Advances in Intelligent Systems and Computing 1433

Luis A. García-Escudero · Alfonso Gordaliza · Agustín Mayo · María Asunción Lubiano Gomez · Maria Angeles Gil · Przemyslaw Grzegorzewski · Olgierd Hryniewicz *Editors*

Building Bridges between Soft and Statistical Methodologies for Data Science

Springer

Advances in Intelligent Systems and Computing

Volume 1433

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within "Advances in Intelligent Systems and Computing" are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Indexed by DBLP, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST).

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

More information about this series at https://link.springer.com/bookseries/11156

Luis A. García-Escudero ·
Alfonso Gordaliza · Agustín Mayo ·
María Asunción Lubiano Gomez ·
Maria Angeles Gil · Przemyslaw Grzegorzewski ·
Olgierd Hryniewicz
Editors

Building Bridges between Soft and Statistical Methodologies for Data Science

Springer

Editors
Luis A. García-Escudero
Departamento de Estadística e I.O.
University of Valladolid
Valladolid, Spain

Alfonso Gordaliza
Departamento de Estadística e I.O.
University of Valladolid
Valladolid, Spain

Agustín Mayo
Departamento de Estadística e I.O.
University of Valladolid
Valladolid, Spain

María Asunción Lubiano Gomez
Departamento de Estadística, I.O. y D.M.
University of Oviedo
Oviedo, Spain

Maria Angeles Gil
Departamento de Estadística, I.O. y D.M.
University of Oviedo
Oviedo, Spain

Przemyslaw Grzegorzewski
Department of Computational Statistics
and Data Analysis
Warsaw University of Technology
Warsaw, Poland

Olgierd Hryniewicz
Department of Stochastic Methods, Systems
Research Institute
Polish Academy of Sciences
Warsaw, Poland

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-031-15508-6 ISBN 978-3-031-15509-3 (eBook)
https://doi.org/10.1007/978-3-031-15509-3

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume is a selection of peer-reviewed papers presented at the 10th International Conference on Soft Methods in Probability and Statistics, SMPS 2022, held in Valladolid (Spain) during September 14–16, 2022. The series of biannual international conferences on Soft Methods in Probability and Statistics (SMPS) started in Warsaw in 2002 and continued in Oviedo (2004), Bristol (2006), Toulouse (2008), Oviedo/Mieres (2010), Konstanz (2012), Warsaw (2014), Rome (2016) and Compiègne (2018), progressively consolidating itself in the international agenda of events in Probability, Statistics and Soft Computing.

The 10th edition of the SMPS conference was initially scheduled for 2020, however, the SARS-COV-2 pandemic obligated to postpone twice the event. Although it could have been held on the initially scheduled date through an online modality, organizers have decided to wait for having a face-to-face activity. For sure, the warm atmosphere associated with the close exchange of scientific discussions is one of the best assets of this conference.

SMPS 2022 has been organized by the Departamento de Estadística e Investigación Operativa and the Instituto de Investigación en Matemáticas (IMUVA) of the University of Valladolid, Spain. Valladolid has a long tradition in research in the field of Probability and Statistics and has successfully welcomed important national and international events in this area, such as the Spanish Conference on Statistics an Operations Research (SEIO 2007), the International Conference on Robust Statistics (ICORS 2011) or the recent conference New Bridges between Mathematics and Data Science (NBMDS 2021). We are convinced that this event will also be a success, thanks to the effort and good work of the entire local committee. We have done our best in order to guarantee the participants in the conference have a comfortable stay in Valladolid. Thus, this will meet their expectations both scientifically and personally and will allow them to enjoy our rich culture and history.

We are grateful to the Executive Board of this conference for allowing us to have the opportunity to hold the event in Valladolid. We are also grateful to the Program Committee members for their support on the scientific aspects of the conference, specifically to all the session organizers. We are grateful to the keynote speakers of the conference, Christophe Croux (EDHEC Business School, France), Francisco Herrera (University of Granada, Spain) and Frank Klawonn (Helmholtz Centre for Infection Research, Germany) for accepting our invitation. We would also like to thank the members of the Local Committee for their valuable contribution to the logistics of the event.

In recent years, we are experiencing a revolution around data analysis motivated by the emergence of new data collection sources due to great technical developments. A by-product of this enormous availability of data is the appearance of a myriad of new data typologies that need to be analyzed. The complexity of these data sets requires the development of new probabilistic and statistical approaches capable of dealing with the difficulties associated with them. Different communities of experts, with very different origins, including mathematicians, statisticians, engineers, computer scientists, biotechnologists, econometricians and psychologists try to respond to these challenges with tools based on their own background. All these varied origins and different approaches motivate the importance of building bridges between all these fields for Data Science.

Soft methods are designed either to address, among others, difficulties related to imprecise or other complex data, or to create/combine alternatives to deal with traditional data. Consequently, they will certainly play an important role in the near future to cope with these current challenges. Furthermore, contaminated data are ubiquitous, and therefore, robust methodologies are required when certain degree of inaccuracy and noise are present in data.

The volume contains more than fifty selected contributions that are clearly useful in establishing those important bridges between soft and statistical methodologies for Data Science. These contributions cover very different and relevant aspects such as imprecise probabilities, information theory, random sets and random fuzzy sets, belief functions, possibility theory, dependence modeling and copulas, clustering, depth concepts, dimensionality reduction of complex data and robustness. The editors are grateful to all the contributing authors, Program Committee members and additional referees who made it possible to put together this interesting volume and preparing such attractive program for the SMPS 2022 conference. The priceless effort and good work of M. Asunción Lubiano as Publication Chair deserve our sincere recognition. Without her excellent work, this publication would not have been possible.

We would like to thank Publishing Editor of the Springer Series of Advances in Intelligent Systems and Computing, Dr. Thomas Ditzinger, as well as Series Editor, Professor Janusz Kacprzyk, and Springer for their dedication to the production of this volume.

May 2022

Luis A. García-Escudero
Alfonso Gordaliza
Agustín Mayo
María Asunción Lubiano Gomez
Maria Angeles Gil
Przemyslaw Grzegorzewski
Olgierd Hryniewicz

Organization

General Chairs

Luis A. García-Escudero	University of Valladolid, Spain
Alfonso Gordaliza	University of Valladolid, Spain
Agustín Mayo-Iscar	University of Valladolid, Spain

Executive Board (Core SMPS Group)

María Ángeles Gil	University of Oviedo, Spain
Przemyslaw Grzegorzewski	Polish Academy of Sciences, Warsaw, Poland
Olgierg Hryniewicz	Polish Academy of Sciences, Warsaw, Poland

