{R}$ such that*

$$E[q(\omega, g) - a(\omega)]^- < \infty \quad \text{for all } g \in \Lambda,$$
$$E[q(\omega, g) - a(\omega)]^+ < \infty \quad \text{for some } g \in \Lambda;$$

Thus, $\gamma(g) = E[q(\omega, g) - a(\omega)]$ is well-defined for all g;

- *Assumption (A-4): There is a $g_0 \in \Lambda$ such that $\gamma(g) > \gamma(g_0)$ for all $g \neq g_0$;*
- *Assumption (A-5): There is a continuous function $b(g) > 0$ such that*
  - *for some integrable h,*

$$\inf_{g \in \Lambda} \frac{q(\omega, g) - a(\omega)}{b(g)} \geq h(\omega);$$

  - *the following condition is satisfied:*

$$\liminf_{g \to \infty} b(g) > \gamma(g_0);$$

  where ∞ denotes here the point at infinity in its one-point compactification;
  - *it is also fulfilled that*

$$E\left[\liminf_{g \to \infty} \frac{q(\omega, g) - a(\omega)}{b(g)}\right] \geq 1;$$

the sequence $\{T_n\}_n$ converges almost surely to g_0.

On the basis of this supporting result, one can state that

Theorem 3.1.8. *Let X be a Hilbert-valued random element associated with a probability space $(\Omega, \mathcal{A}, P)$. Under the assumptions:*

- *ρ is continuous, non-decreasing and subadditive,*
- *Λ, the parameter set, is a locally compact space with a countable basis,*

- *the population M-estimate of location,* $g_P^M[X]$, *exists and it is unique,*
- $E\left[\liminf_{g\to\infty} \frac{\rho(\|X(\omega)-g\|_{\mathbb{H}})-\rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g\|_{\mathbb{H}})+1}\right] \geq 1,$

the sample M-estimate of location, $\widehat{g^M[X]}_n$, *is a strongly consistent estimator of* $g_P^M[X]$.

Proof. To prove the strong consistency, we are going to check that all Huber's conditions required in Lemma 3.1.7 are fulfilled under the assumptions in this theorem. The first condition of being the parameter set Λ a locally compact space with a countable base has been directly imposed in this theorem.

If q is the following real-valued function

$$q : \Omega \times \Lambda \to \mathbb{R}, \quad (\omega, g) \longmapsto q(\omega, g) := \rho(\|X(\omega) - g\|_{\mathbb{H}}),$$

then the other assumptions fulfill. Thus,

- Assuming that $\omega_1, \omega_2, \ldots$ are independent Ω-valued random elements with common probability distribution P, the sequence of functions $\{T_n\}_{n\in\mathbb{N}}$, defined as $T_n(\omega_1, \ldots, \omega_n) = \widehat{g^M[(X(\omega_1), \ldots, X(\omega_n))]}_n$, satisfies that

$$\lim_{n\to\infty}\left[\frac{1}{n}\sum_{i=1}^{n} q(\omega_i, T_n(\omega_1, \ldots, \omega_n)) - \inf_{g\in\Lambda}\frac{1}{n}\sum_{i=1}^{n} q(\omega_i, g)\right] = 0,$$

i.e.,

$$\lim_{n\to\infty}\left[\frac{1}{n}\sum_{i=1}^{n} \rho(\|X(\omega_i) - \widehat{g^M[(X(\omega_1), \ldots, X(\omega_n))]}_n)\|_{\mathbb{H}}) - \inf_{g\in\Lambda}\frac{1}{n}\sum_{i=1}^{n} \rho(\|X(\omega_i) - g\|_{\mathbb{H}})\right] = 0,$$

because of the definition of sample M-estimate of location of a Λ-valued random element (which is also an element of Λ).

- *(A-1)* For each fixed $g_0 \in \Lambda$, the function q_0

$$\begin{array}{rccl} q_0 : & \Omega & \longrightarrow & \mathbb{R} \\ & \omega & \longmapsto & q_0(\omega) = q(\omega, g_0) = \rho(\|X(\omega) - g\|_{\mathbb{H}}) \end{array}$$

is $\mathcal{A}$-measurable (because both the norm and ρ are continuous and X is $\mathcal{A}$-measurable). It is also separable in Doob's sense. Thus, Λ is second-countable (that is to say, it has a countable basis) and that implies the separability. Therefore, Λ contains a countable dense subset, say S. Then, for every open set $U \subset \Lambda$ and every closed interval A, it will be seen that the sets

$$V_1 = \{\omega : q(\omega, g) \in A \text{ for all } g \in U\}, \; V_2 = \{\omega : q(\omega, g) \in A \text{ for all } g \in U \cap S\}$$

coincide. Obviously, $V_1 \subseteq V_2$. On the other hand, by *reductio ad absurdum* suppose that $V_2 \cap V_1^c \neq \emptyset$. If $\omega_0 \in V_2 \cap V_1^c$, then

- Since $\omega_0 \in V_2$, $q(\omega_0, g) \in A$ for all $g \in U \cap S$;
- Since $\omega_0 \in V_1^c$, there exists $g_0 \in U$ such that $q(\omega_0, g_0) \in A^c$. A^c is an open set, so there exists a ball of radius $r > 0$ such that

$$(q(\omega_0, g_0) - r, q(\omega_0, g_0) + r) \subseteq A^c.$$

Notice now that, for a fixed $\omega_0 \in \Omega$, the function

$$q_{\omega_0} : \Lambda \longrightarrow \mathbb{R}, \quad g \longmapsto q_{\omega_0}(g) = q(\omega_0, g) = \rho(\|X(\omega_0) - g\|_{\mathbb{H}})$$

is continuous (since both the norm and ρ are continuous). Therefore, the set $q_{\omega_0}^{-1}(q(\omega_0, g_0) - r, q(\omega_0, g_0) + r)$ is an open subset of Λ and, furthermore, $U \cap q_{\omega_0}^{-1}(q(\omega_0, g_0) - r, q(\omega_0, g_0) + r) \neq \emptyset$ too since g_0 belongs to the intersection. S is a dense set of Λ, so that

$$U \cap q_{\omega_0}^{-1}(q(\omega_0, g_0) - r, q(\omega_0, g_0) + r) \cap S \neq \emptyset.$$

Let $g^* \in U \cap q_{\omega_0}^{-1}(q(\omega_0, g_0) - r, q(\omega_0, g_0) + r) \cap S$. Then, $g^* \in U \cap S$, whence $q(\omega_0, g^*) \in A$. But also,

$$q(\omega_0, g^*) \in (q(\omega_0, g_0) - r, q(\omega_0, g_0) + r) \subset A^c.$$

This is a contradiction, so the conclusion is that $V_2 \subseteq V_1$, and hence both subsets coincide.

- *(A-2)* Indeed, this assumption can be proved for all $\omega \in \Omega$. Let ω be an arbitrary element of Ω and let g_0 be any (fixed) point of Λ.

  First, notice that any sequence of neighborhoods $\{U_n\}_{n\in\mathbb{N}}$ of g_0 with $U_n \downarrow$ satisfies that

$$\left\{\inf_{g\in U_n} q(\omega, g)\right\}_n = \left\{\inf_{g\in U_n} \rho(\|X(\omega) - g\|_{\mathbb{H}})\right\}_n$$

  is a monotonically increasing sequence. Moreover, this sequence is bounded since

$$\inf_{g\in U_n} \rho(\|X(\omega) - g\|_{\mathbb{H}}) \leq \rho(\|X(\omega) - g_0\|_{\mathbb{H}})$$

  for all $n \in \mathbb{N}$ because $g_0 \in \cap_{n=1}^{\infty} U_n$. Therefore, the sequence converges to its supremum, which corresponds to $\rho(\|X(\omega) - g_0\|_{\mathbb{H}})$. By *reductio ad absurdum*, suppose that there is a smaller upper bound

$$c = \rho(\|X(\omega) - g_0\|_{\mathbb{H}}) - \varepsilon,$$

for an arbitrary $\varepsilon > 0$. Let's denote by U_{n_0} a neighborhood of g_0 satisfying that $U_{n_0} \subseteq B(g_0, \varepsilon^*)$, with $\varepsilon^* > 0$ such that $\rho(\varepsilon^*) < \frac{\varepsilon}{2}$. Then, it can be seen that $c < \inf_{g \in U_{n_0}} \rho(\|X(\omega) - g\|_{\mathbb{H}})$, so c cannot be the supremum.

Thus, by using the triangular inequality for norms and the fact that the ρ function is non-decreasing and subadditive,

$$\begin{aligned} \inf_{g \in U_{n_0}} \rho(\|X(\omega) - g\|_{\mathbb{H}}) &\geq \inf_{g \in B(g_0, \varepsilon^*)} \rho(\|X(\omega) - g\|_{\mathbb{H}}) \\ &\geq \inf_{g \in B(g_0, \varepsilon^*)} [\rho(\|X(\omega) - g_0\|_{\mathbb{H}}) - \rho(\|g - g_0\|_{\mathbb{H}})] \\ &= \rho(\|X(\omega) - g_0\|_{\mathbb{H}}) - \sup_{g \in B(g_0, \varepsilon^*)} \rho(\|g - g_0\|_{\mathbb{H}}). \end{aligned}$$

When $g \in B(g_0, \varepsilon^*)$, $\|g - g_0\|_{\mathbb{H}} < \varepsilon^*$ and, by making use again of the non-decreasing property of ρ, one gets that $\rho(\|g - g_0\|_{\mathbb{H}}) \leq \rho(\varepsilon^*) < \varepsilon/2$, whence $\sup_{g \in B(g_0, \varepsilon^*)} \rho(\|g - g_0\|_{\mathbb{H}}) \leq \varepsilon/2 < \varepsilon$ and hence

$$\inf_{g \in U_{n_0}} \rho(\|X(\omega) - g\|_{\mathbb{H}}) > \rho(\|X(\omega) - g_0\|_{\mathbb{H}}) - \varepsilon = c.$$

Now this result is to be extended to general sequences $\{U_n\}_n$. Consider the suprema and the infima radii reached in every neighborhood, namely,

$$r_n = \sup_{g \in U_n} \|g - g_0\|_{\mathbb{H}}, \quad s_n = \inf_{g \in U_n} \|g - g_0\|_{\mathbb{H}}.$$

It is known that $\lim_{n \to \infty} r_n = 0$, since $\{U_n\}_{n \in \mathbb{N}}$ shrinks to $\{g_0\}$. Moreover, $\lim_{n \to \infty} s_n = 0$ since $0 \leq s_n \leq r_n$ for all $n \in \mathbb{N}$.

Let ε be any nonnegative number and consider $\varepsilon^* > 0$ such that $\rho(\varepsilon^*) < \varepsilon/2$. As $\lim_{n \to \infty} r_n = 0$, there exists $n_1 \in \mathbb{N}$ such that for all $n > n_1$, $r_n < \varepsilon^*$. Then, $U_n \subseteq B(g_0, r_n)$ and

$$\begin{aligned} \inf_{g \in U_n} \rho(\|X(\omega) - g\|_{\mathbb{H}}) &\geq \inf_{g \in B(g_0, r_n)} \rho(\|X(\omega) - g\|_{\mathbb{H}}) \\ &\geq \rho(\|X(\omega) - g_0\|_{\mathbb{H}}) - \sup_{g \in B(g_0, r_n)} \rho(\|g - g_0\|_{\mathbb{H}}). \end{aligned}$$

If $g \in B(g_0, r_n)$, then $\|g - g_0\|_{\mathbb{H}} < r_n < \varepsilon^*$ and, as ρ is non-decreasing, $\rho(\|g - g_0\|_{\mathbb{H}}) \leq \rho(\varepsilon^*) < \varepsilon/2$, so that

$$\inf_{g \in U_n} \rho(\|X(\omega) - g\|_{\mathbb{H}}) > \rho(\|X(\omega) - g_0\|_{\mathbb{H}}) - \varepsilon.$$

Analogously, since $\lim_{n \to \infty} s_n = 0$, there exists $n_2 \in \mathbb{N}$ such that for all $n > n_2$, $s_n < \varepsilon^*$. Then, $U_n \supseteq B(g_0, s_n)$ and

$$\inf_{g \in U_n} \rho(\|X(\omega) - g\|_{\mathbb{H}}) \leq \inf_{g \in B(g_0, s_n)} \rho(\|X(\omega) - g\|_{\mathbb{H}})$$

$$\leq \rho(\|X(\omega) - g_0\|_{\mathbb{H}}) + \inf_{g \in B(g_0, s_n)} \rho(\|g - g_0\|_{\mathbb{H}})$$
$$\leq \rho(\|X(\omega) - g_0\|_{\mathbb{H}}) + \sup_{g \in B(g_0, s_n)} \rho(\|g - g_0\|_{\mathbb{H}}).$$

Again, if $g \in B(g_0, s_n)$, $\rho(\|g - g_0\|_{\mathbb{H}}) \leq \rho(\varepsilon^*) < \varepsilon/2$ and

$$\inf_{g \in U_n} \rho(\|X(\omega) - g\|_{\mathbb{H}}) < \rho(\|X(\omega) - g_0\|_{\mathbb{H}}) + \varepsilon.$$

Consequently, for any $\varepsilon > 0$, there exists $n_0 = \max\{n_1, n_2\}$, such that for all $n > n_0$,

$$\rho(\|X(\omega) - g_0\|_{\mathbb{H}}) - \varepsilon < \inf_{g \in U_n} \rho(\|X(\omega) - g\|_{\mathbb{H}}) < \rho(\|X(\omega) - g_0\|_{\mathbb{H}}) + \varepsilon,$$

that is to say,

$$\left| \inf_{g \in U_n} \rho(\|X(\omega) - g\|_{\mathbb{H}}) - \rho(\|X(\omega) - g_0\|_{\mathbb{H}}) \right| < \varepsilon,$$

so the sequence $\left\{ \inf_{g \in U_n} \rho(\|X(\omega) - g\|_{\mathbb{H}}) \right\}_n$ converges to $\rho(\|X(\omega) - g_0\|_{\mathbb{H}})$.

- *(A-3)* Let a be the measurable function (see (*A-1*)):

$$\begin{array}{rccl} a: & \Omega & \longrightarrow & \mathbb{R} \\ & \omega & \longmapsto & \rho(\|X(\omega)\|_{\mathbb{H}}). \end{array}$$

For any arbitrarily fixed $g \in \Lambda$,

$$E[\rho(\|X(\omega) - g\|_{\mathbb{H}}) - a(\omega)]^-$$
$$= \int_\Omega - \min\{\rho(\|X(\omega) - g\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}}), 0\} \, dP(\omega)$$
$$= \int_{\{\omega \in \Omega \,:\, \rho(\|X(\omega)\|_{\mathbb{H}}) > \rho(\|X(\omega) - g\|_{\mathbb{H}})\}} \big[\rho(\|X(\omega)\|_{\mathbb{H}}) - \rho(\|X(\omega) - g\|_{\mathbb{H}})\big] \, dP(\omega).$$

By the triangular inequality and the subadditivity of ρ,

$$E[\rho(\|X(\omega) - g\|_{\mathbb{H}}) - a(\omega)]^-$$
$$\leq \int_{\{\omega \in \Omega \,:\, \rho(\|X(\omega)\|_{\mathbb{H}}) > \rho(\|X(\omega) - g\|_{\mathbb{H}})\}} \big[\rho(\|X(\omega) - g\|_{\mathbb{H}}) + \rho(\|g\|_{\mathbb{H}})$$
$$-\rho(\|X(\omega) - g\|_{\mathbb{H}})\big] \, dP(\omega) = \rho(\|g\|_{\mathbb{H}}) \cdot P\Big(\{\omega \,:\, \rho(\|X(\omega)\|_{\mathbb{H}}) > \rho(\|X(\omega) - g\|_{\mathbb{H}})\}\Big)$$
$$\leq \rho(\|g\|_{\mathbb{H}}) < \infty.$$

Analogously,

$$E[\rho(\|X(\omega) - g\|_{\mathbb{H}}) - a(\omega)]^+$$

$$= \int_\Omega \max\{\rho(\|X(\omega)-g\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}}), 0\}\, dP(\omega)$$
$$= \int_{\{\omega\in\Omega\,:\,\rho(\|X(\omega)\|_{\mathbb{H}})\le\rho(\|X(\omega)-g\|_{\mathbb{H}})\}} \big[\rho(\|X(\omega)-g\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}})\big]\, dP(\omega).$$

By the triangular inequality and the subadditivity of ρ,

$$E[\rho(\|X(\omega)-g\|_{\mathbb{H}}) - a(\omega)]^+$$
$$\le \int_{\{\omega\in\Omega\,:\,\rho(\|X(\omega)\|_{\mathbb{H}})\le\rho(\|X(\omega)-g\|_{\mathbb{H}})\}} \big[\rho(\|X(\omega)\|_{\mathbb{H}}) + \rho(\|g\|_{\mathbb{H}})$$
$$-\rho(\|X(\omega)\|_{\mathbb{H}})\big]\, dP(\omega) = \rho(\|g\|_{\mathbb{H}})\cdot P\Big(\{\omega\,:\,\rho(\|X(\omega)\|_{\mathbb{H}})\le\rho(\|X(\omega)-g\|_{\mathbb{H}})\}\Big)$$
$$\le \rho(\|g\|_{\mathbb{H}}) < \infty.$$

So the second inequality also holds for all $g\in\Lambda$ in this case.

- *(A-4)* The population M-estimate of location exists and it is unique, so that
$$\int_\Omega \rho(\|X(\omega) - g_P^M[X]\|_{\mathbb{H}})dP(\omega) = \min_{g\in\Lambda}\int_\Omega \rho(\|X(\omega)-g\|_{\mathbb{H}}),$$
that is to say,
$$g_P^M[X] = \arg\min_{g\in\Lambda} E[\rho(\|X(\omega)-g\|_{\mathbb{H}})]$$
$$= \arg\min_{g\in\Lambda}\Big(E[\rho(\|X(\omega)-g\|_{\mathbb{H}})] - E[\rho(\|X(\omega)\|_{\mathbb{H}})]\Big)$$
$$= \arg\min_{g\in\Lambda} E\Big[\rho(\|X(\omega)-g\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}})\Big] = \arg\min_{g\in\Lambda}\gamma(g),$$
and $g_0 = g_P^M[X]$ fulfills this assumption.

- *(A-5)* There is a continuous function $b(g) > 0$
$$\begin{array}{rccl} b: & \Lambda & \longrightarrow & \mathbb{R} \\ & g & \longmapsto & b(g) = \rho(\|g\|_{\mathbb{H}}) + 1 \end{array}$$
such that

  – for the integrable function $h(\omega) := -1$,
$$\inf_{g\in\Lambda}\frac{\rho(\|X(\omega)-g\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g\|_{\mathbb{H}})+1} \ge -1,$$
  since by the triangular inequality, $\|X(\omega)-g\|_{\mathbb{H}} \ge \|X(\omega)\|_{\mathbb{H}} - \|g\|_{\mathbb{H}}$, and because of ρ being subadditive
$$\inf_{g\in\Lambda}\frac{\rho(\|X(\omega)-g\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g\|_{\mathbb{H}})+1}$$
$$\ge \inf_{g\in\Lambda}\frac{\rho(\|X(\omega)\|_{\mathbb{H}}) - \rho(\|g\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g\|_{\mathbb{H}})+1} = \inf_{g\in\Lambda}\frac{-\rho(\|g\|_{\mathbb{H}})}{\rho(\|g\|_{\mathbb{H}})+1} \ge -1.$$

– the following condition is satisfied:

$$\liminf_{g\to\infty} b(g) > \gamma(g_0) = E\left[\rho(\|X(\omega)-g_0\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}})\right].$$

Let $\{g_n\}_{n\in\mathbb{N}} \subset \Lambda$ be any sequence with $\lim_{n\to\infty} g_n = \infty$ (and hence, $\lim_{n\to\infty} \|g_n\|_{\mathbb{H}} = \infty$). Then,

$$M = E\left[\rho(\|X(\omega)-g_0\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}})\right] \in \mathbb{R},$$

where g_0 represents the minimum found in *(A-4)*, and let M^* be a real number such that $\rho(M^*) \geq M$. Note that, even if ρ is bounded by C, it is possible to choose this M^* because $M = E\left[\rho(\|X(\omega)-g_0\|_{\mathbb{H}})\right] - E\left[\rho(\|X(\omega)\|_{\mathbb{H}})\right] \leq E\left[\rho(\|X(\omega)-g_0\|_{\mathbb{H}})\right] \leq C$.
Then, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $\|g_n\|_{\mathbb{H}} > M^*$, whence for all $n \geq n_0$,

$$\inf_{k\geq n} b(g_k) = \inf_{k\geq n} \left(\rho(\|g_k\|_{\mathbb{H}}) + 1\right) \geq \rho(M^*) + 1 \geq M + 1.$$

Finally,

$$\liminf_{n\to\infty} b(g_n) = \lim_{n\to\infty} \left(\inf_{k\geq n} b(g_k)\right) \geq M + 1 > M = \gamma(g_0).$$

The third part of Huber's condition *(A-5)* is fulfilled due to the assumption

$$E\left[\liminf_{g\to\infty} \frac{\rho(\|X(\omega)-g\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g\|_{\mathbb{H}}) + 1}\right] \geq 1.$$

□

The following result slightly modifies the statement of Theorem 3.1.8 by replacing the last condition by a sufficient one. More concretely,

Proposition 3.1.9. *Let X be a Hilbert-valued random element associated with a probability space $(\Omega, \mathcal{A}, P)$. The last assumption in Theorem 3.1.8 can be proved to be satisfied when ρ is non-decreasing, subadditive and also unbounded.*

Proof. Under the assumed condition in, in case of ρ is subadditive, one can ensure that

$$E\left[\liminf_{g\to\infty} \frac{\rho(\|X(\omega)-g\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}})}{b(g)}\right] \geq 1.$$

Thus, for any fixed $\omega \in \Omega$,

$$\liminf_{g\to\infty} \frac{\rho(\|X(\omega)-g\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g\|_{\mathbb{H}}) + 1}$$

$$= \lim_{n\to\infty}\left(\inf_{k\geq n}\frac{\rho(\|X(\omega)-g_k\|_{\mathbb{H}})-\rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g_k\|_{\mathbb{H}})+1}\right).$$

The sequence

$$\left\{\inf_{k\geq n}\frac{\rho(\|X(\omega)-g_k\|_{\mathbb{H}})-\rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g_k\|_{\mathbb{H}})+1}\right\}_n$$

is monotonically increasing and it is upper bounded by 1, since for all $k\in\mathbb{N}$, by applying the triangular inequality and the subadditivity of ρ, one gets that

$$\frac{\rho(\|X(\omega)-g_k\|_{\mathbb{H}})-\rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g_k\|_{\mathbb{H}})+1}\leq\frac{\rho(\|g_k\|_{\mathbb{H}})}{\rho(\|g_k\|_{\mathbb{H}})+1}\leq 1.$$

So, it converges to its supremum

$$\lim_{n\to\infty}\left(\inf_{k\geq n}\frac{\rho(\|X(\omega)-g_k\|_{\mathbb{H}})-\rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g_k\|_{\mathbb{H}})+1}\right)$$
$$=\sup_n\left(\inf_{k\geq n}\frac{\rho(\|X(\omega)-g_k\|_{\mathbb{H}})-\rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g_k\|_{\mathbb{H}})+1}\right)$$

We are now going to see that this supremum is at least equal to 1. By *reductio ad absurdum*, suppose that

$$\sup_n\left(\inf_{k\geq n}\frac{\rho(\|X(\omega)-g_k\|_{\mathbb{H}})-\rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g_k\|_{\mathbb{H}})+1}\right)=1-\varepsilon,$$

for some $\varepsilon>0$. One gets then a contradiction because one finds an $n^*\in\mathbb{N}$ such that

$$\inf_{k\geq n^*}\frac{\rho(\|X(\omega)-g_k\|_{\mathbb{H}})-\rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g_k\|_{\mathbb{H}})+1}>1-\varepsilon$$

since for all $k\geq n^*$

$$\frac{\rho(\|X(\omega)-g_k\|_{\mathbb{H}})-\rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g_k\|_{\mathbb{H}})+1}\geq 1-\frac{\varepsilon}{2}>1-\varepsilon.$$

Take, for the fixed arbitrary $\omega\in\Omega$, $M:=\frac{2}{\varepsilon}-1+\frac{4}{\varepsilon}\cdot\rho(\|X(\omega)\|_{\mathbb{H}})\in\mathbb{R}$. Consider, as ρ is unbounded, $M^*>0$ with $\rho(M^*)>M$. Recall that $\lim_{n\to\infty}g_n=\infty$, so there exists $n^*\in\mathbb{N}$ such that for all $n\geq n^*$, $\|g_n\|_{\mathbb{H}}>M^*$. Therefore, by the subadditivity and non-decreasing property of ρ,

$$\rho(\|g_n-X(\omega)\|_{\mathbb{H}})\geq\rho(\|g_n\|_{\mathbb{H}})-\rho(\|X(\omega)\|_{\mathbb{H}})\geq\rho(M^*)-\rho(\|X(\omega)\|_{\mathbb{H}})>M-\rho(\|X(\omega)\|_{\mathbb{H}}).$$

We can easily check that $1-\varepsilon/2$ is a lower bound of the sequence

$$\left\{\frac{\rho(\|X(\omega)-g_k\|_{\mathbb{H}})-\rho(\|X(\omega)\|_{\mathbb{H}})}{\rho(\|g_k\|_{\mathbb{H}})+1}\right\}_{k\geq n^*}.$$

For any $k \geq n^*$,

$$
\begin{aligned}
&\rho(\|X(\omega) - g_k\|_{\mathbb{H}}) - \rho(\|X(\omega)\|_{\mathbb{H}}) \\
&= \left(1 - \frac{\varepsilon}{2}\right) \rho(\|X(\omega) - g_k\|_{\mathbb{H}}) + \frac{\varepsilon}{2}\rho(\|X(\omega) - g_k\|_{\mathbb{H}}) \\
&\quad - \left(1 - \frac{\varepsilon}{2}\right) \rho(\|X(\omega)\|_{\mathbb{H}}) - \frac{\varepsilon}{2}\rho(\|X(\omega)\|_{\mathbb{H}}) \\
&\geq \left(1 - \frac{\varepsilon}{2}\right) \rho(\|g_k\|_{\mathbb{H}}) - \left(1 - \frac{\varepsilon}{2}\right) \rho(\|X(\omega)\|_{\mathbb{H}}) + \frac{\varepsilon}{2}\rho(\|X(\omega) - g_k\|_{\mathbb{H}}) \\
&\quad - \left(1 - \frac{\varepsilon}{2}\right) \rho(\|X(\omega)\|_{\mathbb{H}}) - \frac{\varepsilon}{2}\rho(\|X(\omega)\|_{\mathbb{H}}) \\
&= \left(1 - \frac{\varepsilon}{2}\right) \rho(\|g_k\|_{\mathbb{H}}) + \frac{\varepsilon}{2}\rho(\|X(\omega) - g_k\|_{\mathbb{H}}) - \left(2 - \frac{\varepsilon}{2}\right) \rho(\|X(\omega)\|_{\mathbb{H}}) \\
&> \left(1 - \frac{\varepsilon}{2}\right) \rho(\|g_k\|_{\mathbb{H}}) + \frac{\varepsilon}{2} \left(\frac{2}{\varepsilon} - 1 + \left(\frac{4}{\varepsilon} - 1\right)\rho(\|X(\omega)\|_{\mathbb{H}})\right) \\
&- \left(2 - \frac{\varepsilon}{2}\right) \rho(\|X(\omega)\|_{\mathbb{H}}) = \left(1 - \frac{\varepsilon}{2}\right) \rho(\|g_k\|_{\mathbb{H}}) + 1 - \frac{\varepsilon}{2} = \left(1 - \frac{\varepsilon}{2}\right) (\rho(\|g_k\|_{\mathbb{H}}) + 1).
\end{aligned}
$$

□

The robustness of the M-estimates is now to be analyzed empirically. By means of simulations dealing with functional data, we are going to see that its empirical value for two well-known loss functions fulfilling the required conditions amounts $\frac{1}{n} \cdot \lfloor \frac{n+1}{2} \rfloor$: the Huber loss function (Remark 3.1.1) with $a = 1.345$ (one of the most usual choices for this tuning parameter -see, for instance, Wang *et al.* [213]) and the Hampel loss function [100], which corresponds to

$$
\rho_{a,b,c}(x) = \begin{cases} x^2/2 & \text{if } 0 \leq |x| < a \\ a(|x| - a/2) & \text{if } a \leq |x| < b \\ \dfrac{a(|x| - c)^2}{2(b-c)} + \dfrac{1}{2}a(b + c - a) & \text{if } b \leq |x| < c \\ \dfrac{1}{2}a(b + c - a) & \text{if } c \leq |x|, \end{cases}
$$

where the nonnegative parameters $a < b < c$ allow us to control the degree of supression of the outliers. The smaller their values, the greater this degree. To fix them, the ideas in Kim and Scott [114] have been followed, that is,

1. to choose the initial seed in this case, the mean has been considered;
2. to compute the distances between the observations and the seed;
3. a will be the median, b the 75^{th} percentile and c the 85^{th} percentile of these distances.

Two cases have been considered in the simulation study, namely, an even sample size ($n = 100$) and an odd sample size ($n = 101$). In each situation, a sample of functions have been generated from *Model 1* in Section 2.4 (see p. 74) and M-estimates, using both Huber and Hampel loss functions, have been computed. Afterwards, $i \in \{1, \ldots, n\}$ observations have been highly contaminated (concretely, i functions have been translated 10^7 units) and Huber and Hampel M-estimates have been computed. Distances between them and the non-contaminated estimates for each amount of modified observations are plotted in Figure 3.1. The value in red represents the minimum number of perturbed observations that makes the distance between the non-contaminated and the contaminated corresponding M-estimates increase arbitrarily, i.e., the finite sample breakdown point.

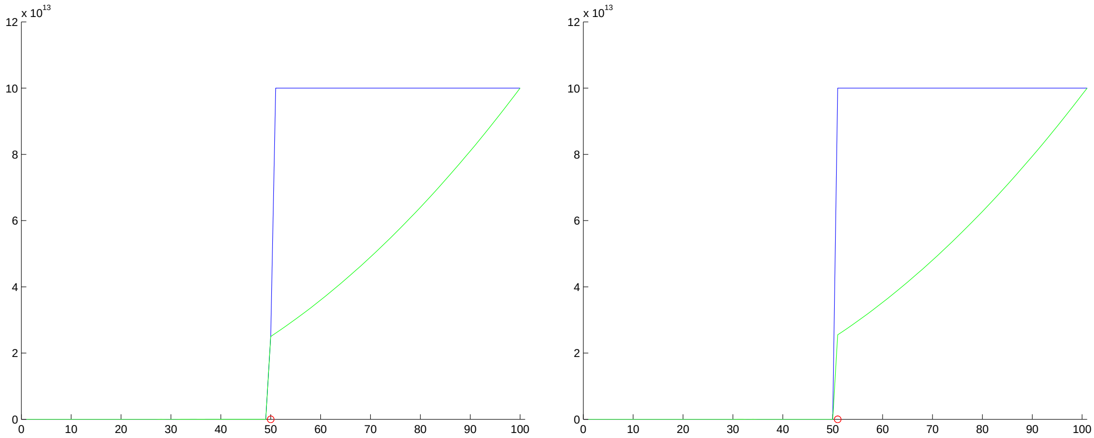


Figure 3.1: Empirical value (in red) obtained for the finite sample breakdown point of the M-estimators for functional data and distances between the non-contaminated and contaminated estimates using the Huber (blue) and Hampel (green) loss functions when the sample size is even (left) and odd (right).

Therefore, it can be checked that the empirical value for the finite sample breakdown point has been 50 for the sample size $n = 100$ and 51 for the sample size $n = 101$ independently of the chosen loss function.

3.1.5 Simulation studies about M-estimates of location for Hilbert space-valued data

The aim of this subsection is to compare the behavior of the M-estimates of location and the mean through some simulation studies dealing with functional data and mimicking the usual scheme for this kind of comparisons. The empirical robustness of the M-estimates of location is also examined.

n	q	M	Estimator	*Model 1*	*Model 2*	*Model 3*	*Model 4*	*Model 5*
50	.05	5	Mean	**.019443** (.019470)	.045089 (.058401)	.100940 (.117672)	.033178 (.033396)	**.020686** (.018855)
			Huber	.023212 (.023754)	**.023310** (.026784)	**.023762** (.024501)	**.022677** (.024247)	.023520 (.024675)
			Hampel	.055884 (.060037)	.044582 (.051090)	.040349 (.049614)	.043753 (.054135)	.048482 (.054478)
80	.05	5	Mean	**.012736** (.014075)	.028164 (.037772)	.080471 (.030194)	.020019 (.025748)	**.013512** (.029245)
			Huber	**.013103** (.013907)	**.015130** (.017157)	**.016623** (.018098)	**.014961** (.014886)	.015280 (.016722)
			Hampel	.028760 (.033564)	.027222 (.030228)	.027940 (.031312)	.028802 (.033663)	.030742 (.034910)
50	.1	5	Mean	**.019443** (.019470)	.074806 (.093568)	.316388 (.268233)	.046796 (.052995)	**.022762** (.022055)
			Huber	.023212 (.023754)	**.026060** (.030354)	**.030726** (.036714)	**.026157** (.028968)	.023460 (.023222)
			Hampel	.055884 (.060037)	.039116 (.044438)	.036172 (.044248)	.041076 (.047470)	.047891 (.053994)
80	.1	5	Mean	**.012736** (.014075)	.046142 (.058836)	.288916 (.213707)	.026058 (.027837)	**.014167** (.012959)
			Huber	.013103 (.013907)	**.017012** (.018024)	**.017706** (.019974)	**.015825** (.018233)	.014523 (.013159)
			Hampel	.028760 (.033564)	.024219 (.026392)	.021057 (.025622)	.025754 (.031279)	.027775 (.031590)
50	.05	25	Mean	**.019443** (.019470)	.598505 (.877871)	2.00705 (2.18651)	.357199 (.452284)	.058920 (.037000)
			Huber	.023212 (.023754)	**.023265** (.024225)	**.026303** (.028609)	**.022207** (.022989)	**.023838** (.023955)
			Hampel	.055884 (.060037)	.035515 (.041124)	.032637 (.041761)	.038481 (.045578)	.045256 (.049924)
80	.05	25	Mean	**.012736** (.014075)	.380566 (.554507)	2.002653 (1.892887)	.214524 (.256844)	.037192 (.021657)
			Huber	**.013103** (.013907)	**.014711** (.016482)	**.014104** (.013847)	**.013355** (.014485)	**.014967** (.016080)
			Hampel	.028760 (.033564)	.023162 (.027398)	.015680 (.017095)	.021675 (.026070)	.030053 (.035918)
50	.1	25	Mean	**.019443** (.019470)	1.43842 (2.09170)	7.67767 (6.20711)	.650730 (.793654)	.104663 (.056480)
			Huber	.023212 (.023754)	**.024824** (.026862)	.024978 (.027394)	**.025024** (.026450)	**.025992** (.028564)
			Hampel	.055884 (.060037)	.030439 (.033359)	**.023353** (.026776)	.033160 (.039795)	.039204 (.043693)
80	.1	25	Mean	**.012736** (.014075)	.721368 (.985035)	7.13882 (4.81952)	.393982 (.482538)	.058355 (.028707)
			Huber	.013103 (.013907)	**.015770** (.017007)	.015915 (.016359)	**.015159** (.016724)	**.016459** (.018256)
			Hampel	.028760 (.033564)	.019124 (.021638)	**.014022** (.014399)	.020958 (.028069)	.025035 (.029074)

Table 3.1: Results of the simulations for functional data to compare the behaviour of the mean and the M-estimates of location using both Huber and Hampel loss functions in each of the *Models 1-5*

n	q	Estimator	*Model 6*	*Model 7*	*Model 8*	*Model 9*
50	.05	Mean	.274843 (.274843)	.318066 (.201209)	**.279586** (.175522)	**.332575** (.221595)
		Huber	**.023834** (.031628)	**.021681** (.023717)	.534812 (.194016)	.534853 (.182469)
		Hampel	.059589 (.079215)	.054867 (.063509)	.557492 (.302287)	.577072 (.306548)
80	.05	Mean	.177057 (.093181)	.202028 (.112033)	**.170289** (.089113)	**.197961** (.103885)
		Huber	**.015659** (.018343)	**.014882** (.018069)	.532897 (.151203)	.535929 (.154593)
		Hampel	.038747 (.047992)	.038024 (.047002)	.563185 (.249383)	.570632 (.254555)
50	.1	Mean	.499788 (.254557)	.558465 (.295773)	**.521980** (.272500)	.590299 (.312491)
		Huber	**.023966** (.027664)	**.023019** (.025300)	.548149 (.190937)	**.539698** (.174565)
		Hampel	.052455 (.063761)	.056005 (.065200)	.593019 (.316606)	.561826 (.295460)
80	.1	Mean	.309428 (.127167)	.357100 (.162768)	**.304456** (.128494)	**.363893** (.155640)
		Huber	**.013608** (.015213)	**.014354** (.016560)	.531472 (.147077)	.539704 (.148112)
		Hampel	.031621 (.038699)	.033124 (.041502)	.552189 (.236073)	.562649 (.251195)

Table 3.2: Results of the simulations for functional data to compare the behaviour of the mean and the M-estimates of location using both Huber and Hampel loss functions in each of the *Models 6-9*

The different models, the value of the parameters and the calculus of the measures coincide with the detailed procedure explained in Section 2.4 (see p. 74). Indeed, notice that Huber and Hampel M-estimates have been computed on the basis of the same generated samples in that section in order to be able to compare them with the trimmed means in the next chapter.

The empirical results obtained can be seen in Tables 3.1 and 3.2. The numbers in parenthesis are the standard errors of the mean squared errors of the estimators and the bold number is the minimum mean squared error obtained in each situation and for each of the models.

For instance, as it happened in the comparison with the trimmed means, independently of the case of study, the best estimator of the population mean is the sample mean in *Model 1*, what is logical because in that model there is no contamination at all.

The estimator chosen is the M-estimator with Huber loss function in almost all the situations in *Models 2-7* (except from half of the cases in *Model 5*, in which the mean was the best option, and two cases in *Model 3*, for which the M-estimate with Hampel loss function behaved better), whereas in *Models 8-9*, the mean was the best estimator in all situations, but one.

One remark is that, apart from the best performance of the Huber loss function with respect to the Hampel loss function in this simulation study, the computation of the latter is slower because the parameters a, b and c depend here on the sample, so they have to be updated not only for every situation and model, but also for every Monte Carlo iteration. Of course, other alternatives to fix these parameters could be analyzed.

3.2 Application of the representer theorem to the set and fuzzy set-valued cases

In order to calculate the M-estimates of location for imprecise-valued random elements, we have to recall that set- and fuzzy- valued data can be identified with functional data thanks to the isometrical embedding of $\mathcal{F}_c(\mathbb{R}^p)$ into a convex cone of a Hilbert space using the support functions, as explained in Section 1.3.1, for certain metrics. More concretely, the norms defined associated with distances D_θ^φ and $\mathfrak{D}_\theta^\varphi$ arise in the usual way from the corresponding inner products, so the results in this section will be applicable to both of them and also to any other metric whose associated norm satisfies the parallelogram law. From now on in this section, D will denote an arbitrary metric on $\mathcal{F}_c(\mathbb{R}^p)$ associated with a norm $\|\cdot\|_D$ (in the Hilbert space in which the space of fuzzy set-valued data is embedded) fulfilling the previous condition and d, an arbitrary metric on $\mathcal{K}_c(\mathbb{R}^p)$ associated with a norm $\|\cdot\|_d$ (in the Hilbert space in which the space of set-valued data is embedded) satisfying the same condition.

Definition 3.2.1. *Consider $\mathcal{F}_c(\mathbb{R}^p)$ with an arbitrary associated metric D. Let $(\Omega, \mathcal{A}, P)$ be a probability space, $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ be an associated random fuzzy vector and ρ be an arbitrary loss function. The* ***population fuzzy M-estimate of location*** *is the element $\widetilde{g}_P^M \in \mathcal{F}_c(\mathbb{R}^p)$ minimizing*

$$J_P(\widetilde{g}) = \int_\Omega \rho(\|s_{\mathcal{X}(\omega)} - s_{\widetilde{g}}\|_D)\, dP(\omega),$$

i.e.,

$$\widetilde{g}_P^M = \arg\min_{\widetilde{g}\in\mathcal{F}_c(\mathbb{R}^p)} E[\rho(D(\mathcal{X}(\omega), \widetilde{g})].$$

In particular, consider $\mathcal{K}_c(\mathbb{R}^p)$ with an arbitrary associated distance d. Let $(\Omega, \mathcal{A}, P)$ be a probability space, $\mathsf{X} : \Omega \to \mathcal{K}_c(\mathbb{R}^p)$ be an associated random compact convex set and ρ be an arbitrary loss function. The ***population set M-estimate of location*** *is the element $\mathsf{g}_P^M \in \mathcal{K}_c(\mathbb{R}^p)$ minimizing*

$$J_P(\mathsf{g}) = \int_\Omega \rho(\|s_{\mathsf{X}(\omega)} - s_{\mathsf{g}}\|_d)\, dP(\omega),$$

i.e.,

$$\mathsf{g}_P^M = \arg \min_{\mathsf{g} \in \mathcal{K}_c(\mathbb{R}^p)} E[\rho(d(\mathsf{X}(\omega), \mathsf{g})].$$

Definition 3.2.2. *Consider $\mathcal{F}_c(\mathbb{R}^p)$ with an arbitrary associated metric D. Let $(\Omega, \mathcal{A}, P)$ be a probability space, $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ be an associated random fuzzy vector and ρ be an arbitrary loss function. The* ***sample fuzzy M-estimate of location*** *is the fuzzy set-valued statistic $\widehat{\widetilde{g}^M[\mathcal{X}]_n}$ such that for each realization from the simple random sample, $\widetilde{\mathbf{x}}_\mathbf{n} = (\widetilde{x}_1, \ldots, \widetilde{x}_n)$, the fuzzy set value(s) $\widehat{\widetilde{g}^M[\widetilde{\mathbf{x}}_\mathbf{n}]}$ is (are) the solution(s) of the following optimization problem:*

$$\min_{\widetilde{g} \in \mathcal{F}_c(\mathbb{R}^p)} J(\widetilde{g}) = \min_{\widetilde{g} \in \mathcal{F}_c(\mathbb{R}^p)} \frac{1}{n} \sum_{i=1}^{n} \rho(D(\widetilde{x}_i, \widetilde{g})).$$

In particular, consider $\mathcal{K}_c(\mathbb{R}^p)$ with an arbitrary associated norm d. Let $(\Omega, \mathcal{A}, P)$ be a probability space, $\mathsf{X} : \Omega \to \mathcal{K}_c(\mathbb{R}^p)$ be an associated random compact convex set and ρ be an arbitrary loss function. The ***sample set M-estimate of location*** *is the fuzzy set-valued statistic $\widehat{\mathsf{g}^M[\mathcal{X}]_n}$ such that for each realization from the simple random sample, $\mathbf{x}_n = (\mathsf{x}_1, \ldots, \mathsf{x}_n)$, the set value(s) $\widehat{\mathsf{g}^M[\mathsf{x}_n]}$ is (are) the solution(s) of the following optimization problem:*

$$\min_{\mathsf{g} \in \mathcal{K}_c(\mathbb{R}^p)} J(\mathsf{g}) = \min_{\mathsf{g} \in \mathcal{K}_c(\mathbb{R}^p)} \frac{1}{n} \sum_{i=1}^{n} \rho(d(\mathsf{x}_i, \mathsf{g})).$$

By particularizing the results in Subsection 3.1.2 to the set- and fuzzy set-valued cases one can briefly state from Theorem 3.1.2 that the sample location M-estimate associated with either a random fuzzy vector or a random compact convex set can be expressed as a convex linear combination of the sample components. Thus,

Theorem 3.2.1. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ be an associated random fuzzy vector and ρ be an arbitrary loss function. Consider $\mathcal{F}_c(\mathbb{R}^p)$ with an arbitrary associated metric D, and let $\widetilde{\mathbf{x}}_\mathbf{n} = (\widetilde{x}_1, \ldots, \widetilde{x}_n)$ be a sample of independent observations from $\mathcal{X}$. Under the assumptions*

- *ρ is non-decreasing, $\rho(0) = 0$ and $\lim_{x \to 0} \rho(x)/x = 0$,*
- *$\phi(0)$ exists and it is finite, and*
- *$\sum_{i=1}^n \phi(D(\widetilde{x}_i, \widehat{\widetilde{g}^M[\widetilde{\mathbf{x}}_\mathbf{n}]})) > 0$,*

the sample M-estimate of location exists and it can be expressed as

$$\widehat{\widetilde{g}^M[\widetilde{\mathbf{x}}_\mathbf{n}]} = \sum_{i=1}^{n} w_i \cdot \widetilde{x}_i,$$

where $w_i \geq 0$ and $\sum_{i=1}^n w_i = 1$. Furthermore, $w_i \propto \phi(D(\widetilde{x}_i, \widehat{\widetilde{g}^M[\widetilde{\mathbf{x}}_\mathbf{n}]}))$.

In particular, if $\mathsf{X} : \Omega \to \mathcal{K}_c(\mathbb{R}^p)$ is a random compact convex set, $\mathcal{K}_c(\mathbb{R}^p)$ is associated with a metric d, and $\mathbf{x}_\mathbf{n} = (\mathsf{x}_1, \ldots, \mathsf{x}_n)$ is a sample of independent observations from X, then under the two first assumptions above and

- $\sum_{i=1}^{n} \phi(d(\mathsf{x}_i, \widehat{\mathsf{g}^M[\mathbf{x_n}]})) > 0$,

the sample M-estimate of location exists and it can be expressed as

$$\widehat{\mathsf{g}^M[\mathbf{x_n}]} = \sum_{i=1}^{n} w_i \cdot \mathsf{x}_i,$$

where $w_i \geq 0$ *and* $\sum_{i=1}^{n} w_i = 1$. *Furthermore,* $w_i \propto \phi(d(\mathsf{x}_i, \widehat{\mathsf{g}^M[\mathbf{x_n}]}))$.

And the necessary condition to minimize J proved in Theorem 3.2.1 is also sufficient by adding another assumption on J, as it can be seen by particularizing Theorem 3.1.3 and using the result in Proposition 3.1.4.

Theorem 3.2.2. *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space,* $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ *be an associated random fuzzy vector and* ρ *be an arbitrary loss function. Consider* $\mathcal{F}_c(\mathbb{R}^p)$ *with an arbitrary associated metric* D, *and let* $\widetilde{\mathbf{x}}_\mathbf{n} = (\widetilde{x}_1, \ldots, \widetilde{x}_n)$ *be a sample of independent observations from* $\mathcal{X}$. *Under the assumptions*

- ρ *is non-decreasing,* $\rho(0) = 0$ *and* $\lim_{x \to 0} \rho(x)/x = 0$,
- $\phi(0)$ *exists and it is finite, and*
- J *is strictly convex (for which sufficient conditions are given by either*
  - •• ρ *is strictly convex and non-decreasing, or*
  - •• ρ *is convex, strictly increasing,* $n \geq 3$ *and* $\mathbf{A} = (\langle s_{\widetilde{x}_i}, s_{\widetilde{x}_j} \rangle_D)_{i,j=1}^{n}$ *is positive definite),*

the following conditions

i) $\widehat{\widetilde{g}^M[\widetilde{\mathbf{x}}_\mathbf{n}]} = \sum_{i=1}^{n} w_i \cdot \widetilde{x}_i$,

ii) $w_i \propto \phi(D(\widetilde{x}_i, \widehat{\widetilde{g}^M[\widetilde{\mathbf{x}}_\mathbf{n}]}))$,

iii) $\sum_{i=1}^{n} w_i = 1$

are sufficient conditions for $\widehat{\widetilde{g}^M[\widetilde{\mathbf{x}}_\mathbf{n}]}$ *to minimize* $J(\widetilde{g}) = \frac{1}{n} \sum_{i=1}^{n} \rho(D(\widetilde{x}_i, \widetilde{g}))$.

In particular, if $\mathsf{X} : \Omega \to \mathcal{K}_c(\mathbb{R}^p)$ *is a random compact convex set,* $\mathcal{K}_c(\mathbb{R}^p)$ *is associated with a metric* d, *and* $\mathbf{x}_{\boldsymbol{n}} = (\mathsf{x}_1, \ldots, \mathsf{x}_n)$ *be a sample of independent observations from* X, *then under the assumptions above the following conditions*

i) $\widehat{\mathsf{g}^M[\mathbf{x}_{\boldsymbol{n}}]} = \sum_{i=1}^{n} w_i \cdot \mathsf{x}_i$,

ii) $w_i \propto \phi(d(\mathsf{x}_i, \widehat{\mathsf{g}^M[\mathbf{x}_{\boldsymbol{n}}]}))$,

iii) $\sum_{i=1}^{n} w_i = 1$

are sufficient conditions for $\widehat{\mathsf{g}^M[\mathbf{x}_{\boldsymbol{n}}]}$ *to minimize* $J(\mathsf{g}) = \frac{1}{n} \sum_{i=1}^{n} \rho(d(\mathsf{x}_i, \mathsf{g}))$.

Remark 3.2.1. It should be clearly remarked that the particularization above has been possible to accomplish, due to the semilinearity of $\mathcal{F}_c(\mathbb{R}^p)$ and $\mathcal{K}_c(\mathbb{R}^p)$, which guarantees that the estimates $\widehat{\widetilde{g}^M[\widetilde{\mathbf{x}}_\mathbf{n}]}$ and $\widehat{\mathbf{g}^M[\mathbf{x}_n]}$ fall in the corresponding parameter spaces, since they can be expressed as convex linear combinations of elements in the spaces $\mathcal{F}_c(\mathbb{R}^p)$ and $\mathcal{K}_c(\mathbb{R}^p)$, respectively.

In computing M-estimates of location, different choices for the loss function ρ satisfying the above-mentioned conditions have been found, most of them being parameterized by some scale factors. In dealing with set- and fuzzy set-valued data, the computation also involves the choice of the metric the norm is based on.

Recall that the required conditions in order to represent the M-estimates of location as weighted linear combinations of the sample elements (the result that allows us to guarantee that the M-estimates of location for fuzzy- and set- valued data are within the convex cone of the isometrical embedding) and to prove the convergence of the algorithm are:

- ρ continuous, non decreasing, $\rho(0) = 0$ and either
  - strictly convex or
  - convex, strictly increasing, $n \geq 3$ and $(\langle s_{\widetilde{x}_i}, s_{\widetilde{x}_j}\rangle)_{j,i=1}^n$ positive definite;
- ρ' exists, continuous and $\lim_{x\to 0} \rho(x)/x = 0$;
- ϕ exists, continuous and $\phi(0) \triangleq \lim\limits_{x\to 0} \phi(x)$ exists and is finite.

The following example illustrates such a computation in case one considers the well-known Huber loss function (Remark 3.1.1) with $a = 1.345$.

Example 3.2.1. Consider the random fuzzy number taking on values $\widetilde{x}_1 = \text{Tra}(0, 2, 3, 4)$, $\widetilde{x}_2 = \text{Tri}(1, 1.5, 2)$ and $\widetilde{x}_3 = \text{Tri}(3, 4, 5)$ (see Figure 3.2 on the left) with the induced probabilities equal to 1/3.

The corresponding M-estimate of location with the ℓ-wabl/ldev/rdev-based L^2 metric and the Huber loss function (Remark 3.1.1) can be determined by means of the algorithm detailed in Section 3.1.3 particularized to random fuzzy numbers:

Step 1. Take as initial estimate $\widetilde{g}^M_{(1)}$ $(k = 1)$ a 'central' value. In this case, as the sample size is 3, the seed has been chosen to be the trapezoidal fuzzy number which is placed in the 'central position', $\text{Tra}(0, 2, 3, 4)$. In future examples or simulations, other robust location measures will be considered as seeds. Fix a tolerance $\varepsilon = 10^{-7}$.

Step 2. Update the weights

$$w_i^{(k)} = \frac{\phi(\mathfrak{D}^\ell_{1/3}(\widetilde{x}_i, \widetilde{g}^M_{(k)}))}{\sum_{j=1}^n \phi(\mathfrak{D}^\ell_{1/3}(\widetilde{x}_j, \widetilde{g}^M_{(k)}))}$$

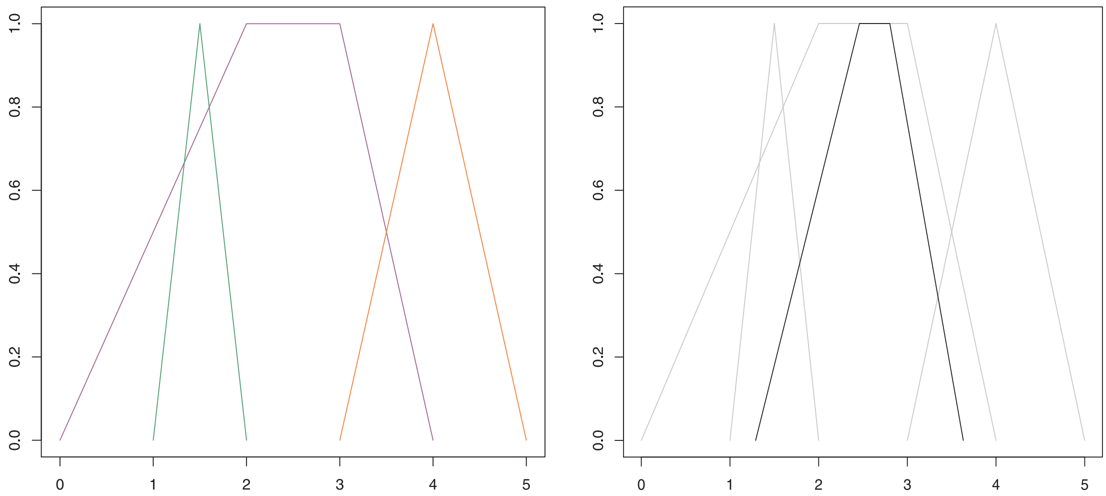


Figure 3.2: The Huber M-estimate of location of a random fuzzy number (using $\mathfrak{D}^{\ell}_{1/3}$) (which does not necessarily corresponds to a value the random fuzzy number takes on)

and the estimate

$$\widetilde{g}^{M}_{(k+1)} = \sum_{i=1}^{n} w_i^{(k)} \cdot \widetilde{x}_i.$$

Step 3. Terminate the algorithm when

$$\frac{|J(\widetilde{g}^{M}_{(k+1)}) - J(\widetilde{g}^{M}_{(k)})|}{J(\widetilde{g}^{M}_{(k)})} < \varepsilon.$$

The final estimate is given by the fuzzy number in Figure 3.2 on the right which, in spite of being unique, does not coincide with any of the three values the random fuzzy number takes on.

Analogously, by using the mid/spr-based L^2 metric and the Huber loss function, the estimate is a fuzzy number which does not coincide with any of the three values of the random fuzzy number, as Figure 3.3 shows.

The resulting values cannot be distinguished at first glance, because the two estimates are really close: $\text{Tra}(1.289103, 2.460193, 2.802372, 3.631283)$ for the metric $\mathfrak{D}^{\ell}_{1/3}$, and $\text{Tra}(1.288801, 2.459921, 2.802161, 3.631041)$ for $D^{\ell}_{1/3}$.

The sufficient conditions over the loss function allowing us to guarantee the existence of sample M-estimates of location as well as their expression as convex linear combinations of the sample elements are satisfied for different interesting choices of the loss function ρ, as it has just been verified. However, there are some other interesting choices of ρ for which such conditions fail and *ad hoc* developments should be considered.

In this respect, if one chooses either $\rho(x) = |x|$ or $\rho(x) = \sqrt{|x|}$, one cannot apply the results in Section 3.1, and those in this subsection either. In the next

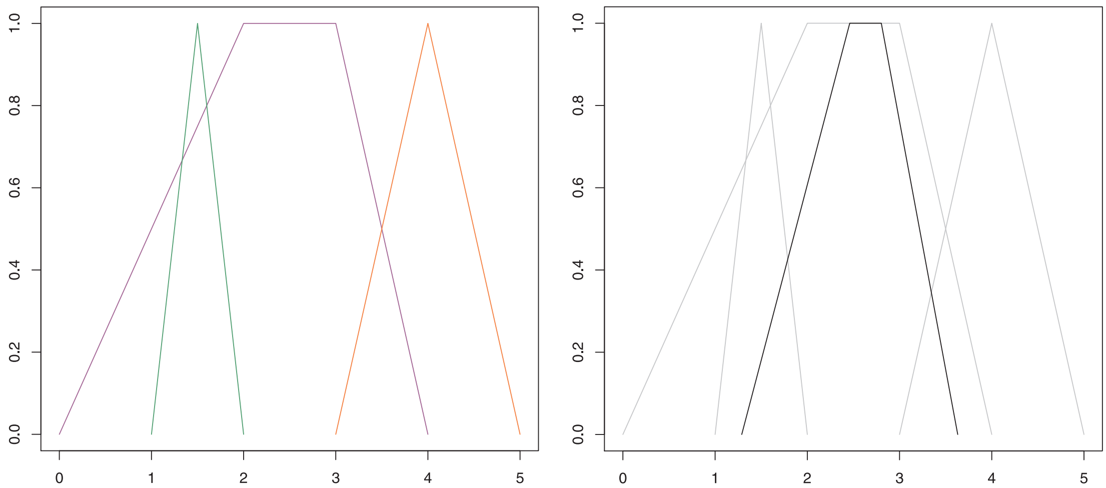


Figure 3.3: The Huber M-estimate of location of a random fuzzy number (using $D^{\ell}_{1/3}$) (which does not necessarily corresponds to a value the random fuzzy number takes on)

three subsections some *ad hoc* developments are to be carried out to get the (exact) M-estimates associated with the first choice in case of dealing with fuzzy number-valued data, and the (approximate) M-estimates with the second choice in case of dealing with interval-valued data.

Since the study of the robustness of the M-estimates only has been analyzed empirically for functional data, it will be repeated now in order to deal with fuzzy-valued data too.

Remark 3.2.2. The empirical computation of the finite sample breakdown point in Section 3.1.4 can be reproduced for interval- and fuzzy number-valued data. For instance, two sample sizes are considered: one even ($n = 100$) and another odd ($n = 101$). For each of these two samples,

- A sample of n fuzzy numbers has been generated in a random way following the CASE 1 non-contaminated distribution presented in *Step 1* of the simulation studies in Section 1.5.

- The M-estimates using both Huber and Hampel loss functions have been computed.

- Afterwards, $i \in \{1, \ldots, n\}$ observations have been highly contaminated, that is, they have been replaced by other observations obtained from the following distributions:

$$X_1 \sim \mathcal{N}(0, 3) + 10^7,$$

$$X_2, X_3, X_4 \sim \chi^2_4 + 100.$$

- Huber and Hampel M-estimates are computed for the contaminated sample.
- D_θ^ℓ distances, with $\theta = 1/3$ and $\theta = 1$, between the (Huber and Hampel) M-estimates for the original and the contaminated samples have been computed for each amount of perturbed observations and plotted in Figures 3.4 and 3.5.
- The value of the finite sample breakdown point is given by the minimum number of perturbed observations for which distance between the non-contaminated and the contaminated corresponding M-estimates is arbitrarily big. This value has been emphasized in red.

In Figure 3.4, the $D_{\theta=1/3}^\ell$ distances between the non-contaminated and the contaminated Huber (blue color) and Hampel (green color) M-estimates have been plotted for the even sample size (left graphical display) and for the odd one (right graphical display). It can be seen that, independently from the considered loss function, the empirical value of the finite sample breakdown point is 50 when $n = 100$ and 51 when $n = 101$, that is, $\frac{1}{n} \cdot \lfloor \frac{n+1}{2} \rfloor$.

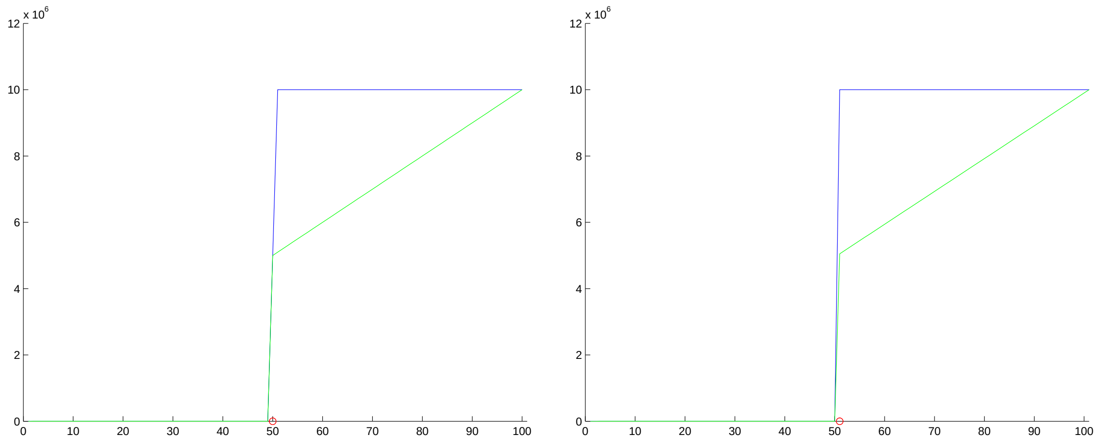


Figure 3.4: Empirical value (in red) obtained for the finite sample breakdown point of the M-estimators for fuzzy number-valued data and distances between the non-contaminated and contaminated estimates using the Huber (blue) and Hampel (green) loss functions and the $D_{\theta=1/3}^\ell$-metric when the sample size is even (left) and odd (right).

In Figure 3.5, the metric used to compute the distances between the non-contaminated and the contaminated Huber and Hampel M-estimates is the $D_{\theta=1}^\ell$ metric. Again, it can be seen that, independently from the considered loss function, the empirical value of the finite sample breakdown point is 50 when $n = 100$ (left graphical display) and 51 when $n = 101$ (right graphical display), that is, $\frac{1}{n} \cdot \lfloor \frac{n+1}{2} \rfloor$.

From Figures 3.4 and 3.5 we can also conclude that the value of the finite sample breakdown point has not depended on the considered metric.

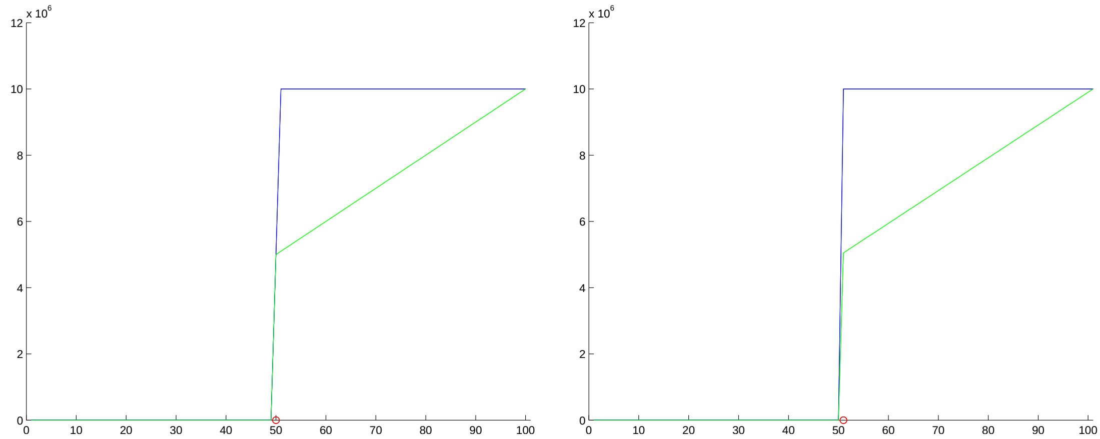


Figure 3.5: Empirical value (in red) obtained for the finite sample breakdown point of the M-estimators for fuzzy number-valued data and distances between the non-contaminated and contaminated estimates using the Huber (blue) and Hampel (green) loss functions and the $D^{\ell}_{\theta=1}$-metric when the sample size is even (left) and odd (right).

An analogue study has been performed for interval-valued data, the same conclusions arising from the outputs. However, as the details of the simulation studies dealing with intervals have not been specified yet, it will not be presented here.

3.3 The 1-norm median for a random fuzzy number

The population and sample 1-norm median for random fuzzy numbers are defined as follows:

Definition 3.3.1. *Given a probability space $(\Omega, \mathcal{A}, P)$ and an associated random fuzzy number $\mathcal{X}$, the* ***population 1-norm median(s)*** *of $\mathcal{X}$ is the fuzzy number(s)*

$$\widetilde{\mathrm{Me}}(\mathcal{X}) = \arg \min_{\widetilde{U} \in \mathcal{F}_c(\mathbb{R})} E\left(\boldsymbol{\rho_1}(\mathcal{X}, \widetilde{U})\right),$$

whenever these expectations exist.

Definition 3.3.2. *Given a probability space $(\Omega, \mathcal{A}, P)$, an associated random fuzzy number $\mathcal{X}$, and a simple random sample $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ from $\mathcal{X}$, the* ***sample 1-norm median(s)*** *of $\mathcal{X}$ is(are) the fuzzy number-valued statistic(s)*

$$\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n = \arg \min_{\widetilde{U} \in \mathcal{F}_c(\mathbb{R})} \frac{1}{n} \sum_{i=1}^{n} \left(\boldsymbol{\rho_1}(\mathcal{X}_i, \widetilde{U})\right).$$

Two key questions at this stage are whether the 1-norm median exists and whether it can be computed easily in practice. A result is first to be established for the population measure. The result guarantees that at least one such median always exists and it is rather easy to compute.

Theorem 3.3.1. *Given a probability space* $(\Omega, \mathcal{A}, P)$ *and an associated random fuzzy number* $\mathcal{X}$*, for any* $\alpha \in [0,1]$*, the fuzzy number* $\widetilde{\mathrm{Me}}(\mathcal{X}) \in \mathcal{F}_c(\mathbb{R})$ *such that*

$$\left(\widetilde{\mathrm{Me}}(\mathcal{X})\right)_\alpha = \left[\mathrm{Me}\big(\inf \mathcal{X}_\alpha\big), \mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)\right],$$

(where in case $\mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$ *or* $\mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)$ *are non-unique the most usual convention, in accordance with which* $\mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$ *and* $\mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)$ *is chosen to be the midpoint of the interval of medians of* $\inf \mathcal{X}_\alpha$ *and* $\sup \mathcal{X}_\alpha$*, respectively, is considered) is a population 1-norm median of* $\mathcal{X}$*.*

Proof. Indeed, on one hand whatever $\alpha \in [0,1]$ may be, $\inf \widetilde{U}_\alpha, \sup \widetilde{U}_\alpha \in \mathbb{R}$, and $\inf \mathcal{X}_\alpha$ and $\sup \mathcal{X}_\alpha$ are real-valued random variables, so for all $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ one has that

$$E\left[\left|\inf \mathcal{X}_\alpha - \mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)\right|\right] \leq E\left[\left|\inf \mathcal{X}_\alpha - \inf \widetilde{U}_\alpha\right|\right],$$
$$E\left[\left|\sup \mathcal{X}_\alpha - \mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)\right|\right] \leq E\left[\left|\sup \mathcal{X}_\alpha - \sup \widetilde{U}_\alpha\right|\right],$$

whence

$$E\left(\boldsymbol{\rho_1}(\mathcal{X}, \widetilde{U})\right) = \frac{1}{2}\int_{[0,1]} E\left[\left|\inf \mathcal{X}_\alpha - \inf \widetilde{U}_\alpha\right|\right] d\ell(\alpha)$$
$$+\frac{1}{2}\int_{[0,1]} E\left[\left|\sup \mathcal{X}_\alpha - \sup \widetilde{U}_\alpha\right|\right] d\ell(\alpha) \geq \frac{1}{2}\int_{[0,1]} E\left[\left|\inf \mathcal{X}_\alpha - \mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)\right|\right] d\ell(\alpha)$$
$$+\frac{1}{2}\int_{[0,1]} E\left[\left|\sup \mathcal{X}_\alpha - \mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)\right|\right] d\ell(\alpha) = E\left(\boldsymbol{\rho_1}(\mathcal{X}, \widetilde{\mathrm{Me}}(\mathcal{X}))\right).$$

On the other hand, intervals $\left[\mathrm{Me}\big(\inf \mathcal{X}_\alpha\big), \mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)\right]$ correspond to the α-levels of a fuzzy number, what can be proved by checking the sufficient conditions in Proposition 1.1.3. Thus, for any $\alpha \in [0,1]$ they are well-defined intervals, because of the considered convention $\inf \mathcal{X}_\alpha \leq \sup \mathcal{X}_\alpha$ entails that $\mathrm{Me}\big(\inf \mathcal{X}_\alpha\big) \leq \mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)$. Moreover, $\mathrm{Me}\big(\inf \mathcal{X}_1\big) \leq \mathrm{Me}\big(\sup \mathcal{X}_1\big)$ ensures that the 1-level is nonempty.

Since $\inf \mathcal{X}_\alpha$ and $\sup \mathcal{X}_\alpha$ are non-decreasing and non-increasing functions of α, respectively, then $\mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$ and $\mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)$ are also non-decreasing and non-increasing, respectively.

One should also verify that $\mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$ and $\mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)$ are left-continuous at every $\alpha \in (0,1]$. If $\{\alpha_n\}_n \uparrow \alpha \in (0,1]$ as $n \to \infty$, then for all element in Ω we have that $\{\inf \mathcal{X}_{\alpha_n}\}_n \uparrow \inf \mathcal{X}_\alpha$ and because of the considered convention the sequence $\left\{\mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big)\right\}_n \uparrow$ is bounded above, $\mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$ being an upper bound. Hence, a limit for this sequence exists and will be denoted by $L_\alpha = \lim_{n\to\infty} \mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big) \leq \mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$.

Furthermore, $L_\alpha = \mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$, since for all $\omega \in \Omega$ we have that

$$0.5 \leq P\left(\inf \mathcal{X}_{\alpha_n} \leq \mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big)\right) \leq P\left(\inf \mathcal{X}_{\alpha_n} \leq L_\alpha\right)$$

and

$$\{(\inf \mathcal{X}_{\alpha_n} \le L_\alpha)\}_n \downarrow \bigcap_{n=1}^{\infty} (\inf \mathcal{X}_{\alpha_n} \le L_\alpha) = (\inf \mathcal{X}_\alpha \le L_\alpha),$$

whence

$$P(\inf \mathcal{X}_\alpha \le L_\alpha) = P\left(\lim_{n\to\infty}\left(\inf \mathcal{X}_{\alpha_n} \le L_\alpha\right)\right) = \lim_{n\to\infty} P\Big(\inf \mathcal{X}_{\alpha_n} \le L_\alpha\Big) \ge 0.5.$$

Following similar arguments,

$$P(\inf \mathcal{X}_\alpha < L_\alpha) = P\left(\bigcup_{n=1}^{\infty}\Big(\inf \mathcal{X}_\alpha < \mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big)\Big)\right)$$
$$= P\left(\lim_{n\to\infty}\Big(\inf \mathcal{X}_\alpha < \mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big)\Big)\right) = \lim_{n\to\infty} P\Big(\inf \mathcal{X}_\alpha < \mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big)\Big)$$
$$\le \lim_{n\to\infty} P\Big(\inf \mathcal{X}_{\alpha_n} < \mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big)\Big) \le 0.5.$$

Consequently, taking into account the considered convention, we have that $L_\alpha \ge \mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$ and, therefore, $L_\alpha = \mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$.

Analogously, if $\{\alpha_n\}_n \uparrow \alpha \in (0,1]$ as $n \to \infty$, it holds that $\{\sup \mathcal{X}_{\alpha_n}\}_n \downarrow \sup \mathcal{X}_\alpha$ and the sequence $\big\{\mathrm{Me}\big(\sup \mathcal{X}_{\alpha_n}\big)\big\}_n \downarrow$ and it is bounded below by $\mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)$ so that there exists $L'_\alpha = \lim_{n\to\infty} \mathrm{Me}\big(\sup \mathcal{X}_{\alpha_n}\big)$ and we can easily prove that $L'_\alpha = \mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)$.

Finally, the right-continuity at 0 of $\mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$ and $\mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)$ should be proved. If $\{\alpha_n\}_n \downarrow 0$ as $n \to \infty$, then for all element in Ω we have that $\{\inf \mathcal{X}_{\alpha_n}\}_n \downarrow \inf \mathcal{X}_0$ and because of the considered convention the sequence $\big\{\mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big)\big\}_n \downarrow$ is bounded below, $\mathrm{Me}\big(\inf \mathcal{X}_0\big)$ being a lower bound. Hence, a limit for this sequence exists and will be denoted by $L_0 = \lim_{n\to\infty} \mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big) \ge \mathrm{Me}\big(\inf \mathcal{X}_0\big)$. Furthermore, $L_0 = \mathrm{Me}\big(\inf \mathcal{X}_0\big)$, since for all $\omega \in \Omega$ we have that

$$\{(\inf \mathcal{X}_{\alpha_n} < L_0)\}_n \uparrow \bigcup_{n=1}^{\infty} (\inf \mathcal{X}_{\alpha_n} < L_0) = (\inf \mathcal{X}_0 < L_0),$$

whence

$$P(\inf \mathcal{X}_0 < L_0) = P\left(\lim_{n\to\infty}\left(\inf \mathcal{X}_{\alpha_n} < L_0\right)\right) = \lim_{n\to\infty} P\Big(\inf \mathcal{X}_{\alpha_n} < L_0\Big)$$
$$\le \lim_{n\to\infty} P\left(\inf \mathcal{X}_{\alpha_n} < \mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big)\right) \le 0.5.$$

Following similar arguments,

$$P(\inf \mathcal{X}_0 \le L_0) = P\left(\bigcap_{n=1}^{\infty}\Big(\inf \mathcal{X}_0 \le \mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big)\Big)\right)$$
$$= P\left(\lim_{n\to\infty}\Big(\inf \mathcal{X}_0 \le \mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big)\Big)\right) = \lim_{n\to\infty} P\Big(\inf \mathcal{X}_0 \le \mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big)\Big)$$
$$\ge \lim_{n\to\infty} P\Big(\inf \mathcal{X}_{\alpha_n} \le \mathrm{Me}\big(\inf \mathcal{X}_{\alpha_n}\big)\Big) \ge 0.5.$$

In an analogous way one can verify the right-continuity at 0 of $\mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)$. □

Similarly, for the sample approach

Theorem 3.3.2. *Given a probability space $(\Omega, \mathcal{A}, P)$, an associated random fuzzy number $\mathcal{X}$, and a simple random sample $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ from $\mathcal{X}$, the fuzzy number-valued statistic*

$$\left(\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n\right)_\alpha = \left[\widehat{\mathrm{Me}(\inf \mathcal{X}_\alpha)}_n, \widehat{\mathrm{Me}(\sup \mathcal{X}_\alpha)}_n\right],$$

where $\widehat{\mathrm{Me}(\inf \mathcal{X}_\alpha)}_n$ *and* $\widehat{\mathrm{Me}(\sup \mathcal{X}_\alpha)}_n$ *denote the sample medians of the corresponding real-valued random variables and with a convention similar to that in Theorem 3.3.1, is a sample 1-norm median.*

Remark 3.3.1. It should be pointed out that with the convention in Theorems 3.3.1 and 3.3.2, it is easy to compute exactly a fuzzy numbered solution of Definitions 3.3.1 and 3.3.2, respectively. However, if we do not consider some valid conventions, then the result can fail. That is, in case $\mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$ or $\mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)$ are non-unique, there are choices for them which do not determine a fuzzy number.

As a counterexample, consider the random fuzzy number $\mathcal{X}$ taking on the triangular values $\widetilde{x}_1 = \mathrm{Tri}(0,1,2)$ and $\widetilde{x}_2 = \mathrm{Tri}(1,2,3)$ both with induced probabilities $P(\mathcal{X} = \widetilde{x}_1) = P(\mathcal{X} = \widetilde{x}_2) = 0.5$; then, for $\alpha = 0.75$ we have that $\mathrm{Me}\big(\inf \mathcal{X}_{0.75}\big)$ is any value in $[0.75, 1.75]$, whereas $\mathrm{Me}\big(\sup \mathcal{X}_{0.75}\big)$ is any value in $[1.25, 2.25]$, so that some choices for the medians of $\inf \mathcal{X}_{0.75}$ and $\sup \mathcal{X}_{0.75}$ would lead to empty α-levels.

To avoid an unnecessary cumbersome checking and to ease the study of the properties of the median, from now on the population and sample 1-norm medians will be assumed to be defined as the unique fuzzy number in Theorems 3.3.1 and 3.3.2, respectively.

Remark 3.3.2. In contrast to the median for random variables, the 1-norm median of a random fuzzy number does not necessarily correspond to one of the values of the random fuzzy number. As an example corroborating this assertion and illustrating the computation of the median we can consider, for instance, the random fuzzy number introduced in Example 3.2.1.

The corresponding 1-norm median can be trivially determined and it is given by the fuzzy number in Figure 3.6 on the right, which does not coincide with any of the three values of the random fuzzy number, and this is not affected by the convention which has not been required in this example. In cases the 1-norm median cannot be that easily obtained, it can be approximated by using a large number of levels, following ideas similar to those by Trutschnig and Lubiano [202].

