INFINITE RANDOM SETS AND APPLICATIONS IN UNCERTAINTY ANALYSIS

(last revision: August 27, 2008)

Dissertation

By

Diego Andrés Alvarez Marín

diegoandresalvarez@gmx.net

Supervisor:

Ao. Univ.-Prof. Dr. Michael Oberguggenberger

Arbeitsbereich für Technische Mathematik am
Institut für Grundlagen der Bauingenieurwissenschaften
Leopold-Franzens Universität Innsbruck

External examiner:

Prof. Dr. George J. Klir

Department of System Science and Industrial Engineering
Thomas J. Watson School of Engineering and Applied Science
Binghamton University

SUBMITTED TO THE FAKULTÄT FÜR BAUINGENIEURWISSENSCHAFTEN
OF THE LEOPOLD-FRANZENS UNIVERSITÄT INNSBRUCK
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF SCIENCE (DR. TECHN.)

INNSBRUCK, AUSTRIA
FEBRUARY, 2007

To all those who hold me in their hearts

Contents

List of Tables

List of Figures

Abstract

The use of random set theory in uncertainty analysis has dealt until now with the use of focal sets with a finite number of elements; this particularization is equivalent to Dempster-Shafer evidence theory. In this thesis a generalization of evidence theory using an infinite number of focal elements is done. In particular, we define and study a type of random sets, that we call random sets of indexable type, which allows the natural modeling of cumulative distribution functions, probability boxes, possibility distributions or families of intervals provided by experts. This makes the method very suitable for uncertainty analyses with parameter uncertainty, providing in this way bounds on the probability of occurrence of events.

The first step in an uncertainty analysis employing finite random sets requires the discrezation of the basic variables, introducing in this way additional uncertainty in the analysis. The proposed method does not perform this step, making the proposed approach not only more accurate but also easier to program.

Conditions are presented under which "sequences of discretizations" of the basic variables converge to an infinite random set. In particular it is seen that the evaluation of the belief and plausibility by means of the finite approach is simply the evaluation of the Lebesgue-Stieltjes integrals of the lower and upper probabilities, employed when using infinite random sets of indexable type, by means of Riemann-Stieltjes sums. It is shown that this approximation is not efficient and that other strategies like Monte-Carlo-simulation-based methods are not only more accurate but also require less computational effort than the classical Dempster-Shafer approach in the estimation of the belief and plausibility.

The role of copulas in the specification of the dependence between basic variables is also analyzed. A connection between the theories of coherent lower and upper probabilities and expectations and infinite random sets is established. Other topics that are analyzed in this thesis include the definition of a measure of nonspecificity for infinite random sets and its use in sensitivity analysis.

The results shown discourage the future use of Dempster-Shafer evidence theory in uncertainty analysis and suggests the techniques developed here, not only because the proposed approach inherits all the nice properties of evidence theory, but also because the proposed approach removes several strong drawbacks that Dempster-Shafer theory has.

Acknowledgements

I would like to thank my tutor Ao. Univ.-Prof. Dr. Michael Oberguggenberger, for his countably infinite suggestions, constant support during this work, and for providing me a really pleasant working environment where creativity easily blooms. Without his help I would, for sure, have not been able to make this work. I will never forget him not only because he taught me that "the set of rational numbers can also be a failure set" :-) (this is a somewhat traumatic concept for an engineer, like me) but also because of the many nice things he taught me about life with his way of being. I am also thankful to Univ.-Ass. Dr. Thomas Fetz, who was always ready for answering questions about math and to Anna Bombasaro for helping me with some of the graphics and for the many chocolates and post cards she gave me :-).

Of course, I am grateful to my family and friends for their constant encouragement, support, never-ending love and understanding through the years.

This research was supported by the Programme Alßan, European Union Programme of High Level Scholarships for Latin America, identification number E03X17491CO.

Innsbruck, Austria Diego Andrés

February, 2007

Chapter 1

Introduction

Calvin: *I think we've got enough information now, don't you?*
Hobbes: *All we have is one "fact" you made up.*
Calvin: *That's plenty. By the time we add an introduction, a few illustrations, and a conclusion, it will look like a graduate thesis. Besides, I've got a secret weapon that will guarantee a good grade! No teacher can resist this! A clear plastic binder! Pretty professional looking, eh?*
Hobbes: *I don't want co-author credit on this, OK?*

Calvin and Hobbes (by Bill Watterson), 1991 Scientific Progress Goes "Boink" on page 25

1.1 Motivation

The classical problem in reliability analysis of structures (see e.g. Ditlevsen and Madsen (1996); Melchers (1999)) is the assessment of the probability of failure of a system. This probability is expressed as

$$P_{\mathrm{f}} = P_{\boldsymbol{X}}(F) = \int_F f_{\boldsymbol{X}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{1.1}$$

where $\boldsymbol{x}$ is the vector of *basic variables* which represents the material, loads and geometric characteristics of the structure, $F = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq 0, \boldsymbol{x} \in X\}$ is the failure region, $g : X \to \mathbb{R}$ is the so-called *limit state function*, which determines if a point $\boldsymbol{x}$ represents a safe ($g(\boldsymbol{x}) > 0$) or unsafe ($g(\boldsymbol{x}) \leq 0$) condition for a structure, $f_{\boldsymbol{X}}$ is the joint probability density function (PDF) of the implied random variables and $X \subseteq \mathbb{R}^n$.

In the last decades methods based on the evaluation of integral (1.1) have become popular and well developed to model and estimate the effect of uncertainties in engineering systems. These methods, although well founded theoretically, suffer the problem that in any given application the available information is usually incomplete and insufficient to define accurately the limit state function (this is called *model uncertainty*) and the joint PDF $f_{\boldsymbol{X}}$ (which is also referred to as *parameter uncertainty*); in consequence the methods in consideration lose applicability (Blockley, 1980, 1999).

For example, Oberguggenberger and Fellin (2002, 2004) showed that the probability of failure of a shallow foundation may fluctuate even by orders of magnitude (they were seen to range between 10^{-11} and 10^{-3}) when different likely PDFs associated to the soil parameters, and estimated using a relatively large number of samples, were considered. In the case of soil mechanics, those PDFs cannot be estimated accurately because of the limited sampling, the discrepancy between different methods of laboratory, the uncertainties in soil models, among other reasons. In addition, sometimes information cannot be expressed in a probabilistic fashion, but in terms of intervals (like it is found in many engineering manuals[1]) or linguistic terms (like the ones expressed by an expert); the reason for this difficulty resides in the fact that we are dealing with two very different classes of uncertainty; this will be clarified in the next section.

1.2 Types of uncertainty and the theories that are appropriate for their analysis

Helton (1997) classified uncertainty into two distinct groups: aleatory and epistemic. On the one hand, *aleatory uncertainty* (also referred to as *random* or *irreducible* or *inherent uncertainty*) is related to the natural variability of the variables involved.[2] Probability theory has proved to be well suited to manage this kind of uncertainty, representing the uncertainty associated to a variable using a PDF or a cumulative distribution function (CDF). Here every observation has to be a single number or vector and no imprecision is allowed. Small quantities of the probability of failure are associated to extreme observations of the input data which are by definition also scarce, in consequence usually observations are not being taken in these extreme regions. On the other hand, *epistemic uncertainty* stems from lack of knowledge or data, therefore it can be reduced when new information is available. Possibility, evidence, interval analysis and random set (RS) theories have shown to be appropriate to deal with this type of subjective uncertainty and in contraposition to aleatory uncertainty, here the information is expressed by means of intervals and linguistic terms.

In this sense, according to the characteristics of the information, the uncertainty quantification strategy is depicted by Bae et al. (2003) as shown in Figure 1.1. Parameter uncertainties are basically aleatory uncertainties, but they can be epistemic when there is not enough data available to construct an accurate cumulative distribution function (CDF). Model form and scenario abstraction uncertainties can come from unexpected failure modes, different choices of solution

[1] see e.g. Rackwitz (2000) who reports interval ranges for the means and standard deviations of soil properties for a range of cohesive and non-cohesive soils.

[2] Other authors have given other classifications of uncertainty. See for instance Klir and Wierman (1998) and the special issue of Reliability Engineering and System Safety edited by Helton and Burmaster (1996)

approaches, inaccurate models of the system, and so on, due to lack of knowledge and information, in consequence these uncertainties are classified as epistemic.

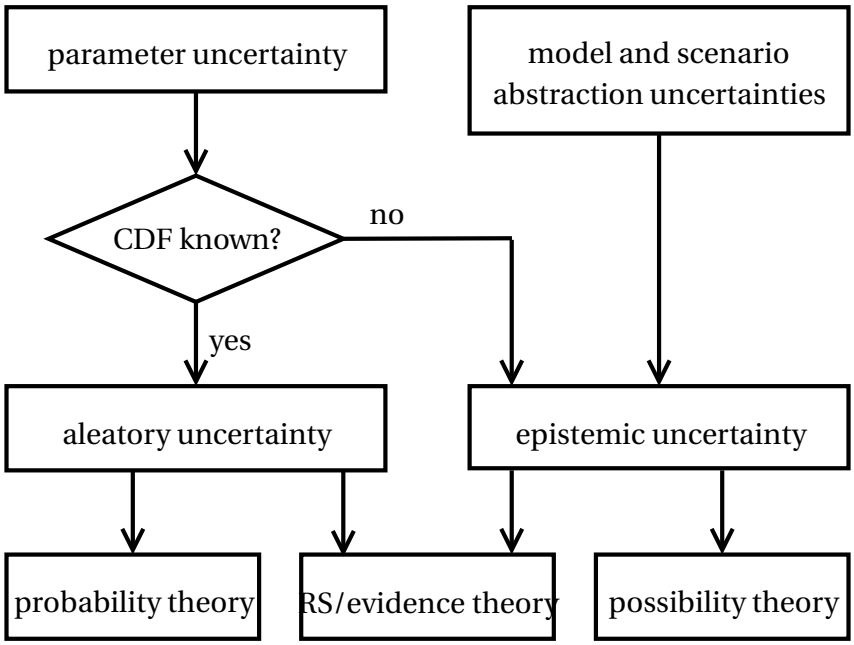


Figure 1.1: Uncertainty quantification techniques (source Bae et al. (2003))

As commented before, there exist several techniques to represent uncertainty (a good overview on these techniques is given in Joslyn and Booker (2004)). *Probability theory* (see e.g. Papoulis and Pillai (2002)) has been for long time the standard approach to represent aleatory uncertainty. In the 1960's, one framework for dealing with epistemic uncertainty was proposed by Zadeh (1965) who introduced the idea of *fuzzy set theory*, which allows us to represent imprecise concepts that are expressed by linguistic terms; for example, the values of a linguistic variable such as resistance can be represented as a fuzzy set taking the values high, moderate and low. Later, in 1967, Arthur Dempster developed a theory for representing epistemic knowledge on probability distributions propagated through multi-valued maps. Afterwards, in 1976, Glenn Shafer extended Dempster's work to what is known nowadays as *Dempster-Shafer evidence theory* or, in short, *evidence theory*. Meanwhile, the idea of *random set* (RS) *theory* appeared in the context the theory of stochastic geometry thanks to the independent works of Kendall (1974) and Matheron (1975), who studied the properties of random variables valued in closed bounded sets of $\mathbb{R}^d$ and of random sets; this approach is an extension of a probability theory to set-valued rather than point-valued maps. When the size of the body of evidence is finite[3], random sets resulted to be mathematically isomorphic to Dempster-Shafer bodies of evidence, although

[3] see section 2.1 for a definition.

with somewhat different semantics. In his doctoral thesis, Sugeno (1974) introduced the idea of fuzzy measure, which was intended to be a direct generalization of probability measures by relaxing the requirement of additivity. Several classes of fuzzy measures were identified, among them the belief and plausibility measures, already available within Dempster-Shafer theory. In 1978, Zadeh introduced the special class of fuzzy measures called possibility measures and in addition suggested a close connection to fuzzy sets. The 1980's and 1990's was a period of consolidation where several open problem within uncertainty theory were solved and relationships between these theories were studied. One of the most outstanding developments during this period was the introduction by Walley (1991) of the *theory of imprecise probabilities*. Imprecise probabilities are defined by sets of probability measures. Lower and upper probability (belief and plausibility) measures, defined in random set (and DS evidence theories), can be seen as special cases of imprecise probabilities. In fact these functions are known to be the lower and upper envelope of the probability measures that dominate them. Other theories for the quantification of the uncertainty have appeared during the last two decades, like for example the transferable belief model (see e.g. Smets and Kennes (1994)), the theory of hints (see e.g. Kohlas and PA (1995)) and the theory of fuzzy randomness (see e.g. Möller and Beer (2004)).

1.3 Advantages of random set and evidence theory

Given the great variety of theories to treat uncertainty, when both epistemic and aleatory uncertainties are present at the same time in an engineering system, sometimes researchers treat them separately by trying to convert one type of uncertainty into another. Random set and evidence theories are well suited to handle epistemic and aleatory uncertainties jointly without making assumptions which are not justified from the data, because it is a general framework that includes interval analysis, probability and possibility theory and probability bounds analysis as special cases. Therefore, they are excellent tools to deal with imprecise data inasmuch as they allow the designer engineer to use data in the format it appears. This is important because usually field observations is expressed by means of intervals, statistical data is expressed through histograms and when there is not enough information one must resort to experts who usually express their opinions either through intervals or in linguistic terms; note that in an elicitation process experts will almost always reach a point where they are indifferent between the possible alternatives.

Another advantage of random set and evidence theory is that they enable the designer to assess both the most pessimistic and optimistic assumptions that can be expected, without being advocated to include in the calculation dangerous and ungrounded suppositions; for example, it is possible to consider situations of total lack of knowledge about the dependence between the implied basic variables, instead of considering the dangerous supposition of independence.

1.4 Some notes about probability theory

According to Oberkampf et al. (2001), a common practice in probability theory when there is little or no closely related experimental data, is to simply pick some familiar CDF to represent the possible ranges of the values that may occur, like the normal or the uniform CDFs. Note that this selection may fit the bulk of the data, but will poorly represent the rare events, which reside in the tails of the CDF and that are the ones that usually induce the failure events, for any realistic sample size (which is usually limited). Also, discrepancy appears when the epistemic uncertainty is treated like random uncertainty. For example, suppose that only the range in which a variable dwells is known. Therefore, according to Laplace's principle of insufficient reason (see e.g. Papoulis and Pillai (2002)), a uniform CDF should be assumed to express total ignorance. Then the probabilistic analysis would exaggeratedly increase the probability of occurrence of the parameters near the ends of the interval. In contrast, random set theory does not need to suppose a CDF when there is not enough information, i.e., it does not make any claim regarding any specific value within the interval or the likelihood of any other value in the interval.

Also, probability theory is based on the three axioms of Kolmogorov, and therefore, for any event A it follows that $P(A) + P(A^c) = 1$, which implies that the absence of evidence of one event must be equal to the evidence for the negation of the event. For random set theory and the imprecise probability theories this supposition is too strong, that is, a weak statement of evidence can result in corroboration for an event, but the evidence makes no conclusions fostering the negation of the event.

1.5 Outline of the thesis

After the foregoing introduction to the types of uncertainties present in an engineering system, and the enumeration of the techniques that may be employed to deal with them, the document continues in Chapter 2 with a a self-contained summary of the main results to be employed in the rest of the document in relation to evidence and random set theories; thereafter, Chapter 3 summarizes some important concepts and theorems related to the theory of copulas; this section includes, in particular, a proof that states that the set of all copulas is compact.

The new results obtained in this dissertation start in Chapter 4 with the concept of infinite random sets (i.e., random sets with infinitely many focal elements); Chapter 5 discusses convergence considerations of finite random sets towards infinite random sets; Chapter 6 deals with dependence considerations when extending the theory to infinite random sets; Chapter 7 presents a characterization of the different types of discretizations of basic variables that are possible in random set theory; it also shows some techniques for the evaluation of the lower

and upper probability integrals appearing in the developed formulation; Chapter 8 deals with the extension of the Hartley-like measure of nonspecificity for finite random sets to infinite random sets; in the course of that chapter, concepts defined in the realm of Dempster-Shafer evidence theory, like joint and marginal random sets and random set inclusion are also generalized. Chapter 9 shows an application of nonspecificity to sensitivity analysis; Chapter 10 presents some final considerations, a summary of the main results developed in this thesis and a pool of ideas that could inspire future research.

The thesis concludes with Appendix A, which contains a literature review of random set and evidence theories in the context of risk and uncertainty assessment in civil and mechanical engineering and with Appendix B, where several ideas that deserve further research are explained.

Part I

The foundations

Chapter 2

Evidence and random set theories

The main idea of evidence and random set theories is that the knowledge of a given problem can be inherently imprecise. These closely related strategies allow the designer to take into account all uncertainties present in the data, in addition to provide an estimator of the upper and lower bounds on the probability of occurrence of an outcome. In the following, these theories are presented.

2.1 Dempster-Shafer evidence theory

This brief introduction to Dempster-Shafer (DS) evidence theory is based, among others, on Dubois and Prade (1991), Klir and Wierman (1998), Oberkampf et al. (2001), Sentz and Ferson (2002) and Joslyn and Booker (2004).

Following Dubois and Prade (1991), let us consider a universal non-empty set X and its power set $\mathscr{P}(X)$. Let

$$\mathscr{F} := \{A_1, A_2, \ldots, A_n\},$$

with $A_i \in \mathscr{P}(X) \setminus \emptyset$, and $A_i \neq A_j$ for $i, j = 1, \ldots, n$ and $i \neq j$, be a family of distinct (possibly overlapping) non-empty subsets of X and

$$m : \mathscr{F} \to [0, 1]$$

be a mapping called the *basic mass assignment* which satisfies the property $\sum_{A_i \in \mathscr{F}} m(A_i) = 1$. A pair $(\mathscr{F}, m)$ is called a *Dempster Shafer body of evidence* or a *Dempster Shafer structure* on X; both terms will be used indistinctively along this document. Helton et al. (2004) calls the triple $(X, \mathscr{F}, m)$ the *evidence space*. Every set $A_i \in \mathscr{F}$ has an associated $m(A_i) > 0$ and is called *focal element*. The collection of all focal elements $\mathscr{F}$ is called *focal set*. Each focal set contains sets of possible values of the variable $x \in X$, and $m(A_i)$ expresses the probability of A_i to be the actual range of x. Klir (1995) interprets $m(A_i)$ as the proportion to which all available and important evidence supports the assertion that a particular element of X belongs *exclusively* to the set A_i and does not imply any additional

claims regarding subsets of A_i. If there is some additional evidence supporting the claim that the element belongs to a subset B of A_i, it must be explicitly stated by another value $m(B)$. The main difference between the basic mass assignment and a probability mass function is that the first one is defined on X while the second one is defined on the power set of X, $\mathscr{P}(X)$. Also, the following properties hold for the basic mass assignment:

- it is not required that $m(X) = 1$;
- $B \subseteq A$ does not necessarily imply that $m(B) \leq m(A)$;
- it is not required $m(A) + m(A^c)$ to be one;
- it is not required $m(A)+m(B)$ to be $m(A\cup B)+m(A\cap B)$, given that $A, B, A\cup B$ and $A\cap B$ belong to $\mathscr{F}$;
- if $m(X) = 1$, it means that $\mathscr{F} = \{X\}$, i.e. X is the unique focal element; according to Klir (1989) this is the case of total ignorance and expresses the situation in which no information is available;
- if there is a unique focal element $A \subseteq X$ such that $m(A) = 1$, then $m(X) = 0$;
- if $n > 1$ (i.e. $|\mathscr{F}| > 1$) then $m(X) < 1$;
- if all elements of $\mathscr{F}$ are singletons (i.e. $|A_i| = 1, i = 1,\ldots,n$), then m is a probability mass function. In this case the focal set $\mathscr{F}$ is called *specific*.

Example 2.1.1. *Consider for instance a DS structure $(\mathscr{F}, m)$ which is formed by gathering the information provided by four different sources (e.g. books, experts, previous analysis, etc.) on the friccion angle of some soil; each of those opinions will form one element A_i of our focal set $\mathscr{F}$. Suppose that $\mathscr{F} := \{A_1 := [20°, 22°], A_2 := [21°, 23°], A_3 := [18°, 20°], A_4 := [20°, 25°]\}$. Note that there exist focal elements that contain conflicting information, and some others that share the same information. The basic mass assignment given to each of those focal elements will represent the importance of each of those opinions in our assessments. Suppose for example that $m(A_1) = 0.4, m(A_1) = 0.2, m(A_1) = 0.1, m(A_1) = 0.3$; this means that we are giving to our first source of information the largest relevance.*

Belief and Plausibility measures

Due to our lack of information, the probability that an element x belongs to a set F, where $F \in \mathscr{P}(X)$, cannot be expressed by a single value of probability, but is bounded from above and below by

$$\mathrm{Bel}_{(\mathscr{F},m)}(F) \leq P(F) \leq \mathrm{Pl}_{(\mathscr{F},m)}(F) \tag{2.1}$$

where belief and plausibility measures Bel and Pl are defined respectively by,

$$\mathrm{Bel}_{(\mathscr{F},m)}(F) := \sum_{i=1}^{n} I[A_i \subseteq F] m(A_i) \tag{2.2}$$

and

$$\mathrm{Pl}_{(\mathscr{F},m)}(F) := \sum_{i=1}^{n} I[A_i \cap F \neq \emptyset] m(A_i) \tag{2.3}$$

for $A_i \in \mathscr{F}$; in words, the *belief* is the sum of the basic assignments of those focal elements contained in F, and whose occurrence *imply* that $x \in F$; the *plausibility* is the sum of the basic assignment of those focal elements which have some element x in common with F and whose occurrence *imply or could imply* that $x \in F$. Strict equalities on (2.1) occur when $\mathscr{F}$ is a specific focal set because in this case m is a probability mass function.

Moreover, when X is a finite set, the basic mass assignment can be restored from the belief or plausibility measures using the so-called Möbius transform,

$$m(F) = \sum_{i=1}^{|\mathscr{F}|} I[A_i \subseteq F](-1)^{|A_i \setminus F|} \mathrm{Bel}_{(\mathscr{F},m)}(A_i)$$

The quantity $\mathrm{Pl}_{(\mathscr{F},m)}(F) - \mathrm{Bel}_{(\mathscr{F},m)}(F)$ is known as the *degree of ignorance* around the proposition $x \in F$. This gap will decrease when new information is available, that means that it will be zero if the epistemic uncertainty vanishes; this topic will be treated in deep in Chapter 9.

A note on notation is necessary. The subindex $\cdot_{(\mathscr{F},m)}$ in $\mathrm{Bel}_{(\mathscr{F},m)}$ and $\mathrm{Pl}_{(\mathscr{F},m)}$ denoting the body of evidence in question will not be indicated when $(\mathscr{F},m)$ is easily identifiable from the context.

The belief and plausibility functions can be identified as dual fuzzy (non-additive) measures (see e.g. Wang and Klir (1992) and Dennenberg (1994)), that is

$$\begin{aligned} \mathrm{Bel}_{(\mathscr{F},m)}(F) &= 1 - \mathrm{Pl}_{(\mathscr{F},m)}(F^{\mathrm{c}}) \\ \mathrm{Pl}_{(\mathscr{F},m)}(F) &= 1 - \mathrm{Bel}_{(\mathscr{F},m)}(F^{\mathrm{c}}) \end{aligned}$$

and the following properties hold:

- *Super-additivity*: $\mathrm{Bel}(A \cup B) \geq \mathrm{Bel}(A) + \mathrm{Bel}(B) - \mathrm{Bel}(A \cap B)$.
- *Sub-additivity*: $\mathrm{Pl}(A \cup B) \leq \mathrm{Pl}(A) + \mathrm{Pl}(B) - \mathrm{Pl}(A \cap B)$.

and therefore, in contrast to the probability measure, where the additivity condition $\mathrm{P}(A)+\mathrm{P}(A^{\mathrm{c}}) = 1$ holds, we have that $\mathrm{Bel}(A)+\mathrm{Bel}(A^{\mathrm{c}}) \leq 1$ and $\mathrm{Pl}(A)+\mathrm{Pl}(A^{\mathrm{c}}) \geq 1$.

Example 2.1.2. *Continuing with example 2.1.1, suppose that we want to know what is the probability that the friction angle resides in the interval $F = [21°, 23°]$ provided the information contained in the DS structure $(\mathscr{F}, m)$. We have that $\mathrm{Bel}_{(\mathscr{F},m)}(F) = 0.2 + 0.3 = 0.5$ and that $\mathrm{Pl}_{(\mathscr{F},m)}(F) = 0.4 + 0.2 + 0.3 = 0.9$. This means that the probability that the friction angle resides in the interval F is upper-bounded by 0.9 and lower-bounded by 0.5.*

2.1.1 Some other concepts

The *support* of the focal set $\mathscr{F}$ is the set $\cup_{A_i \in \mathscr{F}} A_i$. In addition, the *core* of the focal set $\mathscr{F}$, core($\mathscr{F}$), is the set $\cap_{A_i \in \mathscr{F}} A_i$.

A *consistent* focal set is one whose core($\mathscr{F}$) $\neq \emptyset$, that is all evidence claims about the set although imprecise, agree in some information.

A focal set is said to be *consonant* if its focal elements are nested, that is its focal elements can be ordered in a way such that i.e. $A_i \subseteq A_{i+1}, i = 1, \ldots, n-1$, this means that all evidence is almost in complete agreement because each focal element confirms and makes more specific the information contained in the subsequent focal element.

Dempster-Shafer bodies of evidence are closely related to finite random sets. The following section gives a presentation of random set theory.

2.2 Random set theory

The following is a brief introduction to random set theory, after Wolkenhauer (2001) and Bertoluzza et al. (2002).

2.2.1 Generalities on random variables

Let $(\Omega, \sigma_\Omega, P_\Omega)$ be a probability space and (X, σ_X) be a measurable space. A *random variable* X is a $(\sigma_\Omega - \sigma_X)$-measurable mapping

$$X : \Omega \to X; \alpha \mapsto X(\alpha). \tag{2.4}$$

This mapping can be used to generate a probability measure on (X, σ_X) such that the probability space (X, σ_X, P_X) is the mathematical description of the experiment as well as of the original probability space $(\Omega, \sigma_\Omega, P_\Omega)$. This mapping is given by $P_X = P_\Omega \circ X^{-1}$. This means that an event $F \in \sigma_X$ has the probability

$$\begin{aligned} P_X(F) &= P_\Omega \circ X^{-1}(F) \\ &= P_\Omega \left(X^{-1}(F) \right) \\ &= P_\Omega \{ \alpha : X(\alpha) \in F \} \end{aligned} \tag{2.5}$$

for $\alpha \in \Omega$. The benefit of mapping (2.4) arises when (X, σ_X) is a well characterized measurable space where mathematical tools such as Riemann integration are well defined. One of the most commonly used measurable spaces is $(\mathbb{R}, \mathscr{B})$, where $\mathscr{B}$ is the Borel σ-algebra on $\mathbb{R}$; in this case X is called a *numerical random variable*. For example, according to the Kolmogorov axioms, the probability

measure is additive, therefore, depending on whether we have a discrete or continuous random variables, it follows that, for $F \in \sigma_X$,

$$\begin{aligned} P_X(F) &:= \sum_{\alpha \in X^{-1}(F)} P_\Omega(\{\alpha\}) \qquad \text{(discrete case)} \\ &= \int_{X^{-1}(F)} \mathrm{d}P_\Omega(\alpha) \qquad \text{(general case)} \end{aligned}$$

In the last case, $P_X(F)$ can also be written in terms of the expectation operation as

$$P_X(F) = \int_F \mathrm{d}P_X(x) = \int_X I\left[x \in F\right] \mathrm{d}P_X(x) = E_X\left[I\left[x \in F\right]\right]$$

where $x \in X$, and I stands for the *indicator function*, defined by,

$$I[\cdot] = \begin{cases} 1 & \text{if } \cdot \text{ is true} \\ 0 & \text{if } \cdot \text{ is false} \end{cases}$$

2.2.2 Generalities on random sets

Definition 2.2.1. *Let us consider a universal non-empty set X and its power set $\mathscr{P}(X)$. Let $(\Omega, \sigma_\Omega, P_\Omega)$ be a probability space and $(\mathscr{F}, \sigma_{\mathscr{F}})$ be a measurable space where $\mathscr{F} \subseteq \mathscr{P}(X)$. A* random set Γ *is a $(\sigma_\Omega - \sigma_{\mathscr{F}})$-measurable mapping $\Gamma : \Omega \to \mathscr{F}$, $\alpha \mapsto \Gamma(\alpha)$. We will call every $\gamma := \Gamma(\alpha) \in \mathscr{F}$ a* focal element *while $\mathscr{F}$ will be called a* focal set.

In an analogous way to the definition of a random variable, this mapping can be used to generate a probability measure on $(\mathscr{F}, \sigma_{\mathscr{F}})$ given by $P_\Gamma := P_\Omega \circ \Gamma^{-1}$. This means that an event $\mathscr{R} \in \sigma_{\mathscr{F}}$ has the probability

$$P_\Gamma(\mathscr{R}) = P_\Omega\{\alpha : \Gamma(\alpha) \in \mathscr{R}\}. \tag{2.6}$$

In short, a random set is a set-valued random variable. Note that the RS Γ will be called a *finite* or *infinite random set* depending on the cardinality of $\mathscr{F}$ (this in contrast to the definition employed by some other authors when the final space is finite; in fact, some authors do not mention explicitly the initial space).

When all elements of $\mathscr{F}$ are singletons (points), then Γ becomes a random variable, and $\mathscr{F}$ is called *specific* ; in other words, when $\mathscr{F}$ is specific, $\Gamma(\alpha) = X(\alpha)$ and the value of the probability of occurrence of the event F, $P_X(F)$, can be exactly captured by equation (2.5) for any $F \in \sigma_X$. In the case of random sets, it is not possible to know the exact value of $P_X(F)$ but upper and lower bounds of it. Dempster (1967) defined those upper and lower probabilities by the measures

$$\begin{aligned} \mathrm{LP}_{(\mathscr{F}, P_\Gamma)}(F) &:= P_\Omega\{\alpha : \Gamma(\alpha) \subseteq F, \Gamma(\alpha) \neq \emptyset\} \qquad (2.7) \\ &= P_\Gamma\left\{\gamma : \gamma \subseteq F, \gamma \neq \emptyset\right\} \\ \mathrm{UP}_{(\mathscr{F}, P_\Gamma)}(F) &:= P_\Omega\{\alpha : \Gamma(\alpha) \cap F \neq \emptyset\} \qquad (2.8) \\ &= P_\Gamma\left\{\gamma : \gamma \cap F \neq \emptyset\right\} \end{aligned}$$

where

$$\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) \leq P_X(F) \leq \mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F). \tag{2.9}$$

The strict equality in (2.9) occurs when $\mathscr{F}$ is specific.

In the following, we will refer to $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$ respectively as the *lower* and *upper* probability measures of the set F with respect to the random set $(\mathscr{F},P_\Gamma)$. Note also that these measures are a generalization of the belief and plausibility measures (see equations (2.2) and (2.3)) defined by Shafer (1976) in evidence theory, and which are used exclusively when dealing with finite random sets.

It can be shown that lower and upper probability measures are *dual fuzzy measures*, that is

$$\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) = 1 - \mathrm{UP}_{(\mathscr{F},P_\Gamma)}\left(F^{\mathrm{c}}\right)$$
$$\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F) = 1 - \mathrm{LP}_{(\mathscr{F},P_\Gamma)}\left(F^{\mathrm{c}}\right)$$

and also that lower probability is an ∞-monotone Choquet capacity and the upper probability is an ∞-alternating Choquet capacity (see e.g. (Klir, 2006, p.66)).

(Infinite) random sets have been analyzed several other authors. See for example Miranda et al. (2005b,a); Goodman and Nguyen (2002) and references therein.

2.2.3 What is the relationship between Dempster-Shafer bodies of evidence and random sets?

Random set theory is closely related to Dempster-Shafer evidence theory. Indeed, when the cardinality of the body of evidence is finite, (finite) random sets are mathematically isomorphic to Dempster-Shafer bodies of evidence, although with somewhat different semantics. That is, given a body of evidence $(\mathscr{F}_n,m)$ with $\mathscr{F}_n = \{A_1,A_2,\ldots,A_n\}$ and a RS $\Gamma:\Omega\to\mathscr{F}$, then the following relationships appear: $\mathscr{F}\equiv\mathscr{F}_n$, i.e., $A_j\equiv\gamma_j$ for $j=1,2,\ldots,n$ and $m(A_j)\equiv P_\Gamma(\gamma_j)$.

Note 1: In the remainder of this document, we will generally refer to random sets, and will represent them in the infinite case as $(\mathscr{F},P_\Gamma)$ or in the finite case either as $(\mathscr{F},m)$ or as $(\mathscr{F}_n,m)$ when special emphasis in the cardinality of $\mathscr{F}$ is desired. In some cases the superindex i will be employed to denote the index of a marginal RS in a random relation (see Section 2.2.4); for example if $\left(\mathscr{F}^i,m^i\right)$ for $i=1,\ldots,d$ are finite random sets defined on $\mathbb{R}$ then $(\mathscr{F},m) := \times_{i=1}^{d}\left(\mathscr{F}^i,m^i\right)$ is the corresponding joint finite random set, up to some dependence specification; this should be clear from the context. In the finite RS representation, the focal elements will be denoted as A_j for some $j=1,\ldots,n$, while in the infinite case, they will be denoted as γ or as $\Gamma(\alpha)$, except when the RS represents a possibility distribution, in which case the notation will be A_α, to agree with the conventional notation of α-cut. The reader is referred to Section 2.4.1 for details.