Program Committee

Christian Borgelt, Germany, Austria
Paula Brito, Portugal
Giulianella Coletti, Italy
Ana Colubi, Spain, Germany
Inés Couso, Spain
Bernard De Baets, Belgium
Gert De Cooman, Belgium
Thierry Denoeux, France
Sébastien Destercke, France
Francesco Dotto, Italy
Didier Dubois, France
Fabrizio Durante, Italy
Pierpaolo D'Urso, Italy
Erol Eğrioğlu, Turkey
Alessio Farcomeni, Italy
Maria Brigida Ferraro, Italy
Peter Filzmoser, Austria
Alfonso García-Pérez, Spain
Jonathan Garibaldi, UK
Paolo Giordani, Italy
Gil González-Rodríguez, Spain
Francesca Greselin, Italy
Salvatore Ingrassia, Italy
Janusz Kacprzyk, Poland
Frank Klawonn, Germany
Vladik Kreinovich, USA
Rudolf Kruse, Germany
Mark Last, Israel
Jonathan Lawry, UK
Jacek M. Łęski, Poland
Miguel López-Díaz, Spain
M. Asunción Lubiano, Spain
Enrique Miranda, Spain
Susana Montes, Spain

Witold Pedrycz, Canada
Henri Prade, France
Madan L. Puri, USA
Dan A. Ralescu, USA
Marco Riani, Italy
Beatriz Sinova, Spain
Martin Štěpnička, Czech Republic
Pedro Terán, Spain
Valentin Todorov, Austria
Wolfgang Trutschnig, Austria
Stefan Van Aelst, Belgium
Barbara Vantaggi, Italy
Maurizio Vichi, Italy
José Antonio Vilar, Spain
Christian Wagner, UK
Jin Hee Yoon, South Korea
Slawomir Zadrożny, Poland

Additional Referees

Sebastian Fuchs, Austria
Pedro Miranda, Spain
Ignacio Montes, Spain
Arthur Van Camp, UK

Organizing Local Committee

Pedro C. Álvarez-Esteban, Spain
Eustasio del Barrio, Spain
Miguel A. Fernández, Spain
José Luis García-Lapresta, Spain
Raquel González del Pozo, Spain
Paula Gordaliza, Spain
Hristo Inouzhe, Spain
Carlos Matrán, Spain
Julio Pastor (Conference Website Chair), Spain
Azucena Prieto (Administrative Staff), Spain

Publication Chair

M. Asunción Lubiano University of Oviedo, Spain

Contents

Multi-dimensional Maximal Coherent Subsets Made Easy: Illustration on an Estimation Problem

Loïc Adam and Sébastien Destercke(✉)

UMR CNRS 7253 Heudiasyc, Sorbonne Université, Université de Technologie de Compiègne, CS 60319, 60203 Compiègne Cedex, France
{loic.adam,sebastien.destercke}@hds.utc.fr

Abstract. Fusing uncertain pieces of information to obtain a synthetic estimation when those are inconsistent is a difficult task. A particularly appealing solution to solve such conflict or inconsistency is to look at maximal coherent subsets of sources (MCS), and to concentrate on those. Yet, enumerating MCS is a difficult combinatorial task in general, making the use of MCS limited in practice. In this paper, we are interested in the case where the pieces of information are multi-dimensional sets or polytopes. While the problem remains difficult for general polytopes, we show that it can be solved more efficiently for hyperrectangles. We then illustrate how such an approach could be used to estimate linear models in the presence of outliers or in the presence of misspecified model.

1 Introduction

This paper deals with the problem of fusing multiple pieces of information (Dubois et al. 2016). In this problem, handling conflict between contradicting sources of information is one of the most difficult tasks. This is often a mandatory task, including in situations where we want a synthetic estimation from all the sources. Moreover, analyzing the reasons of the contradiction and trying to explain its appearance can be of equal importance, as it can give important insights about the situation.

Dubois et al. (1999) reviewed different methods for aggregating conflict, some requiring additional data like the reliability of the different sources. Yet, having additional data is sometimes difficult or even impossible. A quite appealing way for dealing with contradiction, requiring no additional information, is based on maximal coherent subsets (MCS), which are groups of consistent sources that are as big as possible. MCS have been used in the past both in logic (Manor and Rescher 1970) and in numerical settings (Destercke et al. 2008). Here, we illustrate their application to estimation problems.

Detecting and enumerating MCS are NP-hard problems, with intervals being a well-known exception (Dubois et al. 2000). In this paper, we show that we

L. A. García-Escudero et al. (Eds.): SMPS 2022, AISC 1433, pp. 1–8, 2023.
https://doi.org/10.1007/978-3-031-15509-3_1

can extend this exception to hyperrectangles, which can in turn be used as approximations of polytopes.

In Sect. 2 we further introduce maximal coherent subset and explain the difficulty of enumerating them. We then show in Sect. 3 that it is easy to list the MCSs of a set of axis-aligned hyperrectangles. Lastly, Sect. 4 illustrates our approach on linear regression problems.

2 Maximal Coherent Subsets

As mentioned previously, maximal coherent subsets are in theory a nice solution to manage conflict between information sources. Moreover, they can be used with different structures of information, like polytopes. However, we will show that listing the different MCSs is usually a difficult combinatorial problem.

General Definition: Let us suppose we have a set $\mathcal{S} = \{S_1, ..., S_N\}$ of sources of information providing a subset $S_i \subseteq \mathcal{X}$ of some space $\mathcal{X}$ of information, for which intersection $\cap$ is well-defined. A maximal coherent subset $c \subseteq \{1, \ldots, N\}$ is a list of source indices such that $\cap_{i \in c} S_i \neq \emptyset$, and for any $j \notin c$, $\cap_{i \in c} S_i \cap S_j = \emptyset$, i.e., the subset c of sources is consistent and is maximal with this property.

Example 1. Let us suppose we have a set of sources $\mathcal{S} = \{S_1, ..., S_4\}$ as shown on Fig. 1. As we can see, $\{1, 2\}$ is a coherent subset, but not a maximal coherent subset, as it is possible to add S_3 and have a non-empty intersection. $\{1, 2, 3\}$ is a maximal coherent subset, as S_4 is contradicting S_3 (empty intersection).

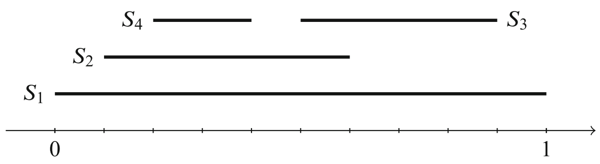

Fig. 1. Visualization of not fully consistent sources $S_i \in [0, ..., 1]$

MCSs of Polytopes: When considering sources of information in the d-dimensional Euclidean space R^d, polytopes are a quite versatile tool to model set-valued information. Such shapes can either be defined through their vertices or extreme points (V-representation) or through a system of linear constraints defining intersections of half-planes (H-representation).

While finding whether two polytopes given by their V-representation intersect is a NP-hard problem (Tiwary 2008), the same problem can be solved easily in H-representation through linear programming. Switching between representations is a NP-complete problem,[1] with some efficient algorithms for specific H-polytopes (Khachiyan et al. 2009).

Finding a single MCS in H-representation is thus easy: we add polytopes one by one, and we have a MCS when it is not possible to add another polytope without having an empty intersection. Checking if a set is a MCS is also easy: we check that the intersection of the corresponding H-polytopes is not empty and maximal. However, listing all the MCSs of a set of polytopes $\mathscr{P} = \{P_1, ..., P_N\}$ requires in the worst case to consider all the subsets of $\mathscr{P}$, thus at most 2^N sets for which we need to check if the intersection is not empty. When the number of polytopes is important, it becomes impossible to list all the MCSs. In Sect. 3, we propose an efficient algorithm to list all the MCSs through an approximation of the polytopes with minimum bounding axis-aligned hyperrectangles.