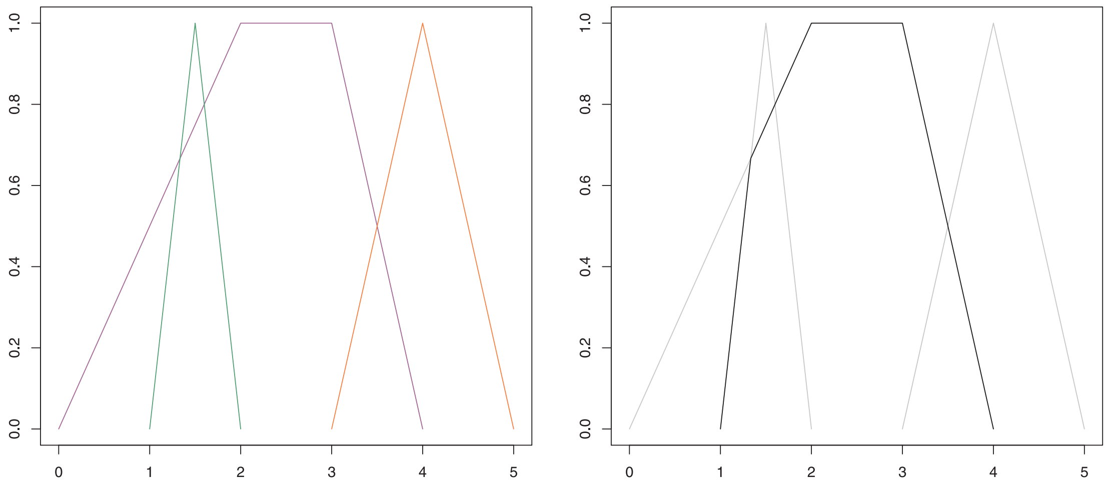


Figure 3.6: Counterexample: the 1-norm median of a random fuzzy number, even if uniquely valued, does not necessarily corresponds to a value the random fuzzy number takes on

3.3.1 Basic properties of the 1-norm median of a random fuzzy number

The 1-norm median of a random fuzzy number preserves most of the basic properties of the median of a random variable, irrespective of the considered version being the population or the sample one. Thus, on the basis of the results in Theorems 3.3.1 and 3.3.2, it can be straightforwardly proved that

Proposition 3.3.3. $\widetilde{\text{Me}}$ *is equivariant under 'linear' transformations, that is, if* $\gamma \in \mathbb{R}$*,* $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ *and* $\mathcal{X}$ *is a random fuzzy number, then*

$$\widetilde{\text{Me}}(\gamma \cdot \mathcal{X} + \widetilde{U}) = \gamma \cdot \widetilde{\text{Me}}(\mathcal{X}) + \widetilde{U}.$$

Consequently, if $\mathcal{X}$ *is a random fuzzy number associated with the probability space* $(\Omega, \mathcal{A}, P)$ *and the distribution of* $\mathcal{X}$ *is degenerate at a fuzzy number* $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ *(i.e.,* $\mathcal{X} = \widetilde{U}$ *a.s.* $[P]$*), then* $\widetilde{\text{Me}}(\mathcal{X}) = \widetilde{U}$*.*

The median of a real-valued random variable is usually defined in two equivalent ways, namely: either as a value minimizing the mean distance to the distribution of the variable through an L^1-type metric or as a 'middle position' value with respect to a specified ranking. Although fuzzy numbers cannot be ranked through a universally acceptable total ordering, it can be verified that the 1-norm median of a random fuzzy number can be also formalized as a 'middle position' value with respect to the *fuzzy max partial order*, whenever this order applies. The fuzzy max order on $\mathcal{F}_c(\mathbb{R})$ was introduced by Dubois and Prade [60], and equivalent definitions were stated by Ramík and Římánek [160] and more recently by Valvis [207]. It is the natural level-wise extension through the inf/sup representation of the product order on $\mathbb{R}^2$, so that $\widetilde{U} \precsim \widetilde{V}$ if and only if for all $\alpha, \lambda \in [0,1]$ one has that

$$\lambda \sup \widetilde{U}_\alpha + (1-\lambda) \inf \widetilde{U}_\alpha \leq \lambda \sup \widetilde{V}_\alpha + (1-\lambda) \inf \widetilde{V}_\alpha.$$

The main practical drawback for this ranking lies in the fact that it only leads to a partial ordering and many fuzzy numbers cannot be compared with it. However, it is often viewed as a widely accepted ranking criterion and as a pattern which should be preserved for any other suggested partial or total ranking.

Proposition 3.3.4. *For any sample or finite population $(\omega_1, \ldots, \omega_n)$ for which the values of a random fuzzy number $\mathcal{X}$ satisfy that*

$$\mathcal{X}(\omega_1) \precsim \ldots \precsim \mathcal{X}(\omega_n)$$

we have that

- *if n is odd, then*
$$\widetilde{\mathrm{Me}}(\mathcal{X}) = \mathcal{X}(\omega_{(n+1)/2}),$$
- *if n is even, then*
$$\widetilde{\mathrm{Me}}(\mathcal{X}) = \frac{1}{2} \cdot \left(\mathcal{X}(\omega_{n/2} + \mathcal{X}(\omega_{(n/2)+1}\right).$$

Remark 3.3.3. It should be pointed out that alternate approaches to the median extension could be stated in terms of total orderings on the space of fuzzy numbers, since one can state a notion of empirical or population distribution function. Nevertheless, in addition of none of these rankings being universally accepted, the formalization of the properties in this subsection and especially those in the next one could be unfeasible.

Another interesting property in examining the adequacy of the 1-norm median for random fuzzy numbers as a central tendency measure is now discussed by considering their behaviour in case of symmetrically distributed random fuzzy numbers.

In the real-valued case a well-known result is that the median of a symmetric random variable coincides with the point the variable is symmetric about whenever it is unique. As for the Aumann-type mean, in case of considering random fuzzy numbers and the 1-norm median this assertion should be slightly modified, due to the involved fuzziness. Thus, the 1-norm median shows a suitable central tendency behaviour since it leads to a fuzzy number which is symmetric about the symmetry point. Moreover, this measure neither necessarily coincides nor corresponds to any of the values the random fuzzy number takes on. Thus,

Proposition 3.3.5. *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space, and let* $\mathcal{X}$ *be a symmetric random fuzzy number about* $c \in \mathbb{R}$. *Then, the 1-norm median of* $\mathcal{X}$ *is a symmetric fuzzy number about* c.

Proof. Since $\mathcal{X} \overset{d}{=} 2c - \mathcal{X}$, then $\widetilde{\mathrm{Me}}(\mathcal{X}) = \widetilde{\mathrm{Me}}(2c - \mathcal{X})$, whence because of the equivariance properties of $\widetilde{\mathrm{Me}}$ under affine transformations, we have that

$$\widetilde{\mathrm{Me}}(\mathcal{X}) = 2c - \widetilde{\mathrm{Me}}(\mathcal{X}).$$

By adding $\widetilde{\mathrm{Me}}(\mathcal{X})$ to the two members in the last equality, $2\,\widetilde{\mathrm{Me}}(\mathcal{X}) = 2c + \widetilde{\mathrm{Me}}(\mathcal{X}) - \widetilde{\mathrm{Me}}(\mathcal{X})$ and, hence,

$$\widetilde{\mathrm{Me}}(\mathcal{X}) = c + \frac{1}{2} \cdot \mathcal{O}_{\widetilde{\mathrm{Me}}(\mathcal{X})},$$

whence for each $\alpha \in [0, 1]$

$$\left(\widetilde{\mathrm{Me}}(\mathcal{X})\right)_\alpha = \left[c - \mathrm{spr}\left(\widetilde{\mathrm{Me}}(\mathcal{X})\right)_\alpha, c + \mathrm{spr}\left(\widetilde{\mathrm{Me}}(\mathcal{X})\right)_\alpha\right],$$

which leads to a symmetric fuzzy number about c. □

The result in Proposition 3.3.5 is now illustrated by computing the two 1-norm medians of the symmetric random fuzzy numbers in Examples 1.4.3 and 1.4.4.

Example 3.3.1. To compute the 1-norm median of the symmetric random fuzzy number about 0.5 in Example 1.4.3 (p. 47) we should take into account that

label	VD	D	SD	N	SA	A	VA
absol. freq.	38	143	207	177	207	143	38
$\inf_\alpha$	0	$\frac{\alpha}{6}$	$\frac{\alpha+1}{6}$	$\frac{\alpha+2}{6}$	$\frac{\alpha+3}{6}$	$\frac{\alpha+4}{6}$	$\frac{\alpha+5}{6}$
$\sup_\alpha$	$\frac{1-\alpha}{6}$	$\frac{2-\alpha}{6}$	$\frac{3-\alpha}{6}$	$\frac{4-\alpha}{6}$	$\frac{5-\alpha}{6}$	$\frac{6-\alpha}{6}$	1

whence, by developing a comparison of the values in each row as a function of α, one can easily conclude that

$$\widetilde{\mathrm{Me}}(\mathcal{X}) = \mathrm{N}.$$

To compute the 1-norm median of the symmetric random fuzzy number about 2 in Example 1.4.4 (p. 49) we should take into account that

label	$\gamma_{(2)}(0)$	$\gamma_{(2)}(1)$	$\gamma_{(2)}(2)$	$\gamma_{(2)}(3)$	$\gamma_{(2)}(4)$
probab.	.0625	.25	.375	.25	.0625
$\inf_\alpha$	$-\sqrt[3]{1-\alpha}$	$1-\sqrt{1-\alpha}$	$1+\alpha$	$3-(1-\alpha)^2$	$4-(1-\alpha)^3$
$\sup_\alpha$	$(1-\alpha)^3$	$1+(1-\alpha)^2$	$3-\alpha$	$3+\sqrt{1-\alpha}$	$4+\sqrt[3]{1-\alpha}$

whence, by developing a comparison of the values in each row as a function of α for the 1-norm median, one can easily conclude that $\widetilde{\mathrm{Me}}(\mathcal{X}) = \gamma_{(2)}(2)$.

Consequently, one can assert that for symmetric random fuzzy numbers about c, both the Aumann-type mean and the 1-norm median are symmetric fuzzy numbers about c, but they do not necessarily coincide even if it is unique.

Examples 1.4.3, 1.4.4 and 3.3.1 show that the Aumann-type mean is not necessarily a value of the symmetric random fuzzy number, even in case the number of different values is an odd one, and the same happens (although not that frequently) with the 1-norm median. Anyway, the examined examples make us think that the behaviour of the 1-norm median seems to be closer to that of the real-valued case than the behaviour of the mean value. An empirical discussion on this point is now carried out.

Thus, in measuring the central tendency for symmetric random fuzzy numbers the 1-norm median behaves in a more suitable and advisable way than the Aumann-type mean. In addition to provide us with more robust estimates than the mean (as it will be seen in the next subsection), the 1-norm median also leads to a fuzzy value which is closer to the one which occupies the 'central position'.

To illustrate this assertion we have considered three different random fuzzy numbers that are symmetric about 0, assumption made for the sake of simplicity and unification although not being relevant.

These symmetric random fuzzy numbers have been obtained by composing the aforementioned characterizing fuzzy representation (González-Rodríguez *et al.* [89]) given by

$$\left(\gamma_{(0)}(x)\right)_\alpha = \begin{cases} \left[x-(1-\alpha)^{1+x}, x+(1-\alpha)^{1/(1+x)}\right] & \text{if } x \geq 0 \\ \\ \left[x-(1-\alpha)^{1/(1-x)}, x+(1-\alpha)^{1-x}\right] & \text{if } x < 0 \end{cases}$$

with three symmetric real-valued random variables: a standard normal $X \sim \mathcal{N}(0,1)$; a uniform $X \sim \mathrm{Uniform}(-0.5, 0.5)$; and a translated binomial $X \sim \mathrm{Bin}(5, 0.5) - 2.5$.

The (population) 1-norm median and Aumann-type mean of each of the random fuzzy numbers $\gamma_{(0)} \circ X$ are now graphically displayed in Figures 3.7-3.9.

Figure 3.7 shows that when the considered random fuzzy number is $\gamma_{(0)} \circ \mathcal{N}(0,1)$, then the 1-norm median coincides with the central position value, whereas the Aumann-type mean is not that close.

Analogously, Figure 3.8 shows that when the considered random fuzzy number is $\gamma_{(0)} \circ \mathrm{Uniform}(-0.5, 0.5)$, then the 1-norm median coincides with the central position value, whereas the Aumann-type mean is not that close.

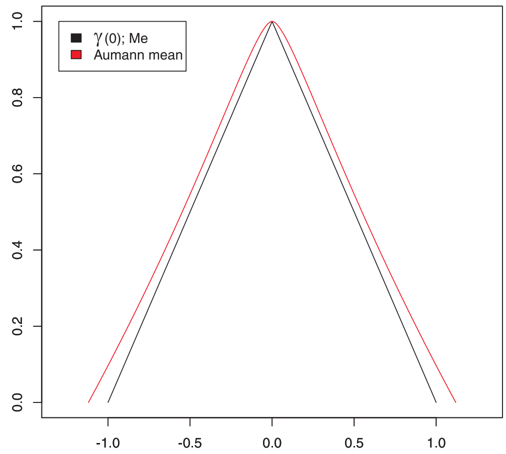


Figure 3.7: Aumann-type mean and 1-norm median $\mathrm{Me} = \gamma_{(0)}(0)$ of the random fuzzy number $\gamma_{(0)} \circ \mathcal{N}(0,1)$

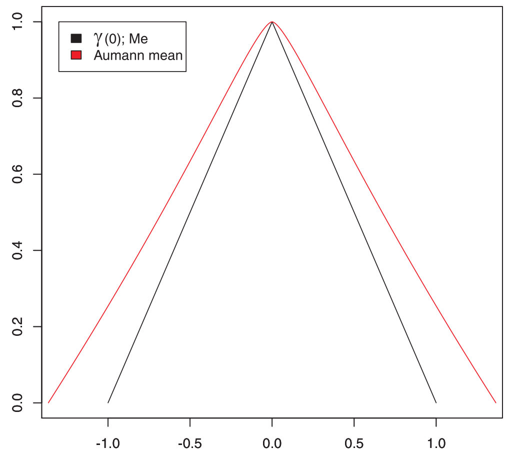


Figure 3.8: Aumann-type mean and 1-norm median $\mathrm{Me} = \gamma_{(0)}(0)$ of the random fuzzy number $\gamma_{(0)} \circ \mathrm{Uniform}(-0.5, 0.5)$

Remark 3.3.4. It should be emphasized that the coincidence $\widetilde{\mathrm{Me}}\big(\gamma_{(0)} \circ X\big) = \gamma_{(0)}(0)$ is not at all casual. Whenever X has a either a symmetric continuous distribution or a discrete one with an odd number of distinct values, then the equality holds. This is due to the fact that both $\inf(\gamma_{(0)} \circ X)_\alpha$ and $\sup(\gamma_{(0)} \circ X)_\alpha$ are strictly increasing functions of X, whence for each $\alpha \in [0,1]$ we have that $\mathrm{Me}\big(\inf(\gamma_{(0)} \circ X)_\alpha\big) = \inf\big(\gamma_{(0)}(0)\big)_\alpha$ and $\mathrm{Me}\big(\sup(\gamma_{(0)} \circ X)_\alpha\big) = \sup\big(\gamma_{(0)}(0)\big)_\alpha$.

The above-mentioned coincidence does not hold in general when the number of distinct values of X is even. In such a case, we cannot properly talk about 'central position' and conventions should be made, so the use of $\gamma_{(0)}(0)$ as the central position value is not completely fair. Anyway, it serves us to illustrate that the behaviour of the 1-norm median in contrast to that of the mean is preserved.

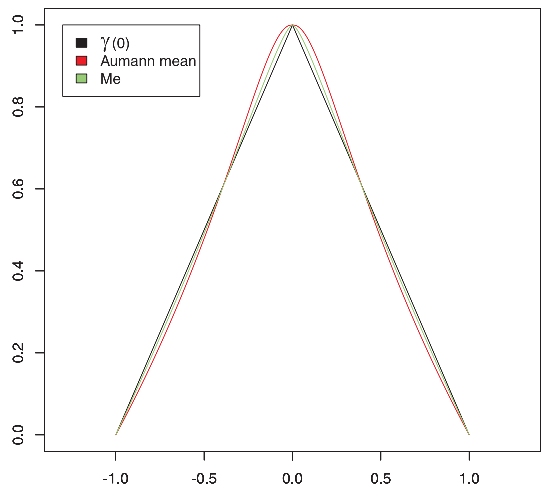


Figure 3.9: Aumann-type mean and 1-norm median of the random fuzzy number (Me) of the random fuzzy number $\gamma_{(0)} \circ [\mathrm{Bin}(5, 0.5) - 2.5]$, and comparison with $\gamma_{(0)}(0)$

In this way, Figure 3.9 shows the scenario when the considered random fuzzy number is $\gamma_{(0)} \circ [\mathrm{Bin}(5, 0.5) - 2.5]$, a random fuzzy number symmetric about 0 and taking on 6 different values. In this case, the 1-norm median is very close to $\gamma_{(0)}(0)$ (which does not exactly correspond to a central position) but the Aumann-type mean is not very close to it.

3.3.2 Consistency and robustness of the sample 1-norm median and comparisons with the sample mean

The inferential behaviour of the 1-norm median of a random fuzzy number is now to be analyzed. In this respect, the $\boldsymbol{\rho_1}$-strong consistency and the finite sample breakdown point to examine its robustness are to be discussed.

As for the real-valued case, under rather mild conditions the sample median is shown to be a strongly consistent estimator of the population median, that is,

Theorem 3.3.6. *Let $\mathcal{X}$ be a random fuzzy number associated with a probability space $(\Omega, \mathcal{A}, P)$ satisfying that for each $\alpha \in [0, 1]$ the real-valued population medians $\mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$ and $\mathrm{Me}\big(\sup \mathcal{X}_\alpha\big)$ exist and they are actually unique (i.e., they are unique without applying the convention in Theorem 3.3.1).*

If $\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n$ denotes the sample median corresponding to a simple random sample $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ from $\mathcal{X}$, and the two sequences of the real-valued sample medians $\big\{\widehat{\mathrm{Me}(\inf \mathcal{X}_\alpha)}_n\big\}_n$ and $\big\{\widehat{\mathrm{Me}(\sup \mathcal{X}_\alpha)}_n\big\}_n$ as functions of α over $[0, 1]$ are both uniformly integrable, then $\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n$ is a strongly consistent estimator of $\widetilde{\mathrm{Me}}(\mathcal{X})$ in $\boldsymbol{\rho_1}$-sense *(and hence in the sense of all the topologically equivalent metrics), i.e.*

$$\lim_{n\to\infty} \boldsymbol{\rho_1}\left(\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n, \widetilde{\mathrm{Me}}(\mathcal{X})\right) = 0 \quad a.s.\,[P].$$

Proof. Indeed,

$$P\left(\lim_{n\to\infty} \boldsymbol{\rho_1}\left(\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n, \widetilde{\mathrm{Me}}(\mathcal{X})\right) = 0\right)$$

$$= P\left(\lim_{n\to\infty}\left(\frac{1}{2}\int_{[0,1]} |\widehat{\mathrm{Me}(\inf \mathcal{X}_\alpha)}_n - \mathrm{Me}(\inf \mathcal{X}_\alpha)|\, d\ell(\alpha)\right.\right.$$

$$\left.\left.+\frac{1}{2}\int_{[0,1]} |\widehat{\mathrm{Me}(\sup \mathcal{X}_\alpha)}_n - \mathrm{Me}(\sup \mathcal{X}_\alpha)|\, d\ell(\alpha)\right) = 0\right)$$

$$= P\left(\left(\lim_{n\to\infty}\int_{[0,1]} |\widehat{\mathrm{Me}(\inf \mathcal{X}_\alpha)}_n - \mathrm{Me}(\inf \mathcal{X}_\alpha)|\, d\ell(\alpha) = 0\right)\right.$$

$$\left.\bigcap\left(\lim_{n\to\infty}\int_{[0,1]} |\widehat{\mathrm{Me}(\sup \mathcal{X}_\alpha)}_n - \mathrm{Me}(\sup \mathcal{X}_\alpha)|\, d\ell(\alpha) = 0\right)\right)$$

Under the assumption of uniqueness for the median of $\inf \mathcal{X}_\alpha$, the sample median is a strongly consistent estimator of the population median, and hence

$$P\left(\lim_{n\to\infty}\left(\widehat{\mathrm{Me}(\inf \mathcal{X}_\alpha)}_n - \mathrm{Me}(\inf \mathcal{X}_\alpha)\right) = 0\right) = 1.$$

On the other hand, assumptions for $\widehat{\mathrm{Me}(\inf \mathcal{X}_\alpha)}_n$ and $\mathrm{Me}\big(\inf \mathcal{X}_\alpha\big)$ guarantee that conditions to apply Vitali's Convergence Theorem are fulfilled, whence

$$P\left(\left(\lim_{n\to\infty}\int_{[0,1]} |\widehat{\mathrm{Me}(\inf \mathcal{X}_\alpha)}_n - \mathrm{Me}(\inf \mathcal{X}_\alpha)|\, d\ell(\alpha) = 0\right)\right) = 1.$$

By following similar arguments, one can prove that

$$P\left(\left(\lim_{n\to\infty}\int_{[0,1]} |\widehat{\mathrm{Me}(\sup \mathcal{X}_\alpha)}_n - \mathrm{Me}(\sup \mathcal{X}_\alpha)|\, d\ell(\alpha) = 0\right)\right) = 1.$$

Consequently,

$$P\left(\lim_{n\to\infty} \boldsymbol{\rho_1}\left(\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n, \widetilde{\mathrm{Me}}(\mathcal{X})\right) = 0\right) = 1. \qquad \square$$

Remark 3.3.5. Assumptions in Theorem 3.3.6 are not very restrictive in practice, since many real-life conditions accomplish them. In this respect, a very common situation involves random fuzzy sets which are bounded (more concretely, their 0-level mappings are bounded random variables), so the conclusions from Theorem 3.3.6 would directly apply.

		CASE 1		CASE 2	
c_P	c_D	$\boldsymbol{\rho_1}$	$\boldsymbol{\rho_2}$	$\boldsymbol{\rho_1}$	$\boldsymbol{\rho_2}$
0	0	1.386439	1.552950	0.827758	1.032440
0.1	0	1.392037	1.564486	0.827719	1.031691
0.1	1	1.393240	1.569681	0.828281	1.037552
0.1	5	1.394710	1.566279	0.829544	1.043102
0.1	10	1.395252	1.568843	0.829840	1.044393
0.1	100	1.395479	1.569227	0.829921	1.045329
0.2	0	1.413077	1.593075	0.827801	1.030994
0.2	1	1.421773	1.602914	0.829706	1.043201
0.2	5	1.431524	1.615051	0.836509	1.056944
0.2	10	1.433923	1.617811	0.837271	1.058855
0.2	100	1.434000	1.617947	0.837488	1.061132
0.4	0	1.558658	1.759157	0.827868	1.028484
0.4	1	1.659079	1.865478	0.840436	1.061503
0.4	5	1.795391	2.014625	0.897220	1.124580
0.4	10	1.862582	2.092274	0.907626	1.138088
0.4	100	1.873241	2.101532	0.909834	1.142909

Table 3.3: Mean distances of the mixed (partially contaminated and non-contaminated) sample 1-norm median to the non-contaminated distribution of a random fuzzy number

The comparative robustness of the sample 1-norm median of a random fuzzy number as an estimator of the population median, in contrast to that of the sample mean of a random fuzzy number as an estimator of the population mean, is now to be discussed.

Before presenting a formal discussion and comparison, the first simulations in Section 1.5 (p. 54) when the mean is replaced by the 1-norm median are to be analyzed. To determine the effect of the contamination on the median of the random fuzzy number $\mathcal{X}$, the mean distance between the non-contaminated 'distribution' and the Monte Carlo approximated 1-norm median is collected in Table 3.3 for the different values of c_p and C_D and CASES 1 and 2. Contrary to the results in Section 1.5, the results in Table 3.3 show that the expected distance between the non-contaminated distribution and the sample median only slightly changes when the amount of contamination is increased, even when the contamination lies far from the non-contaminated distribution.

The analysis of the robustness of the 1-norm median in comparison to the mean is now made through the finite sample breakdown point, quantifying the minimum proportion of sample data which should be perturbed to get an arbitrarily large or small estimator value. Following Donoho and Huber [59], the fsbp of the sample median in a sample of size n from a random fuzzy number $\mathcal{X}$ is given by

$$\mathrm{fsbp}(\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n, \widetilde{\mathbf{x}}_n, \boldsymbol{\rho_1})$$

$$= \frac{1}{n} \min \left\{ k \in \{1, \ldots, n\} \, : \, \sup_{Q_{n,k}} \boldsymbol{\rho_1}(\widehat{\widetilde{\mathrm{Me}}(P_n)}, \widehat{\widetilde{\mathrm{Me}}(Q_{n,k})}) = \infty \right\},$$

where $\widetilde{\mathbf{x}}_n$ denotes the considered sample of n data from the metric space $(\mathcal{F}_c(\mathbb{R}), \boldsymbol{\rho_1})$ in which $\sup_{\widetilde{U},\widetilde{V}\in\mathcal{F}_c(\mathbb{R})} \boldsymbol{\rho_1}(\widetilde{U}, \widetilde{V}) = \infty$, P_n is the empirical distribution of $\widetilde{\mathbf{x}}_n$ and $Q_{n,k}$ is the empirical distribution of sample $\widetilde{\mathbf{y}}_{n,k}$ obtained from the original one $\widetilde{\mathbf{x}}_n$ by perturbing at most k components. Then, we have that

Proposition 3.3.7. *The finite sample breakdown point of the sample 1-norm median from a random fuzzy number $\mathcal{X}$, $\mathrm{fsbp}\big(\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n, \widetilde{\mathbf{x}}_n, \boldsymbol{\rho_1}\big)$, equals*

$$\mathrm{fsbp}\big(\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n, \widetilde{\mathbf{x}}_n, \boldsymbol{\rho_1}\big) = \frac{1}{n} \cdot \lfloor \frac{n+1}{2} \rfloor,$$

where $\lfloor \cdot \rfloor$ denotes the floor function.

Proof. First note that the condition $\sup_{\widetilde{U},\widetilde{V}\in\mathcal{F}_c(\mathbb{R})} \boldsymbol{\rho_1}(\widetilde{U}, \widetilde{V}) = \infty$ is satisfied in this case, since $\boldsymbol{\rho_1}\big(\mathbb{1}_{[n-1,n+1]}, \mathbb{1}_{[-n-1,-n+1]}\big) = 2n$.

Furthermore,

$$\boldsymbol{\rho_1}(\widehat{\widetilde{\mathrm{Me}}(P_n)}, \widehat{\widetilde{\mathrm{Me}}(Q_{n,k})}) \geq \int_{[0,1]} \frac{1}{2} \cdot |\inf\big(\widehat{\widetilde{\mathrm{Me}}(P_n)}\big)_\alpha - \inf\big(\widehat{\widetilde{\mathrm{Me}}(Q_{n,k})}\big)_\alpha| \, d\ell(\alpha)$$

$$= \int_{[0,1]} \frac{1}{2} \cdot |\mathrm{Me}\big(\widehat{\inf(P_n)_\alpha}\big) - \mathrm{Me}\big(\widehat{\inf(Q_{n,k})_\alpha}\big)| \, d\ell(\alpha).$$

Therefore, by recalling the fsbp for the sample median of a real-valued random variable, one can conclude that whenever at least $\lfloor \frac{n+1}{2} \rfloor$ elements $\widetilde{x}_i \in \mathcal{F}_c(\mathbb{R})$ of $\widetilde{\mathbf{x}}_n$ are replaced by other arbitrarily 'large' elements in $\mathcal{F}_c(\mathbb{R})$ so that

$$\sup_{Q_{n,k}} \int_{[0,1]} \frac{1}{2} \cdot |\mathrm{Me}\big(\widehat{\inf(P_n)_\alpha}\big) - \mathrm{Me}\big(\widehat{\inf(Q_{n,k})_\alpha}\big)| \, d\ell(\alpha) = \infty,$$

we have that

$$\sup_{Q_{n,k}} \boldsymbol{\rho_1}(\widehat{\widetilde{\mathrm{Me}}(P_n)}, \widehat{\widetilde{\mathrm{Me}}(Q_{n,k})})$$

$$\geq \frac{1}{2} \cdot \sup_{Q_{n,k}} \int_{[0,1]} \frac{1}{2} \cdot |\mathrm{Me}\big(\widehat{\inf(P_n)_\alpha}\big) - \mathrm{Me}\big(\widehat{\inf(Q_{n,k})_\alpha}\big)| \, d\ell(\alpha) = \infty,$$

whence

$$\mathrm{fsbp}(\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n, \widetilde{\mathbf{x}}_n, \boldsymbol{\rho_1}) \leq \frac{1}{n} \cdot \lfloor \frac{n+1}{2} \rfloor.$$

On the other hand, by using the fsbp for the sample median of a real-valued random variable, we have that for all α

$$\min \left\{ k \in \{1, \ldots, n\} \, : \, \sup_{Q_{n,k}} |\mathrm{Me}\big(\widehat{\inf(P_n)_\alpha}\big) - \mathrm{Me}\big(\widehat{\inf(Q_{n,k})_\alpha}\big)| = \infty \right\} = \lfloor \frac{n+1}{2} \rfloor,$$

$$\min \left\{ k \in \{1, \ldots, n\} \, : \, \sup_{Q_{n,k}} |\mathrm{Me}\big(\widehat{\sup(P_n)_\alpha}\big) - \mathrm{Me}\big(\widehat{\sup(Q_{n,k})_\alpha}\big)| = \infty \right\} = \lfloor \frac{n+1}{2} \rfloor,$$

whence

$$\sup_{Q_{n,\lfloor\frac{n+1}{2}\rfloor-1}} |\mathrm{Me}\big(\widehat{\inf(P_n)_\alpha}\big) - \mathrm{Me}\big(\widehat{\inf(Q_{n,\lfloor\frac{n+1}{2}\rfloor-1})_\alpha}\big)| = M_1 < \infty,$$

$$\sup_{Q_{n,\lfloor\frac{n+1}{2}\rfloor-1}} |\mathrm{Me}\big(\widehat{\sup(P_n)_\alpha}\big) - \mathrm{Me}\big(\widehat{\sup(Q_{n,\lfloor\frac{n+1}{2}\rfloor-1})_\alpha}\big)| = M_2 < \infty,$$

and therefore

$$\sup_{Q_{n,\lfloor\frac{n+1}{2}\rfloor-1}} \boldsymbol{\rho_1}(\widehat{\widetilde{\mathrm{Me}}(P_n)}, \widehat{\widetilde{\mathrm{Me}}(Q_{n,\lfloor\frac{n+1}{2}\rfloor-1})})$$

$$= \sup_{Q_{n,\lfloor\frac{n+1}{2}\rfloor-1}} \left[\frac{1}{2}\int_{[0,1]} |\mathrm{Me}\big(\widehat{\inf(P_n)_\alpha}\big) - \mathrm{Me}\big(\widehat{\inf(Q_{n,\lfloor\frac{n+1}{2}\rfloor-1})_\alpha}\big)|\, d\ell(\alpha)\right.$$

$$\left.+\frac{1}{2}\int_{[0,1]} |\mathrm{Me}\big(\widehat{\inf(P_n)_\alpha}\big) - \mathrm{Me}\big(\widehat{\inf(Q_{n,\lfloor\frac{n+1}{2}\rfloor-1})_\alpha}\big)|\, d\ell(\alpha)\right] \leq \frac{M_1+M_2}{2} < \infty.$$

Consequently,

$$\min\left\{k \in \{1,\ldots,n\} \,:\, \sup_{Q_{n,k}} \boldsymbol{\rho_1}(\widehat{\widetilde{\mathrm{Me}}(P_n)}, \widehat{\widetilde{\mathrm{Me}}(Q_{n,k})}) = \infty\right\} > \lfloor\frac{n+1}{2}\rfloor - 1,$$

whence

$$\mathrm{fsbp}(\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n, \widetilde{\mathbf{x}}_n, \boldsymbol{\rho_1}) \geq \frac{1}{n}\cdot\lfloor\frac{n+1}{2}\rfloor. \qquad \square$$

The following result formalizes the comparison of the robustness of the sample 1-norm median and the sample mean of a random fuzzy number. Thus,

Theorem 3.3.8. *The finite sample breakdown point of the sample mean from a random fuzzy number* $\mathcal{X}$*,* $\mathrm{fsbp}\big(\overline{\mathcal{X}_n}\big)$*, is lower than that for the sample median for sample sizes* $n > 2$*.*

Proof. Indeed, by arguing like for the preceding proposition we have that

$$\mathrm{fsbp}(\overline{\mathcal{X}_n}, \widetilde{\mathbf{x}}_n, \boldsymbol{\rho_1}) = \frac{1}{n},$$

and, consequently,

$$\mathrm{fsbp}(\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n, \widetilde{\mathbf{x}}_n, \boldsymbol{\rho_1}) \geq \frac{n/2}{n} = \frac{1}{2} > \frac{1}{n} = \mathrm{fsbp}(\overline{\mathcal{X}_n}, \widetilde{\mathbf{x}}_n, \boldsymbol{\rho_1}). \qquad \square$$

The sample mean has the lowest possible breakdown point while the sample median can withstand up to 50% of contamination. This huge difference can be also stressed in the fuzzy case. It means that the definition of fuzzy median in this paper succeeds in inheriting the robustness properties of the real valued sample median.

The theoretical conclusion in Theorem 3.3.8 can be corroborated empirically by analyzing the simulations in Section 1.5 and those in Table 3.3. Moreover, and on the basis of these simulations an additional table has been constructed.

Table 3.4 gathers empirical results for the influence of contamination on both the sample mean and 1-norm median, by computing the distances between the mean/median of the non-contaminated sample and the mean/median of the contaminated sample, respectively, for the different values of c_p and C_D and the CASES 1 and 2 in Tables 1.2 and 3.3.

		CASE 1				CASE 2			
		mean		1-norm median		mean		1-norm median	
c_P	c_D	ρ_1	ρ_2	ρ_1	ρ_2	ρ_1	ρ_2	ρ_1	ρ_2
0.0	0	0.004850	0.005446	0.004936	0.005688	0.002796	0.003035	0.001443	0.001726
0.1	0	0.334582	0.340882	0.139385	0.141140	0.011580	0.012257	0.002781	0.003251
0.1	1	0.444295	0.457775	0.153509	0.156870	0.107663	0.120268	0.019206	0.021525
0.1	5	0.901119	0.965272	0.169022	0.179650	0.576996	0.641559	0.040033	0.044311
0.1	10	1.485318	1.614466	0.172593	0.184649	1.177204	1.306176	0.042455	0.046530
0.1	100	12.147356	13.479923	0.175099	0.186307	11.806141	13.156037	0.042852	0.046511
0.2	0	0.668190	0.680722	0.309054	0.313257	0.022772	0.023612	0.005022	0.005626
0.2	1	0.897204	0.924358	0.352139	0.360181	0.216280	0.241804	0.040727	0.045634
0.2	5	1.830108	1.952914	0.398174	0.423064	1.152823	1.283131	0.085405	0.095957
0.2	10	3.010209	3.261031	0.409606	0.436820	2.319878	2.579237	0.090404	0.100944
0.2	100	24.129091	26.793580	0.407285	0.436193	23.572994	26.292482	0.091153	0.100430
0.4	0	1.339551	1.364644	0.806692	0.821507	0.045481	0.046735	0.012593	0.013589
0.4	1	1.821615	1.874279	1.025370	1.052681	0.436245	0.485976	0.100506	0.114768
0.4	5	3.606433	3.860446	1.260487	1.360074	2.290189	2.551476	0.244459	0.271295
0.4	10	5.964314	6.470993	1.374637	1.499189	4.616273	5.133452	0.263837	0.291161
0.4	100	48.697644	54.170428	1.396510	1.515434	46.673034	51.856741	0.267270	0.289860

Table 3.4: Distances between the sample mixed (partially contaminated and non-contaminated) mean/1-norm median to the non-contaminated one for a random fuzzy number

On the basis of these simulations and by comparing the results in Tables 1.2 and Table 3.3, and the results in Table 3.4, one can empirically conclude that

- for a fixed level of contamination c_P, the farther the contaminated distribution from the non-contaminated one, the substantially greater mean $\boldsymbol{\rho_2}$-distance between the approximated mean and the non-contaminated distribution, whereas for the approximated 1-norm median the increase is modest; actually this mean distance asymptotically would only depend on a certain fractile of the non-contaminated distribution;

- for a fixed level of contamination c_P, the farther the contaminated distribution from the non-contaminated one, the substantially greater distance between the contaminated and the non-contaminated means, whereas for the 1-norm medians the increase is not really substantial.

To conclude this section, three remarks should be made. The first one relates to the fact that replacing $\boldsymbol{\rho_1}$ by $\boldsymbol{\rho_1^\varphi}$ does not affect the minimization problem in this section. The second and the third concern the particularization of the 1-norm median to the interval-valued case and its extension to the fuzzy vector-valued case, respectively.

Remark 3.3.6. The minimization problem and solutions in Definitions 3.3.1 and Theorems 3.3.1 and 3.3.2 would remain clearly invariant if one replaces $\boldsymbol{\rho_1}$ by $\boldsymbol{\rho_1^\varphi}$.

Remark 3.3.7. All the notions and results in this section can be trivially particularized to the interval-valued case, so that the 1-norm median of a random interval X is given by the interval value $[\mathrm{Me}(\inf \mathsf{X}), \mathrm{Me}(\sup \mathsf{X})]$.

Remark 3.3.8. The extension of the 1-norm median to the fuzzy vector-valued case cannot be made by extending arguments in Theorem 3.3.1. In other words, by using $\boldsymbol{\rho_1}$ defined on $\mathcal{F}_c(\mathbb{R}^p)$ on the basis of the support function and for a random fuzzy set $\mathcal{X}$, if one attempts to define $\widetilde{\mathrm{Me}}(\mathcal{X})$ as the value in $\mathcal{F}_c(\mathbb{R}^p)$ such that

$$s_{\widetilde{\mathrm{Me}}(\mathcal{X})}(\alpha, \mathbf{u}) = \mathrm{Me}(s_{\mathcal{X}}(\alpha, \mathbf{u}))$$

for all $\alpha \in [0,1]$ and $\mathbf{u} \in \mathbb{S}^{p-1}$, this is not right in general since this real-valued median does not necessarily determine the support function of a fuzzy set. Reasons for the last assertion lie in the fact that, whereas the median of a real-valued random variable preserves monotonicity and continuity, and hence necessary and sufficient conditions characterizing fuzzy numbers by means of the inf/sup representation (see Theorem 3.3.1), it does not preserve subadditivity, so that the necessary and sufficient conditions characterizing fuzzy set values by means of the support function (Proposition 1.1.2) are not generally fulfilled.

3.4 The φ-wabl/ldev/rdev median for a random fuzzy number

One of the main advantages of the 1-norm median in practice is that it can be computed on the basis of the medians of certain real-valued random variables. This makes computations rather easy-to-perform and, mainly, easy to implementing and programming in R or others. At this point, we should indicate that when the involved L^1 metric is replaced by other ones, the minimization problem can become a very difficult task, and often infeasible at least to get the exact solution.

In this section an L^1 metric based on another representation of fuzzy numbers for which there exists a set of sufficient conditions characterizing them is to be considered for purposes similar to those in the preceding section. This L^1 metric extends Hausdorff's one from the interval to the fuzzy number-valued case and corresponds to $\mathscr{D}_\theta^\varphi$ introduced in Definition 1.3.8.

The population and sample φ-wabl/ldev/rdev median for random fuzzy numbers are defined as follows:

Definition 3.4.1. *Given a probability space $(\Omega, \mathcal{A}, P)$, an absolutely continuous probability measure φ on the measurable space $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function on $(0,1)$, and an associated random fuzzy number $\mathcal{X}$, the **population φ-wabl/ldev/rdev median(s)** of $\mathcal{X}$ is the fuzzy number(s)*

$$\widetilde{\mathrm{M}}^\varphi(\mathcal{X}) = \arg \min_{\widetilde{U} \in \mathcal{F}_c(\mathbb{R})} E\left(\mathscr{D}_1^\varphi(\mathcal{X}, \widetilde{U})\right),$$

whenever these expectations exist.

Definition 3.4.2. *Given a probability space $(\Omega, \mathcal{A}, P)$, an absolutely continuous probability measure φ on the measurable space $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function on $(0,1)$, an associated random fuzzy number $\mathcal{X}$, and a simple random sample $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ from $\mathcal{X}$, the **sample φ-wabl/ldev/rdev median(s)** of $\mathcal{X}$ is(are) the fuzzy number-valued statistic(s)*

$$\widehat{\widetilde{\mathrm{M}}^\varphi(\mathcal{X})}_n = \arg \min_{\widetilde{U} \in \mathcal{F}_c(\mathbb{R})} \frac{1}{n} \sum_{i=1}^{n} \left(\mathscr{D}_1^\varphi(\mathcal{X}_i, \widetilde{U})\right).$$

As for the 1-norm median, two key questions at this stage are whether the φ-wabl/ldev/rdev median exists and whether it can be computed easily in practice. A result is first to be established for the population measure. The result guarantees that at least one such median always exists and it is easy to compute.

Theorem 3.4.1. *Given a probability space $(\Omega, \mathcal{A}, P)$, an absolutely continuous probability measure φ on the measurable space $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function on $(0,1)$, and an associated random fuzzy number $\mathcal{X}$, for any $\alpha \in [0,1]$, the fuzzy number $\widetilde{\mathrm{M}}^\varphi(\mathcal{X}) \in \mathcal{F}_c(\mathbb{R})$ such that*

$$\left(\widetilde{\mathrm{M}}^\varphi(\mathcal{X})\right)_\alpha = \left[\mathrm{Me}\big(\mathrm{wabl}^\varphi(\mathcal{X})\big) - \mathrm{Me}\big(\mathrm{ldev}_{\mathcal{X}}^\varphi(\alpha)\big), \mathrm{Me}\big(\mathrm{wabl}^\varphi(\mathcal{X})\big) + \mathrm{Me}\big(\mathrm{rdev}_{\mathcal{X}}^\varphi(\alpha)\big)\right],$$

(where in case $\mathrm{Me}\big(\mathrm{wabl}^\varphi(\mathcal{X})\big)$, $\mathrm{Me}\big(\mathrm{ldev}_{\mathcal{X}}^\varphi(\alpha)\big)$ or $\mathrm{Me}\big(\mathrm{rdev}_{\mathcal{X}}^\varphi(\alpha)\big)$ are non-unique the most usual convention, in accordance with which these real-valued median(s) are chosen to be the midpoint of the interval of medians of is considered) is a population φ-wabl/ldev/rdev median of $\mathcal{X}$.

Proof. Indeed, on one hand, whatever $\alpha \in [0,1]$ and $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ may be, since $\text{ldev}^{\varphi}_{\widetilde{U}}(\alpha), \text{rdev}^{\varphi}_{\widetilde{U}}(\alpha), \text{wabl}^{\varphi}(\widetilde{U}) \in \mathbb{R}$, and $\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha)$, $\text{rdev}^{\varphi}_{\mathcal{X}}(\alpha)$, and $\text{wabl}^{\varphi}(\mathcal{X})$ are real-valued random variables, we have that

$$E\left[|\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha) - \text{Me}\big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha)\big)|\right] \leq E\left[|\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha) - \text{ldev}^{\varphi}_{\widetilde{U}}(\alpha)|\right],$$

$$E\left[|\text{rdev}^{\varphi}_{\mathcal{X}}(\alpha) - \text{Me}\big(\text{rdev}^{\varphi}_{\mathcal{X}}(\alpha)\big)|\right] \leq E\left[|\text{rdev}^{\varphi}_{\mathcal{X}}(\alpha) - \text{rdev}^{\varphi}_{\widetilde{U}}(\alpha)|\right],$$

$$E\left[|\text{wabl}^{\varphi}(\mathcal{X}) - \text{Me}\big(\text{wabl}^{\varphi}(\mathcal{X})\big)|\right] \leq E\left[|\text{wabl}^{\varphi}(\mathcal{X}) - \text{wabl}^{\varphi}(\widetilde{U})|\right],$$

whence

$$E\left(\mathscr{D}^{\varphi}_{1}(\mathcal{X}, \widetilde{U})\right) \geq E\left(\mathscr{D}^{\varphi}_{1}(\mathcal{X}, \widetilde{\text{M}}^{\varphi}(\mathcal{X}))\right).$$

On the other hand, $\widetilde{\text{M}}^{\varphi}(\mathcal{X}) \in \mathcal{F}_c(\mathbb{R})$ since it satisfies the three conditions in Proposition 1.1.4 (p. 8) as one can see now.

Regarding Condition *i*) in Proposition 1.1.4, over all Ω we have that $\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha)$ and $\text{rdev}^{\varphi}_{\mathcal{X}}(\alpha)$ are non-increasing functions of α in $[0,1]$ whence, because of the considered convention, $\text{Me}(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha))$ and $\text{Me}(\text{rdev}^{\varphi}_{\mathcal{X}}(\alpha))$ are non-increasing functions of α in $[0,1]$.

Furthermore, functions $\text{Me}(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha))$ and $\text{Me}(\text{rdev}^{\varphi}_{\mathcal{X}}(\alpha))$ are left-continuous at every $\alpha \in (0,1]$. Indeed, if $\{\alpha_n\}_n \uparrow \alpha \in (0,1]$ as $n \to \infty$, then for all element in Ω we have that $\{\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n)\}_n \downarrow \text{ldev}^{\varphi}_{\mathcal{X}}(\alpha)$ and because of the considered convention the sequence $\big\{\text{Me}\big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n)\big)\big\}_n \downarrow$ is bounded below, $\text{Me}\big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha)\big)$ being a lower bound. Hence, a limit for this sequence exists and will be denoted by $L^{\varphi}_{\alpha} = \lim_{n\to\infty} \text{Me}\big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n)\big) \geq \text{Me}\big(\text{ldev}\,\mathcal{X}_{\alpha}\big)$. Actually, one can prove that $L^{\varphi}_{\alpha} = \text{Me}\big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha)\big)$, since for all $\omega \in \Omega$ we have that

$$0.5 \leq P\Big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n) \geq \text{Me}\big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n)\big)\Big) \leq P\Big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n) \leq L^{\varphi}_{\alpha}\Big)$$

and

$$\Big\{\big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n) \geq L^{\varphi}_{\alpha}\big)\Big\}_n \downarrow \bigcap_{n=1}^{\infty} \big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n) \geq L^{\varphi}_{\alpha}\big) = \big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha) \geq L^{\varphi}_{\alpha}\big),$$

whence

$$\begin{aligned} P\Big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha) \geq L^{\varphi}_{\alpha}\Big) &= P\Big(\lim_{n\to\infty} \big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n) \geq L^{\varphi}_{\alpha}\big)\Big) \\ &= \lim_{n\to\infty} P\Big(\text{ldev}^{\varphi}_{\mathcal{X}} \geq L^{\varphi}_{\alpha}\Big) \geq 0.5. \end{aligned}$$

Following similar arguments,

$$P\Big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha) > L^{\varphi}_{\alpha}\Big) = P\left(\bigcup_{n=1}^{\infty} \Big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha) > \text{Me}\big(\text{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n)\big)\Big)\right)$$

$$= P\left(\lim_{n\to\infty}\Big(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha) > \mathrm{Me}\big(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n)\big)\Big)\right) = \lim_{n\to\infty} P\Big(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha) > \mathrm{Me}\big(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n)\big)\Big)$$
$$\le \lim_{n\to\infty} P\Big(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n) > \mathrm{Me}\big(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha_n)\big)\Big) \le 0.5.$$

Consequently, taking into account the considered convention, we have that $L^{\varphi}_{\alpha} = \mathrm{Me}\big(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha)\big)$.

Analogously, if $\{\alpha_n\}_n \uparrow \alpha \in (0,1]$ as $n \to \infty$, it holds that $\{\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha_n)\}_n \downarrow \mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha)$ and the sequence $\big\{\mathrm{Me}\big(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha_n)\big)\big\}_n \downarrow$ and it is bounded below by $\mathrm{Me}\big(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha)\big)$, so that there exists $L'^{\varphi}_{\alpha} = \lim_{n\to\infty} \mathrm{Me}\big(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha_n)\big)$ and we can easily prove that $L'^{\varphi}_{\alpha} = \mathrm{Me}\big(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha)\big)$.

The right-continuity at 0 of both, $\mathrm{Me}(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha))$ and $\mathrm{Me}(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha))$, can be proved by following similar arguments.

Condition *ii*) holds, since $-\mathrm{ldev}^{\varphi}_{\mathcal{X}}(1) \le \mathrm{rdev}^{\varphi}_{\mathcal{X}}(1)$ over all Ω whence, because of the considered convention, one can guarantee that

$$-\mathrm{ldev}^{\varphi}_{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}(1) = \mathrm{Me}(-\mathrm{ldev}^{\varphi}_{\mathcal{X}}(1)) \le \mathrm{Me}(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(1)) = \mathrm{rdev}^{\varphi}_{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}(1).$$

Finally,

$$\int_{[0,1]} \mathrm{ldev}^{\varphi}_{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}(\alpha)\, d\varphi(\alpha) = \int_{[0,1]} \frac{\mathrm{Me}(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha)) + \mathrm{Me}(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha))}{2}\, d\varphi(\alpha)$$
$$= \int_{[0,1]} \mathrm{rdev}^{\varphi}_{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}(\alpha)\, d\varphi(\alpha),$$

whence Condition *iii*) is clearly fulfilled. □

Similarly, for the sample approach

Theorem 3.4.2. *Given a probability space* $(\Omega, \mathcal{A}, P)$*, an absolutely continuous probability measure* φ *on the measurable space* $([0,1], \mathcal{B}_{[0,1]})$ *with positive mass function on* $(0,1)$*, an associated random fuzzy number* $\mathcal{X}$*, and a simple random sample* $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ *from* $\mathcal{X}$*, the fuzzy number-valued statistic such that for each* $\alpha \in [0,1]$

$$\left(\widehat{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}_n\right)_{\alpha} = \left[\widehat{\mathrm{Me}}(\widehat{\mathrm{wabl}^{\varphi}}\,\mathcal{X})_n - \widehat{\mathrm{Me}}(\widehat{\mathrm{ldev}^{\varphi}_{\mathcal{X}}}(\alpha))_n, \widehat{\mathrm{Me}}(\widehat{\mathrm{wabl}^{\varphi}}\,\mathcal{X})_n + \widehat{\mathrm{Me}}(\widehat{\mathrm{rdev}^{\varphi}_{\mathcal{X}}}(\alpha))_n\right],$$

where $\widehat{\mathrm{Me}(\cdot)}$ *denotes the sample median of the corresponding real-valued random variable and makes use of a convention similar to that in Theorem 3.4.1, is a sample* φ*-wabl/ldev/rdev-median of* $\mathcal{X}$*.*

Remark 3.4.1. It should be emphasized that the φ-wabl/ldev/rdev median has been introduced, and results in Theorems 3.4.1 and 3.4.2 have been developed, on the basis of the L^1 metric $\mathscr{D}^{\varphi}_{\theta}$ when $\theta = 1$. However, the conclusions and proofs in

Theorems 3.4.1 and 3.4.2 do not depend at all of such a choice, in the same way that the Aumann-type mean value of a random fuzzy number does not depend on the possible weight θ of the considered L^2 metrics. Of course, they depend on the choice of φ, as will be illustrated in the second part of Example 3.4.1.

Remark 3.4.2. As for the 1-norm median, the wabl/ldev/rdev ones is not necessarily a value the random fuzzy number takes on, even when we don't need to make use of the convention. As an example corroborating this assertion and illustrating the computation of the median we can consider, the situation in Example 3.2.1 (see Figure 3.10 on the left) with the induced probabilities equal to 1/3.

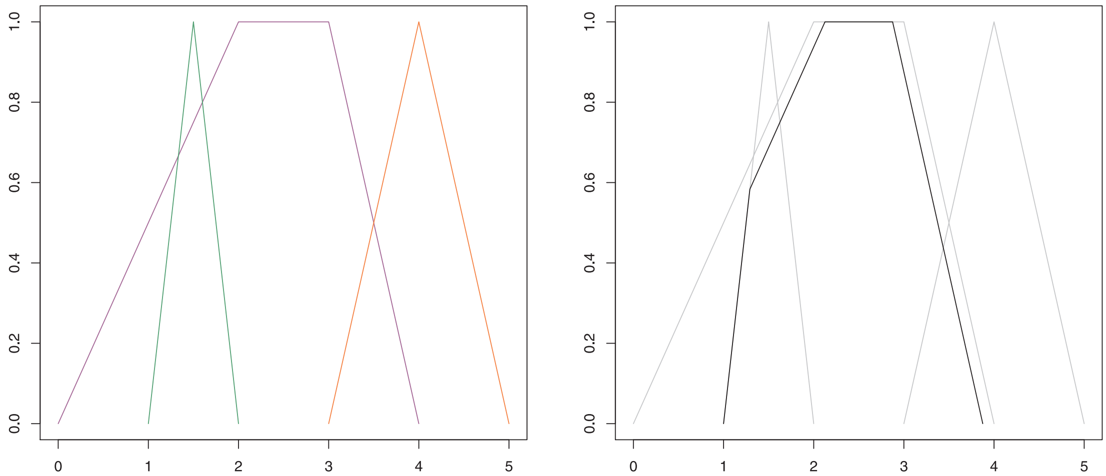


Figure 3.10: Counterexample: the φ-wabl/ldev/rdev median of a random fuzzy number, even if uniquely valued, does not necessarily corresponds to a value the random fuzzy number takes on

The corresponding ℓ-wabl/ldev/rdev can be trivially determined and it is given by the fuzzy number in Figure 3.10 on the right, which does not coincide with any of the three values of the random fuzzy number, and this is not affected by the convention which has not been required in this example. In cases the φ-wabl/ldev/rdev median cannot be immediately derived, it can be approximated by using a large number of levels.

3.4.1 Basic properties of the φ-wabl/ldev/rdev median of a random fuzzy number

The φ-wabl/ldev/rdev median of a random fuzzy number preserves most of the basic properties of the median of a random variable, irrespective of the considered version being the population or the sample one. Thus, on the basis of the results in Theorems 3.4.1 and 3.4.2, it can be straightforwardly proved that

Proposition 3.4.3. $\widetilde{\mathrm{M}}^{\varphi}$ *is equivariant under 'linear' transformations, that is, if* $\gamma \in \mathbb{R}$, $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ *and* $\mathcal{X}$ *is a random fuzzy number, then*

$$\widetilde{\mathrm{M}}^{\varphi}(\gamma \cdot \mathcal{X} + \widetilde{U}) = \gamma \cdot \widetilde{\mathrm{M}}^{\varphi}(\mathcal{X}) + \widetilde{U}.$$

Consequently, if $\mathcal{X}$ *is a random fuzzy number associated with the probability space* $(\Omega, \mathcal{A}, P)$ *and the distribution of* $\mathcal{X}$ *is degenerate at a fuzzy number* $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ *(i.e.,* $\mathcal{X} = \widetilde{U}$ *a.s.* $[P]$*), then* $\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X}) = \widetilde{U}$.

Remark 3.4.3. The φ-wabl/ldev/rdev median of a random fuzzy number cannot be formalized as a 'middle position' value with respect to the fuzzy max partial order. Actually, although one can find a partial orderings w.r.t. which a property similar to that in Proposition 3.3.4 holds for the φ-wabl/ldev/rdev median, it is not as simple and easy-to-interpret as the fuzzy max one. In this way, the φ-wabl/ldev/rdev median can be formalized as a 'middle position' value with respect to the partial ordering which is the level-wise extension through the φ-wabl/ldev/rdev representation of the product order on $\mathbb{R}^3$, so that $\widetilde{U} \precsim_{\mathrm{wlr}} \widetilde{V}$ if and only if $\mathrm{wabl}^{\varphi}(\widetilde{U}) \leq \mathrm{wabl}^{\varphi}(\widetilde{V})$ and for all $\alpha \in [0,1]$ one has that $\mathrm{ldev}^{\varphi}_{\widetilde{U}}(\alpha) \leq \mathrm{ldev}^{\varphi}_{\widetilde{V}}(\alpha)$, $\mathrm{rdev}^{\varphi}_{\widetilde{U}}(\alpha) \leq \mathrm{rdev}^{\varphi}_{\widetilde{V}}(\alpha)$ (or the two last inequalities are just the opposite simultaneously).

Another interesting property in examining the adequacy of the φ-wabl/ldev/rdev median for random fuzzy numbers as a central tendency measure is now discussed by considering their behaviour in case of symmetrically distributed random fuzzy numbers. As for the 1-norm median, the φ-wabl/ldev/rdev median shows a suitable central tendency behaviour since it leads to a fuzzy number which is symmetric about the symmetry point. Moreover, this measure neither necessarily coincides nor corresponds to any of the values the random fuzzy number takes on. Thus,

Proposition 3.4.4. *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space,* φ *be an absolutely continuous probability measure on the measurable space* $([0,1], \mathcal{B}_{[0,1]})$ *with positive mass function on* $(0,1)$*, and let* $\mathcal{X}$ *be a symmetric random fuzzy number about* $c \in \mathbb{R}$*. Then, the* φ*-wabl/ldev/rdev median of* $\mathcal{X}$ *is a symmetric fuzzy number about* c*.*

Proof. Since $\mathcal{X} \stackrel{d}{=} 2c - \mathcal{X}$, then $\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X}) = \widetilde{\mathrm{M}}^{\varphi}(2c - \mathcal{X})$, whence because of the equivariance properties of $\widetilde{\mathrm{M}}^{\varphi}$ under affine transformations, we have that

$$\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X}) = 2c - \widetilde{\mathrm{M}}^{\varphi}(\mathcal{X}).$$

By adding $\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})$ to the two members in the last equality, $2\,\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X}) = 2c + \widetilde{\mathrm{M}}^{\varphi}(\mathcal{X}) - \widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})$ and, hence,

$$\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X}) = c + \frac{1}{2} \cdot \mathcal{O}_{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})},$$

whence for each $\alpha \in [0,1]$

$$\left(\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})\right)_{\alpha} = \left[c - \mathrm{spr}\left(\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})\right)_{\alpha}, c + \mathrm{spr}\left(\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})\right)_{\alpha}\right],$$

which leads to a symmetric fuzzy number about c. □

The result in Proposition 3.4.4 is now illustrated by computing the two ℓ-wabl/ldev/rdev medians of the symmetric random fuzzy numbers in Examples 1.4.3 and 1.4.4.

Example 3.4.1. To compute the ℓ-wabl/ldev/rdev median of the symmetric random fuzzy number about 0.5 in Example 1.4.3 (see also Example 3.3.1, p. 130) we should take into account that

label	VD	D	SD	N	SA	A	VA
absol. freq.	38	143	207	177	207	143	38
wabl^{ℓ}	1/24	4/24	8/24	12/24	16/24	20/24	23/24
$\mathrm{ldev}^{\ell}(\alpha)$	1/24	$(1-\alpha)/6$	$(1-\alpha)/6$	$(1-\alpha)/6$	$(1-\alpha)/6$	$(1-\alpha)/6$	$(3-4\alpha)/24$
$\mathrm{rdev}^{\ell}(\alpha)$	$(3-4\alpha)/24$	$(1-\alpha)/6$	$(1-\alpha)/6$	$(1-\alpha)/6$	$(1-\alpha)/6$	$(1-\alpha)/6$	1/24

whence, by developing a comparison of the values in each row as a function of α, one can easily conclude that

$$\widetilde{\mathrm{M}}^{\ell}(\mathcal{X}) = \mathrm{N}.$$

To compute the ℓ-wabl/ldev/rdev and the $\varphi \equiv \beta(1,500)$-wabl/ldev/rdev medians of the symmetric random fuzzy number about 2 in Example 1.4.4 (see also Example 3.3.1, p. 130) we should take into account that

label	$\gamma_{(2)}(0)$	$\gamma_{(2)}(1)$	$\gamma_{(2)}(2)$	$\gamma_{(2)}(3)$	$\gamma_{(2)}(4)$
probab.	0.0625	0.25	0.375	0.25	.0625
wabl^{ℓ}	$-1/4$	5/6	2	19/6	17/4
$\mathrm{ldev}^{\ell}(\alpha)$	$-\frac{1}{4}+\sqrt[3]{1-\alpha}$	$-\frac{1}{6}+\sqrt{1-\alpha}$	$1-\alpha$	$\frac{1}{6}+(1-\alpha)^2$	$\frac{1}{4}+(1-\alpha)^3$
$\mathrm{rdev}^{\ell}(\alpha)$	$\frac{1}{4}+(1-\alpha)^3$	$\frac{1}{6}+(1-\alpha)^2$	$1-\alpha$	$-\frac{1}{6}+\sqrt{1-\alpha}$	$-\frac{1}{4}+\sqrt[3]{1-\alpha}$

label	$\gamma_{(2)}(0)$	$\gamma_{(2)}(1)$	$\gamma_{(2)}(2)$	$\gamma_{(2)}(3)$	$\gamma_{(2)}(4)$
probab.	0.0625	0.25	0.375	0.25	.0625
$\mathrm{wabl}^{\beta(1,500)}$	-0.0027	0.9985	2	3.0015	4.0023
$\mathrm{ldev}^{\beta(1,500)}(\alpha)$	$-0.0027+\sqrt[3]{1-\alpha}$	$-0.0015+\sqrt{1-\alpha}$	$1-\alpha$	$0.0015+(1-\alpha)^2$	$0.0023+(1-\alpha)^3$
$\mathrm{rdev}^{\beta(1,500)}(\alpha)$	$0.0027+(1-\alpha)^3$	$0.0015+(1-\alpha)^2$	$1-\alpha$	$-0.0015+\sqrt{1-\alpha}$	$-0.0023+\sqrt[3]{1-\alpha}$

whence, the ℓ-wabl/ldev/rdev and the $\varphi \equiv \beta(1,500)$-wabl/ldev/rdev medians of $\gamma_{(2)} \circ \mathrm{Bin}(4, 0.5)$ has been graphically displayed in Figure 3.11, and they are very close to $\gamma_{(2)}(0)$, especially for the second choice for φ.

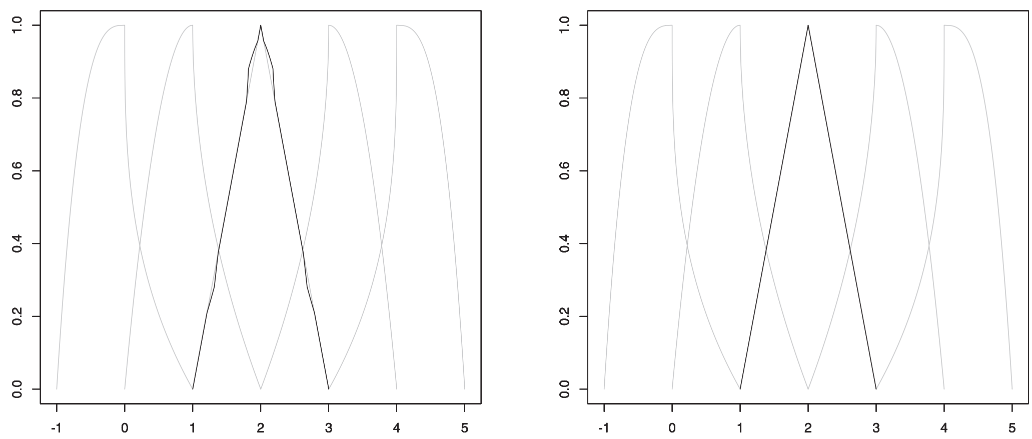


Figure 3.11: The ℓ-wabl/ldev/rdev median (in black on the left) and the $\varphi \equiv \beta(1,500)$-wabl/ldev/rdev median (in black on the right) of the characterizing fuzzy representation of the Bin$(4, 0.5)$ in Example 1.4.4

Consequently, one can assert that for symmetric random fuzzy numbers about c the two main central tendencies (i.e., the Aumann-type mean and the median, defined in accordance with the two approaches based on L^1 metrics between fuzzy numbers in this work) are symmetric fuzzy numbers about c, but they do not necessarily coincide.

To check whether the φ-wabl/ldev/rdev median is mostly closer than the Aumann-type mean (although maybe not that much as the 1-norm median) to the one which occupies the 'central position' in symmetric random fuzzy numbers, the three ones $\gamma_{(0)} \circ \mathcal{N}(0,1)$, $\gamma_{(0)} \circ$ Uniform$(-0.5, 0.5)$ and $\gamma_{(0)} \circ [\text{Bin}(5, 0.5) - 2.5]$ in Subsection 3.3.1 are to be examined.

After representing the (population) 1-norm median, ℓ-wabl/ldev/rdev median and Aumann-type mean of each of the random fuzzy numbers $\gamma_{(0)} \circ X$ graphically, distances between each of the summary measures and the correspondent central position value $\gamma_{(0)}(0)$ have been computed and graphically displayed (as functions of θ when the distance is parameterized). Conclusions are now presented.

Figure 3.12 shows that when the considered random fuzzy number is $\gamma_{(0)} \circ \mathcal{N}(0,1)$, whereas the 1-norm median coincides with the central position value, the ℓ-wabl/ldev/rdev median is quite close to it, but the Aumann-type mean is not that close.

This is also corroborated by computing the D^ℓ_θ-, $\mathscr{D}^\ell_\theta$- and $\boldsymbol{\rho_1}$-distances between each summary measure and $\gamma_{(0)}(0)$, the two first distances as functions of the weighting parameter θ. Thus,

$$\boldsymbol{\rho_1}\left(\widetilde{\text{Me}}\big(\gamma_{(0)} \circ \mathcal{N}(0,1)\big), \gamma_{(0)}(0)\right) = 0,$$
$$\boldsymbol{\rho_1}\left(\widetilde{\text{M}}^\ell\big(\gamma_{(0)} \circ \mathcal{N}(0,1)\big), \gamma_{(0)}(0)\right) = 0.0092,$$
$$\boldsymbol{\rho_1}\left(\widetilde{E}\big(\gamma_{(0)} \circ \mathcal{N}(0,1)\big), \gamma_{(0)}(0)\right) = 0.0530,$$

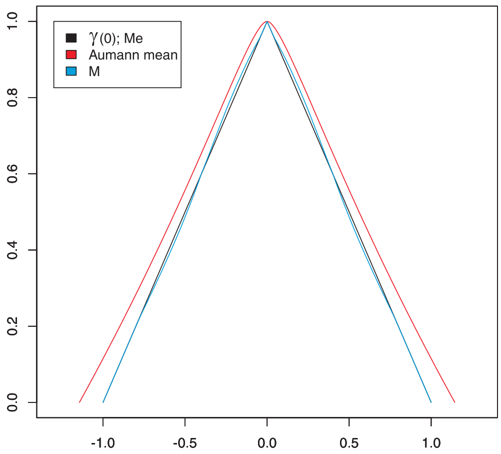


Figure 3.12: Aumann-type mean, ℓ-wabl/ldev/rdev median (M) and 1-norm median (Me) $= \gamma_{(0)}(0)$ of the random fuzzy number $\gamma_{(0)} \circ \mathcal{N}(0,1)$

and the D^ℓ_θ- and the $\mathscr{D}^\ell_\theta$-distances have been displayed in Figure 3.13.

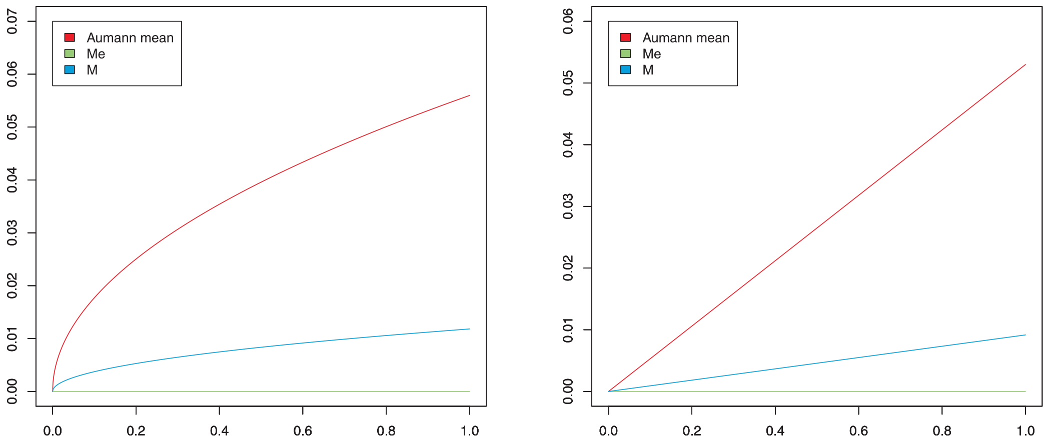


Figure 3.13: D^ℓ_θ-distance (on the left) and the $\mathscr{D}^\ell_\theta$-distance (on the right) between $\gamma_{(0)}(0)$ and the Aumann-type mean, ℓ-wabl/ldev/rdev median (M) and 1-norm median (Me) of the random fuzzy number $\gamma_{(0)} \circ \mathcal{N}(0,1)$ as functions of θ

Analogously, Figure 3.14 shows that when the considered random fuzzy number is $\gamma_{(0)} \circ \text{Uniform}(-0.5, 0.5)$, then the 1-norm median coincides with the central position value, and the ℓ-wabl/ldev/rdev median is quite close to it, whereas the Aumann-type mean is not that close.

The $\mathscr{D}^\ell_\theta$- and the $\mathscr{D}^\ell_\theta$-distances have been displayed in Figure 3.15, and

$$\boldsymbol{\rho_1}\left(\widetilde{\text{Me}}\big(\gamma_{(0)} \circ \text{Uniform}(-0.5, 0.5)\big), \gamma_{(0)}(0)\right) = 0,$$
$$\boldsymbol{\rho_1}\left(\widetilde{\text{M}}^\ell\big(\gamma_{(0)} \circ \text{Uniform}(-0.5, 0.5)\big), \gamma_{(0)}(0)\right) = 0.0032,$$
$$\boldsymbol{\rho_1}\left(\widetilde{E}\big(\gamma_{(0)} \circ \text{Uniform}(-0.5, 0.5)\big), \gamma_{(0)}(0)\right) = 0.1797.$$

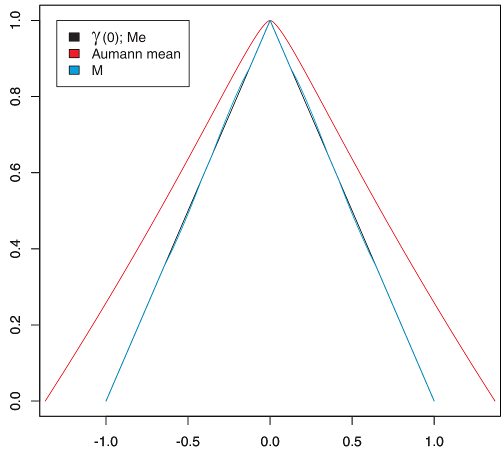


Figure 3.14: Aumann-type mean, ℓ-wabl/ldev/rdev median (M) and 1-norm median Me $= \gamma_{(0)}(0)$ of the random fuzzy number $\gamma_{(0)} \circ \text{Uniform}(-0.5, 0.5)$

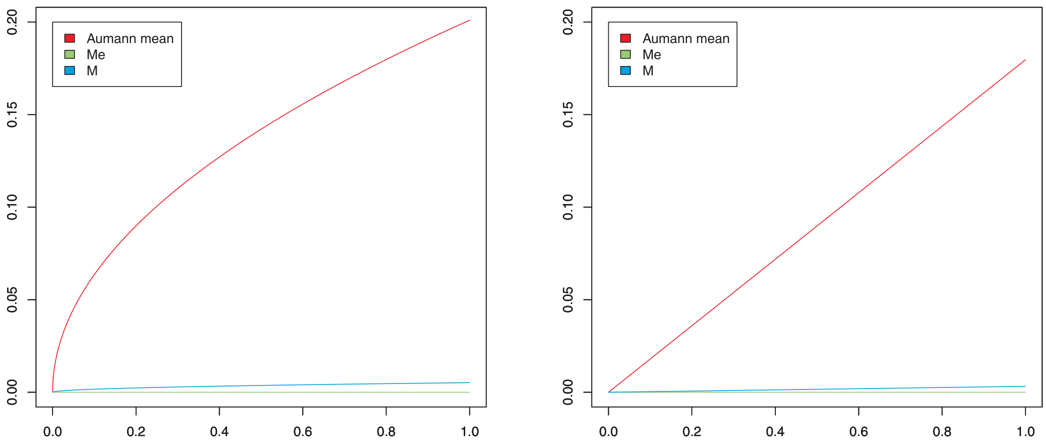


Figure 3.15: D_θ^ℓ-distance (on the left) and the $\mathscr{D}_\theta^\ell$-distance (on the right) between $\gamma_{(0)}(0)$ and the Aumann-type mean, ℓ-wabl/ldev/rdev median (M) and 1-norm median (Me) of the random fuzzy number $\gamma_{(0)} \circ \text{Uniform}(-0.5, 0.5)$ as functions of θ

Finally, Figure 3.16 shows for $\gamma_{(0)} \circ [\text{Bin}(5, 0.5) - 2.5]$ that the 1-norm and ℓ-wabl/ldev/rdev medians do not coincide with $\gamma_{(0)}(0)$, but they are slightly closer to it than the Aumann-type mean, the 1-norm median being the closest one.

Furthermore, D_θ^ℓ-, $\mathscr{D}_\theta^\ell$-distances between each summary measure and $\gamma_{(0)}(0)$, have been graphically displayed in Figure 3.17, and

$$\boldsymbol{\rho_1}\left(\widetilde{\text{Me}}\big(\gamma_{(0)} \circ [\text{Bin}(5, 0.5) - 2.5]\big), \gamma_{(0)}(0)\right) = 0.0108,$$

$$\boldsymbol{\rho_1}\left(\widetilde{\text{M}}^\ell\big(\gamma_{(0)}[\text{Bin}(5, 0.5) - 2.5]\big), \gamma_{(0)}(0)\right) = 0.0224,$$

$$\boldsymbol{\rho_1}\left(\widetilde{E}\big(\gamma_{(0)} \circ [\text{Bin}(5, 0.5) - 2.5]\big), \gamma_{(0)}(0)\right) = 0.0274.$$

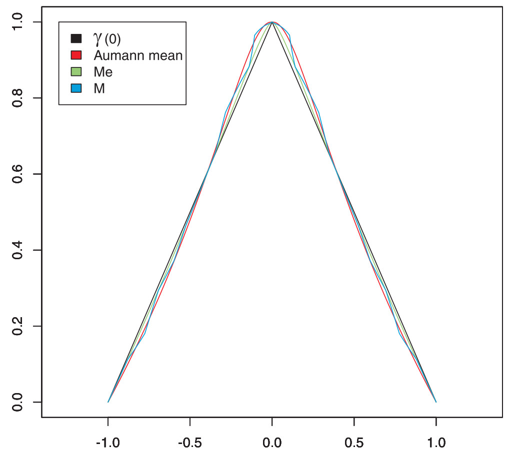


Figure 3.16: Aumann-type mean, ℓ-wabl/ldev/rdev median (M) and 1-norm median Me of the random fuzzy number $\gamma_{(0)} \circ [\text{Bin}(5, 0.5) - 2.5]$, and comparison with $\gamma_{(0)}(0)$

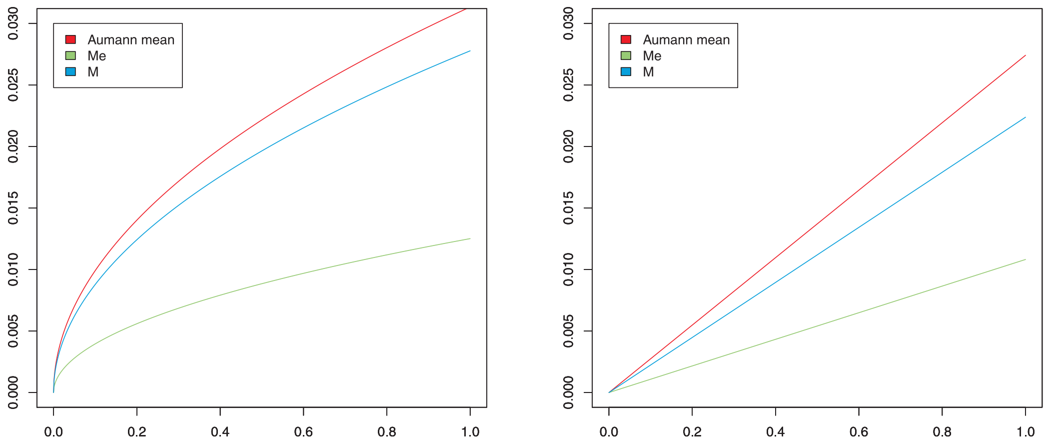


Figure 3.17: D_θ^ℓ-distance (on the left) and the $\mathscr{D}_\theta^\ell$-distance (on the right) between $\gamma_{(0)}(0)$ and the Aumann-type mean, ℓ-wabl/ldev/rdev median (M) and 1-norm median (Me) of the random fuzzy number $\gamma_{(0)} \circ [\text{Bin}(5, 0.5) - 2.5]$ as functions of θ

3.4.2 Consistency and robustness of the sample φ-wabl/ldev/rdev median and comparisons with the sample mean

This section analyzes the $\mathscr{D}_\theta^\varphi$-strong consistency and the finite sample breakdown point of the φ-wabl/ldev/rdev sample median. As for the real-valued case, and the 1-norm median, under rather mild conditions the sample median is shown to be a strongly consistent estimator of the population median, that is,

Theorem 3.4.5. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, φ be an absolutely continuous probability measure on the measurable space $([0,1], \mathcal{B}_{[0,1]})$ with positive mass function on $(0,1)$, and let $\mathcal{X}$ be a random fuzzy number associated with $(\Omega, \mathcal{A}, P)$ and satisfy-*

ing that $\mathrm{Me}(\mathrm{wabl}^{\varphi}(\mathcal{X}))$, $\mathrm{Me}\big(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha)\big)$ *and* $\mathrm{Me}\big(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha)\big)$ *are actually unique (for each* $\alpha \in [0,1]$ *in case of the ones for ldev and rdev).*

If $\widehat{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}_n$ *denotes the sample median corresponding to a simple random sample* $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ *from* $\mathcal{X}$*, and the two sequences of the real-valued sample medians* $\big\{\widehat{\mathrm{Me}(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha))}_n\big\}_n$ *and* $\big\{\widehat{\mathrm{Me}(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha))}_n\big\}_n$ *as functions of* α *over* $[0,1]$ *are both uniformly integrable, then for each* $\theta \in (0,1]$ *the estimator* $\widehat{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}_n$ *is strongly consistent in* $\mathscr{D}^{\varphi}_{\theta}$*-sense (and hence in the sense of all the topologically equivalent metrics), i.e.*

$$\lim_{n\to\infty} \mathscr{D}^{\varphi}_{\theta}\Big(\widehat{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}_n, \widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})\Big) = 0 \quad a.s.\,[P].$$

Proof. Indeed,

$$P\left(\lim_{n\to\infty} \mathscr{D}^{\varphi}_{\theta}\Big(\widehat{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}_n, \widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})\Big) = 0\right)$$

$$= P\Bigg(\lim_{n\to\infty}\Big(|\widehat{\mathrm{Me}(\mathrm{wabl}^{\varphi}(\mathcal{X}))}_n - \mathrm{Me}(\mathrm{wabl}^{\varphi}(\mathcal{X}))|$$

$$+\frac{\theta}{2}\int_{[0,1]} |\widehat{\mathrm{Me}(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha))}_n - \mathrm{Me}(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha))|\, d\varphi(\alpha)$$

$$+\frac{\theta}{2}\int_{[0,1]} |\widehat{\mathrm{Me}(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha))}_n - \mathrm{Me}(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha))|\, d\varphi(\alpha)\Big) = 0\Bigg)$$

$$= P\Bigg(\Big(\lim_{n\to\infty} |\widehat{\mathrm{Me}(\mathrm{wabl}^{\varphi}(\mathcal{X}))}_n - \mathrm{Me}(\mathrm{wabl}^{\varphi}(\mathcal{X}))| = 0\Big)$$

$$\bigcap\Big(\lim_{n\to\infty}\int_{[0,1]} |\widehat{\mathrm{Me}(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha))}_n - \mathrm{Me}(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha))| = 0\Big)$$

$$\bigcap\Big(\lim_{n\to\infty}\int_{[0,1]} |\widehat{\mathrm{Me}(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha))}_n - \mathrm{Me}(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha))| = 0\Big)\Bigg).$$

On one hand,

$$P\Big(\lim_{n\to\infty} |\widehat{\mathrm{Me}(\mathrm{wabl}^{\varphi}(\mathcal{X}))}_n - \mathrm{Me}(\mathrm{wabl}^{\varphi}(\mathcal{X}))| = 0\Big)$$

$$= P\Big(\lim_{n\to\infty}\Big(\widehat{\mathrm{Me}(\mathrm{wabl}^{\varphi}(\mathcal{X}))}_n - \mathrm{Me}(\mathrm{wabl}^{\varphi}(\mathcal{X}))\Big) = 0\Big) = 1,$$

due to the strong consistency of $\widehat{\mathrm{Me}(\mathrm{wabl}^{\varphi}(\mathcal{X}))}_n$.