Note 2: Sometimes a sequence of random sets is required. The ordering in the sequence will be expressed by a superindex in parenthesis; for instance

$\left(\mathscr{F}^{(k),i}, m^{(k),i}\right)$ makes explicit the fact that this is the i-th marginal random set corresponding to the k-th RS $\left(\mathscr{F}^{(k)}, m^{(k)}\right)$ of the sequence.

2.2.4 Random relations for finite random sets

In order to deal with functions of several variables, it is customary to introduce the definition of a random relation. Let $X := \times_{i=1}^{d} X_i$. A random relation on X is a RS $(\mathscr{F}, m)$ on the Cartesian product X given by the combination of the marginal random sets $\left(\mathscr{F}^i, m^i\right)$, where $\mathscr{F}^i = \left\{A_{j_i}^i : j_i = 1, \ldots, n_i\right\}, i = 1, \ldots, d$ by the formula[1], $(\mathscr{F}, m) := \left(A_{j_1,\ldots,j_d} := \times_{i=1}^{d} A_{j_i}^i, m_{j_1,\ldots,j_d} := f\left(m^1, \ldots, m^d\right)\right)$. The function f takes into account the dependence relation between the marginal random sets. For example, when the marginal random sets are independent, random set independence can be used (see Fetz and Oberguggenberger (2004)) and therefore basic mass assignment in the joint space can be obtained as $m\left(A_{j_1,\ldots,j_d}\right) := \prod_{i=1}^{d} m^i\left(A_{j_i}^i\right)$, for all $A_{j_1,\ldots,j_d} \in \mathscr{F}$ i.e, as the product of the basic mass assignments m^i of the marginal random sets.

When nothing is known about the dependence between the basic variables, unknown interaction (see Fetz and Oberguggenberger (2004)) is the most conservative method in reliability analysis, because it contains all possible answers that result from all possible dependency relationships. In this case f is defined as the solution of a linear optimization problem. We will explain this case in Section 2.3.

2.2.5 Extension principle for finite random sets

Given a function $g : X \to Y$ (which represents for example the system response) and a RS $(\mathscr{F}, m)$ one could be interested in the image of $(\mathscr{F}, m)$ through g, i.e. $(\mathscr{R}, \rho)$. This mapped RS can be obtained by the application of the extension principle (Dubois and Prade (1991)):

$$\mathscr{R} := \left\{R_j := g\left(A_i\right) : A_i \in \mathscr{F}\right\} \tag{2.10}$$

$$\rho\left(R_j\right) := \sum_{i=1}^{n} I\left[R_j = g(A_i)\right] m\left(A_i\right) \tag{2.11}$$

Usually, $X \subseteq \mathbb{R}^d$, and $A_i \in \mathscr{F}$ is a d-dimensional box with 2^d vertices obtained as a Cartesian product of finite intervals, i.e. $A_i := I_1 \times \cdots \times I_d$. It must be noted that $|\mathscr{R}| \leq |\mathscr{F}|$ because some focal sets could have the same image through g.

The calculation of the image of the focal elements through the function g (equation (2.10)), when $Y \subseteq \mathbb{R}$, is usually performed by one of the following techniques: the optimization method, the sampling methods, the vertex method and the function approximation method.

[1] Note that here i was employed as a index of the marginal RS, not as an exponent.

The optimization method

If the focal element A_i is connected and compact and g is continuous, then R_j can be calculated as $\left[\min_{x\in A_i} g(x), \max_{x\in A_i} g(x)\right]$. This method is appropriate when g is a nonlinear function of the system parameters. The main drawback of this method is that it requires a high computational effort in a complex and large scale system.

The sampling method

The sampling method consists in assuming a PDF defined on every focal element (the simplest way is to assume a uniform PDF). After generating a desired sample for the assumed PDF, the belief and plausibility can be estimated by evaluating those points in g. If the drawn sample is large enough, this method gives a reliable result. However this method can be computationally expensive when great precision in the evaluation of R_j is required. In Joslyn and Kreinovich (2005) there is a well founded mathematical proof of the convergence of this method.

It is worth noting that the sampling method generates an inner approximation to the image of the focal element. This may, depending on the configuration of the problem, result in wrong estimates of the image.

This method is not computationally efficient in high dimensions because an appropriate sampling requires a large number of samples, that is, this method suffers the *curse of dimensionality* (term coined by Bellman (1961)). This can be better understood using the following example: the volume of a hypersphere of radius r in an d-dimensional space is $V(r,d) = S_d r^d/d$ where S_d is the hyper-surface area of a d-sphere of unit radius. The question is: what is the radius of the hypersphere that contains 5% of the volume of the unit hypersphere? i.e., for which r it holds that $V(r,d)/V(1,d) = 0.05$? The answer is $r = \sqrt[d]{0.05}$. In this case, for $d = 1,\ldots,10$, it follows that $r = 0.0500, 0.2236, 0.3684, 0.4729, 0.5493, 0.6070, 0.6518, 0.6877, 0.7169$ and 0.7411 respectively. So, when the number of dimensions increases, the required radius to cover the same percentage of volume increases. In consequence, the number of points required to sample the space effectively increases as the dimension increases, and thus, the method has the curse of dimensionality.

The vertex method

The sampling and the optimization methods can be infeasible in practice due to their large computational cost. To circumvent this problem the vertex method (Dong and Shah (1987)) can be employed. This technique only checks the vertices of each joint event to find the maximum and minimum responses, however, its application is only valid when the function g is monotonic and continuous with respect to each variable. For non-monotonic and discontinuous functions this method provides erroneous responses. Also its computational cost grows

exponentially with the number of uncertain variables; however, it can reduce the computational cost significantly compared to the sampling and optimization methods.

One can check for a single element how the function increases and once it is known which are the nodes that contain the extremes, the function g has to be calculated only twice for every focal element A_i, since these combinations are always the same for all focal elements A_i.

When the function is not monotonic, this method can be used as an approximation which will become closer to the correct result as the size of the focal elements decreases.

The function approximation method

The main idea of the function approximation method is to fit to the function g a surrogate function $\hat{g}$, called the *response surface*, which approximates g up to some degree of accuracy and which is computationally much easier to evaluate than g. Then the vertex or optimization methods are used, so that the evaluations of g are performed instead on the response surface $\hat{g}$. This method is beneficial because it reduces the computational cost of the application of the extension principle since there is no need to call the system solver g but the surrogate when some computation is required. Note that the accuracy of the algorithm largely depends on the response surface employed. Bae et al. (2004a) have successfully used this approach within evidence theory. Traditional methodologies of response surfaces have employed linear and quadratic polynomial functions; however recently artificial neural networks have been popularized as a powerful nonlinear response surface. See e.g. Papadrakakis et al. (1996) and Hurtado and Alvarez (2001) for an example of its application in the area of reliability assessment of structural systems.

2.2.6 Inclusion of finite random sets

Yager (1986); Delgado and Moral (1987); Dubois and Prade (1991) provided the following definition of inclusion for finite random sets:

Definition 2.2.2 (Dubois and Prade (1991))**.** *The finite random set* $\left(\mathscr{F}^A, m^A\right)$ *is said to be* included *in the finite random set* $\left(\mathscr{F}^B, m^B\right)$*, denoted as* $\left(\mathscr{F}^A, m^A\right) \subseteq \left(\mathscr{F}^B, m^B\right)$ *if and only if the following three conditions hold:*

1. *For all* $A \in \mathscr{F}^A$ *there exists a set* $B \in \mathscr{F}^B$ *such that* $A \subseteq B$.
2. *For all* $B \in \mathscr{F}^B$ *there exists a set* $A \in \mathscr{F}^A$ *such that* $A \subseteq B$.
3. *There exists an* assignment matrix $W : \mathscr{F}^A \times \mathscr{F}^B \to [0,1]$, $\left(A_i, B_j\right) \mapsto$

$W\left(A_i, B_j\right)$ *such that:*

$$m^A(A_i) = \sum_{j=1}^{s} W\left(A_i, B_j\right) \quad \textit{for all } A_i \in \mathscr{F}^A \tag{2.12}$$

$$m^B(B_j) = \sum_{i=1}^{r} W\left(A_i, B_j\right) \quad \textit{for all } B_j \in \mathscr{F}^B \tag{2.13}$$

and $W\left(A_i, B_j\right) = 0$ *if* $A_i \not\subset B_j$*, i.e., if* A_i *is not strictly contained in* B_j*.*

We will generalize this concept to infinite random sets in Section 8.4. The following theorems provide important properties useful in the manipulation of random sets:

Theorem 2.2.3 (Dubois and Prade (1991))**.** *The inclusion of two finite random sets implies the inclusion of the interval* $[\mathrm{Bel}(\cdot), \mathrm{Pl}(\cdot)]$*, i.e.* $\left(\mathscr{F}^A, m^A\right) \subseteq \left(\mathscr{F}^B, m^B\right)$ *implies*

$$\left[\mathrm{Bel}_{\left(\mathscr{F}^A, m^A\right)}(F), \mathrm{Pl}_{\left(\mathscr{F}^A, m^A\right)}(F)\right] \subseteq \left[\mathrm{Bel}_{\left(\mathscr{F}^B, m^B\right)}(F), \mathrm{Pl}_{\left(\mathscr{F}^B, m^B\right)}(F)\right]$$

but the converse is not generally true.

Theorem 2.2.4 (Dubois and Prade (1991))**.** *Let* $\left(\mathscr{F}^A, m^A\right)$ *and* $\left(\mathscr{F}^B, m^B\right)$ *be two finite random sets on* U *and* f *be a function* $U \to Y$*. Let* $\left(\mathscr{R}^A, \rho^A\right)$ *and* $\left(\mathscr{R}^B, \rho^B\right)$ *be the respective images of* $\left(\mathscr{F}^A, m^A\right)$ *and* $\left(\mathscr{F}^B, m^B\right)$ *through* f*, and let* $\left(\mathscr{F}^A, m^A\right) \subseteq \left(\mathscr{F}^B, m^B\right)$*; then* $\left(\mathscr{R}^A, \rho^A\right) \subseteq \left(\mathscr{R}^B, \rho^B\right)$

2.3 Dependence considerations for finite random sets

Couso et al. (1999) and Fetz and Oberguggenberger (2004) identified several kinds of dependence in sets of joint probability measures (they reside inside the focal elements of a joint finite random set) which result as a combination of the marginal sets of probability measures residing in their marginal finite random sets. According to the way the basic mass assignment of the finite random sets and the probability measures on these sets is chosen, different types of dependence such as strong independence, random set independence or unknown interaction appear.

We will neither deal with the topic of how to distribute the probability measures inside the marginal focal elements, nor how to combine them. In consequence we will consider *all* possible probability measures residing in the marginal focal elements and also we will allow *arbitrary dependence* for the combination of those marginal probability measures. When these matters are taken into consideration, *strong independence* plays a main role. However, when these probability measures are allowed to vary freely inside the focal elements, random set independence and unknown interaction must be analyzed.

Random set independence If $(\mathscr{F}, m) := \times_{i=1}^{d} \left(\mathscr{F}^i, m^i\right)$ is the random set formed under *random set independence*, the basic mass assignment corresponding to the focal element $A_{j_1}^1 \times A_{j_2}^2 \times \cdots \times A_{j_d}^d =: A_{j_1,j_2,\ldots,j_d} \in \mathscr{F}$ is $m\left(A_{j_1,j_2,\ldots,j_d}\right) := \prod_{i=1}^{d} m^i\left(A_{j_i}^i\right)$ while for the probability measures inside $A_{j_1,j_2,\ldots,j_d}$ dependent selections are allowed.

Unknown interaction When there is no dependence information about how to assign the basic mass assignment m for the set $A_{j_1,j_2,\ldots,j_d}$ and also when the probability measures on $A_{j_1,j_2,\ldots,j_d}$ are allowed to vary freely, *unknown interaction* is employed. Fetz and Oberguggenberger (2004) developed a linear programming approach to estimate bounds for m. The method can be expressed as follows: suppose d finite random sets $\left(\mathscr{F}^i, m^i\right)$, for $i = 1, \ldots, d$ are given. The combination of these RSs generates the joint random set $(\mathscr{F}, m)$ where $\mathscr{F}$ stands for the set resulting as the direct product $\mathscr{F} := \mathscr{F}^1 \times \mathscr{F}^2 \times \cdots \times \mathscr{F}^d$, i.e., as the system of sets $\left\{A_{j_1,j_2,\ldots,j_d} := A_{j_1}^1 \times A_{j_2}^2 \times \cdots \times A_{j_d}^d\right\}$ and m represents the basic mass assignment, which assigns a mass $m_{j_1,j_2,\ldots,j_d}$ to the joint focal element $A_{j_1,j_2,\ldots,j_d}$, i.e. $m_{j_1,j_2,\ldots,j_d} := m\left(A_{j_1,j_2,\ldots,j_d}\right)$ for all $A_{j_1,j_2,\ldots,j_d} \in \mathscr{F}$, and is calculated using the solution to the linear optimization problems

$$\underline{P}(F) = \min_{m \in \mathscr{M}(\mathscr{F})} \sum_{A_{j_1,j_2,\ldots,j_d} \in \mathscr{F}} I\left[A_{j_1,j_2,\ldots,j_d} \subseteq F\right] m\left(A_{j_1,j_2,\ldots,j_d}\right) \tag{2.14}$$

$$\overline{P}(F) = \max_{m \in \mathscr{M}(\mathscr{F})} \sum_{A_{j_1,j_2,\ldots,j_d} \in \mathscr{F}} I\left[A_{j_1,j_2,\ldots,j_d} \cap F \neq \emptyset\right] m\left(A_{j_1,j_2,\ldots,j_d}\right) \tag{2.15}$$

where $F \subseteq \mathbb{R}^d$ and $\mathscr{M}(\mathscr{F})$ is the set of all basic mass assignments that can be associated to the focal set $\mathscr{F}$, subject to

$$m^i\left(A_{j_i}\right) = \sum_{j_1=1}^{\left|\mathscr{F}^1\right|} \cdots \sum_{j_{i-1}=1}^{\left|\mathscr{F}^{i-1}\right|} \sum_{j_{i+1}=1}^{\left|\mathscr{F}^{i+1}\right|} \cdots \sum_{j_d=1}^{\left|\mathscr{F}^d\right|} m\left(A_{j_1,j_2,\ldots,j_d}\right) \tag{2.16}$$

(the joint mass assignment must have marginals corresponding to the basic mass assignments of the marginal random sets $\left(\mathscr{F}^i, m^i\right), i = 1, \ldots, d$) with $A_{j_i} \in \mathscr{F}^i$ for all $j_i = 1, \ldots, \left|\mathscr{F}^i\right|$, $i = 1, \ldots, d$ and

$$m\left(A_{j_1,j_2,\ldots,j_d}\right) \geq 0 \text{ for all } A_{j_1,j_2,\ldots,j_d} \in \mathscr{F}. \tag{2.17}$$

(all basic mass assignments must be non-negative) provides, under the assumption of lack of information about dependence, bounds on $m\left(A_{j_1,j_2,\ldots,j_d}\right)$, i.e. $\left[\underline{m}\left(A_{j_1,j_2,\ldots,j_d}\right), \overline{m}\left(A_{j_1,j_2,\ldots,j_d}\right)\right]$. Here $\underline{m}$ and $\overline{m}$ are the basic mass assignments associated to the argument that produces the infimum and the supremum in the solution to (2.14) and (2.15) correspondingly. Note that $\underline{m}$ and $\overline{m}$ do depend on F. However, here this does not represent additional complications because F is kept fixed.

Figure 2.1 shows a conceptual representation when $d = 2$, which illustrates condition (2.16). Here the RSs $\left(\mathscr{F}^1, m^1\right)$ and $\left(\mathscr{F}^2, m^2\right)$ are represented by the left and

upper bars respectively, while the joint RS $(\mathscr{F}, m)$ is symbolized by the central rectangle; also, every focal element is represented by a cell within the grid. In addition, the constraint (2.16) for $i = 1,2$ is shown. Note that the position of the cells in the grid has nothing to do with the spatial location of the focal elements on X, but symbolizes its position with regard to the indexing of elements in the focal set.

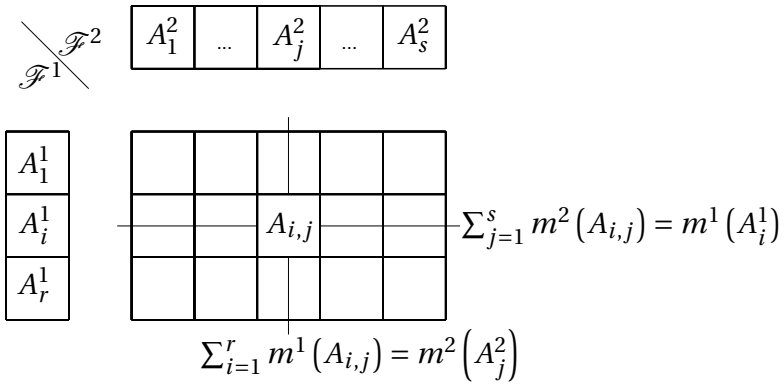


Figure 2.1: Schematic representation for $d = 2$ of condition (2.16): a joint mass assignment m must have marginals corresponding to the basic mass assignment m^1 and m^2 of the marginal RSs $\left(\mathscr{F}^1, m^1\right)$ and $\left(\mathscr{F}^2, m^2\right)$.

This is a good place to comment about the work of Berleant and coworkers (Berleant and Goodman-Strauss (1998); Berleant and Cheng (1998); Berleant and Zhang (2004a,b)) who developed a strategy in the same spirit as the one proposed by Fetz and Oberguggenberger (2004), which they called DEnv: Distribution Envelope determination. Given a finite random set, their approach maximizes and minimizes the probability represented by the left $\overline{F}$ and right $\underline{F}$ bounds of the associated probability box (see definition in Section 2.4.2) for every value of the abscissa.

2.4 Relationship between random set theory and probability, possibility, probability bound analysis and random interval analysis theories

Random sets can be understood as a generalization of CDFs, possibility distributions, intervals and probability boxes. In the following these relationships will be clarified.

2.4.1 Relationship between random sets and possibility distributions

For details about fuzzy sets and possibility theory, the reader is referred elsewhere (see e.g. Dubois and Prade (1988); Klir and Folger (1988); Nguyen and Walker (1996)). A *possibility distribution*[2] (that is a normalized fuzzy set) A of a set X is a mapping $A : X \to [0,1]$, where $\sup_{x \in X} A(x) = 1$. In this case, $A(x)$, for $x \in X$, represents the degree to which x is compatible with the concept represented by A. The α-cut of a membership function is represented by the crisp set $A_\alpha = \{ x \in X : A(x) \geq \alpha \}$ for $\alpha \in (0,1]$.

Let $(\mathscr{F}, \sigma_{\mathscr{F}})$ be a measurable space,where $\mathscr{F} \subseteq \mathscr{P}(X)$; if for every $B \in \mathscr{F}$ there exists a family of subsets $\mathscr{C}_B := \left\{ C : B \subseteq C \in \mathscr{F} \right\}$ such that $\mathscr{C}_B \in \sigma_{\mathscr{F}}$, then the function

$$c_\Gamma(B) := P_\Gamma(\mathscr{C}_B) = P_\Omega \{ \alpha : \Gamma(\alpha) \in \mathscr{C}_B \} = P_\Omega \{ \alpha : B \subseteq \Gamma(\alpha) \} \tag{2.18}$$

provides a measure on X for every $B \subseteq X$ called the *subset coverage function*. Here P_Γ was defined in section 2.2.2. In the particular case when $B = \{x\}$, then $\mathscr{C}_x = \mathscr{C}_{\{x\}} = \left\{ C : x \in C \in \mathscr{F} \right\}$ and (2.18) becomes $c_\Gamma(x) = P_\Omega \{ \alpha : x \in \Gamma(\alpha) \}$ for every $x \in X$ which defines the so called *one point coverage function* of the RS Γ. Note that according to (2.8), $c_\Gamma(x) = \mathrm{Pl}\{x\}$.

Let A be a normalized fuzzy set on X, and let $\tilde{\alpha} : \Omega \to (0,1], \alpha \mapsto \tilde{\alpha}(\alpha)$ be a uniformly distributed random variable on some probability space $(\Omega, \sigma_\Omega, P)$, i.e., $P\{\alpha : \tilde{\alpha}(\alpha) \leq z\} = z$ for $z \in (0,1]$. Then $\tilde{\alpha}$ induces a RS $\Gamma_A(\alpha) = \{ x \in X : A(x) \geq \tilde{\alpha}(\alpha) \}$, which is simply the randomized α-cut set $A_{\tilde{\alpha}(\alpha)}$. This is the way of representing a particular membership function (or possibility distribution) using a RS (see e.g. Goodman and Nguyen (2002)). The associated RS has the one point coverage function

$$\begin{aligned} c_{\Gamma_A}(x) &= P \{ \alpha : x \in \Gamma_A(\alpha) \} \\ &= P \{ \alpha : A(x) \geq \tilde{\alpha}(\alpha) \} = A(x) \end{aligned} \tag{2.19}$$

In words, Γ_A is a RS in X whose one point coverage function coincides with the membership function (possibility distribution) of the fuzzy set A. The one point coverage function in (2.19) defines the *possibility measure* $\mathrm{Pos}_A : X \to [0,1]$ given by

$$\begin{aligned} \mathrm{Pos}_A(K) &:= \sup_{x \in K} \left\{ c_{\Gamma_A}(x) \right\}, \\ &= \sup_{x \in K} \{ A(x) \} \end{aligned} \tag{2.20}$$

for $K \subseteq X$, since it follows from equation (2.19) that $P\{\alpha : \tilde{\alpha}(\alpha) \leq A(x)$ for some $x \in K\} = P\{\alpha : \tilde{\alpha}(\alpha) < \sup\{A(x) : x \in K\}\} = \sup\{A(x) : x \in K\}$. This is a

[2] According to (Klir, 2006, p. 147), the term "possibility distribution A of a set X" is a misnomer, since A does not distribute values between the elements of X; he recommends the name *basic possibility function*; however, we will continue using the expression possibility distribution, since it is the most common term found in the literature.

good point to remember that the *necessity measure* Nec_A is defined by $\mathrm{Nec}_A(K) := 1 - \mathrm{Pos}_A(K^c)$.

It is important to observe that the associated (generated) RS is *consonant*, i.e. nested, in the sense that $\Gamma_A = \{A_\alpha : 0 < \alpha \leq 1\}$ is totally ordered by set inclusion and that the corresponding membership function (possibility distribution) must be unimodal.

In correspondence with the above representation, Dubois and Prade (1991) proposed that a possibility distribution $A : X \to [0,1]$ can be *inner approximated* by a finite consonant random set $(\mathscr{F}_n, m)$ of n focal sets. Given a set $\{\alpha_1 = 1, \alpha_2, \alpha_3 \ldots, \alpha_{n+1} = 0 : \alpha_i > \alpha_j, i > j\}$ of α-cut levels, then the possibility distribution $A(x)$ is equivalent to the consonant random set $(\mathscr{F}_n, m)$ defined by

$$\mathscr{F} = \left\{ A_{\alpha_i} \;\middle|\; i = 1, \ldots, n \right\}$$

and

$$m\left(A_{\alpha_i}\right) = \begin{cases} \alpha_i - \alpha_{i-1} & i = 1, \ldots, n-1 \\ \alpha_n & i = n \end{cases}$$

where $A_{\alpha_i} := \{x : A(x) \geq \alpha_i\}$ is the α-level cut. Usually, the $\alpha-$levels are regularly distributed, so that, $m(A_i) = 1/n$. An *outer approximation* can be defined in an analogous way. In Chapter 7, we will deal with the issue of discretization and approximation of random sets.

2.4.2 Relationship between finite random sets and probability boxes

A *probability box* or *p-box* (term coined by Ferson and Hajagos (2002)) $\left\langle \underline{F}, \overline{F} \right\rangle$ is a class of cumulative distribution functions (CDFs) $\left\{ F : \underline{F} \leq F \leq \overline{F}, F \text{ is a CDF} \right\}$ delimited by upper and lower CDF bounds $\underline{F}$ and $\overline{F} : \mathbb{R} \to [0,1]$. This class of CDFs collectively represents the epistemic uncertainty about the CDF of a random variable. There is a close relationship between probability boxes and random sets. Every RS generates a unique p-box whose constituent CDFs are all those consistent with the evidence. In turn, every p-box generates an equivalence class of random intervals consistent with it (see e.g. Joslyn and Ferson (2004)). A p-box can always be discretized to obtain from it a finite random set that approximates it; this discretization is not unique, because it depends on the conditions applied; for example, Ferson et al. (2003b), Tonon (2004c) and Hall and Lawry (2004) have proposed different techniques to obtain an equivalent RS from a probability box. When $(\mathscr{F}, m)$ is a finite RS defined on $\mathbb{R}$, each $A_i \in \mathscr{F}$ is an interval. In this case, the belief and plausibility of the set $(-\infty, x]$ leads to two limit CDFs (Ferson et al. (2003b)),

$$\overline{F}(x) := \mathrm{Pl}_{(\mathscr{F},m)}((-\infty, x])$$
$$\underline{F}(x) := \mathrm{Bel}_{(\mathscr{F},m)}((-\infty, x])$$

which define the probability box $\left\langle \underline{F}, \overline{F} \right\rangle$ associated with the RS.

Given a probability box $\left\langle \underline{F}, \overline{F} \right\rangle$ such that $\underline{F}$ and $\overline{F}$ are piecewise continuous from the right, the quasi-inverses of $\underline{F}$ and $\overline{F}$ are defined respectively by,

$$\underline{F}^{(-1)}(\alpha) := \inf\left\{ x : \underline{F}(x) \geq \alpha \right\} \tag{2.21}$$

$$\overline{F}^{(-1)}(\alpha) := \inf\left\{ x : \overline{F}(x) \geq \alpha \right\} \tag{2.22}$$

for $\alpha \in (0,1]$. Given a probability box $\left\langle \underline{F}, \overline{F} \right\rangle$, a corresponding RS is given by the infinite RS with focal elements defined by (Joslyn and Ferson (2004))

$$\left\langle \underline{F}, \overline{F} \right\rangle^{-1}(\alpha) := \left[\overline{F}^{(-1)}(\alpha), \underline{F}^{(-1)}(\alpha) \right] \tag{2.23}$$

for all $\alpha \in (0,1]$. Joslyn and Ferson (2004) did not define the "basic mass assignment" associated to (2.23), but we will do it in Section 4.2.1.

Example of the use of probability boxes: Oberguggenberger and Fellin (2005, 2006) cite a list of the friction angles obtained in twenty direct shear tests, as follows:

$$\varphi[^\circ] = \begin{cases} 22, 23.2, 23.4, 24, 24, 24, 24.1, 24.3, 24.4, 24.9, \\ 25, 25.3, 25.5, 25.6, 26, 26.5, 27, 28.5, 29.5, 30. \end{cases} \tag{2.24}$$

Instead of choosing one CDF that fulfills a goodness-of-fit hyphothesis test (like the Kolmogorov-Smirnov (KS) test), Figure 2.2 shows a probability box that contains all the CDFs that do it, provided that the samples are independent and identically distributed.

It is important to comment that the KS test calculates distribution free bounds about the empirical CDF associated to the values φ. Note that as the number of samples tends to infinity, the probability box will tend to the empirical CDF. Also, note that these probability boxes are not sure bounds, but just represent statistical claims of the form: 80% of the time such bounds are constructed they will contain the true CDF.

Finally, a closer analysis to Figure 2.2 reveals that the Kolmogorov-Smirnov test is not very robust in verifying whether the empirical CDF fits the tails of the reference CDF; other goodness-of-fit tests might be used for this purpose. This is an issue that deserves further research.

2.4.3 Relationship between finite random sets and families of intervals

A single interval estimate can be regarded as RS with a unique element A with $m(A) = 1$. When a set of n intervals is available, every interval is considered to be a focal element A_i with a corresponding $m(A_i) = 1/n$. In the case that there

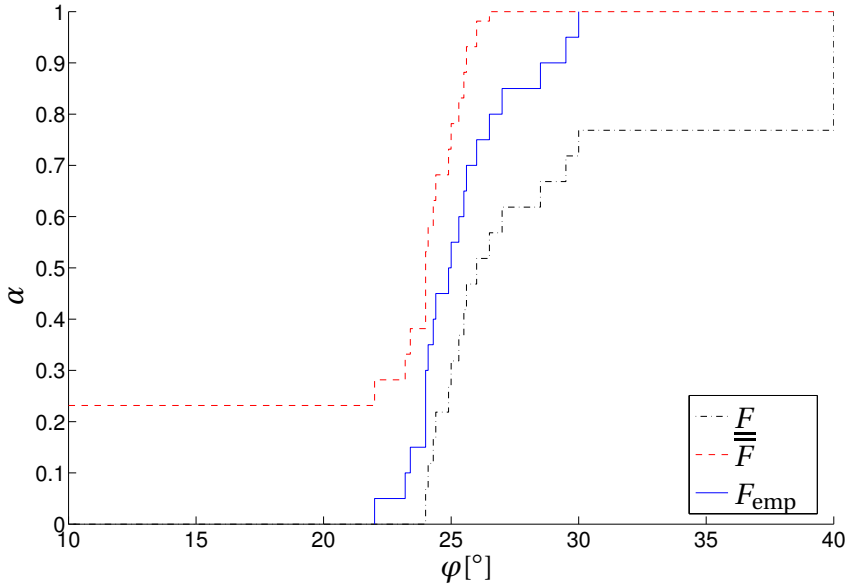


Figure 2.2: Probability box $\langle \underline{F}, \overline{F} \rangle$ that contains all the CDFs that accomplish the KS test for a confidence level of 80% and a sample φ given by (2.24). Here the empirical CDF F_{emp} corresponding to φ is also depicted and the tails of the probability box were truncated to reasonable limits.

is evidence supporting the fact that the occurrence of an interval is more probable than another, then the corresponding basic assignment might be modified accordingly. In conclusion, in this case we form a Dempster-Shafer body of evidence.

2.4.4 Relationship between finite random sets and probability density functions

A PDF $f_X(x)$ can be approximated by a histogram with n discrete intervals. Every interval can be interpreted as a focal set A_i with $m(A_i) = \int_{A_i} f_X(x)\,\mathrm{d}x$. When f_X has an unbounded domain, it is necessary to impose upper and lower bounds on the distribution, for instance, setting the bounds at the 0.005 and 0.995 percentiles of f_X.

2.5 Combination of different sources of information for finite random sets

When two different experts provide different bodies of evidence (say $(\mathscr{F}_r, m_r)$ and $(\mathscr{F}_s, m_s)$) which contain their opinion about a particular situation, it could be necessary to mix these sets into a new body of evidence $\left(\mathscr{F}_{\text{joint}}, m_{\text{joint}}\right)$ which combines their claims. Several authors have proposed different strategies to mix different bodies of evidence; Sentz and Ferson (2002) provide a comprehensive

reference on this topic, however they conclude that this is still an open and unsolved problem within evidence theory. Among the methods they cite, we have the Dempster's rule of combination, which we present in the following.

Dempster's rule of combination. This rule, suggested by Shafer (1976), combines the information using the following formulas:

$$\mathscr{F}_{\text{joint}} := \mathscr{F}_r \cup \mathscr{F}_s$$

and

$$m_{\text{joint}}(A_i) := \frac{\sum_{A_r \cap A_s = A_i} m_r(A_r) m_s(A_s)}{1 - \sum_{A_r \cap A_s = \emptyset} m_r(A_r) m_s(A_s)} \tag{2.25}$$

for all $A_i \neq \emptyset$, $A_r \in \mathscr{F}_r$, $A_s \in \mathscr{F}_s$ and $m_{\text{joint}}(\emptyset) = 0$. In equation (2.25), the second term of the denominator accounts for contradiction or conflict among the information given by independent knowledge sources. Therefore, Dempster's rule disregards every contradiction by standarizing with the complementary degree of contradiction, because it is designed to use consistent opinions from different sources as much as possible. However, according to Zadeh (1984), this normalization can cause a counterintuitive and numerically unstable mixing when the given information from independent sources has a lot of contradictions or conflicts. In consequence, Dempster's rule can be appropriate in a situation where there is some consistency and sufficient agreement among the opinions of different sources (Sentz and Ferson (2002)).

2.6 Application

In practice the method is applied following the next steps:

- Obtain information of the basic variables from multiple sources and convert it to random sets. This information can be given for example in the way of families of intervals, PDFs, possibility distributions, fuzzy sets, parameterized families of PDFs, or probability boxes. This conversion can be done using the techniques presented in section 2.4.
- Combine the different bodies of evidence for each uncertain parameter, using the combination rules like the ones referred in section 2.5.
- Construct the joint random set of basic variables, by using the techniques presented in section 2.2.4. In this step the knowledge of dependence, unknown dependence or independence among the variables must be considered.
- Use the extension principle explained in section 2.2.5 to propagate the input random set through the system solver g.
- Estimate the belief and plausibility functions of the propagated DS structure.

- Estimate the intervals where the probability of the event of interest F resides.

It must be noted that the gap between plausibility and belief can be reduced by employing additional assumptions. However, one must keep in mind that without justifying the assumptions by means of evidence or data, the result would be just a consequence of these suppositions.

Tonon (2004c) contains a nice example on the application of the method. The reader is referred to this paper for additional information.