3 Enumerating the MCSs of Axis-Aligned Hyperrectangles

Enumerating all the MCS of a set of polytopes is a difficult problem in the general case. However, polynomial algorithms (Dubois et al. 2000) exist in the case of intervals. We will show in this section that such results can also be used in the case where we consider a set $\mathscr{H}$ of axis-aligned hyperrectangles, in order to efficiently determine the set of MCS $\mathscr{C}_{\mathscr{H}}$.

We denote by $I^d_{H_i} \in \mathbb{R}$ the projection of H_i onto the dth dimension of the space $\mathbb{R}^D$. We have an important equivalence between the intersection of hyperrectangles and the intersection of their projections:

Proposition 1. *Given a set $\mathscr{H} = \{H_1, ..., H_N\}$ of axis-aligned hyperrectangles in the space $\mathbb{R}^D$, and their projections $I^d_{H_i} \in \mathbb{R}$ onto the different dimensions $d \in \{1, ..., D\}$, we have:*

$$\bigcap_{H_i \in \mathscr{H}} H_i \neq \emptyset \Longleftrightarrow \bigcap_{\mathscr{H}} I^d_{H_i} \neq \emptyset \; \forall d \in \{1, ..., D\}. \tag{1}$$

Proof. To see the equivalence, it is sufficient to observe that $\times_{d=1}^{D} \cap_{i=1}^{N} I^d_{H_i} = \bigcap_{H_i \in \mathscr{H}} H_i$. This means in particular that any point $x \in \mathbb{R}^D$ such that its projection $x^d \in \cap_{i=1}^{N} I^d_{H_i}$ for all $d \in \{1, \ldots, D\}$ will also be in $\bigcap_{H_i \in \mathscr{H}} H_i$. Note that this is only true for axis-aligned hyperrectangles. □

The following corollary, which is merely the negation of Proposition 1, will be useful in further demonstrations.

Corollary 1

$$\bigcap_{\mathscr{H}} H_i = \emptyset \Longleftrightarrow \exists d \in \{1, ..., D\} \; s.t. \; \bigcap_{\mathscr{H}} I^d_{H_i} = \emptyset. \tag{2}$$

[1] Otherwise the two problems would have the same complexity.

In the next proof, we show that the MCS of axis-aligned hyperrectangles can be found exactly by combining the MCS of their projections, which we recall can be found in polynomial time.

Proposition 2. *Given a set $\mathscr{H} = \{H_1, ..., H_N\}$ of axis-aligned hyperrectangles in the space $\mathbb{R}^D$, its set of MCSs $\mathscr{C}_{\mathscr{H}}$ and the sets of MCSs $\mathscr{C}_d$ on their projection on the d-dimension, we have:*

$$\mathscr{C}_{\mathscr{H}} = \left\{\cap_{d=1}^{D} c_d \mid c_d \in \mathscr{C}_d \ \forall d \in \{1, ..., D\}, \cap_{d=1}^{D} c_d \neq \emptyset\right\}. \tag{3}$$

Proof. We proceed by showing a double inclusion for a given MCS $c_h \in \mathscr{C}_{\mathscr{H}}$:

- Let us first show that there exists $c_i \in \left\{\cap_{d=1}^{D} c_d \mid c_d \in \mathscr{C}_d\right\}$ s.t. $c_h \subseteq c_i$, i.e. c_i is an outer approximation of c_h. Because $\cap_{i \in c_h} H_i \neq \emptyset$ (it being a MCS), Proposition 1 tells us that for any dimension d, $\cap_{i \in c_h} I_{H_i}^d \neq \emptyset$, meaning that there will be a MCS $c_d \in \mathscr{C}_d$ such that $c_h \subseteq c_d$. Since this is true for all $d \in \{1, \ldots, D\}$, this means that $c_h \subseteq \cap_{d=1}^{D} c_d$ for some collection of c_d, showing the inclusion.
- To show the other inclusion, we consider a set $c_d, d \in \{1, \ldots, D\}$ of MCSs on dimension d which outer-approximate $c_h \subseteq c_d$, that we know exists from the first part of the proof. We will then demonstrate that $j \notin c_h$ implies $j \notin \cap_{d=1}^{D} c_d$, therefore $\cap_{d=1}^{D} c_d \subseteq c_h$. To see this, simply consider the set of hyperrectangles $H_k, k \in c_h \cup \{j\}$, then by Corollary 1, there will be a dimension d such that $\cap_{k \in c_h \cup \{j\}} I_{H_k}^d = \emptyset$, yet $\cap_{k \in c_h} I_{H_k}^d \neq \emptyset$ (c_h being a MCS). This shows that $j \notin c_h$ implies $j \notin \cap_{d=1}^{D} c_d$. $\square$

Proposition 2 provides us with an easy approach to get MCS: we start by projecting the hyperrectangles onto the different dimensions, in order to obtain intervals. Then we enumerate the MCS on each dimension, which is polynomial. Lastly we determine $\mathscr{C}_{\mathscr{H}}$ as the set of common sources among all the enumerations, i.e., the conjunctive combination (i.e., intersection) of the different sources such that for each element c_h of $\mathscr{C}_{\mathscr{H}}$. By Proposition 2 and Eq. (3), this gives us exactly the set of MCS.

4 Illustration on Linear Estimation

In this section, we illustrate our method on small imprecise linear regression problems with only two dimensions. Note that the purpose of this section is purely illustrative, so as to show the potential usefulness of our result when performing estimation from a logical, set-theoretic standpoint.

Given a data set $\{y_i, x_i\}_{i=1}^{N}$ with a single input and a single output, a linear model assumes that the relationship between the response variable y_i and the input variable x_i is linear:

$$y_i = \beta_0 + \beta_1 x_i \tag{4}$$

Usually, the y_i are considered to be observed with a (normal) noise ϵ, and a statistical regression is performed. In our case, we will adopt a more logical, version space point of view (Mitchell 1982): we assume that data is set-valued, and will consider the linear models consistent with it.

4.1 Estimating Possible Linear Models with MCS and Rectangles

In our setting, we assume that we observe imprecise data points $R_i = ([\underline{y_i}, \overline{y_i}], [\underline{x_i}, \overline{x_i}])$. Given the Cartesian equation of a line $\mathrm{L} = \{(x, y) \mid ax + by = c\}$ and two imprecise points R_i and R_j, finding all the lines that intersect both rectangles is formalized as:

$$\mathscr{L}_{ij} = \{(a, b) \; |(L \cap H_i \neq \emptyset) \wedge (L \cap H_j \neq \emptyset)\} \tag{5}$$

n points can be intersected by a single line if and only if all the corresponding $\mathscr{L}_{ij}$ ($\binom{n}{2}$) in total) have a non-empty intersection, i.e., their indices belong to a single coherent subset. It is maximal if no other points can be intersected by the same line. Listing all the MCSs is hard, as mentioned in Sect. 2, but our results tell us that if we approximate the different $\mathscr{L}_{ij}$ with minimum bounding axis-aligned rectangles, minimal outer approximations of the $\mathscr{L}_{ij}$ with axis-aligned rectangles, then enumerating their MCSs is easily done as shown in Sect. 3.