On the other hand, under the assumption of uniqueness for the medians of $\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha)$ and $\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha)$, the sample medians are strongly consistent estimators of the corresponding population medians, and hence

$$P\Big(\lim_{n\to\infty}\Big(\widehat{\mathrm{Me}(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha))}_n - \mathrm{Me}(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha))\Big) = 0\Big) = 1,$$

$$P\Big(\lim_{n\to\infty}\Big(\widehat{\mathrm{Me}(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha))}_n - \mathrm{Me}(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha))\Big) = 0\Big) = 1.$$

Moreover, assumptions for $\mathrm{Me}(\widehat{\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha)})_n$ and $\mathrm{Me}\big(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha)\big)$ guarantee that conditions to apply Vitali's Convergence Theorem are fulfilled, whence

$$P\left(\left(\lim_{n\to\infty}\int_{[0,1]} \left|\mathrm{Me}(\widehat{\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha)})_n - \mathrm{Me}(\mathrm{ldev}^{\varphi}_{\mathcal{X}}(\alpha))\right| d\varphi(\alpha) = 0\right)\right) = 1.$$

By following similar arguments, one can prove that

$$P\left(\left(\lim_{n\to\infty}\int_{[0,1]} \left|\mathrm{Me}(\widehat{\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha)})_n - \mathrm{Me}(\mathrm{rdev}^{\varphi}_{\mathcal{X}}(\alpha))\right| d\varphi(\alpha) = 0\right)\right) = 1.$$

Consequently,

$$P\left(\lim_{n\to\infty} \mathscr{D}^{\varphi}_{\theta}\left(\widehat{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}_n, \widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})\right) = 0\right). \qquad \square$$

The comparative robustness of the sample ℓ-wabl/ldev/rdev median of a random fuzzy number as an estimator of the population median, in contrast to that of the sample mean and the 1-norm median of a random fuzzy number as an estimator of the population mean and 1-norm median, respectively, is now to be discussed.

Before presenting a formal discussion and comparison, the first simulations in Section 1.5 (p. 54) when the mean is replaced by the ℓ-wabl/ldev/rdev median are to be analyzed. To determine the effect of the contamination on the median of the random fuzzy number $\mathcal{X}$, the mean distance between the non-contaminated 'distribution' and the Monte Carlo approximated ℓ-wabl/ldev/rdev median is collected in Table 3.5 for the different values of c_p and C_D and CASES 1 and 2. The results in Table 3.5 show that, as for the 1-norm median, the expected distance between the non-contaminated distribution and the sample ℓ-wabl/ldev/rdev median only slightly changes when the amount of contamination is increased, even when the contamination lies far from the non-contaminated distribution.

Actually, it seems even slightly more robust than the 1-norm median (see columns for $\boldsymbol{\rho_2}$ in Tables 3.3 and 3.5).

The analysis of the robustness of the φ-wabl/ldev/rdev median in comparison to the mean is now made through the finite sample breakdown point of the sample median in a sample of size n from a random fuzzy number $\mathcal{X}$, which is now given by

$$\begin{aligned}&\mathrm{fsbp}(\widehat{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}_n, \widetilde{\mathbf{x}}_n, \mathscr{D}^{\varphi}_{\theta})\\ &= \frac{1}{n}\min\left\{k\in\{1,\ldots,n\} \,:\, \sup_{Q_{n,k}} \mathscr{D}^{\varphi}_{\theta}(\widehat{\widetilde{\mathrm{M}}^{\varphi}(P_n)}, \widehat{\widetilde{\mathrm{M}}^{\varphi}(Q_{n,k})}) = \infty\right\},\end{aligned}$$

where $\widetilde{\mathbf{x}}_n$ denotes the considered sample of n data from the metric space $(\mathcal{F}_c(\mathbb{R}), \mathscr{D}^{\varphi}_{\theta})$ in which $\sup_{\widetilde{U},\widetilde{V}\in\mathcal{F}_c(\mathbb{R})} \mathscr{D}^{\varphi}_{\theta}(\widetilde{U},\widetilde{V}) = \infty$, P_n is the empirical distribution of $\widetilde{\mathbf{x}}_n$ and $Q_{n,k}$ is the empirical distribution of sample $\widetilde{\mathbf{y}}_{n,k}$ obtained from the original one $\widetilde{\mathbf{x}}_n$ by perturbing at most k components. Then, we have that

		CASE 1		CASE 2	
c_P	c_D	$\mathscr{D}^{\ell}_{1/3}$	$\boldsymbol{\rho_2}$	$\mathscr{D}^{\ell}_{1/3}$	$\boldsymbol{\rho_2}$
0	0	1.227657	1.554372	0.9504636	1.0092032
0.1	0	1.229411	1.553082	0.9506150	1.0128246
0.1	1	1.229757	1.553262	0.9519204	0.9949145
0.1	5	1.231567	1.554583	0.9537212	0.9960642
0.1	10	1.232087	1.554976	0.9543808	0.9967514
0.1	100	1.232150	1.555065	0.9544468	0.9967801
0.2	0	1.236350	1.563680	0.950800	1.016413
0.2	1	1.238210	1.565674	0.957562	0.986686
0.2	5	1.247372	1.572603	0.966244	0.992840
0.2	10	1.250346	1.574932	0.968865	0.994872
0.2	100	1.250324	1.575027	0.968361	0.993515
0.4	0	1.300707	1.711233	0.9519090	1.0267570
0.4	1	1.317032	1.742423	0.9903528	1.0154772
0.4	5	1.383328	1.802162	1.0436144	1.0676748
0.4	10	1.411993	1.817192	1.0674811	1.0911364
0.4	100	1.412609	1.829062	1.0671669	1.0918211

Table 3.5: Mean distances of the mixed (partially contaminated and non-contaminated) sample ℓ-wabl/ldev/rdev median to the non-contaminated distribution of a random fuzzy number

Proposition 3.4.6. *The finite sample breakdown point of the sample φ-wabl/ldev/rdev median from a random fuzzy number $\mathcal{X}$,* $\mathrm{fsbp}\big(\widehat{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})_n}\big)$*, equals*

$$\mathrm{fsbp}(\widehat{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})_n}, \widetilde{\mathbf{x}}_n, \mathscr{D}^{\varphi}_{\theta}) = \frac{1}{n} \cdot \lfloor \frac{n+1}{2} \rfloor.$$

Proof. First note that the condition $\sup_{\widetilde{U},\widetilde{V}\in\mathcal{F}_c(\mathbb{R})} \mathscr{D}^{\varphi}_{\theta}(\widetilde{U},\widetilde{V}) = \infty$ is satisfied in this case, since $\mathscr{D}^{\varphi}_{\theta}\big(\mathbb{1}_{[n-1,n+1]}, \mathbb{1}_{[-n-1,-n+1]}\big) = 2n$ (of course, other examples could be provided for the same purpose).

Furthermore,

$$\mathscr{D}^{\varphi}_{\theta}(\widehat{\widetilde{\mathrm{M}}^{\varphi}(P_n)}, \widehat{\widetilde{\mathrm{M}}^{\varphi}(Q_{n,k})}) \geq |\mathrm{wabl}^{\varphi}(\widehat{\widetilde{\mathrm{M}}^{\varphi}(P_n)}) - \mathrm{wabl}^{\varphi}(\widehat{\widetilde{\mathrm{M}}^{\varphi}(Q_{n,k})})_{\alpha})|$$

$$= |\mathrm{Me}(\widehat{\mathrm{wabl}^{\varphi}(P_n)}) - \mathrm{Me}(\widehat{\mathrm{wabl}^{\varphi}(Q_{n,k})})|.$$

Therefore, by recalling the fsbp for the sample median of a real-valued random variable, one can conclude that whenever at least $\lfloor \frac{n+1}{2} \rfloor$ elements $\widetilde{x}_i \in \mathcal{F}_c(\mathbb{R})$ of $\widetilde{\mathbf{x}}_n$ are replaced by other arbitrarily 'large' elements in $\mathcal{F}_c(\mathbb{R})$ so that

$$\sup_{Q_{n,k}} |\mathrm{Me}(\widehat{\mathrm{wabl}^{\varphi}(P_n)}) - \mathrm{Me}(\widehat{\mathrm{wabl}^{\varphi}(Q_{n,k})})| = \infty,$$

we have that

$$\sup_{Q_{n,k}} \mathscr{D}_\theta^\varphi(\widehat{\widetilde{\mathrm{M}}^\varphi(P_n)}, \widehat{\widetilde{\mathrm{M}}^\varphi(Q_{n,k})}) \geq \sup_{Q_{n,k}} |\mathrm{Me}(\widehat{\mathrm{wabl}^\varphi}(P_n)) - \mathrm{Me}(\widehat{\mathrm{wabl}^\varphi(Q_{n,k})})| = \infty,$$

whence

$$\mathrm{fsbp}(\widehat{\widetilde{\mathrm{M}}^\varphi(\mathcal{X})}_n, \widetilde{\mathbf{x}}_n, \mathscr{D}_\theta^\varphi) \leq \frac{1}{n} \cdot \lfloor \frac{n+1}{2} \rfloor.$$

On the other hand, by using the fsbp for the sample median of a real-valued random variable, we have that for all $\alpha \in [0,1]$

$$\min \left\{ k \in \{1,\ldots,n\} \,:\, \sup_{Q_{n,k}} |\mathrm{Me}(\widehat{\mathrm{wabl}^\varphi}(P_n)) - \mathrm{Me}(\widehat{\mathrm{wabl}^\varphi(Q_{n,k})})| = \infty \right\} = \lfloor \frac{n+1}{2} \rfloor,$$

$$\min \left\{ k \in \{1,\ldots,n\} \,:\, \sup_{Q_{n,k}} |\mathrm{Me}\big(\widehat{\mathrm{ldev}_{P_n}^\varphi}(\alpha)\big) - \mathrm{Me}\big(\widehat{\mathrm{ldev}_{Q_{n,k}}^\varphi}(\alpha)\big)| = \infty \right\} = \lfloor \frac{n+1}{2} \rfloor,$$

$$\min \left\{ k \in \{1,\ldots,n\} \,:\, \sup_{Q_{n,k}} |\mathrm{Me}\big(\widehat{\mathrm{rdev}_{P_n}^\varphi}(\alpha)\big) - \mathrm{Me}\big(\widehat{\mathrm{rdev}_{Q_{n,k}}^\varphi}(\alpha)\big)| = \infty \right\} = \lfloor \frac{n+1}{2} \rfloor,$$

whence for all $\alpha \in [0,1]$

$$\sup_{Q_{n,\lfloor \frac{n+1}{2} \rfloor - 1}} |\mathrm{Me}(\widehat{\mathrm{wabl}^\varphi}(P_n)) - \mathrm{Me}(\widehat{\mathrm{wabl}^\varphi(Q_{n,k})})| = M_1 < \infty,$$

$$\sup_{Q_{n,\lfloor \frac{n+1}{2} \rfloor - 1}} |\mathrm{Me}\big(\widehat{\mathrm{ldev}_{P_n}^\varphi}(\alpha)\big) - \mathrm{Me}\big(\widehat{\mathrm{ldev}_{Q_{n,k}}^\varphi}(\alpha)\big)| = M_2 < \infty,$$

$$\sup_{Q_{n,\lfloor \frac{n+1}{2} \rfloor - 1}} |\mathrm{Me}\big(\widehat{\mathrm{rdev}_{P_n}^\varphi}(\alpha)\big) - \mathrm{Me}\big(\widehat{\mathrm{rdev}_{Q_{n,k}}^\varphi}(\alpha)\big)| = M_3 < \infty,$$

and therefore

$$\begin{aligned}
&\sup_{Q_{n,\lfloor \frac{n+1}{2} \rfloor - 1}} \mathscr{D}_\theta^\varphi(\widehat{\widetilde{\mathrm{M}}^\varphi(P_n)}, \widehat{\widetilde{\mathrm{M}}^\varphi(Q_{n,\lfloor \frac{n+1}{2} \rfloor - 1})}) \\
&= \sup_{Q_{n,\lfloor \frac{n+1}{2} \rfloor - 1}} \Bigg[|\mathrm{Me}(\widehat{\mathrm{wabl}^\varphi}(P_n)) - \mathrm{Me}(\widehat{\mathrm{wabl}^\varphi(Q_{n,k})})| \\
&+ \frac{\theta}{2} \int_{[0,1]} |\mathrm{Me}\big(\widehat{\mathrm{ldev}_{P_n}^\varphi}(\alpha)\big) - \mathrm{Me}\big(\widehat{\mathrm{ldev}_{Q_{n,\lfloor \frac{n+1}{2} \rfloor - 1}}^\varphi}(\alpha)\big)| \, d\varphi(\alpha) \\
&+ \frac{\theta}{2} \int_{[0,1]} |\mathrm{Me}\big(\widehat{\mathrm{ldev}_{P_n}^\varphi}(\alpha)\big) - \mathrm{Me}\big(\widehat{\mathrm{ldev}_{Q_{n,\lfloor \frac{n+1}{2} \rfloor - 1}}^\varphi}(\alpha)\big)| \, d\varphi(\alpha) \Bigg] \\
&\leq M_1 + \frac{M_2 + M_3}{2}\,\theta < \infty.
\end{aligned}$$

Consequently,

$$\min \left\{ k \in \{1,\ldots,n\} \,:\, \sup_{Q_{n,k}} \mathscr{D}_\theta^\varphi(\widehat{\widetilde{\mathrm{M}}^\varphi(P_n)}, \widehat{\widetilde{\mathrm{M}}^\varphi(Q_{n,k})}) = \infty \right\} > \lfloor \frac{n+1}{2} \rfloor - 1,$$

whence

$$\mathrm{fsbp}(\widehat{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}_n, \widetilde{\mathbf{x}}_n, \mathscr{D}_\theta^\varphi) \geq \frac{1}{n} \cdot \lfloor \frac{n+1}{2} \rfloor. \qquad \square$$

The following result formalizes the comparison of the robustness of the sample φ-wabl/ldev/rdev median and the sample mean and 1-norm median of a random fuzzy number. Thus,

Theorem 3.4.7. *The finite sample breakdown point of the sample mean from a random fuzzy number $\mathcal{X}$, $\mathrm{fsbp}\big(\overline{\mathcal{X}_n}\big)$, is lower than that for the sample (either φ-wabl/ldev/rdev or 1-norm) median for sample sizes $n > 2$.*

Proof. Indeed, by arguing like for the preceding proposition we have that

$$\mathrm{fsbp}(\overline{\mathcal{X}_n}, \widetilde{\mathbf{x}}_n, \mathscr{D}_\theta^\varphi) = \mathrm{fsbp}(\overline{\mathcal{X}_n}, \widetilde{\mathbf{x}}_n, \rho_1) = \frac{1}{n},$$

and, consequently,

$$\mathrm{fsbp}(\widehat{\widetilde{\mathrm{M}}^{\varphi}(\mathcal{X})}_n, \widetilde{\mathbf{x}}_n, \mathscr{D}_\theta^\varphi) = \mathrm{fsbp}(\widehat{\widetilde{\mathrm{Me}}(\mathcal{X})}_n, \widetilde{\mathbf{x}}_n, \boldsymbol{\rho_1})$$

$$\geq \frac{n/2}{n} = \frac{1}{2} > \frac{1}{n} = \mathrm{fsbp}(\overline{\mathcal{X}_n}, \widetilde{\mathbf{x}}_n, \boldsymbol{\rho_1}) = \mathrm{fsbp}(\overline{\mathcal{X}_n}, \widetilde{\mathbf{x}}_n, \mathscr{D}_\theta^\varphi). \qquad \square$$

The theoretical conclusion in Theorem 3.4.7 can be corroborated empirically by analyzing the simulations in Section 1.5 and those in Table 3.5. Moreover, and on the basis of these simulations an additional table has been constructed.

Table 3.6 gathers empirical results for the influence of contamination on both the sample mean and median, by computing the distances between the mean/median of the non-contaminated sample and the mean/ℓ-wabl/ldev/rdev median of the contaminated sample, respectively, for the different values of c_p and C_D and the CASES 1 and 2 in Tables 1.2, 3.3 and 3.5.

On the basis of these simulations and by comparing the results in Tables 1.2, 3.3 and 3.5, and the results in Tables 3.3 and 3.5, one can empirically conclude that

- for a fixed level of contamination c_P, the farther the contaminated distribution from the non-contaminated one the substantially greater mean $\boldsymbol{\rho_2}$-distance between the approximated mean and the non-contaminated distribution, whereas for both the approximated 1-norm and ℓ-wabl/ldev/redv medians the increase is modest; actually this mean distance asymptotically would only depend on a certain fractile of the non-contaminated distribution;

		CASE 1				CASE 2			
		mean		ℓ-w/l/r median		mean		ℓ-w/l/r median	
c_P	c_D	$\mathscr{D}^{\ell}_{1/3}$	$\boldsymbol{\rho_2}$	$\mathscr{D}^{\ell}_{1/3}$	$\boldsymbol{\rho_2}$	$\mathscr{D}^{\ell}_{1/3}$	$\boldsymbol{\rho_2}$	$\mathscr{D}^{\ell}_{1/3}$	$\boldsymbol{\rho_2}$
0.0	0	0.004141	0.005446	0.004736	0.005744	0.002935	0.003035	0.004041	0.004609
0.1	0	0.115821	0.340882	0.041959	0.117079	0.006640	0.012257	0.008726	0.017791
0.1	1	0.200101	0.457775	0.057713	0.122703	0.083433	0.120268	0.045510	0.089681
0.1	5	0.549571	0.965272	0.105273	0.144576	0.429103	0.641559	0.088103	0.113566
0.1	10	0.985210	1.614466	0.113848	0.149298	0.863819	1.306176	0.094945	0.117746
0.1	100	8.801379	13.479923	0.112416	0.150624	8.714803	13.156037	0.095313	0.117737
0.2	0	0.224863	0.680722	0.091140	0.266010	0.010471	0.023612	0.014003	0.033147
0.2	1	0.402916	0.924358	0.132474	0.287719	0.168120	0.241804	0.099201	0.197792
0.2	5	1.095563	1.952914	0.234256	0.333697	0.867697	1.283131	0.194468	0.249352
0.2	10	1.972511	3.261031	0.253764	0.350314	1.740925	2.579237	0.210592	0.261186
0.2	100	17.620494	26.793580	0.253381	0.347141	17.588028	26.292482	0.209108	0.262428
0.4	0	0.469059	1.364644	0.275940	0.829457	0.019153	0.046735	0.026856	0.071649
0.4	1	0.797371	1.874279	0.381792	0.890780	0.337019	0.485976	0.231633	0.472210
0.4	5	2.209684	3.860446	0.655998	1.055769	1.720116	2.551476	0.462176	0.611100
0.4	10	3.973298	6.470993	0.726146	1.085354	3.497305	5.133452	0.525197	0.651425
0.4	100	35.370045	54.170428	0.722964	1.114021	35.107081	51.856741	0.521087	0.653929

Table 3.6: Distances between the sample mixed (partially contaminated and non-contaminated) mean/ℓ-wabl/ldev/rdev median to the non-contaminated one for a random fuzzy number

- for a fixed level of contamination c_P, the farther the contaminated distribution from the non-contaminated one, the substantially greater distance between the contaminated and the non-contaminated means, whereas for the 1-norm and the ℓ-wabl/ldev/redv medians the increase is not really substantial;
- by simply comparing the two L^1 medians, the behaviour of the ℓ-wabl/ldev/redv median is slightly more stable than that of the 1-norm median in CASE 1, and the opposite situation arises in CASE 2.

To conclude this section, three remarks should be made. The first and the second ones concern the particularization of the φ-wabl/ldev/rdev median to the interval-valued case and its extension to the fuzzy vector-valued case, respectively.

Remark 3.4.4. All the notions and results in this section can be trivially particularized to the interval-valued case, so that the φ-wabl/ldev/rdev median of a random interval X is given by the interval value $[\mathrm{Me}(\mathrm{mid}\,\mathsf{X})-\mathrm{Me}(\mathrm{spr}\,\mathsf{X}), \mathrm{mid}\,\mathsf{X})+\mathrm{Me}(\mathrm{spr}\,\mathsf{X})]$.

Remark 3.4.5. The extension of the φ-wabl/ldev/rdev median to the fuzzy vector-valued case through the use of an L^1 metric based on the support/Steiner representation of fuzzy vectors cannot be made by extending arguments in Theorem 3.4.1, because of the reasons argued for the extension of the 1-norm median.

Remark 3.4.6. It should be pointed out that, as for the 1-norm median, one of the main advantages of the φ-wabl/ldev/rdev median is that they can be computed on the basis of the medians for certain real-valued random variables. This makes computations rather easy-to-perform and, mainly, easy to implementing and programming in R or others.

At this point, we should indicate that when the involved L^1 metric is replaced by other ones, the minimization problem can become a very difficult task, and often infeasible at least to get the exact solution.

In this respect, if we consider the L^1 metric extending Hausdorff one from $\mathcal{K}_c(\mathbb{R})$ to $\mathcal{F}_c(\mathbb{R})$, and given (Klement *et al.* [116]) by

$$\boldsymbol{d_1}(\widetilde{U},\widetilde{V}) = \int_{[0,1]} \left(\left|\operatorname{mid} \widetilde{U}_\alpha - \operatorname{mid} \widetilde{V}_\alpha\right| + \left|\operatorname{spr} \widetilde{U}_\alpha - \operatorname{spr} \widetilde{V}_\alpha\right| \right) \, d\ell(\alpha)$$

one cannot reason as for $\boldsymbol{\rho_1}$ since there is not a set of sufficient conditions for the mid/spr representation characterizing fuzzy numbers.

More concretely, and arguing as in Remark 3.3.8, if following the solutions for $\boldsymbol{\rho_1}$ one is tempted to use as a possible solution minimizing $E\left(\left[\boldsymbol{d_1}(\mathcal{X},\widetilde{U})\right]\right)$ over $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$, the level-wise solution in Remark 3.4.4

$$\mathrm{M}_\alpha = [\mathrm{Me}(\operatorname{mid} \mathcal{X}_\alpha) - \mathrm{Me}(\operatorname{spr} \mathcal{X}_\alpha), \mathrm{Me}(\operatorname{mid} \mathcal{X}_\alpha) + \mathrm{Me}(\operatorname{spr} \mathcal{X}_\alpha)]$$

for each $\alpha \in [0,1]$, the class $\{\mathrm{M}_\alpha\}_\alpha$ does not define in general a fuzzy number.

As a counterexample illustrating the assertions in Remark 3.4.6, we can consider the following:

Example 3.4.2. Consider a random fuzzy number $\mathcal{X}$ taking with probability 0.2 each of five different values $\widetilde{x}_1$ ($i \in \{1,\ldots,5\}$ which, in accordance with their vertical view, are given by

$$(\operatorname{mid} \widetilde{x}_1)_\alpha = 1 - \alpha/2, \quad (\operatorname{spr} \widetilde{x}_1)_\alpha = 1.1 - \alpha,$$

$$(\operatorname{mid} \widetilde{x}_2)_\alpha = \begin{cases} 0.75 & \text{if } \alpha \leq 0.5 \\ 1.25 - \alpha & \text{otherwise} \end{cases}, \quad (\operatorname{spr} \widetilde{x}_2)_\alpha = 1.1 - \alpha,$$

$$(\operatorname{mid} \widetilde{x}_3)_\alpha = 0.6 + 0.3\,\alpha, (\operatorname{mid} \widetilde{x}_4)_\alpha = 0, \ (\operatorname{mid} \widetilde{x}_5)_\alpha = 2,$$

$$(\operatorname{spr} \widetilde{x}_3)_\alpha = (\operatorname{spr} \widetilde{x}_4)_\alpha = (\operatorname{spr} \widetilde{x}_5)_\alpha = \begin{cases} 0.75 - \alpha & \text{if } \alpha \leq 0.5 \\ 0.4 - 0.3\,\alpha & \text{otherwise,} \end{cases}$$

which are graphically displayed in Figure 3.18

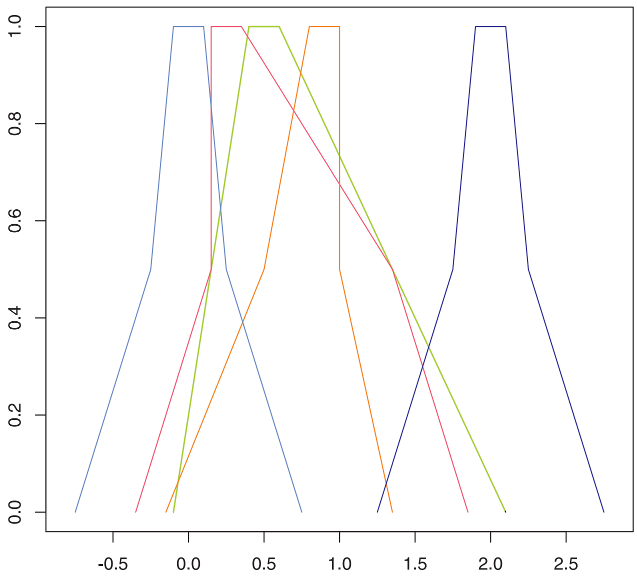


Figure 3.18: Five different values of a random set (that takes them with the same probability)

The value for the mean, the 1-norm median and the ℓ-wabl/ldev/rdev median can be found graphically displayed in Figure 3.19

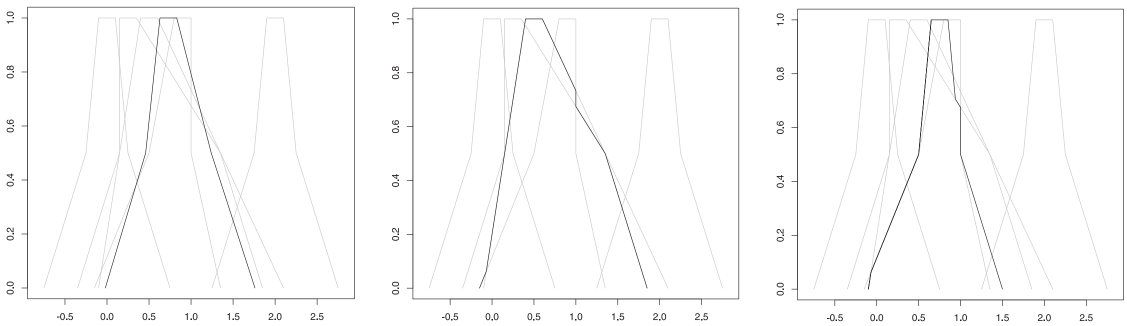


Figure 3.19: Mean, 1-norm median and ℓ-wabl/ldev/rdev median of the random fuzzy number being uniformly distributed on the set of fuzzy number values in Figure 3.18

In this case we have that

$$\mathrm{Me}(\mathrm{mid}\,\mathcal{X}_\alpha) = \begin{cases} 0.75 & \text{if } \alpha \leq 0.5 \\ 1 - \alpha/2 & \text{otherwise,} \end{cases} \quad \mathrm{Me}(\mathrm{spr}\,\mathcal{X}_\alpha) = \begin{cases} 0.75 - \alpha & \text{if } \alpha \leq 0.5 \\ 0.4 - 0.3\,\alpha & \text{otherwise,} \end{cases}$$

whence the intervals $[\mathrm{Me}(\mathrm{mid}\,\mathcal{X}_\alpha) - \mathrm{Me}(\mathrm{spr}\,\mathcal{X}_\alpha), \mathrm{Me}(\mathrm{mid},\mathcal{X}_\alpha) + \mathrm{Me}(\mathrm{spr}\,\mathcal{X}_\alpha)]$ do not lead to a fuzzy number, but to the function in Figure 3.20.

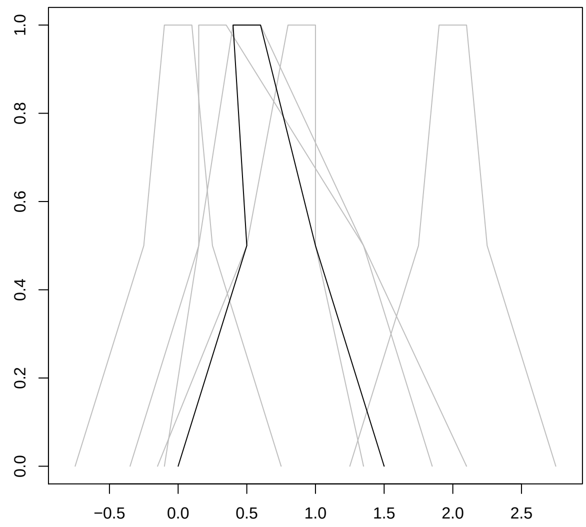


Figure 3.20: Result of representing the fuzzy number (?) with α-levels given by $[\text{Me}(\text{mid}\,\mathcal{X}_\alpha) - \text{Me}(\text{spr}\,\mathcal{X}_\alpha), \text{Me}(\text{mid},\mathcal{X}_\alpha) + \text{Me}(\text{spr}\,\mathcal{X}_\alpha)]$, which is not a fuzzy number

3.5 The spatial median for a random interval

As for the last two sections, the sufficient conditions over the loss function in Section 3.1 allowing us to guarantee the existence of sample M-estimates of location, and their expression as convex linear combinations of the sample elements, are not fulfilled by some interesting choices of ρ, like $\rho(x) = \sqrt{|x|}$. Consequently, one should face this problem by developing *ad hoc* methods.

In addition to the use of L^1 metrics, as those in Sections 3.3 and 3.4, it should be noted that a well-known generalization of the median of real-valued random variables to the multivariate settings is the spatial median or mediancentre (see, for example, Gower [93] or Milasevic and Ducharme [132]), which is based on an L^2-type metric. To develop an appropriate extension of the notion of spatial median to fuzzy number-valued random elements becomes a very hard task. So, for a simpler approximation to the problem, the interval-valued case is first to be examined. As it will be shown, the first empirical conclusions based on synthetic examples lead us to conclude that concerning robustness there is no evidence of major and general advantages of the spatial median versus the particularized 1-norm and wabl/ldev/rdev justifying the effort associated with the extension to the fuzzy number-valued case. Furthermore, the computation of the spatial median is mich more complex than that of the medians in Sections 3.3 and 3.4, which is immediate to perform because of being based in the already implemented tools for the real-valued case.

To adapt the ideas behind the spatial median approach to random intervals, we are going to consider the generalized L^2 metric d_θ introduced by Bertoluzza *et al.* [16] (and expressed by Gil *et al.* [84] in terms of the mid/spr representation of interval values), which is the mid/spr-based L^2 metric between interval values (see p. 22) and also coincides with the support/Steiner-based L^2 metric, $\mathfrak{d}_\theta$, between them (see p. 25). This is a very wide, intuitive, and valuable family of distances for interval values, and the well-known Vitale's L^2 metric (see Vitale [211]), δ_2, on $\mathcal{K}_c(\mathbb{R})$ is a particular element of this family. It also shares an interesting feature with Hausdorff metric on $\mathcal{K}_c(\mathbb{R})$: it weights the squared distances between mids and those between the spreads.

Inspired by the spatial median as extension of the median to higher dimensional Euclidean spaces and even Banach spaces (see Cadre [23]) using an L^2-type metric on $\mathcal{K}_c(\mathbb{R})$ the population and sample d_θ-medians are to be introduced.

Definition 3.5.1. *Given a probability space* $(\Omega, \mathcal{A}, P)$ *and an associated random interval* $\mathsf{X} : \Omega \to \mathcal{K}_c(\mathbb{R})$, *the* ***population*** d_θ-***median(s)*** *of* X *is the interval value* $\mathsf{M}_\theta[\mathsf{X}] \in \mathcal{K}_c(\mathbb{R})$ *such that*

$$\mathsf{M}_\theta[\mathsf{X}] = \arg \min_{K \in \mathcal{K}_c(\mathbb{R})} E\left(d_\theta(\mathsf{X}, K)\right),$$

whenever the involved expectations exist.

Definition 3.5.2. *Given a probability space* $(\Omega, \mathcal{A}, P)$, *an associated random interval* $\mathsf{X} : \Omega \to \mathcal{K}_c(\mathbb{R})$, *and a simple random sample* $(\mathsf{X}_1, \ldots, \mathsf{X}_n)$ *from* X, *the* ***sample*** d_θ-***median(s)*** *of* X *is(are) the interval-valued statistic(s)*

$$\begin{aligned}\widehat{\mathsf{M}_\theta[\mathsf{X}]}_n &= \arg \min_{K \in \mathcal{K}_c(\mathbb{R})} \frac{1}{n} \sum_{i=1}^{n} \left(d_\theta(\mathsf{X}_i, K)\right) \\ &= \arg \min_{(y,z) \in \mathbb{R} \times [0,\infty)} \frac{1}{n} \sum_{i=1}^{n} \sqrt{(\operatorname{mid} \mathsf{X}_i - y)^2 + \theta \cdot (\operatorname{spr} \mathsf{X}_i - z)^2},\end{aligned}$$

where K, y *and* z *depend in fact on* $(\mathsf{X}_1, \ldots, \mathsf{X}_n)$ *(although, for the sake of simplicity, this has been omitted from the notation).*

The use of the d_θ metric in formalizing the median of a random interval allows us to guarantee its existence and to state conditions for its uniqueness. However, one cannot provide in general with an explicit expression for the d_θ-median in a way similar to what has been established for the M-estimates associated with loss function satisfying sufficient conditions in Subsection 3.1.2 and the 1-norm and wabl/ldev/rdev medians in Theorems 3.3.1 and 3.4.1.

The convexity of the objective function to be minimized in Definition 3.5.2 is now to be proven. Thus,

Proposition 3.5.1. *Whatever the sample of independent interval-valued observations* $\mathbf{x_n} = (\mathsf{x}_1, \dots, \mathsf{x}_n)$ *from the random interval* X *may be, the function*

$$\begin{array}{rcl} f: \ \mathbb{D} = \mathbb{R} \times [0, \infty) & \longrightarrow & \mathbb{R} \\ (y, z) & \longmapsto & \displaystyle\frac{1}{n} \sum_{i=1}^{n} \sqrt{(\operatorname{mid} \mathsf{x}_i - y)^2 + \theta \cdot (\operatorname{spr} \mathsf{x}_i - z)^2}, \end{array}$$

is strictly convex unless all the sample points $\{(\operatorname{mid} \mathsf{x}_i, \operatorname{spr} \mathsf{x}_i)\}_{i=1}^n$ *are collinear (in such a situation, the function* f *is convex).*

Proof. Indeed, the dominium $\mathbb{D}$ is convex and given any two different elements $(y_1, z_1), (y_2, z_2) \in \mathbb{D}$ and $\lambda \in (0, 1)$:

$$f\big(\lambda(y_1, z_1) + (1 - \lambda)(y_2, z_2)\big) = f\big(\lambda y_1 + (1 - \lambda) y_2, \lambda z_1 + (1 - \lambda) z_2\big)$$

$$= \frac{1}{n} \sum_{i=1}^{n} \sqrt{\big(\operatorname{mid} \mathsf{x}_i - \lambda y_1 - (1 - \lambda) y_2\big)^2 + \theta\big(\operatorname{spr} \mathsf{x}_i - \lambda z_1 - (1 - \lambda) z_2\big)^2}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \Big[\big(\lambda \cdot \operatorname{mid} \mathsf{x}_i - \lambda y_1 + (1 - \lambda) \operatorname{mid} \mathsf{x}_i - (1 - \lambda) y_2\big)^2$$

$$+ \theta\big(\lambda \cdot \operatorname{spr} \mathsf{x}_i - \lambda z_1 + (1 - \lambda) \operatorname{spr} \mathsf{x}_i - (1 - \lambda) z_2\big)^2\Big]^{1/2}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \Big[\lambda^2\big((\operatorname{mid} \mathsf{x}_i - y_1)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_1)^2\big) + (1 - \lambda)^2\big((\operatorname{mid} \mathsf{x}_i - y_2)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_2)^2\big)$$

$$+ 2\lambda(1 - \lambda)\big((\operatorname{mid} \mathsf{x}_i - y_1)(\operatorname{mid} \mathsf{x}_i - y_2) + \theta(\operatorname{spr} \mathsf{x}_i - z_1)(\operatorname{spr} \mathsf{x}_i - z_2)\big)\Big]^{1/2}.$$

On the other hand,

$$\lambda \cdot f(y_1, z_1) + (1 - \lambda) f(y_2, z_2)$$

$$= \frac{1}{n} \sum_{i=1}^{n} \sqrt{\lambda^2\big((\operatorname{mid} \mathsf{x}_i - y_1)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_1)^2\big)}$$

$$+ \frac{1}{n} \sum_{i=1}^{n} \sqrt{(1 - \lambda)^2\big((\operatorname{mid} \mathsf{x}_i - y_2)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_2)^2\big)}.$$

Since not all the sample points are collinear, then

$$\big((\operatorname{mid} \mathsf{x}_i - y_1)(\operatorname{spr} \mathsf{x}_i - z_2) - (\operatorname{spr} \mathsf{x}_i - z_1)(\operatorname{mid} \mathsf{x}_i - y_2)\big)^2 > 0$$

for at least a subindex $i \in \{1, \dots, n\}$. Consequently,

$$2\theta(\operatorname{mid} \mathsf{x}_i - y_1)(\operatorname{mid} \mathsf{x}_i - y_2)(\operatorname{spr} \mathsf{x}_i - z_1)(\operatorname{spr} \mathsf{x}_i - z_2)$$

$$< \theta(\operatorname{mid} \mathsf{x}_i - y_1)^2(\operatorname{spr} \mathsf{x}_i - z_2)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_1)^2(\operatorname{mid} \mathsf{x}_i - y_2)^2.$$

By adding $(\operatorname{mid} \mathsf{x}_i - y_1)^2(\operatorname{mid} \mathsf{x}_i - y_2)^2 + \theta^2(\operatorname{spr} \mathsf{x}_i - z_1)^2(\operatorname{spr} \mathsf{x}_i - z_2)^2$ to both sides of the inequation and taking square roots we have that

$$\begin{aligned}&(\operatorname{mid} \mathsf{x}_i - y_1)(\operatorname{mid} \mathsf{x}_i - y_2) + \theta(\operatorname{spr} \mathsf{x}_i - z_1)(\operatorname{spr} \mathsf{x}_i - z_2)\\ &< \Big[(\operatorname{mid} \mathsf{x}_i - y_1)^2(\operatorname{mid} \mathsf{x}_i - y_2)^2 + \theta^2(\operatorname{spr} \mathsf{x}_i - z_1)^2(\operatorname{spr} \mathsf{x}_i - z_2)^2\\ &+ \theta(\operatorname{mid} \mathsf{x}_i - y_1)^2(\operatorname{spr} \mathsf{x}_i - z_2)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_1)^2(\operatorname{mid} \mathsf{x}_i - y_2)^2\Big]^{1/2},\end{aligned}$$

whence for $\lambda \in (0, 1)$

$$\begin{aligned}&\lambda(1-\lambda)((\operatorname{mid} \mathsf{x}_i - y_1)(\operatorname{mid} \mathsf{x}_i - y_2) + \theta(\operatorname{spr} \mathsf{x}_i - z_1)(\operatorname{spr} \mathsf{x}_i - z_2))\\ &< \sqrt{\lambda^2[(\operatorname{mid} \mathsf{x}_i - y_1)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_1)^2] \cdot (1-\lambda)^2[(\operatorname{mid} \mathsf{x}_i - y_2)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_2)^2]}.\end{aligned}$$

After adding $\lambda^2[(\operatorname{mid} \mathsf{x}_i - y_1)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_1)^2] + (1-\lambda)^2[(\operatorname{mid} \mathsf{x}_i - y_2)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_2)^2]$ to both sides of the last inequation and by taking square roots, one can conclude that

$$f\big(\lambda(y_1, z_1) + (1-\lambda)(y_2, z_2)\big) < \lambda \cdot f(y_1, z_1) + (1-\lambda) f(y_2, z_2).$$

Of course, if all the sample points are collinear, all the previous strict inequalities reduce simply to inequalities. □

The existence of the sample d_θ-median is now to be discussed.

Theorem 3.5.2. *Given a simple random sample* $(\mathsf{X}_1, \ldots, \mathsf{X}_n)$ *from a random interval* $\mathsf{X} : \Omega \to \mathcal{K}_c(\mathbb{R})$*, the corresponding sample* d_θ*-median always exists. Moreover, the sample* d_θ*-median is unique for any sample realization* $\mathbf{x}_n = (\mathsf{x}_1, \ldots, \mathsf{x}_n)$ *for which the two-dimensional sample points* $\{(\operatorname{mid} \mathsf{x}_i, \operatorname{spr} \mathsf{x}_i)\}_{i=1}^n$ *are not all collinear.*

Proof. First of all, consider the particular case in which all the sample interval-valued data x_i are real numbers. Then, $\operatorname{spr} \mathsf{x}_i = 0$ for all $i \in \{1, \ldots, n\}$. Should this be the case, the aim would be to find the solution of the minimization problem

$$\min_{(y,z) \in \mathbb{R} \times [0,\infty)} \frac{1}{n} \sum_{i=1}^n \sqrt{(\operatorname{mid} \mathsf{x}_i - y)^2 + \theta \cdot z^2}.$$

By taking into account that

$$\begin{aligned}&\min_{(y,z) \in \mathbb{R} \times [0,\infty)} \frac{1}{n} \sum_{i=1}^n \sqrt{(\operatorname{mid} \mathsf{x}_i - y)^2 + \theta \cdot z^2}\\ &\geq \min_{y \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^n \sqrt{(\operatorname{mid} \mathsf{x}_i - y)^2} = \min_{y \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^n |\operatorname{mid} \mathsf{x}_i - y|,\end{aligned}$$

the sample d_θ-median is the interval with mid-point $y = \mathrm{Me}(\{\mathrm{mid}\, \mathsf{x}_i\}_{i=1}^n)$ and spread $z = 0$. As a consequence, the d_θ-median coincides with the median for the real-valued case (i.e., since d_θ extends the Euclidean distance in $\mathbb{R}$, the d_θ-median extends the median of a real-valued random variable).

The case in which at least one sample interval-valued datum is not degenerated at a real number (i.e., there exists i_0 such that $\mathrm{spr}\, \mathsf{x}_{i_0} \neq 0$) is now to be examined. Note that the possibility of proposing a degenerated d_θ-median with spread zero is excluded, so the considered dominium be $\mathbb{D} = \mathbb{R} \times (0, \infty)$, and

$$f\left(y, 2 \cdot \min_{i:\, \mathrm{spr}\, \mathsf{x}_i \neq 0} \{\mathrm{spr}\, \mathsf{x}_i\}\right)$$

$$= \frac{1}{n} \sum_{i=1}^n \sqrt{(\mathrm{mid}\, \mathsf{x}_i - y)^2 + \theta(\mathrm{spr}\, \mathsf{x}_i - 2 \cdot \min_{i:\, \mathrm{spr}\, \mathsf{x}_i \neq 0} \{\mathrm{spr}\, \mathsf{x}_i\})^2} = \frac{1}{n} \sum_{i=1}^n \Bigg[(\mathrm{mid}\, \mathsf{x}_i - y)^2$$

$$+ \theta(\mathrm{spr}\, \mathsf{x}_i)^2 + 4\theta \cdot \left(\min_{i:\, \mathrm{spr}\, \mathsf{x}_i \neq 0} \{\mathrm{spr}\, \mathsf{x}_i\} \right)^2 - 4\theta \cdot \mathrm{spr}\, \mathsf{x}_i \min_{i:\, \mathrm{spr}\, \mathsf{x}_i \neq 0} \{\mathrm{spr}\, \mathsf{x}_i\} \Bigg]^{1/2}$$

$$= \frac{1}{n} \sum_{i=1}^n \sqrt{(\mathrm{mid}\, \mathsf{x}_i - y)^2 + \theta(\mathrm{spr}\, \mathsf{x}_i)^2 + 4\theta \min_{i:\, \mathrm{spr}\, \mathsf{x}_i \neq 0} \{\mathrm{spr}\, \mathsf{x}_i\} [\min_{i:\, \mathrm{spr}\, \mathsf{x}_i \neq 0} \{\mathrm{spr}\, \mathsf{x}_i\} - \mathrm{spr}\, \mathsf{x}_i]}$$

$$< \frac{1}{n} \sum_{i=1}^n \sqrt{(\mathrm{mid}\, \mathsf{x}_i - y)^2 + \theta(\mathrm{spr}\, \mathsf{x}_i)^2} = f(y, 0).$$

By the necessary condition for local maximum and minimum of real-valued functions of several variables, if there exists any (y_0, z_0) interior point of $\mathbb{D} = \mathbb{R} \times (0, \infty)$ for which a local maximum or local minimum of f is achieved, where

$$\begin{aligned} f:\ \mathbb{D} \subset \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (y, z) &\longmapsto \frac{1}{n} \sum_{i=1}^n \sqrt{(\mathrm{mid}\, \mathsf{x}_i - y)^2 + \theta \cdot (\mathrm{spr}\, \mathsf{x}_i - z)^2} \end{aligned}$$

in which the partial derivatives f_y and f_z exist, then $f_y(y_0, z_0) = f_z(y_0, z_0) = 0$. Therefore, if a local minimum is achieved at (y_0, z_0), one of the following situations will hold:

- Either the point associated with the local minimum belongs to the interior of $\mathbb{D}$, and f_y and f_z exist at this point. In this case, the conditions $f_y(y_0, z_0) = f_z(y_0, z_0) = 0$ are satisfied, i.e.,

$$0 = f_y(y_0, z_0) = \sum_{i=1}^n \frac{\mathrm{mid}\, \mathsf{x}_i - y_0}{\sqrt{(\mathrm{mid}\, \mathsf{x}_i - y_0)^2 + \theta(\mathrm{spr}\, \mathsf{x}_i - z_0)^2}},$$

$$0 = f_z(y_0, z_0) = \sum_{i=1}^{n} \frac{\operatorname{spr} \mathsf{x}_i - z_0}{\sqrt{(\operatorname{mid} \mathsf{x}_i - y_0)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_0)^2}}.$$

In fact, $f(y_0, z_0)$ is a local minimum and not a local maximum because of the second derivative criterion for local minimum, that is,

- the partial derivatives

$$f_y(y, z) = \frac{1}{n} \sum_{i=1}^{n} \frac{y - \operatorname{mid} \mathsf{x}_i}{\sqrt{(\operatorname{mid} \mathsf{x}_i - y)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z)^2}},$$

$$f_z(y, z) = \frac{1}{n} \sum_{i=1}^{n} \frac{\theta(z - \operatorname{spr} \mathsf{x}_i)}{\sqrt{(\operatorname{mid} \mathsf{x}_i - y)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z)^2}},$$

  are continuous and have continuous partial derivatives on the domain $\mathbb{D} \setminus \{(\operatorname{mid} \mathsf{x}_i, \operatorname{spr} \mathsf{x}_i)\}_{i=1}^n \subset \mathbb{R}^2$;

- $f_y(y_0, z_0) = f_z(y_0, z_0) = 0$ for an interior point (y_0, z_0) of the domain;

- $f_{yy}(y_0, z_0) = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \frac{\theta(\operatorname{spr} \mathsf{x}_i - z_0)^2}{\sqrt{(\operatorname{mid} \mathsf{x}_i - y_0)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_0)^2)^3}} > 0;$

- finally, by using the Cauchy-Swartz Inequality (strict inequality whenever not all the sample points are not collinear),

$$\begin{aligned} f_{yy} f_{zz}(y_0, z_0) = \frac{\theta^2}{n^2} & \left(\sum_{i=1}^{n} \frac{(\operatorname{spr} \mathsf{x}_i - z_0)^2}{\sqrt{((\operatorname{mid} \mathsf{x}_i - y_0)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_0)^2)^3}} \right) \\ & \cdot \left(\sum_{i=1}^{n} \frac{(y_0 - \operatorname{mid} \mathsf{x}_i)^2}{\sqrt{((\operatorname{mid} \mathsf{x}_i - y_0)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_0)^2)^3}} \right) \\ & > \frac{\theta^2}{n^2} \left(\sum_{i=1}^{n} \frac{\operatorname{spr} \mathsf{x}_i - z_0}{\sqrt[4]{((\operatorname{mid} \mathsf{x}_i - y_0)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_0)^2)^3}} \right. \\ & \qquad \left. \cdot \frac{y_0 - \operatorname{mid} \mathsf{x}_i}{\sqrt[4]{((\operatorname{mid} \mathsf{x}_i - y_0)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_0)^2)^3}} \right) \\ = & \left(\frac{1}{n} \sum_{i=1}^{n} \frac{\theta(y_0 - \operatorname{mid} \mathsf{x}_i)(\operatorname{spr} \mathsf{x}_i - z_0)}{\sqrt{((\operatorname{mid} \mathsf{x}_i - y_0)^2 + \theta(\operatorname{spr} \mathsf{x}_i - z_0)^2)^3}} \right)^2 = f_{yz}^2(y_0, z_0). \end{aligned}$$

  So, all the sufficient conditions of the criterion are fulfilled and f has a local minimum at (y_0, z_0).

- Or the local minimum is achieved at a point belonging to the interior of $\mathbb{D}$ but f_y or f_z do not exist at it. This can only happen if any of the square roots is equal to zero, i.e., if the point we are considering belongs to the sample $\{(\operatorname{mid} \mathsf{x}_i, \operatorname{spr} \mathsf{x}_i)\}_{i=1}^n$.

Of course, the other possible situations cannot hold under the assumed conditions. Thus,

- if $y_0 < \min_{1\leq i\leq n}\{\operatorname{mid} \mathsf{x}_i\}$, then $f_y(y_0, z_0) > 0$ and (y_0, z_0) does not belong to the sample;
- if $y_0 > \max_{1\leq i\leq n}\{\operatorname{mid} \mathsf{x}_i\}$, then $f_y(y_0, z_0) < 0$ and (y_0, z_0) does not belong to the sample;
- if $z_0 < \min_{1\leq i\leq n}\{\operatorname{spr} \mathsf{x}_i\}$, then $f_z(y_0, z_0) > 0$ and (y_0, z_0) does not belong to the sample;
- if $z_0 > \max_{1\leq i\leq n}\{\operatorname{spr} \mathsf{x}_i\}$, then $f_y(y_0, z_0) < 0$ and (y_0, z_0) does not belong to the sample.

Consequently, if (y_0, z_0) exists, it belongs to the rectangle

$$\left[\min_{1\leq i\leq n}\{\operatorname{mid} \mathsf{x}_i\}, \max_{1\leq i\leq n}\{\operatorname{mid} \mathsf{x}_i\}\right] \times \left[\min_{1\leq i\leq n}\{\operatorname{spr} \mathsf{x}_i\}, \max_{1\leq i\leq n}\{\operatorname{spr} \mathsf{x}_i\}\right].$$

As the restriction of the objective function f to this closed and bounded subset of $\mathbb{R}^2$ is continuous, the Weierstrass theorem guarantees that f has at least one minimum within this subset. Indeed, the function f is convex, so any local minimum is also a global minimum. Since the function is strictly convex whenever not all the sample data are collinear, the global minimum is unique in such a case. □

3.5.1 Consistency and robustness of the d_θ-median

In analyzing the inferential behaviour of the sample d_θ-median for random intervals, we are first going to analyze their strong consistency.

Proposition 3.5.3. *Let* X *be a random interval associated with a probability space* $(\Omega, \mathcal{A}, P)$ *and satisfying that the population* d_θ*-median* $\mathsf{M}_\theta[\mathsf{X}]$ *exists and it is unique. Then, the sample* d_θ*-median,* $\widehat{\mathsf{M}_\theta[\mathsf{X}]_n}$*, is a strongly consistent estimator of the population* d_θ*-median in* d_θ*-sense (and hence in the sense of all the topologically equivalent metrics), that is,*

$$\lim_{n\to\infty} d_\theta(\widehat{\mathsf{M}_\theta[\mathsf{X}]_n}, \mathsf{M}_\theta[\mathsf{X}]) = 0 \quad a.s.\ [P].$$

Proof. Recalling that the spatial median is the M-estimate of location for the special choice of $\rho(x) = \sqrt{|x|}$, it is sufficient to check that this loss function satisfies the assumptions required by Theorem 3.1.8 and Proposition 3.1.9, and this can be done easily. The rest of the conditions are fulfilled since the parameter set (the cone $\mathbb{R} \times [0, \infty)$ with the topology induced by the norm associated with the d_θ-metric through the isometrical embedding from $\mathcal{K}_c(\mathbb{R})$) is a locally compact space with a countable basis, $(\Omega, \mathcal{A}, P)$ is a probability space and it is supposed that the population d_θ-median exists and is unique. □

On the other hand, and concerning robustness, following Donoho and Huber [59], the finite sample breakdown point of the sample d_θ-median on a realization of a simple random sample of size n from a random interval X, $\mathbf{x}_n$, is given by

$$\begin{aligned} &\text{fsbp}(\widehat{\mathsf{M}_\theta[\mathsf{X}]}_n, \mathbf{x}_n, d_\theta) \\ &= \frac{1}{n} \min \left\{ k \in \{1, \ldots, n\} \ : \ \sup_{\mathbf{y}_{n,k}} d_\theta(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,k}]}) = \infty \right\}, \end{aligned}$$

with the metric space $(\mathcal{K}_c(\mathbb{R}), d_\theta)$ satisfying that

$$\sup_{K,K' \in \mathcal{K}_c(\mathbb{R})} d_\theta(K, K') = \infty$$

(since $d_\theta([n-1, n+1], [-n-1, -n+1]) = 2n$) and $\widehat{\mathsf{M}_\theta[\mathbf{y}_{n,k}]}$ is the sample median of a sample $\mathbf{y}_{n,k}$ obtained from $\mathbf{x}_n$ by perturbing at most k observations. Then, we have that

Proposition 3.5.4. *The finite sample breakdown point of the sample d_θ-median from a random interval X equals*

$$\text{fsbp}\left(\widehat{\mathsf{M}_\theta[\mathsf{X}]}_n, \mathbf{x}_n, d_\theta\right) = \frac{1}{n} \cdot \lfloor \frac{n+1}{2} \rfloor,$$

where $\lfloor \cdot \rfloor$ denotes the floor function.

Proof. The proof is prompted by some of the ideas by Lopuhaä and Rousseeuw [126] to compute the fsbp of the L_1 estimator in $\mathbb{R}^p$. However, the situation is now different, because d_θ is not translational equivariant.

First, one can see that

$$\text{fsbp}(\widehat{\mathsf{M}_\theta[\mathsf{X}]}_n, \mathbf{x}_n, d_\theta) > \frac{1}{n} \cdot \left(\lfloor \frac{n+1}{2} \rfloor - 1 \right) = \frac{1}{n} \cdot \lfloor \frac{n-1}{2} \rfloor,$$

that is, $\min \left\{ k \in \{1, \ldots, n\} \ : \ \sup_{\mathsf{y}_{n,k}} d_\theta(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,k}]}) = \infty \right\} > \lfloor \frac{n-1}{2} \rfloor$ or, equivalently,

$$\sup_{\mathsf{y}_{n, \lfloor \frac{n-1}{2} \rfloor}} d_\theta \left(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n, \lfloor \frac{n-1}{2} \rfloor}]} \right) < \infty.$$

Thus, by using the triangular inequality, we have that

$$\sup_{\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}} d_\theta\left(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]}\right)$$

$$\leq \sup_{\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}} d_\theta\left(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}, \{0\}\right) + \sup_{\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}} d_\theta\left(\{0\}, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]}\right).$$

The first term, which does not depend on the perturbed sample, is finite. So, it is enough to prove that the other one is also finite. Let $\eta = \max_{1\leq i\leq n} d_\theta(\{0\}, \mathsf{x}_i)$ and let $B(\mathbf{0}, 2\eta)$ be the closed ball with center $\mathbf{0} = (0,0)$ and radius 2η. Let

$$d = \inf_{\substack{(\operatorname{mid} V, \operatorname{spr} V) \in B(\mathbf{0}, 2\eta) \\ V \in \mathcal{K}_c(\mathbb{R})}} d_\theta\left(\widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]}, V\right)$$

denote the distance between $\widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]}$ and $B(\mathbf{0}, 2\eta)$, so that

$$d_\theta\left(\{0\}, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]}\right) \leq d + 2\eta.$$

Then, for each of the $\lfloor\frac{n-1}{2}\rfloor$ replaced y_j's, it holds by using the triangular inequality that

$$\begin{aligned} d_\theta\left(\mathsf{y}_j, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]}\right) &\geq d_\theta(\mathsf{y}_j, \{0\}) - d_\theta\left(\{0\}, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]}\right) \\ &\geq d_\theta(\mathsf{y}_j, \{0\}) - (d + 2\eta). \end{aligned} \tag{3.4}$$

Suppose the distance between $\widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]}$ and $B(\mathbf{0}, 2\eta)$ is large so that $d > 2\eta\lfloor\frac{n-1}{2}\rfloor$. This assumption can be proved to be wrong. Thus, since $(\operatorname{mid}\mathsf{x}_i, \operatorname{spr}\mathsf{x}_i) \in B(\mathbf{0}, \eta)$, for each of the $n - \lfloor\frac{n-1}{2}\rfloor$ original x_l's in $\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}$ we have that

$$d_\theta(\mathsf{x}_l, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]}) \geq \eta + d \geq d_\theta(\{0\}, \mathsf{x}_l) + d. \tag{3.5}$$

From Equations 3.4 and 3.5,

$$\sum_{i=1}^{n} d_\theta(\mathsf{y}_i, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]})$$

$$= \sum_{\substack{i=1 \\ i:\mathsf{y}_i \text{ replaced}}}^{n} d_\theta(\mathsf{y}_i, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]}) + \sum_{\substack{i=1 \\ i:\mathsf{y}_i \text{ original}}}^{n} d_\theta(\mathsf{y}_i, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]})$$

$$\geq \sum_{\substack{i=1 \\ i:\mathsf{y}_i \text{ replaced}}}^{n} (d_\theta(\mathsf{y}_i, \{0\}) - (d + 2\eta)) + \sum_{\substack{i=1 \\ i:\mathsf{y}_i \text{ original}}}^{n} (d_\theta(\{0\}, x_i) + d)$$

$$= \sum_{i=1}^{n} d_\theta(\mathsf{y}_i, \{0\}) - \lfloor\frac{n-1}{2}\rfloor(d + 2\eta) + \sum_{\substack{i=1 \\ i:\mathsf{x}_i \text{ original}}}^{n} d$$

$$= \sum_{i=1}^{n} d_\theta(\mathsf{y}_i, \{0\}) + \left(n - \lfloor\frac{n-1}{2}\rfloor\right) d - \lfloor\frac{n-1}{2}\rfloor(d+2\eta)$$
$$= \sum_{i=1}^{n} d_\theta(\mathsf{y}_i, \{0\}) + nd - \lfloor\frac{n-1}{2}\rfloor d - \lfloor\frac{n-1}{2}\rfloor d - 2\eta\lfloor\frac{n-1}{2}\rfloor$$
$$= \sum_{i=1}^{n} d_\theta(\mathsf{y}_i, \{0\}) + d - 2\eta\lfloor\frac{n-1}{2}\rfloor > \sum_{i=1}^{n} d_\theta(\mathsf{y}_i, \{0\})$$

by using the assumption $d > 2\eta\lfloor\frac{n-1}{2}\rfloor$. But this inequality is not valid because $\widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]}$ minimizes the mean d_θ distance to all the interval-valued data from the sample $\mathsf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}$. Therefore, $d \leq 2\eta\lfloor\frac{n-1}{2}\rfloor$ and the following inequalities are satisfied:

$$\sup_{\mathsf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}} d_\theta\left(\{0\}, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n-1}{2}\rfloor}]}\right) \leq d + 2\eta \leq 2\eta\lfloor\frac{n-1}{2}\rfloor + 2\eta = 2\eta\lfloor\frac{n+1}{2}\rfloor < \infty.$$

The second inequality, $\mathrm{fsbp}(\widehat{\mathsf{M}_\theta[\mathsf{X}]}_n, \mathbf{x}_n, d_\theta) \leq \frac{1}{n}\cdot\lfloor\frac{n+1}{2}\rfloor$, is now to be proved. This is equivalent to see that

$$\min\left\{k \in \{1,\ldots,n\} \; : \; \sup_{\mathsf{y}_{n,k}} d_\theta(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,k}]}) = \infty\right\} \leq \lfloor\frac{n+1}{2}\rfloor,$$

i.e., $\sup_{\mathsf{y}_{n,\lfloor\frac{n+1}{2}\rfloor}} d_\theta\left(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}, \widehat{\mathsf{M}_\theta[\mathbf{y}_{n,\lfloor\frac{n+1}{2}\rfloor}]}\right) = \infty.$

To see this, it is enough to find a corrupted collection $\mathsf{y}^*_{n,\lfloor\frac{n+1}{2}\rfloor}$ (by replacing at most $\lfloor\frac{n+1}{2}\rfloor$ points of $\mathbf{x}_n$) such that $d_\theta\left(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}, \widehat{\mathsf{M}_\theta[\mathbf{y}^*_{n,\lfloor\frac{n+1}{2}\rfloor}]}\right) = \infty$. For this purpose, and for an arbitrary $m \in \mathbb{N}$, $\lfloor\frac{n+1}{2}\rfloor$ observations of $\mathbf{x}_n$ are replaced by the same number of points that are all equal to the value $\mathsf{y}^{(m)} \in \mathcal{K}_c(\mathbb{R})$ with

$$\inf_{\substack{V \in \mathcal{K}_c(\mathbb{R}) \\ (\operatorname{mid} V, \operatorname{spr} V) \in B(\mathbf{0},\eta)}} d_\theta(\mathsf{y}^{(m)}, V) = m \in \mathbb{N}.$$

The new sample $\mathsf{y}^*_{n,\lfloor\frac{n+1}{2}\rfloor}$ contains $q = n - \lfloor\frac{n+1}{2}\rfloor$ of the original points. We are now going to verify that $\widehat{\mathsf{M}_\theta[\mathbf{y}^*_{n,\lfloor\frac{n+1}{2}\rfloor}]} = \mathsf{y}^{(m)}$.

Since $\mathsf{y}^*_{n,\lfloor\frac{n+1}{2}\rfloor} = (\mathsf{x}_1, \ldots, \mathsf{x}_q, \overbrace{\mathsf{y}^{(m)}, \ldots, \mathsf{y}^{(m)}}^{(n-q \text{ times})})$, for any $\mathsf{z} \in \mathcal{K}_c(\mathbb{R}) \setminus \{\mathsf{y}^{(m)}\}$ we have that

$$\sum_{i=1}^{n} d_\theta(\mathsf{y}^*_i, \mathsf{z}) = \sum_{i=1}^{q} d_\theta(\mathsf{x}_i, \mathsf{z}) + \sum_{i=q+1}^{n} d_\theta(\mathsf{y}^{(m)}, \mathsf{z})$$
$$= \sum_{i=1}^{q} d_\theta(\mathsf{x}_i, \mathsf{z}) + (n-q)\cdot d_\theta(\mathsf{y}^{(m)}, \mathsf{z}) \geq \sum_{i=1}^{q} d_\theta(\mathsf{x}_i, \mathsf{z}) + q\cdot d_\theta(\mathsf{y}^{(m)}, \mathsf{z})$$
$$= \sum_{i=1}^{q} \left(d_\theta(\mathsf{x}_i, \mathsf{z}) + d_\theta(\mathsf{y}^{(m)}, \mathsf{z})\right) \geq \sum_{i=1}^{q} d_\theta(\mathsf{x}_i, \mathsf{y}^{(m)})$$

$$= \sum_{i=1}^{q} d_\theta(\mathsf{x}_i, \mathsf{y}^{(m)}) + \sum_{i=q+1}^{n} d_\theta(\mathsf{y}^{(m)}, \mathsf{y}^{(m)}) = \sum_{i=1}^{n} d_\theta(\mathsf{y}_i^*, \mathsf{y}^{(m)}).$$

Therefore, and due to the fact that $\widehat{\mathsf{M}_\theta[\mathbf{x}_n]} \in B(\mathbf{0}, \eta)$,

$$\sup_{\mathsf{y}_{n,\lfloor\frac{n+1}{2}\rfloor}} d_\theta\left(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}, \mathsf{M}_\theta\widehat{[\mathbf{y}_{n,\lfloor\frac{n+1}{2}\rfloor}]}\right) \geq d_\theta\left(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}, \mathsf{M}_\theta\widehat{[\mathbf{y}^*_{n\lfloor\frac{n+1}{2}\rfloor}]}\right)$$

$$= d_\theta(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}, \mathsf{y}^{(m)}) \geq m.$$

Since $m \in \mathbb{N}$ could be chosen to be arbitrarily large

$$\sup_{\mathsf{y}_{n,\lfloor\frac{n+1}{2}\rfloor}} d_\theta\left(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}, \mathsf{M}_\theta\widehat{\left[\mathsf{y}_{n,\lfloor\frac{n+1}{2}\rfloor}\right]}\right) = \infty. \qquad \square$$

The following result formalizes the comparison of the robustness of the sample d_θ-median and the sample mean of a random interval. Thus,

Theorem 3.5.5. *The finite sample breakdown point of the sample mean from a random interval* X, $\mathrm{fsbp}(\overline{\mathsf{X}_n})$, *is lower than that for the sample d_θ-median for sample sizes* $n > 2$.

Proof. Indeed, by arguing like for the preceding proposition we have that

$$\mathrm{fsbp}(\overline{\mathsf{X}_n}, \mathbf{x}_n, d_\theta) = \mathrm{fsbp}(\overline{\mathsf{X}_n}, \mathbf{x}_n, \delta_1) = \frac{1}{n},$$

and, consequently,

$$\mathrm{fsbp}(\widehat{\mathsf{M}_\theta[\mathsf{X}]}_n, \mathbf{x}_n, d_\theta) = \mathrm{fsbp}(\widehat{\mathsf{M}_\theta[\mathsf{X}]}_n, \mathbf{x}_n, \delta_1)$$

$$\geq \frac{n/2}{n} = \frac{1}{2} > \frac{1}{n} = \mathrm{fsbp}(\overline{\mathsf{X}_n}, \mathbf{x}_n, \delta_1) = \mathrm{fsbp}(\overline{\mathsf{X}_n}, \mathbf{x}_n, d_\theta). \qquad \square$$

3.5.2 An algorithm to compute the sample spatial median

A natural algorithm to compute the sample d_θ-median is to be explained and analyzed along this subsection. Weiszfeld [215] (the English translation of this paper and some annotations can be found in Weiszfeld and Plastria [216]) proposed this iterative algorithm for Euclidean spaces and distances, the algorithm being rediscovered some more times in the literature.

The objective function in the minimization problem in Definition 3.5.2 is differentiable at any point of the domain $\mathbb{R}\times(0,\infty)$ but the sample points $\{(\mathrm{mid}\,\mathsf{x}_i, \mathrm{spr}\,\mathsf{x}_i)\}_{i=1}^n$. The minimum will be reached either at a sample point or at the point for which both partial derivatives equal zero, that is to say, at the point (y_0, z_0) satisfying that

$$y_0 = \frac{\displaystyle\sum_{i=1}^{n} \frac{\mathrm{mid}\,\mathsf{x}_i}{\sqrt{(\mathrm{mid}\,\mathsf{x}_i - y_0)^2 + \theta\cdot(\mathrm{spr}\,\mathsf{x}_i - z_0)^2}}}{\displaystyle\sum_{i=1}^{n} \frac{1}{\sqrt{(\mathrm{mid}\,\mathsf{x}_i - y_0)^2 + \theta\cdot(\mathrm{spr}\,\mathsf{x}_i - z_0)^2}}}$$

and

$$z_0 = \frac{\displaystyle\sum_{i=1}^{n} \frac{\operatorname{spr} \mathsf{x}_i}{\sqrt{(\operatorname{mid} \mathsf{x}_i - y_0)^2 + \theta \cdot (\operatorname{spr} \mathsf{x}_i - z_0)^2}}}{\displaystyle\sum_{i=1}^{n} \frac{1}{\sqrt{(\operatorname{mid} \mathsf{x}_i - y_0)^2 + \theta \cdot (\operatorname{spr} \mathsf{x}_i - z_0)^2}}}.$$

An immediate remark after observing these expressions is that the mid-point and the spread of the sample d_θ-median are a weighted mean of the mid-points and the spreads of the intervals in the sample, respectively. The steps of the algorithm used are now to be detailed and, afterwards, its convergence is to be proved.

Step 0. Firstly, compute the mid-points and spreads of the interval-valued data, namely,

$$\operatorname{mid} \mathsf{x}_i = \frac{\inf \mathsf{x}_i + \sup \mathsf{x}_i}{2}, \quad \operatorname{spr} \mathsf{x}_i = \frac{\sup \mathsf{x}_i - \inf \mathsf{x}_i}{2} \quad \text{for } i = 1, \ldots, n.$$

Step 1. Fix the maximum number of iterations, the tolerance of the approximation and set $m = 1$. Moreover, fix a seed $(y_m, z_m) \in \mathbb{R} \times [0, \infty)$ and the weight $\theta > 0$, and calculate the corresponding error

$$\text{Error}_m = \frac{1}{n} \sum_{i=1}^{n} \sqrt{(\operatorname{mid} \mathsf{x}_i - y_m)^2 + \theta \cdot (\operatorname{spr} \mathsf{x}_i - z_m)^2}. \tag{3.6}$$

Step 2. Compute the weights

$$w_i = \frac{\dfrac{1}{\sqrt{(\operatorname{mid} \mathsf{x}_i - y_m)^2 + \theta \cdot (\operatorname{spr} \mathsf{x}_i - z_m)^2}}}{\displaystyle\sum_{j=1}^{n} \frac{1}{\sqrt{(\operatorname{mid} \mathsf{x}_j - y_m)^2 + \theta \cdot (\operatorname{spr} \mathsf{x}_j - z_m)^2}}} \quad \text{for all } i = 1, \ldots, n$$

and update the estimate

$$y_{m+1} = \sum_{i=1}^{n} w_i \cdot \operatorname{mid} \mathsf{x}_i, \quad z_{m+1} = \sum_{i=1}^{n} w_i \cdot \operatorname{spr} \mathsf{x}_i.$$

Step 3. For the new estimate (y_{m+1}, z_{m+1}), compute the corresponding error Error_{m+1} as given by (3.6). If the difference $\text{Error}_m - \text{Error}_{m+1}$ exceeds the specified tolerance and the number of iterations is lower than the maximum, then increase m by 1 and go to *Step 2.* Otherwise, go to *Step 4.*

Step 4. Compare the final error Error_{m+1} obtained in *Step 3* with the errors $\mathrm{Error}(\mathsf{x}_j)$ corresponding to each interval data x_j, where

$$\mathrm{Error}(\mathsf{x}_j) = \frac{1}{n}\sum_{i=1}^{n}\sqrt{(\mathrm{mid}\,\mathsf{x}_i - \mathrm{mid}\,\mathsf{x}_j)^2 + \theta\cdot(\mathrm{spr}\,\mathsf{x}_i - \mathrm{spr}\,\mathsf{x}_j)^2}$$

If $\mathrm{Error}_{m+1} < \min_j \mathrm{Error}(\mathsf{x}_j)$, then retake the solution (y_{m+1}, z_{m+1}). Otherwise, retake the solution $(\mathrm{mid}\,\mathsf{x}_{j_0}, \mathrm{spr}\,\mathsf{x}_{j_0})$ where $j_0 = \arg\min_j \mathrm{Error}(\mathsf{x}_j)$.

The following theorem, proving the convergence of this algorithm, is based on the ideas by Kuhn [121] to show the convergence of the analogous algorithm in computing the spatial median in a Euclidean space (with the Euclidean distance).

Theorem 3.5.6. *Given a realization of a simple random sample from a random interval* X, $\mathbf{x}_n = (\mathsf{x}_1, \ldots, \mathsf{x}_n)$, *and any* $P_0 \in \mathbb{R}\times(0,\infty)\setminus\{(\mathrm{mid}\,\mathsf{x}_i, \mathrm{spr}\,\mathsf{x}_i)\}_{i=1}^n$, *define* $P_r := T^r(P_0) = T(T(\ldots T(P_0)))$ *for* $r = 1, 2, \ldots$, *where*

$$T:\ \mathbb{R}\times(0,\infty)\setminus\{(\mathrm{mid}\,\mathsf{x}_i, \mathrm{spr}\,\mathsf{x}_i)\}_{i=1}^n \longrightarrow \mathbb{R}\times(0,\infty)$$

$$P \longmapsto \left(\frac{\sum_{i=1}^{n}\mathrm{mid}\,\mathsf{x}_i/d_\theta^i(P)}{\sum_{i=1}^{n}1/d_\theta^i(P)}, \frac{\sum_{i=1}^{n}\mathrm{spr}\,\mathsf{x}_i/d_\theta^i(P)}{\sum_{i=1}^{n}1/d_\theta^i(P)}\right),$$

with $d_\theta^i(P)$ *standing for the simplified notation of* $d_\theta(\mathsf{x}_i, P)$. *If no* P_r *is a vertex, then* $\lim_{r\to\infty} P_r = \widehat{\mathsf{M}_\theta[\mathbf{x}_n]}$.

Proof. With the possible exception of P_0, the sequence $\{P_r\}_r$ lies in the convex hull of the vertices, which is a compact set. Hence, by the Bolzano-Weierstrass Theorem, there exists at least one point P and a subsequence $\{P_{r_l}\}_l$ such that $\lim_{l\to\infty} P_{r_l} = P$. To prove the theorem, one should verify that P coincides with the minimum, denoted by $\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}$. All the possible situations can be examined:

- Case $P_{r+1} = T(P_r) = P_r$ for some r.

  Then, the sequence repeats from that point and $P = P_r$. Since P_r is not a vertex and $T(P) = P$ (the necessary condition to be the minimum $\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}$)), P equals $\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}$ because there is only a unique minimum.

- Case $P_{r+1} = T(P_r) \neq P_r$ for all r and P not being a vertex.

  In this case, $f(T(P_r)) < f(P_r)$ and this holds for all r, since $P_r \neq T(P_r)$ and, hence,

$$g(T(P_r)) = \sum_{i=1}^{n}\frac{1}{d_\theta^i(P_r)}(d_\theta^i(T(P_r)))^2 < \sum_{i=1}^{n}\frac{1}{d_\theta^i(P_r)}(d_\theta^i(P_r))^2 = f(P_r)$$

because $T(P_r)$ is the center of gravity of weights $1/d_\theta^i(P_r)$ placed at the vertices $(\operatorname{mid} \mathsf{x}_i, \operatorname{spr} \mathsf{x}_i)$ and, therefore, the unique minimum of the strictly convex function $g(Q) := \sum_{i=1}^n (d_\theta^i(Q))^2/d_\theta^i(P_r)$. On the other hand,

$$g(T(P_r)) = \sum_{i=1}^n \frac{1}{d_\theta^i(P_r)}(d_\theta^i(P_r) + (d_\theta^i(T(P_r)) - d_\theta^i(P_r)))^2$$

$$= \sum_{i=1}^n d_\theta^i(P_r) + 2\sum_{i=1}^n d_\theta^i(T(P_r)) - 2\sum_{i=1}^n d_\theta^i(P_r)$$

$$+ \sum_{i=1}^n \frac{1}{d_\theta^i(P_r)}(d_\theta^i(T(P_r)) - d_\theta^i(P_r))^2$$

$$= f(P_r) + 2f(T(P_r)) - 2f(P_r) + \sum_{i=1}^n \frac{1}{d_\theta^i(P_r)}(d_\theta^i(T(P_r)) - d_\theta^i(P_r))^2.$$

Therefore,

$$2f(T(P_r)) \leq 2f(T(P_r)) + \sum_{i=1}^n \frac{1}{d_\theta^i(P_r)}(d_\theta^i(T(P_r)) - d_\theta^i(P_r))^2 < 2f(P_r).$$

Hence, $f(P_0) > f(P_1) > \ldots > f(P_r) > f(P_{r+1}) > \ldots f(\widehat{\mathsf{M}_\theta[\mathbf{x}_n]})$ and furthermore $\lim_{l\to\infty}(f(P_{r_l}) - f(T(P_{r_l}))) = 0$.