2.7 Imprecise probability theories

"Imprecise probability theories" is a general term used to denote all theories of uncertainty that are useful to measure the indeterminacy present in a given uncertainty problem when there is either ignorance, or insufficient information or both in a) the description of the basic variables present in the problem, and b) the relation between those variables or other factors that define the problem at hand. For example, random set and evidence theory are two techniques under the term imprecise probabilities. Other theories subsumed under this name are the theory of coherent lower and upper probabilities (see e.g. Walley (1991, 1996)), the theory of coherent lower and upper previsions (see e.g. Walley (1991, 2000)), the theory of closed convex sets of probability distributions (see e.g. Kyburg (1987)), the theory of feasible inter-valued probability distributions (see e.g. Weichselberger and Pohlmann (1990); Pan and Klir (1997)), the theory of fuzzy randomness (see e.g. Möller and Beer (2004)), the theory of interval probabilities (see e.g. Weichselberger (2000, 2001)), among others. Here I will make a brief account of the theories of coherent lower and upper probabilities and previsions.

The central idea of the theory of sets of probability measures (or theory of *credal sets*) is to relax the precision inherent in a probabilistic model, and to replace a single probability measure by a set of probability measures $\mathscr{P}$, which is closed and convex. This credal set $\mathscr{P}$ defines the *coherent lower and upper probabilities* of the set F according to the formulas (see e.g. Dempster (1967); Walley (1991))

$$\underline{P}(F) = \inf_{P \in \mathscr{P}} P(F) \tag{2.26}$$

$$\overline{P}(F) = \sup_{P \in \mathscr{P}} P(F) \tag{2.27}$$

respectively. That is, this theory specifies an interval of probabilities $\left[\underline{P}(F), \overline{P}(F)\right]$ where $P(F)$ could reside. The coherent upper and lower probabilities are dual measures, i.e., $\underline{P}(F) = 1 - \overline{P}(F^{\mathrm{c}})$

Another theory closely related is the one of upper and lower expectations. Consider a loss function l and a probability measure $P \in \mathscr{P}$, where $\mathscr{P}$ is a closed and

convex credal set. The value of the expected loss is given by

$$\mathrm{E}[l] = \int l(x)\mathrm{d}P(x).$$

The *lower and upper expectations*, given by (see e.g. Dempster (1967))

$$\underline{\mathrm{E}}[l] = \inf_{P\in\mathscr{P}} \mathrm{E}[l] \tag{2.28}$$

$$\overline{\mathrm{E}}[l] = \sup_{P\in\mathscr{P}} \mathrm{E}[l], \tag{2.29}$$

provide bounds for the expected loss. Note in addition that $\underline{\mathrm{E}}[l] = -\overline{\mathrm{E}}[-l]$.

2.8 Some open problems

Up to the author's knowledge, the developments in random set theory in the field of risk, reliability and uncertainty analysis have been done in the framework of a *finite* random sets. Random sets with an infinite number of focal elements have not been employed in those fields yet. In their papers, many authors have expressed the feeling that this is the direction to go, by making remarks such that the precision in the calculations will increase if a larger number of focal elements or equivalently, more granular discretizations of the basic variables are employed. Note that the discretization step brings errors due to the loss of information, inasmuch as usually the information available is cast in a continuous fashion.

Another area that requires special attention is the development of strategies to evaluate the reliability of engineering systems using random sets and requiring relatively low computational resources. The actual methods have the curse of dimensionality and require a large computational work, making them inefficient when employed in high dimensional systems.

Combination of evidence in random set and imprecise probability theories is still an open issue. Some achievements have been obtained (see e.g. Sentz and Ferson (2002)), but up to now, there is no definitive method to mix different sources of information. This is an important issue, inasmuch as the results of an uncertainty analysis may strongly depend on which combination method is employed. It is important to note that whatever is the combination method employed, its selection must depend on the context of the information to be combined.

As discussed in Section 1.1, model uncertainty and scenario abstraction uncertainty are basically kinds of epistemic uncertainty. It is required to develop methods to deal with these problems using random sets. In fact, according to Dodagoudar and Venkatachalam (2000) and Hall and Lawry (2003) model uncertainty may be more appropriately deal with by a fuzzy representation rather than being represented probabilistically.

The main topic of this dissertation is the analysis of random sets with an infinite number of focal elements. As will be seen in the sequel, this generalization

produces very important theoretical results, closes several important open questions and regards Dempster-Shafer evidence theory to be just a particularization of infinite RSs.

Chapter 3

Copulas

The following is a brief presentation of some key points on the theory of copulas. For a proof of the statements made, the reader is referred to the standard reference on the topic, Nelsen (1999), and the publications therein cited. Also, Section 3.6 contains an original proof of the fact that the set of all copulas $\mathscr{C}$ is compact.

3.1 Preliminary concepts

In this section, concepts like quasi-inverses, H-volumes, increasing functions and multivariate cumulative distribution functions are introduced.

3.1.1 Quasi-inverses

Quasi-inverses are generalizations of the inverse of a function, which are defined even when the function has jump discontinuities. Let F be a nondecreasing function defined on the closed interval $[a,b]$ and let $\bar{\mathbb{R}}$ denote the extended real line, $[-\infty,\infty]$. Then, a *quasi-inverse* of F is any function $F^{(-1)}$ defined on $[F(a),F(b)]$ such that,

- if t is in the range of F, then $F^{(-1)}(t)$ is any number $x \in \mathbb{R}$ such that $F(x) = t$. In other words, for all t in the range of F, $F\left(F^{(-1)}(t)\right) = t$.
- if t is *not* in the range of F, then $F^{(-1)}(t) = \inf\{x \mid F(x) \geq t\} = \sup\{x \mid F(x) \leq t\}$.

When F is a strictly monotonically increasing function, then it has a single quasi-inverse which coincides with the classical concept of the inverse of a function, and which is denoted as F^{-1}.

3.1.2 V_H-volume

Let H be a mapping $H : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^d$. Let $B = \times_{i=1}^{d} [a_i, b_i]$ be a d-box all of whose vertices are in S. The *H-volume of B*, $V_H(B)$, is the *d-th order difference of H on B*,

$$V_H(B) = \Delta_{a_d}^{b_d} \Delta_{a_{d-1}}^{b_{d-1}} \cdots \Delta_{a_1}^{b_1} H \tag{3.1}$$

where the *d-first order differences of H* are defined as,

$$\Delta_{a_k}^{b_k} H = H(t_1, \ldots, t_{k-1}, b_k, t_{k+1}, \ldots, t_d) - H(t_1, \ldots, t_{k-1}, a_k, t_{k+1}, \ldots, t_d) \tag{3.2}$$

for $k = 1, \ldots, d$. More explicitly

$$V_H(B) = \sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1 + \ldots + i_d} H\left(x_{1 i_1}, \ldots, x_{d i_d}\right) \tag{3.3}$$

where $x_{j1} = a_j$ and $x_{j2} = b_j$ for all $j = 1, \ldots, d$ (see e.g. Embrechts et al. (2002)).

For example, when $d = 2$, then $B = [a_1, b_1] \times [a_2, b_2]$ and the H-volume of the rectangle B is given by the second order difference of H on B,

$$V_H(B) = H(a_2, b_2) - H(a_1, b_2) - H(a_2, b_1) + H(a_1, b_1) \tag{3.4}$$

3.1.3 d-increasing functions

Let be a mapping $H : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^d$. H is said to be *d-increasing* if $V_H(B) \geq 0$ for all d-boxes B whose all vertices lie in S.

3.1.4 Multivariate cumulative distribution functions

A d-dimensional cumulative distribution function (d-dimensional CDF), is a function $H : \bar{\mathbb{R}}^d \to [0, 1]$ that satisfies the following properties:

1. H is d-increasing,
2. $H(\boldsymbol{t}) = 0$ for all $\boldsymbol{t} = [t_1, \ldots, t_k, \ldots, t_d]$ in $\bar{\mathbb{R}}^d$ whenever $t_k = -\infty$ for at least one $k = 1, \ldots, d$ and,
3. $H(\infty, \ldots, \infty) = 1$.

3.1.5 Support of a cumulative distribution function

Each joint CDF $H : \bar{\mathbb{R}}^d \to [0, 1]$ induces a probability measure in $\mathbb{R}^d$ given by V_H. The *support of a CDF $F(x)$*, is the complement of the union of all open subsets of $\mathbb{R}^d$ with null V_H-measure.

3.1.6 k-margins

Let $H : S_1 \times S_2 \times \cdots \times S_d \to [0,1]$ be a joint CDF. The function

$$H_j(x) = H\left(\sup S_1, \ldots, \sup S_{j-1}, x, \sup S_{j+1}, \ldots, \sup S_d\right) \tag{3.5}$$

for all $x \in S_j$ is a univariate margin of H, or simply, *margin*. Higher dimensional margins (*k-margins*, $k \geq 2$) are defined by fixing less positions in (3.5).

3.2 Copulas

Copulas are important in probability because they contain all the dependence information between the involved random variables. In this section the definition of copulas and some of their properties will be given.

Definition 3.2.1. *A* copula *is a multivariate CDF $C : [0,1]^d \to [0,1]$ such that each of its marginal CDFs is uniform on the interval $[0,1]$. Alternatively, a copula is any function C with domain $[0,1]^d$ which fulfills the following three conditions:*

1. *C is* grounded, *i.e., for every $\boldsymbol{\alpha} \in [0,1]^d$, $C(\boldsymbol{\alpha}) = 0$ if there exists $i \in \{1,2,\ldots,d\}$ with $\alpha_i = 0$.*
2. *If all coordinates of $\boldsymbol{\alpha} \in [0,1]^d$ are 1 except α_i then $C(\boldsymbol{\alpha}) = \alpha_i$.*
3. *C is d-*increasing, *i.e., for all $[a_1,\ldots,a_i,\ldots,a_d], [b_1,\ldots,b_i,\ldots,b_d] \in [0,1]^d$ such that $a_i \leq b_i$ for $i = 1,\ldots,d$ we have that $V_C(\times_{i=1}^{d}[a_i,b_i]) \geq 0$.*

For example, when $d = 2$, then $B = [a_1,b_1] \times [a_2,b_2]$ and according to (3.4) the condition becomes

$$C(b_1,b_2) - C(a_1,b_2) - C(a_2,b_1) + C(a_1,a_2) \geq 0 \tag{3.6}$$

Observe that C is a special case of a multivariate CDF. It can be shown that for every d-dimensional copula C with $d \geq 3$, each k-margin of C is a k-dimensional copula with $2 \leq k < d$.

It is important to say that the set $\mathscr{C}$ of all copulas in $\mathbb{R}^d$ is convex, inasmuch as $\lambda C' + (1-\lambda)C''$ is also a copula for any $C', C'' \in \mathscr{C}$ and any $\lambda \in [0,1]$.

As a consequence of the Lebesgue's decomposition theorem, any d-dimensional copula can be expressed as,

$$C(u_1,\ldots,u_d) = A_C(u_1,\ldots,u_d) + S_C(u_1,\ldots,u_d)$$

where

$$A_C(u_1,\ldots,u_d) = \int_0^{u_1} \cdots \int_0^{u_d} \frac{\partial^d}{\partial u_1 \cdots \partial u_d} C(u_1,\ldots,u_d)\, \mathrm{d}s_1 \ldots \mathrm{d}s_d$$

and

$$S_C(u_1,\ldots,u_d) = C(u_1,\ldots,u_d) - A_C(u_1,\ldots,u_d).$$

When $C = A_C$ on $[0,1]^d$ then C is said to be *absolutely continuous*; in this case C has corresponding joint probability density function $\frac{\partial^d}{\partial u_1 \cdots \partial u_d} C(u_1,\ldots,u_d)$. Now when $C = S_C$ on $[0,1]^d$ then C is said to be *singular* and $\frac{\partial^d}{\partial u_1 \cdots \partial u_d} C(u_1,\ldots,u_d) = 0$ almost everywhere in $[0,1]^d$. Otherwise, C has an *absolute continuous component* A_C and a *singular component* S_C.

When the support of C is $[0,1]^d$, then we say that the copula has *full support*, also when C is singular, its support has Lebesgue measure zero and conversely; however a copula can have full support without being absolutely continuous.

It will be understood after defining the μ_C measure in Section 5.3.1 that, unlike multivariate CDFs in general, the margins of the copulas are continuous and in consequence a copula does not have individual points in $[0,1]^d$ with a positive μ_C measure.

Copulas are uniformly continuous functions on $[0,1]^d$, as explained by the following theorem:

Theorem 3.2.2. *Let $C : [0,1]^d \to [0,1]$ be a d-dimensional copula. Then, for every $\boldsymbol{u} = [u_1,\ldots,u_k,\ldots,u_d]$ and $\boldsymbol{v} = [v_1,\ldots,v_k,\ldots,v_d]$ in $[0,1]^d$*

$$|C(\boldsymbol{v}) - C(\boldsymbol{u})| \leq \sum_{k=1}^{d} |v_k - u_k|$$

Hence C is uniformly continuous on $[0,1]^d$.

Proof. The proof is contained in Nelsen (1999, p. 40). □

In applications, sometimes the copula is given; sometimes it is unknown, and sometimes, when sufficient observations are available, it is possible to generate *empirical copulas* from the data. This procedure is equivalent to the estimation of a CDF in the univariate case when enough samples are given. For details on empirical copulas, the reader is referred to Nelsen (1999).

3.3 Sklar's theorem

The word copula was introduced in Sklar (1959), although similar ideas and results can be traced back to Hoeffding (1940). Copulas are functions that link marginal CDFs to form the joint CDF of a set of random variables. This relationship is given by Sklar's theorem:

Theorem 3.3.1 (Sklar's theorem)**.** *Let $X_1, X_2, \ldots, X_d$ be random variables with CDFs F_{X_i} for $i = 1, \ldots, d$ and coupled through the joint CDF $F_{X_1 X_2 \ldots X_d}$. Then, there exists a copula C such that*

$$F_{X_1 X_2 \ldots X_d}(\boldsymbol{x}) = C\left(F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_d}(x_d)\right) \tag{3.7}$$

for all $\boldsymbol{x} := [x_1, x_2, \ldots, x_d] \in \mathbb{R}^d$. If all marginal CDFs F_{X_i} for $i = 1, \ldots, d$ are continuous then C is unique; otherwise, C is uniquely determined on $\text{range}(F_1) \times \text{range}(F_2) \times \cdots \times \text{range}(F_d)$. *Conversely, if C is a copula and F_{X_i} for $i = 1, \ldots, d$ are one-dimensional CDFs, then the function $F_{X_1 X_2 \ldots X_d}$ defined by (3.7) is a d-dimensional CDF with marginals F_{X_i} for $i = 1, \ldots, d$.*

Proof. The proof for the two dimensional case is contained in Sklar (1959) while the demonstration for the multivariate case appears in Sklar (1996). □

The inverse relation between the copula and the joint CDF is given by

$$C(\boldsymbol{u}) = F_{X_1 X_2 \ldots X_d}\left(F_{X_1}^{(-1)}(u_1), F_{X_2}^{(-1)}(u_2), \ldots, F_{X_d}^{(-1)}(u_d)\right)$$

for any $\boldsymbol{u} := [u_1, u_2, \ldots, u_d] \in [0,1]^d$, where $F_{X_i}^{(-1)}$ is the quasi-inverse of F_{X_i} for every $i = 1, \ldots, d$.

Note that from Sklar's theorem it can be deduced that for continuous joint CDFs, the univariate margins and the multivariate dependence structure can be separated, and that the dependence structure can be represented by a copula.

A famous copula is the *product copula* $\Pi_d : [0,1]^d \to [0,1]$, defined by

$$C(\boldsymbol{u}) = \Pi_d(\boldsymbol{u}) = \prod_{k=1}^{d} u_k,$$

which models independence of the random variables in consideration (see Theorem 3.4.2).

3.4 The Frechét-Hoeffding bounds

The Frechét-Hoeffding bounds impose upper and lower limits to any copula. In the following they will be described in their two and multivariate version, and then some of their properties will be summarized.

First, we will describe the Frechét-Hoeffding bounds in their two dimensional form:

Theorem 3.4.1. *All two dimensional copulas satisfy the inequality*

$$W(u, v) \leq C(u, v) \leq M(u, v) \tag{3.8}$$

for all $(u,v) \in [0,1] \times [0,1]$ *where* W *and* M *stand for the copulas*

$$W(u,v) = \max(u+v-1,0) \tag{3.9}$$
$$M(u,v) = \min(u,v) \tag{3.10}$$

which are known as the lower and upper Frechét-Hoeffding bounds *respectively.*

Proof. (Taken from Nelsen (1999, p. 8)) Let (u,v) be an arbitrary point in $[0,1] \times [0,1]$. Since $C(u,v) \leq C(u,1) = u$ and $C(u,v) \leq C(1,v) = v$ we have that $C(u,v) \leq \min(u,v)$. For proving the lower bound we make use of the inequality (3.6) with $u_1 = 1$, $v_1 = 1$, $u_2 = u$, and $v_2 = v$, that is, $C(u,v) - C(u,1) - C(1,v) + C(1,1) \geq 0$, which implies, $C(u,v) \geq u + v - 1$, and since $C(u,v) \geq 0$, it follows that $C(u,v) \geq \max(u+v-1,0)$. □

Note that the Frechét-Hoeffding bounds (3.8) can be rewritten in terms of the univariate marginal F_X, F_Y and the joint CDF F_{XY} using (3.9) and (3.10) as

$$\max\big(F_X(x) + F_Y(y) - 1, 0\big) \leq F_{XY}(x,y) \leq \min\big(F_X(x), F_Y(y)\big);$$

this expression gives bounds of the joint CDF in terms of the marginals.

When two random variables are perfectly dependent, their dependence function is represented by the copula M. Analogous statements are valid with opposite dependence and the copula W. The product copula, plays the same role modeling the independence.

The two dimensional copulas M, Π and W can be extended respectively to d dimensional formulas M_d, Π_d and W_d given by:[1]

$$W_d(\boldsymbol{u}) = \max(u_1 + u_2 + \ldots + u_k - d + 1, 0)$$
$$\Pi_d(\boldsymbol{u}) = u_1 u_2 \cdots u_d$$
$$M_d(\boldsymbol{u}) = \min(u_1, u_2, \ldots, u_d)$$

In this case, the functions M_d and Π_d are copulas for $d \geq 2$, whereas the function W_d is only a copula for $d = 2$ and it fails to be a copula for any $d \geq 3$. In any case, there is a d-dimensional version of the Frechét-Hoeffding bounds

$$W_d(\boldsymbol{u}) \leq C(\boldsymbol{u}) \leq M_d(\boldsymbol{u}) \tag{3.11}$$

for every $\boldsymbol{u} \in [0,1]^d$.

Even though the Frechét-Hoeffding lower bound W_d is not a copula for $d \geq 3$, the left hand inequality of (3.11) is best possible in the sense that for any $n \geq 3$ and any $\boldsymbol{u} \in [0,1]^d$, there is a copula C such that $W_d(\boldsymbol{u}) = C(\boldsymbol{u})$.

The following theorem states some important properties of multivariate copulas:

[1] Note that W, Π and M is a shortcut for W_2, Π_2 and M_2.

Theorem 3.4.2. *If $X_1, X_2, \ldots, X_d$ are continuous random variables, then*

- *$X_1, X_2, \ldots, X_d$ are independent if and only if their d-dimensional copula is Π_d*
- *each of the random variables $X_1, X_2, \ldots, X_d$ is almost surely a strictly increasing function of any of the others if and only if the d-dimensional copula of $X_1, X_2, \ldots, X_d$ is M_d.*

Proof. The proof is contained in (Nelsen, 1999, p. 43). □

3.5 Binary operations between random variables and copulas

Let L be a binary operation between two variables x and y. If X and Y are two random variables with marginal CDFs F_X and F_Y and related through the copula C, then the Lebesgue-Stieltjes integral

$$F_Z(z) = \int_{L\{z\}} \mathrm{d}C\left(F_X(x), F_Y(y)\right) \tag{3.12}$$

with

$$L\{z\} = \left\{(x,y) \;\middle|\; x, y \in \mathbb{R}, L(x,y) \leq z\right\}$$

gives the distribution of the operation $Z = L(X,Y)$ in terms of the dependence (copula) function C between X and Y.

When the copula C is unknown, an estimate of $F_Z(z)$ in (3.12) can be given in terms of what is known as *dependency bounds analysis* (see e.g. Williamson (1989)). This technique gives the lower and upper bounds ($\underline{F}_Z$ and $\overline{F}_Z$) on the CDF F_Z when only the marginal CDFs F_X and F_Y are known, that is, it provides the probability box $\langle \underline{F}_Z, \overline{F}_Z \rangle$ which contains F_Z. Note that both $\underline{F}_Z$ and $\overline{F}_Z$ depend on F_X, F_Y and L.

3.6 The set of all copulas $\mathscr{C}$ is compact

In this section we will denote by X a subset of $\mathbb{R}^d$, by Y a subset of $\mathbb{R}$, by $C(X,Y)$ the set of all continuous functions $f : X \to Y$, by $B(X,Y)$ the set of all bounded functions $f : X \to Y$ and by $C_b(X,Y)$ the set of all continuous bounded functions $B(X,Y) \cap C(X,Y)$. Note that these last two sets are metric spaces equipped with the distance function

$$d(f,g) = \sup_{x \in X} |f(x) - g(x)| \tag{3.13}$$

where $f, g \in B(X,Y)$ or $C_b(X,Y)$ respectively.

A set of functions $\Phi \in C(X,Y)$ is said to be *equicontinuous* if for every $\epsilon > 0$ and every $x \in X$, there exists a $\delta > 0$, such that for all $\phi \in \Phi$ and all $x' \in X$ with $\left|x' - x\right| < \delta$ we have $\left|\phi(x) - \phi(x')\right| < \epsilon$. Now, $\Phi \in C(X,Y)$ is said to be *uniformly equicontinuous* if for every $\epsilon > 0$ there exists a $\delta > 0$, such that for all $\phi \in \Phi$ and all $x, x' \in X$ with $\left|x' - x\right| < \delta$ we have $\left|\phi(x) - \phi(x')\right| < \epsilon$.

A metric space X is said to be *totally bounded* (or *pre-compact*) if and only if for any $\epsilon > 0$, there is a set of points $S := \{x_1, x_2, \ldots, x_n\}$ of X such that $X \subseteq \cup_{i=1}^{n} B_\epsilon(x_i)$, where $B_\epsilon(x_i)$ denotes an open ball about x_i with radius ϵ.

In the following I will list some theorems that are required to show that the set of all copulas is compact (see e.g. Rudin (1973); Kolmogorov and Fomin (1970)).

Theorem 3.6.1 (Ascoli's theorem)**.** *Suppose X is a compact space and $\Phi \subseteq C(X,Y)$ is* pointwise bounded*, i.e., $\sup_{x \in X}\{|f(x)|\} < \infty$ for all $f \in \Phi$, and* equicontinuous*. Then Φ is totally bounded in $C(X,Y)$.*

Theorem 3.6.2. *A metric space is compact if and only if it is totally bounded and complete.*

Theorem 3.6.3. *A subspace of a complete space is complete if and only if it is closed.*

Theorem 3.6.4. *If X is a metric space and Y is a complete metric space, then $(B(X,Y), d)$ is a complete metric space.*

Theorem 3.6.5. *The set of all copulas $\mathscr{C}$ is totally bounded under the metric d defined by equation* (3.13).

Proof. It is clear that both $[0,1]^d$ and $[0,1]$ are compact. Also since $\mathscr{C} \subset C_b\left([0,1]^d, [0,1]\right)$ is immediate that $\mathscr{C}$ is pointwise bounded. In addition, according to Theorem 3.2.2, all copulas $C \in \mathscr{C}$ satisfy the Lipschitz condition

$$|C(\boldsymbol{u}) - C(\boldsymbol{v})| \leq \sum_{i=1}^{d} |u_k - v_k|$$

where $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$. Thus, all copulas are uniformly continuous on $[0,1]^d$, and $\mathscr{C}$ is (uniformly) equicontinuous. Therefore, according to Ascoli's theorem, the set of all copulas $\mathscr{C}$ is totally bounded. □

Theorem 3.6.6. *The set of all copulas $\mathscr{C}$ is complete under the metric d defined by equation* (3.13).

Proof. Let $\{C_m\}$ be any Cauchy sequence in $\mathscr{C}$. Then, given any $\epsilon > 0$, there is a N such that for all $m, n > N$ we have that

$$d(C_m, C_n) = \max_{\boldsymbol{\alpha} \in [0,1]^d} |C_m(\boldsymbol{\alpha}) - C_n(\boldsymbol{\alpha})| < \epsilon, \tag{3.14}$$

this implies that for each $\boldsymbol{\alpha} \in [0,1]^d$, $|C_m(\boldsymbol{\alpha}) - C_n(\boldsymbol{\alpha})| < \epsilon$ whenever $m,n > N$. Thus, for each $\boldsymbol{\alpha} \in [0,1]^d$, $\{C_m(\boldsymbol{\alpha})\}$ is a Cauchy sequence of real numbers that converges to $C(\boldsymbol{\alpha})$, inasmuch as $\mathbb{R}$ is complete. Since ϵ is independent of the selection of $\boldsymbol{\alpha} \in [0,1]^d$, then this convergence is uniform.

Making $n \to \infty$ in (3.14), we have that for each $\boldsymbol{\alpha} \in [0,1]^d$,

$$|C_m(\boldsymbol{\alpha}) - C(\boldsymbol{\alpha})| < \epsilon \tag{3.15}$$

whenever $m > N$. This shows that C_m converges uniformly to C on $[0,1]^d$. Finally, we have to show that C is a copula. Therefore, we have to verify that C fulfills the three conditions of Definition 3.2.1. Properties (1) and (2) are obvious. To show property (3), remember that according to equations (3.1) and (3.2),

$$V_{C_m}(B) = \Delta_{a_d}^{b_d} \Delta_{a_{d-1}}^{b_{d-1}} \cdots \Delta_{a_1}^{b_1} C_m$$

where $B := \times_{k=1}^{d} [a_k, b_k]$ is some d-box in $[0,1]^d$ and,

$$\Delta_{a_k}^{b_k} C_m = C_m(t_1, \ldots, t_{k-1}, b_k, t_{k+1}, \ldots, t_d) - C_m(t_1, \ldots, t_{k-1}, a_k, t_{k+1}, \ldots, t_d)$$

for $k = 1, \ldots, d$ and for all $\boldsymbol{\alpha} := [a_1, a_2, \ldots, a_d]$ and $\boldsymbol{\beta} := [b_1, b_2, \ldots, b_d] \in [0,1]^d$ with $\boldsymbol{\alpha} \leq \boldsymbol{\beta}$. The convergence of $\Delta_{a_k}^{b_k} C_m$ to $\Delta_{a_k}^{b_k} C$ follows from (3.15). In consequence $V_{C_m}(B)$ converges to $V_C(B)$ as $m \to \infty$. Now, $V_{C_m}(B) \geq 0$ for all $m = 1, 2, \ldots$ and all d-boxes B in $[0,1]^d$ since all C_m's are copulas. But then this implies that $V_C(B) \geq 0$; in consequence C is also a copula, i.e., $C \in \mathscr{C}$. This proves the completeness of $\mathscr{C}$. □

Corollary 3.6.7. *The set of all copulas $\mathscr{C}$ is closed under the metric d defined by equation* (3.13).

Proof. This follows directly from the application of Theorems 3.6.3, 3.6.4 and 3.6.6 inasmuch as $\mathscr{C} \in B([0,1]^d, [0,1])$. □

Here is the central result of this section,

Theorem 3.6.8. *The set of all copulas $\mathscr{C}$ is compact under the metric d defined by equation* (3.13).

Proof. This follows directly from the application of Theorems 3.6.2, 3.6.5 and 3.6.6. □

Note that the fact that the set of all copulas is compact under the metric (3.13) implies that if a sequence of copulas $\{C_k\}$ converges pointwise, then this sequence converges to a copula C; this convergence is uniform everywhere on $[0,1]^d$.

Part II

From finite to infinite random sets

Chapter 4

Infinite random sets

> ...I commenced with the simplest and most general truths, and thus each truth discovered was a rule available in the discovery of subsequent ones.
>
> René Descartes - A Discourse on the Method

4.1 Introduction

The main aim of this section is to suggest a simulation method for the calculation of the lower and upper probability bounds in the case of infinite random sets and the application of the method to the computation of the bounds for the probability of an event F when the uncertainty is expressed by aleatory and epistemic uncertainty in the representation of the basic variables. Using RS theory, basic variables can be represented as *a*) possibility distributions, *b*) probability boxes, *c*) families of intervals provided by experts or *d*) CDFs, which are in a posterior step converted to a finite RS representation. The drawback of the conversion to a finite RS lies in that some information is lost or modified in the process of conversion, and that coarser discretizations tend to alter more the information than finer discretizations. This is one strong motivation to represent a basic variable as an infinite RS because in this process there will be no loss or modification of the information.

We will see that the proposed method introduces a new geometrical representation of the space of basic variables, that we will call the α-space. This method is an appropriate technique to model the bounds of the probability of failure of structural systems when there is parameter uncertainty in the representation of the basic variables.

The plan of this section is as follows: Section 4.2 introduces the relationship between basic variables and infinite random sets, showing that possibility distributions, probability boxes, families of intervals provided by experts and the CDFs are just special cases of an infinite random set. Section 4.3 explains a Monte Carlo simulation method that can be useful to evaluate the upper and lower probability

integrals derived in Section 4.2. A benchmark example is used in Section 4.4 to demonstrate the advantages and differences of the proposed method compared with the finite approach. Finally, Section 4.5 contains the proofs of several statements made in this chapter.

4.2 Simulation techniques applied to the evaluation of lower and upper probability measures of functionally propagated infinite random sets

Integral (1.1) defines the probability of an event F. Note that this integral can also be written as

$$\begin{aligned} P_{\boldsymbol{X}}(F) &= \int_X I[\boldsymbol{x} \in F] f_{\boldsymbol{X}}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \\ &= \int_X I[\boldsymbol{x} \in F] \, \mathrm{d}F_{\boldsymbol{X}}(\boldsymbol{x}) \\ &= \int_X I[\boldsymbol{x} \in F] \, \mathrm{d}P_{\boldsymbol{X}}(\boldsymbol{x}) \\ &= \mathrm{E}_{\boldsymbol{X}}[I[\boldsymbol{x} \in F]] \end{aligned}$$

provided F is a $P_{\boldsymbol{X}}$-measurable set; here $F_{\boldsymbol{X}}(\boldsymbol{x}) = P_{\boldsymbol{X}}(\boldsymbol{X} \leq \boldsymbol{x})$ and $f_{\boldsymbol{X}}$ is the probability density function (PDF) associated to $F_{\boldsymbol{X}}$. The advantages of the RS representation for the analysis of uncertainty have been discussed in the present document; this representation will allow us to compute, subject to the limitations in the knowledge about the basic variables, the lower and upper bounds on the probability of the event F, $P_{\boldsymbol{X}}(F)$, that is using (2.9), $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) \leq P_{\boldsymbol{X}}(F) \leq \mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$ where

$$\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) = P_\Gamma\left(\gamma : \gamma \subseteq F, \gamma \in \mathscr{F}\right) = \int_{\mathscr{F}} I\left[\gamma \subseteq F\right] \mathrm{d}P_\Gamma(\gamma), \tag{4.1}$$

$$P_{\boldsymbol{X}}(F) = P_{\boldsymbol{X}}(\boldsymbol{x} : \boldsymbol{x} \in F, \boldsymbol{x} \in X) = \int_X I[\boldsymbol{x} \in F] \, \mathrm{d}P_{\boldsymbol{X}}(\boldsymbol{x}), \tag{4.2}$$

$$\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F) = P_\Gamma\left(\gamma : \gamma \cap F \neq \emptyset, \gamma \in \mathscr{F}\right) = \int_{\mathscr{F}} I\left[\gamma \cap F \neq \emptyset\right] \mathrm{d}P_\Gamma(\gamma) \tag{4.3}$$

provided $I\left[\gamma \subseteq F\right]$ and $I\left[\gamma \cap F \neq \emptyset\right]$ are P_Γ-measurable functions of γ. Here γ denotes a focal element of the joint RS $(\mathscr{F}, P_\Gamma)$, which contains the information about the basic variables. The evaluation of integrals (4.1) and (4.3) is not straightforward, so it is better to look for another representation of those integrals. If every focal element could be represented as a single point, then the evaluation of (4.1) and (4.3) could be easier and in addition the available methods used to solve (4.2) could be employed.

In the RS formulation, every basic variable is represented by a RS defined on the real line and whose focal sets are composed of intervals or even points. As will be shown in the following, to every focal element of an *infinite* RS defined on the

real line and represented by possibility distributions, probability boxes, families of intervals or CDFs, one can associate a unique number $\alpha \in (0,1]$ that represents exclusively that focal element, and that induces an ordering relation in the RS (an indexing); for the sake of readability, all proofs will be presented in Section 4.5.

4.2.1 Indexation by α and sampling of the basic variables

In this subsection, for the sake of simplicity in the notation, the sub/super-index i corresponding to the i-th basic variable will be omitted (see Section 4.2.2). Also, we will consider a universal non-empty set $X \subseteq R$, its power set $\mathscr{P}(X)$, the probability space $(\Omega, \sigma_\Omega, P_\Omega)$ with $\Omega = (0,1]$, $\sigma_\Omega = (0,1] \cap \mathscr{B}$ and $P_\Omega = \lambda$ (the Lebesgue measure on the interval $(0,1]$), and the measurable space $(\mathscr{F}, \sigma_\mathscr{F})$ with $\mathscr{F} \subseteq \mathscr{P}(X)$. Note that $F_{\tilde{\alpha}}(\alpha) := P_\Omega(\tilde{\alpha} \leq \alpha) = \alpha$ with $\alpha \in (0,1]$, defines a uniform CDF on $(0,1]$.[1]

Indexation by α of a normalized fuzzy set Let A be a normalized fuzzy set (possibility distribution) with membership function $A(x)$ on $X \subseteq \mathbb{R}$, and $\mathscr{F}$ be the family of all α-cuts A_α, i.e. $\mathscr{F} := \{\gamma := A_\alpha : \alpha \in (0,1]\}$. The possibility distribution A can be represented as an infinite RS $(\mathscr{F}, P_\Gamma)$, $\Gamma(\alpha) = A_\alpha$. Note that the probability measure $P_\Gamma : \sigma_\mathscr{F} \to [0,1]$ is defined by (2.6).