Determining the minimum axis-aligned bounding rectangle $H_{ij} = [\underline{y}_i, \overline{y}_i] \times [\underline{x}_i, \overline{x}_i]$ of $\mathscr{L}_{ij}$, i.e., the minimal volume that is fully enclosing $\mathscr{L}_{ij}$, is equivalent to finding the minimum and maximum values of the parameters a and b. It can be done quite easily, as maximizing (respectively minimizing) a is equivalent to minimizing (respectively maximizing) b.

4.2 Application

We first start with the case where the model is indeed linear, but where one data point is an outlier (box 2), as pictured on Fig. 2. As we can see, the statistical regression model, due to the outlier, does not capture the true model. In contrast, Fig. 3 shows the 5 different MCS (and their intersection) obtained using axis-aligned rectangle approximations. We can see that the biggest MCS is c_3, and it includes the true parameters, while the other MCSs are smaller and point out possible model outliers (i.e., observation 2 is in all the remaining MCSs, but not in the biggest one). Note that such an approach is quite different from standard imprecise regressions, that still uses least-square inspired approaches (Ferraro et al. 2010).

The second illustration considers the case where the observations are not faulty (the model goes through all observations), but where the model assumption is wrong, as we have a piecewise linear regression as shown on Fig. 4, but not a linear one. We have two partitions over x: $[0, 0.5]$ and $[0.5, 1]$. The two ground truths are very different, and a single statistical linear regression (in dotted blue) over the whole domain of x fits poorly to the data.

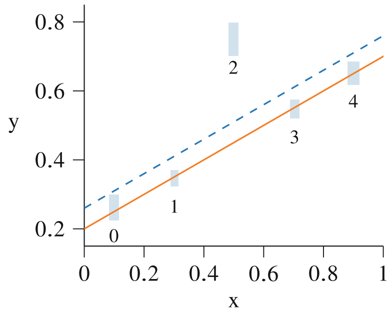

Fig. 2. Example of an imprecise linear regression problem. The dotted line corresponds to the linear regression including the outlier. The continuous line corresponds to the ground truth

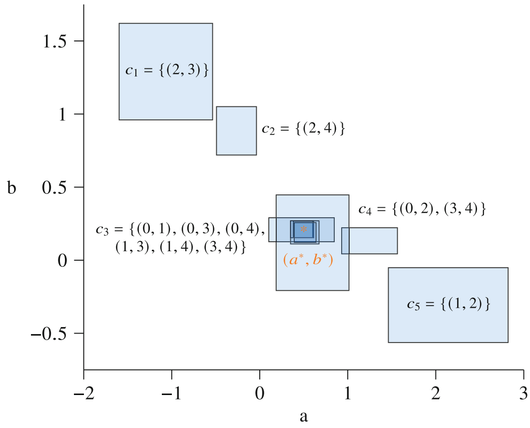

Fig. 3. Approximation with rectangles of the different $\mathscr{L}_{ij}$ from the example shown on Fig. 2

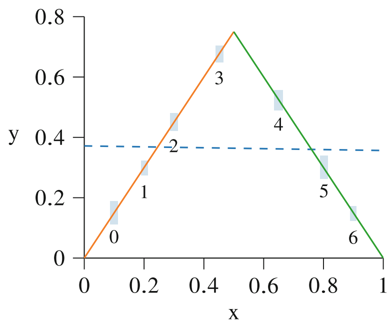

Fig. 4. Example of an imprecise piecewise linear regression problem. The dotted line corresponds to the full linear regression. The continuous lines correspond to the piecewise ground truths

Figure 5 shows what happens when we consider a MCS approach in this case. Instead of having one large MCS and quite smaller ones, we have now two quite large MCSs c_1 and c_2, along with other isolated ones. Those two distinct MCSs indicates that the error is likely in the model, that may only be locally valid. It also shows that it is not clear whether observation 3 belongs to one line or the other, being in both MCS. This ambiguity also cause c_1 to not include the true parameter (a_2^*, b_2^*).

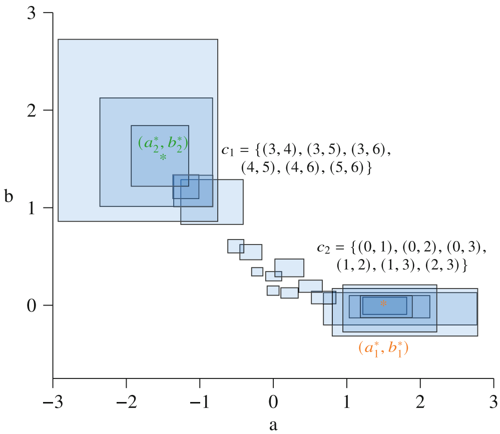

Fig. 5. Approximation with rectangles of the different $\mathscr{L}_{ij}$ from the example shown on Fig. 4

5 Conclusion

MCS is a theoretically very interesting notion for handling conflicting observations or pieces of information, but hard to enumerate in practice. We showed that such an enumeration was easier with hyperrectangles, as we can use known polynomial algorithms on their interval-valued projection to perform it. We illustrated their possible use on estimation problems, showing that their different behaviors could provide useful information (erroneous observations vs erroneous models).

In the future, we plan to perform some comparisons of such estimation approaches with Bayesian approaches in case of model misspecification, similarly to previous works intending to solve inverse problems (Shinde et al. 2021). Another case where our results may be useful is in the repair of inconsistent preference information (Adam and Destercke 2021).

Acknowledgements. This work was done within the PreServe Project, funded by grant ANR-18-CE23-0008 of the national research agency (ANR).

References

Adam, L., Destercke, S.: Possibilistic preference elicitation by minimax regret. In: Uncertainty in Artificial Intelligence, pp. 718–727. PMLR (2021)

Destercke, S., Dubois, D., Chojnacki, E.: Possibilistic information fusion using maximal coherent subsets. IEEE Trans. Fuzzy Syst. **17**(1), 79–92 (2008)

Dubois, D., Fargier, H., Prade, H.: Multiple-sources informations fusion-a practical inconsistency-tolerant approach. In: 8th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU 2000, Madrid, Spain) (2000). https://hal.archives-ouvertes.fr/hal-03405306

Dubois, D., Liu, W., Ma, J., Prade, H.: The basic principles of uncertain information fusion. An organised review of merging rules in different representation frameworks. Inf. Fusion **32**, 12–39 (2016)

Dubois, D., Prade, H., Yager, R.: Merging fuzzy information. In: Bezdek, J.C., Dubois, D., Prade, H. (eds.) Fuzzy Sets in Approximate Reasoning and Information Systems. The Handbooks of Fuzzy Sets Series, vol. 5, pp. 335–401. Springer, Boston (1999). https://doi.org/10.1007/978-1-4615-5243-7_7

Ferraro, M.B., Coppi, R., Rodríguez, G.G., Colubi, A.: A linear regression model for imprecise response. Int. J. Approx. Reason. **51**(7), 759–770 (2010)

Khachiyan, L., Boros, E., Borys, K., Gurvich, V., Elbassioni, K.: Generating all vertices of a polyhedron is hard. In: Twentieth Anniversary Volume, pp. 1–17. Springer, New York (2009). https://doi.org/10.1007/978-0-387-87363-3_17