Consequently, if $P = \lim_{l\to\infty} P_{r_l}$, and by using the continuity of T and f,

$$f(P) - f(T(P)) = f(\lim_{l\to\infty} P_{r_l}) - f(\lim_{l\to\infty} T(P_{r_l}))$$

$$= \lim_{l\to\infty}(f(P_{r_l}) - f(T(P_{r_l}))) = 0,$$

i.e., $f(P) = f(T(P))$. By using again what has just been proven (if $T(P) \neq P$, then $f(T(P)) < f(P)$), it is concluded that $P = T(P)$. Finally, P is not a vertex, so $P = \widehat{\mathsf{M}_\theta[\mathbf{x}_n]}$ like for the first case.

- Case $P = (\operatorname{mid} \mathsf{x}_j, \operatorname{spr} \mathsf{x}_j)$ for some $j \in \{1, \ldots, n\}$.

  Like the cases before, the proof can be adapted from the reasoning in Kuhn [121] to this different situation (d_θ metric, space of non-empty compact intervals and constant weights in the objective function) without extra difficulties. Suppose that $(\operatorname{mid} \mathsf{x}_j, \operatorname{spr} \mathsf{x}_j) \neq \widehat{\mathsf{M}_\theta[\mathbf{x}_n]}$. By the result in Kuhn [121], there exists $\delta > 0$ such that $0 < d_\theta^j(P) \leq \delta$ and that implies $d_\theta^j(T^s(P)) > \delta$ and $d_\theta^j(T^{s-1}(P)) \leq \delta$ for some positive integer s. Consequently, the subsequence $P_{r_l} \to (\operatorname{mid} \mathsf{x}_j, \operatorname{spr} \mathsf{x}_j)$ can be chosen such that $d_\theta^j(T(P_{r_l})) > \delta$ for all l, that is to say, the ratio $d_\theta^j(T(P_{r_l}))/d_\theta^j(P_{r_l})$ is unbounded.

Another result in Kuhn [121] which can be immediately adapted to interval-valued data is the calculus of the following limit:

$$\lim_{P\to(\operatorname{mid}\mathsf{x}_j,\operatorname{spr}\mathsf{x}_j)} \frac{d_\theta^j(T(P))}{d_\theta^j(P)} = \left|\sum_{i\neq j} \frac{(\operatorname{mid}\mathsf{x}_i,\operatorname{spr}\mathsf{x}_i)-(\operatorname{mid}\mathsf{x}_j,\operatorname{spr}\mathsf{x}_j)}{d_\theta^i((\operatorname{mid}\mathsf{x}_j,\operatorname{spr}\mathsf{x}_j))}\right| < \infty.$$

This contradict the considered assumption, so one can immediately conclude that $(\operatorname{mid}\mathsf{x}_j,\operatorname{spr}\mathsf{x}_j)=\widehat{\mathsf{M}_\theta[\mathbf{x}_n]}$. □

3.5.3 Simulation-based comparison between the Aumann mean value and the d_θ-median

The aim of this subsection is to compare synthetically the behaviour of the Aumann mean value and the d_θ-median for interval-valued data.

The simulations have been performed as follows:

Step 1. A sample of size $n = 100$ interval-valued data has been simulated from a random interval X for each of some different situations in such a way that

- to generate the interval-valued data, we have considered two real-valued random variables as follows: $\mathsf{X} = [X_1 - X_2, X_1 + X_2]$, with $X_1 = \operatorname{mid}\mathsf{X}$, $X_2 = \operatorname{spr}\mathsf{X}$;
- each sample is assumed to be split into a subsample of size $n(1-c_p)$ (where c_p denotes the proportion of contamination and is supposed to range in $\{0, 0.1, 0.2, 0.4\}$) associated with a non-contaminated distribution and a subsample of size $n\cdot c_p$ associated with a contaminated one, where an additional contamination role is played by C_D (which measures the relative distance between the distribution of the two subsamples and ranges in $\{0, 1, 5, 10, 100\}$). In total, the 16 situations for different values of c_p and C_D have been considered;
- for each of these situations three cases have been selected, namely, one in which random variables X_i are independent (CASE 1) and two in which they are dependent (CASES 2 and 2'). First, the non contaminated data are generated according to
  - ∗ $\operatorname{mid}\mathsf{X}\sim\mathcal{N}(0,1)$ and $\operatorname{spr}\mathsf{X}\sim\chi_1^2$ for CASE 1,
  - ∗ $\operatorname{mid}\mathsf{X}\sim\mathcal{N}(0,1)$ and $\operatorname{spr}\mathsf{X}\sim\left(\frac{1}{(\operatorname{mid}\mathsf{X})^2+1}\right)^2+0.1\cdot\chi_1^2$ for CASE 2,
  - ∗ $\operatorname{mid}\mathsf{X}\sim\mathcal{N}(0,1)$ and $\operatorname{spr}\mathsf{X}\sim\frac{1}{(\operatorname{mid}\mathsf{X})^2+1}+\sqrt{\chi_1^2}$ for CASE 2'.

  Then, $n\cdot c_p$ data from that sample are contaminated (contaminating either the location or the spread or both of them) using the following distributions:

* $\operatorname{mid} \mathsf{X} \sim \mathcal{N}(0,3) + C_D$ and $\operatorname{spr} \mathsf{X} \sim \chi^2_4 + C_D$ for CASE 1,
* $\operatorname{mid} \mathsf{X} \sim \mathcal{N}(0,3) + C_D$ and $\operatorname{spr} \mathsf{X} \sim \left(\frac{1}{(\operatorname{mid} \mathsf{X})^2+1}\right)^2 + 0.1 \cdot \chi^2_1 + C_D$ for CASE 2,
* $\operatorname{mid} \mathsf{X} \sim \mathcal{N}(0,\sqrt{3}) + C_D$ and $\operatorname{spr} \mathsf{X} \sim \frac{1}{(\operatorname{mid} \mathsf{X})^2+1} + \sqrt{\chi^2_1} + C_D$ for CASE 2'.

As it can be noticed, the difference between CASE 2 and CASE 2' derives from the way of generating the spread of the contaminated observations.

Step 2. Since the population parameters (i.e. the Aumann expected value and population $\boldsymbol{\rho_1}$ and d_θ-medians) cannot be derived analytically for random intervals X according to the above distributions, $N = 1000$ replications of *Step 1* have been considered for the situation $c_p = C_D = 0$ in order to approximate them by using a Monte Carlo approach. These targets appear on the first row of Tables 3.7, 3.8 and 3.9.

Step 3. In order to choose the parameter θ, a sensitivity analysis of the influence of such a choice on the bias

$$d_\theta \left(E[\widehat{M[\mathsf{X}]_n}], M[\mathsf{X}] \right),$$

the variance

$$E\left(d^2_\theta(\widehat{M[\mathsf{X}]_n}, E[\widehat{M[\mathsf{X}]_n}]) \right)$$

and the mean squared error

$$E\left(d^2_\theta(\widehat{M[\mathsf{X}]_n}, M[\mathsf{X}]) \right)$$

of the d_θ-median has been first developed by distinguishing (abscise labels) the 16 different situations in terms of the considered choices of the parameters c_p and C_D. More concretely,

$$1 \equiv (c_p = 0, C_D = 0),\ 2 \equiv (c_p = 0.1, C_D = 0),$$
$$3 \equiv (c_p = 0.1, C_D = 1),\ 4 \equiv (c_p = 0.1, C_D = 5),$$
$$5 \equiv (c_p = 0.1, C_D = 10),\ 6 \equiv (c_p = 0.1, C_D = 100),$$
$$7 \equiv (c_p = 0.2, C_D = 0),\ 8 \equiv (c_p = 0.2, C_D = 1),$$
$$9 \equiv (c_p = 0.2, C_D = 5),\ 10 \equiv (c_p = 0.2, C_D = 10),$$
$$11 \equiv (c_p = 0.2, C_D = 100),\ 12 \equiv (c_p = 0.4, C_D = 0),$$
$$13 \equiv (c_p = 0.4, C_D = 1),\ 14 \equiv (c_p = 0.4, C_D = 5),$$
$$15 \equiv (c_p = 0.4, C_D = 10),\ 16 \equiv (c_p = 0.4, C_D = 100).$$

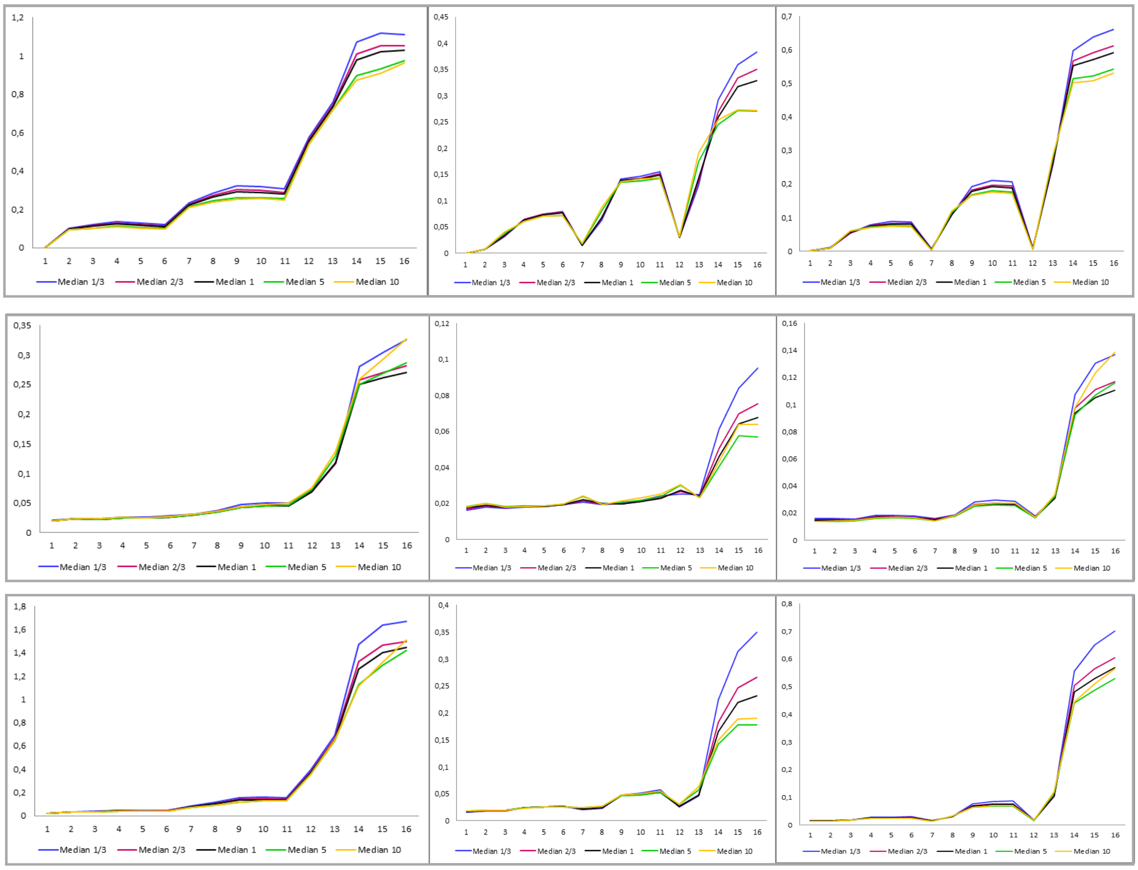


Figure 3.21: Bias (top), variance (middle) and MSE (bottom) of the d_θ-median in CASES 1 (left column), 2 (middle) and 2' (right column) for different choices of θ: 1/3, 2/3, 1, 5 and 10

$N = 1000$ replications of *Step 1* have been considered for all the situations (c_p, C_D), the three measures have been computed for each of them ($\widehat{M[X]}^j_{1000}$ denotes the sample d_θ-median of the j^{th} sample, $j \in \{1, \ldots, 1000\}$) and the Monte Carlo approximation has been given by

$$d_\theta\left(\frac{1}{1000}\sum_{i=1}^{1000} \widehat{M[\mathsf{X}]}^i_{1000}, M[\mathsf{X}]\right)$$

for the bias,

$$\frac{1}{1000}\sum_{i=1}^{1000} d^2_\theta\left(\widehat{M[\mathsf{X}]}^i_{1000}, E[\widehat{M[\mathsf{X}]}_{1000}]\right)$$

for the variance and

$$\frac{1}{1000}\sum_{i=1}^{n} d^2_\theta\left(\widehat{M[\mathsf{X}]}^i_{1000}, M[\mathsf{X}]\right)$$

for the mean squared error.

Figure 3.21 summarizes some of the most noteworthy obtained results.

Although the study has also taken into account other possible values of θ, they have not been included in the graphical displays because they have not visibly improved the ones selected for them.

The value of θ that rather uniformly minimizes the bias, variance and mean squared error is $\theta = 5$. Apart from considering this value, $\theta = 1$ has been chosen too in the computation of the d_θ-median, because this corresponds to $\boldsymbol{\rho_2}$, the analogous L^2-type of the $\boldsymbol{\rho_1}$ distance.

Step 4. $N = 1000$ replications of *Step 1* have been considered for all the situations (c_p, C_D) and the approximated bias, variance and mean squared error have been computed for each location measure (Aumann mean, $\boldsymbol{\rho_1}$-median, $d_{\theta=1}$- and $d_{\theta=5}$-medians) and in every situation (c_p, C_D) following the scheme in *Step 3* and in terms of the three considered metrics: $\boldsymbol{\rho_1}$, $d_{\theta=1}$ and $d_{\theta=5}$.

The results for CASES 1, 2 and 2' are shown in Tables 3.7, 3.8 and 3.9, respectively. The bias, variance and mean squared error have been computed for each location measure and for every situation in terms of the three considered metrics: $\boldsymbol{\rho_1}$ (left), $d_{\theta=1}$ (center) and $d_{\theta=5}$ (right). The minimum value for each choice of c_p and C_D (each row) and each metric has been highlighted in bold letters.

These tables show that the estimates of all the medians are much less influenced by the contamination than the sample (Aumann) mean.

Furthermore,

- in CASE 1 the $d_{\theta=5}$-median is the best choice in terms of bias, mean squared error and for most of the situations c_p and C_D when looking at the variance;

- in CASE 2 the most convenient measure in terms of bias and mean squared error and also of variance is the $\boldsymbol{\rho_1}$-median for most of the choices of c_p and C_D;

- in CASE 2' the $\boldsymbol{\rho_1}$-median is again the best choice in terms of bias, variance and mean squared error in most of the situations.

The simulation study was also developed for a sample size of $n = 5000$, but the conclusions were the same than for the considered sample size.

c_p	c_D	feature	Mean	$\boldsymbol{\rho_1}$-Median	$d_{\theta=1}$-Median	$d_{\theta=5}$-Median
.0	0	Estimate	[−1.00363, .99079]	[−.75301, .73239]	[−.61602, .60027]	[−.55149, .53517]
		Bias	0 ⋆ 0 ⋆ 0	0 ⋆ 0 ⋆ 0	0 ⋆ 0 ⋆ 0	0 ⋆ 0 ⋆ 0
		Variance	.0237 ⋆ .0281 ⋆ .1040	.0222 ⋆ .0272 ⋆ .0789	.0182 ⋆ .0221 ⋆ .0621	**.0179** ⋆ **.0218** ⋆ **.0599**
		MSE	.0237 ⋆ .0281 ⋆ .1040	.0222 ⋆ .0272 ⋆ .0789	.0182 ⋆ .0221 ⋆ .0621	**.0179** ⋆ **.0218** ⋆ **.0599**
.1	0	Estimate	[−1.20792, 1.22446]	[−.82329, .85281]	[−.68828, .71099]	[−.61441, .63869]
		Bias	.2189 ⋆ .2194 ⋆ .4898	.0953 ⋆ .0985 ⋆ .2146	.0914 ⋆ .0934 ⋆ .2054	**.0832** ⋆ **.0856** ⋆ **.1872**
		Variance	.0374 ⋆ .0445 ⋆ .1636	.0264 ⋆ .0323 ⋆ .0981	.0221 ⋆ .0268 ⋆ .0789	**.0218** ⋆ **.0267** ⋆ **.0753**
		MSE	.0827 ⋆ .0927 ⋆ .4035	.0348 ⋆ .0420 ⋆ .1442	.0297 ⋆ .0355 ⋆ .1211	**.0283** ⋆ **.0340** ⋆ **.1103**
.1	1	Estimate	[−1.24314, 1.34182]	[−.84280, .85998]	[−.70045, .72645]	[−.62565, .65218]
		Bias	.2952 ⋆ .3004 ⋆ .6625	.1086 ⋆ .1103 ⋆ .2437	.1053 ⋆ .1073 ⋆ .2363	**.0955** ⋆ **.0979** ⋆ **.2148**
		Variance	.0375 ⋆ .0447 ⋆ .1625	.0275 ⋆ .0332 ⋆ .0981	.0229 ⋆ .0276 ⋆ .0820	**.0226** ⋆ **.0273** ⋆ **.0781**
		MSE	.1208 ⋆ .1350 ⋆ .6016	.0380 ⋆ .0454 ⋆ .1575	.0329 ⋆ .0392 ⋆ .1378	**.0308** ⋆ **.0369** ⋆ **.1243**
.1	5	Estimate	[−1.30759, 1.85151]	[−.81899, .92700]	[−.66771, .79271]	[−.591700, .71218]
		Bias	.5823 ⋆ .6454 ⋆ 1.331	.1302 ⋆ .1453 ⋆ .2983	.1220 ⋆ .1409 ⋆ .2818	**.1086** ⋆ **.1283** ⋆ **.2523**
		Variance	.0776 ⋆ .0985 ⋆ .3425	.0287 ⋆ .0352 ⋆ .1081	.0248 ⋆ .0302 ⋆ .0907	**.0245** ⋆ **.0295** ⋆ **.0862**
		MSE	.4324 ⋆ .5151 ⋆ 2.115	.0467 ⋆ .0563 ⋆ .1971	.0415 ⋆ .0500 ⋆ .1702	**.0381** ⋆ **.0460** ⋆ **.1499**
.1	10	Estimate	[−1.42225, 2.50618]	[−.81255, .93881]	[−.65708, .80637]	[−.58126, .72727]
		Bias	.9670 ⋆ 1.111 ⋆ 2.230	.1329 ⋆ .1519 ⋆ .3062	.1235 ⋆ .1486 ⋆ .2884	**.1109** ⋆ **.1374** ⋆ **.2610**
		Variance	.1587 ⋆ .2087 ⋆ .6513	.0277 ⋆ .0340 ⋆ .1006	**.0243** ⋆ **.0299** ⋆ .0880	.0246 ⋆ .0300 ⋆ **.0851**
		MSE	1.153 ⋆ 1.444 ⋆ 5.627	.0471 ⋆ .0570 ⋆ .1944	.0428 ⋆ .0520 ⋆ .1712	**.0403** ⋆ **.0489** ⋆ **.1532**
.1	100	Estimate	[−3.20663, 13.87819]	[−.80369, .93273]	[−.63829, .80506]	[−.56320, .73136]
		Bias	7.545 ⋆ 9.244 ⋆ 17.69	.1255 ⋆ .1461 ⋆ .2904	.1135 ⋆ .1456 ⋆ .2697	**.1039** ⋆ **.1389** ⋆ **.2500**
		Variance	10.89 ⋆ 15.10 ⋆ 41.48	.0312 ⋆ .0379 ⋆ .1129	.0267 ⋆ .0322 ⋆ .0941	**.0266** ⋆ **.0319** ⋆ **.0893**
		MSE	75.92 ⋆ 100.5 ⋆ 354.6	.0489 ⋆ .0593 ⋆ .1973	.0441 ⋆ .0534 ⋆ .1669	**.0423** ⋆ **.0512** ⋆ **.1519**
.2	0	Estimate	[−1.45244, 1.43712]	[−.96369, .95769]	[−.82747, .82451]	[−.75115, .74535]
		Bias	.4475 ⋆ .4475 ⋆ 1.000	.2179 ⋆ .2181 ⋆ .4874	.2178 ⋆ .2179 ⋆ .4871	**.2049** ⋆ **.2049** ⋆ **.4582**
		Variance	.0583 ⋆ .0681 ⋆ .2658	.0346 ⋆ .0418 ⋆ .1377	.0308 ⋆ .0368 ⋆ .1239	**.0302** ⋆ **.0363** ⋆ **.1197**
		MSE	.2519 ⋆ .2685 ⋆ 1.267	.0770 ⋆ .0894 ⋆ .3753	.0737 ⋆ .0843 ⋆ .3613	**.0679** ⋆ **.0784** ⋆ **.3297**
.2	1	Estimate	[−1.48406, 1.70003]	[−.96885, 1.05911]	[−.83062, .91792]	[−.74210, .83336]
		Bias	.5948 ⋆ .6057 ⋆ 1.335	.2712 ⋆ .2768 ⋆ .6091	.2661 ⋆ .2710 ⋆ .5973	**.2444** ⋆ **.2502** ⋆ **.5491**
		Variance	.0714 ⋆ .0826 ⋆ .3252	.0383 ⋆ .0466 ⋆ .1499	**.0345** ⋆ **.0415** ⋆ **.1435**	.0350 ⋆ .0423 ⋆ .1438
		MSE	.4182 ⋆ .4496 ⋆ 2.107	.1058 ⋆ .1233 ⋆ .5210	.1011 ⋆ .1150 ⋆ .5003	**.0913** ⋆ **.1049** ⋆ **.4454**
.2	5	Estimate	[−1.66475, 2.69900]	[−.91026, 1.16926]	[−.75377, 1.02697]	[−.65478, .93147]
		Bias	1.184 ⋆ 1.295 ⋆ 2.700	.2970 ⋆ .3283 ⋆ .6787	.2822 ⋆ .3170 ⋆ .6474	**.2497** ⋆ **.2895** ⋆ **.5774**
		Variance	.2032 ⋆ .2615 ⋆ .9013	.0447 ⋆ .0551 ⋆ .1727	**.0401** ⋆ .0494 ⋆ .1600	.0402 ⋆ **.0490** ⋆ **.1509**
		MSE	1.641 ⋆ 1.939 ⋆ 8.192	.1368 ⋆ .1629 ⋆ .6334	.1268 ⋆ .1499 ⋆ .5791	**.1110** ⋆ **.1329** ⋆ **.4844**
.2	10	Estimate	[−1.82779, 3.96648]	[−.89244, 1.17874]	[−.71878, 1.05280]	[−.62321, .96513]
		Bias	1.899 ⋆ 2.183 ⋆ 4.382	.2928 ⋆ .3306 ⋆ .6726	.2776 ⋆ .3281 ⋆ .6450	**.2508** ⋆ **.3082** ⋆ **.5888**
		Variance	.5682 ⋆ .7688 ⋆ 2.426	.0484 ⋆ .0596 ⋆ .1848	**.0427** ⋆ **.0525** ⋆ .1682	.0439 ⋆ .0539 ⋆ **.1607**
		MSE	4.440 ⋆ 5.535 ⋆ 21.63	.1406 ⋆ .1689 ⋆ .6373	.1323 ⋆ .1602 ⋆ .5843	**.1231** ⋆ **.1489** ⋆ **.5073**
.2	100	Estimate	[−5.91430, 26.91285]	[−.91153, 1.17310]	[−.71181, 1.05516]	[−.61891, .97184]
		Bias	15.41 ⋆ 18.65 ⋆ 36.03	.2999 ⋆ .3315 ⋆ .6853	.2753 ⋆ .3287 ⋆ .6413	**.2520** ⋆ **.3124** ⋆ **.5930**
		Variance	42.49 ⋆ 57.28 ⋆ 148.9	.0502 ⋆ .0619 ⋆ .1964	.0468 ⋆ **.0571** ⋆ .1836	**.0468** ⋆ .0571 ⋆ **.1717**
		MSE	309.4 ⋆ 405.3 ⋆ 1447	.1450 ⋆ .1719 ⋆ .6661	.1377 ⋆ .1651 ⋆ .5949	**.1288** ⋆ **.1547** ⋆ **.5235**
.4	0	Estimate	[−1.90223, 1.89912]	[−1.27397, 1.27655]	[−1.16865, 1.16930]	[−1.08960, 1.09451]
		Bias	.9034 ⋆ .9034 ⋆ 2.020	**.5325** ⋆ **.5326** ⋆ **1.190**	.5608 ⋆ .5608 ⋆ 1.254	.5487 ⋆ .5488 ⋆ 1.227
		Variance	.1231 ⋆ .1398 ⋆ .5899	**.0692** ⋆ .0826 ⋆ **.3122**	.0701 ⋆ **.0812** ⋆ .3336	.0738 ⋆ .0860 ⋆ .3479
		MSE	.9297 ⋆ .9561 ⋆ 4.671	**.3433** ⋆ **.3663** ⋆ **1.730**	.3796 ⋆ .3958 ⋆ 1.906	.3690 ⋆ .3872 ⋆ 1.853
.4	1	Estimate	[−1.97052, 2.38461]	[−1.32775, 1.56285]	[−1.25645, 1.44984]	[−1.16511, 1.37739]
		Bias	1.180 ⋆ 1.199 ⋆ 2.6479	**.7025** ⋆ **.7141** ⋆ **1.576**	.7450 ⋆ .7523 ⋆ 1.669	.7279 ⋆ .7368 ⋆ 1.631
		Variance	.1761 ⋆ .2054 ⋆ .8566	**.1069** ⋆ **.1270** ⋆ **.4972**	.1224 ⋆ .1394 ⋆ .6038	.1337 ⋆ .1526 ⋆ .6593
		MSE	1.562 ⋆ 1.644 ⋆ 7.868	**.5918** ⋆ **.6370** ⋆ **2.981**	.6755 ⋆ .7053 ⋆ 3.389	.6626 ⋆ .6956 ⋆ 3.321
.4	5	Estimate	[−2.28312, 4.42717]	[−1.25364, 2.16264]	[−1.17114, 1.96028]	[−0.99879, 1.83575]
		Bias	2.357 ⋆ 2.592 ⋆ 5.381	.9654 ⋆ 1.071 ⋆ 2.208	.9575 ⋆ 1.038 ⋆ 2.178	**.8739** ⋆ **.9725** ⋆ **2.002**
		Variance	.7400 ⋆ .9841 ⋆ 3.455	**.2185** ⋆ **.2674** ⋆ **.9993**	.2525 ⋆ .3017 ⋆ 1.246	.2496 ⋆ .3057 ⋆ 1.217
		MSE	6.482 ⋆ 7.707 ⋆ 32.41	1.195 ⋆ 1.415 ⋆ 5.875	1.224 ⋆ 1.380 ⋆ 5.993	**1.099** ⋆ **1.251** ⋆ **5.218**
.4	10	Estimate	[−2.74184, 6.92206]	[−1.22611, 2.31048]	[−1.11704, 2.11321]	[−.91925, 2.01288]
		Bias	3.834 ⋆ 4.370 ⋆ 8.8273	1.025 ⋆ 1.164 ⋆ 2.358	1.006 ⋆ 1.126 ⋆ 2.307	**.9227** ⋆ **1.076** ⋆ **2.136**
		Variance	2.049 ⋆ 2.769 ⋆ 8.568	**.2214** ⋆ **.2793** ⋆ **.8894**	.2558 ⋆ .3231 ⋆ 1.212	.2665 ⋆ .3418 ⋆ 1.187
		MSE	17.66 ⋆ 21.87 ⋆ 86.49	1.337 ⋆ 1.636 ⋆ 6.453	1.368 ⋆ 1.593 ⋆ 6.538	**1.278** ⋆ **1.501** ⋆ **5.753**
.4	100	Estimate	[−11.43411, 52.31994]	[−1.24750, 2.32713]	[−1.09851, 2.19196]	[−.91253, 2.14301]
		Bias	30.87 ⋆ 37.03 ⋆ 72.01	1.044 ⋆ 1.180 ⋆ 2.399	1.037 ⋆ 1.176 ⋆ 2.384	**.9844** ⋆ **1.165** ⋆ **2.287**
		Variance	162.1 ⋆ 217.9 ⋆ 565.4	**.2281** ⋆ **.2897** ⋆ **.9339**	.2798 ⋆ .3578 ⋆ 1.323	.3052 ⋆ .3949 ⋆ 1.343
		MSE	1220 ⋆ 1589 ⋆ 5751	**1.382** ⋆ **1.683** ⋆ 6.692	1.477 ⋆ 1.740 ⋆ 7.009	1.470 ⋆ 1.752 ⋆ **6.577**

Table 3.7: Monte Carlo approximation, bias, variance and mean squared error of the location measures in CASE 1

c_p	c_D	feature	Mean	ρ_1-Median	$d_{\theta=1}$-Median	$d_{\theta=5}$-Median
.0	0	Estimate	[−.60357, .59748]	[−1.04536, 1.04316]	[−.78298, .77889]	[−.70305, .69865]
		Bias	0 ⋆ 0 ⋆ 0	0 ⋆ 0 ⋆ 0	0 ⋆ 0 ⋆ 0	0 ⋆ 0 ⋆ 0
		Variance	**.0102 ⋆ .0111 ⋆ .0161**	.0108 ⋆ .0128 ⋆ .0226	.0182 ⋆ .0196 ⋆ .0275	.0191 ⋆ .0207 ⋆ .0298
		MSE	**.0102 ⋆ .0111 ⋆ .0161**	.0108 ⋆ .0128 ⋆ .0226	.0182 ⋆ .0196 ⋆ .0275	.0191 ⋆ .0207 ⋆ .0298
.1	0	Estimate	[−.59104, .59191]	[−1.03738, 1.03652]	[−.77475, .77271]	[−.69428, .69089]
		Bias	.0084 ⋆ .0084 ⋆ .0189	**.0046 ⋆ .0047 ⋆ .0103**	.0078 ⋆ .0079 ⋆ .0175	.0082 ⋆ .0083 ⋆ .0184
		Variance	.0159 ⋆ .0169 ⋆ **.0222**	**.0129 ⋆ .0152** ⋆ .0262	.0210 ⋆ .0227 ⋆ .0317	.0225 ⋆ .0243 ⋆ .0340
		MSE	.0159 ⋆ .0170 ⋆ **.0226**	**.0128 ⋆ .0152** ⋆ .0263	.0209 ⋆ .0227 ⋆ .0320	.0225 ⋆ .0244 ⋆ .0344
.1	1	Estimate	[−.60679, .72450]	[−1.04268, 1.07424]	[−.78818, .83681]	[−.71048, .76413]
		Bias	.0656 ⋆ .0879 ⋆ .1581	**.0168 ⋆ .0227 ⋆ .0407**	.0309 ⋆ .0392 ⋆ .0733	.0364 ⋆ .0455 ⋆ .0859
		Variance	.0175 ⋆ .0189 ⋆ **.0258**	**.0110 ⋆ .0135** ⋆ .0262	.0198 ⋆ .0216 ⋆ .0309	.0202 ⋆ .0222 ⋆ .0328
		MSE	.0230 ⋆ .0266 ⋆ .0508	**.0119 ⋆ .0140 ⋆ .0279**	.0210 ⋆ .0232 ⋆ .0362	.0218 ⋆ .0243 ⋆ .0402
.1	5	Estimate	[−.68677, 1.21743]	[−1.04988, 1.11860]	[−.77104, .89670]	[−.69004, .81353]
		Bias	.3521 ⋆ .4407 ⋆ .8308	**.0426 ⋆ .0544 ⋆ .1012**	.0626 ⋆ .0816 ⋆ .1327	.0621 ⋆ .0803 ⋆ .1297
		Variance	.0378 ⋆ .0467 ⋆ .1054	**.0075 ⋆ .0102 ⋆ .0234**	.0188 ⋆ .0207 ⋆ .0311	.0190 ⋆ .0210 ⋆ .0327
		MSE	.1839 ⋆ .2410 ⋆ .7957	**.0114 ⋆ .0132 ⋆ .0336**	.0244 ⋆ .0273 ⋆ .0487	.0245 ⋆ .0275 ⋆ .0496
.1	10	Estimate	[−.77077, 1.85755]	[−1.04035, 1.11780]	[−.76427, .90253]	[−.68559, .81644]
		Bias	.7142 ⋆ .8972 ⋆ 1.686	**.0381 ⋆ .0534 ⋆ .0921**	.0689 ⋆ .0863 ⋆ .1349	.0658 ⋆ .0827 ⋆ .1301
		Variance	.1215 ⋆ .1622 ⋆ .3910	**.0087 ⋆ .0116 ⋆ .0259**	.0210 ⋆ .0230 ⋆ .0335	.0211 ⋆ .02332 ⋆ .0349
		MSE	.7326 ⋆ .9672 ⋆ 3.236	**.0126 ⋆ .0145 ⋆ .0344**	.0274 ⋆ .0305 ⋆ .0517	.0270 ⋆ .0301 ⋆ .0519
.1	100	Estimate	[−2.29196, 13.43840]	[−1.04257, 1.11976]	[−.76356, .90700]	[−.68530, .81891]
		Bias	7.265 ⋆ 9.156 ⋆ 17.17	**.0395 ⋆ .0548 ⋆ .0963**	.0715 ⋆ .0895 ⋆ .1398	.0672 ⋆ .0845 ⋆ .1329
		Variance	10.33 ⋆ 14.39 ⋆ 38.85	**.0076 ⋆ .0101 ⋆ .0225**	.0196 ⋆ .0216 ⋆ .0316	.0205 ⋆ .0227 ⋆ .0341
		MSE	73.74 ⋆ 98.23 ⋆ 333.8	**.0114 ⋆ .0131 ⋆ .0318**	.0266 ⋆ .0296 ⋆ .0512	.0268 ⋆ .0298 ⋆ .0518
.2	0	Estimate	[−.58781, .57649]	[−1.03204, 1.02691]	[−.76973, .76150]	[−.68802, .67912]
		Bias	.0178 ⋆ .0187 ⋆ .0402	**.0120 ⋆ .0124 ⋆ .0272**	.0159 ⋆ .0164 ⋆ .0358	.0172 ⋆ .0177 ⋆ .0388
		Variance	.0182 ⋆ .0194 ⋆ **.0254**	**.0143 ⋆ .0166** ⋆ .0286	.0218 ⋆ .0236 ⋆ .0325	.0236 ⋆ .0255 ⋆ .0356
		MSE	.0183 ⋆ .0198 ⋆ **.0271**	**.0140 ⋆ .0168** ⋆ .0294	.0218 ⋆ .0238 ⋆ .0338	.0236 ⋆ .0258 ⋆ .0371
.2	1	Estimate	[−.61501, .83663]	[−1.03871, 1.10449]	[−.79484, .89450]	[−.71970, .83118]
		Bias	.1258 ⋆ .1674 ⋆ .3023	**.0322 ⋆ .0440 ⋆ .0745**	.0631 ⋆ .0803 ⋆ .1496	.0746 ⋆ .0933 ⋆ .1760
		Variance	.0242 ⋆ .0267 ⋆ .0410	**.0105 ⋆ .0136 ⋆ .0297**	.0206 ⋆ .0231 ⋆ .0366	.0204 ⋆ .0232 ⋆ .0391
		MSE	.0454 ⋆ .0548 ⋆ .1324	**.0134 ⋆ .0156 ⋆ .0352**	.0258 ⋆ .0296 ⋆ .0590	.0275 ⋆ .0319 ⋆ .0701
.2	5	Estimate	[−.77078, 1.83803]	[−1.03638, 1.17067]	[−.75380, 1.01836]	[−.67919, .93675]
		Bias	.7044 ⋆ .8835 ⋆ 1.663	**.0665 ⋆ .0909 ⋆ .1536**	.1321 ⋆ .1684 ⋆ .2685	.1292 ⋆ .1678 ⋆ .2721
		Variance	.1188 ⋆ .1555 ⋆ .3681	**.0091 ⋆ .0148 ⋆ .0404**	.0210 ⋆ .0244 ⋆ .0416	.0214 ⋆ .0244 ⋆ .0412
		MSE	.7142 ⋆ .9362 ⋆ 3.133	**.0200 ⋆ .0230 ⋆ .0640**	.0459 ⋆ .0528 ⋆ .1137	.0453 ⋆ .0526 ⋆ .1152
.2	10	Estimate	[−1.04070, 3.10648]	[−1.04170, 1.17202]	[−.75429, 1.02898]	[−.68130, .94562]
		Bias	1.473 ⋆ 1.799 ⋆ 3.453	**.0652 ⋆ .0918 ⋆ .1596**	.1372 ⋆ .1759 ⋆ .2818	.1325 ⋆ .1739 ⋆ .2846
		Variance	.4080 ⋆ .5437 ⋆ 1.364	**.0096 ⋆ .0155 ⋆ .0419**	.0221 ⋆ .0259 ⋆ .0437	.0229 ⋆ .0261 ⋆ .0428
		MSE	2.887 ⋆ 3.781 ⋆ 13.28	**.0207 ⋆ .0239 ⋆ .0674**	.0488 ⋆ .0568 ⋆ .1231	.0482 ⋆ .0563 ⋆ .1238
.2	100	Estimate	[−4.86615, 26.03105]	[−1.03277, 1.17267]	[−.74618, 1.04094]	[−.67600, .95717]
		Bias	14.84 ⋆ 18.23 ⋆ 34.84	**.0693 ⋆ .0924 ⋆ .1533**	.1472 ⋆ .1850 ⋆ .2905	.1410 ⋆ .1824 ⋆ .2947
		Variance	40.36 ⋆ 54.02 ⋆ 137.2	**.0106 ⋆ .0172** ⋆ .0464	.0223 ⋆ .0262 ⋆ .0445	.0230 ⋆ .0263 ⋆ **.0435**
		MSE	293.4 ⋆ 386.4 ⋆ 1351	**.0221 ⋆ .0258 ⋆ .0700**	.0524 ⋆ .0605 ⋆ .1289	.0513 ⋆ .0596 ⋆ .1304
.4	0	Estimate	[−.56563, .56857]	[−1.01743, 1.01583]	[−.75167, .74884]	[−.67084, .66527]
		Bias	.0328 ⋆ .0328 ⋆ .0734	**.0249 ⋆ .0249 ⋆ .0557**	.0312 ⋆ .0313 ⋆ .0699	.0327 ⋆ .0328 ⋆ .0733
		Variance	.0278 ⋆ .0293 ⋆ **.0372**	**.0182 ⋆ .0214** ⋆ .0381	.0261 ⋆ .0281 ⋆ .0391	.0283 ⋆ .0304 ⋆ .0421
		MSE	.0280 ⋆ .0303 ⋆ .0426	**.0177 ⋆ .0221 ⋆ .0412**	.0262 ⋆ .0291 ⋆ .0440	.0285 ⋆ .0315 ⋆ .0474
.4	1	Estimate	[−.65245, 1.07968]	[−1.04341, 1.16200]	[−.83031, 1.02147]	[−.77495, .98004]
		Bias	.2661 ⋆ .3410 ⋆ .6321	**.0611 ⋆ .0847 ⋆ .1487**	.1443 ⋆ .1730 ⋆ .3365	.1766 ⋆ .2044 ⋆ .4082
		Variance	.0405 ⋆ .0472 ⋆ .0856	**.0118 ⋆ .0177 ⋆ .0458**	.0225 ⋆ .0273 ⋆ .0547	.0216 ⋆ .0268 ⋆ .0578
		MSE	.1328 ⋆ .1635 ⋆ .4852	**.0215 ⋆ .0248 ⋆ .0680**	.0489 ⋆ .0572 ⋆ .1680	.0582 ⋆ .0686 ⋆ .2245
.4	5	Estimate	[−1.04593, 3.06623]	[−1.02759, 1.37928]	[−.77180, 1.29272]	[−.71840, 1.18035]
		Bias	1.456 ⋆ 1.772 ⋆ 3.409	**.1752 ⋆ .2385 ⋆ .4021**	.2603 ⋆ .3614 ⋆ .6181	.2485 ⋆ .3395 ⋆ .6020
		Variance	.4217 ⋆ .5584 ⋆ 1.326	**.0382** ⋆ .0590 ⋆ .1780	.0450 ⋆ .0608 ⋆ .1342	.0389 ⋆ **.0486 ⋆ .0917**
		MSE	2.848 ⋆ 3.698 ⋆ 12.94	**.0971 ⋆ .1159 ⋆ .3398**	.1625 ⋆ .1914 ⋆ .5163	.1379 ⋆ .1639 ⋆ .4541
.4	10	Estimate	[−1.44327, 5.54909]	[−.97458, 1.40392]	[−.71332, 1.35805]	[−.67236, 1.22567]
		Bias	2.896 ⋆ 3.549 ⋆ 6.793	**.2140 ⋆ .2600 ⋆ .3935**	.3222 ⋆ .4103 ⋆ .6532	.2770 ⋆ .3719 ⋆ .6202
		Variance	1.634 ⋆ 2.220 ⋆ 5.563	**.0547** ⋆ .0842 ⋆ .2370	.0649 ⋆ .0857 ⋆ .1770	.0592 ⋆ **.0709 ⋆ .1201**
		MSE	11.51 ⋆ 14.82 ⋆ 51.71	**.1291 ⋆ .1518 ⋆ .3919**	.2224 ⋆ .2541 ⋆ .6037	.1830 ⋆ .2093 ⋆ .5049
.4	100	Estimate	[−9.51527, 50.95382]	[−.98241, 1.40652]	[−.70326, 1.38340]	[−.67636, 1.23637]
		Bias	29.63 ⋆ 36.15 ⋆ 69.42	**.2114 ⋆ .2609 ⋆ .4020**	.3399 ⋆ .4290 ⋆ .6769	.2804 ⋆ .3793 ⋆ .6364
		Variance	155.2 ⋆ 206.6 ⋆ 521.3	**.0506** ⋆ .0789 ⋆ .2232	.0655 ⋆ .0856 ⋆ .1746	.0574 ⋆ **.0683 ⋆ .1148**
		MSE	1153 ⋆ 1514 ⋆ 5341	**.1234 ⋆ .1470 ⋆ .3848**	.2324 ⋆ .2697 ⋆ .6328	.1823 ⋆ .2122 ⋆ .5199

Table 3.8: Monte Carlo approximation, bias, variance and mean squared error of the location measures in CASE 2

c_p	c_D	feature	Mean	$\boldsymbol{\rho_1}$-Median	$d_{\theta=1}$-Median	$d_{\theta=5}$-Median
.0	0	Estimate	[−1.45257, 1.45454]	[−1.59709, 1.59680]	[−1.46444, 1.45955]	[−1.41036, 1.40514]
		Bias	0 ⋆ 0 ⋆ 0	0 ⋆ 0 ⋆ 0	0 ⋆ 0 ⋆ 0	0 ⋆ 0 ⋆ 0
		Variance	.0121 ⋆ .0143 ⋆ **.0303**	**.0116** ⋆ **.0139** ⋆ .0332	.0140 ⋆ .0164 ⋆ .0339	.0134 ⋆ .0157 ⋆ .0341
		MSE	.0121 ⋆ .0143 ⋆ **.0303**	**.0116** ⋆ **.0139** ⋆ .0332	.0140 ⋆ .0164 ⋆ .0339	.0134 ⋆ .0157 ⋆ .0341
.1	0	Estimate	[−1.43436, 1.45139]	[−1.58183, 1.59857]	[−1.44531, 1.46310]	[−1.39227, 1.40871]
		Bias	.0106 ⋆ .0130 ⋆ .0250	**.0085** ⋆ **.0108** ⋆ **.0173**	.0113 ⋆ .0137 ⋆ .0207	.0108 ⋆ .0130 ⋆ .0195
		Variance	.0167 ⋆ .0192 ⋆ .0369	**.0125** ⋆ **.0151** ⋆ .0369	.0143 ⋆ .0168 ⋆ **.0353**	.0137 ⋆ .0162 ⋆ .0354
		MSE	.0168 ⋆ .0193 ⋆ .0375	**.0126** ⋆ **.0152** ⋆ .0372	.0144 ⋆ .0170 ⋆ **.0358**	.0138 ⋆ .0164 ⋆ .0358
.1	1	Estimate	[−1.45938, 1.57164]	[−1.61134, 1.65225]	[−1.48605, 1.52761]	[−1.43292, 1.47959]
		Bias	.0619 ⋆ .0829 ⋆ .1491	**.0348** ⋆ **.0404** ⋆ **.0806**	.0448 ⋆ .0504 ⋆ .1029	.0485 ⋆ .0550 ⋆ .1115
		Variance	.0169 ⋆ .0194 ⋆ **.0391**	**.0132** ⋆ **.0161** ⋆ .0408	.0157 ⋆ .0186 ⋆ .0411	.0151 ⋆ .0179 ⋆ .0417
		MSE	.0222 ⋆ .0263 ⋆ .0613	**.0145** ⋆ **.0178** ⋆ **.0473**	.0178 ⋆ .0212 ⋆ .0517	.0175 ⋆ .0209 ⋆ .0541
.1	5	Estimate	[−1.53508, 2.07426]	[−1.62057, 1.71687]	[−1.48329, 1.59478]	[−1.42656, 1.53334]
		Bias	.3511 ⋆ .4420 ⋆ .8297	**.0717** ⋆ **.0865** ⋆ **.1676**	.0770 ⋆ .0965 ⋆ .1818	.0722 ⋆ .0913 ⋆ .1708
		Variance	.0429 ⋆ .0535 ⋆ .1209	**.0151** ⋆ **.0179** ⋆ .0436	.0173 ⋆ .0203 ⋆ **.0432**	.0169 ⋆ .0199 ⋆ .0437
		MSE	.1932 ⋆ .2489 ⋆ .8095	**.0212** ⋆ **.0254** ⋆ **.0717**	.0250 ⋆ .0296 ⋆ .0763	.0237 ⋆ .0282 ⋆ .0729
.1	10	Estimate	[−1.59340, 2.70553]	[−1.60011, 1.72678]	[−1.45983, 1.61214]	[−1.40402, 1.54998]
		Bias	.6959 ⋆ .8901 ⋆ 1.652	**.0665** ⋆ **.0919** ⋆ **.1616**	.0785 ⋆ .1079 ⋆ .1831	.0755 ⋆ .1025 ⋆ .1723
		Variance	.1210 ⋆ .1607 ⋆ .4100	**.0140** ⋆ **.0172** ⋆ .0453	.0157 ⋆ .0189 ⋆ .0431	.0152 ⋆ .0182 ⋆ **.0426**
		MSE	.7218 ⋆ .9531 ⋆ 3.139	**.0210** ⋆ **.0257** ⋆ **.0714**	.0254 ⋆ .0305 ⋆ .0766	.0239 ⋆ .0287 ⋆ .0723
.1	100	Estimate	[−3.48429, 14.28107]	[−1.61363, 1.72107]	[−1.47242, 1.60739]	[−1.41681, 1.54443]
		Bias	7.429 ⋆ 9.182 ⋆ 17.46	**.0704** ⋆ **.0886** ⋆ **.1663**	.0779 ⋆ .1046 ⋆ .1877	.0728 ⋆ .0986 ⋆ .1759
		Variance	9.728 ⋆ 13.31 ⋆ 34.32	**.0142** ⋆ **.0171** ⋆ **.0431**	.0174 ⋆ .0205 ⋆ .0454	.0168 ⋆ .0200 ⋆ .0457
		MSE	72.81 ⋆ 97.63 ⋆ 339.4	**.0207** ⋆ **.0249** ⋆ **.0708**	.0261 ⋆ .0314 ⋆ .0807	.0245 ⋆ .0297 ⋆ .0767
.2	0	Estimate	[−1.43149, 1.44011]	[−1.58705, 1.59286]	[−1.44542, 1.45640]	[−1.39032, 1.40123]
		Bias	.0177 ⋆ .0180 ⋆ .0398	**.0069** ⋆ **.0076** ⋆ **.0159**	.0110 ⋆ .0136 ⋆ .0260	.0119 ⋆ .0144 ⋆ .0279
		Variance	.0197 ⋆ .0225 ⋆ .0415	**.0132** ⋆ **.0159** ⋆ .0393	.0151 ⋆ .0178 ⋆ **.0374**	.0146 ⋆ .0172 ⋆ .0374
		MSE	.0198 ⋆ .0229 ⋆ .0431	**.0133** ⋆ **.0160** ⋆ .0396	.0152 ⋆ .0180 ⋆ **.0380**	.0147 ⋆ .0174 ⋆ .0382
.2	1	Estimate	[−1.49435, 1.68002]	[−1.65003, 1.71026]	[−1.53713, 1.60213]	[−1.48452, 1.55883]
		Bias	.1336 ⋆ .1621 ⋆ .3126	**.0832** ⋆ **.0885** ⋆ **.1884**	.1076 ⋆ .1131 ⋆ .2432	.1139 ⋆ .1206 ⋆ .2578
		Variance	.0241 ⋆ .0276 ⋆ **.0514**	**.0170** ⋆ **.0202** ⋆ .0518	.0189 ⋆ .0222 ⋆ .0516	.0190 ⋆ .0223 ⋆ .0520
		MSE	.0437 ⋆ .0538 ⋆ .1492	**.0232** ⋆ **.0281** ⋆ **.0874**	.0288 ⋆ .0350 ⋆ .1107	.0301 ⋆ .0369 ⋆ .1185
.2	5	Estimate	[−1.65137, 2.66631]	[−1.65785, 1.87142]	[−1.51606, 1.76165]	[−1.44957, 1.69308]
		Bias	.7052 ⋆ .8683 ⋆ 1.656	.1676 ⋆ **.1988** ⋆ .3899	.1768 ⋆ .2167 ⋆ .4148	**.1635** ⋆ .2054 ⋆ **.3863**
		Variance	.1168 ⋆ .1539 ⋆ .3826	**.0209** ⋆ **.0265** ⋆ .0743	.0237 ⋆ .0294 ⋆ .0734	.0229 ⋆ .0282 ⋆ **.0684**
		MSE	.7064 ⋆ .9078 ⋆ 3.126	**.0551** ⋆ **.0661** ⋆ .2263	.0631 ⋆ .0764 ⋆ .2455	.0579 ⋆ .0704 ⋆ **.2177**
.2	10	Estimate	[−1.81346, 3.97370]	[−1.63723, 1.89056]	[−1.49540, 1.79558]	[−1.42822, 1.72534]
		Bias	1.440 ⋆ 1.799 ⋆ 3.396	**.1669** ⋆ **.2096** ⋆ **.3942**	.1834 ⋆ .2386 ⋆ .4377	.1690 ⋆ .2267 ⋆ .4070
		Variance	.4002 ⋆ .5338 ⋆ 1.340	**.0225** ⋆ **.0279** ⋆ .0771	.0257 ⋆ .0316 ⋆ .0786	.0253 ⋆ .0310 ⋆ **.0750**
		MSE	2.863 ⋆ 3.772 ⋆ 12.87	**.0589** ⋆ **.0718** ⋆ **.2326**	.0718 ⋆ .0885 ⋆ .2702	.0669 ⋆ .0824 ⋆ .2407
.2	100	Estimate	[−6.03844, 26.72992]	[−1.64179, 1.88729]	[−1.49645, 1.79678]	[−1.43480, 1.71998]
		Bias	14.93 ⋆ 18.16 ⋆ 34.95	**.1675** ⋆ **.2078** ⋆ **.3943**	.1846 ⋆ .2395 ⋆ .4401	.1696 ⋆ .2232 ⋆ .4061
		Variance	41.43 ⋆ 56.20 ⋆ 144.0	**.0222** ⋆ **.0280** ⋆ .0808	.0263 ⋆ .0328 ⋆ .0839	.0254 ⋆ .0316 ⋆ **.0787**
		MSE	294.6 ⋆ 386.1 ⋆ 1365	**.0591** ⋆ **.0712** ⋆ **.2363**	.0738 ⋆ .0902 ⋆ .2776	.0668 ⋆ .0814 ⋆ .2437
.4	0	Estimate	[−1.41693, 1.42175]	[−1.58826, 1.58947]	[−1.44519, 1.44446]	[−1.38863, 1.38977]
		Bias	.0342 ⋆ .0342 ⋆ .0765	**.0080** ⋆ **.0081** ⋆ **.0180**	.0171 ⋆ .0172 ⋆ .0384	.0185, ⋆ .0188 ⋆ .0415
		Variance	.0292 ⋆ .0326 ⋆ .0535	**.0158** ⋆ **.0190** ⋆ **.0421**	.0196 ⋆ .0225 ⋆ .0424	.0191 ⋆ .0220 ⋆ .0428
		MSE	.0296 ⋆ .0337 ⋆ .0593	**.0158** ⋆ **.0191** ⋆ **.0424**	.0196 ⋆ .0228 ⋆ .0438	.0192 ⋆ .0224 ⋆ .0445
.4	1	Estimate	[−1.49881, 1.92425]	[−1.69753, 1.88527]	[−1.60777, 1.79981]	[−1.55359, 1.77199]
		Bias	.2579 ⋆ .3337 ⋆ .6144	**.1944** ⋆ **.2159** ⋆ **.4448**	.2417 ⋆ .2610 ⋆ .5495	.2550 ⋆ .2784 ⋆ .5811
		Variance	.0461 ⋆ .0542 ⋆ .1055	**.0308** ⋆ **.0375** ⋆ **.1034**	.0328 ⋆ .0403 ⋆ .1112	.0344 ⋆ .0419 ⋆ .1135
		MSE	.1357 ⋆ .1656 ⋆ .4831	**.0706** ⋆ **.0842** ⋆ **.3013**	.0920 ⋆ .1084 ⋆ .4132	.1005 ⋆ .1195 ⋆ .4513
.4	5	Estimate	[−1.85535, 3.91731]	[−1.75626, 2.45461]	[−1.64767, 2.33840]	[−1.54410, 2.25335]
		Bias	1.432 ⋆ 1.764 ⋆ 3.365	.5084 ⋆ .6169 ⋆ 1.189	.5310 ⋆ .6348 ⋆ 1.237	**.4909** ⋆ **.6071** ⋆ **1.154**
		Variance	.4226 ⋆ .5622 ⋆ 1.356	.0839 ⋆ .1134 ⋆ .3590	.0878 ⋆ .1198 ⋆ .3638	**.0821** ⋆ **.1108** ⋆ **.3100**
		MSE	2.842 ⋆ 3.675 ⋆ 12.68	.4069 ⋆ .4940 ⋆ 1.773	.4374 ⋆ .5228 ⋆ 1.894	**.3954** ⋆ **.4795** ⋆ **1.642**
.4	10	Estimate	[−2.43213, 6.41147]	[−1.75240, 2.51641]	[−1.65208, 2.44767]	[−1.53502, 2.35117]
		Bias	2.968 ⋆ 3.572 ⋆ 6.928	.5374 ⋆ **.6594** ⋆ **1.261**	.5878 ⋆ .7111 ⋆ 1.374	**.5353** ⋆ .6747 ⋆ 1.265
		Variance	1.628 ⋆ 2.167 ⋆ 5.269	**.0935** ⋆ **.1291** ⋆ .3782	.1074 ⋆ .1468 ⋆ .4089	.1077 ⋆ .1421 ⋆ **.3523**
		MSE	11.58 ⋆ 14.93 ⋆ 53.27	**.4630** ⋆ **.5640** ⋆ 1.968	.5444 ⋆ .6526 ⋆ 2.297	.4956 ⋆ .5974 ⋆ **1.953**
.4	100	Estimate	[−9.57178, 52.00442]	[−1.70707, 2.51529]	[−1.59118, 2.48420]	[−1.48280, 2.38771]
		Bias	29.33 ⋆ 36.20 ⋆ 68.93	**.5142** ⋆ **.6541** ⋆ **1.218**	.5757 ⋆ .7300 ⋆ 1.363	.5275 ⋆ .6966 ⋆ 1.264
		Variance	152.8 ⋆ 203.9 ⋆ 520.4	**.0837** ⋆ **.1131** ⋆ **.3304**	.1026 ⋆ .1397 ⋆ .3827	.1061 ⋆ .1377 ⋆ .3336
		MSE	1139 ⋆ 1514 ⋆ 5273	**.4331** ⋆ **.5409** ⋆ **1.8160**	.5438 ⋆ .6726 ⋆ 2.241	.5027 ⋆ .6230 ⋆ 1.932

Table 3.9: Monte Carlo approximation, bias, variance and mean squared error of the location measures in CASE 2'

		Spatial		
c_p	c_D	Case 1	Case 2	Case 2'
0.0	0	0.003245	0.001503	0.003451
0.1	0	0.004349	0.003103	0.003176
0.1	1	0.003040	0.002462	0.003303
0.1	5	0.003773	0.002684	0.002879
0.1	10	0.002415	0.004036	0.004215
0.1	100	0.002882	0.003122	0.004140
0.2	0	0.004299	0.003837	0.003091
0.2	1	0.005382	0.004384	0.003277
0.2	5	0.003776	0.004167	0.004741
0.2	10	0.002993	0.005597	0.004991
0.2	100	0.003904	0.005066	0.004936
0.4	0	0.003310	0.004491	0.003683
0.4	1	0.0036030	0.005110	0.003415
0.4	5	0.005399	0.004911	0.004935
0.4	10	0.004122	0.005416	0.004658
0.4	100	0.003382	0.006818	0.004225

Table 3.10: Maximum range of the weights (i.e., the difference between the maximum and minimum allocated weights) in the computation of the spatial median considering 11 different values of θ

Remark 3.5.1. Although the parameter θ is involved in the computation of the spatial median, its influence on the estimation is modest as it has been checked empirically with the following simulation study:

Step 1 is similar to that for the simulations already developed in this subsection;

Step 2. For each situation (c_p and C_D), the spatial median has been approximated by Monte Carlo, using $N = 1000$ replications of *Step 1*, for some different choices of the parameter $\theta \in \{0.1, 1/3, 0.5, 2/3, 0.9, 1, 5, 10, 20, 50, 100\}$.

Step 3. The range of each of the 100 weights (corresponding to each of the 100 observations) has been computed over the 11 possible values of θ. Finally, for each situation (c_p and C_D), the maximum of the 100 ranges is represented in Table 3.10.

It can be seen that, independently from the case, the maximum range obtained is quite small, always smaller than 1/100. Therefore, none of the weights vary very much when θ changes.

One should not worry about this limited influence of θ on the estimation of the spatial median, since the Aumann mean and the wabl/ldev/rdev median, in spite of being defined by means of a distance that also involves the theta parameter, do not depend on theta.

3.6 Illustrative application to real-life example

A real-life example involving interval-valued data is first to be considered, aiming to compare the behaviour of the Aumann mean value and the different approaches to extend the median in this chapter.

Example 3.6.1. The IBEX 35 is the benchmark stock market index of the Bolsa de Madrid, Spain's principal stock exchange. Administered and calculated by Sociedad de Bolsas (a subsidiary of the company which runs Spain's securities markets including the Bolsa de Madrid), it is composed of the 35 securities listed on the Stock Exchange Interconnection System of the four Spanish Stock Exchanges, which were most liquid during the control period pursuant to the terms of the regulation. For more details, visit the web page of the Bolsa de Madrid (http://www.bolsamadrid.es/ing/aspx/Portada/Portada.aspx).

In this example, the daily fluctuation of the IBEX 35 during 6 months (from the 14^{th} June to the 14^{th} December of 2012) has been considered. To measure such fluctuation, the minimum and the maximum values achieved by the index every day along this period of time have been recorded. Table 3.11 shows the data obtained and Figure 3.22 represents the daily fluctuation along the considered period of time.

By using the algorithm proposed in this section, the sample d_θ-median has been calculated:

$$\mathsf{M}^{(1)}_{\theta=1/3}[\mathsf{X}] = [7598.308, 7740.571],$$

and also the Aumann-type mean value and the location measures defined in this chapter (M-estimates of location, 1-norm median and φ-wabl/ldev/rdev median):

$$\mathsf{E}^{(1)}[\mathsf{X}] = [7338.07, 7502.121],$$

$$\mathsf{Me}^{(1)}[\mathsf{X}] = [7594.94, 7733.85],$$

$$\mathsf{M}^{\varphi(1)}[\mathsf{X}] = [7603.375, 7742.275],$$

$${\mathsf{g}^{\mathsf{M}}_{\text{Huber},\,\theta=1/3}}^{(1)}[\mathsf{X}] = [7598.022, 7740.281],$$

$${\mathsf{g}^{\mathsf{M}}_{\text{Hampel},\,\theta=1/3}}^{(1)}[\mathsf{X}] = [7665.158, 7804.915].$$

The M-estimates of location have been computed by means of two well-known loss functions satisfying the assumptions of the representer theorem (Theorem 3.2.1), the Huber and the Hampel loss functions.

First, the considered loss function will be Huber's one (Remark 3.1.1) with $a = 1.345$. The function ρ_a, its derivative ρ'_a and ϕ_a corresponding to the quotient of ρ'_a and the identity function are displayed in Figure 3.23 (top left).

The second one is Hampel's loss function, introduced in Subsection 3.1.5.

Date	Low	High	Date	Low	High	Date	Low	High
14/06/2012	6558.6	6696.4	15/08/2012	7066.8	7146.7	16/10/2012	7741.6	7941.8
15/06/2012	6659.0	6829.8	16/08/2012	7123.3	7419.3	17/10/2012	7987.0	8131.9
18/06/2012	6503.3	6862.6	17/08/2012	7452.5	7612.1	18/10/2012	8035.5	8156.6
19/06/2012	6479.4	6717.4	20/08/2012	7387.4	7645.0	19/10/2012	7880.9	8070.6
20/06/2012	6668.2	6806.9	21/08/2012	7430.4	7558.4	22/10/2012	7841.5	7955.7
21/06/2012	6675.4	6914.9	22/08/2012	7326.4	7516.0	23/10/2012	7699.8	7886.6
22/06/2012	6697.1	6960.0	23/08/2012	7176.9	7427.0	24/10/2012	7653.2	7819.7
25/06/2012	6612.3	6857.5	24/08/2012	7183.0	7323.1	25/10/2012	7766.3	7851.3
26/06/2012	6512.3	6708.3	27/08/2012	7216.8	7398.9	26/10/2012	7665.6	7803.7
27/06/2012	6513.0	6666.9	28/08/2012	7284.2	7404.1	29/10/2012	7684.5	7779.7
28/06/2012	6593.5	6724.5	29/08/2012	7274.3	7365.4	30/10/2012	7755.9	7844.0
29/06/2012	6872.5	7102.2	30/08/2012	7178.8	7304.7	31/10/2012	7830.9	7939.9
02/07/2012	7036.9	7178.4	31/08/2012	7161.4	7424.7	01/11/2012	7790.3	7928.8
03/07/2012	7122.2	7219.5	03/09/2012	7351.2	7442.1	02/11/2012	7834.6	7995.9
04/07/2012	7116.2	7202.7	04/09/2012	7444.2	7542.8	05/11/2012	7793.3	7894.7
05/07/2012	6912.2	7180.8	05/09/2012	7421.6	7562.4	06/11/2012	7800.9	7868.3
06/07/2012	6726.7	6921.2	06/09/2012	7533.0	7864.9	07/11/2012	7638.5	7918.6
09/07/2012	6611.6	6785.7	07/09/2012	7834.4	8027.0	08/11/2012	7606.4	7718.0
10/07/2012	6638.2	6812.2	10/09/2012	7796.9	7890.8	09/11/2012	7496.0	7674.0
11/07/2012	6679.0	6829.7	11/09/2012	7730.0	7933.4	12/11/2012	7548.1	7633.3
12/07/2012	6600.8	6755.3	12/09/2012	7928.7	8076.6	13/11/2012	7490.6	7715.5
13/07/2012	6563.3	6687.5	13/09/2012	7865.1	7968.3	14/11/2012	7642.8	7756.5
16/07/2012	6485.1	6659.2	14/09/2012	8087.5	8231.0	15/11/2012	7610.4	7744.3
17/07/2012	6529.7	6642.6	17/09/2012	8049.6	8156.6	16/11/2012	7588.2	7721.8
18/07/2012	6497.2	6607.8	18/09/2012	7949.1	8099.7	19/11/2012	7601.7	7775.1
19/07/2012	6579.5	6682.5	19/09/2012	8031.5	8149.4	20/11/2012	7695.5	7787.2
20/07/2012	6232.6	6668.1	20/09/2012	7961.4	8088.0	21/11/2012	7716.8	7825.4
23/07/2012	5905.3	6240.5	21/09/2012	8035.9	8230.7	22/11/2012	7809.0	7894.4
24/07/2012	5950.8	6254.6	24/09/2012	8082.9	8182.1	23/11/2012	7825.1	7915.6
25/07/2012	5939.4	6093.1	25/09/2012	8117.5	8197.9	26/11/2012	7842.6	7893.0
26/07/2012	5955.0	6368.8	26/09/2012	7840.4	8082.5	27/11/2012	7819.7	7955.2
27/07/2012	6244.2	6617.6	27/09/2012	7790.6	7914.8	28/11/2012	7753.7	7854.9
30/07/2012	6589.2	6805.5	28/09/2012	7678.7	7939.2	29/11/2012	7883.8	7975.5
31/07/2012	6691.1	6913.7	01/10/2012	7703.0	7838.5	30/11/2012	7922.7	7989.6
01/08/2012	6569.2	6787.2	02/10/2012	7710.2	7913.7	03/12/2012	7874.1	8027.8
02/08/2012	6364.8	6864.6	03/10/2012	7802.3	7900.8	04/12/2012	7870.9	7945.7
03/08/2012	6296.1	6755.7	04/10/2012	7784.4	7913.3	05/12/2012	7844.7	7976.7
06/08/2012	6738.4	7061.9	05/10/2012	7825.5	7972.8	06/12/2012	7841.8	7947.8
07/08/2012	7013.0	7218.7	08/10/2012	7855.2	7913.1	07/12/2012	7813.1	7942.6
08/08/2012	7039.8	7232.2	09/10/2012	7734.9	7901.5	10/12/2012	7670.8	7804.4
09/08/2012	7026.6	7235.8	10/10/2012	7664.5	7765.1	11/12/2012	7806.2	7927.0
10/08/2012	6962.2	7079.9	11/10/2012	7565.5	7777.7	12/12/2012	7927.3	7988.3
13/08/2012	6991.9	7119.8	12/10/2012	7652.4	7791.5	13/12/2012	7977.3	8037.0
14/08/2012	7074.3	7153.1	15/10/2012	7624.7	7723.4	14/12/2012	8003.6	8045.3

Table 3.11: Daily fluctuation of the IBEX 35 Index during 6 months (from 14/06/2012 to 14/12/2012) obtained from Bolsa de Madrid

To fix the three parameters involved in the Hampel loss function, the ideas in Kim and Scott [114] have been followed, that is,

1. to choose the initial seed in this case, the 1-norm median has been considered;
2. to compute the distances between the observations and the seed,

   $d_i = d_{\theta=1/3}(\mathsf{x}_i, \mathsf{Me}[\mathsf{X}])$ for all $i = 1, \ldots, n$;
3. a will be the median of $\{d_i\}_{i=1}^n$, b the 75^{th} percentile of $\{d_i\}_{i=1}^n$ and c, the 85^{th} percentile of $\{d_i\}_{i=1}^n$.

The distance involved in the algorithm to compute the M-estimates has been d_θ, because this choice makes the isometrical embedding to lie in a Hilbert space. The

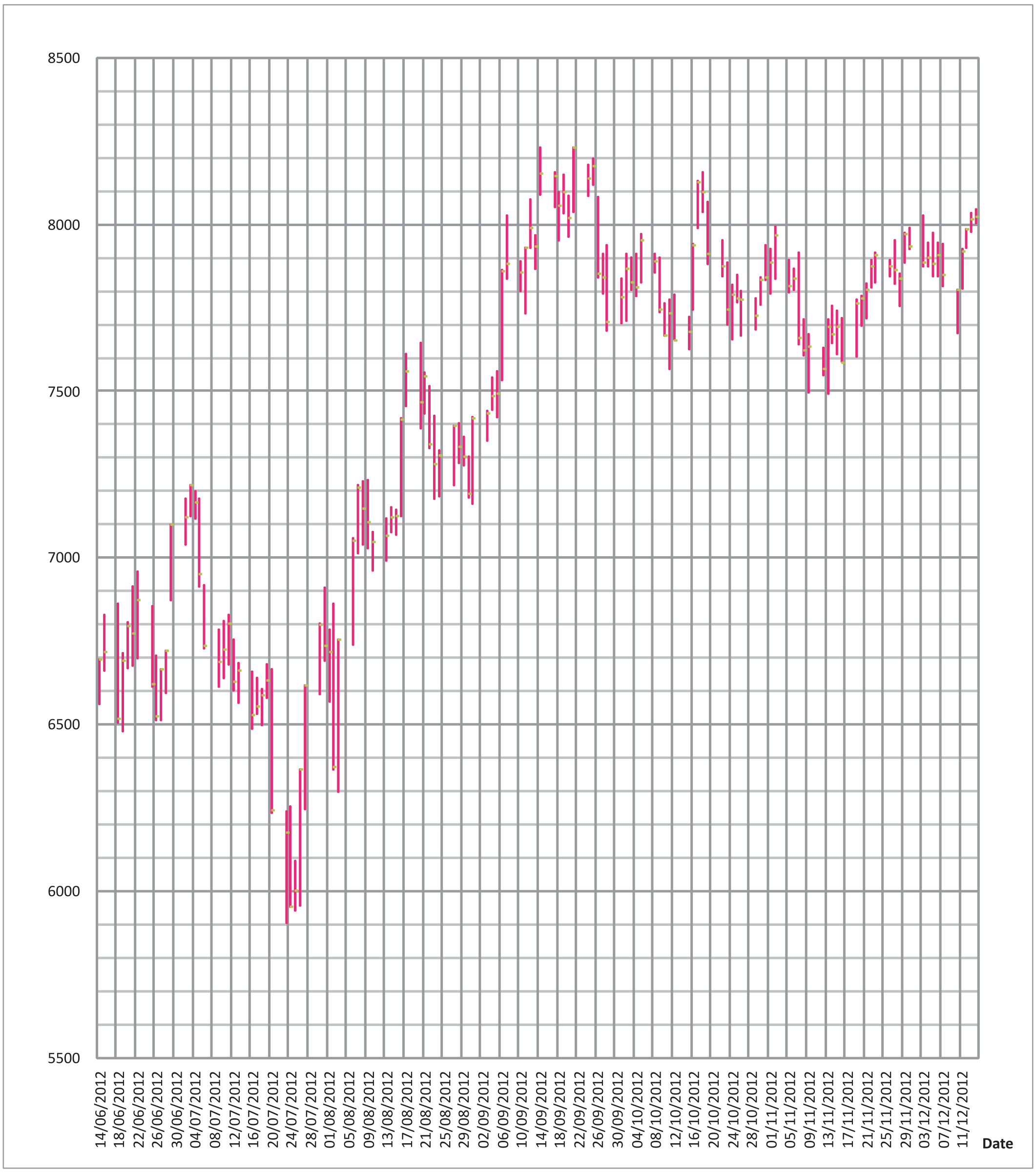


Figure 3.22: The daily fluctuation of the IBEX 35 Index from 14/06/2012 to 14/12/2012

parameter θ has been assumed to be equal to 1/3. So, we obtain that $a = 315.9961$, $b = 699.6943$ and $c = 984.5077$ and the graphics of $\rho_{a,b,c}$, its derivative and $\phi_{a,b,c}$ are displayed in Figure 3.23.

Observe in Figure 3.22 that at the end of July and the beginning of August the daily values of the IBEX 35 Index were much lower (looking at their midpoint they are outliers) and more variable (and looking at their spread too) than the remaining days of the period of time considered. Other days with a huge increase or decrease (big spread) can be also found, but at least these outliers have too much influence on the Aumann-type mean value, but not on the others location measures. Therefore, these last intervals will have a midpoint and a spread more similar to the data set

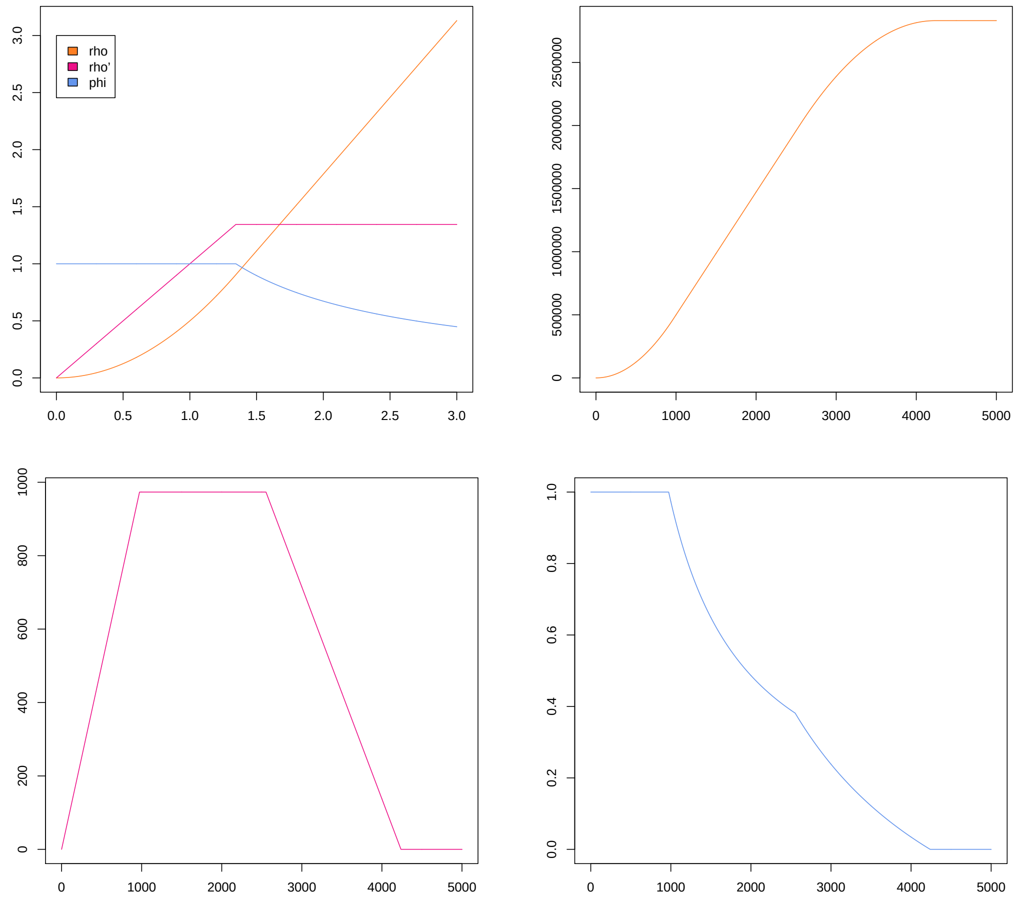


Figure 3.23: Huber loss function $\rho_{1.345}$, its derivative $\rho'_{1.345}$ and $\phi_{1.345}$ (top left), Hampel loss function $\rho_{a,b,c}$ (top right), its derivative $\rho'_{a,b,c}$ (bottom left) and $\phi_{a,b,c}$ (bottom right) with $a = 973.2789$, $b = 2552.638$ and $c = 4238.018$

(bigger midpoint and smaller spread than the Aumann-type mean value).

By removing, for instance, the most extreme values (those achieved on 23rd, 24th, 25th, 26th and 27th July, 2nd and 3rd August), the Aumann-type mean value is not so far from the sample medians:

$$\mathsf{M}^{(2)}_{\theta=1/3}[\mathsf{X}] = [7624.197, 7765.431],$$

$$\mathsf{E}^{(2)}[\mathsf{X}] = [7407.758, 7560.681],$$

$$\mathsf{Me}^{(2)}[\mathsf{X}] = [7624.7, 7775.1],$$

$$\mathsf{M}^{\varphi(2)}[\mathsf{X}] = [7630.05, 7767.85],$$

$${\mathsf{g}^{\mathsf{M}}_{\text{Huber}\,\theta=1/3}}^{(2)}[\mathsf{X}] = [7624.615, 7765.657],$$

$${\mathsf{g}^{\mathsf{M}}_{\text{Hampel}\,\theta=1/3}}^{(2)}[\mathsf{X}] = [7689.12, 7827.5].$$

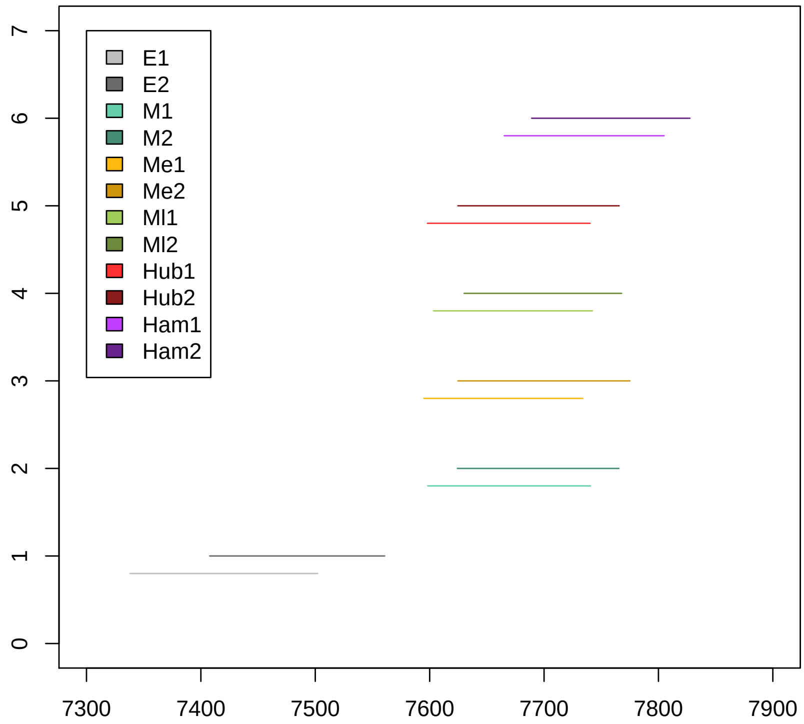


Figure 3.24: From bottom to top: Aumann-type mean, $d_{\theta=1/3}$-median, 1-norm median, φ-wabl/ldev/rdev median, Huber M-estimate and Hampel M-estimate (from each of them, the original estimate (bottom) and the estimate after the removal of the outliers (top))

Notice that for the new computation of the Hampel M-estimate, the new values of the three parameters are $a = 262.8862$, $b = 592.2556$ and $c = 958.2533$. In Figure 3.24, all the estimates before (light color) and after (dark color) the removal of the outliers are shown, in order to compare the robustness of the estimators visually.