For every α drawn at random from $F_{\tilde{\alpha}}$, there corresponds a unique α-cut A_α and viceversa; in other words, there is a one to one relationship between A_α and α. This is the reason why the subindex α of A_α must be preserved because there exist cases were several α-cuts contain the same collection of elements, and in consequence this α makes the distinction between them.

Observe that in this case α induces an ordering in $\mathscr{F}$ such that, $\alpha_i \leq \alpha_j$ if $A_i \supseteq A_j$. Finally, it is shown in Lemma (4.5.1), that the lower and upper probability measures of any subset F of X with regard to the infinite RS $(\mathscr{F}, P_\Gamma)$ is equal to the necessity Nec and possibility Pos of the set F with respect to the normalized fuzzy set A, i.e., for all $F \subseteq X$, $\mathrm{Nec}_A(F) = \mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{Pos}_A(F) = \mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$.

Sampling a focal element from a normalized fuzzy set The sampling of a focal element from a basic variable represented by a normalized fuzzy set consists in drawing a realization of $\tilde{\alpha}$ from $F_{\tilde{\alpha}}$ and picking the corresponding α-cut A_α. Note that the lower and upper probability measures of a finite sample $(\mathscr{F}_n, m)$ will converge almost surely to the necessity Nec and possibility Pos of the set F with respect to the normalized fuzzy set A, as $n \to \infty$ that is, $\mathrm{Nec}_A(F) = \lim_{n\to\infty} \mathrm{LP}_{(\mathscr{F}_n,m)}(F)$ and $\mathrm{Pos}_A(F) = \lim_{n\to\infty} \mathrm{UP}_{(\mathscr{F}_n,m)}(F)$ for all $F \subseteq X$. This is shown in Lemma 4.5.2. Figure 4.1b gives a graphical representation of this sampling.

[1] Note: the tilde ˜ will be employed through this chapter to distinguish between the random variable $\tilde{\alpha}$ and its realization α.

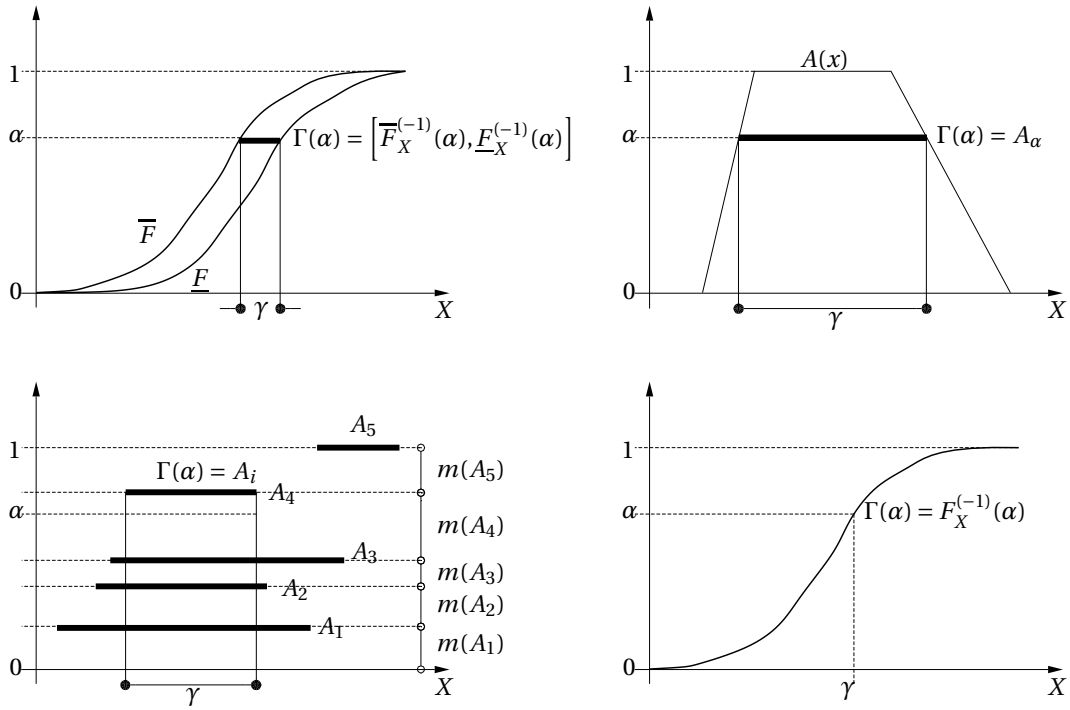


Figure 4.1: Sampling of focal elements. a) from a probability box. b) from a normalized fuzzy set (possibility distribution). c) from a family of intervals. d) from a CDF.

Indexation by α of a probability box Ferson et al. (2003b) proposed a method to approximate a probability box by a finite RS. Here, we want to propose a method to represent a probability box as an infinite RS; this suggestion is based on the definition of the inverse of a p-box proposed by Joslyn and Ferson (2004) and summarized in Section 2.4.2. Let $\left\langle \underline{F}, \overline{F} \right\rangle$ be a probability box defined on a subset of $\mathbb{R}$ and let us define

$$\mathscr{F} = \left\{ \gamma := \left\langle \underline{F}, \overline{F} \right\rangle^{-1}(\alpha) := \left[\overline{F}^{(-1)}(\alpha), \underline{F}^{(-1)}(\alpha) \right] : \alpha \in (0,1] \right\}, \tag{4.4}$$

where $\underline{F}^{(-1)}(\alpha)$ and $\overline{F}^{(-1)}(\alpha)$ are given by (2.21) and (2.22) respectively. The probability box $\left\langle \underline{F}, \overline{F} \right\rangle$ can be represented as an infinite RS $(\mathscr{F}, P_\Gamma), \Gamma(\alpha) = \left\langle \underline{F}, \overline{F} \right\rangle^{-1}(\alpha)$. The probability measure P_Γ is defined in an analogous way to the case of normalized fuzzy sets, by equation (2.6). It must be noted that for every α drawn at random from $F_{\tilde{\alpha}}$, there corresponds a unique focal element $\left\langle \underline{F}, \overline{F} \right\rangle^{-1}(\alpha)$. This relationship is one to one if the subindex $\langle \cdot, \cdot \rangle^{-1}(\alpha)$ is conserved.

Observe that in this particular case, α induces a partial ordering in $\mathscr{F}$ such that if $[a_1, b_1]_{\alpha_1}$ and $[a_2, b_2]_{\alpha_2}$ are elements of $\mathscr{F}$, then it follows that if $\alpha_1 < \alpha_2$ then $a_1 \leq a_2$ and $b_1 \leq b_2$.

In Lemma 4.5.3, it is shown that the lower and upper probability measures of $(-\infty, x]$ with respect to $(\mathscr{F}, P_\Gamma)$ is equal to $\underline{F}(x)$ and $\overline{F}(x)$ respectively, that is, $\underline{F}(x) = \mathrm{LP}_{(\mathscr{F}, P_\Gamma)}((-\infty, x])$ and $\overline{F}(x) = \mathrm{UP}_{(\mathscr{F}, P_\Gamma)}((-\infty, x])$ for all $x \in X$.

Sampling a focal element from a probability box Ferson et al. (2003b) stated that it is required to develop techniques for sampling from a probability box. In the following, an algorithm is proposed based on the considerations given above. The inversion method (see e.g. Rubinstein (1981)) allows to draw a number distributed according to a particular CDF. This method can be applied to sample from a probability box and consists in sampling an α from $F_{\tilde{\alpha}}$ and then retrieve the associated interval $\left\langle \underline{F}, \overline{F} \right\rangle^{-1}(\alpha)$. This interval will be considered as the drawn focal element inasmuch as it contains the samples for all the CDFs in the probability box, i.e., $\left\langle \underline{F}, \overline{F} \right\rangle^{-1}(\alpha) = \left\{ x : F(x) = \alpha, F \in \left\langle \underline{F}, \overline{F} \right\rangle \right\}$.

In Lemma 4.5.4, it is shown that when an infinite number of focal sets is sampled, the lower and upper probability measures of $(-\infty, x]$ with respect to the sampled RS will converge almost surely to $\underline{F}(\alpha)$ and $\overline{F}(\alpha)$ respectively, i.e., for all $x \in X$, $\underline{F}(x) = \lim_{n\to\infty} \mathrm{LP}_{(\mathscr{F}_n, m)}((-\infty, x])$ and $\overline{F}(x) = \lim_{n\to\infty} \mathrm{UP}_{(\mathscr{F}_n, m)}((-\infty, x])$. Figure 4.1a makes a graphical representation of the sampling from a probability box.

Indexation by α of a CDF When a basic variable is expressed as a random variable on $X \subseteq \mathbb{R}$, the probability law of the random variable can be expressed using a CDF, and is given by $F_X(x) = P_\Gamma(X \leq x)$ such that $x \in X$ given some probability measure P_Γ. The CDF function F_X has a quasi-inverse given by $F_X^{(-1)}$. It is well known that if $\tilde{\alpha}$ is a uniformly distributed random variable on $(0,1]$, then $X := F_X^{(-1)}(\tilde{\alpha})$ is distributed according to F_X, or equivalently, if X is a random variable with continuous CDF F_X, then $F_X(x)$ is a realization of a random variable uniformly distributed on $(0,1]$. In consequence, if $\mathscr{F} = \left\{ x \,\middle|\, x := F_X^{(-1)}(\alpha), \alpha \in (0,1] \right\}$, then the random variable X can be represented by a RS $(\mathscr{F}, P_\Gamma)$, $\Gamma(\alpha) = F_X^{(-1)}(\alpha)$; note that in this case the focal set $\mathscr{F}$ is specific.

Note that if α_i has an associated x_i for $i = 1,2$ and if $\alpha_1 < \alpha_2$, then $x_1 \leq x_2$. Observe also that this is a particular case of the probability box when $\underline{F} = \overline{F}$.

Sampling a focal element from a CDF As in the above cases, a focal element of a RS can be sampled by drawing an α from a uniform CDF on $(0,1]$, $F_{\tilde{\alpha}}$ and selecting the associated focal element from $\mathscr{F}$ (see Figure 4.1d).

Indexation by α of a finite family of intervals In practice, engineers only provide finite families of intervals with a corresponding a priori information about the confidence of their opinions on those intervals; in addition, a histogram belongs to this category. This information is contained in a finite RS $(\mathscr{F}_s, m')$.

In order to define an indexation by α, we have to induce in $(\mathscr{F}_s, m')$ an ordering. If $\{[a_i, b_i] \text{ for } i = 1, \ldots, s\}$ are the enumeration of the focal elements of $\mathscr{F}_s$, this family of intervals can be naturally sorted by the criteria: $[a_i, b_i] \leq [a_j, b_j]$ if $a_i < a_j$ or $(a_i = a_j$ and $b_i \leq b_j)$. This can be performed using a standard sorting algorithm (e.g. the quicksort algorithm) by using the appropriate comparison function. The purpose of this sorting is to induce a unique and reproducible ordering

in the family of intervals; this is required because sometimes families of intervals do not have a natural ordering structure. Therefore, if two different analyses employing different sortings are made using the algorithms explained in Section 4.3, the analyst will obtain the same bounds on $P_{\boldsymbol{X}}(F)$ but different F_{LP} and F_{UP} regions (see Section 4.2.3) and also different copulas C will be required to describe the dependence relationship between the basic variables. Other sortings, like for example sorting according to the basic mass assignment are possible, but this would alter the natural representation that for example, histograms have. Also, if several focal elements have the same basic mass assignment, it would be unclear how to sort them. Finally, in Section 5.3, are given more motivations of why this sorting is required in order to have an exact specification of the copula associated to P_{Γ}.

Suppose that after applying the above criteria, the reordering of the family of focal sets is given by the subindexes $i_1, i_2, \dots i_s$. The finite random set $(\mathscr{F}_s, m')$ can be represented as the infinite random set $\Gamma(\alpha) := A^*(\alpha)$ for $\alpha \in (0,1]$ where $A^*(\alpha)$ is the focal element for which $\sum_{k=i_1}^{i_{j-1}} m'(A_k) < \alpha \leq \sum_{k=1}^{i_j} m'(A_k)$ where k will take successively values from the list $i_1, i_2, \dots i_s$; here by convention $\sum_{k=i_1}^{0} m'(A_k) = 0$.

Sampling a focal element from a finite family of intervals Similarly to the cases described before, a focal element of $\mathscr{F}$ can be sampled (to form the RS $(\mathscr{F}_n, m)$) by drawing an α from $F_{\tilde{\alpha}}$, a uniformly distributed CDF on $(0,1]$ and then selecting the element $\Gamma(\alpha)$. An illustration of this kind of sampling is sketched in Figure 4.1c.

It is shown in Lemma 4.5.7, that the lower and upper probability measures with respect to the sampled RS $(\mathscr{F}_n, m)$ converges almost surely to the belief and plausibility with regard to the DS structure $(\mathscr{F}_s, m')$ i.e., for all $F \subseteq X$, $\mathrm{Bel}_{(\mathscr{F}_s, m')}(F) = \lim_{n\to\infty} \mathrm{LP}_{(\mathscr{F}_n, m)}(F)$ and $\mathrm{Pl}_{(\mathscr{F}_s, m')}(F) = \lim_{n\to\infty} \mathrm{UP}_{(\mathscr{F}_n, m)}(F)$.

4.2.2 Combination of focal elements: random relations on infinite random sets

After sampling each basic variable, a combination of the sampled focal elements is carried out. Usually, the joint focal elements are given by $\times_{i=1}^{d} \gamma_i$ where $\gamma_i := \Gamma^i(\alpha_i)$ are the sampled focal elements from every basic variable. Some of these γ_i are intervals, some other, points. Inasmuch as every sample of a basic variable can be represented by γ_i or by the corresponding α_i, the joint focal element can be represented either by the hypercube $\boldsymbol{\gamma} := \times_{i=1}^{d} \gamma_i \subseteq X$ or by the point $\boldsymbol{\alpha} := [\alpha_1, \alpha_2, \dots, \alpha_d] \in (0,1]^d$. Those two representations will be called *the X- and the α-representation* respectively, and $(0,1]^d$ will be referred to as the *α-space* (see Figure 4.2).

The joint focal set $\mathscr{F}$ will be defined then by

$$\mathscr{F} = \left\{ \boldsymbol{\gamma} : \boldsymbol{\gamma} := \times_{i=1}^{d} \gamma_i, \gamma_i := \Gamma^i(\alpha_i), \text{ for } \alpha_i \in (0,1] \text{ and } i = 1, \dots, d \right\}$$

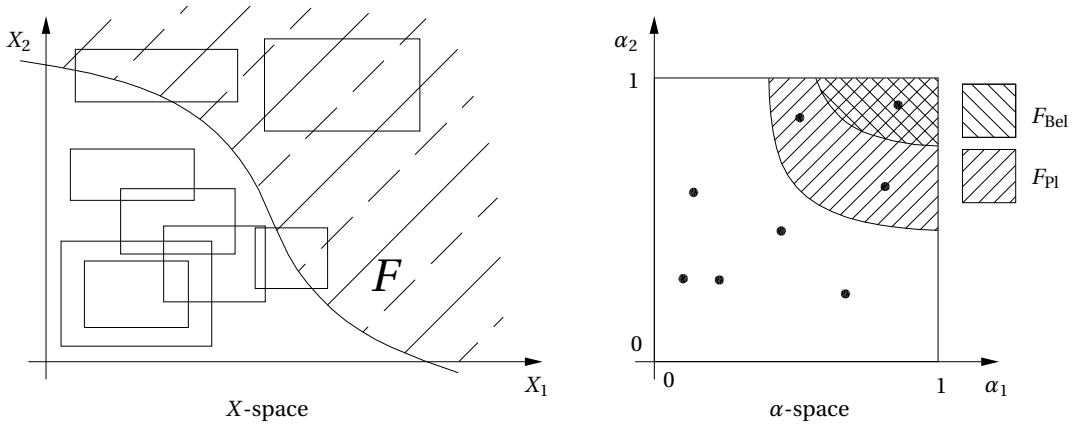


Figure 4.2: Alternative and equivalent representations of a focal element a) in X-space b) in α-space

and $\Gamma(\boldsymbol{\alpha}) := \times_{i=1}^{d} \Gamma^i(\alpha_i)$.

In the α-space there exists a joint CDF $F_{\tilde{\alpha}_1,\ldots,\tilde{\alpha}_d}$ which is defined according to the rules of probability of combination of CDFs, that is, $F_{\tilde{\alpha}_1,\ldots,\tilde{\alpha}_d}(\alpha_1,\ldots,\alpha_d) = C(F_{\tilde{\alpha}_1}(\alpha_1),\ldots,F_{\tilde{\alpha}_d}(\alpha_d))$ where d is the number of basic variables in consideration, $F_{\tilde{\alpha}_i}(\alpha_i) = \alpha_i$ is the uniform CDF defined on $(0,1]$ associated to the RS representation of the i-th basic variable for $i = 1,\ldots,d$, and C is a function that joins the marginal CDFs $F_{\tilde{\alpha}_1},\ldots F_{\tilde{\alpha}_d}$. Since $F_{\tilde{\alpha}_1},F_{\tilde{\alpha}_2},\ldots,F_{\tilde{\alpha}_d}$ are uniform CDFs on the interval $(0,1]$, we have that C is a copula, according to Sklar's theorem (Theorem 3.3.1). Recall that a copula is a probability distribution on a unit cube $(0,1]^d$ all whose marginal distributions are uniform on the interval $(0,1]$. Observe that $F_{\tilde{\alpha}_1,\ldots,\tilde{\alpha}_d} = C$ and in consequence $F_{\tilde{\alpha}_1,\ldots,\tilde{\alpha}_d}$ is a copula.

For instance, when it is assumed that all basic variables are independent, the product copula is used, $C(\alpha_1,\ldots,\alpha_d) = \prod_{i=1}^{d} \alpha_i$ for $\alpha_i \in (0,1]$ and $i = 1,\ldots,d$. To define the associated P_Γ of the joint focal set $\mathscr{F}$ in the X-space, we make use of the copula C on the α-space. The associated $P_\Gamma : \sigma_{\mathscr{F}} \to [0,1]$ is induced as:

$$P_\Gamma(\{\boldsymbol{\gamma} : \boldsymbol{\gamma} := \Gamma(\boldsymbol{\alpha}), \boldsymbol{\alpha} \in G\}) = \int_G \mathrm{d}C(\alpha_1,\ldots,\alpha_d)$$

where $\sigma_{\mathscr{F}}$ is a σ-algebra on $\mathscr{F}$, $G \subseteq (0,1]^d \cap \mathscr{B}^d$ contains the points $\boldsymbol{\alpha}$ corresponding to the focal elements which will be evaluated in the integral and $\{\boldsymbol{\gamma} : \boldsymbol{\gamma} := \Gamma(\boldsymbol{\alpha}), \boldsymbol{\alpha} \in G\}$ is an element of $\sigma_{\mathscr{F}}$.

4.2.3 An alternative representation of the lower and upper probability measure integrals

We must bear in mind that our objective is to develop an efficient method for the evaluation of $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$ according to equations (4.1) and (4.3) respectively. The computation of those integrals is not straightforward, however, using the proposed representation of the RS in the α-space, those integrals can

be rewritten as the Stieltjes integrals,

$$\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) = \underbrace{\int_{0^+}^{1}\cdots\int_{0^+}^{1}}_{d-\text{times}} I\left[[\alpha_1,\ldots,\alpha_d]\in F_{\mathrm{LP}}\right]\mathrm{d}C(\alpha_1,\ldots,\alpha_d) \tag{4.5}$$

$$\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F) = \underbrace{\int_{0^+}^{1}\cdots\int_{0^+}^{1}}_{d-\text{times}} I\left[[\alpha_1,\ldots,\alpha_d]\in F_{\mathrm{UP}}\right]\mathrm{d}C(\alpha_1,\ldots,\alpha_d) \tag{4.6}$$

provided that F_{LP} and F_{UP} are μ_C-measurable sets, where μ_C is the Lebesgue-Stieltjes measure associated to the copula C (see Section 5.3.1). The meaning of F_{LP} and F_{UP} is described in the following. Remember that every element of $\mathscr{F}$ can be represented as a point in the α-space, i.e. in $(0,1]^d$. In equation (4.1), $I\left[\boldsymbol{\gamma}\subseteq F\right]$ takes the value 1 when the focal element $\boldsymbol{\gamma}$ is totally contained in the set F; otherwise it takes 0. This is equivalent to say that there is a region of the d-box $\Omega := (0,1]^d$ that contains the sets $F_{\mathrm{LP}} := \{\alpha\in\Omega : \Gamma(\alpha)\subseteq F, \Gamma(\alpha)\neq\emptyset\}$ and $F_{\mathrm{UP}} := \{\alpha\in\Omega : \Gamma(\alpha)\cap F\neq\emptyset\}$ which are respectively composed of all those points whose corresponding focal elements are completely contained in the set F or have in common at least one point with F; some authors like Nguyen (1978) call these sets the *lower* and *upper inverses of* F respectively. Note that $I\left[\boldsymbol{\alpha}\in F_{\mathrm{LP}}\right]$ is equivalent to $I\left[\boldsymbol{\gamma}\subseteq F\right]$ of equation (4.1). Similar considerations apply to the evaluation of the upper probability measure by means of integral (4.3), and therefore, $I\left[\boldsymbol{\alpha}\in F_{\mathrm{UP}}\right] \equiv I\left[\boldsymbol{\gamma}\cap F\neq\emptyset\right]$. Since the set $\{\boldsymbol{\gamma} : \boldsymbol{\gamma}\subseteq F, \boldsymbol{\gamma}\in\mathscr{F}\}$ is contained in the set $\{\boldsymbol{\gamma} : \boldsymbol{\gamma}\cap F\neq\emptyset, \boldsymbol{\gamma}\in\mathscr{F}\}$ it is clear that $F_{\mathrm{LP}}\subseteq F_{\mathrm{UP}}$.

Now, with regard to the evaluation of the lower probability measure, since in the α-space all focal sets of $\mathscr{F}$ are represented by a point, and since there is a region F_{LP} in equation (1.1) which can be understood as a failure region, many of the algorithms developed to evaluate (1.1), and which only consider the sign of $g(\boldsymbol{x})$ (like for example importance sampling) may be employed in the evaluation of $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ by means of (4.5). The same considerations hold for the evaluation of the upper probability measure $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$ according to (4.6). The chosen algorithm will select some key points in $(0,1]^d$ which must be examined as to whether they belong to F_{LP}, F_{UP} or not. This is verified using one of the traditional methods like the vertex, sampling, optimization or response surface method, already mentioned in section 2.2.5.

Observe that in the case that all basic variables are random, $F_{\mathrm{LP}} = F_{\mathrm{UP}}$.

This is a good point to mention that the fact that the lower and upper probability measures can be expressed as Stieltjes integrals can also be deduced from the results in De Cooman et al. (2005) and that some definitions of random sets have been proposed via some measurability conditions on the upper and lower inverses (see for instance Himmelberg (1975)).

4.3 Sampling from an infinite random set

Since the analytical solution of equations (4.5) and (4.6) may be difficult or even impossible, Monte Carlo simulation (MCS) techniques can help us to approximate $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$.

Given an infinite RS $(\mathscr{F},P_\Gamma)$ which represents for example a probability box or a possibility distribution, MCS would just draw a representative sample from the it. This process is analogous to drawing a sample from a given CDF and can be done using the following algorithm:

Algorithm 1: Procedure to sample n points from the infinite RS $(\mathscr{F},P_\Gamma)$

- For $j=1$ to n do
  - Sample a point $\boldsymbol{\alpha}_j \in (0,1]^d$ from the copula C. Nelsen (1999) provides methods to perform sampling from copulas.
  - For every element of $\boldsymbol{\alpha}_j$, namely α_i^j for $i=1,\ldots,d$, use the methods explained in Section 4.2 to obtain the focal element $\gamma_i^j := \Gamma^i(\alpha_i^j)$.
  - Form the joint focal element $A_j := \times_{i=1}^d \gamma_i^j$.
- Form the finite RS $(\mathscr{F}_n,m)$ where $\mathscr{F}_n = \{A_1,\ldots,A_n\}$ and $m(A_j)=1/n$ for $j=1,\ldots,n$.

Let $(\mathscr{F}_n,m)$ be a sample of $(\mathscr{F},P_\Gamma)$ which contains n elements. In particular, $\mathscr{F}_n=\{A_1,A_2,\ldots,A_n\}$. The lower and upper probability measures of the sample are given by equations (2.2) and (2.3). Since $(\mathscr{F}_n,m)$ was randomly sampled from $(\mathscr{F},P_\Gamma)$, it happens that m has equal weight over the elements sampled, i.e.

$$m(A_j)=\frac{1}{|\mathscr{F}_n|}=\frac{1}{n} \tag{4.7}$$

for all $j=1,\ldots,n$. Notice that $\sum_{j=1}^n m(A_j)=1$. Now, rewriting equations (2.2) and (2.3) using (4.7), we have that

$$\mathrm{LP}_{(\mathscr{F}_n,m)}(F)=\frac{1}{n}\sum_{j=1}^{n} I\left[A_j\subseteq F\right] \tag{4.8}$$

$$\mathrm{UP}_{(\mathscr{F}_n,m)}(F)=\frac{1}{n}\sum_{j=1}^{n} I\left[A_j\cap F\neq\emptyset\right] \tag{4.9}$$

If we take the expected value of $\mathrm{LP}_{(\mathscr{F}_n,m)}(F)$ and $\mathrm{UP}_{(\mathscr{F}_n,m)}(F)$ with regard to the

probability measure P_Γ, we have that,

$$\begin{aligned}
\mathrm{E}\left[\mathrm{LP}_{(\mathscr{F}_n,m)}(F)\right] &= \int_{\mathscr{F}} \frac{1}{n}\sum_{j=1}^{n} I\left[A_j \subseteq F\right] \mathrm{d}P_\Gamma(A_j) \\
&= \frac{1}{n}\sum_{j=1}^{n} \int_{\mathscr{F}} I\left[A_j \subseteq F\right] \mathrm{d}P_\Gamma(A_j) \\
&= \frac{1}{n}\sum_{j=1}^{n} \mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) \\
&= \mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F).
\end{aligned}$$

A similar reasoning can be done with the upper probability measure and then,

$$\mathrm{E}\left[\mathrm{UP}_{(\mathscr{F}_n,m)}(F)\right] = \mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F).$$

This shows that $\mathrm{LP}_{(\mathscr{F}_n,m)}(F)$ and $\mathrm{UP}_{(\mathscr{F}_n,m)}(F)$ are unbiased estimators of $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$ respectively.

Now, we would like to show that when the number of random samples goes to infinity,

$$\sum_{j=1}^{n} I\left[A_j \in \mathscr{S}\right] m(A_j) \xrightarrow{\text{a.s.}} P_\Gamma(\mathscr{S})$$

for all $\mathscr{S} \in \sigma_{\mathscr{F}}$ as $n \to \infty$. This follows directly from (4.7) and Borel's strong law of large numbers (see e.g. Loève (1977)), i.e. $\sum_{j=1}^{n} I\left[A_j \in \mathscr{S}\right]/n \xrightarrow{\text{a.s.}} P_\Gamma(\mathscr{S})$ as $n \to \infty$.

Using the last result, we can state that

Theorem 4.3.1. *Let $(\mathscr{F}, P_\Gamma)$ be an infinite random set defined on X and $(\mathscr{F}_n, m)$ a sample from it. The lower (upper) probability measure of the RS $(\mathscr{F}_n, m)$ converges as $n \to \infty$ almost surely to the lower (upper) probability measure of the RS $(\mathscr{F}, P_\Gamma)$, i.e.*

$$\begin{aligned}
\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) &= \lim_{n\to\infty} \mathrm{LP}_{(\mathscr{F}_n,m)}(F) \\
\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F) &= \lim_{n\to\infty} \mathrm{UP}_{(\mathscr{F}_n,m)}(F)
\end{aligned}$$

almost surely for all $F \in \mathscr{P}(X)$.

In Section 7.3.2 we will analize again this result.

In conclusion the following algorithm helps us to obtain an unbiased estimator of $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$ by means of direct MCS.

Algorithm 2: Procedure to estimate $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$ from a finite sample

- Use Algorithm 1 to obtain n samples from $(\mathscr{F}, P_\Gamma)$.
- Using one of the methods explained in Section 2.2.5 (for example the optimization method), check if every focal element $A_j, j = 1,\ldots,n$ is totally contained in F (i.e. $A_j \subseteq F$) or A_j shares points with F (i.e. $A_j \cap F \neq \emptyset$). In the first case $\boldsymbol{\alpha}_j \in F_{\mathrm{LP}}$ and in the second $\boldsymbol{\alpha}_j \in F_{\mathrm{UP}}$.

- Use equations (4.8) and (4.9) to estimate $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$.

4.4 Example

To test the proposed approach, the "Challenge Problem B" proposed by Oberkampf et al. (2004) was solved. For the sake of completeness, the formulation of this problem will be repeated here. Consider the linear mass-spring-damper system subjected to a forcing function $Y\cos(\omega t)$ and depicted in Figure 4.3. The system has a mass m, a stiffness constant k, a damping constant c and

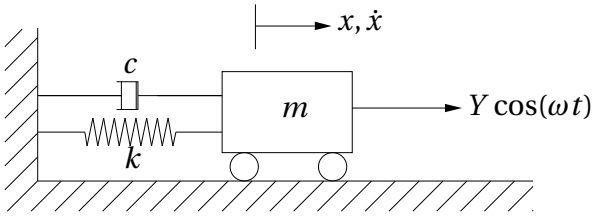


Figure 4.3: Mass-spring-damper system acted on by an excitation function.

the load has an oscillation frequency ω. The task for this problem is to estimate the uncertainty in the steady-state magnification factor D_S, which is defined as the ratio of the amplitude of the steady-state response of the system to the static displacement of the system, i.e.

$$D_S = \frac{k}{\sqrt{\left(k - m\omega^2\right)^2 + (c\omega)^2}} \tag{4.10}$$

To solve the problem we have to use exclusively the information provided, and we have to avoid any extra supposition on the data given. If it is so, they must be clearly specified. The parameters m, k, c and ω are independent, that is, the knowledge about the value of one parameter implies nothing about the value of the other. The information for each parameter is as follows:

- Parameter m. It is given by a triangular PDF defined on the interval $[m_{\min}, m_{\max}] = [10, 12]$ and with mode $m_{\mathrm{mod}} = 11$.
- Parameter k. It is stated by three equally credible and independent sources of information. Sources agree on that k is given by a triangular PDF, however each of them gives a closed interval for the different parameters $m_{\min}$, m_{mod} and $m_{\max}$, i.e.:
  - Source 1: $m_{\min} = [90, 100]$, $m_{\mathrm{mod}} = [150, 160]$ and $m_{\max} = [200, 210]$.
  - Source 2: $m_{\min} = [80, 110]$, $m_{\mathrm{mod}} = [140, 170]$ and $m_{\max} = [200, 220]$.
  - Source 3: $m_{\min} = [60, 120]$, $m_{\mathrm{mod}} = [120, 180]$ and $m_{\max} = [190, 230]$.
- Parameter c. Three equally credible and independent sources of information are available. Each source provided an interval for c, as follows:

  - Source 1: $m = [5, 10]$.
  - Source 2: $m = [15, 20]$.
  - Source 3: $m = [25, 25]$.

- Parameter ω. It is modeled by a triangular PDF defined on the interval $[m_{\min}, m_{\max}]$ and with mode m_{mod}. The values of $m_{\min}$, m_{mod} and $m_{\max}$ are given respectively by the intervals $[2, 2.3]$, $[2.5, 2.7]$ and $[3.0, 3.5]$.

Note that the external amplitude Y does not appear on equation (4.10).

Some remarks are necessary on the implementation of the proposed approach. The parameter m was modeled simply as a triangular CDF $T(10, 11, 12)$; here $T(a, b, c)$ stands for the formulation of a triangular CDF corresponding to the triangular PDF $t(x_1, x_2, x_3)$ with lower limit x_1, mode x_2 and upper limit x_3. For modeling k, the information provided by every source was represented by a probability box and in a further step, they were combined using the intersection rule for aggregation of p-boxes (see Ferson et al. (2003b)), i.e. intersection($\langle T(90, 150, 200), T(100, 160, 210)\rangle$, $\langle T(80, 140, 200), T(110, 170, 220)\rangle$, $\langle T(60, 120, 190), T(120, 180, 230)\rangle$), which turns to be $\langle T(60, 120, 190), T(100, 160, 210)\rangle$. The parameter c was modeled as a finite RS with focal sets $[5, 10]$, $[15, 20]$ and $[25, 25]$, every one of them with a basic mass assignment of $1/3$. Finally, the parameter ω was modeled as a probability box $\langle T(2, 2.5, 3.0), T(2.3, 2.7, 3.5)\rangle$. The image of all focal elements was calculated using the optimization method (see Section 2.2.5), inasmuch as this is the most accurate of the methods to estimate the image of the focal elements. Since the basic variables are considered to be independent, the product copula was employed to model dependence.