Manor, R., Rescher, N.: On inference from inconsistent premises. Theor. Decis. **1**, 179–219 (1970)

Mitchell, T.M.: Generalization as search. Artif. Intell. **18**(2), 203–226 (1982)

Shinde, K., Feissel, P., Destercke, S.: Dealing with inconsistent measurements in inverse problems: set-based approach. Int. J. Uncertain. Quantif. **11**(3), 59–84 (2021)

Tiwary, H.R.: On the hardness of computing intersection, union and Minkowski sum of polytopes. Discrete Comput. Geom. **40**(3), 469–479 (2008)

On Convergence in Distribution of Fuzzy Random Variables

Miriam Alonso de la Fuente(✉) and Pedro Terán

Escuela Politécnica de Ingeniería, Departamento de Estadística e I.O. y D.M., Universidad de Oviedo, Oviedo, Spain
{alonsofmiriam,teranpedro}@uniovi.es

Abstract. We study whether convergence in distribution of fuzzy random variables, defined as the weak convergence of their probability distributions, is consistent with the additional structure of spaces of fuzzy sets. Positive results are obtained which reinforce the viability of that definition.

1 Introduction

Convergence in distribution is one of the most useful notions of convergence for random variables, notably because of its role in the central limit theorem. It is typically defined as follows: $\xi_n \to \xi$ in distribution when $F_{\xi_n}(x) \to F_{\xi}(x)$ (where F_{ξ_n} and F_{ξ} are the respective cumulative distribution functions) for each point of continuity x of F_{ξ}.

Throughout the years, a number of proposals trying to extend the notion of a cumulative distribution function to fuzzy random variables have been made. Without judging their usefulness for specific problems, it is fair to say (Terán 2012) that they fail to have the theoretical properties that make the cumulative distribution function important in the case of random variables and random vectors. Specifically, they do not determine the probability distribution of the fuzzy random variable. Thus they are not useful to study convergence in distribution of fuzzy random variables.

Since a fuzzy random variable can be equivalently described (Krätschmer 2001) as a random element of a metric space of fuzzy sets (endowed with any of the d_p-metrics), it is possible to study probability distributions of fuzzy random variables using the general theory of probability distributions in metric spaces (e.g., Billingsley 1968). This approach was taken by the authors in recent papers (Alonso de la Fuente and Terán 2022a,b). In particular, it provides a way to define convergence in distribution as being tantamount to *weak convergence* of the probability distributions (since the Helly–Bray theorem and its converse establish the equivalence of those two notions for ordinary random variables).

The theoretical properties of weak convergence are well understood (see Billingsley 1968). But spaces of fuzzy sets have more structure than a generic

L. A. García-Escudero et al. (Eds.): SMPS 2022, AISC 1433, pp. 9–15, 2023.
https://doi.org/10.1007/978-3-031-15509-3_2

metric space, which raises the question whether defining convergence in distribution via weak convergence works well with that additional structure.

In this contribution, we show that convergence in distribution of fuzzy random variables with convex values can be studied using the support function embedding into an L^p-type space (i.e., convergence in distribution of the fuzzy random variables and of their support functions are equivalent). We also show that this type of convergence is consistent with some known structures in the space of fuzzy sets. A sequence of trapezoidal fuzzy random variables converges in distribution if and only if the vertices of the trapezoid converge jointly as a 4-dimensional random vector. A sequence of random vectors converges in distribution if and only if their indicator functions converge as fuzzy random variables. Finally, we show a consistency result between convergence and the sum and product by scalars which parallels the corresponding property of ordinary random variables.

2 Preliminaries

Let $\mathscr{F}_c(\mathbb{R}^d)$ be the space of fuzzy sets $U : \mathbb{R}^d \to [0,1]$ whose α-cuts U_α are non-empty compact convex subsets of $\mathbb{R}^d$. Every fuzzy set $U \in \mathscr{F}_c(\mathbb{R}^d)$ is uniquely determined by its support function

$$\begin{aligned} s_U : [0,1] \times \mathbb{S}^{d-1} &\to \mathbb{R} \\ (r,\alpha) &\mapsto s_U(r,\alpha) = \sup_{x \in U_\alpha} \langle r, x \rangle \end{aligned}$$

where $\mathbb{S}^{d-1}$ denotes the unit sphere in $\mathbb{R}^d$.

For each $p \in [1,\infty)$, the metric d_p in $\mathscr{F}_c(\mathbb{R}^d)$, introduced by Klement et al. (1986) and Puri and Ralescu (1986), is defined by

$$d_p(U,V) = \left[\int_{[0,1]} (d_H(U_\alpha, V_\alpha))^p \, d\alpha \right]^{1/p}.$$

The metric ρ_p is defined by

$$\rho_p(U,V) = \left[\int_{[0,1]} \int_{S^{d-1}} |s_U(r,\alpha) - s_V(r,\alpha)|^p \, dr \, d\alpha \right]^{1/p}.$$

Denote by $\mathscr{K}_c(\mathbb{R}^d)$ the space of non-empty compact convex subsets of $\mathbb{R}^d$. Given a probability space $(\Omega, \mathscr{A}, P)$, a mapping $X : \Omega \to \mathscr{K}_c(\mathbb{R}^d)$ is called *random set* (also a *random compact convex set* in the literature) if X is measurable with respect to the Borel σ-algebra $\mathscr{B}_{\mathscr{K}_c(\mathbb{R}^d)}$ generated by the topology of the Hausdorff metric.

Definition 1. A mapping $X : (\Omega, \mathscr{A}, P) \to \mathscr{F}_c(\mathbb{R}^p)$ is called a *fuzzy random variable* if $X_\alpha : \omega \mapsto X(\omega)_\alpha$ is a random compact set for each $\alpha \in [0,1]$.

Denote by σ_L the smallest σ-algebra that makes the mappings $U \in \mathscr{F}_c(\mathbb{R}^d) \mapsto U_\alpha \in \mathscr{K}_c(\mathbb{R}^d)$ measurable. Thus a fuzzy random variable is the same thing as a $(\mathscr{A}, \sigma_L)$-measurable mapping. A sequence of probability measures $\{P_n\}_n$ on σ_L is said to *converge weakly in* d_p to a probability measure P if

$$\int f dP_n \to \int f dP$$

for every $f : \mathscr{F}_c(\mathbb{R}^d) \to \mathbb{R}$ which is d_p-continuous and bounded. A sequence $\{X_n\}_n$ of fuzzy random variables converges *weakly* or *in distribution in* d_p to a fuzzy random variable X if their distributions P_{X_n} converge weakly to P_X, namely

$$E[f(X_n)] \to E[f(X)]$$

for each bounded d_p-continuous function $f : \mathscr{F}_c(\mathbb{R}^d) \to \mathbb{R}$.

The Lebesgue measure in $[0, 1]$ will be denoted by ℓ. The following results will be used in the sequel.