It can be checked that, although outliers make all the measures be farther, the Aumann-type mean value is strongly perturbed by them, whereas the medians and M-estimates are more robust to these 'extreme' observations:

$$d_{\theta=1/3}(\mathsf{M}^{(1)}_{\theta=1/3}[\mathsf{X}], \mathsf{M}^{(2)}_{\theta=1/3}[\mathsf{X}]) = 25.37624,$$

$$d_{\theta=1/3}(\mathsf{E}^{(1)}[\mathsf{X}], \mathsf{E}^{(2)}[\mathsf{X}]) = 64.20441,$$

$$d_{\theta=1/3}(\mathsf{Me}^{(1)}[\mathsf{X}], \mathsf{Me}^{(2)}[\mathsf{X}]) = 35.65489,$$

$$d_{\theta=1/3}(\mathsf{M}^{\varphi(1)}[\mathsf{X}], \mathsf{M}^{\varphi(2)}[\mathsf{X}]) = 26.12693,$$

$$d_{\theta=1/3}({\mathsf{g}^{\mathsf{M}}_{\text{Huber}\,\theta=1/3}}^{(1)}[\mathsf{X}], {\mathsf{g}^{\mathsf{M}}_{\text{Huber}\,\theta=1/3}}^{(2)}[\mathsf{X}]) = 25.98687,$$

$$d_{\theta=1/3}({\mathsf{g}^{\mathsf{M}}_{\text{Hampel}\,\theta=1/3}}^{(1)}[\mathsf{X}], {\mathsf{g}^{\mathsf{M}}_{\text{Hampel}\,\theta=1/3}}^{(2)}[\mathsf{X}]) = 23.27689.$$

		Huber			Hampel		
c_p	c_D	Case 1	Case 2	Case 2’	Case 1	Case 2	Case 2’
.0	0	0.000745	0.000418	0.000675	0.00055542	0.00048121	0.00052294
0.1	0	0.005638	0.000649	0.000697	0.00491613	0.00501037	0.00535308
0.1	1	0.007181	0.006603	0.003999	0.00644601	0.00569424	0.00569457
0.1	5	0.005461	0.007320	0.007301	0.00544466	0.00236087	0.00300744
0.1	10	0.003215	0.004077	0.004002	0.00544466	0.00236087	0.00300744
0.1	100	0.000700	0.000446	0.000775	0.00054933	0.00033402	0.00050894
0.2	0	0.005207	0.000469	0.000773	0.0053485	0.0047826	0.00490651
0.2	1	0.007106	0.006329	0.003608	0.00585253	0.00652674	0.00570962
0.2	5	0.005930	0.008328	0.008073	0.00817223	0.00735479	0.00787581
0.2	10	0.003754	0.004691	0.004658	0.00833543	0.00598578	0.00687753
0.2	100	0.000919	0.000495	0.000761	0.00060516	0.00037831	0.00042655
0.4	0	0.003676	0.000573	0.001148	0.0058407	0.00479858	0.00467297
0.4	1	0.005320	0.005271	0.002449	0.00639711	0.0061476	0.00533589
0.4	5	0.007067	0.009577	0.009073	0.00990732	0.01076321	0.01039308
0.4	10	0.004897	0.006376	0.006138	0.01194383	0.01120573	0.01204855
0.4	100	0.001937	0.000867	0.001146	0.0026441	0.00144126	0.00200212

Table 3.12: Maximum range of the weights (i.e., the difference between the maximum and minimum allocated weights) in the computation of the M-estimator with Huber (left) and Hampel (right) loss functions, considering 11 different values of θ

Remark 3.6.1. Note that a study on the variation of the weights involved in the computation of the M-estimators when changing the parameter θ can be made as in Remark 3.5.1. Both Huber and Hampel loss functions will be considered and the three different cases have already been detailed in Subsection 3.5.3.

The outputs obtained for the M-estimator based on the Huber loss function are contained in Table 3.12 (left columns) and those for the M-estimator using the Hampel loss function, in Table 3.12 (right columns).

As it happened when studying the spatial median, it can be checked that the maximum range is really small in all situations. Recall that the range for each of the 100 weights is the difference between the maximum and minimum corresponding values obtained for them all over the 11 possible choices of the parameter θ.

For the M-estimator computed using the Huber loss function, the conclusions coincide with those given for the d_θ-median: all the ranges are smaller than 1/100, so it has been shown empirically that the parameter has little influence on the estimation of the spatial median. Note that when the chosen loss function is Hampel's one, the results are similar. In all but a pair of situations, the maximum range remains smaller than 1/100. Moreover, in those two situations, the maximum range is only slightly higher than the reference 1/100.

A real-life example involving fuzzy-valued data is now to be considered, aiming to compare the behaviour of the Aumann-type mean value and the different approaches to extend the median in this chapter.

Example 3.6.2. TIMSS (Trends in International Mathematics and Science Study) is an international assessment of mathematics and science at the fourth and eighth grades that has been conducted every four years since 1995, with the most recent assessment in 2011. Countries and regional benchmarking entities could participate in the fourth grade assessment, the eighth grade assessment, or both.

PIRLS (Progress in International Reading Literacy Study) is an international assessment of reading comprehension at the fourth grade that has been conducted every five years since 2001.

In 2011, the TIMSS and PIRLS data collection schedules came into alignment for the first time in the history of these international assessments. This provided countries with the opportunity to assess their fourth grade students in three fundamental curricular areas: mathematics, science, and reading. 34 countries and three benchmarking entities took advantage of this unique opportunity to assess the same students in all three subjects. Taken together, the fourth grade students in these 34 countries and three benchmarking participants have achievement data in the three core academic areas (reading, mathematics, and science) accompanied by an extensive array of background questionnaire data about the home, school, and classroom contexts for learning these three subjects.

In general, participating countries use TIMSS and PIRLS in various ways to explore educational issues, including: monitoring system-level achievement trends in a global context, establishing achievement goals and standards for educational improvement, stimulating curriculum reform, improving teaching and learning through research and analysis of the data, conducting related studies. More information about can be found in http://timss.bc.edu/.

In 2011, the Spanish Institute of Educational Evaluation (INEE) has commissioned a member of the SMIRE Research Group in the Department of Statistics, OR and DM of the University of Oviedo in Spain (Prof. Norberto Corral) to develop such a data analysis with data collected through some of the TIMSS/PIRLS questionnaires conducted in Spanish schools (see Corral *et al.* [44] for a summary of conclusions). These questionnaires are standard ones in which concern responses, and most of the responses have to be chosen among those in a Likert scale with 4 points (namely, DISAGREE A LOT, DISAGREE A LITTLE, AGREE A LITTLE and AGREE A LOT).

As indicated in De la Rosa de Sáa *et al.* [53], for purposes of analyzing and summarizing the obtained responses, they are traditionally viewed as values of linguistic variables, and they are often encoded by means of consecutive integer numbers, in spite of the many concerns associated with such an encoding. One of the crucial drawbacks are related to the fact that descriptive and inferential statistics which can be developed with the responses from one of Likert scale-based questionnaires are quite limited, even in case the responses are encoded in terms of integer numbers.

The fuzzy rating scale has been introduced (Hesketh *et al.* [102]) as an approach allowing to combine a free-response format with a fuzzy valuation. In the fuzzy rating scale, along a continuous line between two end-points

- a respondent selects or draws a 'representative position/interval' of the respondent rating (i.e., the set of points which she/he considers to be fully compatible with such a rating),
- and the respondent also indicates 'latitudes of acceptance' on either side by determining the highest and lowest possible positions for the respondent rating (i.e., the set of points which she/he considers to be compatible to some extent with such a rating).

As it has been shown De la Rosa de Sáa *et al.* [53], one can achieve more accurate conclusions, as well as explore and exploit more information, by considering the fuzzy rating scale than Likert-based and even fuzzy linguistic ones.

As outlined in De la Rosa de Sáa *et al.* [53], although fuzzy rating scale-based questionnaires are not exactly friendly-to-use ones, a minor training is usually enough to respond; this can become a shortcoming in cases a quick response is required (say, in surveys at street, by phone, etc.), but to make non-experts to understand the rudiments to respond these questionnaires just a short time is needed (see Hesketh *et al.* [102]).

To show how this fuzzy rating scale works by means of a simple example, some of the questions for the Student questionnaire TIMSS/PIRLS (see http://timss.bc.edu/timss2011/downloads/T11_StuQ_4.pdf) have been adapted in accordance with this scale, and the questionnaire has been conducted on the fourth grade students of the Colegio San Ignacio in Oviedo-Asturias (Spain).

The questions have been formulated with a double-type response (namely, Likert scale and fuzzy rating scale-based).

Data from three of these adapted questions (in fact those related to Maths) are now to be analyzed, and some of the M-estimates of response location in this chapter are to be computed.

First, the training of the students has been carried out by providing them with some instructions to fill out the questionnaire. Since drawing a trapezoidal fuzzy set is difficult to explain to students at this level, because they do not have yet the required background about real-valued functions, we have made use of the notion of trapezium, which can be trivially identified with that of a trapezoidal fuzzy set. No remarkable problems have been found either in the training or in the obtained responses being coherent and plausible.

The questionnaire has been designed in both paper-and-pencil and computerized formats, so that teachers of the students have been finally the ones deciding about the way to fill out them. In this case, teachers have decided 24 of the students fill out the paper-and-pencil format and 44 of them complete the computerized version, and this will become certainly interesting for future studies beyond the scope of this work.

The computerized format (in Spanish) has been designed by Professor Carlos Carleos from the Department of Statistics, OR and DM of the University of Oviedo and it can be found in http://carleos.epv.uniovi.es:8080/ (see Figure 3.25 for an example of a question from the computerized version of such a questionnaire).

The students have followed the instructions they have received which have been gathered in the guideline in Figures 3.26 and 3.27.

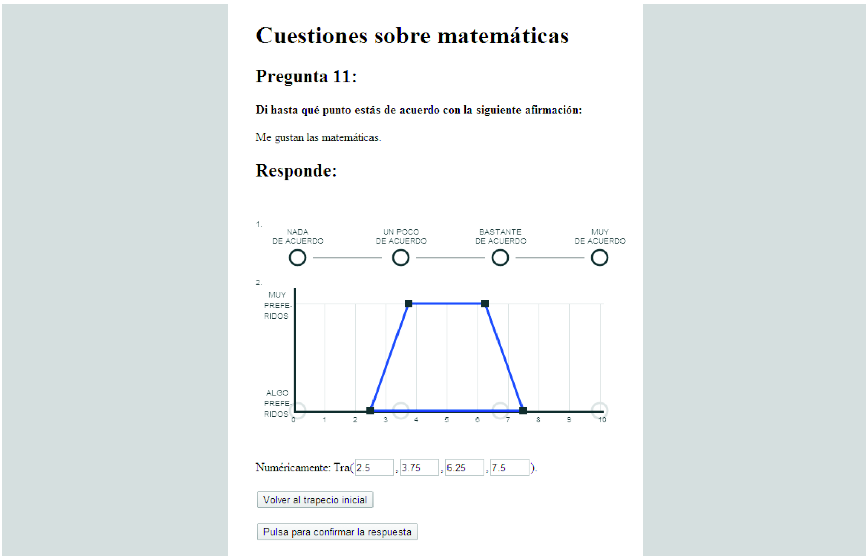
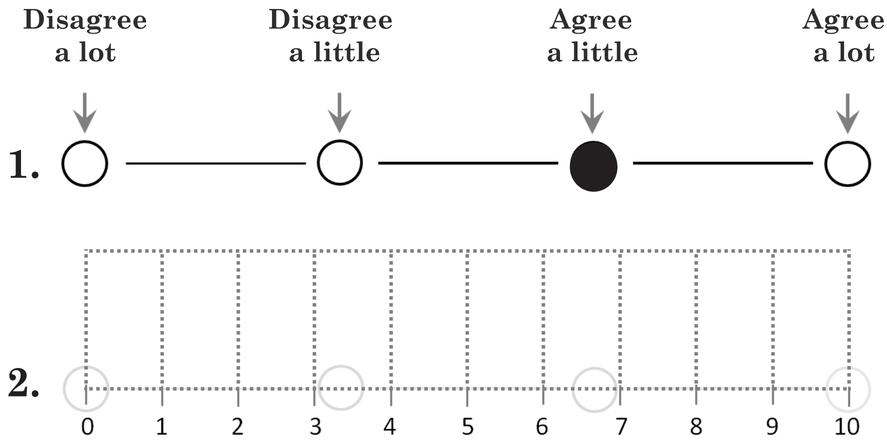


Figure 3.25: Example of a question in the computerized version (in Spanish) of the Likert and fuzzy rating scale-based questionnaire

Directions

In this questionnaire, you will find questions about you and what you think. For each question, you should first choose the answer you think is the best according to responses in Point **1**. Later, you should draw a trapezium describing your response in more detail as it is now explained in connection with Point **2**.

- Read each question carefully, and choose the answer in 1 you think is the best to describe your opinion. Fill in the circle under your answer.



- Read each question carefully, and draw the answer in **2** so that it is a trapezium (in particular, it can be a triangle) so that
  - The upper base (with altitude equal to 1) is the line segment with end points between 0 and 10 that better describes your rating on your opinion.

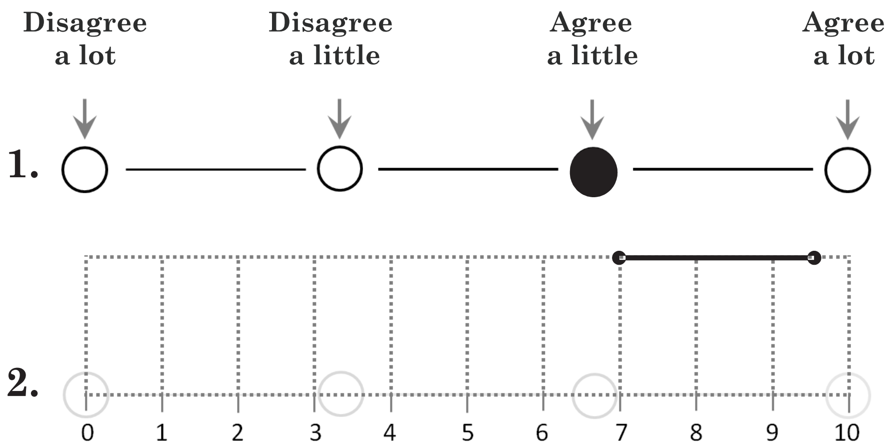
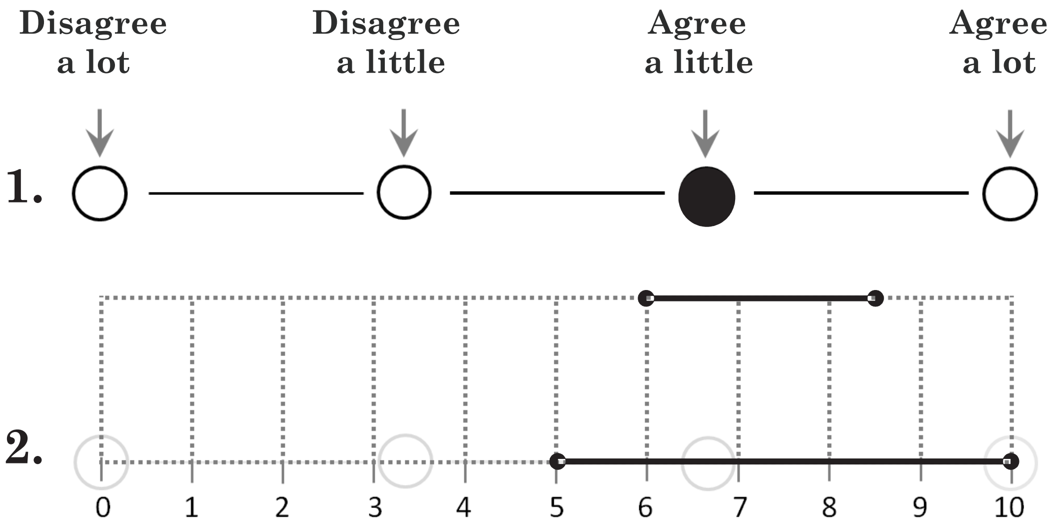
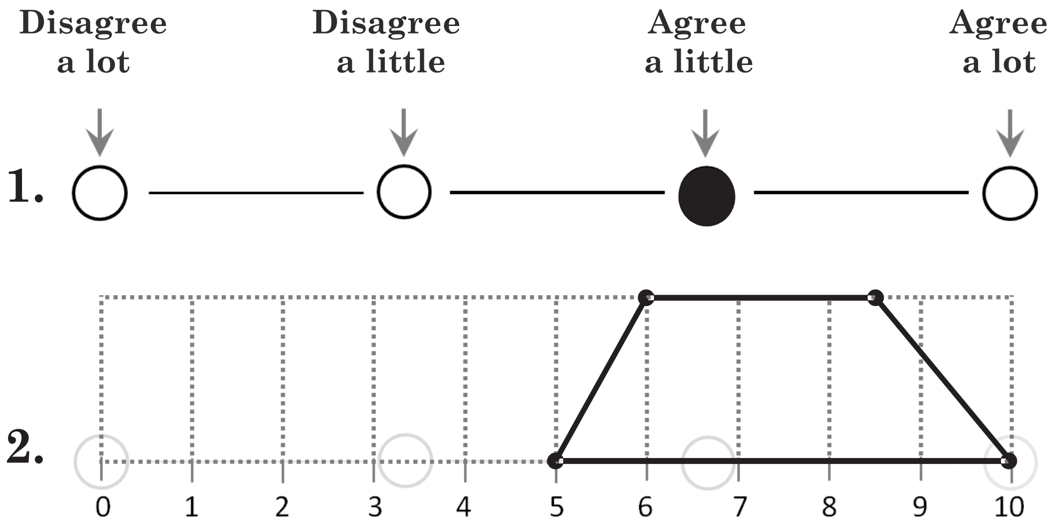


Figure 3.26: Directions to fill out the double-type (Likert scale and fuzzy rating scale-based) response questionnaire (1st page)

- The lower base is the line segment with end points between 0 and 10 that describes to some extent your rating on your opinion.



- The legs of the trapezium can be immediately drawn



When one can only answer through **1**, many people would like having chance to choose something 'in between' two given responses. **2** allows us to use a more flexible and to indicate not only the "fully preferred values", but also those being "preferred to some extent".

Figure 3.27: Directions to fill out the double-type (Likert scale and fuzzy rating scale-based) response questionnaire (2nd page)

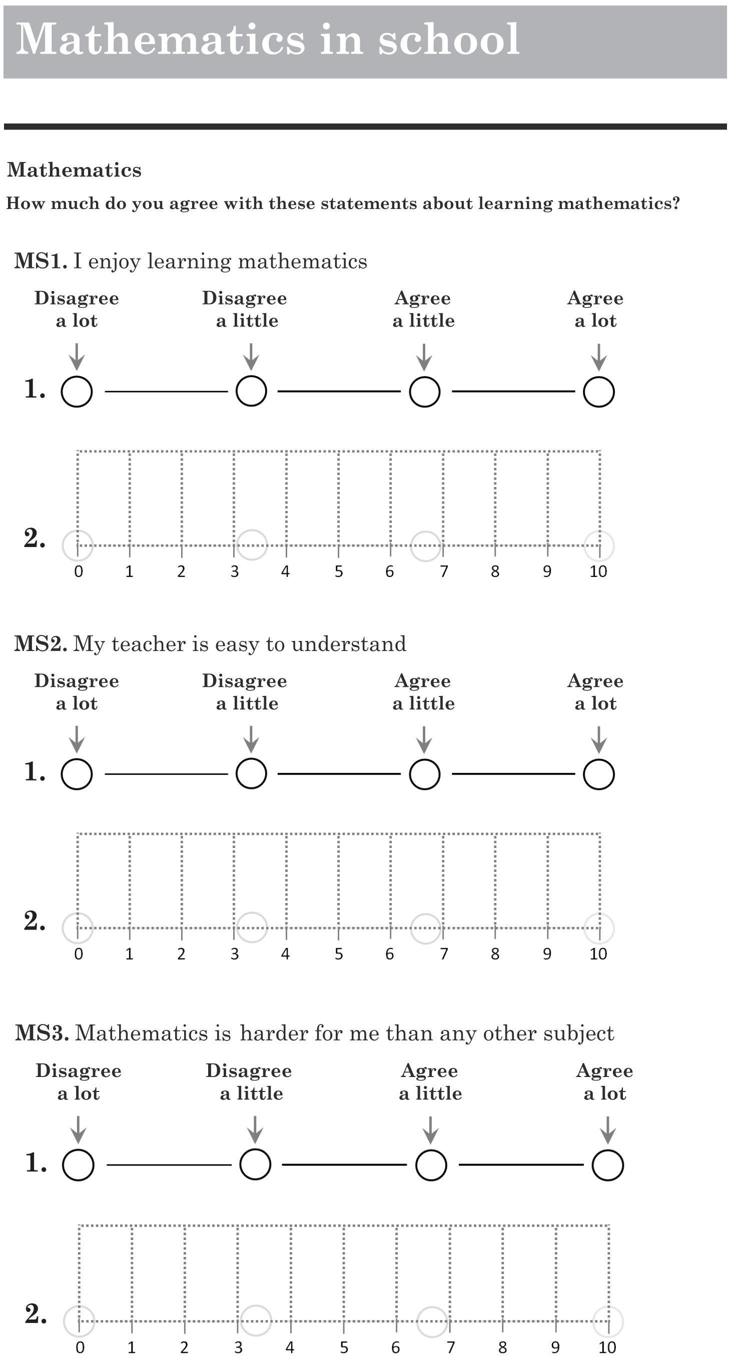


Figure 3.28: The three questions about mathematics to fill out in the double-response questionnaire

The three selected questions to analyze with the M-estimates in this chapter refer to the agreement with three statements about mathematics. Questions have not been modified with respect to the original questionnaire, but simply the second way to respond has been added. The paper-and-pencil format corresponding to these three questions are graphically displayed in Figure 3.28.

Data from the 68 fourth grade students from Colegio San Ignacio have been collected in Table 3.13. Note that MS2 has been only answered by 67 students.

MS1				MS2				MS3			
$\inf \widetilde{A}_0$	$\sup \widetilde{A}_0$	$\sup \widetilde{A}_1$	$\inf \widetilde{A}_1$	$\inf \widetilde{A}_0$	$\sup \widetilde{A}_0$	$\sup \widetilde{A}_1$	$\inf \widetilde{A}_1$	$\inf \widetilde{A}_0$	$\sup \widetilde{A}_0$	$\sup \widetilde{A}_1$	$\inf \widetilde{A}_1$
2.35	3	4	4.8	8.4	9	10	10	5.2	6	7	7
5.65	6.2	7.475	9.8	9.2	9.2	9.975	9.98	0	0	0	2.95
9.85	9.85	9.875	9.975	8.925	9.95	9.95	9.95	6	6.425	7.35	7.875
8.025	9.05	9.9	9.975	0.1	0.1	0.625	1.98	0	0	0.775	1.225
3.525	3.75	6.25	6.725	3.5	3.55	6.25	7.5	0	0	2.15	2.15
9	9.225	9.775	10	10	10	10	10	10	10	10	10
2.5	2.975	5.3	5.35	4.375	5.175	7.5	7.5	7.9	7.95	8.7	8.7
2	2.45	3	3	3	3	3.45	4	9	9	9.4	10
3.1	4	4	4.5	8.6	10	10	10	9	10	10	10
4.325	5.025	7.925	8.5	1.75	2.5	3.675	3.68	4.975	4.975	5.325	5.4
7.525	7.525	7.55	9.025	6.975	6.975	7.5	7.5	7.6	7.6	8.35	8.65
9	9	10	10	10	10	10	10	2	2	4	4
9.4	10	10	10	8.9	9.4	10	10	0	0	0	0.45
5.025	5.95	7.025	8.95	7.3	8.05	9.575	10	9.975	10	10	10
5.75	5.775	9.55	9.875	5.95	6	9.2	10	0	0	1.575	1.575
6.85	8	10	10	6.75	7.025	9.975	9.98	2.225	2.225	3.125	3.125
4.025	5.75	8.725	10	4.9	4.9	8.45	9.98	1.9	1.95	3	3.15
4.2	4.925	6.975	7.2	4.4	4.725	6.25	7.8	4.875	5.05	5.45	5.625
3.75	3.75	7.5	7.5	6.225	6.25	7.5	7.5	6.15	6.15	6.75	6.75
5.85	7.025	9.05	9.1	0	0.125	2.05	2.55	3.45	3.45	4.425	4.425
3.1	3.25	3.85	4.5	10	10	10	10	2.5	3.2	3.3	4.45
6	6.7	7.2	8	8.7	9.4	10	10	3	3.6	4.2	5.05
6.1	6.4	6.75	7.1	9	10	10	10	3	3	3	3
10	10	10	10	9.975	9.975	9.975	9.98	0	0.6	1.25	1.65
9	9.5	9.5	10	8	8.5	8.5	9	0	0.5	0.5	1
2.4	3	3.65	3.65	6	6	6.6	7.7	2.5	3	3.6	3.6
6	6.15	6.55	7	8.1	8.2	8.6	9	3	3.2	3.6	4.2
2.5	2.95	6.25	7.5	3.4	4.825	9.95	9.95	10	10	10	10
9.975	10	10	10	9.975	9.975	10	10	10	10	10	10
2.975	3.05	10	10	3	3	7.95	7.95	6.975	6.975	7.925	7.925
2.5	3.75	6.25	7.5	2.5	3.75	6.25	7.5	0	0	2.575	2.575
3.8	4.25	5.5	6	9.6	9.8	10	10	6	6.45	7.4	8
4.6	4.75	5.15	5.35	9.2	9.8	10	10	2.35	2.8	3.25	3.5
6.2	6.4	6.85	7.1	5.2	5.4	5.65	6	3.15	3.4	3.6	4
3.05	4.05	7.95	9.025	8.725	8.95	9.7	10	0	0.625	2.725	2.75
8	9.15	10	10	8	9	10	10	0	0	1	2
9.925	9.95	10	10	7	7.025	8.9	8.98	4.925	5.025	5.95	6.3
0	0.025	0.025	0.025	9.975	9.975	9.975	10	10	10	10	10
2.925	2.975	5.95	5.975	9.45	9.45	9.925	10	0	0.825	2.425	2.425
0	1.125	1.2	1.275	2.5	3.75	3.9	5.45	0	0.325	1.475	1.475
3.7	3.75	7.225	7.25	6.9	8.175	9.225	9.98	5.15	5.35	6.15	6.15
3.825	4.9	6.05	6.725	6.7	7.775	8.9	10	8.55	8.85	9.625	10
8.975	8.975	8.975	10	3.175	5.025	7.5	9.95	0	0	0	0.725
10	10	10	10	10	10	10	10	0	0	0	0
10	10	10	10	8.05	8.65	10	10	10	10	10	10
6	6.65	7.25	7.25	8	8.5	9.2	9.2	7	7.4	8.2	8.4
2	2	5	5	5	6	6.125	8	4.05	4.05	4.7	4.775
8.975	8.975	9.975	9.975	9.025	9.025	9.95	9.95	10	10	10	10
2.5	2.975	5.5	6.5	8	8.5	9.85	9.88	0	0.85	1.5	1.825
4.85	5	7.05	7.875	7.95	9	10	10	1.6	1.825	2.425	3.075
3.075	3.1	4	7.5	9.325	9.375	10	10	3.125	3.275	3.7	4.05
0.975	3.875	4.075	4.075	3.975	4.925	6.875	6.93	9.9	9.9	10	10
6.675	6.675	6.675	6.7	0.225	3	6.875	9.9	0	0	1.125	1.125
7	7	8	9	7	8	9	9	6	6	7	8
8	8.3	8.55	9	9	10	10	10	1	1.8	2.35	3.1
7.925	7.95	8	8	6.075	6.15	9.05	9.05	0	0.075	1	1.35
9	10	10	10	8	10	10	10	0	0	0	0
8.3	9.3	9.8	10					0	0.4	0.95	1.75
0.05	0.05	0.075	0.075	9.025	9.025	9.95	9.95	10	10	10	10
1.45	1.95	4.95	5.725	5.6	6.7	9.15	10	8.8	8.8	9.5	9.575
2.9	3.75	6.25	7.8	9.85	9.85	9.9	9.9	4.6	6.15	6.15	6.85
9.875	9.95	9.95	9.975	4.225	5.7	7.025	8.9	3.6	3.925	4.575	4.575
2.5	4.075	7.175	8.15	5.825	5.85	9.875	9.95	3.875	3.875	5.6	5.6
2.5	2.55	4.275	4.3	2.5	4.625	4.625	6.9	0	0.25	1.025	1.025
8	8.025	9.8	9.975	9.8	9.8	10	10	10	10	10	10
8.55	9.15	9.7	10	8.6	9.15	9.75	10	0.3	0.45	1.15	1.5
3.5	4.2	5	5.45	5.1	6	6.75	7.3	5.5	6.1	6.9	7.4
2.5	2.5	5.1	7.5	10	10	10	10	6.325	6.925	7.175	7.65

Table 3.13: Fuzzy rating scale-based responses given by 4th grade students in Colegio San Ignacio (Oviedo, Spain)

This table includes for each of the three questions the 4-tuple corresponding to each fuzzy rating scale-based answer $\widetilde{A}$, that is, $\inf \widetilde{A}_0$, $\sup \widetilde{A}_0$, $\sup \widetilde{A}_1$ and $\inf \widetilde{A}_1$.

The M-estimates of the response location have been as follows (see Figures 3.29, 3.30 and 3.31):

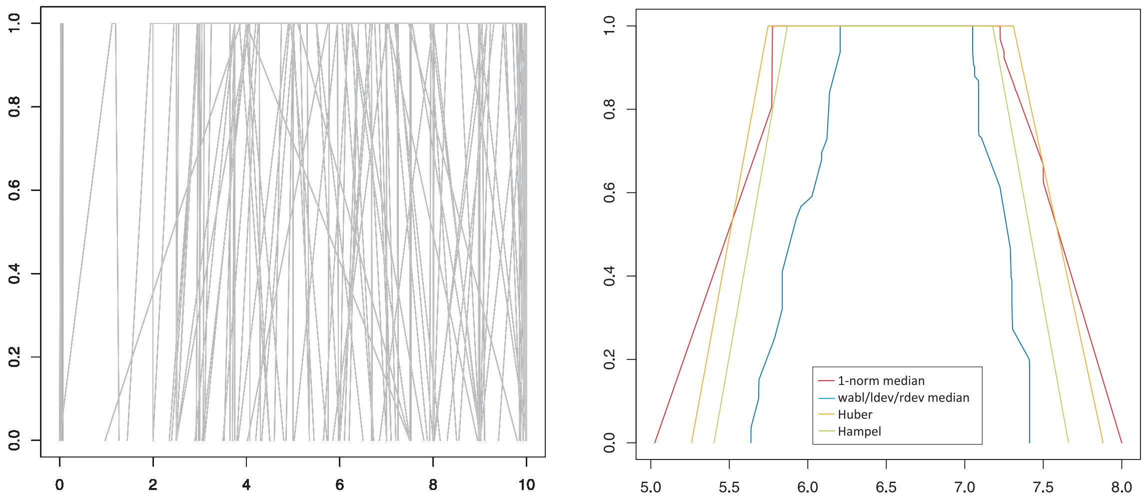


Figure 3.29: Sample fuzzy data and location M-estimates of the 68 fuzzy rating scale-based responses to Question MS1

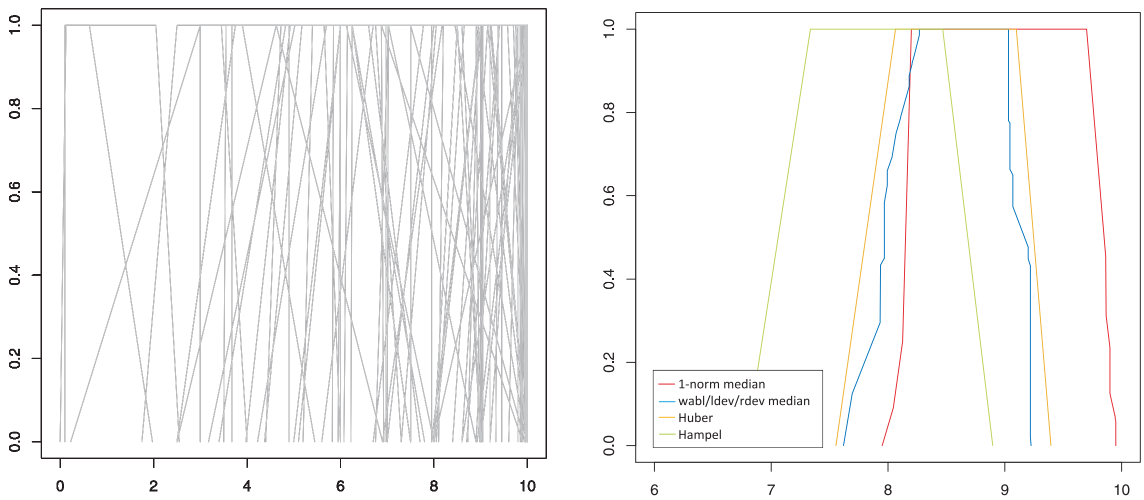


Figure 3.30: Sample fuzzy data and location M-estimates of the 67 fuzzy rating scale-based responses to Question MS2

Remark 3.6.2. The computations for the Huber (with $a = 1.345$) and Hample M-estimates (where the parameters have been fixed like in Example 3.6.1 with the initial seed being the 1-norm median) have been based on the $\mathfrak{D}^{\ell}_{1/3}$ metric, although no perceptible differences have been found with those based on $D^{\ell}_{1/3}$.

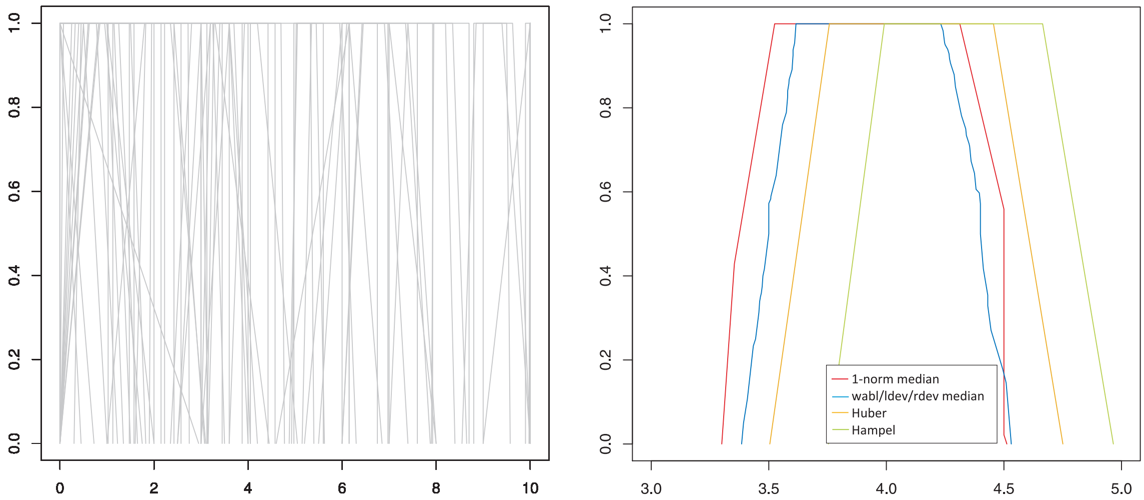


Figure 3.31: Sample fuzzy data and location M-estimates of the 68 fuzzy rating scale-based responses to Question MS3

One can easily check that, because of the conditions concerning the existence and expression of the M-estimates as a convex linear combination of the fuzzy trapezoidal data in the sample sample of size 68 are fulfilled, both estimates preserve the trapezoidal shape of the sample data.

It is interesting to comment that the medians of the Likert-type responses after integer encoding have been AGREE A LITTLE, AGREE A LOT and AGREE A LITTLE for MS1, MS2 and MS3, respectively. These would correspond to $6.\widehat{6}$, 10 and $6.\widehat{6}$, respectively, in the re-scaling to [0,10]. So, as already pointed out by De la Rosa de Sáa *et al.* [53], the fuzzy rating scale-based ones offer richer nuances and expressiveness, this being especially evident in this example for Question MS3.

On the other hand, it should be remarked that the way to proceed in this example follows accurately the path suggested by Zadeh, who has coined it as the *"precisiation of the imprecise"* (see Zadeh [221]).

To illustrate the computation of M-estimates of location for functional data, the following example is considered:

Example 3.6.3. The considered data set consists of $n = 472$ radar waves registered by the satellite Topex/Poseidon around an area of 25 kilometers upon the Amazon River, with the aim of use them for altimetric and hydrological purposes. To represent the curves, their discretized version is obtained from a partition of $t = 70$ moments in time, i.e., for all $i = 1, \ldots, n$, $\mathbf{X}_i = (X_i(t_1), \ldots, X_i(t_{70}))$. The data set and this brief description have been obtained from the web page http://www.math.univ-toulouse.fr/staph/npfda/npfda-datasets.html and deeper information can be found in

Frappart [76].

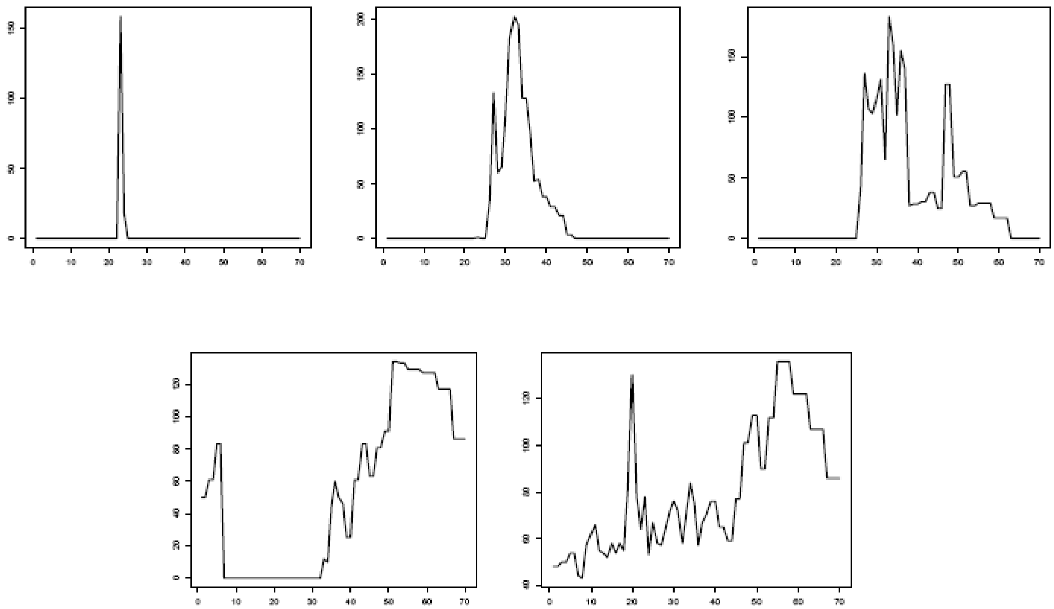


Figure 3.32: Waveforms numbers 21 (top left), 3 (top middle), 1 (top right), 5 (bottom left) and 4 (bottom right)

As it is outlined in the web page and shown in Figure 3.32, there are different kinds of waves, namely,

- curves with one heavy peak, like curve number 21;
- curves with one less heavy peak, like curve 3;
- curves that look to have more than one peak, like curve 1;
- curves that look as they had no really peak, like number 5;
- 'flat noised curves', e.g., observation number 4

and so on.

The considered loss function will be Huber's one with $a = 1.345$, and Hampel's one where the parameters have been fixed like in Example 3.6.1 but the initial seed is now the 0.2-trimmed mean.

The estimates obtained for the mean, the trimmed mean (with the often used trimming proportion of 0.2) and the M-estimators with Huber and Hampel loss functions are plotted in Figure 3.33, where the difference between the mean, more influenced by the perturbations of the curves, and the rest of measures, with a more robust behavior, is evident.

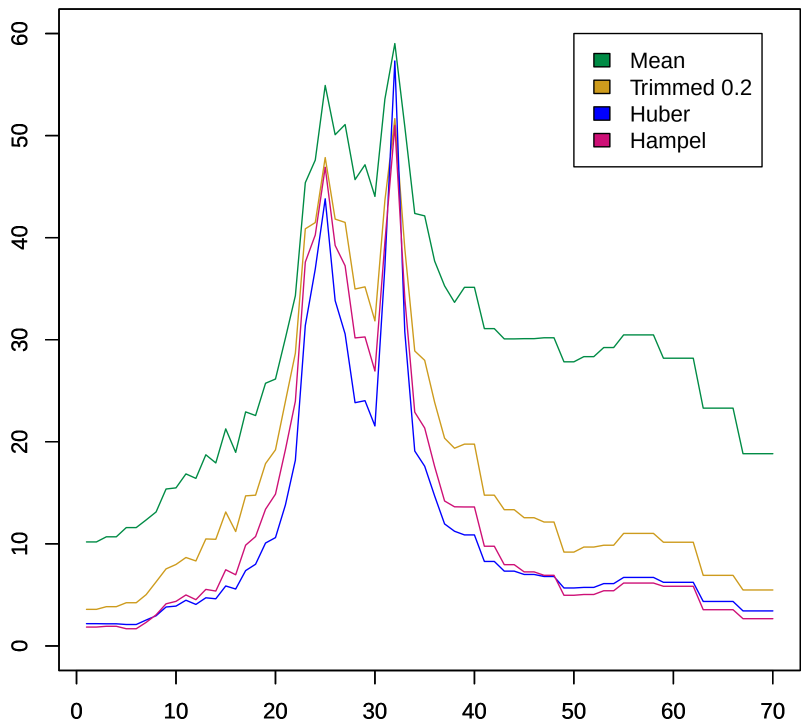


Figure 3.33: The estimates for the mean, the 0.2-trimmed mean, the Huber M-estimator and the Hampel M-estimator in Example 3.6.3

3.7 Concluding remarks of this chapter

This chapter has been devoted to establish M-estimates of location for imprecise-valued data. For this purpose, some recent developments combining ideas in kernel density estimation with those from classical M-estimation have been adapted to develop M-estimates of location for Hilbert space-valued data. Necessary and sufficient conditions for the existence and expression of the M-estimates, as well as an algorithm to approximate them, have been also adapted. These conditions have been shown to be especially interesting in particularizing to imprecise-valued data, due to the semi-linearity of the spaces of imprecise values we have considered in this work, since one can guarantee they lead to estimates within the parameter space.

Although for many valuable loss functions, the above-mentioned necessary and sufficient conditions are fulfilled, there are some other useful and reasonable ones for which they do not. For some of the most outstanding ones among the latter, *ad hoc* developments have been performed. In this way, two new M-estimates have been suggested for fuzzy number-valued data on the basis of L^1 metrics introduced in Chapter 1, and another M-estimate for interval-valued data which extends the spatial median has been proposed. Their properties have been also analyzed in detail.

The main contribution in the Chapter refers to

- the adapted studies of M-estimation on Hilbert spaces, along with the wide (but not general enough) conditions to be fulfilled for the existence and convenient expression of the solutions (the convenience being intended in terms of its particularization to the imprecise-valued case leading to a valid estimate);
- the new approaches involving a situation which is not covered by the conditions before, and based on L^1 metrics between imprecise values, which lead to exactly and easily computable M-estimates and showing also the main properties;
- the new approach involving another situation which is not covered by the conditions before, and based on and L^2 metric between interval values, which lead to approximate M-estimates sharing the main properties, especially those related to robustness;
- the comparison with the mean leads for all the approaches to better estimates of location.

The ideas and results in this chapter have been gathered in three published/accepted papers (Sinova *et al.* [173, 180, 182]), one manuscript already prepared to be submitted (Sinova and Van Aelst [189]) and eleven communications to conferences (Sinova *et al.* [171, 172, 174, 175, 178, 181, 183, 184, 186, 187, 188]). Some other ones are being prepared in connection with the M-estimates under the conditions in Sections 3.1 and 3.2.

Chapter 4

Comparative simulation studies between location estimates for imprecise-valued data

In Chapters 2 and 3 different approaches to the robust measurement/estimation of location for imprecise-valued random elements or data sets have been either introduced or adapted. The robustness for all the approaches have been compared with that of the (Hilbert space, Aumann or Aumann-type) mean from a theoretical perspective, through the finite sample breakdown point, and from an empirical point of view through simulations.

In this chapter, a comparative simulation is set up to check the robustness of the approaches in Chapters 2 and 3 when dealing with the most common types of set-, fuzzy set- and Hilbert space-valued data, namely, interval-, fuzzy number- and functional-valued data.

Section 4.1 is to be devoted to discuss the interval case, Section 4.2 is addressed to discuss the fuzzy numbers one, and Section 4.3 corresponds to the functional case. The chapter ends with Section 4.4, in which the information obtained from these empirical developments is summarized.

4.1 Comparative simulations for interval-valued data

The simulations in this section have been carried out by following basically the ideas in those included in the preceding chapters. Now, four different studies have been conducted varying the sample size ($n = 100, n = 10000$), the non-contaminated (symmetric and asymmetric) and the contaminated distributions.

For each of the four studies, the comparisons have concerned the following location measures/estimates: trimmed means, Huber and Hampel M-estimates (using the d_θ-metric), 1-norm median, generalized Hausdorff median corresponding to the interval-valued particularization of the wabl/ldev/rdev median and the spatial median, where in all of them θ is assumed to range in $\{1/3, 1\}$. When $n = 100$, the medians and trimmed means based on the well-known halfspace and simplicial depths (see Tukey [205, 206] and Liu [124], respectively) have been also considered. The reason not to include them in the comparisons when $n = 10000$ is that the computation of depths is usually hard when the sample size increases.

For each of the measures/estimates, the characteristics in Subsection 3.5.3 have been determined, namely, the Monte Carlo approximation of the estimate, the bias, the variance and the mean squared error of the estimate.

The general scheme of the four studies has been as follows:

Step 1. A sample of n interval-valued data has been simulated from a random interval X for each of some different situations in such a way that

- to generate the interval-valued data, we have considered two real-valued random variables as follows: $\mathsf{X} = [X_1 - X_2, X_1 + X_2]$, with $X_1 = \operatorname{mid} \mathsf{X}$, $X_2 = \operatorname{spr} \mathsf{X}$ or, alternatively, two order real-valued statistics $X_{(1)}$ and $X_{(2)}$ such that $\mathsf{X} = [X_{(1)}, X_{(2)}]$, i.e., $X_{(1)} = \inf \mathsf{X}$, $X_{(2)} = \sup \mathsf{X}$;
- each sample is assumed to be split into a subsample of size $n(1 - c_p)$ (where c_p denotes the proportion of contamination and is supposed to range in $\{0, 0.1, 0.2, 0.4\}$) associated with a non-contaminated distribution and a subsample of size $n \cdot c_p$ associated with a contaminated one, where an additional contamination role is played by C_D (which measures the relative distance between the distribution of the two subsamples and ranges in $\{0, 1, 5, 10, 100\}$);
- 16 situations for different values of c_p and C_D have been considered for simulations and for each of these situations two cases have been selected, namely, one in which random variables X_i (or $X_{(i)}$) are independent (CASES 1 and 3) and another one in which they are dependent (CASES 2, 2' and 4).

Step 2. $N = 1000$ replications of *Step 1* have been considered for the situation $c_p = C_D = 0$ in order to approximate the population measures by using a Monte Carlo approach.

Step 3. $N = 1000$ replications of *Step 1* have been considered for all the situations (c_p, C_D) and the approximated estimates, bias, variance and mean squared error have been computed for each location measure.

Study 1

In this first study, the choices correspond to:

- $n = 100$;
- CASE 1 assumes that
  - •• $X_1 \sim \mathcal{N}(0,1)$ and $X_2 \sim \chi_1^2$ for the non-contaminated subsample,
  - •• $X_1 \sim \mathcal{N}(0,3) + C_D$ and $X_2 \sim \chi_4^2 + C_D$ for the contaminated subsample,

  whereas CASE 2 assumes that
  - •• $X_1 \sim \mathcal{N}(0,1)$ and $X_2 \sim 1/(X_1^2+1)^2 + 0.1 \cdot \chi_1^2$ for the non-contaminated subsample,
  - •• $X_1 \sim \mathcal{N}(0,3) + C_D$ and $X_2 \sim 1/(X_1^2+1)^2 + 0.1 \cdot \chi_1^2 + C_D$ for the contaminated subsample.

  and CASE 2' assumes that
  - •• $X_1 \sim \mathcal{N}(0,1)$ and $X_2 \sim 1/(X_1^2+1)^2 + \sqrt{\chi_1^2}$ for the non-contaminated subsample,
  - •• $X_1 \sim \mathcal{N}(0,3)+C_D$ and $X_2 \sim 1/(X_1^2+1)^2+\sqrt{\chi_1^2}+C_D$ for the contaminated subsample.

The estimates of all the considered measures in this simulation study will be presented through Tables 4.1-4.3, whereas the outputs for the bias, the variance and the mean squared error will be summarized in Table 4.7.

Study 2

In this second study, the choices correspond to:

- $n = 10000$;
- In CASES 1, 2 and 2' the distributions for X_1 and X_2 in the non-contaminated and the contaminated samples coincide with those for Study 1.

The estimates of all the considered measures in this simulation study will be presented through Tables 4.4-4.6, whereas the outputs for the bias, the variance and the mean squared error will be summarized in Table 4.8.

Study 3

In this third study, the choices correspond to:

- $n = 100$;
- CASE 3 assumes that
  - •• $X_{(1)}, X_{(2)} \sim \text{Beta}(5,1)$ (they are simply chosen at random and ordered) for the non-contaminated subsample,
  - •• $X_{(1)}, X_{(2)} \sim \text{Beta}(1, C_D + 1)$ for the contaminated subsample,

  whereas CASE 4 assumes that
  - •• $X_1 \sim \text{Beta}(5,1)$ and $X_2 \sim \text{Uniform}[0, \min\{X_1, 1 - X_1\}]$ for the non-contaminated subsample,
  - •• $X_1 \sim \text{Beta}(1, C_D + 1)$ and $X_2 \sim \min\{X_1, 1 - X_1\} \cdot \text{Beta}(1, C_D + 1)$ for the contaminated subsample.

The estimates of all the considered measures in this simulation study will be presented through Tables 4.9-4.10, whereas the outputs for the bias, the variance and the mean squared error will be summarized in Table 4.13.

Study 4

In this fourth study, the choices correspond to:

- $n = 10000$;
- In CASES 3 and 4 the distributions for $X_{(1)}$, $X_{(2)}$, X_1 and X_2 in the non-contaminated and contaminated samples coincide with those for Study 3.

The estimates of all the considered measures in this simulation study will be presented through Tables 4.11-4.12, whereas the outputs for the bias, the variance and the mean squared error will be summarized in Table 4.14.

To avoid excessive information in the outputs of the simulations, the details about bias, variance and mean squared error can be found in the following link: http://bellman.ciencias.uniovi.es/SMIRE/Intsimul.html

Concerning **Study 1** and **Study 2**, the estimates are given by

c_p	c_D	mean	trimmed ($\theta = 1/3$)	trimmed ($\theta = 1$)	Hausdorff median	1-norm median
0	0	[0,0000 , 0,0000]	[0,0000 , 0,0000]	[0,0000 , 0,0000]	[0,0000 , 0,0000]	[0,0000 , 0,0000]
0,1	0	[0,2197 , 0,2198]	[0,0894 , 0,0895]	[0,0862 , 0,0862]	[0,0925 , 0,0927]	[0,0978 , 0,0979]
0,1	1	[0,2949 , 0,2994]	[0,0975 , 0,0978]	[0,0957 , 0,0959]	[0,1026 , 0,1046]	[0,1171 , 0,1185]
0,1	5	[0,5795 , 0,6379]	[0,0748 , 0,0749]	[0,0954 , 0,0954]	[0,0957 , 0,1140]	[0,1205 , 0,1290]
0,1	10	[0,9306 , 1,0857]	[0,0719 , 0,0720]	[0,0952 , 0,0953]	[0,0881 , 0,1135]	[0,1122 , 0,1229]
0,1	100	[7,5097 , 9,2922]	[0,0818 , 0,0819]	[0,1025 , 0,1026]	[0,0971 , 0,1234]	[0,1276 , 0,1390]
0,2	0	[0,4384 , 0,4385]	[0,2067 , 0,2067]	[0,2007 , 0,2008]	[0,2089 , 0,2089]	[0,2157 , 0,2158]
0,2	1	[0,5911 , 0,5984]	[0,2382 , 0,2382]	[0,2356 , 0,2356]	[0,2307 , 0,2325]	[0,2631 , 0,2647]
0,2	5	[1,1568 , 1,2723]	[0,2621 , 0,2628]	[0,3260 , 0,3281]	[0,2270 , 0,2694]	[0,2942 , 0,3201]
0,2	10	[1,9214 , 2,1943]	[0,3090 , 0,3091]	[0,4002 , 0,4007]	[0,2315 , 0,2777]	[0,3000 , 0,3283]
0,2	100	[15,1008 , 18,3581]	[0,3235 , 0,3235]	[0,4386 , 0,4387]	[0,2299 , 0,2710]	[0,2964 , 0,3222]
0,4	0	[0,8845 , 0,8845]	[0,2778 , 0,2778]	[0,2255 , 0,2256]	[0,5404 , 0,5404]	[0,5137 , 0,5138]
0,4	1	[1,2136 , 1,2305]	[0,3356 , 0,3356]	[0,2623 , 0,2623]	[0,7644 , 0,7685]	[0,7159 , 0,7237]
0,4	5	[2,3496 , 2,5724]	[0,2471 , 0,2475]	[0,2663 , 0,2672]	[0,8026 , 0,8693]	[0,9707 , 1,0652]
0,4	10	[3,8902 , 4,3671]	[0,2838 , 0,2839]	[0,3166 , 0,3166]	[0,8494 , 0,9233]	[1,0513 , 1,1665]
0,4	100	[30,5828 , 36,7510]	[0,2782 , 0,2782]	[0,3139 , 0,3139]	[0,8008 , 0,8864]	[1,0343 , 1,1632]

c_p	c_D	spatial ($\theta = 1/3$)	spatial ($\theta = 1$)	Tukey	Liu	trimmed Tukey
0	0	[0,0000 , 0,0000]	[0,0000 , 0,0000]	[0,0000 , 0,0000]	[0,0000 , 0,0000]	[0,0000 , 0,0000]
0,1	0	[0,1042 , 0,1044]	[0,0981 , 0,0982]	[0,1167 , 0,1168]	[0,1160 , 0,1163]	[0,1736 , 0,1738]
0,1	1	[0,1238 , 0,1254]	[0,1142 , 0,1158]	[0,1369 , 0,1394]	[0,1332 , 0,1349]	[0,2256 , 0,2271]
0,1	5	[0,1299 , 0,1421]	[0,1159 , 0,1285]	[0,1368 , 0,1486]	[0,1371 , 0,1466]	[0,3328 , 0,3472]
0,1	10	[0,1149 , 0,1346]	[0,1052 , 0,1248]	[0,1218 , 0,1416]	[0,1210 , 0,1353]	[0,3892 , 0,4181]
0,1	100	[0,1239 , 0,1483]	[0,1133 , 0,1384]	[0,1345 , 0,1584]	[0,1343 , 0,1547]	[1,7100 , 1,9098]
0,2	0	[0,2307 , 0,2307]	[0,2193 , 0,2193]	[0,2517 , 0,2517]	[0,2478 , 0,2478]	[0,3630 , 0,3631]
0,2	1	[0,2810 , 0,2825]	[0,2601 , 0,2617]	[0,3237 , 0,3249]	[0,3222 , 0,3236]	[0,4812 , 0,4835]
0,2	5	[0,3140 , 0,3449]	[0,2814 , 0,3139]	[0,3351 , 0,3649]	[0,3497 , 0,3766]	[0,8520 , 0,9002]
0,2	10	[0,3172 , 0,3565]	[0,2864 , 0,3279]	[0,3387 , 0,3763]	[0,3491 , 0,3830]	[1,3219 , 1,4376]
0,2	100	[0,3013 , 0,3439]	[0,2749 , 0,3212]	[0,3129 , 0,3511]	[0,3185 , 0,3591]	[8,8525 , 10,2364]
0,4	0	[0,5625 , 0,5625]	[0,5452 , 0,5453]	[0,6285 , 0,6285]	[0,6112 , 0,6112]	[0,7426 , 0,7426]
0,4	1	[0,7876 , 0,7926]	[0,7644 , 0,7701]	[0,8869 , 0,8923]	[0,8688 , 0,8736]	[1,0452 , 1,0518]
0,4	5	[1,0499 , 1,1154]	[0,9621 , 1,0367]	[1,2054 , 1,2591]	[1,2225 , 1,2763]	[1,9265 , 2,0203]
0,4	10	[1,1533 , 1,2379]	[1,0520 , 1,1508]	[1,4328 , 1,4981]	[1,4095 , 1,4781]	[3,0377 , 3,2359]
0,4	100	[1,1081 , 1,2245]	[1,0205 , 1,1578]	[1,2086 , 1,2967]	[1,3052 , 1,3976]	[21,2461 , 23,8563]

c_p	c_D	trimmed Liu	Huber ($\theta = 1/3$)	Huber ($\theta = 1$)	Hampel ($\theta = 1/3$)	Hampel ($\theta = 1$)
0	0	[0,0000 , 0,0000]	[0,0000 , 0,0000]	[0,0000 , 0,0000]	[0,0000 , 0,0000]	[0,0000 , 0,0000]
0,1	0	[0,1779 , 0,1781]	[0,1262 , 0,1263]	[0,1006 , 0,1008]	[0,0799 , 0,0800]	[0,0714 , 0,0715]
0,1	1	[0,2342 , 0,2358]	[0,1622 , 0,1637]	[0,1232 , 0,1250]	[0,0841 , 0,0843]	[0,0753 , 0,0755]
0,1	5	[0,3414 , 0,3592]	[0,1801 , 0,1937]	[0,1272 , 0,1431]	[0,0659 , 0,0659]	[0,0813 , 0,0813]
0,1	10	[0,3929 , 0,4284]	[0,1656 , 0,1868]	[0,1176 , 0,1420]	[0,0631 , 0,0634]	[0,0832 , 0,0833]
0,1	100	[1,1373 , 1,3251]	[0,1706 , 0,1950]	[0,1228 , 0,1520]	[0,0706 , 0,0706]	[0,0898 , 0,0898]
0,2	0	[0,3720 , 0,3721]	[0,2700 , 0,2701]	[0,2161 , 0,2162]	[0,1799 , 0,1799]	[0,1625 , 0,1626]
0,2	1	[0,4900 , 0,4926]	[0,3388 , 0,3406]	[0,2603 , 0,2625]	[0,1986 , 0,1986]	[0,1792 , 0,1792]
0,2	5	[0,8784 , 0,9322]	[0,4079 , 0,4400]	[0,2876 , 0,3266]	[0,1870 , 0,1880]	[0,2080 , 0,2099]
0,2	10	[1,3684 , 1,4998]	[0,4090 , 0,4526]	[0,2908 , 0,3431]	[0,1997 , 0,1998]	[0,2345 , 0,2351]
0,2	100	[9,2485 , 10,7854]	[0,3946 , 0,4425]	[0,2819 , 0,3407]	[0,2101 , 0,2102]	[0,2522 , 0,2522]
0,4	0	[0,7313 , 0,7313]	[0,5879 , 0,5880]	[0,4971 , 0,4972]	[0,4675 , 0,4675]	[0,4202 , 0,4202]
0,4	1	[1,0406 , 1,0480]	[0,7991 , 0,8050]	[0,6937 , 0,7006]	[0,6172 , 0,6177]	[0,5520 , 0,5531]
0,4	5	[1,9146 , 2,0156]	[1,1112 , 1,1889]	[0,9074 , 0,9987]	[0,9338 , 0,9458]	[0,7186 , 0,7553]
0,4	10	[3,0450 , 3,2539]	[1,2111 , 1,3132]	[0,9985 , 1,1194]	[1,1134 , 1,1258]	[0,7970 , 0,8538]
0,4	100	[21,6476 , 24,4015]	[1,1767 , 1,3175]	[0,9695 , 1,1374]	[1,1564 , 1,1685]	[0,8078 , 0,8771]

Table 4.1: Estimates of the location in Study 1-CASE 1

c_p	c_D	mean	trimmed ($\theta = 1/3$)	trimmed ($\theta = 1$)	Hausdorff median	1-norm median
0	0	[-0,6044 , 0,5962]	[-0,7114 , 0,6991]	[-0,7102 , 0,7012]	[-0,5910 , 0,5742]	[-1,0499 , 1,0368]
0,1	0	[-0,5901 , 0,5919]	[-0,6918 , 0,6857]	[-0,6919 , 0,6869]	[-0,5690 , 0,5679]	[-1,0393 , 1,0378]
0,1	1	[-0,6073 , 0,7201]	[-0,7344 , 0,7534]	[-0,7254 , 0,7495]	[-0,6140 , 0,6572]	[-1,0397 , 1,0766]
0,1	5	[-0,6776 , 1,2281]	[-0,6423 , 0,6767]	[-0,6416 , 0,6775]	[-0,5673 , 0,7076]	[-1,0420 , 1,1216]
0,1	10	[-0,7617 , 1,8724]	[-0,6571 , 0,6643]	[-0,6591 , 0,6628]	[-0,5662 , 0,7107]	[-1,0443 , 1,1215]
0,1	100	[-2,2977 , 13,3144]	[-0,6571 , 0,6598]	[-0,6568 , 0,6612]	[-0,5605 , 0,7084]	[-1,0425 , 1,1200]
0,2	0	[-0,5917 , 0,5770]	[-0,6821 , 0,6648]	[-0,6832 , 0,6648]	[-0,5669 , 0,5515]	[-1,0364 , 1,0239]
0,2	1	[-0,6251 , 0,8323]	[-0,7839 , 0,8001]	[-0,7810 , 0,7961]	[-0,6662 , 0,7379]	[-1,0492 , 1,1008]
0,2	5	[-0,7859 , 1,8372]	[-0,6155 , 0,6244]	[-0,6047 , 0,6141]	[-0,5704 , 0,8358]	[-1,0478 , 1,1679]
0,2	10	[-1,0072 , 3,0819]	[-0,6028 , 0,6033]	[-0,6027 , 0,6035]	[-0,5430 , 0,8546]	[-1,0378 , 1,1708]
0,2	100	[-4,8637 , 25,6411]	[-0,5983 , 0,6066]	[-0,5982 , 0,6067]	[-0,5456 , 0,8552]	[-1,0381 , 1,1711]
0,4	0	[-0,5787 , 0,5573]	[-0,7951 , 0,7789]	[-0,7967 , 0,7838]	[-0,5428 , 0,5239]	[-1,0241 , 1,0073]
0,4	1	[-0,6558 , 1,0661]	[-1,0282 , 1,0589]	[-0,9855 , 1,0260]	[-0,7800 , 0,9433]	[-1,0441 , 1,1567]
0,4	5	[-1,0100 , 3,0385]	[-0,6268 , 0,6799]	[-0,6259 , 0,6814]	[-0,5385 , 1,1835]	[-1,0034 , 1,3769]
0,4	10	[-1,5111 , 5,5613]	[-0,6545 , 0,6377]	[-0,6549 , 0,6376]	[-0,5091 , 1,2287]	[-0,9985 , 1,3993]
0,4	100	[-10,3903 , 51,1549]	[-0,6452 , 0,6364]	[-0,6457 , 0,6365]	[-0,4946 , 1,2483]	[-0,9963 , 1,4069]

c_p	c_D	spatial ($\theta = 1/3$)	spatial ($\theta = 1$)	Tukey	Liu	trimmed Tukey
0	0	[-0,8414 , 0,8230]	[-0,7909 , 0,7724]	[-0,7205 , 0,6857]	[-0,6362 , 0,6143]	[-0,6199 , 0,0000]
0,1	0	[-0,8266 , 0,8248]	[-0,7747 , 0,7735]	[-0,7101 , 0,6795]	[-0,6362 , 0,5996]	[-0,6075 , 0,1738]
0,1	1	[-0,8346 , 0,8845]	[-0,7839 , 0,8376]	[-0,6748 , 0,8160]	[-0,5852 , 0,7684]	[-0,6271 , 0,2271]
0,1	5	[-0,8146 , 0,9554]	[-0,7619 , 0,9022]	[-0,6418 , 0,8987]	[-0,5465 , 0,8438]	[-0,6197 , 0,3472]
0,1	10	[-0,8175 , 0,9654]	[-0,7653 , 0,9119]	[-0,6509 , 0,9110]	[-0,5469 , 0,8544]	[-0,6191 , 0,4181]
0,1	100	[-0,8143 , 0,9633]	[-0,7635 , 0,9079]	[-0,6571 , 0,8930]	[-0,5555 , 0,8351]	[-0,6336 , 1,9098]
0,2	0	[-0,8261 , 0,8105]	[-0,7752 , 0,7588]	[-0,7220 , 0,6504]	[-0,6593 , 0,5628]	[-0,6069 , 0,3631]
0,2	1	[-0,8531 , 0,9318]	[-0,8044 , 0,8890]	[-0,6786 , 0,9163]	[-0,6114 , 0,8617]	[-0,6469 , 0,4835]
0,2	5	[-0,8182 , 1,0645]	[-0,7721 , 1,0125]	[-0,7139 , 0,9992]	[-0,6123 , 0,9671]	[-0,7282 , 0,9002]
0,2	10	[-0,7882 , 1,0839]	[-0,7459 , 1,0310]	[-0,6852 , 1,0147]	[-0,5814 , 0,9818]	[-0,8193 , 1,4376]
0,2	100	[-0,7908 , 1,0896]	[-0,7504 , 1,0359]	[-0,6855 , 1,0166]	[-0,5755 , 0,9898]	[-2,5106 , 10,2364]
0,4	0	[-0,8120 , 0,7920]	[-0,7609 , 0,7405]	[-0,6793 , 0,6513]	[-0,6338 , 0,5608]	[-0,6116 , 0,7426]
0,4	1	[-0,8766 , 1,0437]	[-0,8366 , 1,0081]	[-0,7474 , 1,0438]	[-0,7040 , 1,0238]	[-0,7155 , 1,0518]
0,4	5	[-0,8137 , 1,4125]	[-0,7484 , 1,2934]	[-0,7538 , 1,3045]	[-0,7381 , 1,3001]	[-0,9121 , 2,0203]
0,4	10	[-0,8046 , 1,4994]	[-0,7360 , 1,3473]	[-0,7875 , 1,3185]	[-0,7591 , 1,3197]	[-1,1699 , 3,2359]
0,4	100	[-0,7903 , 1,5511]	[-0,7278 , 1,3779]	[-0,7903 , 1,2776]	[-0,7720 , 1,3024]	[-5,5993 , 23,8563]

c_p	c_D	trimmed Liu	Huber ($\theta = 1/3$)	Huber ($\theta = 1$)	Hampel ($\theta = 1/3$)	Hampel ($\theta = 1$)
0	0	[-0,6186 , 0,6124]	[-0,6237 , 0,6157]	[-0,6255 , 0,6175]	[-0,7372 , 0,7260]	[-0,7177 , 0,7072]
0,1	0	[-0,6060 , 0,6060]	[-0,6147 , 0,6136]	[-0,6164 , 0,6153]	[-0,7188 , 0,7147]	[-0,7005 , 0,6977]
0,1	1	[-0,6264 , 0,6637]	[-0,6511 , 0,7057]	[-0,6520 , 0,7066]	[-0,7594 , 0,7783]	[-0,7307 , 0,7488]
0,1	5	[-0,6241 , 0,6890]	[-0,6835 , 0,8553]	[-0,6393 , 0,8000]	[-0,6770 , 0,7026]	[-0,6602 , 0,6850]
0,1	10	[-0,6255 , 0,6925]	[-0,6826 , 0,8678]	[-0,6388 , 0,8091]	[-0,6915 , 0,6929]	[-0,6742 , 0,6758]
0,1	100	[-0,6123 , 0,7005]	[-0,6822 , 0,8635]	[-0,6389 , 0,8039]	[-0,6913 , 0,6870]	[-0,6740 , 0,6701]
0,2	0	[-0,6074 , 0,5916]	[-0,6184 , 0,6016]	[-0,6200 , 0,6033]	[-0,7112 , 0,6945]	[-0,6960 , 0,6770]
0,2	1	[-0,6358 , 0,7527]	[-0,6952 , 0,7878]	[-0,6961 , 0,7896]	[-0,8130 , 0,8253]	[-0,7825 , 0,7957]
0,2	5	[-0,6888 , 1,1250]	[-0,7955 , 1,1266]	[-0,6908 , 1,0022]	[-0,6556 , 0,6600]	[-0,6323 , 0,6485]
0,2	10	[-0,7308 , 1,6282]	[-0,7832 , 1,1701]	[-0,6737 , 1,0320]	[-0,6493 , 0,6563]	[-0,6272 , 0,6468]
0,2	100	[-1,7315 , 9,8117]	[-0,7817 , 1,1894]	[-0,6695 , 1,0453]	[-0,6503 , 0,6599]	[-0,6315 , 0,6431]
0,4	0	[-0,6025 , 0,5872]	[-0,6071 , 0,5915]	[-0,6087 , 0,5930]	[-0,6813 , 0,6701]	[-0,6682 , 0,6558]
0,4	1	[-0,6838 , 0,9216]	[-0,7773 , 0,9765]	[-0,7793 , 0,9812]	[-0,9188 , 0,9619]	[-0,9036 , 0,9508]
0,4	5	[-0,8564 , 1,7241]	[-1,1053 , 1,9458]	[-0,8237 , 1,6439]	[-1,2713 , 1,4865]	[-0,5875 , 0,8700]
0,4	10	[-1,0354 , 2,8694]	[-1,1195 , 2,1378]	[-0,8046 , 1,7670]	[-1,3723 , 1,6332]	[-0,6104 , 0,9872]
0,4	100	[-4,1640 , 21,9958]	[-1,1224 , 2,2419]	[-0,8012 , 1,8275]	[-1,3642 , 1,6439]	[-0,6158 , 1,0367]

Table 4.2: Estimates of the location in Study 1 - CASE 2

c_p	c_D	mean	trimmed ($\theta = 1/3$)	trimmed ($\theta = 1$)	Hausdorff median	1-norm median
0	0	[-1,2947 , 1,3015]	[-1,3956 , 1,3981]	[-1,3447 , 1,3444]	[-1,2095 , 1,2150]	[-1,4633 , 1,4652]
0,1	0	[-1,2939 , 1,2809]	[-1,3841 , 1,3777]	[-1,3375 , 1,3293]	[-1,2089 , 1,1953]	[-1,4636 , 1,4560]
0,1	1	[-1,3066 , 1,4150]	[-1,4291 , 1,4374]	[-1,3765 , 1,3853]	[-1,2503 , 1,2901]	[-1,4778 , 1,5158]
0,1	5	[-1,3715 , 1,9204]	[-1,3435 , 1,3694]	[-1,3123 , 1,3367]	[-1,2045 , 1,3488]	[-1,4812 , 1,5848]
0,1	10	[-1,4764 , 2,5614]	[-1,3538 , 1,3526]	[-1,3223 , 1,3219]	[-1,2090 , 1,3510]	[-1,4862 , 1,5856]
0,1	100	[-3,2593 , 14,0864]	[-1,3591 , 1,3531]	[-1,3262 , 1,3201]	[-1,2126 , 1,3472]	[-1,4832 , 1,5823]
0,2	0	[-1,2774 , 1,2759]	[-1,3596 , 1,3701]	[-1,3204 , 1,3309]	[-1,1904 , 1,1973]	[-1,4600 , 1,4601]
0,2	1	[-1,3201 , 1,5388]	[-1,4747 , 1,4868]	[-1,4205 , 1,4381]	[-1,2953 , 1,3748]	[-1,5038 , 1,5754]
0,2	5	[-1,4828 , 2,5413]	[-1,3110 , 1,3445]	[-1,2812 , 1,3179]	[-1,2302 , 1,5059]	[-1,5116 , 1,7405]
0,2	10	[-1,6984 , 3,7884]	[-1,3050 , 1,3031]	[-1,3031 , 1,3013]	[-1,2213 , 1,5227]	[-1,5140 , 1,7540]
0,2	100	[-5,5409 , 26,4624]	[-1,3000 , 1,3009]	[-1,2990 , 1,2995]	[-1,2184 , 1,5162]	[-1,5093 , 1,7524]
0,4	0	[-1,2665 , 1,2708]	[-1,4485 , 1,4311]	[-1,3690 , 1,3604]	[-1,1830 , 1,1858]	[-1,4572 , 1,4629]
0,4	1	[-1,3691 , 1,7643]	[-1,6746 , 1,6900]	[-1,5907 , 1,6097]	[-1,4081 , 1,5585]	[-1,5785 , 1,7359]
0,4	5	[-1,6821 , 3,7885]	[-1,3330 , 1,3657]	[-1,3030 , 1,3429]	[-1,3332 , 1,9915]	[-1,6147 , 2,3166]
0,4	10	[-2,1175 , 6,2707]	[-1,3496 , 1,3419]	[-1,3254 , 1,3167]	[-1,2789 , 2,0488]	[-1,5922 , 2,3640]
0,4	100	[-10,4824 , 51,7643]	[-1,3496 , 1,3447]	[-1,3247 , 1,3230]	[-1,2995 , 2,0596]	[-1,5973 , 2,3628]

c_p	c_D	spatial ($\theta = 1/3$)	spatial ($\theta = 1$)	Tukey	Liu	trimmed Tukey
0	0	[-1,3794 , 1,3844]	[-1,3268 , 1,3321]	[-1,3156 , 1,3300]	[-1,2973 , 1,3059]	[-1,2869 , 1,2906]
0,1	0	[-1,3799 , 1,3678]	[-1,3274 , 1,3162]	[-1,3111 , 1,3067]	[-1,2956 , 1,2820]	[-1,2869 , 1,2764]
0,1	1	[-1,4030 , 1,4472]	[-1,3509 , 1,3977]	[-1,3388 , 1,3936]	[-1,3178 , 1,3760]	[-1,3139 , 1,3767]
0,1	5	[-1,3926 , 1,5325]	[-1,3362 , 1,4707]	[-1,3307 , 1,4834]	[-1,3107 , 1,4534]	[-1,3277 , 1,5013]
0,1	10	[-1,3984 , 1,5387]	[-1,3412 , 1,4764]	[-1,3352 , 1,4861]	[-1,3222 , 1,4577]	[-1,3332 , 1,5540]
0,1	100	[-1,4009 , 1,5356]	[-1,3441 , 1,4724]	[-1,3592 , 1,4715]	[-1,3327 , 1,4475]	[-1,4755 , 2,2835]
0,2	0	[-1,3661 , 1,3728]	[-1,3145 , 1,3206]	[-1,2944 , 1,3077]	[-1,2749 , 1,2891]	[-1,2729 , 1,2721]
0,2	1	[-1,4373 , 1,5190]	[-1,3863 , 1,4704]	[-1,3753 , 1,4704]	[-1,3462 , 1,4581]	[-1,3448 , 1,4739]
0,2	5	[-1,4363 , 1,7092]	[-1,3724 , 1,6394]	[-1,3678 , 1,6697]	[-1,3532 , 1,6422]	[-1,4594 , 1,9944]
0,2	10	[-1,4325 , 1,7350]	[-1,3679 , 1,6607]	[-1,3629 , 1,6730]	[-1,3599 , 1,6567]	[-1,5850 , 2,5718]
0,2	100	[-1,4319 , 1,7396]	[-1,3666 , 1,6638]	[-1,3804 , 1,6739]	[-1,3617 , 1,6616]	[-3,8129 , 12,6211]
0,4	0	[-1,3616 , 1,3666]	[-1,3091 , 1,3158]	[-1,2764 , 1,2896]	[-1,2603 , 1,2629]	[-1,2610 , 1,2703]
0,4	1	[-1,5339 , 1,6930]	[-1,4841 , 1,6522]	[-1,4565 , 1,6529]	[-1,4435 , 1,6408]	[-1,4398 , 1,6580]
0,4	5	[-1,6051 , 2,2880]	[-1,5104 , 2,2006]	[-1,5841 , 2,3154]	[-1,5875 , 2,3172]	[-1,7136 , 2,7976]
0,4	10	[-1,5873 , 2,4157]	[-1,4824 , 2,3056]	[-1,5952 , 2,4305]	[-1,5879 , 2,4456]	[-2,0093 , 4,0747]
0,4	100	[-1,5979 , 2,4714]	[-1,4982 , 2,3496]	[-1,5715 , 2,3370]	[-1,5546 , 2,3942]	[-8,2275 , 27,4212]

c_p	c_D	trimmed Liu	Huber ($\theta = 1/3$)	Huber ($\theta = 1$)	Hampel ($\theta = 1/3$)	Hampel ($\theta = 1$)
0	0	[-1,2818 , 1,2860]	[-1,3152 , 1,3225]	[-1,3062 , 1,3130]	[-1,4128 , 1,4148]	[-1,3481 , 1,3504]
0,1	0	[-1,2790 , 1,2719]	[-1,3164 , 1,3066]	[-1,3067 , 1,2970]	[-1,4037 , 1,3919]	[-1,3428 , 1,3349]
0,1	1	[-1,3104 , 1,3756]	[-1,3502 , 1,4012]	[-1,3380 , 1,3898]	[-1,4468 , 1,4525]	[-1,3813 , 1,3897]
0,1	5	[-1,3202 , 1,4929]	[-1,3819 , 1,5520]	[-1,3330 , 1,4945]	[-1,3678 , 1,3887]	[-1,3215 , 1,3399]
0,1	10	[-1,3101 , 1,5128]	[-1,3831 , 1,5651]	[-1,3331 , 1,5032]	[-1,3773 , 1,3744]	[-1,3300 , 1,3281]
0,1	100	[-1,2980 , 1,6033]	[-1,3904 , 1,5667]	[-1,3394 , 1,5027]	[-1,3831 , 1,3710]	[-1,3340 , 1,3241]
0,2	0	[-1,2646 , 1,2715]	[-1,3016 , 1,3061]	[-1,2923 , 1,2973]	[-1,3797 , 1,3903]	[-1,3259 , 1,3380]
0,2	1	[-1,3436 , 1,4793]	[-1,3911 , 1,4917]	[-1,3763 , 1,4782]	[-1,4928 , 1,5009]	[-1,4278 , 1,4397]
0,2	5	[-1,4474 , 1,9831]	[-1,4896 , 1,8422]	[-1,3839 , 1,7230]	[-1,3402 , 1,3693]	[-1,2927 , 1,3354]
0,2	10	[-1,5467 , 2,5591]	[-1,4883 , 1,8813]	[-1,3786 , 1,7506]	[-1,3474 , 1,3533]	[-1,3073 , 1,3275]
0,2	100	[-3,3423 , 12,2810]	[-1,4863 , 1,8924]	[-1,3756 , 1,7556]	[-1,3466 , 1,3478]	[-1,3125 , 1,3146]
0,4	0	[-1,2546 , 1,2617]	[-1,2959 , 1,3054]	[-1,2865 , 1,2956]	[-1,3628 , 1,3733]	[-1,3169 , 1,3274]
0,4	1	[-1,4346 , 1,6619]	[-1,4869 , 1,6770]	[-1,4699 , 1,6644]	[-1,6114 , 1,6406]	[-1,5550 , 1,5887]
0,4	5	[-1,6911 , 2,7736]	[-1,8045 , 2,6757]	[-1,5511 , 2,4102]	[-1,9195 , 2,1312]	[-1,4001 , 1,7003]
0,4	10	[-1,9334 , 4,0590]	[-1,7987 , 2,8677]	[-1,5156 , 2,5413]	[-1,9229 , 2,2091]	[-1,3781 , 1,8080]
0,4	100	[-6,9431 , 26,9526]	[-1,8225 , 2,9667]	[-1,5314 , 2,6012]	[-1,9713 , 2,2729]	[-1,3786 , 1,8597]