For comparison reasons, the problem was solved using the same strategy employed in Tonon (2004c) for four different discretizations, namely 5, 10, 20 and 30 elements for each basic variable. The upper and lower CDFs of the system response D_s are shown in Figure 4.4, including a detail of the tails of the CDFs in Figures 4.5 and Figures 4.6.

The proposed method can produce these curves by means of a direct Monte Carlo simulation, however, in this case, only the lower and upper probability bounds for a given region $F = [D_s, \infty)$ were estimated. For the sake of comparison, the values $D_s = 2.0, 2.5$ and 3.0 were chosen. In consequence, the failure regions (sets F) were modeled by $[2.0, \infty)$, $[2.5, \infty)$, and $[3.0, \infty)$. Tables 4.1 and 4.2 show the lower and upper probability bounds obtained by the methodology of Tonon (2004c) and by the proposed one respectively. The results show wide intervals containing $P_{\boldsymbol{X}}(F)$. This should not be taken as an argument against random set theory though. What it does show is the danger, even in simple problems, of assuming precise parameters in order to obtain a unique value of $P_{\boldsymbol{X}}(F)$ at the end. The breadth of those intervals can be reduced if additional information about the basic variables is obtained; this issue is studied in Chapter 9.

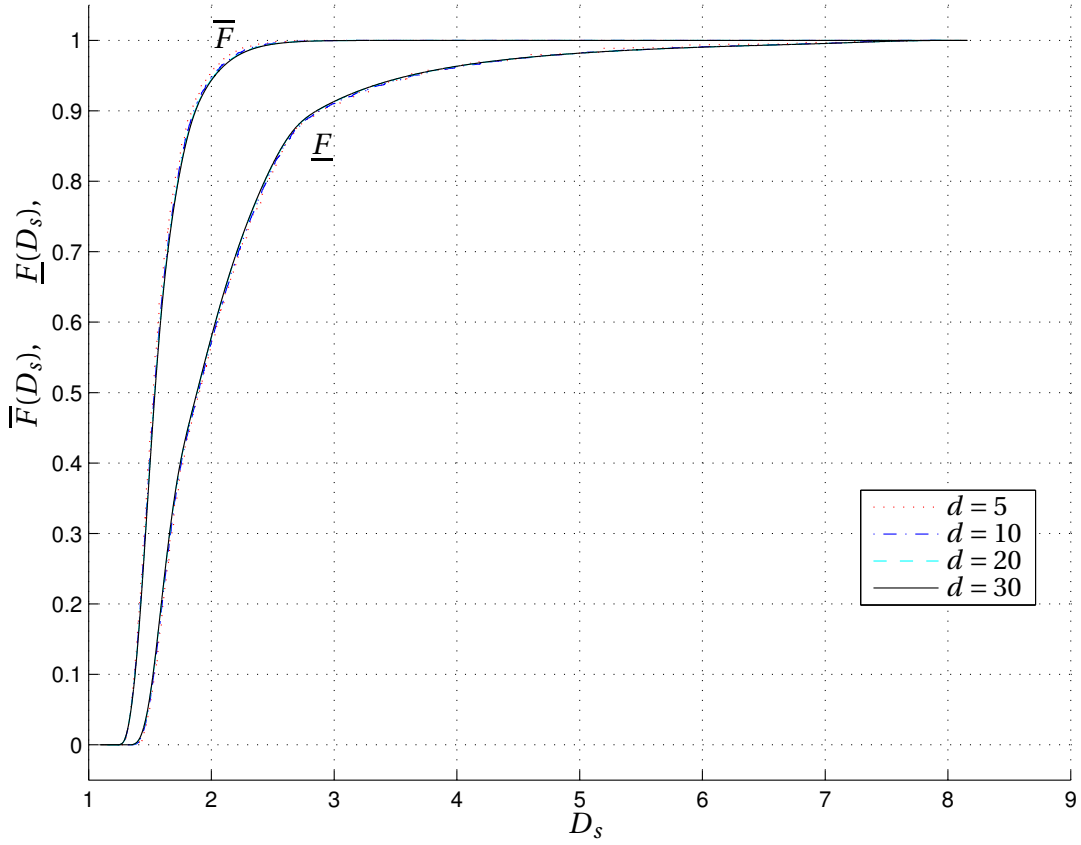


Figure 4.4: Upper and lower CDFs of the system response, for $D_s = 2.0, 2.5$ and 3.0.

Table 4.1: Lower and upper probability bounds of the region $F = [D_s, \infty)$ obtained by applying a methodology of Dempster-Shafer evidence theory, following the same strategy employed in Tonon (2004c). Here n corresponds to the number of discretizations for each basic variable and N_{elem} the number of focal elements evaluated.

		$D_s = 2.0$	$D_s = 2.5$	$D_s = 3.0$	N_{elem}
$n = 5$	LP(F)	0.04389	0.00216	0	1275
	UP(F)	0.42832	0.18890	0.09023	
$n = 10$	LP(F)	0.05275	0.00464	0.00033	16500
	UP(F)	0.42786	0.18328	0.08976	
$n = 20$	LP(F)	0.05568	0.00523	0.00040	259200
	UP(F)	0.42174	0.17638	0.08720	
$n = 30$	LP(F)	0.05680	0.00555	0.00047	1312200
	UP(F)	0.42103	0.17580	0.08678	

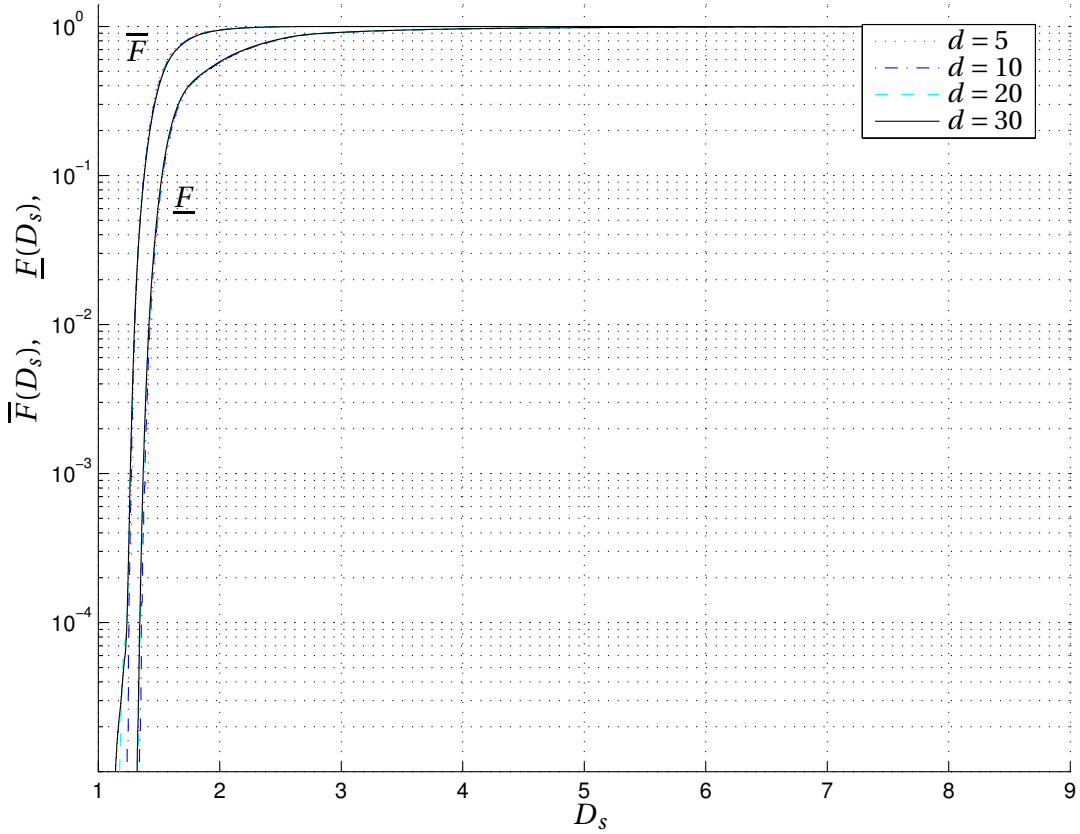


Figure 4.5: Upper and lower CDFs of the system response, for $D_s = 2.0, 2.5$ and 3.0. Detail of the left tail of Figure 4.4

Table 4.2: Lower and upper probability bounds of the region $F = [D_s, \infty)$ obtained by applying the methodology of infinite random sets. The values shown were calculated using 100000 simulations (focal element evaluations) of a direct Monte Carlo simulation.

D_s	Direct MCS	
	LP	UP
2.0	0.0652	0.4252
2.5	0.0111	0.1809
3.0	0.0015	0.1001

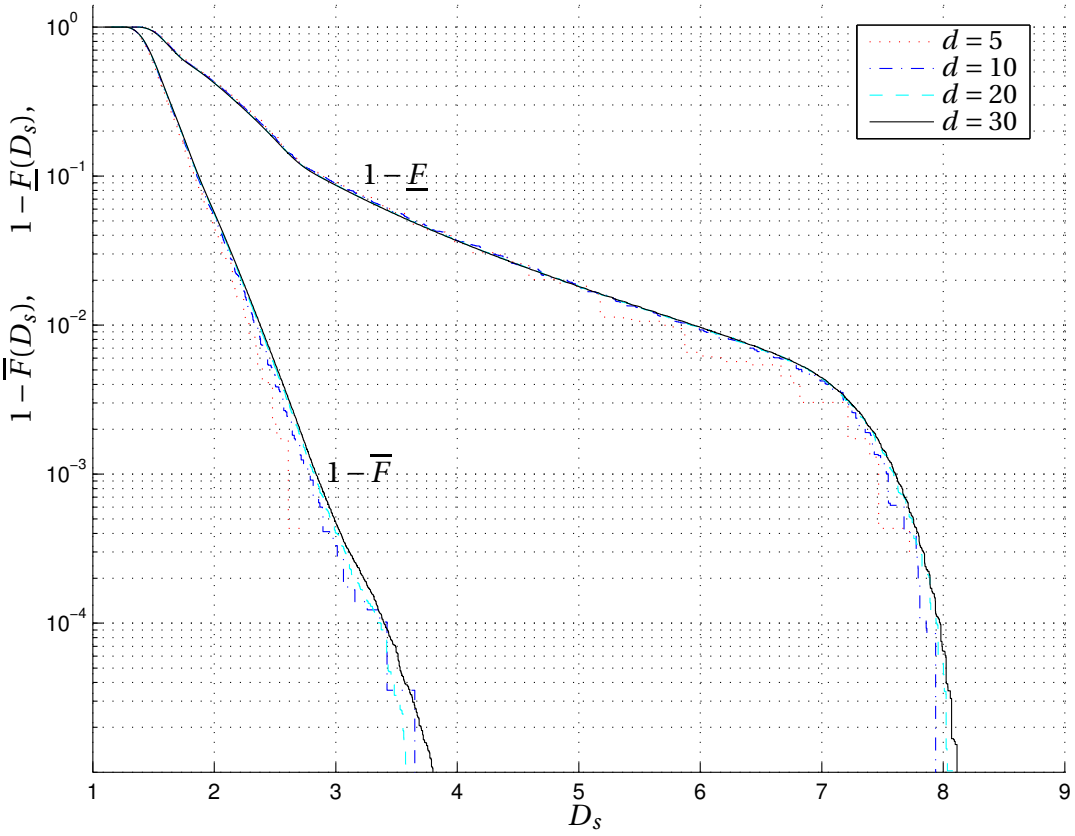


Figure 4.6: Upper and lower CDFs of the system response, for $D_s = 2.0, 2.5$ and 3.0. Detail of the right tail of Figure 4.4

Notice that in the methodology of Dempster-Shafer evidence theory the tails of the upper and lower CDFs of the system response are highly sensitive to the degree of the discretization of each random variable. Their precision improves by increasing the number of focal elements to be evaluated. In this sense, methodologies like the one employed by Bae et al. (2003, 2004a,b) are not efficient when small lower and upper probability measures of the set F must be calculated. The results obtained with the proposed approach do not use a discretization of the basic variables, and so are free of the error that could be introduced by the discretization. Since $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$ where estimated by MCS, the precision depends in this case on the number of simulations employed, which were 100000.

Note also, that since the sources of information in the finite case for k were mixed using the Dempster combination rule (see e.g. Shafer (1976); Sentz and Ferson (2002)), and in the infinite case with the intersection rule, the values of these cases are not comparable, but are similar in magnitude.

In relation to the proposed algorithm, according to Section 4.2.3, the region F_{LP} is contained in the region F_{UP}. This is graphically shown in Figure 4.7. This con-

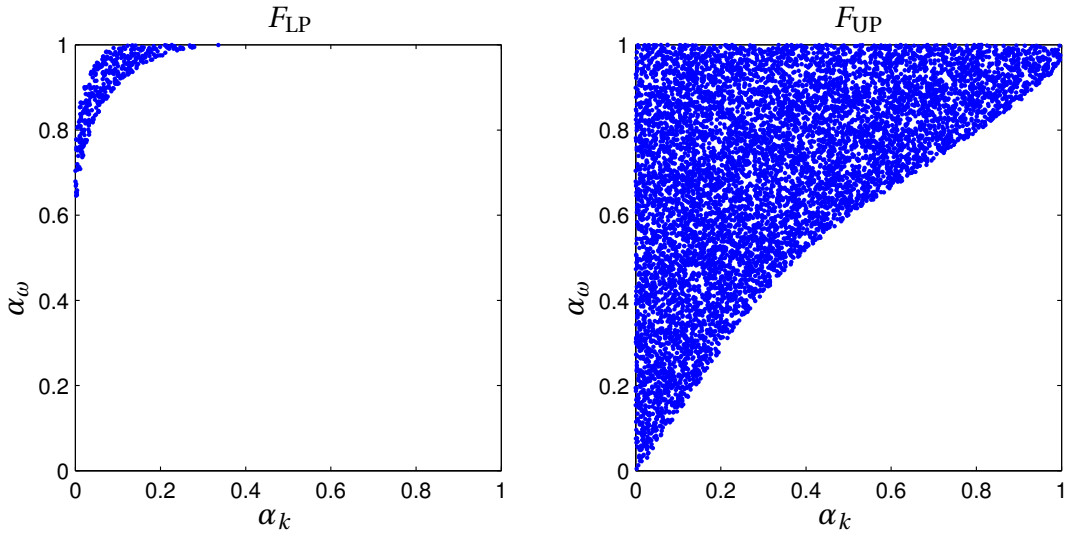


Figure 4.7: Regions F_{LP} and F_{UP}. These graphics, in the α-space, were calculated from the example by means of 20000 Monte Carlo simulations, setting $\alpha_m = 0.95$ and $\alpha_c = 0.20$. In this case the failure region was defined by $D_s = 2.8$, and in consequence, $\mathrm{LP}(F) = 0.02015$ and $\mathrm{UP}(F) = 0.45045$.

firms the relation shown in Figure 4.2.

In conclusion, using an infinite number of focal elements we have obtained a unified approach towards the treatment of probability boxes, possibility distributions and Dempster-Shafer structures, without employing discretizations of the basic variables of any kind.

4.5 Proofs

This section contains the demonstration of some results that where postulated in Section 4.2.

Lemma 4.5.1. *Let $A : X \to [0,1]$ be a possibility distribution and $(\mathscr{F}, P_\Gamma)$ be its representation as an infinite RS defined on $X \subseteq \mathbb{R}$. The lower and upper probability measures of any subset F of X with regard to the RS $(\mathscr{F}, P_\Gamma)$ is equal to the necessity* Nec *and possibility* Pos *of the set F with respect to the possibility distribution A, i.e.,*

$$\mathrm{Nec}_A(F) = \mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) \tag{4.11}$$

$$\mathrm{Pos}_A(F) = \mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F) \tag{4.12}$$

for all $F \subseteq X$.

Proof. Let's recall that equation (2.20) defines what is a possibility measure,

$$\mathrm{Pos}_A(F) = \sup_{x \in F} \{A(x)\}$$

Let $\mathscr{G} = \{A_\alpha : F \cap A_\alpha \neq \emptyset, A_\alpha \in \mathscr{F}\}$ where $A_\alpha = \{x \in X : A(x) \geq \alpha\}$. Since $\alpha_i \leq \alpha_j$ holds if and only if $A_i \supseteq A_j$ and since $\mathscr{F}$ is consonant, there exists an $\alpha^* \in (0,1]$ such that $A_{\alpha^*} \in \mathscr{G}$ is contained in all elements of $\mathscr{G}$; in other words, α^* is the α associated to the focal set $A_\alpha \in \mathscr{G}$ that is contained in all focal sets of $\mathscr{G}$, and therefore α^* is the largest α of the α-s associated to the focal sets of $\mathscr{G}$, i.e., $\alpha \leq \alpha^*$ for $A_\alpha \in \mathscr{G}$. Now, for a focal set $A_\alpha \in \mathscr{G}$ we have that $F \cap A_\alpha \neq \emptyset$; this implies that there exists an x such that $x \in F$, $x \in A_\alpha$, and $A(x) \geq \alpha$. But now, $\mathrm{Pos}_A(F) = \sup_{x \in F}\{A(x)\} \geq \alpha$, and since this holds for all α-s such that $A_\alpha \in \mathscr{G}$, then $\mathrm{Pos}_A(F) \geq \alpha^*$. Now, $\mathrm{Pos}_A(F) \geq A(x)$ for all $x \in F$. On the other hand, for all $\epsilon > 0$, there exists an $x \in F$ such that $\mathrm{Pos}_A(F) - \epsilon < A(x)$. Also, there exists an α such that $A(x) = \alpha$ and $A_\alpha \in \mathscr{G}$. Thus, $\alpha \leq \alpha^*$ and therefore, $\mathrm{Pos}_A(F) - \epsilon < \alpha^*$, and since ϵ is arbitrary, $\mathrm{Pos}_A(F) \leq \alpha^*$. Thus, $\alpha^* = \mathrm{Pos}_A(F)$, i.e., α^* is the possibility of F with regard to the normalized fuzzy set A.

Now, according to (4.3),

$$\begin{aligned} \mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F) &= \int_{\mathscr{F}} I\left[A_\gamma \cap F \neq \emptyset\right] \mathrm{d}P_\Gamma(A_\gamma) \\ &= \int_{\mathscr{G}} \mathrm{d}P_\Gamma(A_\gamma) \\ &= P_\Gamma(\mathscr{G}) \end{aligned} \tag{4.13}$$

Now, let $G = \{\alpha : A_\alpha \in \mathscr{G}\}$. Then from (4.13), and using (2.6), we have

$$\begin{aligned} \mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F) &= P_\Omega(G) \\ &= \alpha^* \end{aligned}$$

Finally using the fact that $\alpha^* = \mathrm{Pos}_A(F)$, equation (4.12) follows.

The proof of equation (4.11) is straightforward, considering the fact that the necessity and the lower probability measure are dual fuzzy measures of the possibility and upper probability measure respectively, i.e.,

$$\begin{aligned}\mathrm{Nec}_A(F) &= 1 - \mathrm{Pos}_A(F^c)\\ \mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) &= 1 - \mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F^c).\end{aligned}$$

□

Lemma 4.5.2. *Let $A : X \to [0,1]$ be a possibility distribution and $(\mathscr{F}, P_\Gamma)$ be its representation as an infinite RS defined of $X \subseteq \mathbb{R}$ and $(\mathscr{F}_n, m)$ be a finite sample of this RS with n elements. The lower and upper probability measure of the set $F \subseteq X$ with respect to $(\mathscr{F}_n, m)$ converge almost surely to the necessity* Nec *and possibility* Pos *of the set F with respect to the possibility distributions A, that is,*

$$\begin{aligned}\mathrm{Nec}_A(F) &= \lim_{n\to\infty} \mathrm{LP}_{(\mathscr{F}_n,m)}(F)\\ \mathrm{Pos}_A(F) &= \lim_{n\to\infty} \mathrm{UP}_{(\mathscr{F}_n,m)}(F)\end{aligned}$$

almost surely for all $F \subseteq X$.

Proof. This result follows immediately from the application of Lemma 4.5.1 and Theorem 4.3.1 □

Lemma 4.5.3. *Let $\langle \underline{F}, \overline{F}\rangle$ be a probability box and $(\mathscr{F}, P_\Gamma)$ be its representation as an infinite RS both defined on $X \subseteq \mathbb{R}$. The lower and upper probability measures of $(-\infty, x]$ with respect to $(\mathscr{F}, P_\Gamma)$ is equal to $\underline{F}(x)$ and $\overline{F}(x)$ respectively, that is,*

$$\underline{F}(x) = \mathrm{LP}_{(\mathscr{F},P_\Gamma)}((-\infty, x]) \tag{4.14}$$

and

$$\overline{F}(x) = \mathrm{UP}_{(\mathscr{F},P_\Gamma)}((-\infty, x]) \tag{4.15}$$

for all $x \in X$.

Proof. To show equation (4.14) we make use of the fact that according to equation (4.1),

$$\mathrm{LP}_{(\mathscr{F},P_\Gamma)}((-\infty, x]) = \int_{\mathscr{F}} I\left[\gamma \subseteq (-\infty, x]\right] \mathrm{d}P_\Gamma(\gamma), \tag{4.16}$$

and that according to (4.4),

$$\gamma = \left[\overline{F}^{(-1)}(\alpha), \underline{F}^{(-1)}(\alpha)\right]_\alpha \tag{4.17}$$

for all $\alpha \in (0,1]$. Since $\gamma \subseteq (-\infty, x]$ implies that $\underline{F}^{(-1)}(\alpha) \leq x$, or equivalently $\alpha \leq \underline{F}(x)$ (since $\underline{F}$ is monotone increasing), then rewriting (4.16) in the α-representation

$$\begin{aligned}\mathrm{LP}_{(\mathscr{F},P_\Gamma)}((-\infty, x]) &= \int_{(0,1]} I\left[\alpha \leq \underline{F}(x)\right] \mathrm{d}P_\Omega(\alpha)\\ \mathrm{LP}_{(\mathscr{F},P_\Gamma)}((-\infty, x]) &= \int_{(0,\underline{F}(x)]} \mathrm{d}P_\Omega(\alpha)\\ &= P_\Omega\left((0, \underline{F}(x)]\right)\\ &= \underline{F}(x).\end{aligned}$$

To show (4.15) we make use of the fact that according to equation (4.3),

$$\mathrm{UP}_{(\mathscr{F},P_\Gamma)}((-\infty,x]) = \int_{\mathscr{F}} I\left[\gamma \cap (-\infty,x] \neq \emptyset\right] \mathrm{d}P_\Gamma(\gamma), \tag{4.18}$$

According to (4.17), $\gamma \cap (-\infty,x] \neq \emptyset$ implies that $\overline{F}^{(-1)}(\alpha) \leq x$, or equivalently $\alpha \leq \overline{F}(x)$ (since $\overline{F}$ is monotone increasing), then rewriting (4.18) in the α-representation

$$\begin{aligned}
\mathrm{UP}_{(\mathscr{F},P_\Gamma)}((-\infty,x]) &= \int_{(0,1]} I\left[\alpha \leq \overline{F}(x)\right] \mathrm{d}P_\Omega(\alpha) \\
\mathrm{UP}_{(\mathscr{F},P_\Gamma)}((-\infty,x]) &= \int_{(0,\overline{F}(x)]} \mathrm{d}P_\Omega(\alpha) \\
&= P_\Omega\left((0,\overline{F}(x)]\right) \\
&= \overline{F}(x)
\end{aligned}$$

since P is a measure that generates the uniform distribution on $(0,1]$. □

Lemma 4.5.4. *Let $\langle \underline{F}, \overline{F} \rangle$ be a probability box, $(\mathscr{F},P_\Gamma)$ be its representation as an infinite RS defined of $X \subseteq \mathbb{R}$ and $(\mathscr{F}_n,m)$ be a finite sample of this RS with n elements. In the limit, when an infinite number of focal sets is sampled, the lower and upper probability measure of $(-\infty,x]$ with respect to the sampled RS converge almost surely to $\underline{F}(\alpha)$ and $\overline{F}(\alpha)$ respectively, that is,*

$$\begin{aligned}
\underline{F}(x) &= \lim_{n\to\infty} \mathrm{LP}_{(\mathscr{F}_n,m)}((-\infty,x]) \\
\overline{F}(x) &= \lim_{n\to\infty} \mathrm{UP}_{(\mathscr{F}_n,m)}((-\infty,x])
\end{aligned}$$

almost surely for all $x \in X$, at which $\underline{F}$ and $\overline{F}$ are continuous respectively.

Proof. This result follows immediately from the application of Lemma 4.5.3 and Theorem 4.3.1. □

Lemma 4.5.5. *Let F_X be a CDF and $(\mathscr{F},P_\Gamma)$ be its representation as an infinite RS defined of $X \subseteq \mathbb{R}$. Then both lower and upper probability measures of $(-\infty,x]$ with respect to $(\mathscr{F},P_\Gamma)$ are equal to $F_X(x)$, i.e.,*

$$F_X(x) = \mathrm{LP}_{(\mathscr{F},P_\Gamma)}((-\infty,x]) = \mathrm{UP}_{(\mathscr{F},P_\Gamma)}((-\infty,x]) \tag{4.19}$$

for all $x \in X$.

Proof. A CDF is a special case of a probability box $\langle \underline{F}, \overline{F} \rangle$ when $\overline{F} = \underline{F}$. In this case (4.19), follows directly from Lemma 4.5.3. □

Lemma 4.5.6. *Let F_X be a CDF and $(\mathscr{F},P_\Gamma)$ be an infinite RS defined on $X \subseteq \mathbb{R}$ and $(\mathscr{F}_n,m)$ be a finite sample of this RS with n elements. Then, both lower and upper probability measures of $(-\infty,x]$ with respect to the RS formed by the finite sample $(\mathscr{F}_n,m)$ converge almost surely to $F_X(x)$ for all $x \in X$, i.e.,*

$$F_X(x) = \lim_{n\to\infty} \mathrm{LP}_{(\mathscr{F}_n,m)}((-\infty,x]) = \lim_{n\to\infty} \mathrm{UP}_{(\mathscr{F}_n,m)}((-\infty,x]) \tag{4.20}$$

almost surely for all $x \in X$ at which $\underline{F}$ and $\overline{F}$ are continuous.

Proof. A CDF is a special case of a probability box $\langle \underline{F}, \overline{F} \rangle$ when $\overline{F} = \underline{F}$. In this case (4.20), follows directly from Lemma 4.5.4. □

Lemma 4.5.7. *Let $(\mathscr{F}_s, m')$ be a DS structure defined on $X \subseteq \mathbb{R}$ and $(\mathscr{F}_n, m)$ be a finite sample of this DS strcuture with n elements. Then, the lower and upper probability with respect to the sampled RS $(\mathscr{F}_n, m)$ converges almost surely to the belief and plausibility with regard to the RS $(\mathscr{F}_s, m')$ i.e.,*

$$\mathrm{Bel}_{(\mathscr{F}_s, m')}(F) = \lim_{n \to \infty} \mathrm{LP}_{(\mathscr{F}_n, m)}(F)$$
$$\mathrm{Pl}_{(\mathscr{F}_s, m')}(F) = \lim_{n \to \infty} \mathrm{UP}_{(\mathscr{F}_n, m)}(F)$$

for all $F \subseteq X$.

Proof. This follows inmediately from the application of Theorem 4.3.1 and the fact that $(\mathscr{F}_s, m') \equiv (\mathscr{F}, P_\Gamma)$. □

Chapter 5

From finite to infinite random sets: convergence considerations

The use of finite RSs (or equivalently Dempster-Shafer bodies of evidence) requires the discretization of the basic variables; it is accepted within the community working in random set and evidence theories in the framework of uncertainty and reliability analysis that when the number of focal elements employed in the representation of a given basic variable increases, the quality of the representation of that variable will also increase, and in consequence better estimators of the lower and upper probabilities of an event F will be obtained. In this chapter we will show that when the level of discretization tends to an infinite number of focal elements, the finite RS in consideration will converge to an infinite RS.

When several basic variables are employed, RS theory requires in addition, some dependence information which is usually given in terms of copulas. In this chapter, we will also analyze the role of copulas in the specification of dependence within infinite RSs.

The plan of this chapter is as follows: sections 5.1 and 5.2 introduce the concept of partitions for finite random sets in the uni- and multi-dimensional case. The role of copulas in the specification of the dependence information is analyzed in section 5.3; section 5.4 defines a special class of infinite RSs, namely those of indexable type, while sections 5.5 and 5.6 discuses respectively refinements and convergence considerations for finite random sets.

5.1 Partitions and refinements

A *partition* of an interval $(a,b]$ is a finite sequence $a = x_0 < x_1 < \cdots < x_n = b$. Each $(x_{j-1}, x_j], j = 1,\ldots,n$ is called a *subinterval* of the partition. The set $\mathscr{P} = \{(x_0,x_1],(x_1,x_2],\ldots,(x_{n-1},x_n]\}$ is sometimes used to denote the partition and $\|\mathscr{P}\| := \max_{i=1}^{n}(x_i - x_{i-1})$ denotes the *norm of the partition*.

A *refinement* of the partition $\mathscr{P}' = \{(x_0,x_1],(x_1,x_2],\ldots,(x_{n-1},x_n]\}$ is a partition

$\mathscr{P}'' = \{(y_0, y_1], (y_1, y_2], \ldots, (y_{m-1}, y_m]\}$, $x_0 = y_0, x_n = y_m$ such that for every $i = 0, \ldots, n$ there is an integer $r(i)$ such that $x_i = y_{r(i)}$; this will be denoted henceforth by $\mathscr{P}' \prec \mathscr{P}''$, since refinement induces a partial ordering. In other words, to make a refinement, one cuts the subintervals into smaller pieces and does not remove any cut.

The definition of a partition and refinement of a partition can be generalized to higher dimensions. Here a box $\times_{i=1}^{d}(a_i, b_i]$ will be decomposed into a disjoint set of boxes, that when refined will be decomposed into boxes as well.

5.2 α-representation of Dempster-Shafer structures: partitions of $(0,1]^d$

Recapitulating from Chapter 4, the *α-representation* of an infinite random set $(\mathscr{F}, P_\Gamma) := \times_{i=1}^{d}\left(\mathscr{F}^i, P_\Gamma^i\right)$ (up to some dependence specification) is a representation of the points $(\alpha_1, \alpha_2, \ldots, \alpha_d) \in (0,1]^d$ which correspond to the focal elements $\Gamma^1(\alpha_1) \times \Gamma^2(\alpha_2) \times \cdots \times \Gamma^d(\alpha_d) \in \mathscr{F}$ for $\Gamma^i(\alpha_i) \in \mathscr{F}^i, i = 1, 2, \ldots, d$. This geometrical depiction allows to see easily which focal elements contribute to the lower or upper probabilities of an event F.

A DS structure $(\mathscr{F}_n, m)$ defined on $\mathbb{R}$ induces a partition $\mathscr{P} = \{(0, \alpha_1], (\alpha_1, \alpha_2], \ldots, (\alpha_{n-1}, \alpha_n]\}, \alpha_n = 1$ of the interval $(0,1]$ where $\alpha_k = \sum_{j=1}^{k} m(A_j)$ for $k = 1, \ldots, n$.We will refer to this partition as *the partition of the interval* $(0,1]$ *associated to* $(\mathscr{F}_n, m)$. This partition forms the *α-representation* of $(\mathscr{F}_n, m)$ which is shown in Figure 5.1; Note that there is a one-to-one relationship between

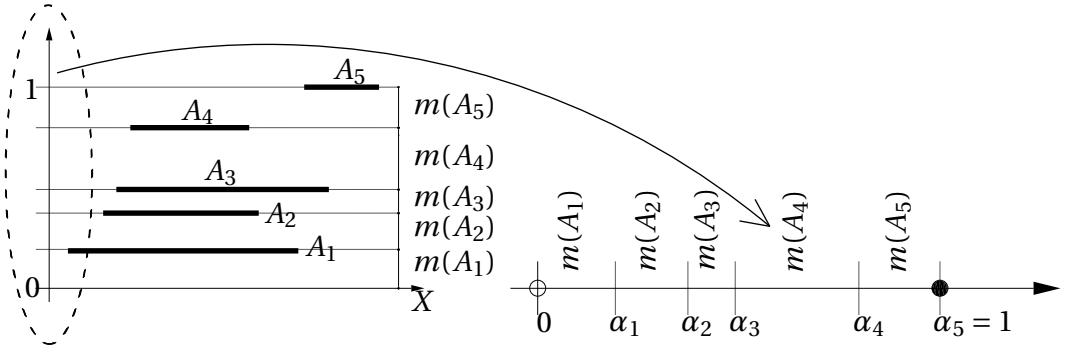


Figure 5.1: Schematic representation a one-dimensional DS structure $(\mathscr{F}, m)$ in the α-space.

$A_j \in \mathscr{F}_n$ and the subinterval $(\alpha_{j-1}, \alpha_j]$, which we will denote by $A_j \leftrightarrow (\alpha_{j-1}, \alpha_j]$ or simply by $A_j \leftrightarrow \alpha_j$; this is possible since $m(A_j) > 0$.