Lemma 1 (Billingsley 1968, Theorem 2.1). *Let* $\mathbb{E}$ *be a metric space,* P *a probability measure and* $\{P_n\}_n$ *a sequence of probabilities in* $(\mathbb{E}, \mathscr{B}_\mathbb{E})$*. Then* $P_n \to P$ *weakly if and only if for every open set* G *we have* $\liminf_{n\to\infty} P_n(G) \ge P(G)$*.*

Lemma 2 (Alonso de la Fuente and Terán 2022a, Theorem 3.5). *Let* $p \in [1, \infty)$*. Let* P_n, P *be probability measures on* σ_L*, such that* $P_n \to P$ *weakly. Then there exist fuzzy random variables* $X_n, X : ([0, 1], \mathscr{B}_{[0,1]}, \mathbb{P}) \to (\mathscr{F}_c(\mathbb{R}^d), d_p)$*, such that*

(a) The distributions of X_n *and* X *are* P_n *and* P*, respectively.*
(b) $X_n(t) \to X(t)$ *in* d_p *for every* $t \in [0, 1]$*.*

Lemma 3 (Alonso de la Fuente and Terán 2022a, Theorem 5.1). *Let* X_n *and* X *be fuzzy random variables such that* $X_n \to X$ *in distribution in* d_p*. If* $f : \mathscr{F}_c(\mathbb{R}^d) \to \mathscr{F}_c(\mathbb{R}^d)$ *is a* P_X*-almost surely continuous function, then* $f(X_n) \to f(X)$ *weakly in* d_p*.*

Lemma 4 (Parthasarathy 1967, Corollary 3.3, p. 22). *If* $\mathbb{E}$ *is a Borel subset of a complete separable metric space* X *and* φ *is a one-one measurable map of* $\mathbb{E}$ *into a separable metric space* Y*, then* $\varphi(\mathbb{E})$ *is a Borel subset of* Y*,* $\mathbb{E}$ *and* $\varphi(\mathbb{E})$ *are isomorphic as measurable spaces and* φ *is an isomorphism.*

Recall that a trapezoidal fuzzy number $Tra(a, b, c, d)$ has the following expression:

$$U(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \le x < b \\ 1 & \text{if } b \le x \le c \\ \frac{d-x}{d-c} & \text{if } c < x \le d \\ 0 & \text{if } x > d \end{cases}$$

We will denote the space of trapezoidal fuzzy numbers by $\mathscr{F}_c^{tra}(\mathbb{R})$.

3 Main Results

Our first result states that d_p-convergence in distribution of fuzzy random variables is equivalent with the convergence obtained by embedding them into an L^p-type space. Note that this is not an immediate consequence of the embedding.

Theorem 1. *Let $p \in [1, \infty)$. Let X_n, X be fuzzy random variables. Then the following conditions are equivalent.*

1. *$X_n \to X$ in distribution in $(\mathscr{F}_c(\mathbb{R}^d), d_p)$.*
2. *$s_{X_n} \to s_X$ in distribution in $L^p(\mathbb{S}^{d-1} \times [0,1], \lambda \otimes \ell)$,*

where λ denotes the uniform measure in $\mathbb{S}^{d-1}$.

Proof. Denote by φ the mapping given by

$$\begin{aligned} \varphi : (\mathscr{F}_c(\mathbb{R}^d), \rho_p) &\to L^p(\mathbb{S}^{d-1} \times [0,1], \lambda \otimes \ell) \\ U &\mapsto s_U. \end{aligned}$$

By, e.g., Krätschmer (2006, p. 444), φ is an isometry.

Let $Y_n, Y : ([0,1], \mathscr{B}_{[0,1]}, \ell) \to (\mathscr{F}_c(\mathbb{R}^d), d_p)$ be the fuzzy random variables given by Lemma 2. We have $Y_n(t) \to Y(t)$ in d_p for all $t \in [0,1]$, $P_{X_n} = \ell_{Y_n}$ and $P_X = \ell_Y$. Since d_p and ρ_p are topologically equivalent (Diamond and Kloeden 1994, Proposition 7.4.5, p. 65), $\rho_p(Y_n(t), Y(t)) \to 0$ for every $t \in [0,1]$.

Set $s_{Y_n} = \varphi \circ Y_n$ and $s_Y = \varphi \circ Y$. Since φ is an isometry, we have $s_{Y_n}(t) \to s_Y(t)$. By Alonso de la Fuente and Terán (2022a, Proposition 5.4), $(\mathscr{F}_c(\mathbb{R}^d), d_p)$ is a Lusin space, hence it is Borel measurable in every metric space it embeds into (see Frolík 1970, Proposition 7.11). There follows that φ is Borel measurable and thus s_{Y_n} and s_Y are random elements of $L^p(\mathbb{S}^{d-1} \times [0,1], \lambda \otimes \ell)$.

We need to check $\ell_{s_{Y_n}} = P_{s_{X_n}}$. For any measurable subset A of $L^p(\mathbb{S}^{d-1} \times [0,1], \lambda \otimes \ell)$,

$$P_{s_{X_n}}(A) = P(\{\omega \in \Omega : s_{X_n}(\omega) \in A\}) = P(\{\omega \in \Omega : (\varphi \circ X_n)(\omega) \in A\})$$

$$= P(\{\omega \in \Omega : X_n(\omega) \in \varphi^{-1}(A)\}) = \ell(\{t \in [0,1] : Y_n(t) \in \varphi^{-1}(A)\})$$

$$= \ell(\{t \in [0,1] : (\varphi \circ Y_n)(t) \in A\}) = \ell(\{t \in [0,1] : s_{Y_n}(t) \in A\}) = \ell_{s_{Y_n}}(A).$$

Analogously, $\ell_{s_Y} = P_{s_X}$. Since $s_{Y_n} \to s_Y$ almost surely, by Kallenberg (2002, Lemma 4.2) almost sure convergence implies convergence in probability and by Kallenberg (2002, Lemma 4.7) convergence in probability implies weak convergence $\ell_{s_{Y_n}} \to \ell_{s_Y}$. In conclusion, $P_{s_{X_n}} \to P_{s_X}$ weakly, that is, $s_{X_n} \to s_X$ in distribution.

For the converse, notice that $P_{X_n} = P_{s_{X_n}} \circ \varphi$ and $P_X = P_{s_X} \circ \varphi$. Since φ is an isometry, for any open set G of $\mathscr{F}_c(\mathbb{R}^d)$ there exists an open set $\mathbf{G}$ of $L^p(\mathbb{S}^{d-1} \times [0,1], \lambda \otimes \ell)$ such that $\varphi(G) = \mathbf{G} \cap \varphi(\mathscr{F}_c(\mathbb{R}^d))$. Then

$$\liminf_{n\to\infty} P_{s_{X_n}} \circ \varphi(G) = \liminf_{n\to\infty} P_{s_{X_n}}(\varphi(G)) = \liminf_{n\to\infty} P_{s_{X_n}}(\mathbf{G} \cap \varphi(\mathscr{F}_c(\mathbb{R}^d)))$$

$$= \liminf_{n\to\infty} P_{s_{X_n}}(\mathbf{G}) \geq P_{s_X}(\mathbf{G}) = P_{s_X}(\mathbf{G} \cap \varphi(\mathscr{F}_c(\mathbb{R}^d))) = P_{s_X}(\varphi(G)) = P_{s_X} \circ \varphi(G)$$

by Lemma 1 and knowing that s_{X_n} and s_X take on values in $\varphi(\mathscr{F}_c(\mathbb{R}^d))$. Again by Lemma 1, $\liminf_{n\to\infty} P_{X_n} \circ \varphi(G) \geq P_X \circ \varphi(G)$ yields $X_n \to X$ in distribution in d_p. □

Therefore this type of convergence can indeed be studied using support functions. Another question concerns the relationship between convergence of fuzzy random variables taking on values in parametric families of fuzzy sets (in this case, trapezoidal fuzzy sets but the study could be extended to other families) and convergence in distribution of their defining parameters. The content of the following lemma is intuitively clear although its proof is not trivial. For space reasons we skip the proof, which may appear elsewhere.