Table 4.3: Estimates of the location in Study 1 - CASE 2'

c_p	c_D	mean	trimmed ($\theta = 1/3$)	trimmed ($\theta = 1$)
0	0	[-1,0003 , 1,0005]	[-0,6618 , 0,6609]	[-0,5416 , 0,5412]
0,1	0	[-1,2267 , 1,2267]	[-0,7471 , 0,7472]	[-0,6205 , 0,6213]
0,1	1	[-1,2516 , 1,3504]	[-0,7543 , 0,7598]	[-0,6266 , 0,6333]
0,1	5	[-1,3458 , 1,8467]	[-0,7349 , 0,7398]	[-0,6302 , 0,6377]
0,1	10	[-1,4617 , 2,4681]	[-0,7407 , 0,7408]	[-0,6394 , 0,6395]
0,1	100	[-4,0463 , 13,6796]	[-0,7415 , 0,7407]	[-0,6400 , 0,6401]
0,2	0	[-1,4536 , 1,4537]	[-0,8651 , 0,8641]	[-0,7350 , 0,7338]
0,2	1	[-1,4984 , 1,7022]	[-0,8940 , 0,9076]	[-0,7650 , 0,7850]
0,2	5	[-1,7043 , 2,7006]	[-0,9100 , 0,9422]	[-0,8333 , 0,9136]
0,2	10	[-1,9264 , 3,9641]	[-0,9707 , 0,9855]	[-0,9198 , 0,9792]
0,2	100	[-6,4039 , 26,4565]	[-1,0003 , 1,0000]	[-1,0003 , 1,0000]
0,4	0	[-1,8987 , 1,8980]	[-0,7806 , 0,7815]	[-0,6036 , 0,6044]
0,4	1	[-1,9933 , 2,3986]	[-0,8131 , 0,8348]	[-0,6187 , 0,6446]
0,4	5	[-2,4369 , 4,3930]	[-0,7286 , 0,7658]	[-0,6174 , 0,6703]
0,4	10	[-2,8299 , 6,9172]	[-0,7718 , 0,7743]	[-0,6822 , 0,6867]
0,4	100	[-11,9526 , 52,2807]	[-0,7732 , 0,7732]	[-0,6861 , 0,6861]

c_p	c_D	Hausdorff median	1-norm median	spatial ($\theta = 1/3$)	spatial ($\theta = 1$)
0	0	[-0,4552 , 0,4552]	[-0,7384 , 0,7384]	[-0,6538 , 0,6538]	[-0,6031 , 0,6031]
0,1	0	[-0,5423 , 0,5426]	[-0,8410 , 0,8416]	[-0,7621 , 0,7623]	[-0,7032 , 0,7034]
0,1	1	[-0,5325 , 0,5663]	[-0,8440 , 0,8794]	[-0,7649 , 0,7994]	[-0,7016 , 0,7365]
0,1	5	[-0,4884 , 0,6085]	[-0,8201 , 0,9236]	[-0,7347 , 0,8528]	[-0,6676 , 0,7835]
0,1	10	[-0,4793 , 0,6144]	[-0,8107 , 0,9290]	[-0,7206 , 0,8607]	[-0,6540 , 0,7926]
0,1	100	[-0,4857 , 0,6145]	[-0,8179 , 0,9301]	[-0,7198 , 0,8613]	[-0,6524 , 0,7945]
0,2	0	[-0,6602 , 0,6596]	[-0,9639 , 0,9636]	[-0,8944 , 0,8937]	[-0,8282 , 0,8275]
0,2	1	[-0,6459 , 0,7190]	[-0,9722 , 1,0518]	[-0,9066 , 0,9824]	[-0,8317 , 0,9086]
0,2	5	[-0,5545 , 0,8126]	[-0,9278 , 1,1729]	[-0,8563 , 1,1208]	[-0,7664 , 1,0287]
0,2	10	[-0,5333 , 0,8298]	[-0,9088 , 1,1888]	[-0,8232 , 1,1437]	[-0,7357 , 1,0551]
0,2	100	[-0,5357 , 0,8284]	[-0,9090 , 1,1891]	[-0,8044 , 1,1434]	[-0,7162 , 1,0596]
0,4	0	[-1,0044 , 1,0037]	[-1,2769 , 1,2772]	[-1,2391 , 1,2389]	[-1,1667 , 1,1665]
0,4	1	[-1,1168 , 1,2795]	[-1,3387 , 1,5607]	[-1,3459 , 1,5271]	[-1,2628 , 1,4540]
0,4	5	[-0,9615 , 1,5709]	[-1,3100 , 2,1757]	[-1,4078 , 2,1158]	[-1,2360 , 1,9719]
0,4	10	[-0,8819 , 1,6379]	[-1,2441 , 2,3190]	[-1,3262 , 2,2692]	[-1,1460 , 2,1312]
0,4	100	[-0,9033 , 1,6537]	[-1,2688 , 2,3228]	[-1,2938 , 2,3274]	[-1,1139 , 2,2037]

c_p	c_D	Huber ($\theta = 1/3$)	Huber ($\theta = 1$)	Hampel ($\theta = 1/3$)	Hampel ($\theta = 1$)
0	0	[-0,8604 , 0,8604]	[-0,7507 , 0,7507]	[-0,6241 , 0,6237]	[-0,4996 , 0,4992]
0,1	0	[-0,9916 , 0,9917]	[-0,8522 , 0,8523]	[-0,6989 , 0,6989]	[-0,5648 , 0,5653]
0,1	1	[-1,0037 , 1,0445]	[-0,8527 , 0,8923]	[-0,7019 , 0,7079]	[-0,5664 , 0,5729]
0,1	5	[-0,9813 , 1,1230]	[-0,8156 , 0,9490]	[-0,6870 , 0,6923]	[-0,5753 , 0,5827]
0,1	10	[-0,9652 , 1,1337]	[-0,8009 , 0,9606]	[-0,6928 , 0,6934]	[-0,5848 , 0,5849]
0,1	100	[-0,9653 , 1,1354]	[-0,8000 , 0,9630]	[-0,6933 , 0,6935]	[-0,5856 , 0,5856]
0,2	0	[-1,1379 , 1,1374]	[-0,9694 , 0,9688]	[-0,8017 , 0,8006]	[-0,6574 , 0,6562]
0,2	1	[-1,1624 , 1,2509]	[-0,9717 , 1,0579]	[-0,8191 , 0,8321]	[-0,6705 , 0,6872]
0,2	5	[-1,1349 , 1,4469]	[-0,9007 , 1,1968]	[-0,7958 , 0,8257]	[-0,6779 , 0,7320]
0,2	10	[-1,0967 , 1,4780]	[-0,8664 , 1,2292]	[-0,8200 , 0,8331]	[-0,7149 , 0,7516]
0,2	100	[-1,0774 , 1,4801]	[-0,8475 , 1,2365]	[-0,8361 , 0,8377]	[-0,7499 , 0,7542]
0,4	0	[-1,4655 , 1,4651]	[-1,2571 , 1,2569]	[-1,0977 , 1,0974]	[-0,9228 , 0,9225]
0,4	1	[-1,5489 , 1,7515]	[-1,3266 , 1,5322]	[-1,2039 , 1,2614]	[-0,9992 , 1,0826]
0,4	5	[-1,6264 , 2,4161]	[-1,2958 , 2,0868]	[-1,4608 , 1,7393]	[-1,0202 , 1,4704]
0,4	10	[-1,5230 , 2,5876]	[-1,1911 , 2,2559]	[-1,5515 , 1,9162]	[-0,9521 , 1,6143]
0,4	100	[-1,4870 , 2,6589]	[-1,1548 , 2,3338]	[-1,6457 , 1,9672]	[-0,9848 , 1,6567]

Table 4.4: Estimates of the location in Study 2 - CASE 1

c_p	c_D	mean	trimmed ($\theta = 1/3$)	trimmed ($\theta = 1$)
0	0	[-0,5998 , 0,6004]	[-0,7067 , 0,7073]	[-0,7071 , 0,7074]
0,1	0	[-0,5929 , 0,5921]	[-0,6914 , 0,6911]	[-0,6918 , 0,6914]
0,1	1	[-0,6162 , 0,7171]	[-0,7493 , 0,7553]	[-0,7444 , 0,7506]
0,1	5	[-0,6977 , 1,2154]	[-0,6560 , 0,6609]	[-0,6562 , 0,6615]
0,1	10	[-0,8464 , 1,8256]	[-0,6573 , 0,6575]	[-0,6575 , 0,6578]
0,1	100	[-3,0848 , 13,1638]	[-0,6572 , 0,6571]	[-0,6576 , 0,6573]
0,2	0	[-0,5855 , 0,5855]	[-0,6774 , 0,6762]	[-0,6779 , 0,6766]
0,2	1	[-0,6308 , 0,8352]	[-0,7934 , 0,8080]	[-0,7912 , 0,8060]
0,2	5	[-0,8379 , 1,8233]	[-0,5981 , 0,6207]	[-0,5939 , 0,6171]
0,2	10	[-1,1274 , 3,0817]	[-0,5992 , 0,6016]	[-0,5992 , 0,6017]
0,2	100	[-5,7732 , 25,4170]	[-0,5999 , 0,5999]	[-0,5999 , 0,5999]
0,4	0	[-0,5712 , 0,5700]	[-0,8012 , 0,8028]	[-0,8008 , 0,8016]
0,4	1	[-0,6677 , 1,0697]	[-1,0606 , 1,0839]	[-1,0491 , 1,0728]
0,4	5	[-1,0692 , 3,0624]	[-0,6364 , 0,6664]	[-0,6365 , 0,6673]
0,4	10	[-1,5941 , 5,5766]	[-0,6410 , 0,6439]	[-0,6412 , 0,6442]
0,4	100	[-10,3318 , 50,6535]	[-0,6423 , 0,6425]	[-0,6425 , 0,6428]

c_p	c_D	Hausdorff median	1-norm median	spatial ($\theta = 1/3$)	spatial ($\theta = 1$)
0	0	[-0,5797 , 0,5805]	[-1,0781 , 1,0788]	[-0,8380 , 0,8388]	[-0,7859 , 0,7868]
0,1	0	[-0,5697 , 0,5692]	[-1,0746 , 1,0744]	[-0,8324 , 0,8322]	[-0,7801 , 0,7798]
0,1	1	[-0,6187 , 0,6533]	[-1,0871 , 1,1142]	[-0,8544 , 0,8899]	[-0,8009 , 0,8416]
0,1	5	[-0,5722 , 0,6978]	[-1,0871 , 1,1432]	[-0,8393 , 0,9613]	[-0,7783 , 0,9067]
0,1	10	[-0,5708 , 0,7027]	[-1,0878 , 1,1437]	[-0,8405 , 0,9701]	[-0,7788 , 0,9138]
0,1	100	[-0,5691 , 0,7048]	[-1,0871 , 1,1437]	[-0,8390 , 0,9732]	[-0,7781 , 0,9165]
0,2	0	[-0,5599 , 0,5587]	[-1,0711 , 1,0700]	[-0,8271 , 0,8259]	[-0,7745 , 0,7733]
0,2	1	[-0,6671 , 0,7404]	[-1,0911 , 1,1382]	[-0,8691 , 0,9446]	[-0,8155 , 0,9002]
0,2	5	[-0,5772 , 0,8331]	[-1,0786 , 1,1686]	[-0,8352 , 1,0790]	[-0,7796 , 1,0253]
0,2	10	[-0,5660 , 0,8494]	[-1,0787 , 1,1709]	[-0,8246 , 1,0967]	[-0,7740 , 1,0412]
0,2	100	[-0,5614 , 0,8484]	[-1,0716 , 1,1709]	[-0,8185 , 1,1005]	[-0,7689 , 1,0450]
0,4	0	[-0,5360 , 0,5365]	[-1,0598 , 1,0603]	[-0,8117 , 0,8121]	[-0,7587 , 0,7591]
0,4	1	[-0,7960 , 0,9573]	[-1,0849 , 1,1644]	[-0,8969 , 1,0622]	[-0,8503 , 1,0262]
0,4	5	[-0,5671 , 1,1928]	[-1,0413 , 1,3759]	[-0,8555 , 1,4236]	[-0,7798 , 1,2962]
0,4	10	[-0,5145 , 1,2468]	[-1,0225 , 1,4009]	[-0,8293 , 1,5208]	[-0,7511 , 1,3602]
0,4	100	[-0,5027 , 1,2504]	[-0,9973 , 1,4011]	[-0,8111 , 1,5608]	[-0,7340 , 1,3833]

c_p	c_D	Huber ($\theta = 1/3$)	Huber ($\theta = 1$)	Hampel ($\theta = 1/3$)	Hampel ($\theta = 1$)
0	0	[-0,6196 , 0,6202]	[-0,6214 , 0,6220]	[-0,7319 , 0,7325]	[-0,7102 , 0,7107]
0,1	0	[-0,6159 , 0,6152]	[-0,6176 , 0,6170]	[-0,7181 , 0,7178]	[-0,6979 , 0,6975]
0,1	1	[-0,6614 , 0,7048]	[-0,6621 , 0,7061]	[-0,7737 , 0,7821]	[-0,7423 , 0,7501]
0,1	5	[-0,7007 , 0,8537]	[-0,6510 , 0,7955]	[-0,6850 , 0,6924]	[-0,6660 , 0,6735]
0,1	10	[-0,7063 , 0,8710]	[-0,6520 , 0,8058]	[-0,6867 , 0,6894]	[-0,6681 , 0,6703]
0,1	100	[-0,7050 , 0,8748]	[-0,6517 , 0,8087]	[-0,6869 , 0,6888]	[-0,6683 , 0,6697]
0,2	0	[-0,6117 , 0,6109]	[-0,6134 , 0,6125]	[-0,7055 , 0,7043]	[-0,6866 , 0,6854]
0,2	1	[-0,7017 , 0,7934]	[-0,7025 , 0,7954]	[-0,8207 , 0,8370]	[-0,7893 , 0,8063]
0,2	5	[-0,8224 , 1,1419]	[-0,7049 , 1,0093]	[-0,6446 , 0,6639]	[-0,6208 , 0,6529]
0,2	10	[-0,8230 , 1,1878]	[-0,6983 , 1,0391]	[-0,6479 , 0,6550]	[-0,6249 , 0,6461]
0,2	100	[-0,8192 , 1,1974]	[-0,6944 , 1,0432]	[-0,6520 , 0,6554]	[-0,6334 , 0,6372]
0,4	0	[-0,6015 , 0,6016]	[-0,6031 , 0,6031]	[-0,6778 , 0,6785]	[-0,6620 , 0,6625]
0,4	1	[-0,7865 , 0,9855]	[-0,7889 , 0,9904]	[-0,9303 , 0,9761]	[-0,9175 , 0,9669]
0,4	5	[-1,1551 , 1,9705]	[-0,8613 , 1,6572]	[-1,3990 , 1,5979]	[-0,5915 , 0,8604]
0,4	10	[-1,1555 , 2,1715]	[-0,8245 , 1,7863]	[-1,4481 , 1,7198]	[-0,5834 , 0,9762]
0,4	100	[-1,1416 , 2,2496]	[-0,8064 , 1,8308]	[-1,4106 , 1,7003]	[-0,5956 , 1,0378]

Table 4.5: Estimates of the location in Study 2 - CASE 2

c_p	c_D	mean	trimmed ($\theta = 1/3$)	trimmed ($\theta = 1$)
0	0	[-1,2981 , 1,2973]	[-1,3998 , 1,3994]	[-1,3478 , 1,3471]
0,1	0	[-1,2903 , 1,2901]	[-1,3864 , 1,3862]	[-1,3393 , 1,3392]
0,1	1	[-1,3159 , 1,4147]	[-1,4404 , 1,4465]	[-1,3841 , 1,3910]
0,1	5	[-1,3955 , 1,9135]	[-1,3538 , 1,3601]	[-1,3212 , 1,3277]
0,1	10	[-1,5622 , 2,5296]	[-1,3550 , 1,3552]	[-1,3235 , 1,3239]
0,1	100	[-3,9820 , 13,7923]	[-1,3553 , 1,3553]	[-1,3242 , 1,3241]
0,2	0	[-1,2828 , 1,2838]	[-1,3721 , 1,3735]	[-1,3303 , 1,3317]
0,2	1	[-1,3313 , 1,5324]	[-1,4842 , 1,5001]	[-1,4284 , 1,4462]
0,2	5	[-1,5524 , 2,5278]	[-1,3135 , 1,3356]	[-1,2879 , 1,3126]
0,2	10	[-1,7305 , 3,7884]	[-1,2967 , 1,2998]	[-1,2964 , 1,2996]
0,2	100	[-6,8962 , 26,1227]	[-1,2985 , 1,2981]	[-1,2985 , 1,2981]
0,4	0	[-1,2675 , 1,2670]	[-1,4608 , 1,4583]	[-1,3792 , 1,3773]
0,4	1	[-1,3736 , 1,7676]	[-1,6983 , 1,7201]	[-1,6040 , 1,6301]
0,4	5	[-1,7405 , 3,7759]	[-1,3347 , 1,3639]	[-1,3035 , 1,3369]
0,4	10	[-2,3565 , 6,2391]	[-1,3400 , 1,3420]	[-1,3157 , 1,3180]
0,4	100	[-11,2494 , 51,3700]	[-1,3406 , 1,3410]	[-1,3166 , 1,3171]

c_p	c_D	Hausdorff median	1-norm median	spatial ($\theta = 1/3$)	spatial ($\theta = 1$)
0	0	[-1,2114 , 1,2105]	[-1,4642 , 1,4634]	[-1,3845 , 1,3837]	[-1,3306 , 1,3297]
0,1	0	[-1,2044 , 1,2039]	[-1,4624 , 1,4624]	[-1,3799 , 1,3797]	[-1,3261 , 1,3259]
0,1	1	[-1,2535 , 1,2874]	[-1,4860 , 1,5166]	[-1,4149 , 1,4495]	[-1,3604 , 1,3968]
0,1	5	[-1,2168 , 1,3420]	[-1,4888 , 1,5855]	[-1,4110 , 1,5316]	[-1,3484 , 1,4686]
0,1	10	[-1,2166 , 1,3468]	[-1,4933 , 1,5890]	[-1,4151 , 1,5414]	[-1,3514 , 1,4761]
0,1	100	[-1,2162 , 1,3481]	[-1,4925 , 1,5893]	[-1,4148 , 1,5445]	[-1,3511 , 1,4786]
0,2	0	[-1,1971 , 1,1985]	[-1,4608 , 1,4614]	[-1,3751 , 1,3763]	[-1,3214 , 1,3225]
0,2	1	[-1,2988 , 1,3721]	[-1,5094 , 1,5773]	[-1,4479 , 1,5225]	[-1,3934 , 1,4721]
0,2	5	[-1,2469 , 1,4995]	[-1,5367 , 1,7436]	[-1,4684 , 1,7131]	[-1,3950 , 1,6388]
0,2	10	[-1,2176 , 1,5186]	[-1,5173 , 1,7527]	[-1,4422 , 1,7387]	[-1,3695 , 1,6625]
0,2	100	[-1,2329 , 1,5129]	[-1,5324 , 1,7532]	[-1,4635 , 1,7468]	[-1,3864 , 1,6652]
0,4	0	[-1,1825 , 1,1819]	[-1,4575 , 1,4568]	[-1,3654 , 1,3646]	[-1,3116 , 1,3110]
0,4	1	[-1,4078 , 1,5661]	[-1,5790 , 1,7423]	[-1,5379 , 1,7024]	[-1,4852 , 1,6603]
0,4	5	[-1,3500 , 1,9908]	[-1,6454 , 2,3160]	[-1,6389 , 2,2948]	[-1,5354 , 2,2006]
0,4	10	[-1,3251 , 2,0403]	[-1,6492 , 2,3771]	[-1,6660 , 2,4368]	[-1,5395 , 2,3108]
0,4	100	[-1,3009 , 2,0382]	[-1,6264 , 2,3783]	[-1,6367 , 2,4838]	[-1,5141 , 2,3497]

c_p	c_D	Huber ($\theta = 1/3$)	Huber ($\theta = 1$)	Hampel ($\theta = 1/3$)	Hampel ($\theta = 1$)
0	0	[-1,3189 , 1,3182]	[-1,3099 , 1,3091]	[-1,4179 , 1,4175]	[-1,3560 , 1,3554]
0,1	0	[-1,3143 , 1,3141]	[-1,3050 , 1,3048]	[-1,4059 , 1,4056]	[-1,3479 , 1,3477]
0,1	1	[-1,3605 , 1,4031]	[-1,3476 , 1,3914]	[-1,4563 , 1,4651]	[-1,3911 , 1,3992]
0,1	5	[-1,4011 , 1,5557]	[-1,3457 , 1,4949]	[-1,3754 , 1,3845]	[-1,3289 , 1,3374]
0,1	10	[-1,4103 , 1,5736]	[-1,3498 , 1,5054]	[-1,3767 , 1,3802]	[-1,3312 , 1,3336]
0,1	100	[-1,4106 , 1,5778]	[-1,3506 , 1,5087]	[-1,3772 , 1,3802]	[-1,3319 , 1,3338]
0,2	0	[-1,3093 , 1,3105]	[-1,2999 , 1,3010]	[-1,3931 , 1,3944]	[-1,3392 , 1,3407]
0,2	1	[-1,4003 , 1,4914]	[-1,3850 , 1,4785]	[-1,4992 , 1,5185]	[-1,4345 , 1,4538]
0,2	5	[-1,5319 , 1,8502]	[-1,4150 , 1,7225]	[-1,3446 , 1,3653]	[-1,2990 , 1,3314]
0,2	10	[-1,5002 , 1,8905]	[-1,3828 , 1,7547]	[-1,3418 , 1,3512]	[-1,3014 , 1,3266]
0,2	100	[-1,5340 , 1,9063]	[-1,4068 , 1,7587]	[-1,3447 , 1,3487]	[-1,3118 , 1,3153]
0,4	0	[-1,2996 , 1,2986]	[-1,2897 , 1,2888]	[-1,3679 , 1,3665]	[-1,3234 , 1,3219]
0,4	1	[-1,4884 , 1,6846]	[-1,4711 , 1,6727]	[-1,6109 , 1,6574]	[-1,5544 , 1,6065]
0,4	5	[-1,8474 , 2,6848]	[-1,5847 , 2,4106]	[-1,9225 , 2,1278]	[-1,4001 , 1,6912]
0,4	10	[-1,8996 , 2,9007]	[-1,5845 , 2,5494]	[-1,9332 , 2,2029]	[-1,3705 , 1,7857]
0,4	100	[-1,8678 , 2,9805]	[-1,5523 , 2,5983]	[-1,9916 , 2,2813]	[-1,3806 , 1,8592]

Table 4.6: Estimates of the location in Study 2 - CASE 2'

For the bias, variance and mean square error, the conclusions for the three cases in Studies 1 and 2 are summarized in Tables 4.7 and 4.8.

STUDY 1		CASE 1	CASE 2	CASE 2'
Bias	$c_p \leq 0.2$	**Hampel**	**Hampel** $(\theta = 1)$	**Hampel** $(\theta = 1)$
	$c_p = 0.4$	**trimmed**	**trimmed** $(\theta = 1)$	**trimmed** $(\theta = 1)$
	$c_D = 0$		**1-norm median**	**1-norm median**
	Dispersion	none	low	low
Variance	$c_p = 0$	**Huber** $(\theta = 1)$	**mean**	**1-norm median**
	$c_p \leq 0.2$	**Huber** **Hampel** $(\theta = 1)$	**1-norm median** **trimmed Liu**	**1-norm median** **trimmed** $(\theta = 1)$
	$c_p = 0.4$	**trimmed** $(\theta = 1)$	**trimmed** $(\theta = 1)$	**trimmed** $(\theta = 1)$
	$c_D = 0$			
	Dispersion	none	medium	none
MSE	$c_p = 0$	**Huber** $(\theta = 1)$	**mean**	**1-norm median**
	$c_p \leq 0.2$	**Hampel**	**1-norm median** **trimmed Liu** **Hampel** $(\theta = 1)$	**1-norm median**
	$c_p = 0.4$	**trimmed**	**trimmed** $(\theta = 1/3)$	**trimmed** $(\theta = 1)$
	$c_D = 0$			
	Dispersion	none	medium	low

Table 4.7: Summary of the main conclusions from Study 1: the most suitable (if any) location measures/estimates

STUDY 2		CASE 1	CASE 2	CASE 2'
Bias	$c_p \leq 0.2$	**Hampel**	**1-norm median** **Hampel** $(\theta = 1)$	**Hampel** $(\theta = 1)$
	$c_p = 0.4$	**trimmed**	**1-norm median** **trimmed** $(\theta = 1)$	**trimmed** $(\theta = 1)$
	$c_D = 0$			**1-norm median**
	Dispersion	none	medium	medium
Variance	$c_p = 0$	**Huber** $(\theta = 1)$	**mean**	**1-norm median**
	$c_p \leq 0.2$	**Hampel**	**Huber** **Hampel** $(\theta = 1)$	**Hampel** **trimmed** $(\theta = 1)$
	$c_p = 0.4$	**trimmed**	**trimmed** $(\theta = 1)$	**trimmed** $(\theta = 1)$
	$c_D = 0/1$		**1-norm median**	**1-norm median**
	Dispersion	none	none	low
MSE	$c_p = 0$	**Huber** $(\theta = 1)$	**mean**	**1-norm median**
	$c_p \leq 0.2$	**Hampel**	**Hampel** $(\theta = 1)$	**Hampel** $(\theta = 1)$
	$c_p = 0.4$	**trimmed**	**trimmed** $(\theta = 1)$	**trimmed** $(\theta = 1)$
	$c_D = 0/1$		**1-norm median**	**1-norm median**
	Dispersion	none	medium	low

Table 4.8: Summary of the main conclusions from Study 2: the most suitable (if any) location measures/estimates

Concerning **Study 3** and **Study 4**, the estimates are given by

c_p	c_D	mean	trimmed ($\theta = 1/3$)	trimmed ($\theta = 1$)	Hausdorff median	1-norm median
0	0	[0,7582 , 0,9094]	[0,8080 , 0,9311]	[0,8117 , 0,9286]	[0,7912 , 0,9088]	[0,7825 , 0,9331]
0,1	0	[0,7265 , 0,8975]	[0,7960 , 0,9261]	[0,7985 , 0,9235]	[0,7762 , 0,9027]	[0,7675 , 0,9277]
0,1	1	[0,7115 , 0,8895]	[0,7915 , 0,9249]	[0,7936 , 0,9220]	[0,7713 , 0,9009]	[0,7617 , 0,9259]
0,1	5	[0,6965 , 0,8804]	[0,7912 , 0,9243]	[0,7931 , 0,9217]	[0,7718 , 0,8996]	[0,7614 , 0,9259]
0,1	10	[0,6913 , 0,8773]	[0,7909 , 0,9244]	[0,7926 , 0,9218]	[0,7718 , 0,8984]	[0,7606 , 0,9257]
0,1	100	[0,6848 , 0,8747]	[0,7904 , 0,9239]	[0,7921 , 0,9214]	[0,7735 , 0,8974]	[0,7603 , 0,9259]
0,2	0	[0,6941 , 0,8861]	[0,7790 , 0,9200]	[0,7809 , 0,9172]	[0,7570 , 0,8951]	[0,7475 , 0,9213]
0,2	1	[0,6673 , 0,8718]	[0,7685 , 0,9159]	[0,7700 , 0,9134]	[0,7478 , 0,8908]	[0,7383 , 0,9176]
0,2	5	[0,6344 , 0,8565]	[0,7595 , 0,9127]	[0,7604 , 0,9110]	[0,7442 , 0,8881]	[0,7328 , 0,9173]
0,2	10	[0,6241 , 0,8506]	[0,7586 , 0,9114]	[0,7593 , 0,9102]	[0,7470 , 0,8867]	[0,7328 , 0,9177]
0,2	100	[0,6114 , 0,8431]	[0,7593 , 0,9103]	[0,7596 , 0,9098]	[0,7497 , 0,8833]	[0,7336 , 0,9171]
0,4	0	[0,6296 , 0,8630]	[0,7973 , 0,9270]	[0,8067 , 0,9256]	[0,7065 , 0,8743]	[0,6957 , 0,9072]
0,4	1	[0,5739 , 0,8367]	[0,7884 , 0,9242]	[0,7932 , 0,9216]	[0,6677 , 0,8557]	[0,6514 , 0,8964]
0,4	5	[0,5069 , 0,8014]	[0,7834 , 0,9213]	[0,7852 , 0,9188]	[0,6440 , 0,8365]	[0,6187 , 0,8935]
0,4	10	[0,4871 , 0,7973]	[0,7838 , 0,9217]	[0,7853 , 0,9194]	[0,6437 , 0,8333]	[0,6162 , 0,8946]
0,4	100	[0,4605 , 0,7820]	[0,7837 , 0,9210]	[0,7852 , 0,9188]	[0,6494 , 0,8297]	[0,6168 , 0,8925]

c_p	c_D	spatial ($\theta = 1/3$)	spatial ($\theta = 1$)	Tukey	Liu	trimmed Tukey
0	0	[0,7836 , 0,9217]	[0,7876 , 0,9212]	[0,7960 , 0,9254]	[0,7961 , 0,9254]	[0,7708 , 0,9150]
0,1	0	[0,7685 , 0,9158]	[0,7725 , 0,9154]	[0,7809 , 0,9196]	[0,7822 , 0,9201]	[0,7490 , 0,9069]
0,1	1	[0,7632 , 0,9141]	[0,7673 , 0,9136]	[0,7770 , 0,9183]	[0,7774 , 0,9184]	[0,7405 , 0,9028]
0,1	5	[0,7628 , 0,9139]	[0,7668 , 0,9134]	[0,7764 , 0,9184]	[0,7764 , 0,9185]	[0,7356 , 0,8998]
0,1	10	[0,7619 , 0,9135]	[0,7659 , 0,9131]	[0,7755 , 0,9185]	[0,7756 , 0,9180]	[0,7332 , 0,8984]
0,1	100	[0,7618 , 0,9137]	[0,7660 , 0,9132]	[0,7751 , 0,9188]	[0,7762 , 0,9189]	[0,7327 , 0,8974]
0,2	0	[0,7487 , 0,9089]	[0,7529 , 0,9086]	[0,7627 , 0,9138]	[0,7622 , 0,9133]	[0,7231 , 0,8982]
0,2	1	[0,7392 , 0,9049]	[0,7435 , 0,9045]	[0,7555 , 0,9103]	[0,7557 , 0,9107]	[0,7024 , 0,8880]
0,2	5	[0,7336 , 0,9038]	[0,7379 , 0,9037]	[0,7505 , 0,9105]	[0,7487 , 0,9102]	[0,6754 , 0,8769]
0,2	10	[0,7339 , 0,9044]	[0,7382 , 0,9041]	[0,7508 , 0,9113]	[0,7497 , 0,9113]	[0,6685 , 0,8731]
0,2	100	[0,7344 , 0,9040]	[0,7389 , 0,9038]	[0,7507 , 0,9117]	[0,7489 , 0,9109]	[0,6607 , 0,8685]
0,4	0	[0,6965 , 0,8912]	[0,7018 , 0,8911]	[0,7167 , 0,8996]	[0,7155 , 0,8989]	[0,6830 , 0,8859]
0,4	1	[0,6540 , 0,8771]	[0,6604 , 0,8768]	[0,6829 , 0,8871]	[0,6819 , 0,8869]	[0,6327 , 0,8655]
0,4	5	[0,6218 , 0,8682]	[0,6301 , 0,8680]	[0,6613 , 0,8814]	[0,6584 , 0,8801]	[0,5683 , 0,8373]
0,4	10	[0,6179 , 0,8687]	[0,6265 , 0,8692]	[0,6568 , 0,8844]	[0,6548 , 0,8841]	[0,5486 , 0,8347]
0,4	100	[0,6197 , 0,8677]	[0,6280 , 0,8684]	[0,6562 , 0,8849]	[0,6504 , 0,8833]	[0,5239 , 0,8226]

c_p	c_D	trimmed Liu	Huber ($\theta = 1/3$)	Huber ($\theta = 1$)	Hampel ($\theta = 1/3$)	Hampel ($\theta = 1$)
0	0	[0,7695 , 0,9146]	[0,7582 , 0,9094]	[0,7582 , 0,9094]	[0,8145 , 0,9336]	[0,8205 , 0,9316]
0,1	0	[0,7459 , 0,9055]	[0,7265 , 0,8975]	[0,7265 , 0,8975]	[0,8038 , 0,9291]	[0,8086 , 0,9269]
0,1	1	[0,7372 , 0,9013]	[0,7115 , 0,8895]	[0,7115 , 0,8895]	[0,8005 , 0,9283]	[0,8045 , 0,9259]
0,1	5	[0,7330 , 0,8984]	[0,6965 , 0,8804]	[0,6965 , 0,8804]	[0,7990 , 0,9273]	[0,8030 , 0,9252]
0,1	10	[0,7324 , 0,8975]	[0,6913 , 0,8773]	[0,6913 , 0,8773]	[0,7987 , 0,9274]	[0,8026 , 0,9252]
0,1	100	[0,7350 , 0,8960]	[0,6848 , 0,8747]	[0,6848 , 0,8747]	[0,7983 , 0,9271]	[0,8023 , 0,9249]
0,2	0	[0,7195 , 0,8962]	[0,6941 , 0,8861]	[0,6941 , 0,8861]	[0,7886 , 0,9236]	[0,7930 , 0,9213]
0,2	1	[0,6976 , 0,8856]	[0,6673 , 0,8718]	[0,6673 , 0,8718]	[0,7806 , 0,9204]	[0,7842 , 0,9181]
0,2	5	[0,6683 , 0,8733]	[0,6344 , 0,8565]	[0,6344 , 0,8565]	[0,7765 , 0,9191]	[0,7789 , 0,9171]
0,2	10	[0,6601 , 0,8684]	[0,6241 , 0,8506]	[0,6241 , 0,8506]	[0,7770 , 0,9189]	[0,7790 , 0,9171]
0,2	100	[0,6521 , 0,8617]	[0,6114 , 0,8431]	[0,6114 , 0,8431]	[0,7783 , 0,9189]	[0,7802 , 0,9172]
0,4	0	[0,6805 , 0,8842]	[0,6296 , 0,8630]	[0,6296 , 0,8630]	[0,7492 , 0,9097]	[0,7531 , 0,9067]
0,4	1	[0,6271 , 0,8620]	[0,5739 , 0,8367]	[0,5739 , 0,8367]	[0,7211 , 0,9028]	[0,7257 , 0,8992]
0,4	5	[0,5570 , 0,8311]	[0,5069 , 0,8014]	[0,5069 , 0,8014]	[0,6997 , 0,9044]	[0,7110 , 0,9016]
0,4	10	[0,5339 , 0,8267]	[0,4871 , 0,7973]	[0,4871 , 0,7973]	[0,7079 , 0,9077]	[0,7252 , 0,9061]
0,4	100	[0,5064 , 0,8116]	[0,4605 , 0,7820]	[0,4605 , 0,7820]	[0,7228 , 0,9087]	[0,7386 , 0,9082]

Table 4.9: Estimates of the location in Study 3 - CASE 3

c_p	c_D	mean	trimmed ($\theta = 1/3$)	trimmed ($\theta = 1$)	Hausdorff median	1-norm median
0	0	[0,7526 , 0,9141]	[0,8363 , 0,9434]	[0,8376 , 0,9408]	[0,8203 , 0,9198]	[0,8104 , 0,9496]
0,1	0	[0,7337 , 0,8977]	[0,8265 , 0,9397]	[0,8277 , 0,9370]	[0,8121 , 0,9134]	[0,7990 , 0,9460]
0,1	1	[0,7261 , 0,8854]	[0,8226 , 0,9375]	[0,8240 , 0,9347]	[0,8102 , 0,9083]	[0,7954 , 0,9428]
0,1	5	[0,7189 , 0,8726]	[0,8220 , 0,9361]	[0,8233 , 0,9335]	[0,8129 , 0,9043]	[0,7955 , 0,9414]
0,1	10	[0,7177 , 0,8677]	[0,8235 , 0,9360]	[0,8249 , 0,9333]	[0,8163 , 0,9033]	[0,7980 , 0,9411]
0,1	100	[0,7136 , 0,8633]	[0,8230 , 0,9357]	[0,8244 , 0,9330]	[0,8157 , 0,9025]	[0,7969 , 0,9413]
0,2	0	[0,7145 , 0,8815]	[0,8141 , 0,9356]	[0,8156 , 0,9327]	[0,8024 , 0,9058]	[0,7864 , 0,9413]
0,2	1	[0,7031 , 0,8594]	[0,8084 , 0,9300]	[0,8098 , 0,9272]	[0,8018 , 0,8969]	[0,7824 , 0,9350]
0,2	5	[0,6891 , 0,8331]	[0,8048 , 0,9271]	[0,8061 , 0,9247]	[0,8058 , 0,8874]	[0,7806 , 0,9321]
0,2	10	[0,6844 , 0,8246]	[0,8056 , 0,9263]	[0,8068 , 0,9239]	[0,8102 , 0,8850]	[0,7829 , 0,9309]
0,2	100	[0,6800 , 0,8176]	[0,8073 , 0,9252]	[0,8086 , 0,9227]	[0,8115 , 0,8830]	[0,7844 , 0,9307]
0,4	0	[0,6801 , 0,8506]	[0,8735 , 0,9563]	[0,8757 , 0,9537]	[0,7811 , 0,8884]	[0,7585 , 0,9299]
0,4	1	[0,6571 , 0,8081]	[0,8671 , 0,9491]	[0,8687 , 0,9465]	[0,7710 , 0,8626]	[0,7382 , 0,9107]
0,4	5	[0,6335 , 0,7597]	[0,8670 , 0,9444]	[0,8689 , 0,9419]	[0,7764 , 0,8407]	[0,7327 , 0,8985]
0,4	10	[0,6190 , 0,7403]	[0,8681 , 0,9432]	[0,8699 , 0,9407]	[0,7815 , 0,8357]	[0,7341 , 0,8964]
0,4	100	[0,6038 , 0,7188]	[0,8675 , 0,9416]	[0,8693 , 0,9389]	[0,7842 , 0,8269]	[0,7328 , 0,8931]

c_p	c_D	spatial ($\theta = 1/3$)	spatial ($\theta = 1$)	Tukey	Liu	trimmed Tukey
0	0	[0,8115 , 0,9352]	[0,8167 , 0,9353]	[0,8369 , 0,9448]	[0,8368 , 0,9446]	[0,7815 , 0,9252]
0,1	0	[0,8016 , 0,9303]	[0,8070 , 0,9301]	[0,8288 , 0,9393]	[0,8294 , 0,9394]	[0,7680 , 0,9145]
0,1	1	[0,7989 , 0,9266]	[0,8043 , 0,9262]	[0,8293 , 0,9361]	[0,8294 , 0,9363]	[0,7635 , 0,9066]
0,1	5	[0,7992 , 0,9246]	[0,8048 , 0,9241]	[0,8296 , 0,9341]	[0,8306 , 0,9344]	[0,7604 , 0,8983]
0,1	10	[0,8016 , 0,9245]	[0,8072 , 0,9240]	[0,8302 , 0,9335]	[0,8298 , 0,9334]	[0,7618 , 0,8961]
0,1	100	[0,8006 , 0,9244]	[0,8063 , 0,9239]	[0,8302 , 0,9332]	[0,8305 , 0,9337]	[0,7597 , 0,8960]
0,2	0	[0,7902 , 0,9246]	[0,7957 , 0,9240]	[0,8214 , 0,9340]	[0,8220 , 0,9342]	[0,7534 , 0,9031]
0,2	1	[0,7878 , 0,9175]	[0,7934 , 0,9166]	[0,8211 , 0,9264]	[0,8210 , 0,9264]	[0,7445 , 0,8851]
0,2	5	[0,7877 , 0,9126]	[0,7937 , 0,9114]	[0,8221 , 0,9201]	[0,8223 , 0,9204]	[0,7338 , 0,8628]
0,2	10	[0,7898 , 0,9118]	[0,7960 , 0,9105]	[0,8261 , 0,9198]	[0,8247 , 0,9195]	[0,7314 , 0,8561]
0,2	100	[0,7908 , 0,9110]	[0,7972 , 0,9098]	[0,8247 , 0,9185]	[0,8240 , 0,9187]	[0,7285 , 0,8524]
0,4	0	[0,7655 , 0,9104]	[0,7714 , 0,9090]	[0,7985 , 0,9174]	[0,7996 , 0,9177]	[0,7528 , 0,8929]
0,4	1	[0,7505 , 0,8888]	[0,7563 , 0,8863]	[0,7850 , 0,8904]	[0,7865 , 0,8906]	[0,7316 , 0,8554]
0,4	5	[0,7501 , 0,8724]	[0,7559 , 0,8686]	[0,7755 , 0,8580]	[0,7720 , 0,8529]	[0,7083 , 0,8061]
0,4	10	[0,7524 , 0,8700]	[0,7582 , 0,8659]	[0,7755 , 0,8518]	[0,7727 , 0,8468]	[0,6999 , 0,7906]
0,4	100	[0,7510 , 0,8662]	[0,7572 , 0,8621]	[0,7744 , 0,8484]	[0,7714 , 0,8440]	[0,6857 , 0,7722]

c_p	c_D	trimmed Liu	Huber ($\theta = 1/3$)	Huber ($\theta = 1$)	Hampel ($\theta = 1/3$)	Hampel ($\theta = 1$)
0	0	[0,7806 , 0,9246]	[0,7526 , 0,9141]	[0,7526 , 0,9141]	[0,8541 , 0,9494]	[0,8566 , 0,9473]
0,1	0	[0,7661 , 0,9129]	[0,7337 , 0,8977]	[0,7337 , 0,8977]	[0,8456 , 0,9464]	[0,8479 , 0,9440]
0,1	1	[0,7599 , 0,9037]	[0,7261 , 0,8854]	[0,7261 , 0,8854]	[0,8426 , 0,9443]	[0,8450 , 0,9420]
0,1	5	[0,7552 , 0,8936]	[0,7189 , 0,8726]	[0,7189 , 0,8726]	[0,8415 , 0,9429]	[0,8440 , 0,9407]
0,1	10	[0,7561 , 0,8905]	[0,7177 , 0,8677]	[0,7177 , 0,8677]	[0,8428 , 0,9427]	[0,8455 , 0,9405]
0,1	100	[0,7559 , 0,8929]	[0,7136 , 0,8633]	[0,7136 , 0,8633]	[0,8428 , 0,9425]	[0,8452 , 0,9402]
0,2	0	[0,7508 , 0,9012]	[0,7145 , 0,8815]	[0,7145 , 0,8815]	[0,8351 , 0,9428]	[0,8378 , 0,9404]
0,2	1	[0,7401 , 0,8816]	[0,7031 , 0,8594]	[0,7031 , 0,8594]	[0,8311 , 0,9383]	[0,8337 , 0,9360]
0,2	5	[0,7285 , 0,8579]	[0,6891 , 0,8331]	[0,6891 , 0,8331]	[0,8277 , 0,9353]	[0,8306 , 0,9332]
0,2	10	[0,7254 , 0,8502]	[0,6844 , 0,8246]	[0,6844 , 0,8246]	[0,8284 , 0,9345]	[0,8315 , 0,9326]
0,2	100	[0,7250 , 0,8495]	[0,6800 , 0,8176]	[0,6800 , 0,8176]	[0,8295 , 0,9333]	[0,8324 , 0,9312]
0,4	0	[0,7532 , 0,8927]	[0,6801 , 0,8506]	[0,6801 , 0,8506]	[0,8131 , 0,9341]	[0,8159 , 0,9313]
0,4	1	[0,7313 , 0,8551]	[0,6571 , 0,8081]	[0,6571 , 0,8081]	[0,7998 , 0,9196]	[0,8023 , 0,9171]
0,4	5	[0,7057 , 0,8038]	[0,6335 , 0,7597]	[0,6335 , 0,7597]	[0,8007 , 0,9110]	[0,8020 , 0,9081]
0,4	10	[0,6963 , 0,7866]	[0,6190 , 0,7403]	[0,6190 , 0,7403]	[0,8049 , 0,9127]	[0,8065 , 0,9101]
0,4	100	[0,6805 , 0,7676]	[0,6038 , 0,7188]	[0,6038 , 0,7188]	[0,8108 , 0,9166]	[0,8135 , 0,9151]

Table 4.10: Estimates of the location in Study 3 - CASE 4

c_p	c_D	mean	trimmed ($\theta = 1/3$)	trimmed ($\theta = 1$)
0	0	[0,7577 , 0,9092]	[0,8106 , 0,9317]	[0,8132 , 0,9287]
0,1	0	[0,7254 , 0,8977]	[0,7973 , 0,9268]	[0,7989 , 0,9236]
0,1	1	[0,7114 , 0,8913]	[0,7929 , 0,9253]	[0,7945 , 0,9222]
0,1	5	[0,6943 , 0,8831]	[0,7911 , 0,9244]	[0,7925 , 0,9216]
0,1	10	[0,6893 , 0,8816]	[0,7912 , 0,9245]	[0,7926 , 0,9216]
0,1	100	[0,6827 , 0,8785]	[0,7912 , 0,9244]	[0,7925 , 0,9216]
0,2	0	[0,6935 , 0,8864]	[0,7798 , 0,9203]	[0,7813 , 0,9171]
0,2	1	[0,6649 , 0,8728]	[0,7672 , 0,9159]	[0,7686 , 0,9131]
0,2	5	[0,6313 , 0,8584]	[0,7582 , 0,9126]	[0,7590 , 0,9108]
0,2	10	[0,6209 , 0,8544]	[0,7573 , 0,9113]	[0,7579 , 0,9100]
0,2	100	[0,6078 , 0,8471]	[0,7573 , 0,9095]	[0,7575 , 0,9091]
0,4	0	[0,6295 , 0,8634]	[0,8064 , 0,9299]	[0,8135 , 0,9278]
0,4	1	[0,5718 , 0,8360]	[0,7932 , 0,9257]	[0,7958 , 0,9221]
0,4	5	[0,5055 , 0,8107]	[0,7843 , 0,9222]	[0,7857 , 0,9194]
0,4	10	[0,4838 , 0,7996]	[0,7839 , 0,9216]	[0,7852 , 0,9191]
0,4	100	[0,4580 , 0,7857]	[0,7841 , 0,9216]	[0,7853 , 0,9191]

c_p	c_D	Hausdorff median	1-norm median	spatial ($\theta = 1/3$)	spatial ($\theta = 1$)
0	0	[0,7910 , 0,9081]	[0,7822 , 0,9330]	[0,7832 , 0,9212]	[0,7873 , 0,9207]
0,1	0	[0,7754 , 0,9021]	[0,7665 , 0,9279]	[0,7673 , 0,9156]	[0,7715 , 0,9152]
0,1	1	[0,7712 , 0,9007]	[0,7619 , 0,9264]	[0,7630 , 0,9142]	[0,7671 , 0,9137]
0,1	5	[0,7703 , 0,8997]	[0,7606 , 0,9260]	[0,7614 , 0,9138]	[0,7655 , 0,9135]
0,1	10	[0,7712 , 0,8990]	[0,7607 , 0,9262]	[0,7615 , 0,9140]	[0,7657 , 0,9137]
0,1	100	[0,7721 , 0,8980]	[0,7606 , 0,9261]	[0,7615 , 0,9140]	[0,7657 , 0,9137]
0,2	0	[0,7568 , 0,8949]	[0,7475 , 0,9219]	[0,7482 , 0,9090]	[0,7526 , 0,9087]
0,2	1	[0,7458 , 0,8906]	[0,7358 , 0,9183]	[0,7368 , 0,9051]	[0,7411 , 0,9047]
0,2	5	[0,7429 , 0,8881]	[0,7317 , 0,9182]	[0,7320 , 0,9043]	[0,7365 , 0,9042]
0,2	10	[0,7442 , 0,8868]	[0,7315 , 0,9180]	[0,7319 , 0,9045]	[0,7363 , 0,9044]
0,2	100	[0,7467 , 0,8842]	[0,7315 , 0,9177]	[0,7320 , 0,9042]	[0,7366 , 0,9041]
0,4	0	[0,7076 , 0,8741]	[0,6966 , 0,9070]	[0,6973 , 0,8912]	[0,7025 , 0,8911]
0,4	1	[0,6666 , 0,8548]	[0,6507 , 0,8963]	[0,6526 , 0,8764]	[0,6590 , 0,8761]
0,4	5	[0,6372 , 0,8392]	[0,6151 , 0,8953]	[0,6160 , 0,8700]	[0,6244 , 0,8701]
0,4	10	[0,6390 , 0,8347]	[0,6138 , 0,8943]	[0,6140 , 0,8686]	[0,6224 , 0,8692]
0,4	100	[0,6439 , 0,8298]	[0,6138 , 0,8942]	[0,6142 , 0,8678]	[0,6229 , 0,8687]

c_p	c_D	Huber ($\theta = 1/3$)	Huber ($\theta = 1$)	Hampel ($\theta = 1/3$)	Hampel ($\theta = 1$)
0	0	[0,7577 , 0,9092]	[0,7577 , 0,9092]	[0,8159 , 0,9337]	[0,8212 , 0,9315]
0,1	0	[0,7254 , 0,8977]	[0,7254 , 0,8977]	[0,8040 , 0,9293]	[0,8084 , 0,9269]
0,1	1	[0,7114 , 0,8913]	[0,7114 , 0,8913]	[0,8003 , 0,9280]	[0,8046 , 0,9257]
0,1	5	[0,6943 , 0,8831]	[0,6943 , 0,8831]	[0,7985 , 0,9273]	[0,8024 , 0,9250]
0,1	10	[0,6893 , 0,8816]	[0,6893 , 0,8816]	[0,7984 , 0,9272]	[0,8023 , 0,9250]
0,1	100	[0,6827 , 0,8785]	[0,6827 , 0,8785]	[0,7983 , 0,9272]	[0,8023 , 0,9250]
0,2	0	[0,6935 , 0,8864]	[0,6935 , 0,8864]	[0,7885 , 0,9236]	[0,7927 , 0,9211]
0,2	1	[0,6649 , 0,8728]	[0,6649 , 0,8728]	[0,7788 , 0,9202]	[0,7824 , 0,9179]
0,2	5	[0,6313 , 0,8584]	[0,6313 , 0,8584]	[0,7759 , 0,9191]	[0,7782 , 0,9171]
0,2	10	[0,6209 , 0,8544]	[0,6209 , 0,8544]	[0,7769 , 0,9191]	[0,7788 , 0,9171]
0,2	100	[0,6078 , 0,8471]	[0,6078 , 0,8471]	[0,7776 , 0,9188]	[0,7794 , 0,9170]
0,4	0	[0,6295 , 0,8634]	[0,6295 , 0,8634]	[0,7495 , 0,9095]	[0,7537 , 0,9066]
0,4	1	[0,5718 , 0,8360]	[0,5718 , 0,8360]	[0,7198 , 0,9021]	[0,7245 , 0,8985]
0,4	5	[0,5055 , 0,8107]	[0,5055 , 0,8107]	[0,7004 , 0,9054]	[0,7110 , 0,9023]
0,4	10	[0,4838 , 0,7996]	[0,4838 , 0,7996]	[0,7097 , 0,9074]	[0,7248 , 0,9056]
0,4	100	[0,4580 , 0,7857]	[0,4580 , 0,7857]	[0,7286 , 0,9093]	[0,7395 , 0,9087]

Table 4.11: Estimates of the location in Study 4 - CASE 3

c_p	c_D	mean	trimmed ($\theta = 1/3$)	trimmed ($\theta = 1$)
0	0	[0,7528 , 0,9141]	[0,8375 , 0,9437]	[0,8388 , 0,9411]
0,1	0	[0,7345 , 0,8981]	[0,8277 , 0,9404]	[0,8290 , 0,9376]
0,1	1	[0,7284 , 0,8875]	[0,8251 , 0,9380]	[0,8265 , 0,9354]
0,1	5	[0,7223 , 0,8748]	[0,8253 , 0,9368]	[0,8266 , 0,9341]
0,1	10	[0,7210 , 0,8721]	[0,8262 , 0,9365]	[0,8275 , 0,9339]
0,1	100	[0,7182 , 0,8675]	[0,8265 , 0,9360]	[0,8279 , 0,9333]
0,2	0	[0,7165 , 0,8825]	[0,8164 , 0,9364]	[0,8179 , 0,9334]
0,2	1	[0,7033 , 0,8598]	[0,8093 , 0,9303]	[0,8108 , 0,9275]
0,2	5	[0,6905 , 0,8345]	[0,8072 , 0,9273]	[0,8086 , 0,9247]
0,2	10	[0,6883 , 0,8289]	[0,8088 , 0,9269]	[0,8101 , 0,9244]
0,2	100	[0,6853 , 0,8223]	[0,8113 , 0,9261]	[0,8126 , 0,9236]
0,4	0	[0,6813 , 0,8516]	[0,8769 , 0,9575]	[0,8781 , 0,9545]
0,4	1	[0,6587 , 0,8098]	[0,8703 , 0,9500]	[0,8715 , 0,9474]
0,4	5	[0,6289 , 0,7552]	[0,8691 , 0,9452]	[0,8704 , 0,9426]
0,4	10	[0,6214 , 0,7417]	[0,8706 , 0,9439]	[0,8720 , 0,9412]
0,4	100	[0,6111 , 0,7248]	[0,8724 , 0,9426]	[0,8739 , 0,9397]

c_p	c_D	Hausdorff median	1-norm median	spatial ($\theta = 1/3$)	spatial ($\theta = 1$)
0	0	[0,8214 , 0,9200]	[0,8112 , 0,9501]	[0,8123 , 0,9355]	[0,8174 , 0,9355]
0,1	0	[0,8134 , 0,9137]	[0,8005 , 0,9465]	[0,8028 , 0,9307]	[0,8081 , 0,9304]
0,1	1	[0,8128 , 0,9095]	[0,7982 , 0,9438]	[0,8012 , 0,9276]	[0,8066 , 0,9271]
0,1	5	[0,8164 , 0,9053]	[0,7994 , 0,9424]	[0,8028 , 0,9256]	[0,8083 , 0,9250]
0,1	10	[0,8180 , 0,9044]	[0,8004 , 0,9424]	[0,8037 , 0,9254]	[0,8094 , 0,9249]
0,1	100	[0,8188 , 0,9035]	[0,8011 , 0,9421]	[0,8042 , 0,9251]	[0,8100 , 0,9247]
0,2	0	[0,8044 , 0,9065]	[0,7885 , 0,9422]	[0,7921 , 0,9253]	[0,7977 , 0,9246]
0,2	1	[0,8020 , 0,8966]	[0,7822 , 0,9355]	[0,7876 , 0,9175]	[0,7932 , 0,9165]
0,2	5	[0,8085 , 0,8883]	[0,7837 , 0,9323]	[0,7902 , 0,9130]	[0,7961 , 0,9116]
0,2	10	[0,8121 , 0,8861]	[0,7861 , 0,9321]	[0,7924 , 0,9125]	[0,7985 , 0,9113]
0,2	100	[0,8148 , 0,8847]	[0,7889 , 0,9318]	[0,7946 , 0,9121]	[0,8008 , 0,9111]
0,4	0	[0,7832 , 0,8885]	[0,7600 , 0,9312]	[0,7670 , 0,9112]	[0,7730 , 0,9098]
0,4	1	[0,7735 , 0,8639]	[0,7422 , 0,9120]	[0,7534 , 0,8899]	[0,7593 , 0,8877]
0,4	5	[0,7779 , 0,8411]	[0,7336 , 0,8988]	[0,7510 , 0,8726]	[0,7566 , 0,8688]
0,4	10	[0,7835 , 0,8363]	[0,7369 , 0,8970]	[0,7543 , 0,8703]	[0,7601 , 0,8663]
0,4	100	[0,7902 , 0,8301]	[0,7410 , 0,8952]	[0,7577 , 0,8681]	[0,7638 , 0,8643]

c_p	c_D	Huber ($\theta = 1/3$)	Huber ($\theta = 1$)	Hampel ($\theta = 1/3$)	Hampel ($\theta = 1$)
0	0	[0,7528 , 0,9141]	[0,7528 , 0,9141]	[0,8557 , 0,9499]	[0,8580 , 0,9477]
0,1	0	[0,7345 , 0,8981]	[0,7345 , 0,8981]	[0,8470 , 0,9469]	[0,8493 , 0,9446]
0,1	1	[0,7284 , 0,8875]	[0,7284 , 0,8875]	[0,8449 , 0,9449]	[0,8472 , 0,9427]
0,1	5	[0,7223 , 0,8748]	[0,7223 , 0,8748]	[0,8448 , 0,9436]	[0,8471 , 0,9414]
0,1	10	[0,7210 , 0,8721]	[0,7210 , 0,8721]	[0,8455 , 0,9434]	[0,8479 , 0,9411]
0,1	100	[0,7182 , 0,8675]	[0,7182 , 0,8675]	[0,8457 , 0,9429]	[0,8481 , 0,9406]
0,2	0	[0,7165 , 0,8825]	[0,7165 , 0,8825]	[0,8371 , 0,9435]	[0,8395 , 0,9410]
0,2	1	[0,7033 , 0,8598]	[0,7033 , 0,8598]	[0,8313 , 0,9384]	[0,8338 , 0,9361]
0,2	5	[0,6905 , 0,8345]	[0,6905 , 0,8345]	[0,8298 , 0,9354]	[0,8325 , 0,9333]
0,2	10	[0,6883 , 0,8289]	[0,6883 , 0,8289]	[0,8311 , 0,9349]	[0,8338 , 0,9328]
0,2	100	[0,6853 , 0,8223]	[0,6853 , 0,8223]	[0,8326 , 0,9340]	[0,8353 , 0,9319]
0,4	0	[0,6813 , 0,8516]	[0,6813 , 0,8516]	[0,8147 , 0,9348]	[0,8174 , 0,9319]
0,4	1	[0,6587 , 0,8098]	[0,6587 , 0,8098]	[0,8019 , 0,9203]	[0,8045 , 0,9178]
0,4	5	[0,6289 , 0,7552]	[0,6289 , 0,7552]	[0,7998 , 0,9101]	[0,8010 , 0,9072]
0,4	10	[0,6214 , 0,7417]	[0,6214 , 0,7417]	[0,8061 , 0,9125]	[0,8077 , 0,9099]
0,4	100	[0,6111 , 0,7248]	[0,6111 , 0,7248]	[0,8156 , 0,9172]	[0,8182 , 0,9156]

Table 4.12: Estimates of the location in Study 4 - CASE 4

STUDY 3		CASE 3	CASE 4
Bias	$c_p \leq 0.2$	**Hampel** $(\theta = 1/3)$	**trimmed Liu**
	$c_p = 0.4$	**trimmed** $(\theta = 1/3)$	**trimmed** $(\theta = 1/3)$
	$c_D = 0$		
	Dispersion	none	high
Variance	$c_p = 0$	**mean** **Huber**	**Hampel** $(\theta = 1)$
	$c_p \leq 0.2$	**trimmed** $(\theta = 1)$	**Hampel** $(\theta = 1)$
	$c_p = 0.4$	**trimmed** $(\theta = 1)$	**trimmed** $(\theta = 1)$
	$c_D = 0$		
	Dispersion	high	low
MSE	$c_p = 0$	**mean** **Huber**	**Hampel** $(\theta = 1)$
	$c_p \leq 0.2$	**trimmed** **Hampel** $(\theta = 1/3)$	**Hampel** $(\theta = 1)$
	$c_p = 0.4$	**trimmed** $(\theta = 1/3)$	**trimmed** $(\theta = 1/3)$
	$c_D = 0$	**1-norm median**	
	Dispersion	medium	none

Table 4.13: Summary of the main conclusions from Study 3: the most suitable (if any) location measures/estimates

STUDY 4		CASE 3	CASE 4
Bias	$c_p \leq 0.2$	**Hampel** $(\theta = 1/3)$	**Hampel**
	$c_p = 0.4$	**trimmed** $(\theta = 1/3)$	**trimmed**
	$c_D = 0$		
	Dispersion	none	medium
Variance	$c_p = 0$	**trimmed** $(\theta = 1)$	**Hampel** $(\theta = 1)$
	$c_p \leq 0.2$	**trimmed** **Hampel**	**Hampel** $(\theta = 1)$ **Hausdorff**
	$c_p = 0.4$	**trimmed** $(\theta = 1/3)$	**trimmed**
	$c_D = 0$		
	Dispersion	high	high
MSE	$c_p = 0$	**trimmed** $(\theta = 1)$	**Hampel** $(\theta = 1)$
	$c_p \leq 0.2$	**Hampel** $(\theta = 1/3)$	**Hampel**
	$c_p = 0.4$	**trimmed** $(\theta = 1/3)$	**trimmed** $(\theta = 1/3)$
	$c_D = 0$		
	Dispersion	none	medium

Table 4.14: Summary of the main conclusions from Study 4: the most suitable (if any) location measures/estimates

On the basis of the conclusions gathered in Tables 4.7, 4.8, 4.13 and 4.14, one can conclude that there is no uniformly most appropriate location estimate. Actually, the outputs seem to depend much more on the distribution considered for the non-contaminated and contaminated distributions, or the involved case, than on the sample size. A rather general assertion is that M-estimates behave better for lower and moderate levels of contamination, whereas trimmed means are more convenient for very high contamination level.

4.2 Comparative simulations for fuzzy number-valued data

The simulations in this section also mimic the studies presented in the previous chapters, adapting the four different situations analyzed in Section 4.1 to the fuzzy valued-case. Therefore, two choices of the sample size ($n = 100, n = 10000$) and different non-contaminated (symmetric and asymmetric) and contaminated distributions have been considered. Only trapezoidal fuzzy numbers have been considered in order to ease the computation.

Dealing with fuzzy number-valued data, the comparisons have concerned the following location measures/estimates: trimmed means, Huber and Hampel M-estimates (using the D_θ^ℓ metric), 1-norm median and wabl/ldev/rdev-median, where in all of them θ is assumed to range in $\{1/3, 1\}$.

For each of the measures/estimates, the Monte Carlo approximation of the estimate, the bias, the variance and the mean squared error of the estimate have been determined.

The general scheme of the four studies is the same detailed in Section 4.1, just replacing the generation of interval-valued data in the first point of *Step 1* by the procedure that will be explained now.

Step 1. A sample of n trapezoidal fuzzy number-valued data has been simulated from a random fuzzy number $\mathcal{X}$ for each of some different situations in such a way that

- to generate the trapezoidal fuzzy data, we have considered four real-valued random variables as follows: $\mathcal{X} = \mathrm{Tra}(X_1 - X_2 - X_3, X_1 - X_2, X_1 + X_2, X_1 + X_2 + X_4)$, with $X_1 = \mathrm{mid}\,\mathcal{X}_1$, $X_2 = \mathrm{spr}\,\mathcal{X}_1$, $X_3 = \inf \mathcal{X}_1 - \inf \mathcal{X}_0$ and $X_4 = \sup \mathcal{X}_0 - \sup \mathcal{X}_1$ or, alternatively, four order real-valued statistics $X_{(1)}$, $X_{(2)}$, $X_{(3)}$ and $X_{(4)}$ such that $\mathcal{X} = [X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}]$, i.e., $X_{(1)} = \inf \mathcal{X}_0$, $X_{(2)} = \inf \mathcal{X}_1$, $X_{(3)} = \sup \mathcal{X}_1$ and $X_{(4)} = \sup \mathcal{X}_0$;

The rest of *Step 1* and both *Step 2* and *Step 3* coincide with those presented in the previous section.

The choices of the non contaminated and contaminated distributions in each study will be specified now. Notice that they are the natural adaption of the ones considered for the interval-valued case, so differences between the behavior of the estimates dealing with these two kinds of imprecise data become apparent.

Study 1

In this first study, the choices correspond to:

- $n = 100$;
- CASE 1 assumes that
  - •• $X_1 \sim \mathcal{N}(0,1)$ and $X_2, X_3, X_4 \sim \chi_1^2$ for the non-contaminated subsample,
  - •• $X_1 \sim \mathcal{N}(0,3) + C_D$ and $X_2, X_3, X_4 \sim \chi_4^2 + C_D$ for the contaminated subsample,

  whereas CASE 2 assumes that
  - •• $X_1 \sim \mathcal{N}(0,1)$ and $X_2, X_3, X_4 \sim 1/(X_1^2+1)^2 + 0.1 \cdot \chi_1^2$ for the non-contaminated subsample,
  - •• $X_1 \sim \mathcal{N}(0,3) + C_D$ and $X_2, X_3, X_4 \sim 1/(X_1^2+1)^2 + 0.1 \cdot \chi_1^2 + C_D$ for the contaminated subsample.

  and CASE 2' assumes that
  - •• $X_1 \sim \mathcal{N}(0,1)$ and $X_2, X_3, X_4 \sim 1/(X_1^2+1)^2 + \sqrt{\chi_1^2}$ for the non-contaminated subsample,
  - •• $X_1 \sim \mathcal{N}(0,3) + C_D$ and $X_2, X_3, X_4 \sim 1/(X_1^2+1)^2 + \sqrt{\chi_1^2} + C_D$ for the contaminated subsample.

The simulations in this study have been presented through Figures 4.1-4.6 and Table 4.15.

Study 2

In this second study, the choices correspond to:

- $n = 10000$;
- In CASES 1, 2 and 2' the distributions for X_1, X_2, X_3 and X_4 in the no-contaminated and the contaminated samples coincide with those for Study 1.

The simulations in this study have been presented through Figures 4.7-4.12 and Table 4.16.