Now, consider the DS structure $(\mathscr{F}_n, m) := \times_{i=1}^{d}\left(\mathscr{F}_{n_i}^i, m^i\right)$ defined as the random relation of d unidimensional DS structures, up to some dependence specification. Every DS structure $\left(\mathscr{F}_{n_i}^i, m^i\right)$ has an associated partition of the interval $(0,1]$, which forms a system of sets $\mathscr{P}^i$ formed by the subintervals of

the partition, i.e., $\mathscr{P}^i = \left\{\left(\alpha_0^i, \alpha_1^i\right], \left(\alpha_1^i, \alpha_2^i\right], \ldots, \left(\alpha_{n_i-1}^i, \alpha_{n_i}^i\right]\right\}$. The direct product $\mathscr{P} = \mathscr{P}^1 \times \mathscr{P}^2 \times \cdots \times \mathscr{P}^d$ will form a rectangular grid of $(0,1]^d$, composed by the disjoint boxes $B_{j_1,j_2,\ldots,j_d} := \times_{i=1}^d \left(\alpha_{j_i-1}^i, \alpha_{j_i}^i\right]$ where $j_i = 1,\ldots,n_i$; this partition of $(0,1]^d$ forms the *α-representation* of $(\mathscr{F}_n, m)$ which is shown in Figure 5.2 in the two dimensional case; note that there is a one-to-one association between

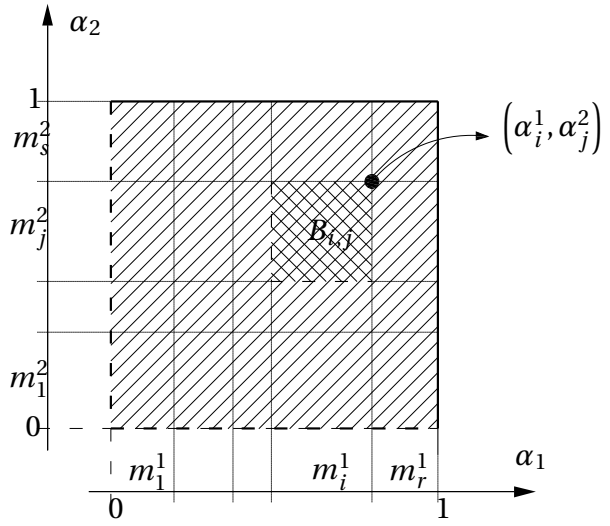


Figure 5.2: Schematic representation a two-dimensional DS structure $(\mathscr{F}, m)$ in the α-space.

$A_{j_1,j_2,\ldots,j_d} \in \mathscr{F}_n$ (here $A_{j_1,j_2,\ldots,j_d} := \times_{i=1}^d A_{j_i}^i$ and $A_{j_i}^i \in \mathscr{F}_{n_i}^i$) and $B_{j_1,j_2,\ldots,j_d}$ which we will denote by $A_{j_1,j_2,\ldots,j_d} \leftrightarrow B_{j_1,j_2,\ldots,j_d}$ or just by the lattice point $\left(\alpha_{j_1}^1, \alpha_{j_2}^2, \ldots, \alpha_{j_d}^d\right)$, i.e. $A_{j_1,j_2,\ldots,j_d} \leftrightarrow \left(\alpha_{j_1}^1, \alpha_{j_2}^2, \ldots, \alpha_{j_d}^d\right)$.

Finally, we will define the image of $\emptyset$ in the X-space to be $\emptyset$ in the α-space and viceversa, i.e., $\emptyset \leftrightarrow \emptyset$.

5.3 Copulas and Dempster-Shafer structures

Another topic that requires particular attention is the role of copulas in the specification of the dependence in between RSs. It was shown in Section 4.2.2 that copulas may specify the dependence information between marginal infinite random sets. In the next lines the relation between copulas and DS structures will be clarified.

5.3.1 Relation between copulas and Dempster-Shafer structures

Ferson et al. (2004, p. 68) showed that the dependence information between basic variables modelled by DS structures can be specified by a copula C. In

fact, they showed that the basic mass assignment associated to the joint focal element $A_{j_1,j_2,\ldots,j_d}$ can be constructed by the V_C volume of its associated box $B_{j_1,j_2,\ldots,j_d}$, that is, $m\left(A_{j_1,j_2,\ldots,j_d}\right) = V_C\left(B_{j_1,j_2,\ldots,j_d}\right)$ where the V_C-volume of the d-box $B_{j_1,j_2,\ldots,j_d} := \times_{i=1}^{d}\left(\alpha^i_{j_i-1}, \alpha^i_{j_i}\right]$ is defined by the d-th order difference of C on $B_{j_1,j_2,\ldots,j_d}$,

$$V_C\left(B_{j_1,j_2,\ldots,j_d}\right) = \Delta_{\alpha^d_{j_d-1}}^{\alpha^d_{j_d}} \Delta_{\alpha^{d-1}_{j_{d-1}-1}}^{\alpha^{d-1}_{j_{d-1}}} \cdots \Delta_{\alpha^1_{j_1-1}}^{\alpha^1_{j_1}} C$$

where

$$\Delta_{\alpha^i_{j_i-1}}^{\alpha^i_{j_i}} C = C\left(\alpha_1,\ldots,\alpha_{i-1},\alpha^i_{j_i},\alpha_{i+1},\ldots,\alpha_d\right) - C\left(\alpha_1,\ldots,\alpha_{i-1},\alpha^i_{j_i-1},\alpha_{i+1},\ldots,\alpha_d\right)$$

for $i = 1,\ldots,d$.

Inasmuch as C is a joint CDF, it follows that V_C denotes the probability of any d-box B, and therefore $V_C(B) \in [0,1]$. All d-boxes in $[0,1]^d$ form a semiring $\mathscr{S}$ that can be extended to a ring $\mathscr{R}$ of all elementary sets that result from the finite union of disjoint elements of $\mathscr{S}$. The measure generated by V_C can be extended by Carathéodory's extension theorem, to the σ-algebra generated by $\mathscr{R}$. In particular that σ-algebra contains all Borel subsets of $[0,1]^d$, that is $[0,1]^d \cap \mathscr{B}^d$. The extension of V_C is unique and is called the *(Lebesgue-) Stieltjes measure* μ_C *corresponding to* C and C is referred to either as the *generating function of* μ_C (in the terminology of Kolmogorov and Fomin (1970)) or as the *distribution function of* μ_C (in the terminology of Loève (1977)), in consequence

$$m\left(A_{j_1,j_2,\ldots,j_d}\right) = \mu_C\left(B_{j_1,j_2,\ldots,j_d}\right). \tag{5.1}$$

Since m is equivalent to V_C and P_Γ is equivalent to μ_C (sometimes denoted in this document as $P_\Gamma \equiv \mu_C$), then P_Γ can also be regarded as the extension of m defined on the σ-algebra $\sigma_{\mathscr{F}}$.

Given a DS structure $(\mathscr{F}, m)$ and its associated partition $\mathscr{P}$ of the α-space, it is possible to observe that there is a family of copulas $\mathscr{C}'$ such that for every $C \in \mathscr{C}'$ we have that equation (5.1) holds, i.e., $m\left(A_{j_1,j_2,\ldots,j_d}\right) = \mu_C\left(B_{j_1,j_2,\ldots,j_d}\right)$ for all $C \in \mathscr{C}'$. In that family $\mathscr{C}'$ there is a unique copula C that can be exactly defined by the information provided by $(\mathscr{F}, m)$, namely

$$C(\boldsymbol{\alpha}) = \int_{\times_{i=1}^{d}[0,\alpha_i]} M(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} \tag{5.2}$$

where $M : [0,1]^d \to [0,1]$ stands for the simple function

$$M(\boldsymbol{\alpha}) = \begin{cases} m\left(A_{j_1,j_2,\ldots,j_d}\right) & \text{if } \boldsymbol{\alpha} \in B_{j_1,j_2,\ldots,j_d}, A_{j_1,j_2,\ldots,j_d} \leftrightarrow B_{j_1,j_2,\ldots,j_d} \\ 0 & \text{otherwise} \end{cases} \tag{5.3}$$

which represents in this case the joint PDF in the α-space associated to C and $(\mathscr{F}, m)$. We will say that this C is the *copula obtained from* $(\mathscr{F}, m)$. Note that this C is absolutely continuous.

5.4 Particularization: random sets of indexable type

The formulation of random sets given in Definition 2.2.1 is very general. We will limit our discussion of random sets to those which represent possibility distributions, probability boxes, CDFs and Dempster-Shafer structures or their joint combinations given a dependence relationship specified by the copula C. On the one hand, in the unidimensional case every focal element $\Gamma(\alpha)$ of $(\mathscr{F}, P_\Gamma)$ defined on $X \subseteq \mathbb{R}$ can be represented by its $\alpha \in (0,1]$ (see Sections 2.4 and 2.2.3). On the other hand, when d basic variables are employed, $X \subseteq \mathbb{R}^d$ and it is required to pick a vector $\boldsymbol{\alpha} \in (0,1]^d$ (as already discussed in Chapter 4). Since there is a one-to-one relationship between $\mathscr{F}$ and $(0,1]^d$, we can represent this relationship taking $\Omega := (0,1]^d$ in Definition 2.2.1. In this case Γ is one-to-one and therefore its inverse $\Gamma^{-1} : \mathscr{F} \to \Omega$ exists. Furthermore, in the same Definition, $P_\Omega(G)$ will be specified by the Lebesgue-Stieljes measure $\mu_C(G)$ of any set $G \in \sigma_\Omega$ and $\sigma_\Omega := (0,1]^d \cap \mathscr{B}^d$, i.e, $P_\Omega := \mu_C$, where C is the copula that specifies the dependence information. Note that in the unidimensional case P_Ω is a probability measure on $\mathbb{R}$ corresponding to the uniform CDF on $(0,1]$, i.e. $F_\alpha(\alpha) = P_\Omega(\alpha \leq \alpha) = \alpha$, for $\alpha \in (0,1]$, and $P_\Omega = P$ in the formulations of Sections 2.4 and 2.2.3. We will say that the random sets with $\Omega := (0,1]^d$, $\sigma_\Omega := (0,1]^d \cap \mathscr{B}^d$, and $P_\Omega := \mu_C$ are of *indexable type*.

As discussed in Section 4.2.3, using the α-representation of random sets of indexable type, the hypercube $(0,1]^d$ contains the regions F_{LP} and F_{UP} which are respectively composed of all those points whose corresponding focal elements are completely contained in the set F or share at least one point with F; it was also shown in Section 4.2.3 that $F_{\text{LP}} \subseteq F_{\text{UP}}$ and that both sets do not depend on the copula C that relates the basic variables $\alpha_1, \ldots, \alpha_d$; in this case, the lower (2.7) and upper (2.8) probabilities of a set F can be computed by the integrals

$$\text{LP}_{(\mathscr{F}, P_\Gamma)}(F) = \int_{(0,1]^d} I\left[\boldsymbol{\alpha} \in F_{\text{LP}}\right] \mathrm{d}C\left(\boldsymbol{\alpha}\right) \tag{5.4}$$

$$\text{UP}_{(\mathscr{F}, P_\Gamma)}(F) = \int_{(0,1]^d} I\left[\boldsymbol{\alpha} \in F_{\text{UP}}\right] \mathrm{d}C\left(\boldsymbol{\alpha}\right) \tag{5.5}$$

provided F_{LP} and F_{UP} are μ_C-measurable sets.

Unidimensional infinite random sets of indexable type have already been analyzed for example by Miranda et al. (2005a, Section 4.1).

5.5 Refinements of finite random sets

In the following, I will define the term refinement for unidimensional finite random sets ($d = 1$) and then I will extend this definition to include joint finite random sets ($d > 1$).

Let $(\mathscr{F}'_r, m')$ and $(\mathscr{F}''_s, m'')$ be two finite random sets defined on $\mathbb{R}$. Suppose

that $(\mathscr{F}''_s, m'')$ has an associated partition $\mathscr{P}''$ of the interval $(0,1]$ which is a refinement of the partition $\mathscr{P}'$ of $(0,1]$ associated to $(\mathscr{F}'_r, m')$, i.e., $\mathscr{P}' \prec \mathscr{P}''$. Suppose also that the subinterval $(\alpha'_{i-1}, \alpha'_i] \in \mathscr{P}'$ was partitioned into k subintervals $(\alpha''_j, \alpha''_{j+1}], (\alpha''_{j+1}, \alpha''_{j+2}], \ldots, (\alpha''_{j+k-1}, \alpha''_{j+k}] \in \mathscr{P}''$; if the focal element $A'_i \in \mathscr{F}'_r$, associated to α'_i, i.e. $A'_i \leftrightarrow \alpha'_i$, contains the union of the focal elements $A''_{i^*} \leftrightarrow \alpha''_{i^*}$ for $i^* = j+1, \ldots, j+k$, i.e.,

$$A'_i \supseteq \cup_{i^*=j+1}^{j+k} A''_{i^*} \tag{5.6}$$

then we will say that $A''_{j+1}, A''_{j+2}, \ldots, A''_{j+k}$ is a *refinement by elimination* (of information) of A'_i; now if the focal element $A'_i \leftrightarrow \alpha'_i$ is such that

$$A'_i \subseteq \cup_{i^*=j+1}^{j+k} A''_{i^*}$$

then we will say that $A''_{j+1}, A''_{j+2}, \ldots, A''_{j+k}$ is a *refinement by addition* (of information) of A'_i. When the focal elements of $(\mathscr{F}'_r, m')$ are refined either by elimination or by addition or by both, then we will say that $(\mathscr{F}''_s, m'')$ is a *refinement* of $(\mathscr{F}'_r, m')$ and we will denote it by $(\mathscr{F}'_r, m') \prec (\mathscr{F}''_s, m'')$. Note that $\mathscr{F}'_r \subseteq \mathscr{F}''_s$ whenever $r \leq s$ does not always hold. Moreover, $m'(A'_i) = \sum_{i^*=j+1}^{j+k} m''(A''_{i^*}) = \alpha'_i - \alpha'_{i-1} = \alpha''_{j+k} - \alpha''_j$.

Figure 5.3 pictures refinements by elimination and by addition, for finite random sets that represent possibility distributions, probability boxes and Dempster-Shafer structures.

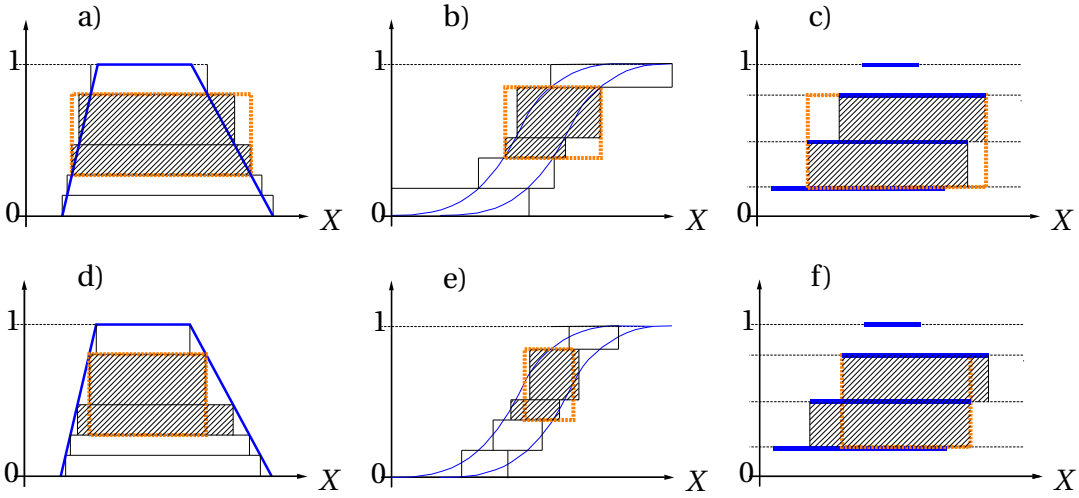


Figure 5.3: Refinement by elimination (Figures a, b and c) and by addition (Figures d, e and f). Here every focal element is represented by a box. The side parallel to the X axis represents the interval associated to the focal element, while the length of the other side represents its basic mass assignment. In each of the figures it is possible to see the approximation of a possibility distribution (Figures a and d), probability boxes (Figures b and e) and Dempster-Shafer structures (Figures c and f) by a finite RS defined on $X \subseteq \mathbb{R}$. Also one particular focal element, the unrefined one, is highlighted (boldline), together with the focal elements belonging to the associated refinement (hatched).

As can be seen in Figure 5.3, on the one hand, refinement by elimination creates new focal elements by eliminating the information of the original focal element and by distributing the original basic mass assignment into the new focal elements; on the other hand, refinement by addition creates new focal elements by adding information to the original focal element and by distributing the original basic mass assignment into the new focal elements, as well. A refinement of a random set can have both types of refinement of the focal elements, however, we will restrict our exposition to refinements by elimination because (5.6) is a necessary condition in the formulation that will be proposed in this chapter. Using both types of refinements is left for future research.

This is a good point to mention the work of Tonon (2004c), who proposed two methods to discretize probability boxes, namely the outer discretization method (ODM, see Figure 5.3b) and the averaging discretization method (ADM, see Figure 5.3e). If a discretization done by the ADM (correspondingly ODM) is refined using again the same method, then the focal elements will be refined by elimination (correspondingly addition). The reader is also referred to the work of Hall and Lawry (2004) for other discretization algorithms.

Now, let us consider the joint finite random sets $(\mathscr{F}'_r, m') := \times_{i=1}^{d} (\mathscr{F}'^i_{r_i}, m'^i)$ and $(\mathscr{F}''_s, m'') := \times_{i=1}^{d} (\mathscr{F}''^i_{s_i}, m''^i)$ defined on $\mathbb{R}^d$. Wi will say that $(\mathscr{F}''_s, m'')$ is a *refinement* of $(\mathscr{F}'_r, m')$ (denoted by $(\mathscr{F}'_r, m') \prec (\mathscr{F}''_s, m'')$) if $(\mathscr{F}''^i_{s_i}, m''^i)$ is a refinement of $(\mathscr{F}'^i_{r_i}, m'^i)$ for $i \in I \subseteq \{1, \ldots, d\}$ while $(\mathscr{F}'^i_{r_i}, m'^i) = (\mathscr{F}''^i_{s_i}, m''^i)$ for $i \in \{1, \ldots, d\} \setminus I$. In other words, a refinement of any of the marginal random sets $(\mathscr{F}'^i_{r_i}, m'^i)$ will imply a refinement of the joint random set $(\mathscr{F}'_r, m')$. Note that the refinement of random sets induces a partial ordering between them. Also, a refinement of the joint random set $(\mathscr{F}'_r, m')$ will imply a refinement of the associated rectangular grid $\mathscr{P}'$ of $(0,1]^d$.

5.6 Convergence of a sequence of finite random sets

Before defining convergence of a sequence of infinite random sets, it is convenient to review some terms that will be employed in this section.

Definition 5.6.1. *Let A and B be two compact subsets of the metric space $\mathbb{R}^d$. Then the* Hausdorff distance *between A and B is defined by*

$$d_H(A,B) := \max\left\{ \sup_{x\in A} \inf_{y\in B} \rho(x,y), \sup_{y\in B} \inf_{x\in A} \rho(x,y) \right\}$$

given a metric $\rho(\cdot,\cdot)$ on $\mathbb{R}^d$.

Note that the Hausdorff distance between not-necessarily-closed subsets can be defined as the Hausdorff distance between their closures. It gives in this case a pseudo-metric on $\mathbb{R}^d$, since the Hausdorff distance between any two sets with the same closure is zero.

Suppose we are given a RS of indexable type Γ (or equivalently $(\mathscr{F}, P_\Gamma)$) defined on the probability space $(\Omega, \sigma_\Omega, \mu_C)$. Let

$$\Gamma_1, \Gamma_2, \ldots, \Gamma_k, \ldots, \tag{5.7}$$

or equivalently

$$(\mathscr{F}_{r_1}^{(1)}, m^{(1)}), (\mathscr{F}_{r_2}^{(2)}, m^{(2)}), \ldots, (\mathscr{F}_{r_k}^{(k)}, m^{(k)}), \ldots, \tag{5.8}$$

be a sequence of finite random sets which are also defined on the same probability space as Γ. Then the sequence (5.7) is said to *converge almost surely* (or simply *converge*) to Γ if

$$\mu_C\left(\omega \in \Omega : \lim_{k\to\infty} d_H(\Gamma(\omega), \Gamma_k(\omega)) = 0\right) = 1 \tag{5.9}$$

(see e.g. (Nguyen, 2006, p. 157–159)). To avoid complications in the convergence of the focal elements, we will consider all of them to be, from now on, closed subsets of $\mathbb{R}^d$.

If (5.7) is a sequence of every time refining finite RSs which are defined on the same probability space as $(\mathscr{F}, P_\Gamma)$, i.e., if

$$(\mathscr{F}_{r_1}^{(1)}, m^{(1)}) \prec (\mathscr{F}_{r_2}^{(2)}, m^{(2)}) \prec \cdots \prec (\mathscr{F}_{r_k}^{(k)}, m^{(k)}) \prec \cdots, \tag{5.10}$$

then the associated sequence $\mathscr{P}_{(1)}, \mathscr{P}_{(2)}, \ldots, \mathscr{P}_{(k)}, \ldots$ of partitions of $(0,1]^d$ is such that $\mathscr{P}_{(k)}$ is a refinement of $\mathscr{P}_{(k-1)}$, $\mathscr{P}_{(k)} := \mathscr{P}_{(1)}^1 \times \mathscr{P}_{(2)}^2 \times \cdots \times \mathscr{P}_{(k)}^d$, denoted by $\mathscr{P}_{(k-1)} \prec \mathscr{P}_{(k)}$, and $\|\mathscr{P}_{(k)}^i\| \to 0$ as $k \to \infty$ for $i = 1, \ldots, d$; in addition, (5.9) can be replaced by the following condition: "for every sequence of boxes $B^{(1)} \supseteq B^{(2)} \supseteq B^{(k)} \supseteq \cdots \supseteq \boldsymbol{\alpha}$ such that $B^{(k)} \in \mathscr{P}_{(k)}$ and $\boldsymbol{\alpha} \in (0,1]^d$, we have that the associated sequence of focal elements $A^{(1)}, A^{(2)}, \ldots, A^{(k)} \ldots$ with $A^{(k)} \leftrightarrow B^{(k)}$ and $A^{(k)} \in \mathscr{F}_{r_k}^{(k)}$ converges in the sense of Hausdorff to $\Gamma(\boldsymbol{\alpha}) \in \mathscr{F}$ for almost all (with regard to μ_C) $\boldsymbol{\alpha} \in (0,1]^d$."

Since the random sets (5.8) and $(\mathscr{F}, P_\Gamma)$ are all defined on $(\Omega, \sigma_\Omega, \mu_C)$ then $m(A^{(k)}) = \mu_C(B^{(k)})$ for all $A^{(k)} \in \mathscr{F}_{r_k}^{(k)}$ and $B^k \in \mathscr{P}^{(k)}$, $A^{(k)} \leftrightarrow B^{(k)}$, $k = 1, 2, \ldots$, and thus the corresponding sequence of copulas $C_1, C_2, \ldots, C_k, \ldots$, such that C_k is the copula obtained from $(\mathscr{F}_{r_k}^{(k)}, m^{(k)})$ (see Section 5.3), converges to the copula C that generates the measure P_Γ by the relation $P_\Gamma \equiv \mu_C$.

It is clear that

$$\mu_{C_k}(B) = \mu_C(B) \tag{5.11}$$

for all $B \in \mathscr{P}^{(k)}$ and all $k = 1, 2, \ldots$.

When a possibility distribution, probability box or CDF is given and it is desired to make computations with them, techniques in random set theory sometimes require to approximate the basic variable by a finite RS $(\mathscr{F}_{r_k}^{(k)}, m^{(k)})$. This approximation can be improved by refining $(\mathscr{F}_{r_k}^{(k)}, m^{(k)})$ in a way that resembles as close as possible the original basic variable.

The following theorem is useful is this context:

Theorem 5.6.2. *Let* $(\mathscr{F}_{r_1}^{(1)}, m^{(1)}) \prec (\mathscr{F}_{r_2}^{(2)}, m^{(2)}) \prec \cdots \prec (\mathscr{F}_{r_k}^{(k)}, m^{(k)}) \prec \cdots$ *be a sequence of every-time-refining finite RS of indexable type defined on* $X \subseteq \mathbb{R}^d$ *which converges to* $(\mathscr{F}, P_\Gamma)$. *Here* $m^{(1)}, m^{(2)}, \ldots, m^{(k)}, \ldots$ *and* P_Γ *are probability measures generated by the copula* C, *i.e.,* $m^{(k)} \equiv \mu_C$ *for all* $k = 1, 2, \ldots$ *and* $P_\Gamma \equiv \mu_C$. *Let* $f : \mathscr{P}(X) \to [0, \infty)$ *be a bounded and continuous function with regard to the Hausdorff metric. Then the sequence*

$$\left\{ \sum_{A^{(k)} \in \mathscr{F}^{(k)}} f\left(A^{(k)}\right) m^{(k)}\left(A^{(k)}\right) \right\}_{k \in \mathbb{N}} \tag{5.12}$$

converges to the unique limit $\int_{(0,1]^d} f(\Gamma(\boldsymbol{\alpha}))\, \mathrm{d}C(\boldsymbol{\alpha})$ *as* $k \to \infty$.

Proof. To each $(\mathscr{F}_{r_k}^{(k)}, m^{(k)})$ corresponds a partition of the α-space $\mathscr{P}_{(k)}$, so that for every $A^{(k)} \in \mathscr{F}^{(k)}$ there exists a corresponding $B^{(k)} \in \mathscr{P}_{(k)}$, i.e., $A^{(k)} \leftrightarrow B^{(k)}$. Then the sequence (5.12) can be rewritten as $\left\{ \sum_{B^{(k)} \in \mathscr{P}_{(k)}} f\left(\Gamma\left(B^{(k)}\right)\right) \mu_C\left(B^{(k)}\right) \right\}_{k \in \mathbb{N}}$. Since $|\mathscr{P}_{(k)}| < \infty$, we have that f takes values on a finite set. The sequence of simple functions $\left\{ f^{(k)}(\boldsymbol{\alpha}) := f\left(\Gamma\left(B^{(k)}\right)\right) \text{ given that } \boldsymbol{\alpha} \in B^{(k)} \text{ for } k = 1, 2, \ldots, |\mathscr{P}_{(k)}| \right\}$ converges to f as $k \to \infty$, inasmuch as $B^{(k)} \supseteq B^{(k+1)}$ for all $k \in \mathbb{N}$ and f is continuous with respect to the Hausdorff metric. Then $f^{(k)}(\boldsymbol{\alpha}) < M$ for all $k \in \mathbb{N}$ and all $\boldsymbol{\alpha} \in (0,1]^d$. Since $\int_{(0,1]^d} M\, \mathrm{d}\mu_C(\boldsymbol{\alpha}) = M < \infty$ Lebesgue's dominated convergence theorem implies that f is integrable on $(0,1]^d$ and $\lim_{k\to\infty} \int_{(0,1]^d} f^{(k)}(\boldsymbol{\alpha})\, \mathrm{d}\mu_C(\boldsymbol{\alpha}) = \int_{(0,1]^d} f(\Gamma(\boldsymbol{\alpha}))\, \mathrm{d}\mu_C(\boldsymbol{\alpha}) = \int_{(0,1]^d} f(\Gamma(\boldsymbol{\alpha}))\, \mathrm{d}C(\boldsymbol{\alpha})$. □

Example: Suppose that the basic variable in consideration is specified by a possibility distribution A, which can be represented as an infinite RS $(\mathscr{F}, P_\Gamma)$; remember that $A(x) = P_\Gamma\{\gamma : x \in \gamma, \gamma \in \mathscr{F}\}$. If the sequence of possibility distributions $A_1, A_2, \ldots, A_k, \ldots$ associated to (5.10), where $A_k(x) = \sum_{A_i^{(k)} \in \mathscr{F}^{(k)}} I[x \in A_i^{(k)}]\, m^{(k)}(A_i^{(k)})$ converges to the possibility distribution A represented by $(\mathscr{F}, P_\Gamma)$, then it can be shown that $A(x) = \lim_{k\to\infty} A_k(x)$ for almost all $x \in X$. In fact, since the problem is unidimensional, $m^{(k)}, k = 1, 2, \ldots$ and P_Γ are related to the uniform distribution on $(0,1]$. Also, let $f_x(\gamma) := I[x \in \gamma]$ which is obviously bounded and continuous with regard to the Hausdorff metric. Using Theorem 5.6.2, the sequence of $A_k(x)$ for $k = 1, 2, \ldots$ will converge to $\int_{(0,1]} f_x(\Gamma(\alpha))\, \mathrm{d}\alpha = \int_{(0,1]} I[x \in \Gamma(\alpha)]\, \mathrm{d}\alpha = P_\Gamma\left\{\gamma : x \in \gamma, \gamma \in \mathscr{F}\right\}$, which is equal to $A(x)$, and $x \in X$ is any point where A is continuous.

Example: Suppose that the basic variable in consideration is specified by a probability box $\langle \underline{F}, \overline{F} \rangle$, which is represented as an infinite RS $(\mathscr{F}, P_\Gamma)$. In this case $\underline{F}(x) = \mathrm{LP}_{(\mathscr{F}, P_\Gamma)}((-\infty, x])$ and $\overline{F}(x) = \mathrm{UP}_{(\mathscr{F}, P_\Gamma)}((-\infty, x])$. If the sequence of probability boxes $\langle \underline{F}_1, \overline{F}_1 \rangle, \langle \underline{F}_2, \overline{F}_2 \rangle, \ldots, \langle \underline{F}_k, \overline{F}_k \rangle, \ldots$ associated to the sequence of finite random sets (5.10), where $\underline{F}_k(x) = \mathrm{LP}_{(\mathscr{F}_{r_k}^{(k)}, m^{(k)})}((-\infty, x])$ and $\overline{F}_k(x) = \mathrm{UP}_{(\mathscr{F}_{r_k}^{(k)}, m^{(k)})}((-\infty, x])$, converges to $(\mathscr{F}, P_\Gamma)$ which represents the probability box $\langle \underline{F}, \overline{F} \rangle$ then it can be shown that $\underline{F}_k \xrightarrow{\mathrm{c}} \underline{F}$ and $\overline{F}_k \xrightarrow{\mathrm{c}} \overline{F}$ as $k \to \infty$ (a sequence

of CDFs $\{F_n\}$ defined on $\mathbb{R}^d$ is said to *converge completely*, if and only if there exists a CDF F such that $F_n(\boldsymbol{x}) \to F(\boldsymbol{x})$ for every $\boldsymbol{x}$ at which the limit function F is continuous; this is written as $F_n \xrightarrow{\text{c}} F$). In fact, like in the case of possibility distributions, since the problem is unidimensional, $m^{(k)}, k = 1,2,\ldots$ and P_Γ are also related to the uniform distribution on $(0,1]$. To show that $\underline{F}_k$ converges completely to $\underline{F}$ as k tends to infinity, we define $f_x(\gamma) := I[\gamma \subseteq (-\infty, x]]$ which is bounded and continuous with regard to the Hausdorff metric. According to Theorem 5.6.2, the sequence of functions $\underline{F}_k(x) = \text{LP}_{(\mathscr{F}_{r_k}^{(k)}, m^{(k)})}((-\infty, x]) = \sum_{A^{(k)} \in \mathscr{F}^{(k)}} f_x\left(A^{(k)}\right) m^{(k)}\left(A^{(k)}\right)$ for $k = 1,2,\ldots$ will converge to $\int_{(0,1]} f_x(\Gamma(\alpha))\,\mathrm{d}\alpha = \int_{(0,1]} I\left[\Gamma(\alpha) \subseteq (-\infty, x]\right]\mathrm{d}\alpha = P_\Gamma\left\{\gamma : \gamma \subseteq (-\infty, x], \gamma \in \mathscr{F}\right\}$, which is equal to $\underline{F}(x) = \text{LP}_{(\mathscr{F}, P_\Gamma)}((-\infty, x])$, where $x \in X$ is any point of continuity of $\underline{F}$. A one-dimensional CDF can have at the most an countable infinite number of discontinuities; in consequence $\underline{F}(x) = \lim_{k\to\infty} \underline{F}_k(x)$ for almost all $x \in X$. Hence $\underline{F}_k \xrightarrow{\text{c}} \underline{F}$. Similar steps must be followed to show that $\overline{F}_k$ converges completely to $\overline{F}$.