Lemma 5. *Let $p \in [1, \infty)$. Let $U_n, U \in \mathscr{F}_c^{tra}(\mathbb{R})$. If $U_n \to U$ in d_p, then the sequence $\{\|(U_n)_0\|\}_n$ is bounded.*

Theorem 2. *Let $p \in [1, \infty)$. Let X_n be $Tra(X_{n,1}, X_{n,2}, X_{n,3}, X_{n,4})$ where $X_{n,1} \leq X_{n,2} \leq X_{n,3} \leq X_{n,4}$ are random variables, and analogously $X = Tra(X_1, X_2, X_3, X_4)$. Then $X_n \to X$ in distribution in d_p if and only if, as random vectors in $\mathbb{R}^4$, $(X_{n,1}, X_{n,2}, X_{n,3}, X_{n,4}) \to (X_1, X_2, X_3, X_4)$ in distribution.*

Proof. Set $A = \{(u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : u_1 \leq u_2 \leq u_3 \leq u_4\}$. The mapping

$$\varphi : A \to (\mathscr{F}_c(\mathbb{R}), d_p)$$
$$(u_1, u_2, u_3, u_4) \mapsto Tra(u_1, u_2, u_3, u_4)$$

is injective. Let us show that φ is continuous. Let $(u_{n,1}, u_{n,2}, u_{n,3}, u_{n,4}) \to (u_1, u_2, u_3, u_4)$ in $\mathbb{R}^4$. Denote by U_n the fuzzy set $Tra(u_{n,1}, u_{n,2}, u_{n,3}, u_{n,4})$ and by U the fuzzy set $Tra(u_1, u_2, u_3, u_4)$.

Now let $\alpha \in [0, 1]$,

$$d_H(U_{n_\alpha}, U_\alpha) = \max\{|\inf(U_n)_\alpha - \inf U_\alpha|, |\sup(U_n)_\alpha - \sup U_\alpha|\}$$
$$= \max\{|(1-\alpha)\inf(U_n)_0 + \alpha \inf(U_n)_1 - (1-\alpha)\inf U_0 - \alpha \inf U_1|,$$
$$|\alpha \sup(U_n)_1 + (1-\alpha)\sup(U_n)_0 - \alpha \sup U_1 - (1-\alpha)\sup U_0|\}$$
$$= \max\{|(1-\alpha)(u_{n,1} - u_1) + \alpha(u_{n,2} - u_2)|, |\alpha(u_{n,3} - u_3) + (1-\alpha)(u_{n,4} - u_4)|\}$$
$$\leq \max\{|u_{n,1} - u_1|, |u_{n,2} - u_2|, |u_{n,3} - u_3|, |u_{n,4} - u_4|\}.$$

Since the last term is the max distance between both vectors in $\mathbb{R}^4$ and is independent of α, indeed it bounds $d_p(U_n, U)$, making φ be d_p-continuous.

We will establish now two further facts which will be used in the proof. Firstly, since A is closed in $\mathbb{R}^4$, it is complete and separable. Moreover $(\mathscr{F}_c(\mathbb{R}), d_p)$ is separable, hence by Lemma 4 the image $\varphi(A) = \mathscr{F}_c^{tra}(\mathbb{R})$ is Borel measurable.

Secondly, set

$$A_{a,b} = \{(u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : a \leq u_1 \leq u_2 \leq u_3 \leq u_4 \leq b\} \subseteq A$$

for each $a, b \in \mathbb{R}$. Since $A_{a,b}$ is compact, φ is continuous and $\varphi(A) = \mathscr{F}_c^{tra}(\mathbb{R})$ is a Hausdorff space, the restriction $\varphi|_{A_{a,b}}$ is a homeomorphism (Joshi 1983, Corollary 2.4, p. 169).

($\Rightarrow$) By Lemma 2, there exist fuzzy random variables Y_n, Y such that $Y_n(t) \to Y(t)$ in d_p for each $t \in [0,1]$, $\ell_{Y_n} = P_{X_n}$ and $\ell_Y = P_X$. Since

$$\ell_{Y_n}(\mathscr{F}_c^{tra}(\mathbb{R})) = P_{X_n}(\mathscr{F}_c^{tra}(\mathbb{R})) = 1,$$

Y_n and Y are almost surely trapezoidal fuzzy sets. For clarity, we assume without loss of generality that all $Y_n(t), Y(t)$ are trapezoidal fuzzy sets (otherwise it would suffice to modify the value of those variables in a null set, which would not change their probability distributions). Set $Tra(Y_{n,1}, Y_{n,2}, Y_{n,3}, Y_{n,4}) = Y_n$, $Tra(Y_1, Y_2, Y_3, Y_4) = Y$ and let us show that $(Y_{n,1}, Y_{n,2}, Y_{n,3}, Y_{n,4})$ converges in distribution to (Y_1, Y_2, Y_3, Y_4).

By Lemma 5, each sequence $\{\|(Y_n)_0(t)\|\}_n$ is bounded by some constant M_t. Therefore $(Y_{n,1}(t), Y_{n,2}(t), Y_{n,3}(t), Y_{n,4}(t)) \in A_{-M_t,M_t}$ for all $t \in [0,1]$. By the homeomorphism between A_{-M_t,M_t} and $\varphi(A_{-M_t,M_t})$, the 4-dimensional vector converges to $(Y_1(t), Y_2(t), Y_3(t), Y_4(t))$. Almost sure convergence of those vectors implies their convergence in distribution. To finish the proof, we just need to check $\ell_{(Y_{1_n},\ldots,Y_{4_n})} = P_{(X_{1_n},\ldots,X_{4_n})}$. For any Borel subset $B \subseteq \mathbb{R}^4$,

$$\ell_{Y_{1,n},\ldots,Y_{4,n}}(B) = \ell(\{t \in [0,1] : (Y_{1,n}, \ldots, Y_{4,n})(t) \in A \cap B\})$$

$$= \ell(\{t \in [0,1] : Y_n(t) \in \varphi(A \cap B)\}) = P(\{\omega \in \Omega : X_n(\omega) \in \varphi(A \cap B)\})$$

$$= P(\{\omega \in \Omega : (X_{1,n}, \ldots, X_{4,n})(\omega) \in A \cap B\}) = P_{X_{1,n},\ldots,X_{4,n}}(B).$$

Analogously, $\ell_{Y_1,\ldots,Y_4} = P_{X_1,\ldots,X_4}$.