Concerning **Study 1** and **Study 2**, the estimates are given by

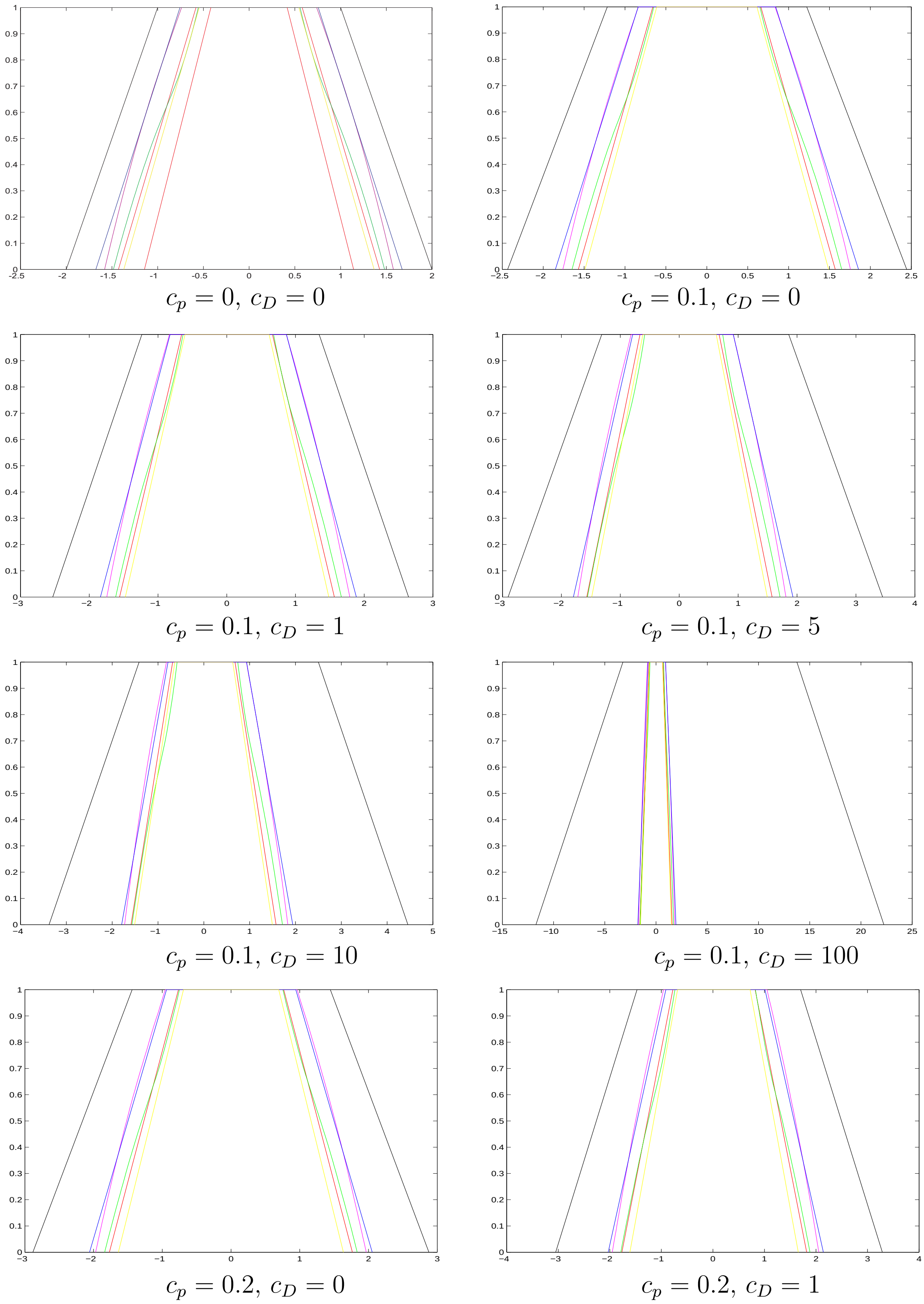
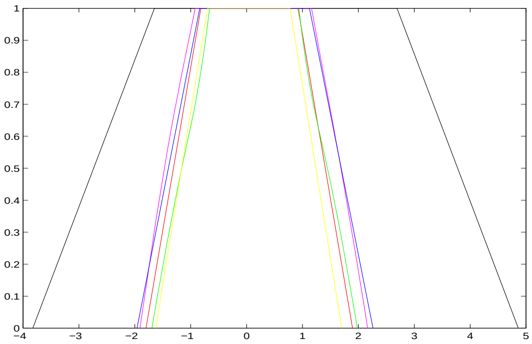
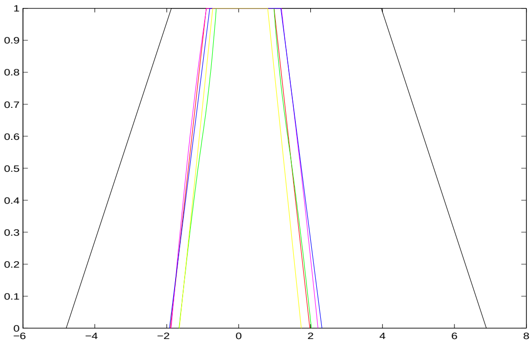


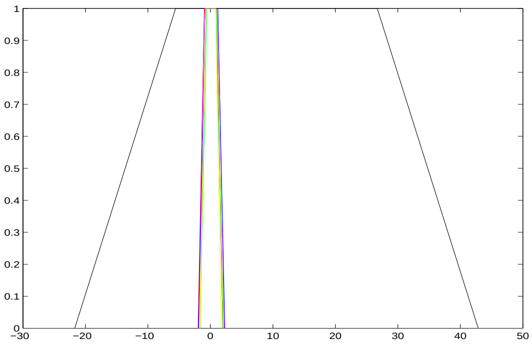
Figure 4.1: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 1 - CASE 1



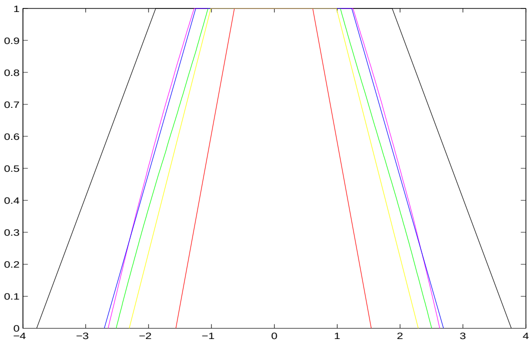
$c_p = 0.2,\ c_D = 5$



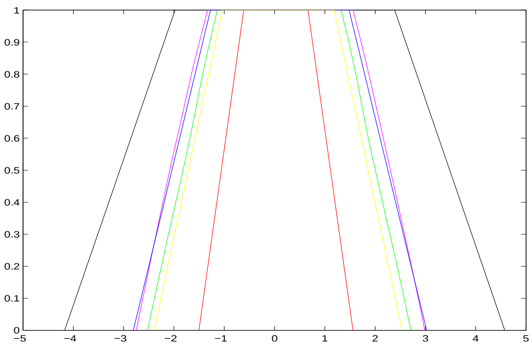
$c_p = 0.2,\ c_D = 10$



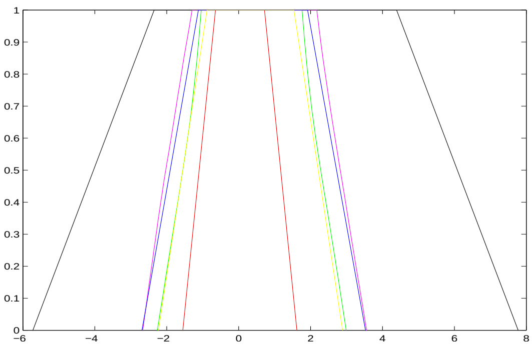
$c_p = 0.2,\ c_D = 100$



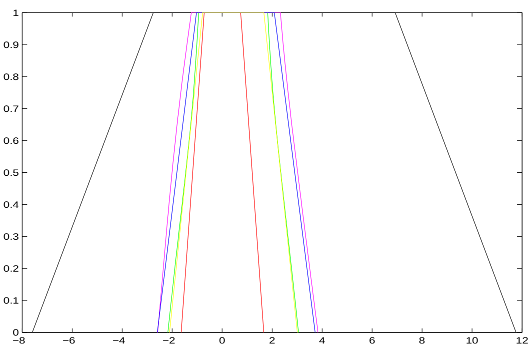
$c_p = 0.4,\ c_D = 0$



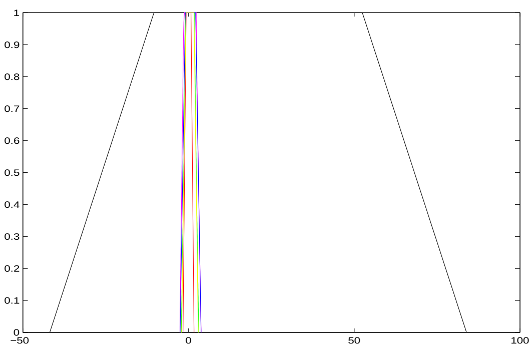
$c_p = 0.4,\ c_D = 1$



$c_p = 0.4,\ c_D = 5$



$c_p = 0.4,\ c_D = 10$



$c_p = 0.4,\ c_D = 100$

Figure 4.2: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 1 - CASE 1

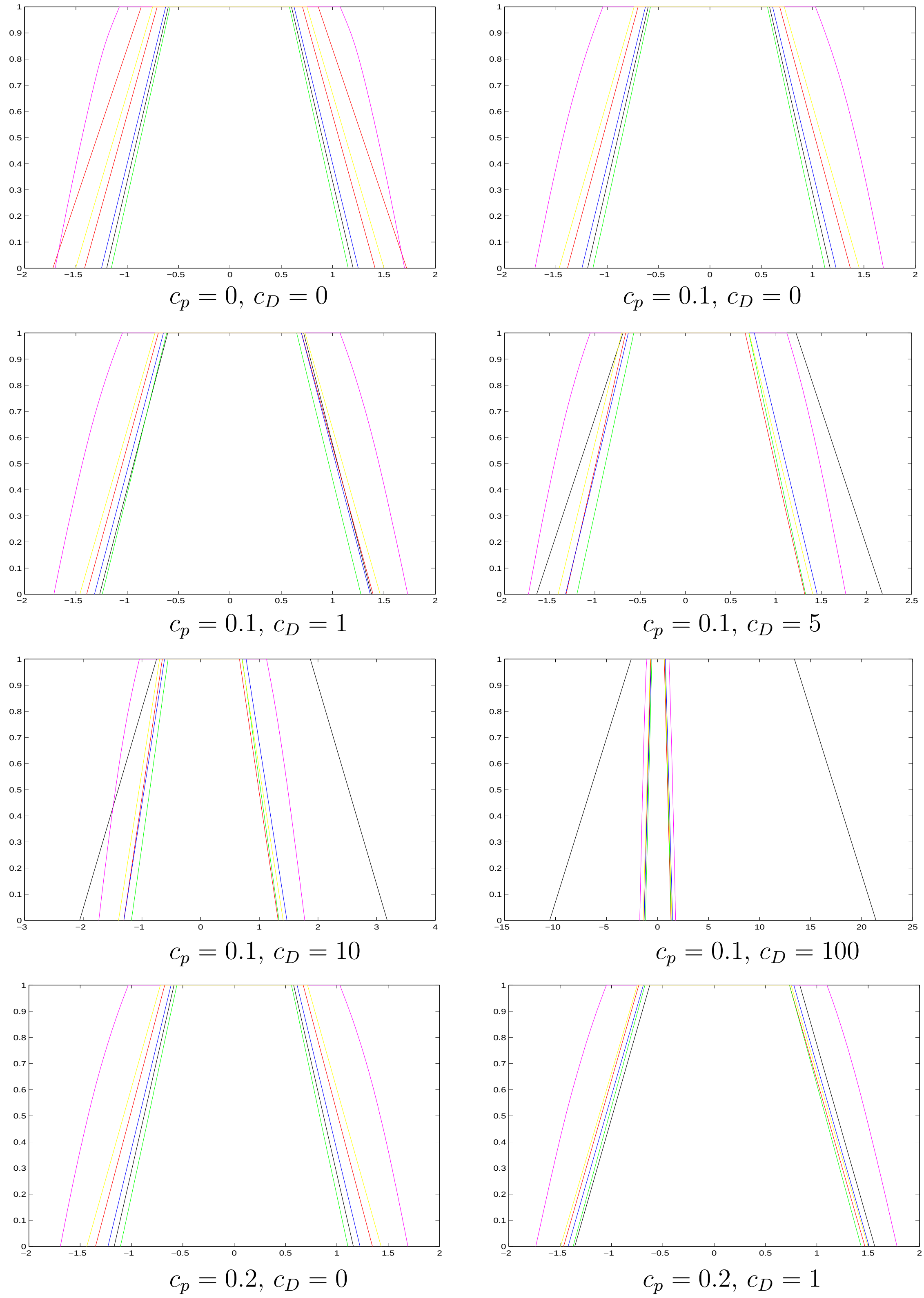


Figure 4.3: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 1 - CASE 2

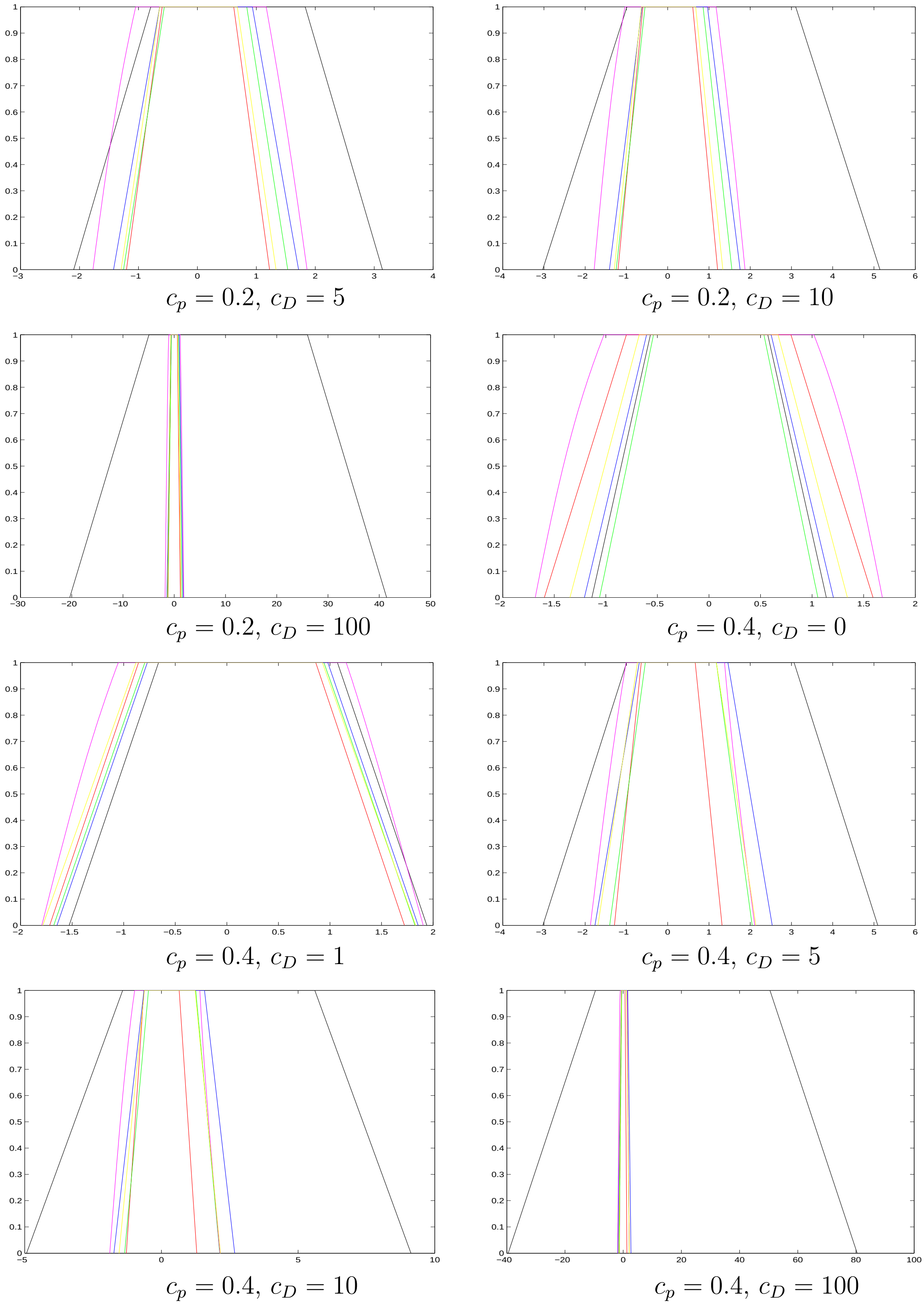


Figure 4.4: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 1 - CASE 2

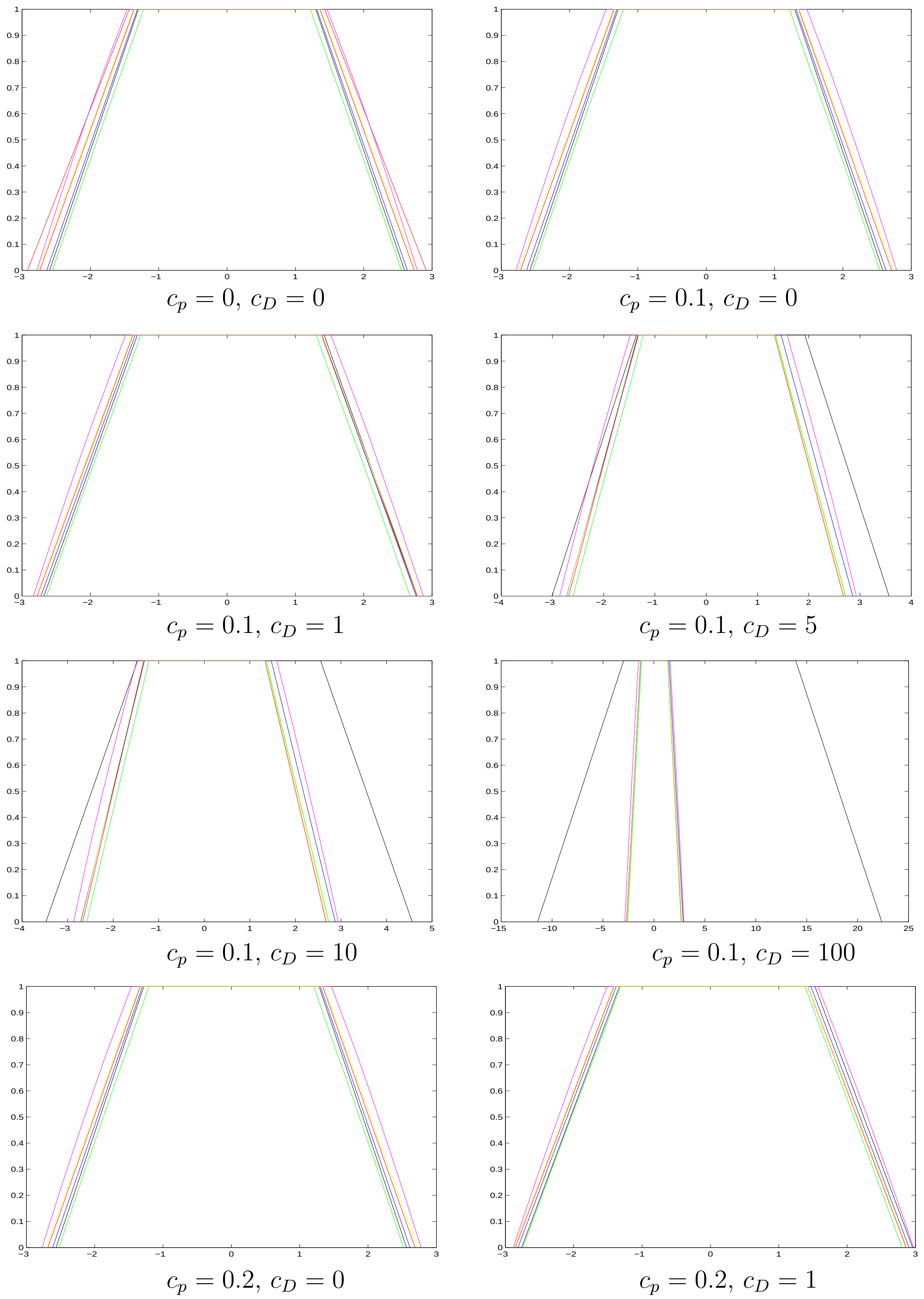


Figure 4.5: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 1 - CASE 2'

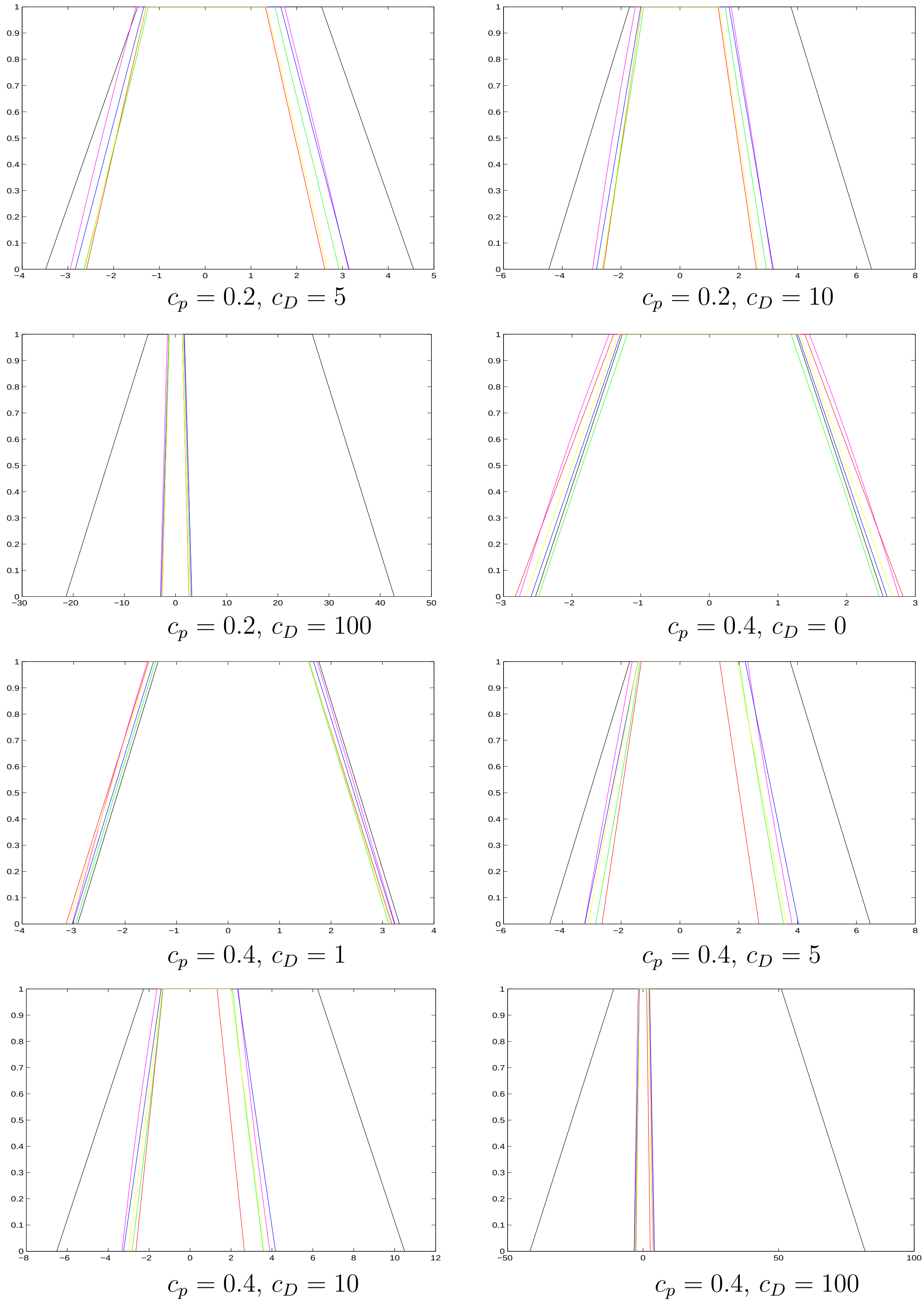


Figure 4.6: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 1 - CASE 2'

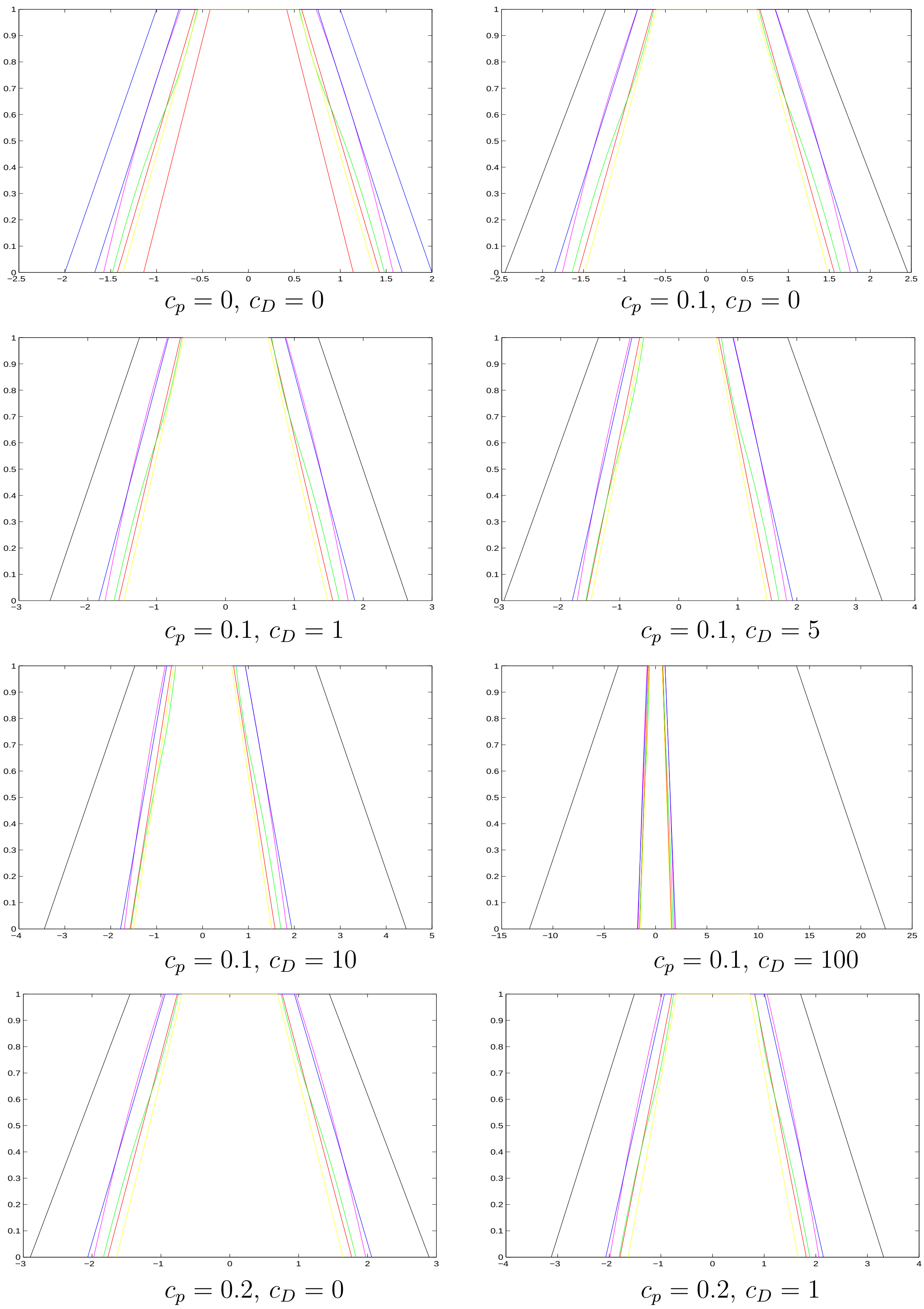


Figure 4.7: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 2 - CASE 1

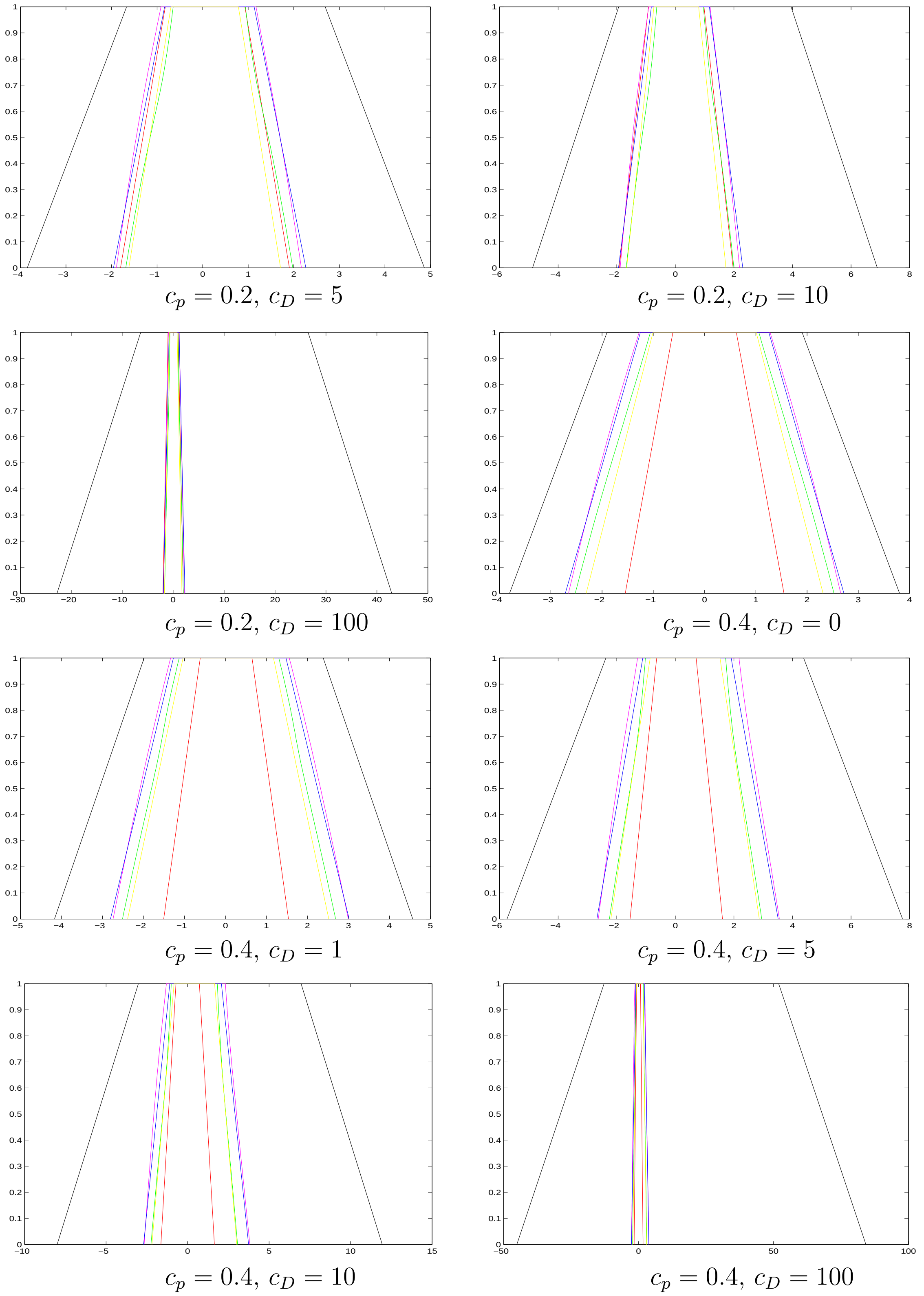


Figure 4.8: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 2 - CASE 1

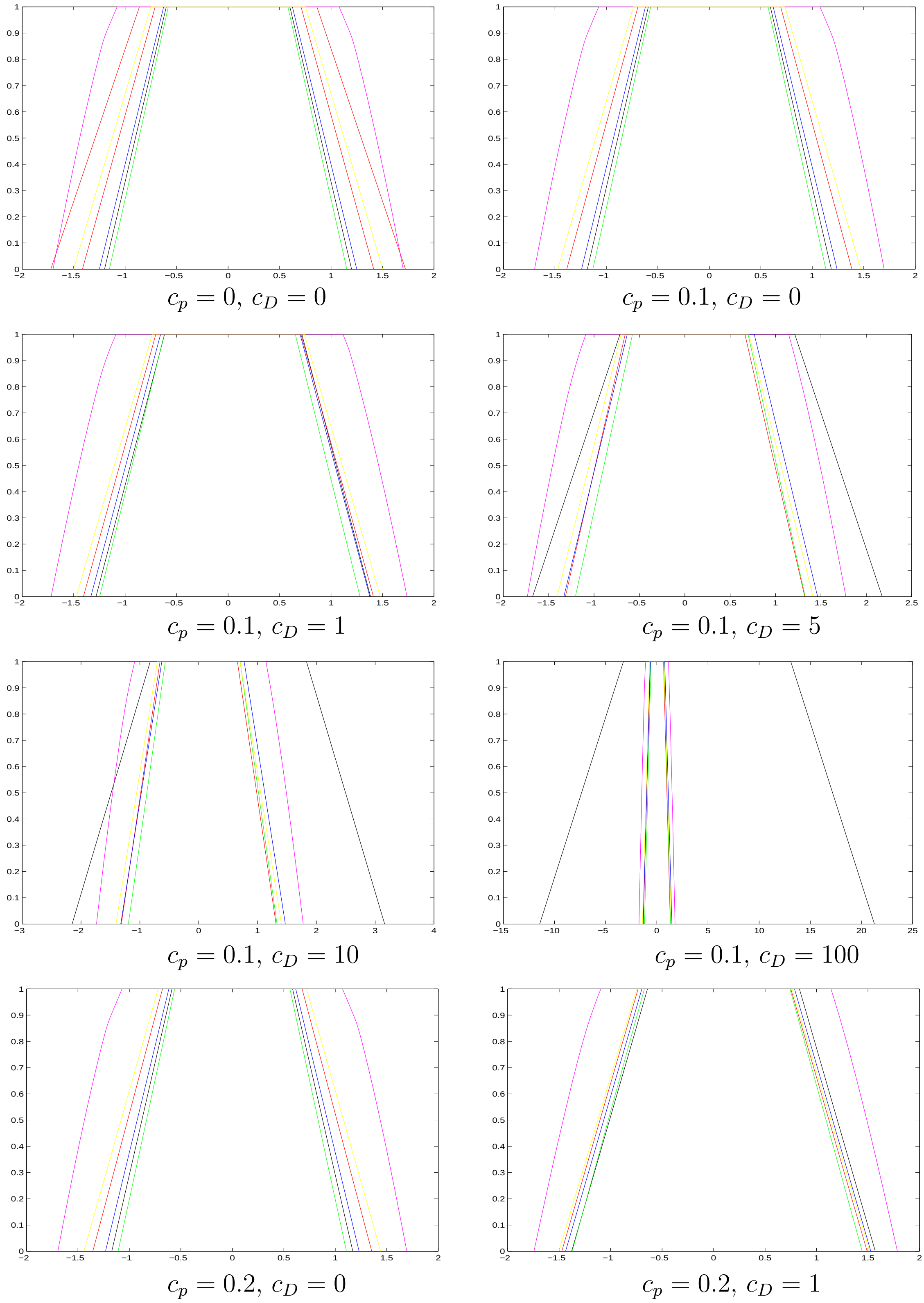


Figure 4.9: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 2 - CASE 2

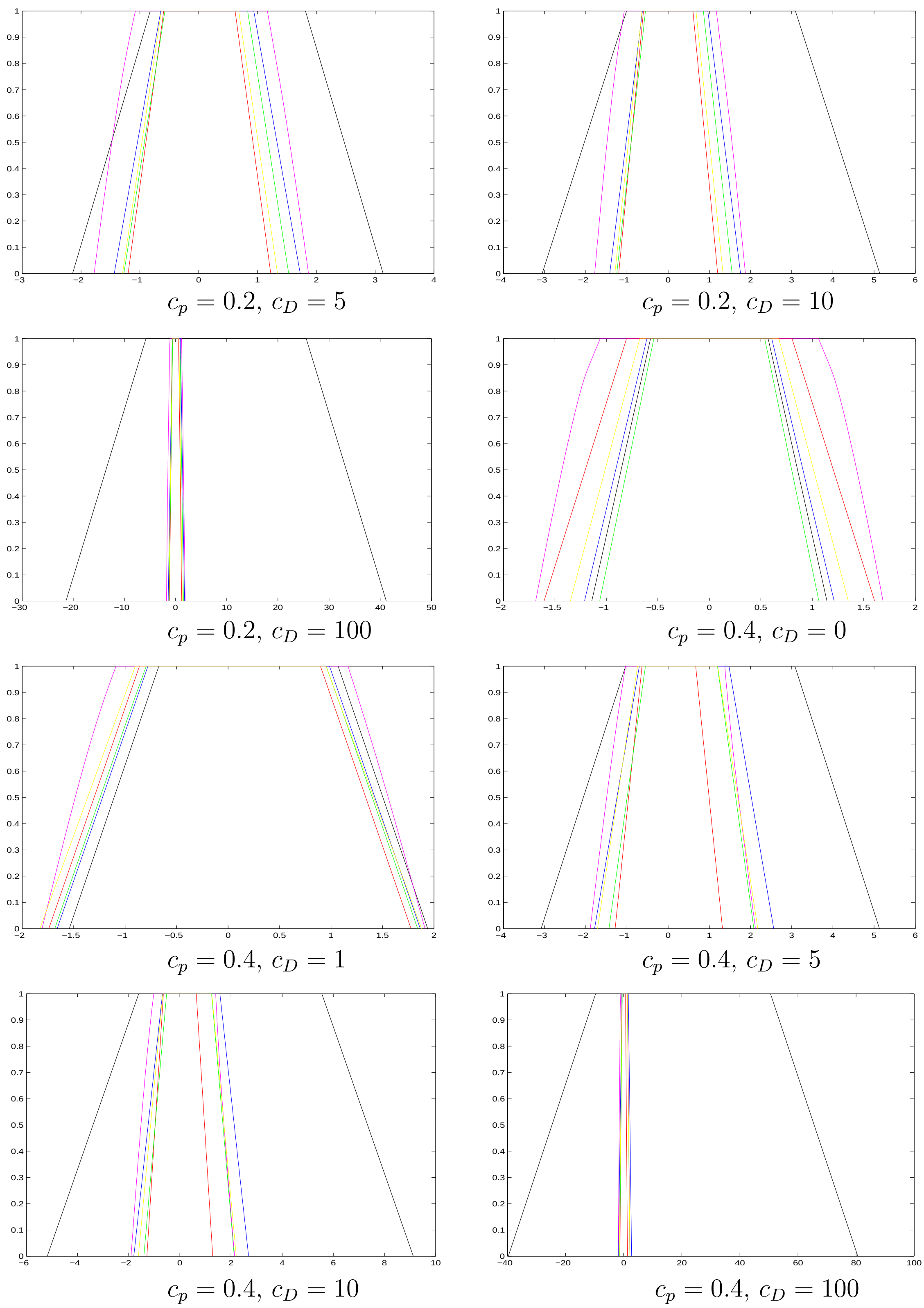


Figure 4.10: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 2 - CASE 2

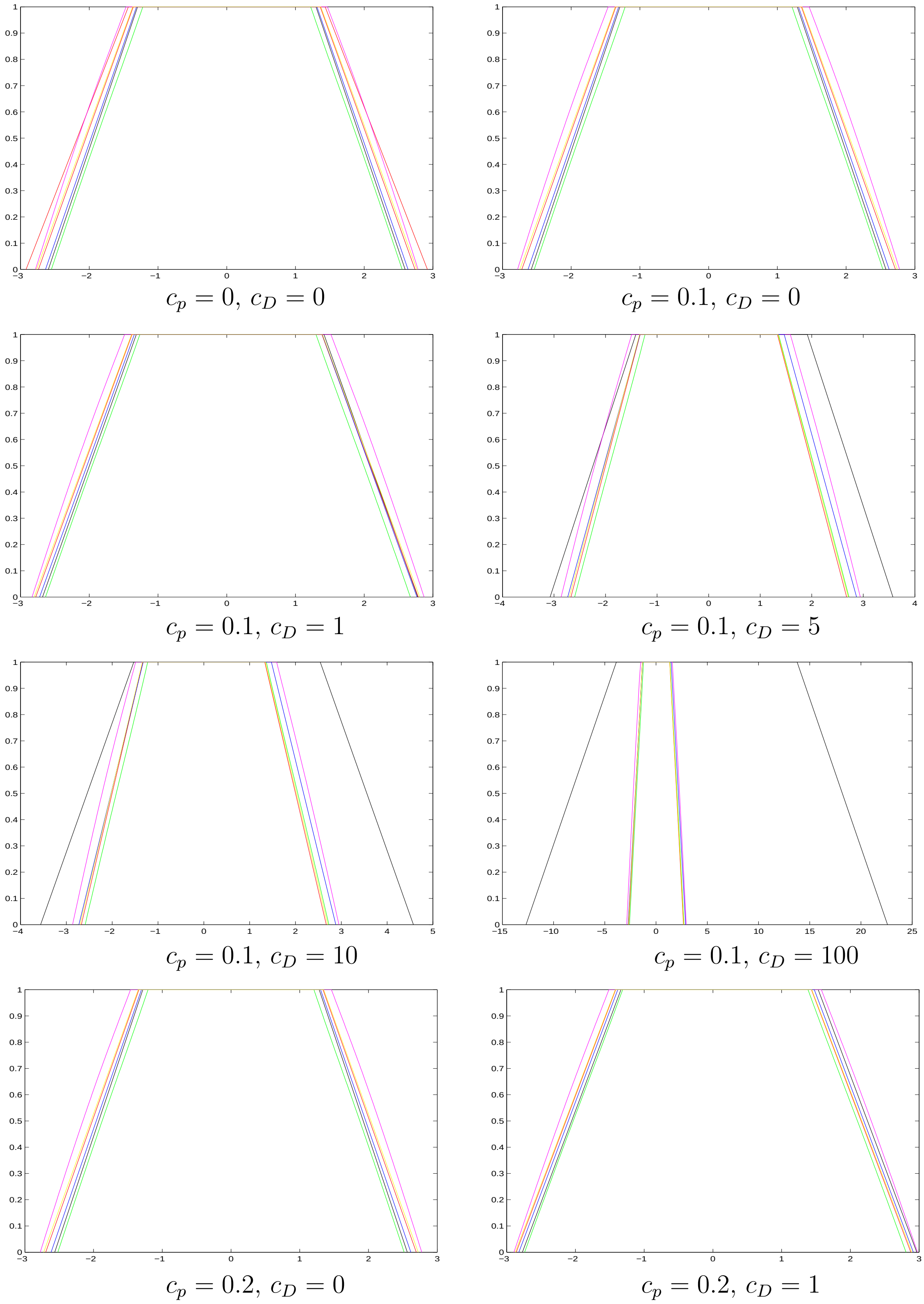
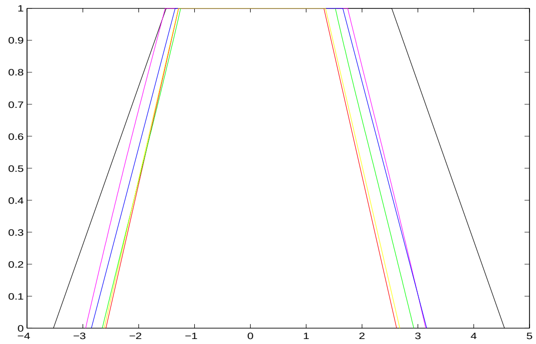
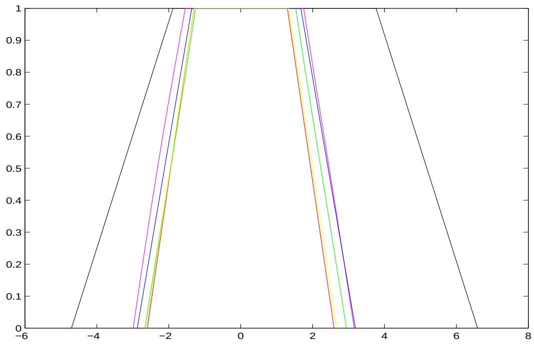


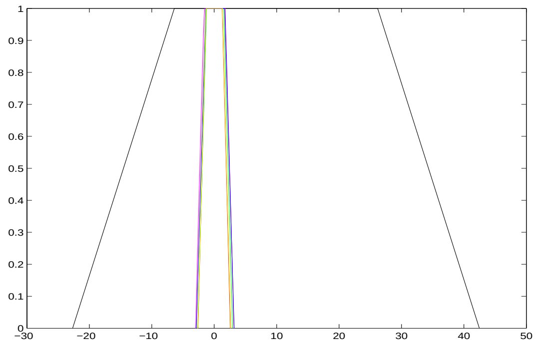
Figure 4.11: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 2 - CASE 2'



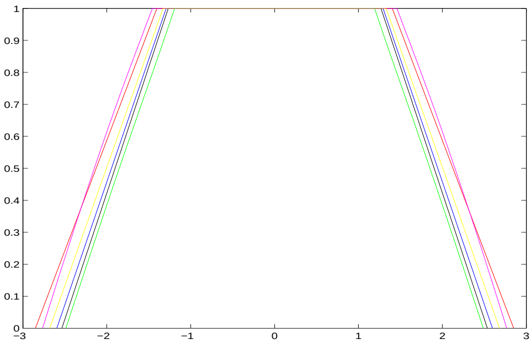
$c_p = 0.2,\ c_D = 5$



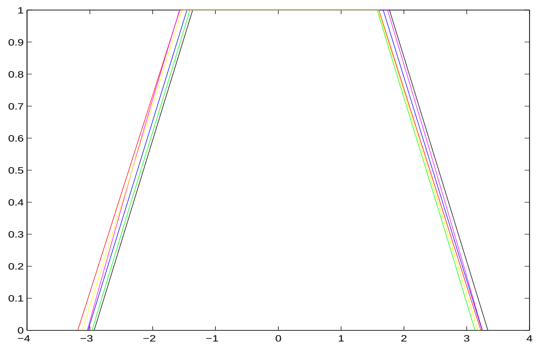
$c_p = 0.2,\ c_D = 10$



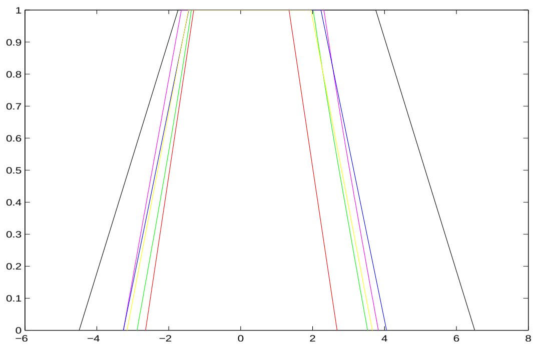
$c_p = 0.2,\ c_D = 100$



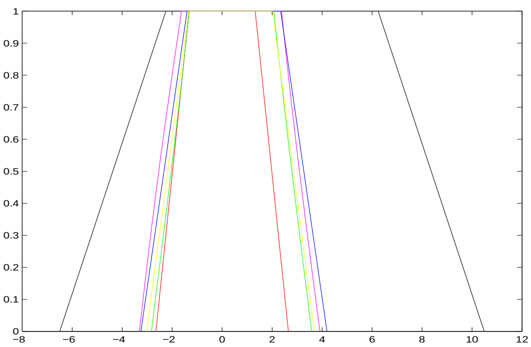
$c_p = 0.4,\ c_D = 0$



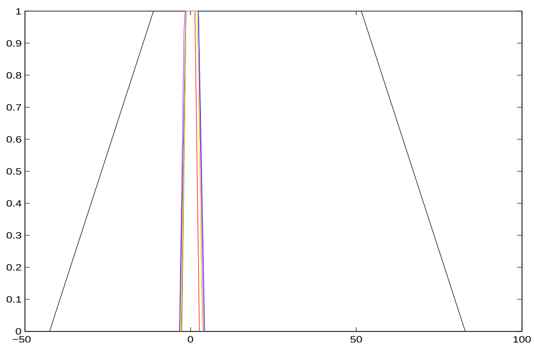
$c_p = 0.4,\ c_D = 1$



$c_p = 0.4,\ c_D = 5$



$c_p = 0.4,\ c_D = 10$



$c_p = 0.4,\ c_D = 100$

Figure 4.12: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 2 - CASE 2’

For the bias, variance and mean square error, the conclusions for the three cases in Studies 1 and 2 are summarized in Tables 4.15 and 4.16. For more details about the outputs, visit the link http://bellman.ciencias.uniovi.es/SMIRE/Fuzsimul.html.

STUDY 1		**CASE 1**	**CASE 2**	**CASE 2'**
Bias	$c_p \leq 0.2$	**Hampel**	**1-norm median**	**Hampel** $(\theta = 1)$
	$c_p = 0.4$	**trimmed**	**1-norm median**	**trimmed** $(\theta = 1)$
	$c_D = 0$		**Huber** $(\theta = 1/3)$	**1-norm median**
	Dispersion	none	low	low
Variance	$c_p = 0$	**1-norm median**	**1-norm median**	**1-norm median** **mean**
	$c_p \leq 0.2$	**1-norm median** **Hampel** $(\theta = 1)$	**1-norm median**	**1-norm median**
	$c_p = 0.4$	**wabl median** **trimmed** $(\theta = 1)$	**1-norm median**	**trimmed** $(\theta = 1)$
	$c_D = 0$			
	Dispersion	low	low	high
MSE	$c_p = 0$	**Huber** $(\theta = 1)$ **1-norm median**	**1-norm median**	**1-norm median** **mean**
	$c_p \leq 0.2$	**Hampel**	**1-norm median**	**1-norm median** **Hampel** $(\theta = 1)$
	$c_p = 0.4$	**trimmed**	**1-norm median**	**trimmed** $(\theta = 1)$
	$c_D = 0$			
	Dispersion	low	low	high

Table 4.15: Summary of the main conclusions from Study 1: the most suitable (if any) location measures/estimates

STUDY 2		**CASE 1**	**CASE 2**	**CASE 2'**
Bias	$c_p \leq 0.2$	**Hampel**	**1-norm median**	**Hampel** $(\theta = 1)$
	$c_p = 0.4$	**trimmed**	**1-norm median**	**trimmed** $(\theta = 1)$
	$c_D = 0$			**1-norm median**
	Dispersion	none	medium	low
Variance	$c_p = 0$	**Huber** $(\theta = 1)$	**1-norm median**	**1-norm median**
	$c_p \leq 0.2$	**Hampel**	**1-norm median** **Hampel** $(\theta = 1)$ **trimmed** $(\theta = 1/3)$	**1-norm median** **trimmed** **Hampel** $(\theta = 1)$
	$c_p = 0.4$	**trimmed**	**trimmed**	**trimmed**
	$c_D = 0/1$		**1-norm median**	**1-norm median**
	Dispersion	medium	low	medium
MSE	$c_p = 0$	**Huber** $(\theta = 1)$	**1-norm median**	**1-norm median**
	$c_p \leq 0.2$	**Hampel**	**1-norm median**	**Hampel** $(\theta = 1)$
	$c_p = 0.4$	**trimmed**	**1-norm median**	**trimmed** $(\theta = 1)$
	$c_D = 0/1$		**1-norm median**	**1-norm median**
	Dispersion	none	low	low

Table 4.16: Summary of the main conclusions from Study 2: the most suitable (if any) location measures/estimates

Study 3

In this third study, the choices correspond to:

- $n = 100$;
- CASE 3 assumes that
  - •• $X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)} \sim \text{Beta}(5, 1)$ (they are simply chosen at random and ordered) for the non-contaminated subsample,
  - •• $X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)} \sim \text{Beta}(1, C_D + 1)$ for the contaminated subsample,

  whereas CASE 4 assumes that
  - •• $X_1 \sim \text{Beta}(5, 1)$, $X_2 \sim \text{Uniform}[0, \min\{X_1, 1-X_1\}]$, $X_3 \sim \text{Uniform}[0, X_1 - X_2]$ and $X_4 \sim \text{Uniform}[0, 1 - X_1 - X_2]$ for the non-contaminated subsample,
  - •• $X_1 \sim \text{Beta}(1, C_D + 1)$, $X_2 \sim \min\{X_1, 1 - X_1\} \cdot \text{Beta}(1, C_D + 1)$, $X_3 \sim (X_1 - X_2) \cdot \text{Beta}(1, C_D + 1)$ and $X_4 \sim (1 - X_1 - X_2) \cdot \text{Beta}(1, C_D + 1)$ for the contaminated subsample.

The simulations in this study have been presented through Figures 4.13-4.16 and Table 4.17.

Study 4

In this fourth study, the choices correspond to:

- $n = 10000$;
- In CASES 3 and 4 the distributions for $X_{(1)}$, $X_{(2)}$, $X_{(3)}$, $X_{(4)}$, X_1, X_2, X_3 and X_4 in the non-contaminated and contaminated samples coincide with those for Study 3.

The simulations in this study have been presented through Figures 4.17-4.20 and Table 4.18.

Concerning **Study 3** and **Study 4**, the estimates are given by

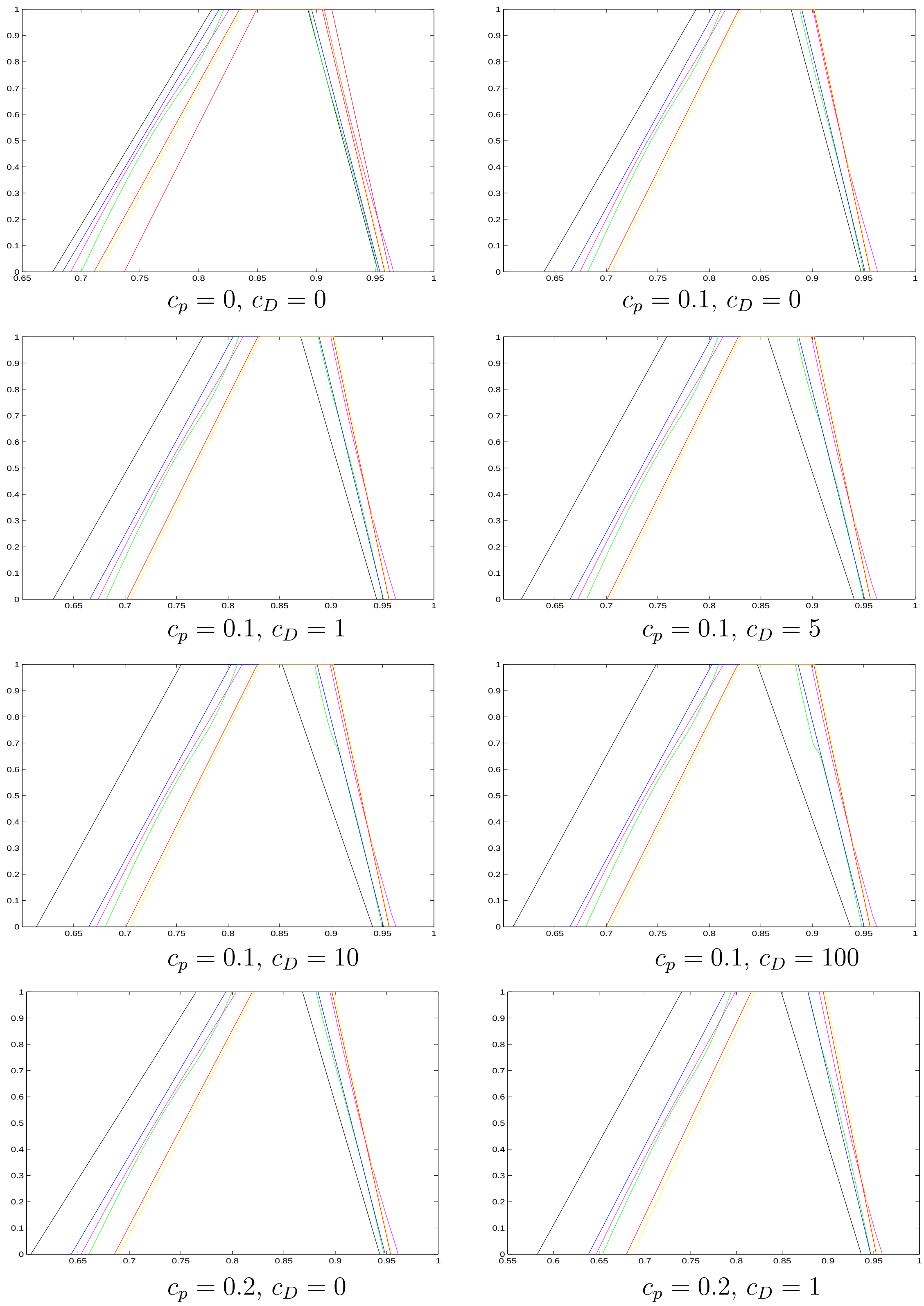


Figure 4.13: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 3 - CASE 3

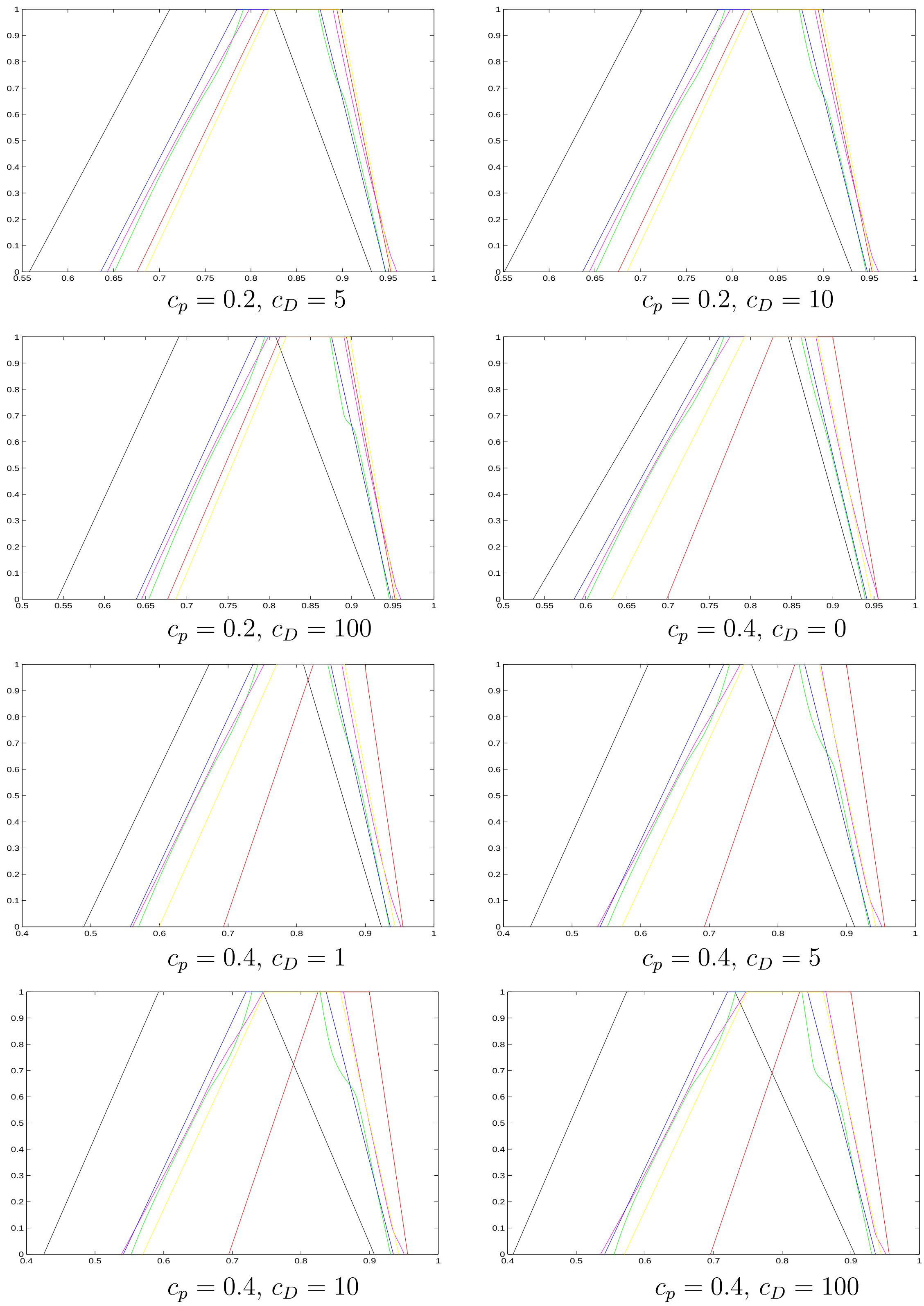


Figure 4.14: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 3 - CASE 3

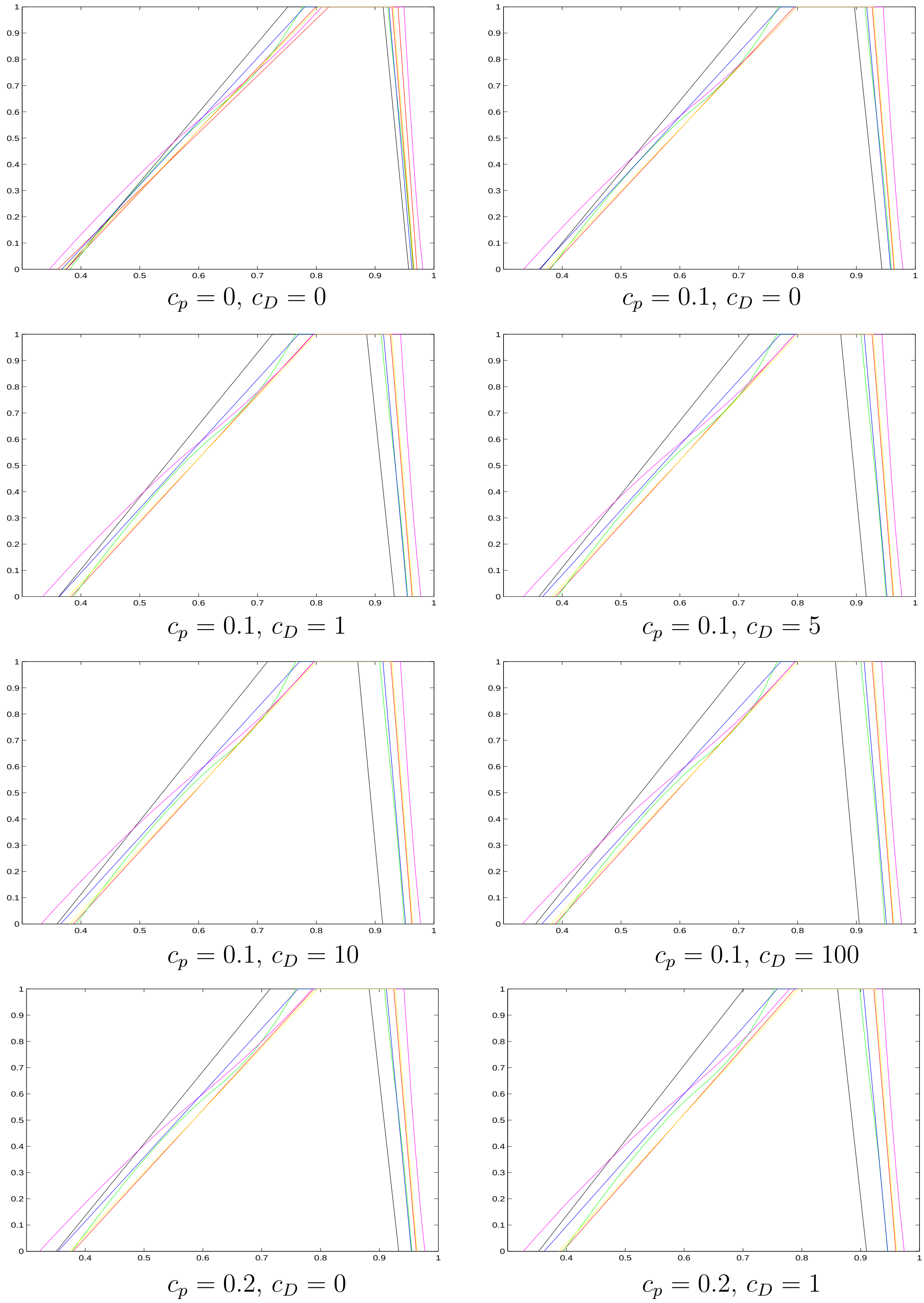
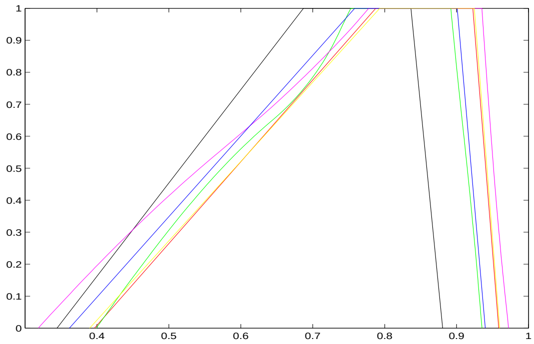
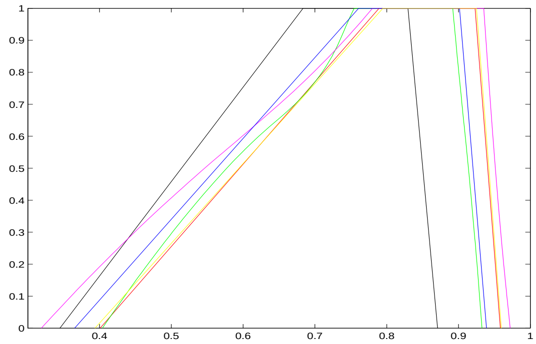


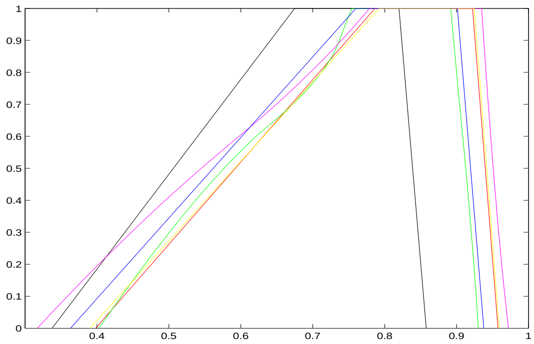
Figure 4.15: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 3 - CASE 4



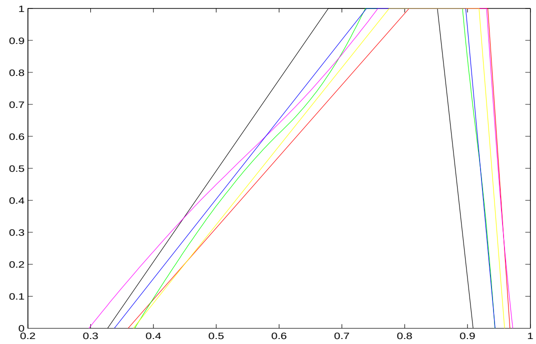
$c_p = 0.2,\ c_D = 5$



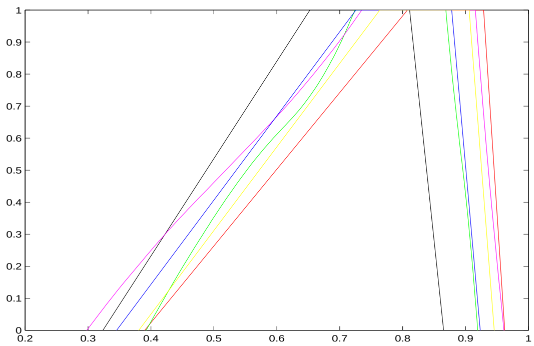
$c_p = 0.2,\ c_D = 10$



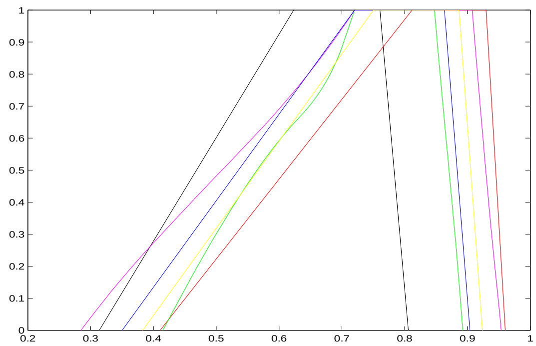
$c_p = 0.2,\ c_D = 100$



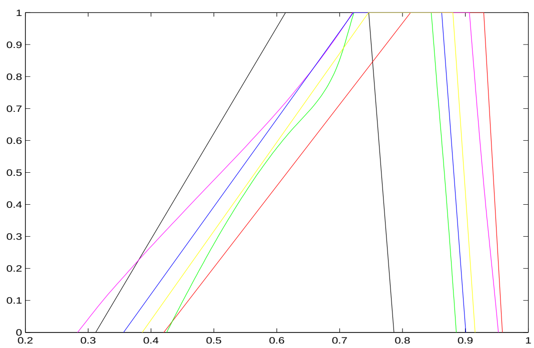
$c_p = 0.4,\ c_D = 0$



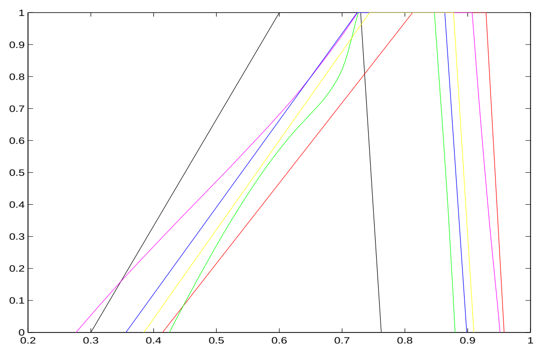
$c_p = 0.4,\ c_D = 1$



$c_p = 0.4,\ c_D = 5$



$c_p = 0.4,\ c_D = 10$



$c_p = 0.4,\ c_D = 100$

Figure 4.16: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 3 - CASE 4

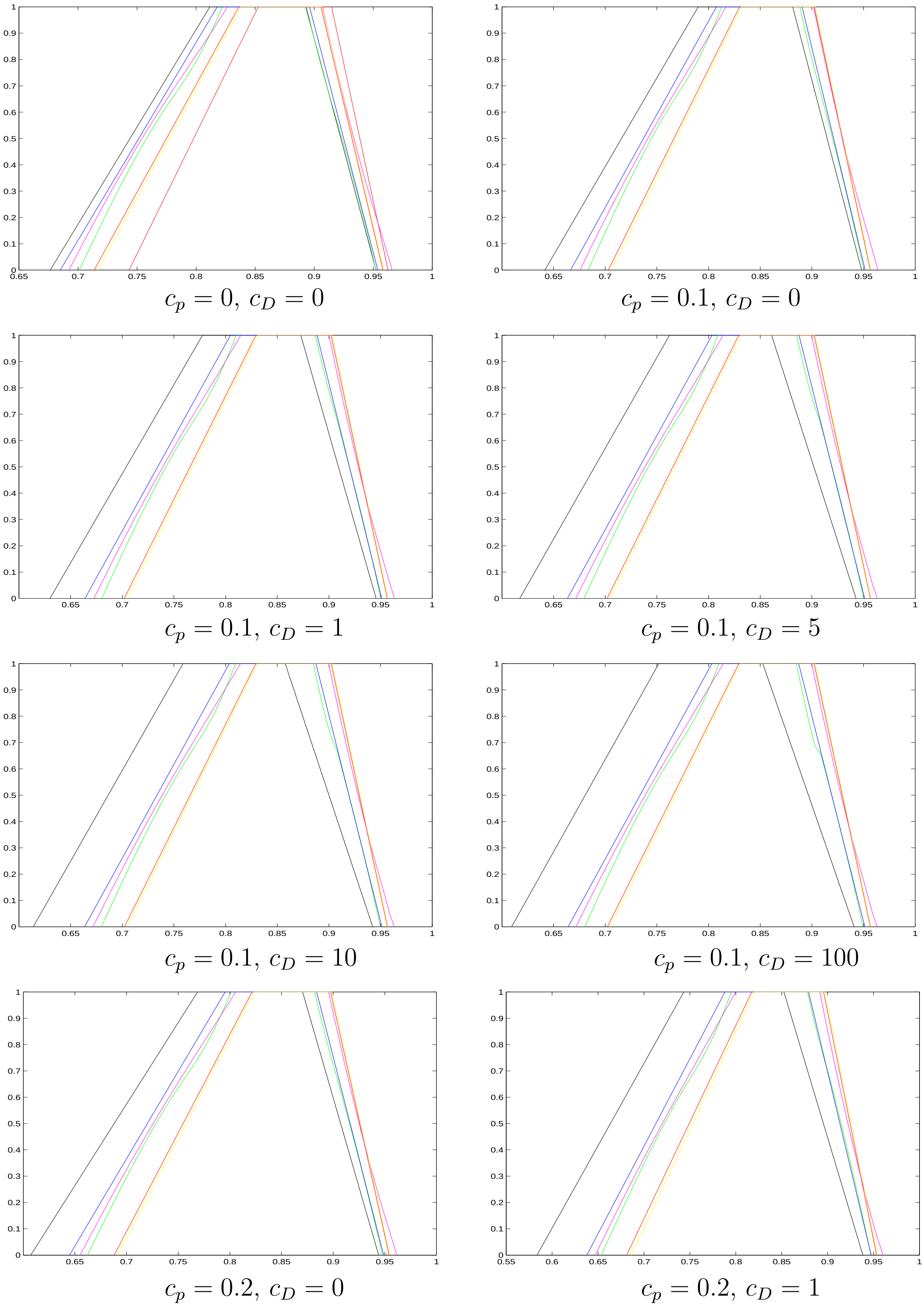


Figure 4.17: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 4 - CASE 3

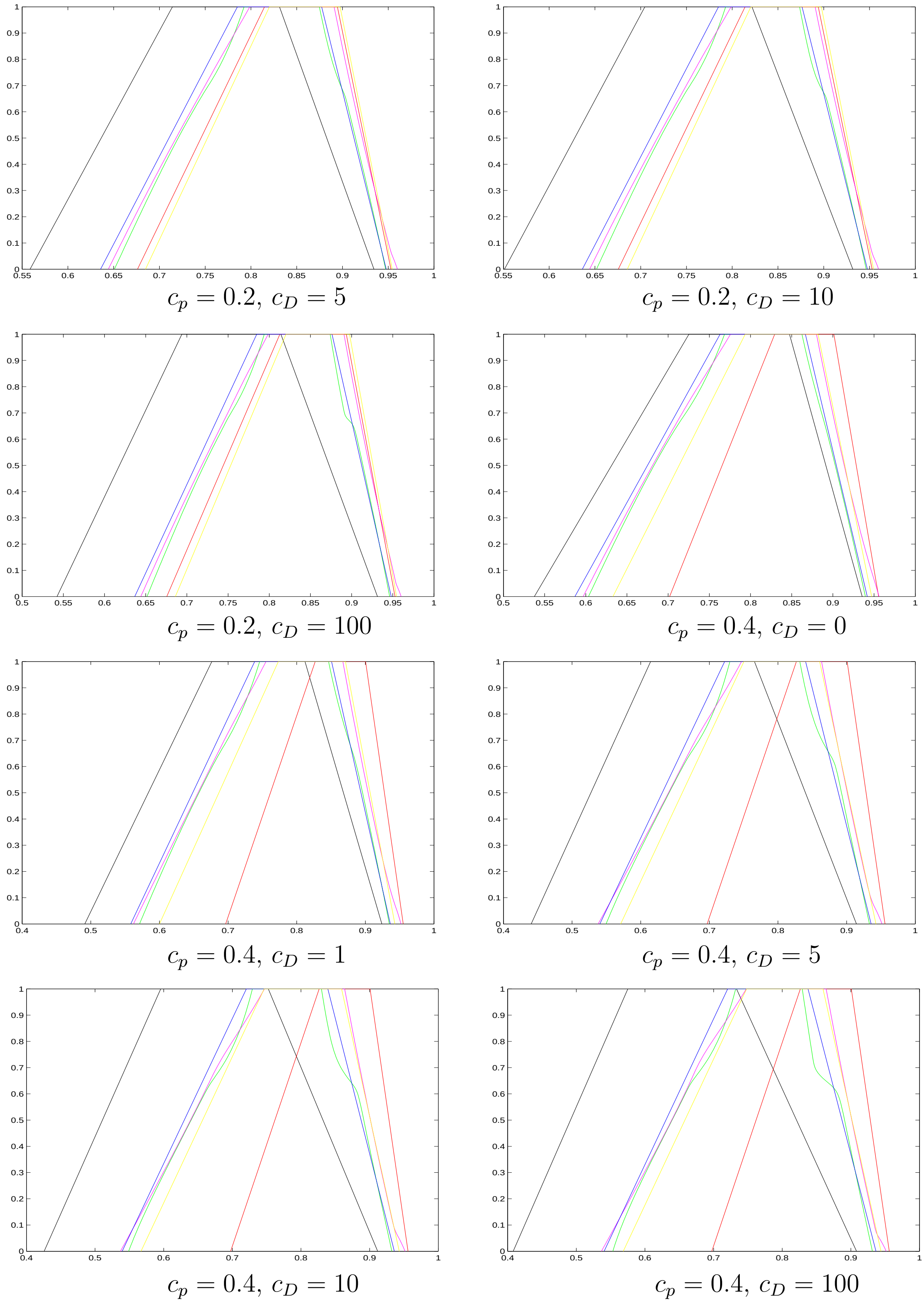


Figure 4.18: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 4 - CASE 3

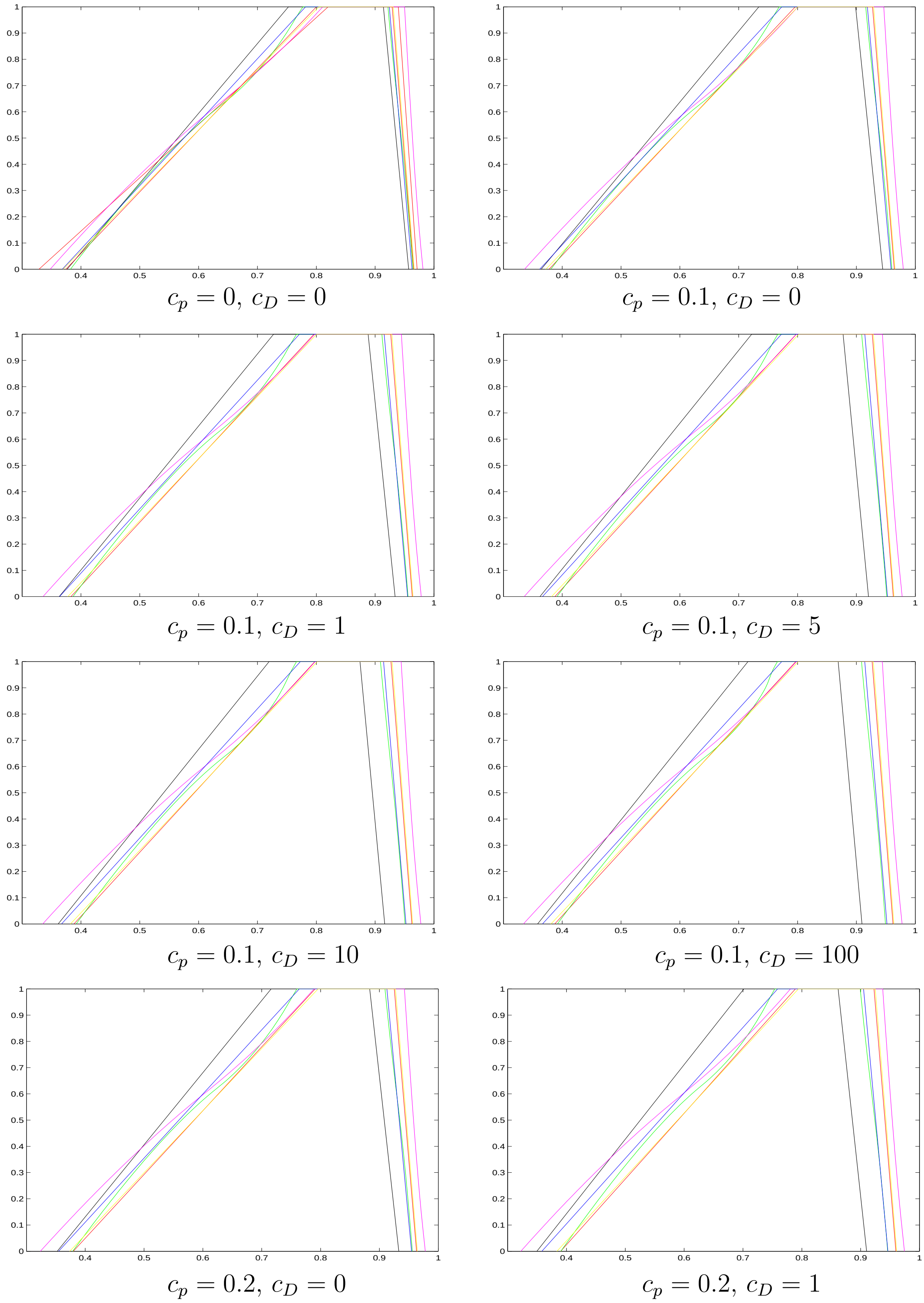
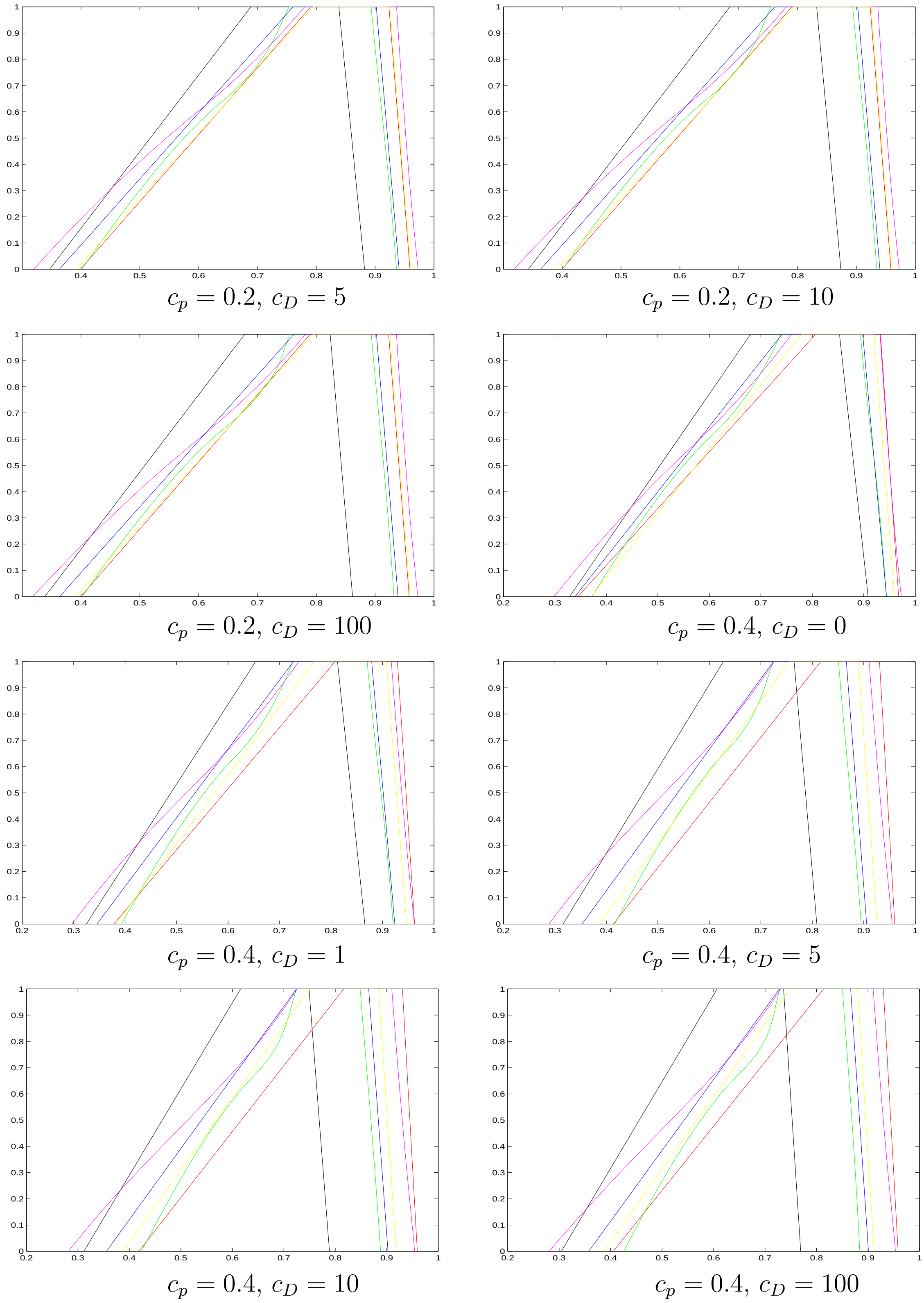


Figure 4.19: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 4 - CASE 4



$c_p = 0.2,\ c_D = 5$

$c_p = 0.2,\ c_D = 10$

$c_p = 0.2,\ c_D = 100$

$c_p = 0.4,\ c_D = 0$

$c_p = 0.4,\ c_D = 1$

$c_p = 0.4,\ c_D = 5$

$c_p = 0.4,\ c_D = 10$

$c_p = 0.4,\ c_D = 100$

Figure 4.20: Monte Carlo estimates of different location measures (mean, trimmed, Huber, Hampel, 1-norm and wabl/ldev/rdev) from the simulated fuzzy data in Study 4 - CASE 4

STUDY 3		**CASE 3**	**CASE 4**
Bias	$c_p \leq 0.2$	**Hampel** $(\theta = 1/3)$ **trimmed** $(\theta = 1/3)$	**Hampel** $(\theta = 1/3)$
	$c_p = 0.4$	**trimmed** $(\theta = 1/3)$	**trimmed** $(\theta = 1/3)$
	$c_D = 0$		
	Dispersion	none	low
Variance	$c_p = 0$	**mean** **Huber** $(\theta = 1/3)$	**1-norm median**
	$c_p \leq 0.2$	**trimmed** $(\theta = 1)$ **wabl median**	**1-norm median** **Huber** $(\theta = 1/3)$
	$c_p = 0.4$	**trimmed** $(\theta = 1)$	**wabl median** **trimmed** $(\theta = 1/3)$
	$c_D = 0$		
	Dispersion	high	high
MSE	$c_p = 0$	**mean** **Huber** $(\theta = 1/3)$	**mean** **1-norm median**
	$c_p \leq 0.2$	**trimmed** $(\theta = 1/3)$ **Hampel** $(\theta = 1/3)$	**Huber** $(\theta = 1/3)$ **Hampel** $(\theta = 1/3)$ **wabl median**
	$c_p = 0.4$	**trimmed** $(\theta = 1/3)$	**trimmed** $(\theta = 1/3)$ **wabl median**
	$c_D = 0$		
	Dispersion	medium	high

Table 4.17: Summary of the main conclusions from Study 3: the most suitable (if any) location measures/estimates

STUDY 4		**CASE 3**	**CASE 4**
Bias	$c_p \leq 0.2$	**Hampel** $(\theta = 1/3)$ **trimmed** $(\theta = 1/3)$	**Hampel** $(\theta = 1/3)$
	$c_p = 0.4$	**trimmed** $(\theta = 1/3)$	**trimmed** $(\theta = 1/3)$
	$c_D = 0$		
	Dispersion	none	low
Variance	$c_p = 0$	**1-norm median**	**wabl median**
	$c_p \leq 0.2$	**trimmed** $(\theta = 1)$ **Hampel** $(\theta = 1/3)$ **1-norm median**	**Hampel** **trimmed** $(\theta = 1)$ **1-norm median**
	$c_p = 0.4$	**trimmed** $(\theta = 1)$	**trimmed** $(\theta = 1/3)$ **1-norm median**
	$c_D = 0$		
	Dispersion	medium	high
MSE	$c_p = 0$	**wabl median**	**mean** **wabl median**
	$c_p \leq 0.2$	**Hampel** $(\theta = 1/3)$ **trimmed** $(\theta = 1/3)$	**Hampel**
	$c_p = 0.4$	**trimmed** $(\theta = 1/3)$	**trimmed** $(\theta = 1/3)$
	$c_D = 0$		
	Dispersion	low	medium

Table 4.18: Summary of the main conclusions from Study 4: the most suitable (if any) location measures/estimates

For the bias, variance and mean square error, the conclusions for the two cases in Studies 3 and 4 are summarized in Tables 4.17 and 4.18.