Let $(\mathscr{F}_{r_k}^{(k)}, m^{(k)})$ be a finite RS defined on $\mathbb{R}^d$, and let $\mathscr{P}_{(1)}, \mathscr{P}_{(2)}, \ldots, \mathscr{P}_{(k)}, \ldots$ be its associated partitions of $(0,1]^d$. Let $\varphi_k^{\text{LP}} : (0,1]^d \to \{0,1\}$ be the simple function that associates a label to each box of the partition in the following way: if $A_{j_1,j_2,\ldots,j_d}^{(k)} \in \mathscr{F}_{r_k}^{(k)}$ is the focal element associated to the box $B_{j_1,j_2,\ldots,j_d}^{(k)} := \times_{i=1}^{d}\left(\alpha_{j_i-1}^{(k),i}, \alpha_{j_i}^{(k),i}\right] \in \mathscr{P}_{(k)}$, i.e. $A_{j_1,j_2,\ldots,j_d}^{(k)} \leftrightarrow B_{j_1,j_2,\ldots,j_d}^{(k)}$, then

$$\varphi_k^{\text{LP}}(\boldsymbol{\alpha}) := \begin{cases} 1 & \text{if } A_{j_1,j_2,\ldots,j_d}^{(k)} \subseteq F \text{ and } \boldsymbol{\alpha} \in B_{j_1,j_2,\ldots,j_d}^{(k)} \\ 0 & \text{otherwise.} \end{cases} \tag{5.13}$$

An analogous definition can be done for the simple function $\varphi_k^{\text{UP}} : (0,1]^d \to \{0,1\}$, namely,

$$\varphi_k^{\text{UP}}(\boldsymbol{\alpha}) := \begin{cases} 1 & \text{if } A_{j_1,j_2,\ldots,j_d}^{(k)} \cap F \neq \emptyset \text{ and } \boldsymbol{\alpha} \in B_{j_1,j_2,\ldots,j_d}^{(k)} \\ 0 & \text{otherwise.} \end{cases} \tag{5.14}$$

Theorem 5.6.3. *Let*

$$(\mathscr{F}_{r_1}^{(1)}, m^{(1)}) \prec (\mathscr{F}_{r_2}^{(2)}, m^{(2)}) \prec \cdots \prec (\mathscr{F}_{r_k}^{(k)}, m^{(k)}) \prec \cdots \tag{5.15}$$

be a sequence of finite RS that are defined on the same probability space as $(\mathscr{F}, P_\Gamma)$ and that are refined every time by elimination, converging in this way to $(\mathscr{F}, P_\Gamma)$. Let $\varphi_1^{\text{LP}}, \varphi_2^{\text{LP}}, \ldots, \varphi_k^{\text{LP}}, \ldots$ and $\varphi_1^{\text{UP}}, \varphi_2^{\text{UP}}, \ldots, \varphi_k^{\text{UP}}, \ldots$ be the sequences of functions associated to the partitions $\mathscr{P}_{(1)}, \mathscr{P}_{(2)}, \ldots, \mathscr{P}_{(3)}, \ldots$ defined by F and those finite random sets. Then the following is true:

1. *$\varphi_1^{\text{LP}} \leq \varphi_2^{\text{LP}} \leq \cdots \leq \varphi_k^{\text{LP}} \leq \cdots$ and this sequence converges to a function φ^{LP} as $k \to \infty$. Also $\varphi_1^{\text{UP}} \geq \varphi_2^{\text{UP}} \geq \cdots \geq \varphi_k^{\text{UP}} \geq \cdots$ and this sequence converges to a function φ^{UP} as $k \to \infty$.*

2. *If φ_k^{LP} and φ_k^{UP} are the functions associated to the random set $(\mathscr{F}_{r_k}^{(k)}, m^{(k)})$*

then

$$\mathrm{LP}_{(\mathscr{F}_{r_k}^{(k)},m^{(k)})}(F)=\int_{(0,1]^d}\varphi_k^{\mathrm{LP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_{C_k}(\boldsymbol{\alpha}) \tag{5.16}$$

$$\mathrm{UP}_{(\mathscr{F}_{r_k}^{(k)},m^{(k)})}(F)=\int_{(0,1]^d}\varphi_k^{\mathrm{UP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_{C_k}(\boldsymbol{\alpha}). \tag{5.17}$$

where C_k is the copula associated to $m^{(k)}$ by equation (5.2).

3.

$$\int_{(0,1]^d}\varphi^{\mathrm{LP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_C(\boldsymbol{\alpha})=\lim_{k\to\infty}\int_{(0,1]^d}\varphi_k^{\mathrm{LP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_{C_k}(\boldsymbol{\alpha}) \tag{5.18}$$

$$\int_{(0,1]^d}\varphi^{\mathrm{UP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_C(\boldsymbol{\alpha})=\lim_{k\to\infty}\int_{(0,1]^d}\varphi_k^{\mathrm{UP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_{C_k}(\boldsymbol{\alpha}). \tag{5.19}$$

4. $\mathrm{supp}(\varphi^{\mathrm{LP}})\subseteq F_{\mathrm{LP}}$ *and* $F_{\mathrm{UP}}\subseteq\mathrm{supp}(\varphi^{\mathrm{UP}})$. *Also, for a) any absolutely continuous copula C or b) for any copula C with a singular component such that $\mu_C\left(F_{\mathrm{LP}}\setminus\mathrm{supp}(\varphi^{\mathrm{LP}})\right)=0$ (correspondingly $\mu_C\left(\mathrm{supp}(\varphi^{\mathrm{UP}})\setminus F_{\mathrm{UP}}\right)=0$) we have that*[1]*, $\mu_C\left(F_{\mathrm{LP}}\triangle\mathrm{supp}(\varphi^{\mathrm{LP}})\right)=0$, (correspondingly $\mu_C\left(F_{\mathrm{UP}}\triangle\mathrm{supp}(\varphi^{\mathrm{UP}})\right)=0$) and*

$$\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)=\int_{(0,1]^d}\varphi^{\mathrm{LP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_C(\boldsymbol{\alpha})=\int_{(0,1]^d}I[\boldsymbol{\alpha}\in F_{\mathrm{LP}}]\,\mathrm{d}C(\boldsymbol{\alpha}) \tag{5.20}$$

or correspondingly,

$$\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)=\int_{(0,1]^d}\varphi^{\mathrm{UP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_C(\boldsymbol{\alpha})=\int_{(0,1]^d}I[\boldsymbol{\alpha}\in F_{\mathrm{UP}}]\,\mathrm{d}C(\boldsymbol{\alpha}) \tag{5.21}$$

provided F_{LP} and F_{UP} are μ_C-measurable sets.

Proof. If $(\mathscr{F}_{r_k}^{(k)},m^{(k)})\prec(\mathscr{F}_{r_{k+1}}^{(k+1)},m^{(k+1)})$ then $\mathscr{P}_{(k)}\prec\mathscr{P}_{(k+1)}$. Consider $A^{(k)}\in\mathscr{F}_{r_k}^{(k)}$ and its associated $B^{(k)}\in\mathscr{P}_{(k)}$. Suppose also that $A_1^{(k+1)},A_2^{(k+1)},\ldots,A_s^{(k+1)}\in\mathscr{F}_{r_k}^{(k+1)}$ is the refinement by elimination of $A^{(k)}$, i.e., $\cup_{i=1}^s A_i^{(k+1)}\subseteq A^{(k)}$, by equation (5.6). If $B_1^{(k+1)},\ldots,B_s^{(k+1)}$ are the associated disjoint boxes in $\mathscr{P}_{(k+1)}$, then $B^{(k)}=\cup_{i=1}^s B_i^{(k+1)}$. Now in (5.13) if $A^{(k)}\subseteq F$, then $A_i^{(k+1)}\subseteq F$ for any $i=1,\ldots,s$. Also, since $A^{(k)}\not\subseteq F$ does not necessarily imply that $A_i^{(k+1)}\not\subseteq F$ then $\varphi_k^{\mathrm{LP}}\leq\varphi_{k+1}^{\mathrm{LP}}$. The fact that $\varphi_k^{\mathrm{LP}}\leq 1$ for all k, implies that the sequence $\varphi_1^{\mathrm{LP}}\leq\varphi_2^{\mathrm{LP}}\leq\cdots\leq\varphi_k^{\mathrm{LP}}\leq\cdots$ converges everywhere to some φ^{LP} as $k\to\infty$ (see Figure 5.4).

Also, $A^{(k)}\cap F\neq\emptyset$ it does not necessarily imply that $A_i^{(k+1)}\cap F\neq\emptyset$ for any $i=1,\ldots,s$, but $A^{(k)}\cap F=\emptyset$ does imply $A_i^{(k+1)}\cap F=\emptyset$. Therefore, $\varphi_k^{\mathrm{UP}}\geq\varphi_{k+1}^{\mathrm{UP}}$. Since $\varphi_k^{\mathrm{UP}}\geq 0$ for all k, then the sequence $\varphi_1^{\mathrm{UP}}\geq\varphi_2^{\mathrm{UP}}\geq\cdots\geq\varphi_k^{\mathrm{UP}}\geq\cdots$ converges everywhere to some φ^{UP} as $k\to\infty$. This shows Part 1.

[1] The symbol $\triangle$ denotes the *symmetric difference* of two sets, i.e., $A\triangle B:=(A\cup B)\setminus(A\cap B)$.

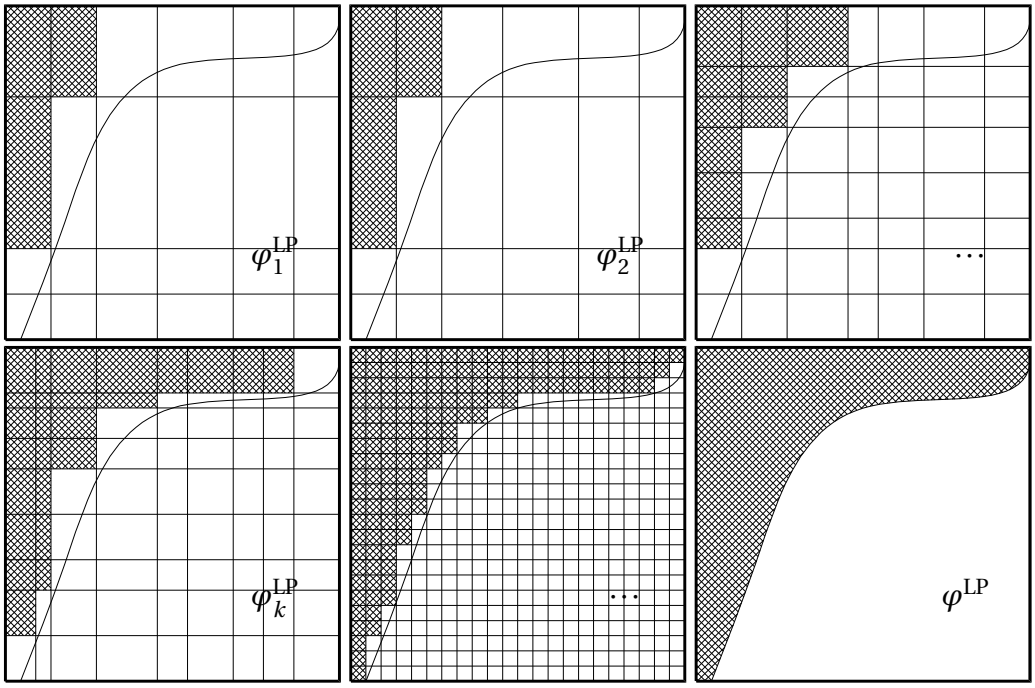


Figure 5.4: Sequence $\varphi_1^{\mathrm{LP}}, \varphi_2^{\mathrm{LP}}, \ldots, \varphi_k^{\mathrm{LP}}, \ldots$ converging to φ^{LP}.

Equations

$$\mathrm{LP}_{(\mathscr{F}_{r_k}^{(k)}, m^{(k)})}(F) = \sum_{j_1}\sum_{j_2}\cdots\sum_{j_d} I\left[A_{j_1,j_2,\ldots,j_d}^{(k)} \subseteq F\right] m^{(k)}\left(A_{j_1,j_2,\ldots,j_d}^{(k)}\right) \tag{5.22}$$

$$\mathrm{UP}_{(\mathscr{F}_{r_k}^{(k)}, m^{(k)})}(F) = \sum_{j_1}\sum_{j_2}\cdots\sum_{j_d} I\left[A_{j_1,j_2,\ldots,j_d}^{(k)} \cap F \neq \emptyset\right] m^{(k)}\left(A_{j_1,j_2,\ldots,j_d}^{(k)}\right) \tag{5.23}$$

follow directly from (2.2) and (2.3). These equations can be restated in the α-space taking into account that $A_{j_1,j_2,\ldots,j_d}^{(k)} \in \mathscr{F}_{r_k}^{(k)}$ has an associated box $B_{j_1,j_2,\ldots,j_d}^{(k)} \in \mathscr{P}_{(k)}$. Note that if $A_{j_1,j_2,\ldots,j_d}^{(k)} \subseteq F$, then the image of $B_{j_1,j_2,\ldots,j_d}^{(k)}$ through φ_k^{LP} equals 1, otherwise equals 0; similar considerations are done for the upper probability; then, using (5.1), equations (5.22) and (5.23) will become

$$\mathrm{LP}_{(\mathscr{F}_{r_k}^{(k)}, m^{(k)})}(F) = \sum_{j_1}\sum_{j_2}\cdots\sum_{j_d} \varphi_k^{\mathrm{LP}}\left(B_{j_1,j_2,\ldots,j_d}^{(k)}\right) \mu_{C_k}\left(B_{j_1,j_2,\ldots,j_d}^{(k)}\right) \tag{5.24}$$

$$\mathrm{UP}_{(\mathscr{F}_{r_k}^{(k)}, m^{(k)})}(F) = \sum_{j_1}\sum_{j_2}\cdots\sum_{j_d} \varphi_k^{\mathrm{UP}}\left(B_{j_1,j_2,\ldots,j_d}^{(k)}\right) \mu_{C_k}\left(B_{j_1,j_2,\ldots,j_d}^{(k)}\right) \tag{5.25}$$

According to the definition of the Lebesgue integral of a simple function equations (5.24) and (5.25) become (5.16) and (5.17) respectively. This shows Part 2.

In accordance with (5.11) we can replace μ_{C_k} by μ_C in equations (5.24) and (5.25) yielding

$$\int_{(0,1]^d} \varphi_k^{\mathrm{LP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_C(\boldsymbol{\alpha}) = \int_{(0,1]^d} \varphi_k^{\mathrm{LP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_{C_k}(\boldsymbol{\alpha})$$

$$\int_{(0,1]^d} \varphi_k^{\mathrm{UP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_C(\boldsymbol{\alpha}) = \int_{(0,1]^d} \varphi_k^{\mathrm{UP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_{C_k}(\boldsymbol{\alpha}).$$

following the steps indicated in the proof of Part 2. Note that $|\varphi_k^{\mathrm{LP}}| \leq 1$ for all $k = 1,2,\ldots$; hence using Lebesgue's dominated convergence theorem, we have that φ^{LP} is integrable on $(0,1]^d$ and (5.18) follows. To show (5.19), similar steps should be performed. This shows Part 3.

Finally, it is required to show that $\mu_C\left(F_{\mathrm{LP}} \triangle \mathrm{supp}(\varphi^{\mathrm{LP}})\right) = 0$. This follows from the definition of convergence of a sequence of finite RSs to an infinite RS. In fact, for almost all (with respect to the μ_C-measure) $\boldsymbol{\alpha} \in (0,1]^d$ such that $\boldsymbol{\alpha} \in F_{\mathrm{LP}}$ there exists a sequence of boxes

$$B^{(1)} \supseteq B^{(2)} \supseteq \cdots \supseteq B^{(k)} \supseteq \cdots \supseteq \boldsymbol{\alpha} \tag{5.26}$$

such that $B^{(k)} \in \mathscr{P}_{(k)}$ with associated focal elements $A^{(1)}, A^{(2)}, \ldots, A^{(k)}, \ldots$ where $A^{(k)} \leftrightarrows B^{(k)}$ and $A^{(k)} \in \mathscr{F}_{r_k}^{(k)}$. This sequence converges to $\Gamma(\boldsymbol{\alpha}) \subseteq F$. Since $\Gamma(\boldsymbol{\alpha}) \subseteq F$ we have that $\varphi^{\mathrm{LP}}(\boldsymbol{\alpha}) = 1$, and therefore $\boldsymbol{\alpha} \in \mathrm{supp}(\varphi^{\mathrm{LP}})$. This implies that $\mathrm{supp}(\varphi^{\mathrm{LP}}) \subseteq F_{\mathrm{LP}}$ and that $\mu_C\left(F_{\mathrm{LP}} \setminus \mathrm{supp}(\varphi^{\mathrm{LP}})\right) = 0$. Note that if $\boldsymbol{\alpha} \in \partial F_{\mathrm{LP}}$ (where $\partial F_{\mathrm{LP}} := \mathrm{closure}(F_{\mathrm{LP}}) \setminus \mathrm{interior}(F_{\mathrm{LP}})$) then sometimes the approximating sequence (5.26) does not exist. In this case $F_{\mathrm{LP}} \setminus \mathrm{supp}(\varphi^{\mathrm{LP}})$ has Lebesgue measure zero, but if C has a singular component, then it may happen that $\mu_C\left(F_{\mathrm{LP}} \setminus \mathrm{supp}(\varphi^{\mathrm{LP}})\right) > 0$, otherwise $\mu_C\left(F_{\mathrm{LP}} \setminus \mathrm{supp}(\varphi^{\mathrm{LP}})\right) = 0$ (see Remark 5.6.5); this will not happen if the copula is absolutely continuous.

Now, $\boldsymbol{\alpha} \in \mathrm{supp}(\varphi^{\mathrm{LP}})$ implies that $\varphi^{\mathrm{LP}}(\boldsymbol{\alpha}) = 1$, which means that there exists an integer $M > 0$ such that $\varphi_k^{\mathrm{LP}}(\boldsymbol{\alpha}) = 1$ for all $k > M$. In consequence, all related $A^{(k)} \leftrightarrows B^{(k)}, \boldsymbol{\alpha} \in B^{(k)}$ for $k > M$ are contained in F. In consequence, those $A^{(k)}$ converge to $\Gamma(\boldsymbol{\alpha}) \subseteq F$, as $k \to \infty$; this means that $\boldsymbol{\alpha} \in F_{\mathrm{LP}}$. This holds for all $\boldsymbol{\alpha} \in \mathrm{supp}(\varphi^{\mathrm{LP}})$.

Finally, since $\mathrm{supp}(\varphi^{\mathrm{LP}}) \setminus F_{\mathrm{LP}}$ and $F_{\mathrm{LP}} \setminus \mathrm{supp}(\varphi^{\mathrm{LP}})$ are disjoint and μ_C is additive then $\mu_C\left(F_{\mathrm{LP}} \triangle \mathrm{supp}(\varphi^{\mathrm{LP}})\right) = 0$. We can make similar considerations to show that $F_{\mathrm{UP}} \subseteq \mathrm{supp}(\varphi^{\mathrm{UP}})$ and that $\mu_C\left(F_{\mathrm{UP}} \triangle \mathrm{supp}(\varphi^{\mathrm{UP}})\right) = 0$.

Equation (5.20) follows from the fact that the functions $g_1(\boldsymbol{\alpha}) = \varphi^{\mathrm{LP}}(\boldsymbol{\alpha})$ and $g_2(\boldsymbol{\alpha}) = I[\boldsymbol{\alpha} \in F_{\mathrm{LP}}]$ are *equivalent with respect to the measure* μ_C, (i.e. they coincide almost everywhere), and that $\mathrm{d}C(\boldsymbol{\alpha})$ is just the symbol employed to denote $\mathrm{d}\mu_C(\boldsymbol{\alpha})$. Similar reasonings can be done with regard to integral and (5.21). This shows Part 4. □

Corollary 5.6.4. *If $(\mathscr{F}', m')$ is a refinement by elimination of $(\mathscr{F}'', m'')$ being both finite RS defined on $\mathbb{R}^d$, and m' and m'' generated by the same copula C, we have that $[\mathrm{LP}_{(\mathscr{F}',m')(F)}, \mathrm{UP}_{(\mathscr{F}',m')}(F)] \subseteq [\mathrm{LP}_{(\mathscr{F}'',m'')(F)}, \mathrm{UP}_{(\mathscr{F}'',m'')}(F)]$ for all $F \subseteq \mathbb{R}^d$.*

Remark 5.6.5. Consider two basic variables X_1 and X_2 described by the triangular possibility distributions $X_1 = X_2 = \mathrm{TriangPos}(0,0,1)$; consider also a closed set $F = \{(x,y) : y \leq x, x, y \in [0,1]\}$ (see Figure 5.5a). Notice that in this case, the corresponding set F_{LP} is also a closed set (see Figure 5.5b). If in addition, X_1 and X_2 have a perfect negative dependence, then they are related by a copula $W(u,v) := \max(u+v-1, 0)$ (see e.g. Nelsen (1999)); this copula is singular and has the support shown in Figure 5.5c. Now if we form a sequence of finite RSs (5.15),

as required by Theorem 5.6.3, then we will obtain in the limit the set $\mathrm{supp}(\varphi^{\mathrm{LP}})$ (see Figure 5.5d). Note that the Lebesgue measure of the set $F_{\mathrm{LP}} \triangle \mathrm{supp}(\varphi^{\mathrm{LP}})$ is zero, that $\mu_C(\mathrm{supp}(\varphi^{\mathrm{LP}})) = 0$ while $\mu_C(F_{\mathrm{LP}}) = 1$ and $\mu_C\left(F_{\mathrm{LP}} \triangle \mathrm{supp}(\varphi^{\mathrm{LP}})\right) = 1 \neq 0$. However if F is open, then F_{LP} will not be closed and $\mu_C\left(F_{\mathrm{LP}} \triangle \mathrm{supp}(\varphi^{\mathrm{LP}})\right) = 0$. This example illustrates the need of checking whether $\mu_C\left(F_{\mathrm{LP}} \setminus \mathrm{supp}(\varphi^{\mathrm{LP}})\right) = 0$ or not.

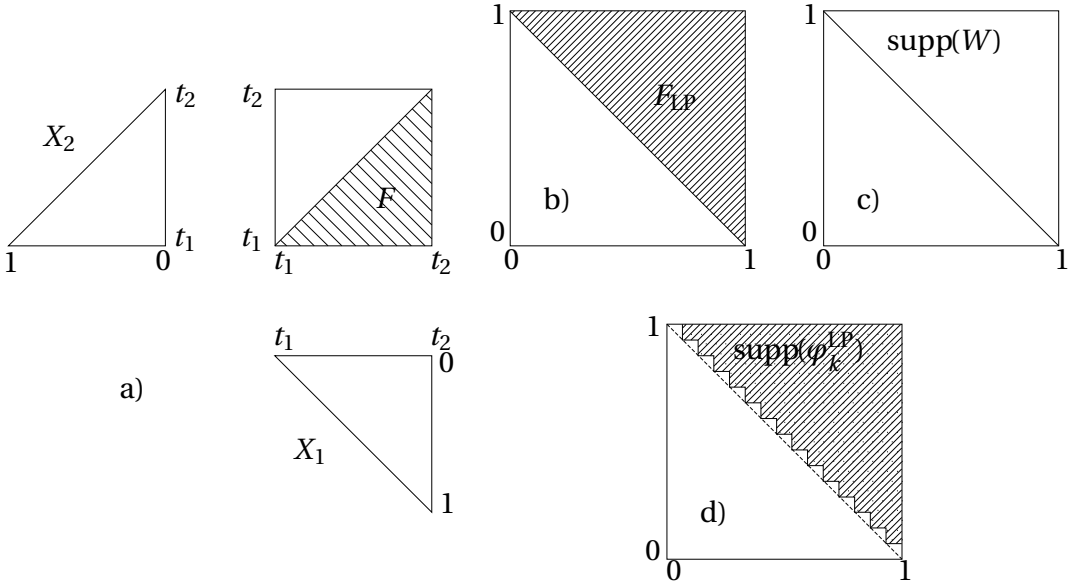


Figure 5.5: Figure considered in Remark 5.6.5. Here Figure d shows also some set φ_k^{LP}.

Chapter 6

From finite to infinite random sets: dependence considerations

In this chapter some dependence considerations in infinite random sets are analyzed, taking as an starting point the analogous for finite random sets, already discussed in Section 2.3. Finally, a relation between infinite random sets and the theories of coherent lower and upper probabilities and expectations is discussed.

6.1 Dependence considerations for infinite random sets

It has been discussed in this document that the copula C models the dependence information. Therefore its analysis requires special attention. With regard to the copula C, we have three possible cases: *a*) we perfectly know it, *b*) we have some information about the copula, or *c*) we don't have any information at all. In the former case, the calculation of (5.4) and (5.5) can be performed using the simulation methods proposed in Chapter 7; in the second and third cases, it is required to develop new strategies to account for this lack of information. These are the problems that will be pursued in this section.

6.1.1 ...when the dependence information is completely known

In this case, the dependence information is modelled by a copula C. Of special importance is the *product copula* Π, i.e.,

$$\Pi(\alpha_1, \alpha_2, \ldots, \alpha_d) = \prod_{i=1}^{d} \alpha_i$$

for $\alpha_i \in [0,1]$ which is used to model non-interactivity between marginal RSs. In fact, the probability measures related to the marginal infinite RSs $\left(\mathscr{F}^X, P_\Gamma^X\right)$ and

$\left(\mathscr{F}^Y, P_\Gamma^Y\right)$ of the random set $\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)$ are called *non-interactive* if and only if for all $\mathscr{G}^X \in \sigma_{\mathscr{F}}^X$, $\mathscr{G}^Y \in \sigma_{\mathscr{F}}^Y$ and $\mathscr{G}^{XY} \in \sigma_{\mathscr{F}}^{XY}$ we have that

$$P_\Gamma^{XY}\left(\mathscr{G}^{XY}\right) = \begin{cases} P_\Gamma^X\left(\mathscr{G}^X\right) P_\Gamma^Y\left(\mathscr{G}^Y\right) & \text{if } \mathscr{G}^{XY} = \mathscr{G}^X \times \mathscr{G}^Y \\ 0 & \text{otherwise} \end{cases} \tag{6.1}$$

This concept is equivalent in the finite case to *random set independence* (see e.g. Couso et al. (1999); Fetz and Oberguggenberger (2004)); in fact, in this case $m\left(A_i^X \times A_j^Y\right) = m^i\left(A_i^X\right) m^j\left(A_j^Y\right)$ for all $i = 1, \ldots, \left|\mathscr{F}^X\right|$ and $j = 1, \ldots, \left|\mathscr{F}^Y\right|$. Since $\mu_{C^{XY}} \equiv P_\Gamma^{XY}$, $\mu_{C^X} \equiv P_\Gamma^X$ and $\mu_{C^Y} \equiv P_\Gamma^Y$, we have from (6.1) that $\mu_{C^{XY}} = \mu_{C^X} \otimes \mu_{C^Y}$, i.e., C^X and C^Y are related by the product copula in order to form the joint copula C^{XY}, that is, $C^{XY} = \Pi\left(C^X, C^Y\right) = C^X C^Y$. In consequence the product copula Π is used if and only if the random sets $\left(\mathscr{F}^X, P_\Gamma^X\right)$ and $\left(\mathscr{F}^Y, P_\Gamma^Y\right)$ are noninteractive (or RS independent).

6.1.2 ...when there is no information at all about the dependence between basic variables: unknown interaction

Here it is natural to ask if the notion of unknown interaction in finite RSs (see e.g. Couso et al. (1999); Fetz and Oberguggenberger (2004)) can be generalized to infinite random sets.

Consider the optimization problem expressed by equations (2.14) to (2.17) (and rewritten here for convenience), i.e., given the joint random set $\left(\mathscr{F}^{(k)}, m^{(k)}\right) := \times_{i=1}^d \left(\mathscr{F}^{(k),i}, m^{(k),i}\right)$, then

$$\underline{P_{\mathrm{f}}}^{(k)}(F) = \min_{m \in \mathscr{M}\left(\mathscr{F}^{(k)}\right)} \sum_{A_{j_1,\ldots,j_d}^{(k)} \in \mathscr{F}^{(k)}} I\left[A_{j_1,\ldots,j_d}^{(k)} \subseteq F\right] m\left(A_{j_1,\ldots,j_d}^{(k)}\right) \tag{6.2}$$

$$\overline{P_{\mathrm{f}}}^{(k)}(F) = \max_{m \in \mathscr{M}\left(\mathscr{F}^{(k)}\right)} \sum_{A_{j_1,\ldots,j_d}^{(k)} \in \mathscr{F}^{(k)}} I\left[A_{j_1,\ldots,j_d}^{(k)} \cap F \neq \emptyset\right] m\left(A_{j_1,\ldots,j_d}^{(k)}\right) \tag{6.3}$$

where $F \subseteq \mathbb{R}^d$ subject to

$$m^{(k),i}\left(A_{j_i}^{(k)}\right) = \sum_{j_1=1}^{\left|\mathscr{F}^{(k),1}\right|} \cdots \sum_{j_{i-1}=1}^{\left|\mathscr{F}^{(k),i-1}\right|} \sum_{j_{i+1}=1}^{\left|\mathscr{F}^{(k),i+1}\right|} \cdots \sum_{j_d=1}^{\left|\mathscr{F}^{(k),d}\right|} m\left(A_{j_1,\ldots,j_d}^{(k)}\right) \tag{6.4}$$

with $A_{j_i}^{(k)} \in \mathscr{F}^{(k),i}$ for all $j_i = 1, \ldots, \left|\mathscr{F}^{(k),i}\right|$, $i = 1, \ldots, d$ and

$$m\left(A_{j_1,\ldots,j_d}^{(k)}\right) \geq 0 \text{ for all } A_{j_1,\ldots,j_d}^{(k)} \in \mathscr{F}^{(k)}. \tag{6.5}$$

The solutions of the optimization problems (6.2) and (6.3) are associated with two basic mass assignments that are just the argument of the minimum and of the maximum of them (not the infimum and the supremum because the sum is performed on a finite set of elements); they will be called henceforth $\underline{m}^{(k)}$ and

$\overline{m}^{(k)}$ respectively. Note that there is no guarantee that $\underline{m}^{(k)}$ and $\overline{m}^{(k)}$ are unique, in other words, there may exist two sets $\underline{\mathscr{M}}^{(k)} \subseteq \mathscr{M}\left(\mathscr{F}^{(k)}\right)$ and $\overline{\mathscr{M}}^{(k)} \subseteq \mathscr{M}\left(\mathscr{F}^{(k)}\right)$ such that for all $\underline{m}^{(k)} \in \underline{\mathscr{M}}^{(k)}$ and for all $\overline{m}^{(k)} \in \overline{\mathscr{M}}^{(k)}$ we obtain $\underline{P_{\mathrm{f}}}^{(k)}(F)$ and $\overline{P_{\mathrm{f}}}^{(k)}(F)$ when replacing them in equations (6.2) and (6.3) respectively.

Since $m\left(A^{(k)}_{j_1,\ldots,j_d}\right) = \mu_C\left(B^{(k)}_{j_1,\ldots,j_d}\right)$, for $A^{(k)}_{j_1,\ldots,j_d} \leftrightarrow B^{(k)}_{j_1,\ldots,j_d}$, we can associate each of the basic mass assignments $m \in \mathscr{M}\left(\mathscr{F}^{(k)}\right)$ with a function M defined by (5.3). This function represents a joint PDF in the α-space associated to some unique copula C by equation (5.2). We will define the set $\mathscr{C}_k$ as the set of all copulas resulting from the integration by means of (5.2) of the $M's$ related to each of the elements m of $\mathscr{M}\left(\mathscr{F}^{(k)}\right)$; note that $\mathscr{C}_k \subseteq \mathscr{C}$.

The optimization problems discussed above can be recast using the α-representation of a finite RS, discussed in Section 5.2, and the relationship between equations (5.22) and (5.23) with (5.24) and (5.25) yielding

$$\underline{P_{\mathrm{f}}}^{(k)}(F) = \min_{C\in\mathscr{C}_k} \sum_{B^{(k)}_{j_1,\ldots,j_d}\in\mathscr{P}^{(k)}} \varphi_k^{\mathrm{LP}}\left(B^{(k)}_{j_1,\ldots,j_d}\right)\mu_C\left(B^{(k)}_{j_1,\ldots,j_d}\right) \tag{6.6}$$

$$\overline{P_{\mathrm{f}}}^{(k)}(F) = \max_{C\in\mathscr{C}_k} \sum_{B^{(k)}_{j_1,\ldots,j_d}\in\mathscr{P}^{(k)}} \varphi_k^{\mathrm{UP}}\left(B^{(k)}_{j_1,\ldots,j_d}\right)\mu_C\left(B^{(k)}_{j_1,\ldots,j_d}\right) \tag{6.7}$$

where $\mathscr{P}^{(k)}$ is the grid associated to $\left(\mathscr{F}^{(k)}, m^{(k)}\right)$. Equations (6.4) and (6.5) are always fulfilled since all copulas have uniform marginals and in consequence

$$\alpha^{(k),i}_{j_i} - \alpha^{(k),i}_{j_i-1} = \sum_{j_1=1}^{\left|\mathscr{F}^{(k),1}\right|} \cdots \sum_{j_{i-1}=1}^{\left|\mathscr{F}^{(k),i-1}\right|} \sum_{j_{i+1}=1}^{\left|\mathscr{F}^{(k),i+1}\right|} \cdots \sum_{j_d=1}^{\left|\mathscr{F}^{(k),d}\right|} \mu_C\left(B^{(k)}_{j_1,\ldots,j_d}\right)$$

for all $i = 1,\ldots,d$ and all copulas are d-increasing by definition, i.e. $\mu_C(R) \geq 0$ for every $R \in (0,1]^d \cap \mathscr{B}^d$.

The solutions of the optimization problems (6.6) and (6.7) are associated with two copulas that are just the argument of the minimum and of the maximum of them; they will be called in the following $\underline{C}_k$ and $\overline{C}_k$ respectively. Similarly to what happens with $\underline{m}^{(k)}$ and $\overline{m}^{(k)}$, $\underline{C}_k$ and $\overline{C}_k$ may not be unique, but just elements of some sets $\underline{\mathscr{C}}_k \subseteq \mathscr{C}_k$ and $\overline{\mathscr{C}}_k \subseteq \mathscr{C}_k$ respectively for which (6.6) and (6.7) reach their optimum values.

Note in addition that $\underline{\mathscr{M}}^{(k)}$, $\overline{\mathscr{M}}^{(k)}$, $\underline{\mathscr{C}}_k$ and $\overline{\mathscr{C}}_k$ depend on the set F; this does not introduce additional difficulties to our formulation inasmuch F is kept "constant" in our analysis.