($\Leftarrow$) By the Skorokhod representation theorem in $\mathbb{R}^4$, there exist random vectors $(Y_{n,1}, Y_{n,2}, Y_{n,3}, Y_{n,4}), (Y_1, Y_2, Y_3, Y_4)$ such that $\ell_{(Y_{n,1},Y_{n,2},Y_{n,3},Y_{n,4})} = P_{(X_{n,1},X_{n,2},X_{n,3},X_{n,4})}$, $\ell_{(Y_1,Y_2,Y_3,Y_4)} = P_{(X_1,X_2,X_3,X_4)}$ and $(Y_{n,1}, Y_{n,2}, Y_{n,3}, Y_{n,4})$ converges to (Y_1, Y_2, Y_3, Y_4) pointwise. Set

$$Y_n = Tra(Y_{n,1}, Y_{n,2}, Y_{n,3}, Y_{n,4}), Y = Tra(Y_1, Y_2, Y_3, Y_4).$$

By the continuity of φ, $Y_n(t) \to Y(t)$ in d_p for each $t \in [0,1]$. By Lemmas 4.2 and 4.7 in Kallenberg (2002), almost sure convergence implies convergence in distribution. Finally, one shows like before $\ell_{Y_n} = P_{X_n}$ and $\ell_Y = P_X$, whence $X_n \to X$ in distribution in d_p. □

Since a random variable ξ can be identified with the trapezoidal fuzzy set $Tra(\xi, \xi, \xi, \xi)$, which is the indicator function $I_{\{\xi\}}$, the following corollary holds.

Corollary 1. *Let ξ_n, ξ be random variables. Then $\xi_n \to \xi$ in distribution if and only if $I_{\{\xi_n\}} \to I_{\{\xi\}}$ in distribution in d_p.*

The following proposition is analogous to an important property of convergence in distribution for random variables. It states that convergence is compatible with the operations in $\mathscr{F}_c(\mathbb{R}^d)$.

Proposition 1. *Let X_n, X be fuzzy random variables such that $X_n \to X$ in distribution in d_p. Then*

1. *For every $U \in \mathscr{F}_c(\mathbb{R}^d)$, we have $X_n + U \to X + U$ in distribution in d_p.*
2. *For every $a \in \mathbb{R}$, we have $aX_n \to aX$ in distribution in d_p.*

Proof. Since the mappings $V \in \mathscr{F}_c(\mathbb{R}^d) \mapsto V + U$ and $V \in \mathscr{F}_c(\mathbb{R}^d) \mapsto aV$ are d_p-continuous, we obtain the result with an application of the continuous mapping theorem (Lemma 3). □

Remark 1. It is not true, in general, that $X_n + Y \to X + Y$ in distribution in d_p provided $X_n \to X$ in distribution. That fails even for random variables.

We close the paper by pointing out another parallel with ordinary random variables: if the limit is a degenerate fuzzy random variable U, then $X_n \to U$ in distribution in d_p if and only if $X_n \to U$ in probability in d_p (by an application of Kallenberg 2002, Lemma 4.7).

Acknowledgements. Research in this paper was partially funded by grants and fellowships from Spain (PID2019-104486GB-I00), the Principality of Asturias (SV-PA-21-AYUD/2021/50897 and PA-21-PF-BP20-112), and the University of Oviedo (PAPI-20-PF-21). Their contribution is gratefully acknowledged.

References

Alonso de la Fuente, M., Terán, P.: Some results on convergence and distributions of fuzzy random variables. Fuzzy Sets Syst. **435**, 149–163 (2022)

Alonso de la Fuente, M., Terán, P.: Convergence theorems for random elements in convex combination spaces (2022b, submitted for publication)

Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)

Diamond, P., Kloeden, P.: Metric Spaces of Fuzzy Sets. World Scientific, Singapore (1994)

Frolík, Z.: A survey of separable descriptive theory of sets and spaces. Czechoslov. Math. J. **20**, 406–467 (1970)

Joshi, K.D.: Introduction to General Topology. New Age, New Delhi (1983)

Klement, E.P., Puri, M.L., Ralescu, D.A.: Limit theorems for fuzzy random variables. Proc. R. Soc. Lond. Ser. A **407**, 171–182 (1986)

Kallenberg, O.: Foundations of Modern Probability, 2nd edn. Springer, New York (2002). https://doi.org/10.1007/978-3-030-61871-1

Krätschmer, V.: A unified approach to fuzzy random variables. Fuzzy Sets Syst. **123**, 1–9 (2001)

Krätschmer, V.: Integrals of random fuzzy sets. TEST **15**, 433–469 (2006). https://doi.org/10.1007/BF02607061

Parthasarathy, K.R.: Probability Measures on Metric Spaces. Academic Press, New York (1967)

Puri, M.L., Ralescu, D.A.: Fuzzy random variables. J. Math. Anal. Appl. **114**, 409–422 (1986)

Terán, P.: Characterizing the distribution of fuzzy random variables. In: Book of Abstracts of the 5th International Conference ERCIM Working Group on Computing and Statistics, CFE-ERCIM, Oviedo, Spain, p. 22 (2012)

A Fuzzy Survival Tree (FST)

Jorge Luis Andrade(✉) and José Luis Valencia

Universidad Complutense, 28040 Madrid, Spain
jorandra@ucm.es, joseval@estad.ucm.es

Abstract. A fuzzy survival tree (FST) is introduced as an alternative proposal for survival tree learning. Fuzzy logic theory and Harrell's index (c-index) are combined as a new rule in node splitting. The introduction of fuzzy sets in tree learning improves FST performance and provides robustness to the algorithm when data are missing. FST performance improves significantly over other tree-based machine learning algorithms as demonstrated in public clinical datasets.

1 Introduction

Right-censored data have been commonly analysed using the Cox regression model (Cox 1972) under the proportional hazards assumption. Recently, studies have used algorithms such as the random survival forest (RSF). It's proposed by Ishwaran et al. (2008) as an extension of the random forest algorithm (Breiman 2001) to the right-censored survival problem. In the learning process at each tree, the maximization of the log-rank statistical test is used for nodes splitting. Ishwaran et al. (2008) demonstrates that the RSF performance is higher than the Cox model and its use is recommended when the assumptions of proportional risks are not met (Omurlu et al. 2009) or when the effect of the explanatory variables is nonlinear.

In decision trees for classification and regression problems, Ferri et al. (2002) and Lee (2019) proposed including the performance metric, the area under the curve (AUC), in the base learner. In trees for right-censored data analysis, the inclusion of the c-index (the AUC equivalent) is proposed by Schmid et al. (2016) with good results.

Recently, fuzzy set theory has been included in learning decision trees (Zhai et al. 2018; Mitra et al. 2002; Olaru and Wehenkel 2003, etc.) because datasets can contain missing values, noise in the class or outlier elements, etc. Fuzzy logic provides the flexibility to deal with these types of datasets without affecting the performance of the algorithm.

This article introduces fuzzy survival tree (FST). This new algorithm presents as a novelty the inclusion of fuzzy logic in combination with the c-index in the learning process of survival trees for splitting at each node.

L. A. García-Escudero et al. (Eds.): SMPS 2022, AISC 1433, pp. 16–23, 2023.
https://doi.org/10.1007/978-3-031-15509-3_3