On the basis of the conclusions gathered in Tables 4.15, 4.16, 4.17 and 4.18, one can conclude that there is no uniformly most appropriate location estimate. Actually, the outputs seem to depend even more on the distribution considered for the non-contaminated and contaminated distributions, or the involved case, than on the sample size. A rather general assertion is that the 1-norm median is the best choice in many cases of Study 1 and Study 2 in terms of any of the considered measures (bias, variance or mean square error), above all in Case 2. In the rest of cases and studies, the best estimate is not as clear as in Case 2 or in the general comments given for interval-valued data, but the Huber and Hampel M-estimators still behave well for small amounts of contamination and the trimmed means when this proportion of contamination is increased. In Cases 3 and 4, with asymmetric non-contaminated distribution and fuzzy numbers having 0-levels contained in the interval $[0, 1]$, the distinction between the advantages of using these estimates in situations of small or big amounts of contamination is not as evident.

The comparisons in this section have only involved trapezoidal fuzzy numbers. For practical purposes, and to ease both the drawing and the computing processes, the fuzzy rating method has been usually applied by considering trapezoidal fuzzy numbers, but this is not at all mandatory. Actually, Pedrycz [148], Grzegorzewski [94, 95, 96], Grzegorzewski and Pasternak-Winiarska [97], Ban *et al.* [12] and others have provided with different arguments to employ triangular or trapezoidal fuzzy numbers or approximations preserving ambiguity, expected interval, etc.

Anyway, as it has been recently outlined in [21] *"... in connection with the shape of fuzzy values (i.e. with subjectivity in drawing them) and from a statistical side, there is an open problem which must be addressed: a deep sensitivity analysis on the choice of the fuzzy sets considered to express the experimental data with respect to the statistical conclusions drawn by applying the statistical methodology in* [20]. *Although sporadically we have made analyses in the course of some studies, we have not yet attempted to develop a unifying wide analysis to be submitted for publication."* A first attempt in this respect is gathered in the contribution by Lubiano *et al.* [127].

4.3 Comparative simulations for functional data

In connection with the comparative analysis for functional-valued data, simulations have already been presented in Section 2.4 (see p. 74) and Subsection 3.1.5 (see

p. 114), and no new simulation developments have been conducted.

The comparison for *Models 1* to *5* from Tables 2.4 (p. 78) and 3.1 (p. 115) leads to the conclusions in Table 4.19.

Model (1-5)	**Contamination**	**Best choice**
Model 1	$M = 5/25$	**mean**
Model 2	$M = 5/25$	**Huber**
Model 3	$M = 5/25$	**Huber**
Model 4	$M = 5/25$	**Huber**
Model 5	$M = 5$	**mean**
	$M = 25$	**Huber**

Table 4.19: Summary of the main conclusions from *Models 1-5* in Tables 2.4 and 3.1: the most suitable location measures/estimates

The comparison for *Models 6* to *9* from Tables 2.5 (p. 79) and 3.2 (p. 116) leads to the conclusions in Table 4.20.

Model (6-9)	**Contamination**	**Best choice**
Model 6	$M = 5/25$	**Huber**
Model 7	$M = 5/25$	**Huber**
Model 8	$M = 5/25$	**trimmed (ETMA)**
Model 9	$M = 5/25$	**trimmed (ETMA)**

Table 4.20: Summary of the main conclusions from *Models 6-9* in Tables 2.5 and 3.2: the most suitable location measures/estimates

Consequently, as a rather general conclusion one can state that there is no uniformly best choice but M-estimates (in some few cases with low values of M) behave better than trimmed means in *Models 1-7*, whereas trimmed means behave better otherwise.

4.4 Concluding remarks of this chapter

In this section, the location measures introduced along this work have been compared in the interval-, fuzzy- and functional-valued cases in terms of not only estimates, but also their bias, variance and mean squared errors. Although there is not any estimate behaving uniformly better, both the Huber and Hampel M-estimators and the trimmed means seem to be the best options in many of the studied situations. Indeed, Huber and Hampel M-estimators behave better for small amounts of contamination and the benefit of trimmed means is usually better with big amounts of contamination. Other measures, like the 1-norm median, become the most appropriate ones especially in certain situations of the fuzzy-valued case, but this advantage is apparently related to the distribution of the samples.

Final conclusions and open problems

In describing some of the most immediate open problems from this work, one should distinguish among the new challenges which could be addressed and those which directly derives from the developments in the work.

Among the new challenges we can mention the following:

- Robust approaches to measure location for imprecise-valued data have to be supplemented by robust approaches to measure scale, probably in a simultaneous way.
- Robust location estimates should be also complemented with robust testing hypothesis procedures concerning the locations estimates.

Among the future directions in connection with the studies already collected in this work, we can mention the following:

- To determine the influence function of different robust location measures in Chapters 2 and 3.
- In connection with the trimmed means, the solutions for non ideal conditions could be examined by determining the maximum bias in the simulation settings too. The theoretical study of how the maximum possible bias could be obtained would be a natural and important future step.
- Regarding M-estimates of location coming from the particularization of those for Hilbert space-valued data, different choices for the loss function fulfilling the suitable assumptions and the involved L^2 distances could be examined and compared.
- Concerning M-estimates of location based on the φ-wabl/ldev/rdev representation of fuzzy numbers and the associated L^1 metric, they have been shown to depend on the choice of φ, so a sensitivity analysis about the choice of this measure as well as how this choice can affect the robustness of the corresponding M-estimate would be interesting.

- With respect to the M-estimates of location extending the spatial median, it would be interesting to attempt the extension to fuzzy-valued data, and to develop comparisons like those in Section 4.2.

Conclusiones finales y problemas abiertos

Al describir los problemas abiertos más inmediatos en relación con esta memoria, conviene distinguir entre los nuevos retos que interesaría abordar y los que se derivan directamente de los desarrollos recogidos en el trabajo.

Entre los primeros cabe destacar los siguientes:

- Los estudios robustos sobre la medición de la tendencia central para datos con valores imprecisos deben complementarse con estudios robustos de la medición de la dispersión, probablemente simultaneando ambos.
- Otro complemento para las estimaciones robustas de la tendencia central pueden ser los procedimientos de contraste de hipótesis sobre tales estimaciones.

Entre las futuras líneas de investigación estrechamente ligadas a los estudios llevados a cabo en esta memoria, deben mencionarse los siguientes:

- La determinación de la función de influencia de las distintas medidas/estimaciones de tendencia central de los Capítulos 2 y 3.
- Por lo que se refiere a las medias recortadas, hay que examinar en el marco de los estudios de simulación las soluciones en condiciones no ideales con el fin de determinar el sesgo máximo. También sería interesante y natural realizar un análisis teórico de cómo puede alcanzarse el mayor sesgo posible.
- En relación con las M-estimaciones de tendencia central que se obtienen por la particularización de las correspondientes a datos con valores en espacios de Hilbert, pueden examinarse y compararse las basadas en elecciones diferentes tanto de la función de pérdida que satisfaga las suposiciones indicadas como de las métricas tipo L^2 involucradas.
- Por lo que concierne a las M-estimaciones de tendencia central basadas en la representación φ-wabl/ldev/rdev de números fuzzy y la métrica L^1 asociada,

se ha comprobado que dependen de φ, por lo que sería conveniente desarrollar un análisis de sensibilidad acerca de la elección de esa medida de ponderación y una discusión de cómo tal elección puede afectar a la robustez de las M-estimaciones correspondientes.

- Con respecto a las M-estimaciones de tendencia central que extienden la media espacial, sería interesante abordar su extensión a datos con valores fuzzy, y desarrollar así comparaciones como las recogidas dentro de la Sección 4.2.

Finale conclusies en open problemen

Bij de beschrijving van de meest voor de hand liggende open problemen gerelateerd aan dit werk moet men een onderscheid maken tussen de nieuwe uitdagingen die aangepakt kunnen worden en deze die direct afstammen van de ontwikkelingen in dit werk.

Bij de nieuwe uitdagingen kunnen we de volgende vernoemen:

- Robuuste procedures om de locatie van niet-precieze gegevens te bepalen moeten aangevuld worden met robuuste procedures om de schaal te meten, waarschijnlijk gelijktijdig met de locatie.
- Robuuste locatieschattingen moeten ook aangevuld worden met robuuste procedures voor hypothesetoetsen met betrekking tot de locatie op basis van deze schattingen.

Als richtingen voor toekomstig onderzoek in verband met de studies verzameld in dit werk, kunnen we de volgende vermelden:

- De invloedsfunctie bepalen van de verschillende robuuste locatieschatters in Hoofdstukken 2 en 3.
- In verband met de getrimde gemiddelden kunnen de oplossingen in niet ideale situaties nog verder onderzocht worden door het bepalen van de maximum bias in de simulatie settings. Theoretisch onderzoek om na te gaan hoe de maximum bias bereikt kan worden is een natuurlijke en belangrijke toekomstige stap.
- Met betrekking tot M-schatters voor locatie afkomstig van deze schatters voor Hilbertruimte-waardige gegevens kunnen verschillende verliesfuncties die voldoen aan de voorwaarden en de betrokken L^2 afstanden onderzocht en vergeleken worden.
- Voor de M-schatters voor locatie gebaseerd op de φ-wabl/ldev/rdev representatie voor vaaggetallen en de bijbehorende L^1 metriek werd aangetoond dat ze afhangen van de keuze van φ, zodat het interessant zou zijn om via een sensitiviteitsanalyse na te gaan wat het effect is van de keuze van φ en hoe deze keuze de robuustheid van de overeenkomstige M-schatter beïnvloedt.

- Wat betreft de M-schatter of locatie als uitbreiding van de spatiale mediaan, zou het interessant zijn om deze te proberen uitbreiden naar vaagwaardige gegevens en om vergelijkingen zoals in Sectie 4.2 uit te voeren.

Appendix. Proofs of the new results on preliminary and supporting tools

Proof of Proposition 1.1.4. (p. 8)

First, consider $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$. Because of the properties stated in Proposition 1.1.3, the functions $\mathrm{ldev}^{\varphi}_{\widetilde{U}}$ and $\mathrm{rdev}^{\varphi}_{\widetilde{U}}$ should be left-continuous functions of α on $(0,1]$ and right-continuous at 0. They should also be non-increasing functions of α on $[0,1]$. Moreover,

$$\mathrm{rdev}^{\varphi}_{\widetilde{U}}(1) \geq -\mathrm{ldev}^{\varphi}_{\widetilde{U}}(1), \quad \int_{[0,1]} \mathrm{ldev}^{\varphi}_{\widetilde{U}}(\alpha)\, d\varphi(\alpha) = \int_{[0,1]} \mathrm{rdev}^{\varphi}_{\widetilde{U}}(\alpha)\, d\varphi(\alpha) \geq 0.$$

Since $\mathrm{wabl}^{\varphi}(\widetilde{U}) \in \mathbb{R}$ and for all $\alpha \in [0,1]$ it holds that

$$\widetilde{U}_\alpha = \left[\mathrm{wabl}^{\varphi}(\widetilde{U}) - \mathrm{ldev}^{\varphi}_{\widetilde{U}}(\alpha), \mathrm{wabl}^{\varphi}(\widetilde{U}) + \mathrm{rdev}^{\varphi}_{\widetilde{U}}(\alpha)\right],$$

it follows that the real number $\mathrm{wabl}^{\varphi}(\widetilde{U})$ and the two functions $\mathrm{ldev}^{\varphi}_{\widetilde{U}}$ and $\mathrm{rdev}^{\varphi}_{\widetilde{U}}$ indeed satisfy Conditions $i) - iii)$.

On the other hand, given $m \in \mathbb{R}$, if $l^* : [0,1] \to \mathbb{R}$, $r^* : [0,1] \to \mathbb{R}$ are mappings satisfying Conditions $i) - ii)$, then, the functions $l : [0,1] \to \mathbb{R}$ and $r : [0,1] \to \mathbb{R}$ given by

$$l(\alpha) = l^*(\alpha) - m, \ r(\alpha) = m + r^*(\alpha)$$

satisfy that l and r are left-continuous non-increasing functions in $(0,1]$ and right-continuous at $\alpha = 0$, and $-l(1) = m - l^*(1) \leq m + r^*(1) = r(1)$, whence Proposition 1.1.3 ensures that there exists a unique $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ such that for $\alpha \in [0,1]$

$$\widetilde{U}_\alpha = [-l(\alpha), r(\alpha)] = [m - l^*(\alpha), m + r^*(\alpha)].$$

Furthermore, if there is an absolutely continuous probability measure φ on $([0,1], \mathcal{B}_{[0,1]})$, with positive mass function on $(0,1)$ and $\int_{[0,1]} l^*(\alpha)\, d\varphi(\alpha) = \int_{[0,1]} r^*(\alpha)\, d\varphi(\alpha)$, then

$$\int_{[0,1]} l(\alpha)\, d\varphi(\alpha) = m - \int_{[0,1]} l^*(\alpha)\, d\varphi(\alpha)$$

$$= m - \int_{[0,1]} r^*(\alpha)\, d\varphi(\alpha) = 2m - \int_{[0,1]} r(\alpha)\, d\varphi(\alpha).$$

Hence,

$$m = \int_{[0,1]} \frac{l(\alpha)+r(\alpha)}{2}\, d\varphi(\alpha) = \int_{[0,1]} \operatorname{mid} \widetilde{U}_\alpha \, d\varphi(\alpha) = \operatorname{wabl}^\varphi(\widetilde{U}).$$

Moreover, for all $\alpha \in [0,1]$

$$\operatorname{ldev}^\varphi_{\widetilde{U}}(\alpha) = \operatorname{wabl}^\varphi(\widetilde{U}) - \inf \widetilde{U}_\alpha = m - l(\alpha) = l^*(\alpha),$$

$$\operatorname{rdev}^\varphi_{\widetilde{U}}(\alpha) = \sup \widetilde{U}_\alpha - \operatorname{wabl}^\varphi(\widetilde{U}) = r(\alpha) - m = r^*(\alpha).$$

□

Proof of Proposition 1.3.2. (p. 26)

Indeed, $\mathfrak{D}^\varphi_\theta$ satisfies

- the nonnegativity (or separation axiom); it is trivial that $\mathfrak{D}^\varphi_\theta(\widetilde{U},\widetilde{V}) \geq 0$ whatever $\widetilde{U},\widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$ may be;
- the identity of indiscernibles (or coincidence axiom); actually, a necessary and sufficient condition for $\mathfrak{D}^\varphi_\theta(\widetilde{U},\widetilde{V})$ to vanish is that $\boldsymbol{\rho}^\varphi_{\mathbf{2}}(\widetilde{U},\widetilde{V})$ also vanishes whence, because of $\boldsymbol{\rho}^\varphi_{\mathbf{2}}$ being a metric, $\widetilde{U} = \widetilde{V}$;
- the symmetry; it is trivial that $\mathfrak{D}^\varphi_\theta(\widetilde{U},\widetilde{V}) = \mathfrak{D}^\varphi_\theta(\widetilde{V},\widetilde{U})$ whatever $\widetilde{U},\widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$ may be;
- the subadditivity (or triangle inequality); if $\widetilde{U},\widetilde{V},\widetilde{W} \in \mathcal{F}_c(\mathbb{R}^p)$, then since $\|\cdot\|$ is a norm and $\boldsymbol{\rho}^\varphi_{\mathbf{2}}$ is a metric,

$$\left[\mathfrak{D}^\varphi_\theta(\widetilde{U},\widetilde{V})\right]^2 \leq (1-\theta)\left[\|\mathbf{S}^\varphi(\widetilde{U}) - \mathbf{S}^\varphi(\widetilde{W})\| + \|\mathbf{S}^\varphi(\widetilde{W}) - \mathbf{S}^\varphi(\widetilde{V})\|\right]^2$$

$$+\,\theta\left[\boldsymbol{\rho}^\varphi_{\mathbf{2}}(\widetilde{U},\widetilde{W}) + \boldsymbol{\rho}^\varphi_{\mathbf{2}}(\widetilde{W},\widetilde{V})\right]^2 = \left[\mathfrak{D}^\varphi_\theta(\widetilde{U},\widetilde{W})\right]^2 + \left[\mathfrak{D}^\varphi_\theta(\widetilde{W},\widetilde{V})\right]^2$$

$$+2(1-\theta)\|\mathbf{S}^\varphi(\widetilde{U}) - \mathbf{S}^\varphi(\widetilde{W})\| \cdot \|\mathbf{S}^\varphi(\widetilde{W}) - \mathbf{S}^\varphi(\widetilde{V})\| + 2\,\theta\, \boldsymbol{\rho}^\varphi_{\mathbf{2}}(\widetilde{U},\widetilde{W}) \cdot \boldsymbol{\rho}^\varphi_{\mathbf{2}}(\widetilde{W},\widetilde{V}),$$

and, due to the fact that

$$\left[(1-\theta)\|\mathbf{S}^\varphi(\widetilde{U}) - \mathbf{S}^\varphi(\widetilde{W})\| \cdot \|\mathbf{S}^\varphi(\widetilde{W}) - \mathbf{S}^\varphi(\widetilde{V})\| + \theta\, \boldsymbol{\rho}^\varphi_{\mathbf{2}}(\widetilde{U},\widetilde{W}) \cdot \boldsymbol{\rho}^\varphi_{\mathbf{2}}(\widetilde{W},\widetilde{V})\right]^2$$

$$= \left[\mathfrak{D}^\varphi_\theta(\widetilde{U},\widetilde{W})\right]^2 \cdot \left[\mathfrak{D}^\varphi_\theta(\widetilde{W},\widetilde{V})\right]^2$$

$$-\,\theta(1-\theta)\left[\|\mathbf{S}^\varphi(\widetilde{U}) - \mathbf{S}^\varphi(\widetilde{W})\| \cdot \boldsymbol{\rho}^\varphi_{\mathbf{2}}(\widetilde{W},\widetilde{V}) - \|\mathbf{S}^\varphi(\widetilde{W}) - \mathbf{S}^\varphi(\widetilde{V})\| \cdot \boldsymbol{\rho}^\varphi_{\mathbf{2}}(\widetilde{U},\widetilde{W})\right]^2$$

$$\leq \left[\mathfrak{D}_\theta^\varphi(\widetilde{U},\widetilde{W})\right]^2 \cdot \left[\mathfrak{D}_\theta^\varphi(\widetilde{W},\widetilde{V})\right]^2,$$

we have that

$$\left[\mathfrak{D}_\theta^\varphi(\widetilde{U},\widetilde{V})\right]^2 \leq \left[\mathfrak{D}_\theta^\varphi(\widetilde{U},\widetilde{W}) + \mathfrak{D}_\theta^\varphi(\widetilde{W},\widetilde{V})\right]^2. \qquad \square$$

Proof of Proposition 1.3.4. (p. 27)

Indeed,

$$\mathfrak{D}_\theta^\varphi(\widetilde{U},\widetilde{V}) = \sqrt{(1-\theta)\|\mathbf{S}^\varphi(\widetilde{U}) - \mathbf{S}^\varphi(\widetilde{V})\|^2 + \theta\left[\boldsymbol{\rho}_\mathbf{2}^\varphi(\widetilde{U},\widetilde{V})\right]^2}$$

$$\geq \sqrt{\theta\left[\boldsymbol{\rho}_\mathbf{2}^\varphi(\widetilde{U},\widetilde{V})\right]^2} = \sqrt{\theta}\cdot\boldsymbol{\rho}_\mathbf{2}^\varphi(\widetilde{U},\widetilde{V}).$$

On the other hand,

$$\|\mathbf{S}^\varphi(\widetilde{U}) - \mathbf{S}^\varphi(\widetilde{V})\|^2 = \left\|\int_{[0,1]\times\mathbb{S}^{p-1}} \mathbf{u}\cdot\left[s_{\widetilde{U}}(\alpha,\mathbf{u}) - s_{\widetilde{V}}(\alpha,\mathbf{u})\right]\, d\lambda_p(\mathbf{u})\, d\varphi(\alpha)\right\|^2$$

$$\leq \int_{[0,1]\times\mathbb{S}^{p-1}} \left\|\mathbf{u}\cdot\left[s_{\widetilde{U}}(\alpha,\mathbf{u}) - s_{\widetilde{V}}(\alpha,\mathbf{u})\right]\right\|^2 d\lambda_p(\mathbf{u})\, d\varphi(\alpha)$$

$$= \int_{[0,1]\times\mathbb{S}^{p-1}} \|s_{\widetilde{U}}(\alpha,\mathbf{u}) - s_{\widetilde{V}}(\alpha,\mathbf{u})\|^2\, d\lambda_p(\mathbf{u})\, d\varphi(\alpha) = \left[\boldsymbol{\rho}_\mathbf{2}^\varphi(\widetilde{U},\widetilde{V})\right]^2,$$

whence

$$\mathfrak{D}_\theta^\varphi(\widetilde{U},\widetilde{V}) \leq \boldsymbol{\rho}_\mathbf{2}^\varphi(\widetilde{U},\widetilde{V}). \qquad \square$$

Proof of Proposition 1.3.5. (p. 32)

Indeed, in case $p = 1$ we have that

$$\begin{aligned} \mathbf{S}^\varphi(\widetilde{U}) &= \mathrm{wabl}^\varphi\, \widetilde{U}, \\ -\mathbf{dev}_{\widetilde{U}}^\varphi(\alpha,-1) &= \mathrm{wabl}^\varphi\, \widetilde{U} - \inf \widetilde{U}_\alpha = \mathrm{ldev}_{\widetilde{U}}^\varphi(\alpha), \\ \mathbf{dev}_{\widetilde{U}}^\varphi(\alpha,1) &= \sup \widetilde{U}_\alpha - \mathrm{wabl}^\varphi\, \widetilde{U} = \mathrm{rdev}_{\widetilde{U}}^\varphi(\alpha). \end{aligned} \qquad \square$$

Proof of Proposition 1.3.6. (p. 33)

Indeed, for each $\alpha \in [0,1]$

$$\frac{1}{2}\int_{[0,1]} \left[\left(\mathrm{wabl}^\varphi(\widetilde{U}) - \xi\cdot\mathrm{ldev}_{\widetilde{U}}^\varphi(\alpha)\right) - \left(\mathrm{wabl}^\varphi(\widetilde{V}) - \xi\cdot\mathrm{ldev}_{\widetilde{V}}^\varphi(\alpha)\right)\right]^2 d\nu(\xi)$$

$$+\frac{1}{2}\int_{[0,1]}\left[\left(\operatorname{wabl}^{\varphi}(\widetilde{U})+\xi\cdot \operatorname{rdev}^{\varphi}_{\widetilde{U}}(\alpha)\right)-\left(\operatorname{wabl}^{\varphi}(\widetilde{V})+\xi\cdot \operatorname{rdev}^{\varphi}_{\widetilde{V}}(\alpha)\right)\right]^2 d\nu(\xi)$$

$$=\left[\operatorname{wabl}^{\varphi}(\widetilde{U})-\operatorname{wabl}^{\varphi}(\widetilde{V})\right]^2+\frac{1}{2}\left[\operatorname{ldev}^{\varphi}_{\widetilde{U}}(\alpha)-\operatorname{ldev}^{\varphi}_{\widetilde{V}}(\alpha)\right]^2\cdot\int_{[0,1]}\xi^2\, d\nu(\xi)$$

$$+\frac{1}{2}\left[\operatorname{rdev}^{\varphi}_{\widetilde{U}}(\alpha)-\operatorname{ldev}^{\varphi}_{\widetilde{U}}(\alpha)\right]^2\cdot\int_{[0,1]}\xi^2\, d\nu(\xi)-\left[\operatorname{wabl}^{\varphi}(\widetilde{U})-\operatorname{wabl}^{\varphi}(\widetilde{V})\right]^2\cdot\int_{[0,1]}2\xi\, d\nu(\xi)$$

$$+\left[\operatorname{wabl}^{\varphi}(\widetilde{U})-\operatorname{wabl}^{\varphi}(\widetilde{V})\right]\cdot\left[\operatorname{mid}\widetilde{U}_\alpha-\operatorname{mid}\widetilde{V}_\alpha\right]\cdot\int_{[0,1]}2\xi\, d\nu(\xi),$$

whence $\mathcal{D}^{\varphi}_{\vartheta}(\widetilde{U},\widetilde{V})=\mathfrak{D}^{\varphi}_{\theta_\nu}(\widetilde{U},\widetilde{V})$ with $\theta_\nu=\int_{[0,1]}\xi^2\, d\nu(\xi)$.

Consequently, for any non-degenerate symmetric probability measure ϑ on $[-1,1]$ there is a probability measure ν on $[0,1]$ and non-degenerate at 0 such that $\vartheta(\xi) = .5\cdot\nu(-\xi)+.5\cdot\nu(\xi)$ and $\mathcal{D}^{\varphi}_{\vartheta}(\widetilde{U},\widetilde{V})=\mathfrak{D}^{\varphi}_{\theta_\nu}(\widetilde{U},\widetilde{V})$ with $\theta_\nu=\int_{[0,1]}\xi^2\, d\nu(\xi)$. Conversely, for any parameter value $\theta\in(0,1]$ there are probability measures ν_θ on $[0,1]$ and non-degenerate at 0 such that $\mathcal{D}^{\varphi}_{\vartheta_\theta}(\widetilde{U},\widetilde{V})=\mathfrak{D}^{\varphi}_{\theta}(\widetilde{U},\widetilde{V})$ for $\vartheta_\theta(\xi)=.5\cdot\nu_\theta(-\xi)+.5\cdot\nu_\theta(\xi)$. For instance, we can consider ν_θ to be the Bernoulli distribution with parameter $\{(\sqrt{1+4\theta}-1)/2\}$ or the $\operatorname{Beta}\left((\sqrt{\theta^2+8\theta}+\theta)/2,1\right)$ distribution, etc. □

Proof of Proposition 1.3.9. (p. 39)

Indeed, because of the properties for the absolute value we can conclude for each α that

$$|\boldsymbol{v}^{\varphi}_{\widetilde{U}}(\alpha)-\boldsymbol{v}^{\varphi}_{\widetilde{U}}(\alpha)|^1_\theta=|\operatorname{wabl}^{\varphi}(\widetilde{U})-\operatorname{wabl}^{\varphi}(\widetilde{V})|$$

$$+\frac{\theta}{2}\cdot|\operatorname{wabl}^{\varphi}(\widetilde{U})-\inf\widetilde{U}_\alpha-\operatorname{wabl}^{\varphi}(\widetilde{V})+\inf\widetilde{V}_\alpha|+\frac{\theta}{2}\cdot|\sup\widetilde{U}_\alpha-\operatorname{wabl}^{\varphi}(\widetilde{U})-\sup\widetilde{V}_\alpha+\operatorname{wabl}^{\varphi}(\widetilde{V})|.$$

Therefore, on one hand

$$|\boldsymbol{v}^{\varphi}_{\widetilde{U}}(\alpha)-\boldsymbol{v}^{\varphi}_{\widetilde{U}}(\alpha)|^1_\theta\geq|\operatorname{wabl}^{\varphi}(\widetilde{U})-\operatorname{wabl}^{\varphi}(\widetilde{V})|$$

$$+\frac{\theta}{2}\cdot|\inf\widetilde{U}_\alpha-\inf\widetilde{V}_\alpha|-\frac{\theta}{2}\cdot|\operatorname{wabl}^{\varphi}(\widetilde{U})-\operatorname{wabl}^{\varphi}(\widetilde{V})|$$

$$+\frac{\theta}{2}\cdot|\sup\widetilde{U}_\alpha-\sup\widetilde{V}_\alpha|-\frac{\theta}{2}\cdot|\operatorname{wabl}^{\varphi}(\widetilde{U})-\operatorname{wabl}^{\varphi}(\widetilde{V})|$$

$$=(1-\theta)\cdot|\operatorname{wabl}^{\varphi}(\widetilde{U})-\operatorname{wabl}^{\varphi}(\widetilde{V})|+\theta\cdot\left[\frac{1}{2}\cdot|\inf\widetilde{U}_\alpha-\inf\widetilde{V}_\alpha|+\frac{1}{2}\cdot|\sup\widetilde{U}_\alpha-\sup\widetilde{V}_\alpha|\right]$$

$$\geq\theta\cdot\left[\frac{1}{2}\cdot|\inf\widetilde{U}_\alpha-\inf\widetilde{V}_\alpha|+\frac{1}{2}\cdot|\sup\widetilde{U}_\alpha-\sup\widetilde{V}_\alpha|\right],$$

whence one derives the first inequality.

On the other hand, one can conclude for each α that

$$|\boldsymbol{v}^{\varphi}_{\widetilde{U}}(\alpha)-\boldsymbol{v}^{\varphi}_{\widetilde{U}}(\alpha)|^1_\theta\leq|\operatorname{wabl}^{\varphi}(\widetilde{U})-\operatorname{wabl}^{\varphi}(\widetilde{V})|$$

$$+\frac{\theta}{2}\cdot|\inf\widetilde{U}_\alpha-\inf\widetilde{V}_\alpha|+\frac{\theta}{2}\cdot|\mathrm{wabl}^\varphi(\widetilde{U})-\mathrm{wabl}^\varphi(\widetilde{V})|$$

$$+\frac{\theta}{2}\cdot|\sup\widetilde{U}_\alpha-\sup\widetilde{V}_\alpha|+\frac{\theta}{2}\cdot|\mathrm{wabl}^\varphi(\widetilde{U})-\mathrm{wabl}^\varphi(\widetilde{V})|$$

$$=(1+\theta)\cdot|\mathrm{wabl}^\varphi(\widetilde{U})-\mathrm{wabl}^\varphi(\widetilde{V})|+\theta\cdot\left[\frac{1}{2}\cdot|\inf\widetilde{U}_\alpha-\inf\widetilde{V}_\alpha|+\frac{1}{2}\cdot|\sup\widetilde{U}_\alpha-\sup\widetilde{V}_\alpha|\right],$$

whence

$$|\mathrm{wabl}^\varphi(\widetilde{U})-\mathrm{wabl}^\varphi(\widetilde{V})|=\left|\int_{[0,1]}\left[\mathrm{mid}\,\widetilde{U}_\alpha-\mathrm{mid}\,\widetilde{V}_\alpha\right]\,d\varphi(\alpha)\right|$$

$$\leq\int_{[0,1]}|\mathrm{mid}\,\widetilde{U}_\alpha-\mathrm{mid}\,\widetilde{V}_\alpha|\,d\varphi(\alpha)\leq\int_{[0,1]}d_H(\widetilde{U}_\alpha,\widetilde{U}_\alpha)\,d\varphi(\alpha)=\boldsymbol{d}_{\mathbf{1}}^{\varphi}(\widetilde{U},\widetilde{V}),$$

and due to the fact that $\boldsymbol{d}_{\mathbf{1}}^{\varphi}(\widetilde{U},\widetilde{V})\leq 2\cdot\boldsymbol{\rho}_{\mathbf{1}}^{\varphi}(\widetilde{U},\widetilde{V})$, one can easily derive the second inequality. □

Proof of Proposition 1.4.8. (p. 46)

Indeed,

$$E\left(\left[\mathfrak{D}_\theta^\varphi(\mathcal{X},\widetilde{U})\right]^2\right)=(1-\theta)\,\mathsf{E}\left(\left\|\mathbf{S}^\varphi(\mathcal{X})-\mathbf{S}^\varphi(\widetilde{U})\right\|^2\right)+\theta\,E\left(\left[\boldsymbol{\rho}_{\mathbf{2}}^{\varphi}(\mathcal{X},\widetilde{U})\right]^2\right).$$

On one hand, it is well-known that $\widetilde{E}(\mathcal{X})$ is the Fréchet expectation associated with $\boldsymbol{\rho}_{\mathbf{2}}^{\varphi}$, that is,

$$\widetilde{E}(\mathcal{X})=\arg\min_{\widetilde{U}\in\mathcal{F}_c^*(\mathbb{R}^p)}E\left(\left[\boldsymbol{\rho}_{\mathbf{2}}^{\varphi}(\mathcal{X},\widetilde{U})\right]^2\right).$$

On the other hand, if $\mathcal{X}$ is a random fuzzy vector, then $\mathbf{S}^\varphi(\mathcal{X})$ is a random vector (see, for instance, Aletti and Bongiorno [3], in case φ is the Lebesgue measure on $([0,1],\mathcal{B}_{[0,1]})$, result that can be straightforwardly extended for any φ). Furthermore,

$$\mathsf{E}(\mathbf{S}^\varphi(\mathcal{X}))=\arg\min_{\widetilde{U}\in\mathcal{F}_c^*(\mathbb{R}^p)}\mathsf{E}\left(\left\|\mathbf{S}^\varphi(\mathcal{X})-\mathbf{S}^\varphi(\widetilde{U})\right\|^2\right)$$

and, because of the sufficient conditions allowing us to apply Fubini's Theorem, we have that

$$\mathsf{E}(\mathbf{S}^\varphi(\mathcal{X}))=\int_{[0,1]\times\mathbb{S}^{p-1}}\mathsf{E}\big(\mathbf{u}\cdot s_{\mathcal{X}}(\alpha,\mathbf{u})\big)\,d\lambda_p(\mathbf{u})\,d\varphi(\alpha)$$

$$=\int_{[0,1]\times\mathbb{S}^{p-1}}\mathbf{u}\cdot E\big(s_{\mathcal{X}}(\alpha,\mathbf{u})\big)\,d\lambda_p(\mathbf{u})\,d\varphi(\alpha).$$

Since $E(s_{\mathcal{X}}(\cdot,\cdot))=s_{\widetilde{E}(\mathcal{X})}(\cdot,\cdot)$, then

$$\mathsf{E}(\mathbf{S}^\varphi(\mathcal{X}))=\int_{[0,1]\times\mathbb{S}^{p-1}}\mathbf{u}\cdot s_{\widetilde{E}(\mathcal{X})}(\alpha,\mathbf{u})\,d\lambda_p(\mathbf{u})\,d\varphi(\alpha)=\mathbf{S}^\varphi(\widetilde{E}(\mathcal{X})),$$

whence the result is proved. □

Proof of Proposition 1.4.9. (p. 49)

Indeed, whatever $\alpha \in [0,1]$ and the compact interval K may be, we have that

$$\mathcal{X}_\alpha^{-1}(K) = \{\omega \in \Omega \,:\, \big(\mathcal{X}(\omega)\big)_\alpha = K\}$$

$$= \bigcup_{\widetilde{x} \in \mathcal{X}(\Omega)\,:\,\widetilde{x}_\alpha = K} \{\omega \in \Omega \,:\, \mathcal{X}(\omega) = \widetilde{x}\} = \bigcup_{\widetilde{x} \in \mathcal{X}(\Omega)\,:\,\widetilde{x}_\alpha = K} \mathcal{X}_\alpha^{-1}(\{\widetilde{x}\}),$$

$$\begin{aligned}(2c - \mathcal{X})_\alpha^{-1}(K) &= \{\omega \in \Omega \,:\, \big(2c - \mathcal{X}(\omega)\big)_\alpha = K\} \\ &= \bigcup_{\widetilde{x} \in (2c-\mathcal{X})(\Omega)\,:\,\widetilde{x}_\alpha = K} \{\omega \in \Omega \,:\, (2c - \mathcal{X})(\omega) = \widetilde{x}\} \\ &= \bigcup_{\widetilde{x} \in (2c-\mathcal{X})(\Omega)\,:\,\widetilde{x}_\alpha = K} \{\omega \in \Omega \,:\, \mathcal{X}(\omega) = 2c - \widetilde{x}\} \\ &= \bigcup_{\widetilde{x} \in (2c-\mathcal{X})(\Omega)\,:\,\widetilde{x}_\alpha = K} \mathcal{X}_\alpha^{-1}(\{2c - \widetilde{x}\}),\end{aligned}$$

whence, because of the symmetry of $\mathcal{X}$ about c, $\mathcal{X}_\alpha^{-1}(K) \overset{a.s.[P]}{=} (2c - \mathcal{X})_\alpha^{-1}(K)$ and, in consequence, $\mathcal{X}_\alpha$ and $(2c - \mathcal{X})_\alpha$ are identically distributed. □

Proof of Proposition 1.4.10. (p. 51)

Since $\mathcal{X} \overset{d}{=} 2c - \mathcal{X}$, then $\widetilde{E}(\mathcal{X}) = \widetilde{E}(2c - \mathcal{X})$. Because of the equivariance of the Aumann-type mean value under affine transformations, we have that

$$\widetilde{E}(\mathcal{X}) = 2c - \widetilde{E}(\mathcal{X}).$$

By adding $\widetilde{E}(\mathcal{X})$ to both members in the last equality,

$$2\,\widetilde{E}(\mathcal{X}) = 2c + \ \widetilde{E}(\mathcal{X}) - \ \widetilde{E}(\mathcal{X})$$

and, hence,

$$\widetilde{E}(\mathcal{X}) = c + \frac{1}{2} \cdot \mathcal{O}_{\widetilde{E}(\mathcal{X})},$$

that is, for all $\alpha \in [0,1]$

$$\big(\widetilde{E}(\mathcal{X})\big)_\alpha = \big[c - \operatorname{spr}\big(\widetilde{E}(\mathcal{X})\big)_\alpha, c + \operatorname{spr}\big(\widetilde{E}(\mathcal{X})\big)_\alpha\big].$$

Consequently, $\widetilde{E}(\mathcal{X})$ is a symmetric fuzzy number about c. □

Bibliography

[1] Abbasbandy S, Amirfakhrian M (2006) The nearest approximation of a fuzzy quantity in parametric form. *Appl. Math. Comput.* **172**: 624–632

[2] Abbasbandy S, Asady B (2004) The nearest trapezoidal fuzzy number to a fuzzy quantity. *Appl. Math. Comput.* **156**: 381–386

[3] Aletti G, Bongiorno EN (2013) A decomposition theorem for fuzzy set-valued random variables. *Fuzzy Sets Syst.* **219**: 98–112

[4] Arribas-Gil A, Müller HG (2014) Pairwise dynamic time warping for event data. *Comp. Stat. Data Anal.* **69**: 255–268

[5] Artstein Z, Vitale R (1975) A Strong Law of Large Numbers for random compact sets. *Ann. Probab.* **3**(5): 879–882

[6] Aumann RJ (1965) Integrals of set-valued functions. *J. Math. Anal. Appl.* **12**: 1–12

[7] Ayala G, López-Díaz M (2009) The simplex dispersion ordering and its application to the evaluation of human corneal endothelia. *J. Mult. Anal.* **100**(7): 1447–1464

[8] Ayala G, López-Díaz MC, López-Díaz M, Martínez-Costa L (2012) Studying hypertension in ocular fundus images using Hausdorff dispersion ordering. *Math. Med. Biol.* **29**(2): 131–143

[9] Ayala G, Sebastian R, Díaz ME, Díaz E, Zoncu R, Toomre D (2006) Analysis of spatially and temporally overlapping events with application to image sequences. *IEEE Trans. Patt. Anal. Mach. Intel.* **28**(10): 1707–1712

[10] Báez-Sánchez AD, Moretti AC, Rojas-Medar MA (2012) On polygonal fuzzy sets and numbers. *Fuzzy Sets Syst.* **209**:54–65

[11] Ban AI, Coroianu L (2012) Nearest interval, triangular and trapezoidal approximation of a fuzzy number preserving ambiguity. *Int. J. Approx. Reas.* **53**: 805–836

[12] Ban A, Coroianu L, Grzegorzewski P (2011) Trapezoidal approximation and aggregation. *Fuzzy Sets Syst.* **177**: 45–59

[13] Bârsan T, Tiba D (2006) One hundred years since the introduction of the set distance by Dimitrie Pompeiu. In: *IFIP International Federation for Information Processing* (Ceragioli F, Dontchev A, Furuta H, Marti K, Pandolfi L, Eds), System Modeling and Optimization Vol. 199. Springer, Boston: pp. 35–39

[14] Bera UK, Maiti MK, Maiti M (2012) Inventory model with fuzzy lead-time and dynamic demand over finite time horizon using a multi-objective genetic algorithm. *Comput. Math. Appl.* **64**: 1822–1838

[15] Beresteanu A, Molinari F (2008). Asymptotic properties for a class of partially identified models. *Econometrica* **76**(4): 763–814

[16] Bertoluzza C, Corral N, Salas A (1995) On a new class of distances between fuzzy numbers. *Math & Soft Comput.* **2**: 71–84

[17] Bianco AM, Boente G, Rodrigues IM (2013) Robust tests in generalized linear models with missing responses. *Comp. Stat. Data Anal.* **65**: 80-97

[18] Blanco-Fernández A, Casals MR, Colubi A, Coppi R, Corral N, de la Rosa de Sáa S, D'Urso PP, Ferraro MB, García-Bárzana M, Gil MA, Giordani P, González-Rodríguez G, López MT, Lubiano MA, Montenegro M, Nakama T, Ramos-Guajardo AB, Sinova B, Trutschnig W (2013) Arithmetic and distance-based approach to the statistical analysis of imprecisely valued data. In: *Towards Advanced Data Analysis by Combining Soft Computing and Statistics* (Borgelt C, Gil MA, Sousa JMC, Verleysen M, Eds), Stud. in Fuzz. Soft Comp. Vol. 285. Springer, Heidelberg: pp. 1–18

[19] Blanco-Fernández A, Casals MR, Colubi A, Corral N, García-Bárzana M, Gil MA, González-Rodríguez G, López MT, Lubiano MA, Montenegro M, Ramos-Guajardo AB, De la Rosa de Sáa S, Sinova B (2013) Random fuzzy sets: a mathematical tool to develop statistical fuzzy data analysis. *Iran. J. Fuzzy Syst.* **10**(2): 1–28

[20] Blanco-Fernández A, Casals MR, Colubi A, Corral N, García-Bárzana M, Gil MA, González-Rodríguez G, López MT, Lubiano MA, Montenegro M, Ramos-Guajardo AB, De la Rosa de Sáa S, Sinova B (2014) A distance-based statistical analysis of fuzzy number-valued data. *Int. J. Approx. Reas.* (in press, doi:10.1016/j.ijar.2013.09.020).

[21] Blanco-Fernández A, Casals MR, Colubi A, Corral N, García-Bárzana M, Gil MA, González-Rodríguez G, López MT, Lubiano MA, Montenegro M, Ramos-Guajardo AB, De la Rosa de Sáa S, Sinova B (2014) Rejoinder on "A distance-based statistical analysis of fuzzy number-valued data". *Int. J. Approx. Reas.* (in press, doi:10.1016/j.ijar.2014.04.003).

[22] Butnariu D, Navara M, Vetterlein T (2005) Linear space of fuzzy vectors. In: *Fuzzy Logics and Related Structures* (Gottwald S, Hájek O, Höhle U, Klement EP, Eds). Universitatsdir. Johannes Kepler Univ., Linz: pp. 23–26

[23] Cadre B (2001) Convergent estimators for the L_1-median of a Banach valued random variable. *Statistics* **35**: 509–521

[24] Cappelli C, D'Urso P, Di Iorio F (2013) Change point analysis of imprecise time series. *Fuzzy Sets Syst.* **225**: 23–38

[25] Casals MR, Corral N, Gil MA, López MT, Lubiano MA, Montenegro M, Naval G, Salas A (2013) Bertoluzza *et al.*'s metric as a basis for analyzing fuzzy data. *Metron* **71**(3): 307–322

[26] Cascos I, Molchanov I (2007) Multivariate risks and depth-trimmed regions. *Finan. Stoch.* **11**(3): 373–397

[27] Castaing C, Valadier M (1977) *Convex Analysis and Measurable Multifunctions.* Lect. Notes Math. Vol. 580, Springer, Berlin

[28] Celmiņš A (1987) Least Squares model fitting to fuzzy vector data. *Fuzzy Sets Syst.* **22**: 245–269

[29] Chernozhukov V, Hong H, Tamer E (2007) Estimation and confidence regions for parameter sets in econometric models. *Econometrica* **75**(5): 1243–1284

[30] Chou JR (2012) A linguistic evaluation approach for universal design. *Inform. Sci.* **190**: 76–94

[31] Colubi A, Domínguez-Menchero JS, López-Díaz M, Körner R (2001) A method to derive strong laws of large numbers for random upper semicontinuous functions. *Stat. Prob. Lett.* **53**(3): 269–275

[32] Colubi A, Domínguez-Menchero JS, López-Díaz M, Ralescu DA (2001) On the formalization of fuzzy random variables. *Inform. Sci.* **133**: 3–6

[33] Colubi A, Domínguez-Menchero JS, López-Díaz M, Ralescu DA (2002) A $D_E[0,1]$ representation of random upper semicontinuous functions. *Proc. Amer. Math. Soc.* **130**: 3237–3242

[34] Colubi A, González-Rodríguez G, Gil MA, Trutschnig W (2011) Nonparametric criteria for supervised classification of fuzzy data. *Int. J. Approx. Reas.* **52**(9): 1272–1282

[35] Colubi A, López-Díaz M, Domínguez-Menchero JS, Gil MA (1999) A generalized Strong Law of Large Numbers. *Prob. Theor. Rel. Fields* **114**: 401–417

[36] Cooper WW, Park KS, Yu G (1999) IDEA and AR-IDEA: Models for dealing with imprecise data in DEA. *Manag. Sci.* **45**: 597–607

[37] Coppi R, D'Urso P (2003) Three-way fuzzy clustering models for LR fuzzy time trajectories. *Comp. Stat. Data Anal.* **43**: 149–177

[38] Coppi R, D'Urso P (2006) Fuzzy unsupervised classification of multivariate time trajectories with the Shannon entropy regularization. *Comp. Stat. Data Anal.* **50**(6): 1452–1477

[39] Coppi R, D'Urso P, Giordani P (2006) Component models for fuzzy data. *Psychometrika* **71**: 733–761

[40] Coppi R, D'Urso P, Giordani P (2012) Fuzzy and possibilistic clustering for fuzzy data. *Comp. Stat. Data Anal.* **56**(4): 915–927

[41] Coppi R, Gil MA, Kiers HAL (2006) The fuzzy approach to statistical analysis. *Comp. Stat. Data Anal.* **51**(1): 1–14

[42] Coroianu L (2012) Lipschitz functions and fuzzy number approximations. *Fuzzy Sets Syst.* **200**: 116–135

[43] Coroianu L, Gal SG, Bede B (2014) Approximation of fuzzy numbers by max-product Bernstein operators. *Fuzzy Sets Syst.* (in press, doi:10.1016/j.fss.2013.04.010)

[44] Corral-Blanco N, Zurbano-Fernández E, Blanco-Fernández A, García-Honrado I, Ramos-Guajardo AB (2013) Structure of the family educational environment: its influence on performance and differential performance. In: *PIRLS - TIMSS 2011 International Study on Progress in Reading Comprehension, Mathematics and Sciences IEA. Volume II. Spanish Report. Secondary Analysis*, Ministerio de Educación, Cultura y Deporte, Instituto Nacional de Evaluación Educativa: pp. 9–31 (http://www.mecd.gob.es/dctm/inee/internacional/pirlstimss2011eng-1.pdf?documentId=0901e72b81825be3)

[45] Cressie N, Laslett GM (1987) Random set theory and problems of modeling. *SIAM Rev.* **29**: 557–574

[46] Cuesta-Albertos JA, Fraiman R (2006) Impartial trimmed means for functional data. In: *Data Depth: Robust Multivariate Statistical Analysis, Computational Geometry and Applications* (Liu R, Serfling R, Souvaine D, Eds), Amer. Math. Soc. in DIMACS Series Vol. 72: pp. 121–145

[47] Cuesta-Albertos JA, Gordaliza A, Matrán C (1997) Trimmed k-means: an attempt to robustify quantizers. *Ann. Stat.* **25**: 553–576

[48] Cuesta-Albertos JA, Matrán C (1988) The strong law of large numbers for k-means and best possible nets of Banach valued random variables. *Prob. Theor. Rel. Fields* **78**(4): 523–534

[49] Cuesta-Albertos JA, Nieto-Reyes A (2008) The random Tukey depth. *Comp. Stat. Data Anal.* **52**(11): 4979–4988

[50] Cuevas A, Febrero M, Fraiman R (2007) Robust estimation and classification for functional data via projection-based depth notions. *Comput. Statist.* **22**: 481–496

[51] Cuevas A, Fraiman R (2009) On depth measures and dual statistics. A methodology for dealing with general data. *J. Multivar. Anal.* **100**: 753–766

[52] De Campos L, González A (1989) A subjective approach for ranking fuzzy numbers. *Fuzzy Sets Syst.* **29**: 145–153

[53] De la Rosa de Sáa S, Gil MA, González-Rodríguez G, López MT, Lubiano MA (2014) Fuzzy rating scale-based questionnaires and their statistical analysis. *IEEE Trans. Fuzzy Syst.* (in press, doi:10.1109/TFUZZ.2014.2307895)

[54] Diamond P, Kloeden P (1990) Metric spaces of fuzzy sets. *Fuzzy Sets Syst.* **35**: 241–249

[55] Diamond P, Kloeden P (1993) The parametrization of fuzzy sets by single-valued mappings. In: *Fuzzy Logic* (Lowen R, Roubens M, Eds), Theory and Decision Library Vol. 12, Springer (formerly Kluwer), Dordrecht: pp. 95–101

[56] Diamond P, Kloeden P (1994) *Metric Spaces of Fuzzy Sets: Theory and Applications.* World Sci., Singapore

[57] Diamond P, Körner R (1997) Extended fuzzy linear models and least squares estimates. *Comp. Math. Appl.* **33**(9): 15–32

[58] Díaz E, Sebastian R, Ayala G, Díaz ME, Zoncu R, Toomre D, Gasman S (2008) Measuring spatiotemporal dependencies in bivariate temporal random sets with applications to Cell Biology. *IEEE Trans. Patt. Anal. Mach. Intel.* **30**(9): 1659–1671

[59] Donoho DL, Huber PJ (1983) The notion of breakdown point. In: *A Festschrift for Erich L. Lehmann* (Bickel PJ, Doksum K, Hodges JL Jr, Eds). Wadsworth, Belmont: pp. 157–184

[60] Dubois D, Prade H (1980) Systems of linear fuzzy constraints. *Fuzzy Sets Syst.* **3**: 37–48

[61] D'Urso P (2003) Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data. *Comp. Stat. Data Anal.* **42**(1–2): 47–72

[62] D'Urso P (2007) Fuzzy clustering of fuzzy data. In: *Advances in Fuzzy Clustering and Its Applications* (de Oliveira V, Pedrycz W, Eds). J. Wiley & Sons, New York: pp. 155–192

[63] D'Urso P, Giordani P (2006) A weighted fuzzy c-means clustering model for fuzzy data. *Comp. Stat. Data Anal.* **50**(6): 1496–1523

[64] D'Urso P, Massari R, Santoro A (2010) A class of fuzzy clusterwise regression models. *Inform. Sci.* **180**: 4737–4762

[65] D'Urso P, Massari R, Santoro A (2011) Robust fuzzy regression analysis. *Inform. Sci.* **181**: 4154–4174

[66] D'Urso P, Santoro A (2006) Goodness of fit and variable selection in the fuzzy multiple linear regression. *Fuzzy Sets Syst.* **157**: 2627–2647

[67] D'Urso P, Santoro A (2006) Fuzzy clusterwise regression analysis with symmetrical fuzzy output variables. *Comp. Stat. Data Anal.* 51:287–313

[68] Febrero-Bande M, González-Manteiga W (2013) Generalized additive models for functional data. *Test* **22**(2): 278–292

[69] Féron R (1976) Ensembles aléatoires flous. *C.R. Acad. Sc. Paris A* **282**: 903–906

[70] Féron R (1976) Ensembles flous attachés à un ensemble aléatoire flou. *Publ. Econometr.* **9**: 51–66

[71] Féron R (1979) Sur les notions de distance et d'ecart dans une structure floue et leurs applications aux ensembles aléatoires flous. *C.R. Acad. Sc. Paris A* **289**: 35–38

[72] Ferraro MB, Colubi A, González-Rodríguez G, Coppi R (2011) A determination coefficient for a linear regression model with imprecise response. *Environmetrics* **22**: 516–529

[73] Ferraro MB, Giordani P (2013) On possibilistic clustering with repulsion constraints for imprecise data. *Inform. Sci.* **245**: 63–75

[74] Ferraty F, Sued M, Vieu P (2013) Mean estimation with data missing at random for functional covariables. *Statistics* **47**(4): 688–706

[75] Fraiman R, Muniz G (2001) Trimmed means for functional data. *Test.* **10**(2): 419–440

[76] Frappart F (2003) Catalogue des formes d'onde de l'altimètre Topex/Poséidon sur le bassin amazonien. *Technical Report*, CNES, Toulouse, France

[77] Fréchet M (1948) Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. L'Inst. H. Poincaré* **10**: 215–310

[78] Fréchet M (1950) Conferencias sobre los elementos aleatorios de naturaleza cualquiera. *Trab. Estadística* **1**: 157–181

[79] García D, Lubiano MA, Alonso MC (2001) Estimating the expected value of fuzzy random variables in the stratified random sampling from finite populations. *Inform. Sci.* **138**: 165–184

[80] García-García D, Santos-Rodríguez R (2011) Sphere packing for clustering sets of vectors in feature space. In: *Proc. 2011 IEEE Int. Conf. Acoust., Speech Signal Proc. (ICASSP)*, Prague: pp. 2092–2095

[81] García-García D, Santos-Rodríguez R, Parrado-Hernández E (2012) Sphere packing for clustering sets of vectors in feature space. In: *Proc. 2012 IEEE Int. Works. Mach.-Lear. Sign. Proc.*, Santander: pp. 1–6

[82] Ghosh PK, Kumar KV (1998) Support function representation of convex bodies, its application in geometric computing, and some related representations. *Comp. Vis. Im. Under.* **72**(3): 379–403

[83] Gil MA, Colubi A, Terán (2014) Random fuzzy sets: why, when, how. *BEIO* **30**(1): 5–29

[84] Gil MA, Lubiano MA, Montenegro M, López-García MT (2002) Least squares fitting of an affine function and strength of association for interval data. *Metrika* **56**: 97–111

[85] Gil MA, Montenegro M, González-Rodríguez G, Colubi A, Casals MR (2006) Bootstrap approach to the multi-sample test of means with imprecise data *Comp. Stat. Data Anal.* **51**(1): 148–162

[86] Giné E, Hann M, Zinn J (1983). Limit theorems for random sets: an application of Probability in Banach spaces. In: *Probability in Banach Spaces IV*, Lect. Notes Math. Vol. 990. Springer, Berlin: pp. 112-135

[87] Goetschel RJr, Voxman W (1986) Elementary fuzzy calculus. *Fuzzy Sets Syst.* **18**: 31–43

[88] González-Rodríguez G, Blanco A, Colubi A, Lubiano MA (2009) Estimation of a simple linear regression model for fuzzy random variables. *Fuzzy Sets Syst.* **160**: 357–370

[89] González-Rodríguez G, Colubi A, Gil MA (2006) A fuzzy representation of random variables: an operational tool in exploratory analysis and hypothesis testing. *Comput. Stat. Data Anal.* **51**(1): 163–176

[90] González-Rodríguez G, Colubi A, Gil MA (2012) Fuzzy data treated as functional data. A one-way ANOVA test approach. *Comp. Stat. Data Anal.* **56**(4): 943–955

[91] González-Rodríguez G, Montenegro M, Colubi A, Gil MA (2006) Bootstrap techniques and fuzzy random variables: Synergy in hypothesis testing with fuzzy data. *Fuzzy Sets Syst.* **157**: 2608–2613

[92] González-Rodríguez G, Sinova B, Colubi A, Van Aelst S (2014) On the trimmed means for generalized space-valued data and applications. Submitted

[93] Gower JC (1974) Algorithm AS 78: The mediancentre. *Appl. Statist.* **23**: 466–470

[94] Grzegorzewski P (2008) Trapezoidal approximations of fuzzy numbers preserving the expected interval - algorithms and properties. *Fuzzy Sets Syst.* **159**: 1354–1364

[95] Grzegorzewski P (2010) Algorithms for trapezoidal approximations of fuzzy numbers preserving the expected interval. In: *Foundations of Reasoning under Uncertainty.* Springer, Berlin: pp. 85–98

[96] Grzegorzewski P (2013) Fuzzy number approximation via shadowed sets. *Inform. Sci.* **225**: 35–46

[97] Grzegorzewski P, Pasternak-Winiarska K (2011) Trapezoidal approximations of fuzzy numbers with restrictions on the support and core. In: *Proc. 7th Conf. EUSFLAT-2011 and LFA-2011.* Paris, Atlantis Press: pp. 749–756

[98] Guillaume S, Charnomordic B, Loisel P (2013) Fuzzy partitions: a way to integrate expert knowledge into distance calculations. *Inform. Sci.* **245**: 76–95

[99] Hampel FR (1971) A general qualitative definition of robustness. *Ann. Math. Stat.* **42**(6): 1887–1896

[100] Hampel FR (1974) The influence curve and its role in robust estimation. *J. Amer. Stat. Assoc.* **69**: 383–393

[101] Hausdorff F (1914) *Grundzuege der Mengenlehre.* Viet, Leipzig

[102] Hesketh T, Pryor R, Hesketh B (1988) An application of a computerized fuzzy graphic rating scale to the psychological measurement of individual differences. *Int. J. Man–Mach. Stud.* **29**: 21–35

[103] Hess C (1999) The distribution of unbounded random sets and the multivalued strong law of large numbers in nonreflexive Banach spaces. *J. Convex Anal.* **6**(1): 163–182

[104] Hiai F, Umegaki H (1977) Integrals, conditional expectations, and martingales of multivalued junctions. *J. Multiv. Anal.* **7**: 149–182

[105] Hsu BM, Shu MH, Chen BS (2011) Evaluating lifetime performance for the Pareto model with censored and imprecise information. *J. Stat. Comput. Simul.* **81**(12): 1817–1833

[106] Huber PJ (1964) Robust estimation of a location parameter. *Ann. Math. Statist.* **35**(1): 73–101

[107] Huber PJ (1967) The behavior of maximum likelihood estimates under non-standard conditions. In: *Proc. 5th Berkeley Symp. Math. Stat. & Prob. 1*: pp. 221–233

[108] Huber PJ (1981) *Robust Statistics.* Wiley, New York

[109] Huttenlocher DP, Klanderman GA, Rucklidge WJ (1993) Comparing images using the Hausdorff distance. *IEEE Trans. Patt. Anal. Mach. Intel.* **15**(9): 850–863

[110] Jacques J, Preda C (2014) Model-based clustering for multivariate functional data. *Comp. Stat. Data Anal.* **71**: 92–106

[111] Kaval K, Molchanov I (2006) Link-save trading. *J. Math. Econ.* **42**: 710–728

[112] Kim JS (2011) *Kernel Methods for Classification with Irregularly Sampled and Contaminated Data.* PhD Thesis, University of Michigan (http://deepblue.lib.umich.edu/bitstream/handle/2027.42/89858/stannum_1.pdf?sequence=1)

[113] Kim JS, Scott CD (2011) On the robustness of kernel density M-estimators. In: *Proc. 28th Int. Conf. Mach. Lear.*, Bellevue, Washington: pp. 697–704

[114] Kim JS, Scott CD (2012) Robust kernel density estimation. *J. Mach. Learn. Res.* **13**: 2529–2565

[115] Kim YK (2002) Measurability for fuzzy valued functions. *Fuzzy Sets Syst.* **129**: 105–109

[116] Klement EP, Puri ML, Ralescu DA (1986) Limit theorems for fuzzy random variables. *Proc. R. Soc. Lond. A* **407**: 171–182

[117] Körner R (1997) On the variance of fuzzy random variables. *Fuzzy Sets Syst.* **92**: 83–93

[118] Körner R (2000) An asymptotic α-test for the expectation of random fuzzy variables. *J. Stat. Plann. Infer.* **83**: 331–346

[119] Körner R, Näther W (1998) Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates. *Inform. Sci.* **109**: 95–118

[120] Krätschmer V (2006) Integrals of random fuzzy sets. *Test* **15**(2): 433–469

[121] Kuhn HW (1973) A note on Fermat's problem. *Math. Prog.* **4**: 98–107

[122] Liang J, Navara M, Vetterlein T (2005) Different representations of fuzzy vectors. In: *Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (Sossai C, Chemello G, Eds), Lect. Notes Comp. Sci. Vol. 5590. Springer, Berlin: pp. 700–711

[123] Lin TH (2014) Model selection information criteria in latent class models with missing data and contingency question. *J. Stat. Comput. Simul.* **84**(1): 159–170

[124] Liu RY (1988) On a notion of simplicial depth. In: *Proc. Natl. Acad. Sci.*, USA: pp. 1732–1734

[125] López-Pintado S, Romo J (2009) On the concept of depth for functional data. *J. Amer. Statist. Assoc.* **104**(486): 718–734

[126] Lopuhaä HP, Rousseeuw PJ (1991) Breakdown points of affine equivariant estimators of multivariate location and covariance matrices. *Ann. Stat.* **19**: 229–248

[127] Lubiano MA, De la Rosa de Sáa S, Sinova B, Gil MA (2014) Empirical sensitivity analysis on the influence of the shape of fuzzy data on the estimation of some statistical measures. In: *Strengthening Links Between Data Analysis and Soft Computing* (Grzegorzewski P, Gagolewski M, Hyniewicz O, Gil MA, Eds). Springer, Heildelber: in press

[128] Luenberger DG (1997) *Optimization by Vector Space Methods.* Wiley-Interscience, New York

[129] Lyashenko NN (1982) Limit theorems for sums of independent, compact, random subsets of Euclidean space. *J. Soviet Math.* **20**(3): 2187–2196

[130] Matheron G (1975) *Random Sets and Integral Geometry.* J. Wiley & Sons, New York

[131] McClure DE, Vitale RA (1975) Polygonal approximation of plane convex bodies. *J. Math. Anal. Appl.* **51**: 326–358

[132] Milasevic P, Ducharme GR (1987) Uniqueness of the spatial median. *Ann. Stat.* **15**: 1332-1333

[133] Ming M (1993) On embedding problems of fuzzy number spaces: Part 4. *Fuzzy Sets Syst.* **58**: 185–193

[134] Minkowski H (1903) Volumen und Oberfläche. *Math. Ann.* **57**: 447–495

[135] Molchanov I (1993) Strong law of large numbers for unions of random closed sets. *Stoch. Proc. Appl.* **46**(2): 199–212

[136] Molchanov I (1993) *Limit Theorems for Unions of Random Closed Sets.* Lect. Notes Math. Vol. 1561. Springer, Berlin

[137] Molchanov I (1999) On Strong Laws of Large Numbers for random upper semicontinuous functions. *J. Math. Anal. Appl.* **235**: 349–355

[138] Molchanov I (2005) *Theory of Random Sets. Probability and its Applications.* Springer, Berlin

[139] Molchanov I (2012) *Lectures on Random Sets and their Applications in Economics and Finance.* CEMMAP-ESRC Master course (http://www.cemmap.ac.uk/resources/molchanov_mc/lectures_sets.pdf).

[140] Montenegro M, Casals MR, Lubiano MA, Gil MA (2001) Two-sample hypothesis tests of means of a fuzzy random variable. *Inform. Sci.* **133**: 89–100

[141] Montenegro M, Colubi A, Casals MR, Gil MA (2004) Asymptotic and Bootstrap techniques for testing the expected value of a fuzzy random variable. *Metrika* **59**: 31–49

[142] Nasibov EN (1989) To linear transformations with fuzzy arguments. *Trans. Acad. Sciences Azerbaijan, Ser. Phys-Tecn. & Math. Sci.* **6**: 164–169

[143] Nasibov EN, Baskan O, Mert A (2005) A learning algorithm for level sets weights in weighted level-based averaging method. *Fuzzy Optim. Dec. Making* **4**: 279–291

[144] Näther W (1997) Linear statistical inference for random fuzzy data. *Statistics* **29**: 221–240

[145] Näther W (2006) Regression with fuzzy random data. *Comp. Stat. Data Anal.* **51**: 915–927

[146] Nguyen HT (1978) A note on the extension principle for fuzzy sets. *J. Math. Anal. Appl.* **64**: 369–380

[147] Quang NV, Thuan NT (2012) Strong Laws of Large Numbers for adapted arrays of set-valued and fuzzy-valued random variables in Banach space. *Fuzzy Sets Syst.* **209**: 14–32

[148] Pedrycz W (1994) Why triangular membership functions? *Fuzzy Sets Syst.* **64**(1): 21–30

[149] Petit-Renaud S, Denoeux T (2004) Nonparametric regression analysis of uncertain and imprecise data using belief functions. *Int. J. Approx. Reas.* **35**: 1–28

[150] Phillis YA, Kouikoglou VS (2009) *Fuzzy Measurement of Sustainability.* Nova Sci. Pub., New York

[151] Pompeiu D (1905) Sur la continuité des fonctions de variables complexes (Thèse). *Ann. Fac. Sci. Toulouse* **7**: 264–315

[152] Porcel C, Tejeda-Lorente A, Martínez MA, Herrera-Viedma E (2012) A hybrid recommender system for the selective dissemination of research resources in a Technology Transfer Office. *Inform. Sci.* **184**: 1–19

[153] Proske FN, Puri ML (2003) A strong law of large numbers for generalized random sets from the viewpoint of the theory of empirical processes. *Proc. Amer. Math. Soc.* **131**(9): 2937–2944

[154] Puri ML, Ralescu DA (1983) Differentials of fuzzy functions. *J. Math. Anal. Appl.* **91**(2): 552–558

[155] Puri ML, Ralescu DA (1983) Strong Law of Large Numbers for Banach space valued random sets. *Ann. Probab.* **11**(1): 1–226

[156] Puri ML, Ralescu DA (1985) The concept of normality for fuzzy random variables. *Ann. Probab.* **11**: 1373–1379

[157] Puri ML, Ralescu DA (1986) Fuzzy random variables. *J. Math. Anal. Appl.* **114**: 409–422

[158] Rousseeuw PJ, Van Driessen K (2006) Computing LTS regression for large data sets. *Data Min. Know. Disc.* **12**(1): 29–45

[159] Rådström H (1952) An embedding theorem for spaces of convex sets. *Proc. Amer. Math. Soc.* **3**: 165–169

[160] Ramík J, Římánek J (1985) Inequality relation between fuzzy numbers and its use in fuzzy optimization. *Fuzzy Sets Syst.* **16**: 123–138

[161] Ramos-Guajardo AB, Lubiano MA (2012) K-Sample tests for equality of variances of random fuzzy sets. *Comp. Stat. Data Anal.* **56**: 956–966

[162] Ramsay JO, Silverman BW (2006) *Functional Data Analysis* (2nd ed). Springer, New York

[163] Robbins EE (1944) On the measure of a random set. *Ann. Math. Stat.* **14**: 70–74

[164] Robbins EE (1945) On the measure of a random set II. *Ann. Math. Stat.* **15**: 342–347

[165] Salski A (2007) Fuzzy clustering of fuzzy ecological data. *Ecol. Inform.* **2**: 262–269

[166] Schneider R (1971) On Steiner points of convex bodies. *Isr. J. Math.* **9**: 241–249

[167] Schneider R (1993) *Convex Bodies: The Brunn-Minkowski Theory.* Camb. Univ. Press, Cambridge

[168] Serrano-Guerrero J, Herrera-Viedma E, Olivas JA, Cerezo A, Romero FP (2011) A google wave-based fuzzy recommender system to disseminate information in University Digital Libraries 2.0. *Inform. Sci.* **181**: 1503–1516

[169] Setnes M, van Nauta Lemkea HR, Kaymak U (1998) Fuzzy arithmetic-based interpolative reasoning for nonlinear dynamic fuzzy systems. *Eng. Appl. Artif. Intel.* **11**: 781–789

[170] Sinha SK, Kaushal A, Xiao W (2014) Inference for longitudinal data with nonignorable nonmonotone missing responses. *Comp. Stat. Data Anal.* **72**: 77–91

[171] Sinova B, Casals MR, Colubi A, Gil MA (2010) The median of a random interval. In: *Combining Soft Computing and Statistical Methods in Data Analysis* (Borgelt C, González-Rodríguez G, Trutschnig W, Lubiano MA, Gil MA, Grzegorzewski P, Hryniewicz O, Eds), Adv. Intel. Soft Comp. Vol. 77. Springer, Berlin: pp. 575–583

[172] Sinova B, Casals MR, Gil MA (2013) Symmetry of the fuzzy-valued mean and median of symmetric random fuzzy numbers. In: *Prog. & Abst. 6th Int. Conf. ERCIM WG Comp. & Stat.*, London: p. 111

[173] Sinova B, Casals MR, Gil MA (2014) Central tendency for symmetric random fuzzy numbers. *Inform. Sci.* (in press, doi:10.1016/j.ins.2014.03.077)

[174] Sinova B, Colubi A, Gil MA (2009) Sensitivity analysis in estimating linear regression between interval data. In: *Abst. 2nd Workshop ERCIM WG Comp. & Stat. (ERCIM 09)*, Limassol: p. 36

[175] Sinova B, Colubi A, Gil MA, Sanz G (2010) Mediana de un conjunto aleatorio basada en una métrica tipo Hausdorff. In: *Res. XXXII Cong. Nac. Estad. Invest. Oper. (SEIO 2010)*, A Coruña: p. 112

[176] Sinova B, Colubi A, Gil MA, Van Aelst S (2011) The mids/ldev/rdev characterization of a fuzzy number. Some statistical applications. In: *Abst. 5th CSDA Int. Conf. CFE & 4rd Int. Conf. ERCIM WG Comp. & Stat.*, London: p. 111

[177] Sinova B, Colubi A, González-Rodríguez G, Van Aelst S (2013) The trimmed mean for random fuzzy numbers. In: *Book Abst. Int. Conf. Rob. Stat. ICORS11*, Valladolid: pp. 72–73

[178] Sinova B, Colubi A, González-Rodríguez G, Van Aelst S (2013) Comparación empírica de la robustez de varios conceptos de medias recortadas en el caso fuzzy. In: *Progr. XXXIII Cong. Nac. Est. Inv. Oper. SEIO'2013*, Castellón: p. 38

[179] Sinova B, De la Rosa de Sáa, Casals MR, Gil, MA, Salas AS (2014) The Mean Square Error of a random fuzzy vector based on the support function and the Steiner point. *Fuzzy Sets Syst.* submitted

[180] Sinova B, De la Rosa de Sáa S, Gil MA (2013) A generalized L^1-type metric between fuzzy numbers for an approach to central tendency of fuzzy data. *Inform. Sci.* **242**: 22–34

[181] Sinova B, Gil MA, Colubi A, González-Rodríguez G, Van Aelst S (2010) An approach to the median of a random fuzzy set. In: *Abst. 4th CSDA Int. Conf. CFE & 3rd Int. Conf. ERCIM WG Comp. & Stat.*, London: p. 97

[182] Sinova B, Gil MA, Colubi A, Van Aelst S (2012) The median of a random fuzzy number. The 1-norm distance approach. *Fuzzy Sets Syst.* **200**: 99–115

[183] Sinova B, Gil MA, González-Rodríguez G, Van Aelst S (2012) Empirical comparison of the robustness of two L1 type medians for random fuzzy numbers. In: *Abst. 6th CSDA Int. Conf. CFE & 5rd Int. Conf. ERCIM WG Comp. & Stat.*, Oviedo: p. 9

[184] Sinova B, Gil MA, González-Rodríguez G, Van Aelst S (2012) The computation of the sample d_θ-median for random intervals. In: *Book of Abst. 20th Int. Conf. Comp. Stat.*, Limassol: p. 14

[185] Sinova B, Gil MA, López MT, Van Aelst S (2014) A parameterized L^2 metric between fuzzy numbers and its parameter interpretation. *Fuzzy Sets Syst.* **245**: 101–115

[186] Sinova B, González-Rodríguez G, Van Aelst S (2013) An alternative approach to the median of a random interval using an L^2 metric. In: *Sinergies of Soft Computing and Statistics for Intelligent Data Analysis* (Kruse R, Berthold MR, Moewes C, Gil MA, Grzegorzewski P, Hryniewicz O, Eds), Adv. Intel. Syst. Comp. Vol. 190. Springer, Berlin: pp. 273–281

[187] Sinova B, Lubiano MA, Trutschnig W (2011) A function to calculate the median of a sample of fuzzy numbers in R. In: *Abst. 58th Session Int. Stat. Inst. (ISI 2011)*, Dublin: CPS065–06

[188] Sinova B, Van Aelst S (2013) Comparing the medians of a random interval defined by means of two different L^1 metrics. In: *Towards Advanced Data Analysis by Combining Soft Computing and Statistics* (Borgelt C, Gil MA, Sousa JMC, Verleysen M, Eds), Stud. Fuzz. Soft Comp. Vol. 285. Springer, Berlin: pp. pp. 75–86

[189] Sinova B, Van Aelst S (2014) A spatial interval-valued median for random intervals. To be submitted

[190] Steinwart I, Christmann A (2008) *Support Vector Machines.* Springer, New York

[191] Sugano N (2006) Fuzzy set theoretical approach to a chromatic relevant color on the natural color system. *Int. J. Innov. Comp., Inform. Contr.* **2**(1): 193–203

[192] Sugano N (2001) Color-naming system using fuzzy set theoretical approach. In: *2001 IEEE Int. Fuzzy Syst. Conf. Proc.*, Melbourne: pp. 81–84

[193] Sugano N (2011) Fuzzy set theoretical approach to the tone triangular system. *J. Computers* **6**(11): 2345–2356

[194] Sugano N, Komatsuzaki S, Ono H, Chiba Y (2009) Fuzzy set theoretical analysis of human membership values on the color triangle. *J. Computers* **4**(7): 593–600

[195] Terán P (2003) A strong law of large numbers for random upper semicontinuous functions under exchangeability conditions. *Stat. Prob. Lett.* **65**(3): 251–258

[196] Terán P (2006) On Borel measurability and large deviations for fuzzy random variables. *Fuzzy Sets Syst.* **157**: 2558–2568

[197] Terán P (2008) Strong law of large numbers for t-normed arithmetics. *Fuzzy Sets Syst.* **159**: 343–360

[198] Terán P (2008) On a uniform law of large numbers for random sets and subdifferentials of random functions. *Stat. Prob. Lett.* **78**(1): 42–49

[199] Terán P (2013) Algebraic, metric and probabilistic properties of convex combinations based on the t-normed extension principle: The Strong Law of Large Numbers. *Fuzzy Sets Syst.* **223**: 1–25

[200] Terán P, Molchanov I (2006) The law of large numbers in a metric space with a convex combination operation. *J. Theor. Prob.* **19**(4): 875–898

[201] Trutschnig W, González-Rodríguez G, Colubi A, Gil MA (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

[202] Trutschnig W, Lubiano MA (2014-last update) SAFD: Statistical Analysis of Fuzzy Data (http://cran.r-project.org/web/packages/SAFD/index.html)

[203] Trutschnig W, Lubiano MA, Lastra J (2013) SAFD - An R package for Statistical Analysis of Fuzzy Data. In: *Towards Advanced Data Analysis by Combining Soft Computing and Statistics* (Borgelt C, Gil MA, Sousa JMC, Verleysen M, Eds), Stud. Fuzz. Soft Comp. Vol. 285. Springer, Heidelberg: pp. 107–118

[204] Tukey JW (1948) Some elementary problems of importance to small sample practice. *Human Biol.* **20**: 205–214

[205] Tukey JW (1974) T6: Order Statistics. In: *Mimeographed notes for Statistics*, Princeton Univ.

[206] Tukey JW (1974) Address to International Congress of Mathematicians, Vancouver.

[207] Valvis E (2009) A new linear ordering of fuzzy numbers on subsets of $\mathcal{F}(\mathbb{R})$. *Fuzzy Optim. Decis. Making* **8**: 141–163

[208] Vandermeulen RA, Scott CD (2013) Consistency of robust kernel density estimators. *JMLR: Works. and Conf. Proc.* **30**: 1–24

[209] Vetterlein T, Navara M (2005) The Steiner centroid of fuzzy sets. In: *Proc. IFSA'2005*, Pekin: pp. 1256–1258

[210] Vetterlein T, Navara M (2006) Defuzzification using Steiner points. *Fuzzy Sets Syst.* **157**: 1455–1462

[211] Vitale RA (1985) L_p metrics for compact, convex sets. *J. Approx. Theor.* **45**: 280–287

[212] Vitale RA (1988) An alternate formulation of mean value for random geometric figures. *J. Microscopy* **151**: 197–204

[213] Wang YG, Lin X, Zhu M, Bai Z (2007) Robust estimation using the Huber function with a data-dependent tuning constant. *J. Comp. Graph. Stat.* **16**(2): 1–14

[214] Weil W (1973) Ein approximationssatz für konvexe Körper. *Manusc. Math.* **8**: 335–362

[215] Weiszfeld E (1937) Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tôhoku Math. J.* **43**: 355–386

[216] Weiszfeld E, Plastria F (2009) On the point for which the sum of the distances to n given points is minimum. *Ann. Oper. Res.* **167**: 7–41

[217] Wood ATA, Chan G (1994) Simulation of stationary Gaussian processes in $\mathcal{C}[0,1]$. *J. Comput. Graph. Stat.* **3**: 409–432

[218] Yager RR (1981) A procedure for ordering fuzzy subsets of the unit interval. *Inform. Sci.* **24**: 143–161

[219] Zadeh LA (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

[220] Zadeh LA (1996) Fuzzy logic = computing with words. *IEEE Trans. Fuzzy Syst.* **4**(2): 103–111

[221] Zadeh LA (2008) Is there a need for fuzzy logic? *Inform. Sci.* **178**: 2751–2779

[222] Zerafat Angiz LM, Emrouznejad A, Mustafa A (2012). Fuzzy data envelopment analysis: A discrete approach. *Expert Syst. Appl.* **36**(3): 2263–2269

[223] Zhao PY, Tang ML, Tang NS (2013) Robust estimation of distribution functions and quantiles with non-ignorable missing data. *Canad. J. Stat.* **41**(4): 575–595

[224] Zhang X, Chen K, Shou L, Chen G, Gao Y, Tan KL (2012) Efficient processing of probabilistic set-containment queries on uncertain set-valued data. *Inform. Sci.* **196**: 97–117

Subject index

Symbols index