Equations (6.6) and (6.7) can be rewritten as

$$\underline{P_{\mathrm{f}}}^{(k)}(F) = \min_{C\in\mathscr{C}_k} \int_{(0,1]^d} \varphi_k^{\mathrm{LP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_C(\boldsymbol{\alpha}) \tag{6.8}$$

$$\overline{P_{\mathrm{f}}}^{(k)}(F) = \max_{C\in\mathscr{C}_k} \int_{(0,1]^d} \varphi_k^{\mathrm{UP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_C(\boldsymbol{\alpha}) \tag{6.9}$$

since φ_k^{LP} and φ_k^{UP} are simple functions and $\mathscr{P}^{(k)}$ contains only elementary sets. It is left as an open problem to describe under which conditions the solution of the optimization problems (6.8) and (6.9) for $k = 1, 2, \ldots$ converge, as $k \to \infty$, to the solution of the optimization problems

$$\underline{P_{\mathrm{f}}}(F) = \inf_{C \in \mathscr{C}} \int_{(0,1]^d} \varphi^{\mathrm{LP}}(\boldsymbol{\alpha}) \, \mathrm{d}\mu_C(\boldsymbol{\alpha}) \tag{6.10}$$

and

$$\overline{P_{\mathrm{f}}}(F) = \sup_{C \in \mathscr{C}} \int_{(0,1]^d} \varphi^{\mathrm{UP}}(\boldsymbol{\alpha}) \, \mathrm{d}\mu_C(\boldsymbol{\alpha}) \tag{6.11}$$

The interval $\left[\underline{P_{\mathrm{f}}}(F), \overline{P_{\mathrm{f}}}(F)\right]$ will contain $P_{\mathrm{f}}(F)$ no matter which type of dependence exists between the basic variables, and in addition it will produce the tightest possible bounds in agreement with the available information.

Finally note that equations (6.10) and (6.11) are linear programming problems in infinite dimensions with infinite constraints (see e.g. Anderson and Nash (1987)). After the solution of these optimization problems, one will obtain some families of copulas $\underline{\mathscr{C}}$ and $\overline{\mathscr{C}}$ that produce the extreme values of $P_{\mathrm{f}}(F)$. The analysis of those CDFs may give us intrinsic information about which are the most critical dependence conditions between the implied basic variables.

6.1.3 ...when there is partial information about the dependence between the basic variables

Consider the following example from civil engineering: imagine a concrete structure subject to wind and earthquake loads. Here there can be easily distinguished three groups of basic variables: one related to the concrete properties, another related to the wind loads and another group modeling the earthquake. Clearly these three groups are completely independent, because the earthquake loads have nothing to do with the concrete properties or the wind loads. We can use this dependence information by applying in a first step a *stratification strategy* to separate the basic variables in groups that are clearly mutually independent; these groups are easy to recognize because all variables within one particular group are intended to model similar information. Suppose then that there were distinguished M groups of variables, each group containing i_k variables, for $i = 1, \ldots, M$. In this way, we have that $\sum_{k=1}^{M} i_k = d$. Since these groups are independent and each group has a known or unknown copula that associates the variables within the group, namely $C_1, C_2, \ldots, C_M$, then there exists a "super"-copula C that associates those groups through the product copula, that is,

$$C(\boldsymbol{\alpha}) = \prod_{k=1}^{M} C_k\left(\boldsymbol{\alpha}_{i_k}\right) \tag{6.12}$$

where $\boldsymbol{\alpha} \equiv \left[\boldsymbol{\alpha}_{i_1}, \ldots, \boldsymbol{\alpha}_{i_M}\right]$. The information given by equation (6.12), can be taken into account by including this equation as a constraint to the optimization problems (6.10) and (6.11). In an analogous way, any other knowledge about dependence can be just modelled as constraints to the family of copulas $\mathscr{C}$ in the optimization problems defined in Section 6.1.2.

6.2 Relation of equations (6.10) and (6.11) with the theories of coherent lower and upper probabilities and expectations

The solution to the case of unknown interaction for infinite random sets is related to the theory of imprecise probabilities and the theory of sets of probability measures (see e.g. Walley (1991)).

Note that equations (6.10) and (6.11) can be written as

$$\underline{P_{\mathrm{f}}}(F) = \inf_{C \in \mathscr{C}} \mathrm{LP}_{(\mathscr{F},\mu_C)}(F)$$

and

$$\overline{P_{\mathrm{f}}}(F) = \sup_{C \in \mathscr{C}} \mathrm{UP}_{(\mathscr{F},\mu_C)}(F)$$

respectively, where $\mathrm{LP}_{(\mathscr{F},\mu_C)}(F) := \int_{(0,1]^d} I\left[\boldsymbol{\alpha} \in F_{\mathrm{LP}}\right] \mathrm{d}C(\boldsymbol{\alpha})$ and $\mathrm{UP}_{(\mathscr{F},\mu_C)}(F) := \int_{(0,1]^d} I\left[\boldsymbol{\alpha} \in F_{\mathrm{UP}}\right] \mathrm{d}C(\boldsymbol{\alpha})$; in consequence, they have correspondingly the form

$$\underline{P_f}(F) = \inf_{C \in \mathscr{C}} \mu_C(F_{\mathrm{LP}}) \tag{6.13}$$

$$\overline{P_f}(F) = \sup_{C \in \mathscr{C}} \mu_C(F_{\mathrm{UP}}); \tag{6.14}$$

in fact, when $F_{\mathrm{LP}} = F_{\mathrm{UP}}$ (for example when all basic variables are random), they are a particular case of equations (2.26) and (2.27). Now, the set $\{\mu_C : C \in \mathscr{C}\}$ is contained in the set of all probability measures, therefore RS theory is a particularization of the theory of coherent upper and lower probabilities when we are dealing exclusively with random variables. Observe however that equations (6.13) and (6.14) use two different sets, namely F_{LP} and F_{UP}, to specify uncertainty. The relation between those two sets when $\mu_C(F_{\mathrm{LP}} \triangle F_{\mathrm{UP}}) \neq 0$ and equations (2.26) and (2.27) is not clear in this context.

With regard to the theory of upper and lower expectations, it can be seen that (2.28) and (2.29) have also a close relation with equations (6.10) and (6.11) since they can be rewritten as

$$\underline{P_f}(F) = \inf_{C \in \mathscr{C}} E_C\left[I[\boldsymbol{\alpha} \in F_{\mathrm{LP}}]\right]$$

$$\overline{P_f}(F) = \sup_{C \in \mathscr{C}} E_C\left[I[\boldsymbol{\alpha} \in F_{\mathrm{UP}}]\right]$$

where $E_C[f]$ stands for the expected value of f with respect to the CDF C, i.e., $E_C[f] := \int_{[0,1]^d} f(\boldsymbol{\alpha})\,\mathrm{d}C(\boldsymbol{\alpha})$. It can be seen that they are a particularization of equations (2.28) and (2.29) for $P \in \{\mu_C : C \in \mathscr{C}\}$ a 0-1 loss function l and $F_{\mathrm{LP}} = F_{\mathrm{UP}}$. In consequence, RS theory is also a particularization of the theory of upper and lower expectations when we are dealing exclusively with random variables.

Chapter 7

Discretizations and approximations of solutions for infinite random sets of indexable type

7.1 Introduction

In the framework of Dempster-Shafer evidence theory, authors like Wilson (2000), have on the one hand made remarks on how to diminish the computational load in the calculation of the belief and plausibility of a set F with relation to some Dempster-Shafer structure $(\mathscr{F}_n, m)$, in the context of how to draw basically those focal elements $A_i \in \mathscr{F}_n$ that contribute the most to the belief and plausibility; on the other hand, some other authors (see e.g. Hall and Lawry (2004); Tonon (2004c)) have discussed how to discretize the basic variables, such that the belief and plausibility results will contain the best possible bounds on the uncertainty of some event.

In this chapter, we will deal with both topics. First, we will characterize the different possible types of discretization of basic variables, concluding that the outer discretization is the most conservative way of discretizing the basic variables, because it does not entail loss of information. Also, we will analyze the classical *modi operandi* employed in Dempster-Shafer evidence theory, and we will conclude that the belief and plausibility are just a special type of Riemann-Stieltjes approximations of the Stieltjes integrals that are used to calculate the lower and upper probabilities in the framework of infinite random sets of indexable type (i.e., integrals (5.4) and (5.5)). We will also characterize the different possible types of discretization of basic variables, concluding that the outer discretization is the most conservative way of discretizing the basic variables. In addition, we will analyze the classical *modi operandi* employed in Dempster-Shafer evidence theory, and we will conclude that belief and plausibility are just a special type of Riemann-Stieltjes approximations of the Stieltjes integrals that are used

to calculate the lower and upper probabilities in the framework of infinite random sets of indexable type. Finally, we will study the numerical overhead of the discretization method and the Monte-Carlo-based method proposed in Chapter 4, concluding that, specially in high dimensions, the later is preferable not only because we will avoid the discretization step of the basic variables, but also because we can speed up the evaluation of the lower and upper probabilities by orders of magnitude.

The main result of this chapter is that it is preferable to work directly with infinite random sets of indexable type than with Dempster-Sahfer structures, not only because we will avoid the discretization step of the basic variables, but also because we can speed up the evaluation of the lower and upper probabilities by using the advances in the theory of multidimensional integration, like Monte Carlo simulation.

The plan of this chapter is as follows: in section 7.2, we will deal with discretizations of infinite random sets of indexable type. In section 7.3, we compare the efficiency of different methods employed in the solution of the lower and upper probability integrals, including their approximation by Riemann-Stieltjes sums and the use of Monte Carlo simulation techniques. Finally, section 7.4 presents a numerical example, showing the effectivity of the method.

7.2 Discretization of infinite random sets of indexable type

The use of Dempster-Shafer structures, or equivalently finite random sets requires the discretization of every basic variable in the cases when they represent possibility distributions, probability boxes or continuous CDFs.

A basic variable can be discretized in several ways, which can be gathered into three main groups, outer, inner and intermediate discretization. Suppose that the basic variable X defined on $\mathbb{R}^d$, and represented by the infinite RS of indexable type $(\mathscr{F}, P_\Gamma)$, is to be discretized into a DS structure $(\mathscr{F}_n, m)$. Given a partition $\mathscr{P}$ of $(0,1]^d$, and a box $B_i \in \mathscr{P}$, the focal element $A_i \in \mathscr{F}_n$ such that $A_i \leftrightarrow B_i$, may be defined by one of the three methods of *focal approximation* that are proposed in the following:

- *Outer focal approximation*: in this case

$$A_i := A_i^* = \bigcup_{\alpha \in B_i} \Gamma(\alpha); \tag{7.1}$$

- *Inner focal approximation*: in this case

$$A_i := A_{i*} = \bigcap_{\alpha \in B_i} \Gamma(\alpha) \tag{7.2}$$

  whenever $A_{i*} \neq \emptyset$.

- *Intermediate focal approximation*: here

$$A_{i*} \subseteq A_i \subset A_i^*; \tag{7.3}$$

and $A_i \neq \emptyset$.

In all three cases A_i is m-measurable with

$$m(A_i) = V_C(B_i) > 0. \tag{7.4}$$

We will say that the DS structure $(\mathscr{F}_n, m)$ is an *outer discretization* of $(\mathscr{F}, P_\Gamma)$ if all focal elements were obtained by outer focal approximations; if at least one focal element $A_i \in \mathscr{F}_n$ was defined by an intermediate focal approximation, then we will say that the discretization is *intermediate*. One special case of the intermediate discretization is the *inner discretization*, which happens when all focal elements were defined by inner focal approximations. Figure 7.1 illustrates these types of discretization for the one dimensional case.

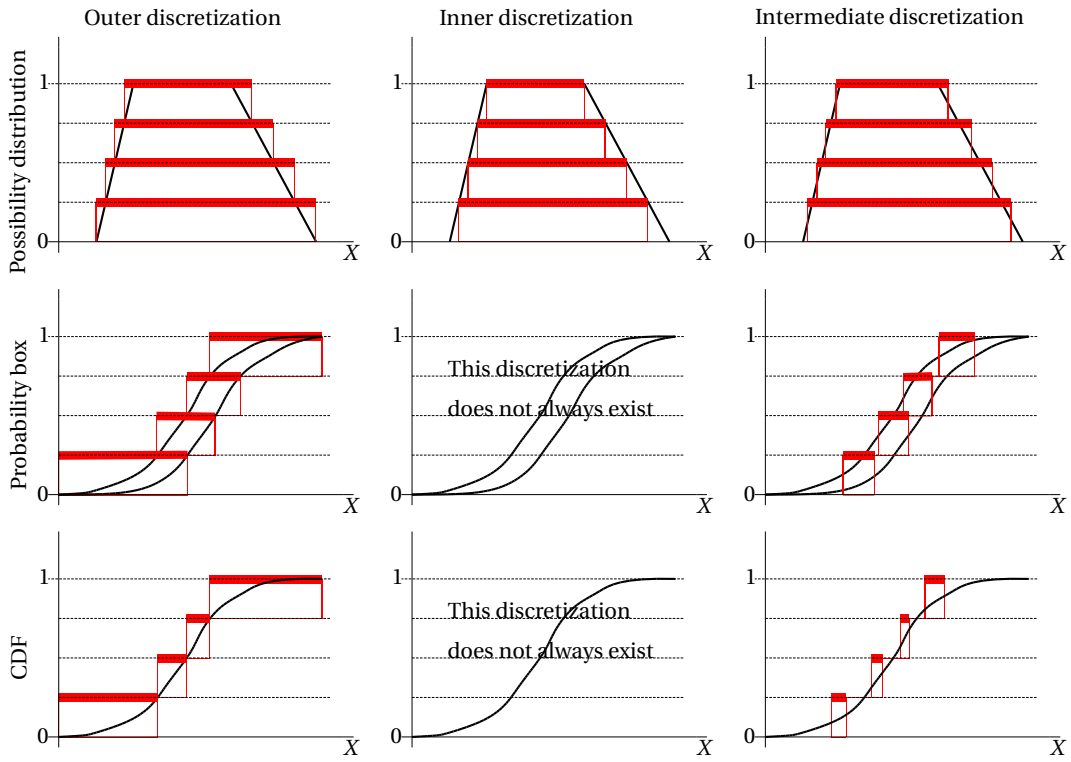


Figure 7.1: Outer, inner and intermediate discretization of CDFs, probability boxes and possibility distributions. Note that the inner discretization of probability boxes sometimes does not exist. The inner discretization of CDFs only exists when the "random" variable is a constant. In this Figure we employ boxes to define the discretization of the basic variable. The side of the boxes parallel to the X axis, and shown here with a thick line, denotes the focal element A_i obtained in the discretization. The other side of the box depicts the basic mass assignment m corresponding to the element, i.e., $m(A_i)$.

Observe that according to equations (7.2) and (7.3) sometimes A_i could be an empty set. In this case the discretization in consideration does not exist inasmuch as $A_i = \emptyset$ is represented in the α-space also by an empty set (see Section 5.2), namely $B_i = \emptyset$, and according to (7.4), then $\mu_C(\emptyset) > 0$, which contradicts the Kolmogorov's second axiom of probability. Also note that A_i does not necessarily belong to $\mathscr{F}$.

It is a good point to comment that the outer discretization of probability boxes coincides with the outer discretization method proposed by Tonon (2004c). Also his averaging discretization method is a particularization of the intermediate discretization of a probability box; also, the upper and lower approximation of possibility distributions described for example by Baudrit et al. (2006), correspond respectively to the outer and inner discretizations of possibility distributions, as shown in Figure 7.1.

We want to analyze what is the relation between the focal elements of $\mathscr{F}_n$ and the regions F_{LP} and F_{UP} for the outer and intermediate types of focal approximation. Then four cases must be analyzed, as shown in Table 7.1; this table shows that

	outer focal approx.	**intermediate focal approx.**
F_{LP}	If $A_i \subseteq F$ then for all $\alpha \in B_i$, $\Gamma(\alpha) \subseteq F$ otherwise $\Gamma(\alpha)$ may or may not be contained in F	In this case, independently of the fact that $A_i \subseteq F$ or $A_i \nsubseteq F$ we have that $\Gamma(\alpha)$ may or may not be contained in F, for all $\alpha \in B_i$
F_{UP}	If $A_i \cap F \neq \emptyset$ then $\Gamma(\alpha)$ may or may not be intersected with F otherwise $\Gamma(\alpha) \cap F = \emptyset$, for all $\alpha \in B_i$	In this case, independently of the fact that $A_i \cap F \neq \emptyset$ or $A_i \cap F = \emptyset$ we have that $\Gamma(\alpha)$ may or may not be intersected with F, for all $\alpha \in B_i$

Table 7.1: Four possible cases of focal approximation of a focal element. The Table depicts the relation between the focal elements of the discretization $(\mathscr{F}_n, m)$ and the regions F_{LP} and F_{UP} generated by $(\mathscr{F}, P_\Gamma)$ and the set F. Here $\alpha \in B_i$ and $A_i \leftrightarrow B_i$.

when a focal element is created by the outer focal approximation, then there are certain relations between its associated box in the α-space and the sets F_{LP} and F_{UP}.

Note that the outer discretization is the most conservative way to discretize a basic variable, inasmuch as the following result holds:

Theorem 7.2.1. *Let $(\mathscr{F}, P_\Gamma)$ be a random set of indexable type and let the DS structure $(\mathscr{F}_n, m)$ be any of its outer discretizations. Then*

$$\left[\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F), \mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)\right] \subseteq \left[\mathrm{Bel}_{(\mathscr{F}_n,m)}(F), \mathrm{Pl}_{(\mathscr{F}_n,m)}(F)\right].$$

Proof. Given some partition $\mathscr{P}$ of $(0,1]^d$, let $(\mathscr{F}_n, m)$ be the outer discretization of $(\mathscr{F}, P_\Gamma)$ associated to $\mathscr{P}$; here $n := |\mathscr{P}|$. Let φ_*^{LP} and φ_*^{UP} be two simple functions

defined on $(0,1]^d$ with range $\{0,1\}$ defined correspondingly by

$$\varphi_*^{\mathrm{LP}}(\boldsymbol{\alpha}) := \begin{cases} 1 & \text{if } A_i \subseteq F \text{ and } B_i \ni \boldsymbol{\alpha} \\ 0 & \text{otherwise.} \end{cases} \tag{7.5}$$

and

$$\varphi_*^{\mathrm{UP}}(\boldsymbol{\alpha}) := \begin{cases} 1 & \text{if } A_i \cap F \neq \emptyset \text{ and } B_i \ni \boldsymbol{\alpha} \\ 0 & \text{otherwise.} \end{cases}$$

Here $A_i \in \mathscr{F}_n$, $B_i \in \mathscr{P}$ and $A_i \leftrightarrow B_i$. Note that the images of B_i, through the above-defined functions, fulfill the relations

$$\varphi_*^{\mathrm{LP}}(B_i) = I[A_i \subseteq F] \tag{7.6}$$

and

$$\varphi_*^{\mathrm{UP}}(B_i) = I[A_i \cap F \neq \emptyset].$$

The fact that $\varphi_*^{\mathrm{LP}}(\boldsymbol{\alpha}) \leq I[\boldsymbol{\alpha} \in F_{\mathrm{LP}}]$ and $\varphi_*^{\mathrm{UP}}(\boldsymbol{\alpha}) \geq I[\boldsymbol{\alpha} \in F_{\mathrm{UP}}]$ for all $\boldsymbol{\alpha} \in (0,1]^d$ (see Figure 7.2), follows easily taking into account that $\varphi_*^{\mathrm{LP}}(\boldsymbol{\alpha}) = 1$ if $\boldsymbol{\alpha} \in B_i$ and $A_i \subseteq F$; however since $\Gamma(\boldsymbol{\alpha}) \subseteq A_i$, by means of equation (7.1), then $\Gamma(\boldsymbol{\alpha}) \subseteq F$, and therefore $I[\boldsymbol{\alpha} \in F_{\mathrm{LP}}] = 1$. Also $I[\boldsymbol{\alpha} \in F_{\mathrm{UP}}] = 1$ if $\Gamma(\boldsymbol{\alpha}) \cap F \neq \emptyset$; but then this implies that $A_i \cap F \neq \emptyset$, for the A_i corresponding to the $B_i \ni \boldsymbol{\alpha}$, and therefore $\varphi_*^{\mathrm{UP}}(\boldsymbol{\alpha}) = 1$. The

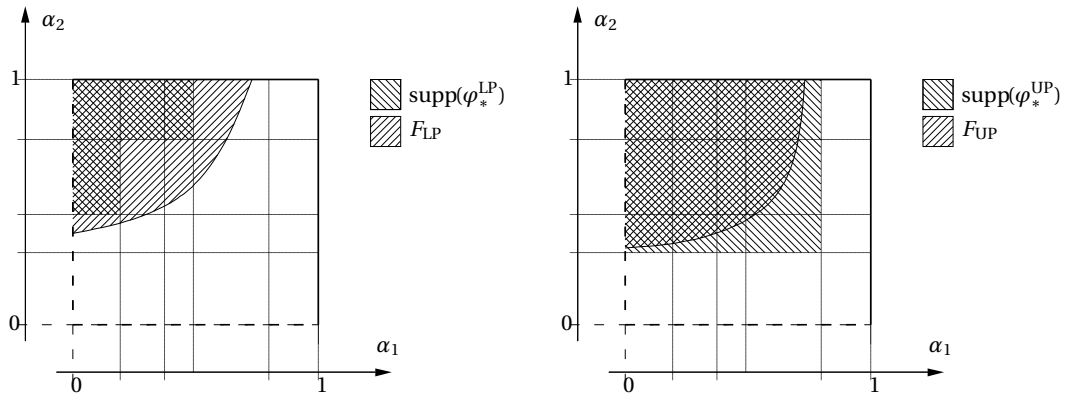


Figure 7.2: Outer discretizations in the α-space. The left picture shows $\mathrm{supp}(\varphi_*^{\mathrm{LP}})$ and F_{LP}; note that $\mathrm{supp}(\varphi_*^{\mathrm{LP}}) \subseteq F_{\mathrm{LP}}$. The right one shows $\mathrm{supp}(\varphi_*^{\mathrm{UP}})$ and F_{UP}; note that $F_{\mathrm{UP}} \subseteq \mathrm{supp}(\varphi_*^{\mathrm{UP}})$.

result that $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) \geq \mathrm{Bel}_{(\mathscr{F}_n,m)}(F)$ follows from the fact that

$$\begin{aligned}
\int_{(0,1]^d} I[\boldsymbol{\alpha} \in F_{\mathrm{LP}}]\,\mathrm{d}C(\boldsymbol{\alpha}) &= \int_{(0,1]^d} I[\boldsymbol{\alpha} \in F_{\mathrm{LP}}]\,\mathrm{d}\mu_C(\boldsymbol{\alpha}) \\
&\geq \int_{(0,1]^d} \varphi_*^{\mathrm{LP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_C(\boldsymbol{\alpha}) \\
&= \sum_{B_i \in \mathscr{P}} \int_{B_i} \varphi_*^{\mathrm{LP}}(\boldsymbol{\alpha})\,\mathrm{d}\mu_C(\boldsymbol{\alpha}) \qquad (7.7) \\
&= \sum_{B_i \in \mathscr{P}} \varphi_*^{\mathrm{LP}}(B_i)\mu_C(B_i) \qquad (7.8) \\
&= \sum_{i=1}^{n} I[A_i \subseteq F]m(A_i) \qquad (7.9)
\end{aligned}$$

The third equality follows from the fact that φ_*^{LP} is, according to (5.13), constant on B_i, and the last one from (5.1) and (7.6) since $A_i \leftrightarrow B_i$. Similar considerations can be done to show that $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F) \leq \mathrm{Pl}_{(\mathscr{F}_n,m)}(F)$. □

NOTE: The same result has been derived in Tonon (2007), using the fact that a random set can be interpreted as a credal set, and thereafter as a probability box $\langle \underline{F}, \overline{F} \rangle$. However, the demonstrations of Tonon (2007) require in addition $\underline{F}$ and $\overline{F}$ to be continuous and strictly monotonically increasing, assumption which is not required inhere, making the proposed solution of more general nature. Finally, as a hint, the above result might also be deduced as well using the notion of natural extensions of Walley (1991).

7.3 Approximation of the lower and upper probability integrals

In most of the cases integrals (5.4) and (5.5) are impossible to solve analytically. Therefore, we require numerical methods to approximate those integrals, like Riemann-Stieltjes sums, or Monte Carlo simulation techniques. Both methods are analyzed in the following.

7.3.1 Approximation of the lower and upper probability integrals using Riemann-Stieltjes sums

Suppose we are given a partition $\mathscr{P}$ of $(0,1]^d$ and a d-dimensional CDF H; let f be a function defined on $(0,1]^d$. For each $B_i \in \mathscr{P}$, let ξ_i be an arbitrary point of B_i. Then any sum of the form

$$S(f,\mathscr{P}) := \sum_{B_i \in \mathscr{P}} f(\xi_i)\mu_H(B_i) \qquad (7.10)$$

is called a *Riemann-Stieltjes sum* of f relative to $\mathscr{P}$ with respect to the CDF H (here we are considering just a special case of the Riemann-Stieltjes sum. More general conditions on H are possible).

For any kind of discretization, given the partition $\mathscr{P}$ of $(0,1]^d$, let $f := \varphi_*^{\mathrm{LP}}$, defined by equation (7.5), and H be the copula C. Then

$$\mathrm{Bel}_{(\mathscr{F}_n,m)}(F) = \sum_{B_i \in \mathscr{P}} \varphi_*^{\mathrm{LP}}(\boldsymbol{\xi}_i)\mu_C(B_i)$$

is a Riemann-Stieltjes sum that approximates integral (5.4), where $\boldsymbol{\xi}_i \in B_i$. This follows from equations (7.7), (7.8) and (7.9) and the fact that the selection of $\boldsymbol{\xi}_i \in B_i$ is irrelevant inasmuch as φ_*^{LP} is constant on B_i. Analogously, if $f := \varphi_*^{\mathrm{UP}}$, as defined by equation (5.14), then

$$\mathrm{Pl}_{(\mathscr{F}_n,m)}(F) = \sum_{B_i \in \mathscr{P}} \varphi_*^{\mathrm{UP}}(\boldsymbol{\xi}_i)\mu_C(B_i)$$

approximates integral (5.5). In other words, the belief and plausibility of a set F with respect to the outer discretization $(\mathscr{F}_n, m)$ of an infinite random set $(\mathscr{F}, P_\Gamma)$ are simply Riemann-Stieltjes sums that approximate integrals (5.4) and (5.5) respectively.

$I[\boldsymbol{\alpha} \in F_{\mathrm{LP}}]$ is a discontinuous function on the boundary ∂F_{LP} of F_{LP} . Therefore, the size of the partition $\mathscr{P}$ together with the regularity of the boundary ∂F_{LP} may have a substantial impact on the discretization error of Riemann-Stieltjes sums. In fact

$$\left| \int_{(0,1]^d} I[\boldsymbol{\alpha} \in F_{\mathrm{LP}}]\, \mathrm{d}C(\boldsymbol{\alpha}) - \sum_{B_i \in \mathscr{P}} \varphi_*^{\mathrm{LP}}\left(\boldsymbol{\xi}_i\right) \mu_C(B_i) \right| \leq \sum_{B_i \in \mathscr{P}} I[B_i \cap \partial F_{\mathrm{LP}} \neq \emptyset]\, \mu_C(B_i)$$

for any kind of approximation φ_*^{LP} and for any $\boldsymbol{\xi}_i \in B_i$. Similar considerations can be done with regard to the calculation of the plausibility.

With regard to equation (7.10), suppose that $C = \Pi$ (remember that in Section 6.1.1 we showed that the copula Π is useful to model random set independence), then μ_C reduces to the Lebesgue measure and then the Riemann-Stieltjes sums become Riemann sums. According to Krommer and Ueberhuber (1998), the convergence of multivariate integration rules based on simple d-dimensional Riemann sums is unacceptably slow for most practical problems.

Suppose now that in addition to $C = \Pi$, f is a continuous function whose second derivatives $\partial^2 f/\partial\alpha_i^2$ for $i = 1,\ldots,d$ are also continuous on $(0,1]^d$ and $\mathscr{P}$ is a *regular grid*, i.e. it is composed of hypercubes of the same size. The trapezoidal rule can be used to integrate f and in this case, if N is the number of discretizations in every variable, then the accuracy of the approximation is of the order $\mathscr{O}(n^{-2/d})$ (see e.g. Niederreiter (1992)), where $n := N^d$ is the number of evaluations of f required. Thus, this algorithm is efficient only in very low dimensions. For example, according to Niederreiter (1992), to guarantee an absolute error of less than 0.01, we must roughly use 10^d nodes, which clearly shows the curse of dimensionality in the method. If $\partial^4 f/\partial\alpha_i^4$ is continuous for $i = 1,\ldots,d$, changing the method to Simpson's rule of integration, does not improve much things, inasmuch as in this case the error bound will be of the order $\mathscr{O}(n^{-4/d})$.

Another trouble with the Riemann-Stieltjes sums strategy is that one has to decide beforehand how fine the discretization $\mathscr{P}$ that defines $(\mathscr{F}_n, m)$ is; then one is compelled to evaluating all of the focal elements A_i in $\mathscr{F}_n$. In other words, with a grid it is not convenient to sample until some convergence or termination criterion is met. Monte Carlo simulation techniques overcome this problem.

7.3.2 Approximation of the lower and upper probability integrals using Monte Carlo simulation techniques

For decades, Monte Carlo simulation (MCS) techniques have been reliable and robust methods to estimate integrals when their analytical solution is too complicated (see e.g. Refs. Fishman (1996); Rubinstein and Kroese (2007)). Their performance depends on how the points $\boldsymbol{\alpha_i} \in (0,1]^d, i = 1,\ldots,n$ are sampled. These points can be chosen for example by simple Monte Carlo sampling, Latin hypercube sampling, importance sampling, quasi Monte Carlo sampling, among others.

Simple Monte Carlo sampling

Section 4.3 shows how to approximate integrals (5.4) and (5.5) by means of simple Monte Carlo sampling. Basically, the method consists in sampling n points from the copula C, namely $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ (Nelsen (1999) provides methods to do it), and then retrieve the corresponding focal elements $A_i := \Gamma(\boldsymbol{\alpha}_i), i = 1,\ldots,n$ from $\mathscr{F}$, to form the focal set $\mathscr{F}_n$. Thereafter to each of those focal elements in $\mathscr{F}_n$, a basic mass assignment of $m(A_i) := 1/n$ is given. Then integrals (5.4) and (5.5) are estimated by computing the belief and plausibility of F with respect to $(\mathscr{F}_n, m)$ using (2.2) and (2.3), that is,

$$\hat{\mathrm{LP}}_{(\mathscr{F},P_\Gamma)}(F) = \mathrm{Bel}_{(\mathscr{F}_n,m)}(F) = \frac{1}{n}\sum_{j=1}^{n} I\left[A_j \subseteq F\right] = \frac{1}{n}\sum_{j=1}^{n} I\left[\boldsymbol{\alpha}_j \in F_{\mathrm{LP}}\right]$$

$$\hat{\mathrm{UP}}_{(\mathscr{F},P_\Gamma)}(F) = \mathrm{Pl}_{(\mathscr{F}_n,m)}(F) = \frac{1}{n}\sum_{j=1}^{n} I\left[A_j \cap F \neq \emptyset\right] = \frac{1}{n}\sum_{j=1}^{n} I\left[\boldsymbol{\alpha}_j \in F_{\mathrm{UP}}\right]$$

Also, in Section 4.3 we show that $\mathrm{Bel}_{(\mathscr{F}_n,m)}(F)$ and $\mathrm{Pl}_{(\mathscr{F}_n,m)}(F)$ are unbiased estimators of $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$ respectively.

Let C be a copula; let $f : \Omega \to \mathbb{R}$ be any random variable defined in the probability space $(\Omega, \sigma_\Omega, \mu_C)$, where $\Omega := (0,1]^d$, $\sigma_\Omega := (0,1]^d \cap \mathscr{B}^d$ and μ_C is the Lebesgue-Stieltjes measure corresponding to the copula C. The expected value of the random variable f is defined by

$$E(f) = \int_\Omega f(\boldsymbol{\alpha})\, \mathrm{d}\mu_C(\boldsymbol{\alpha}).$$

provided f is μ_C-integrable. Using the Monte Carlo method, this value can be estimated by

$$E(f, W_n) = \frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{\alpha}_i)$$

where $W_n := \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n\} \in \Omega$ is a set of samples from Ω which are independent and identically distributed according to the copula C.

By the strong law of large numbers, if n tends to infinity, then $E(f, W_n)$ will tend to $E(f)$ μ_C^∞-almost everywhere, i.e.,

$$\mu_C^\infty\left(\boldsymbol{\omega} \in \Omega^\infty : \lim_{n\to\infty} E_n(f, \boldsymbol{\omega}) = E(f)\right) = 1$$

where $\boldsymbol{\omega} := (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n, \ldots) \in \Omega^\infty =: \times_{i=1}^{\infty}\Omega$, $E_n(f, \boldsymbol{\omega}) := E(f, \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}) = E(f, W_n)$ and μ_C^∞ is the product measure of countable many copies of μ_C, i.e., $\mu_C^\infty = \bigotimes_{i=1}^{\infty}\mu_C$.

We would like to estimate the error in the approximation of $E(f)$ by $E(f, W_n)$ for some sample W_n. The variance of the random variable f,

$$\sigma^2(f) := \int_\Omega \big(f(\boldsymbol{\alpha}) - E(f)\big)^2 \,\mathrm{d}\mu_C(\boldsymbol{\alpha}),$$

is finite provided that $\int_\Omega |f(\boldsymbol{\alpha})|^2 \,\mathrm{d}\mu_C(\boldsymbol{\alpha}) < \infty$; therefore, for any $n \geq 1$, it follows that (Niederreiter (1992, p. 4)),

$$\int_\Omega \cdots \int_\Omega \left(\frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{\alpha}_i) - E(f)\right)^2 \mathrm{d}\mu_C(\boldsymbol{\alpha}_1)\ldots\mathrm{d}\mu_C(\boldsymbol{\alpha}_n) = \frac{\sigma^2(f)}{n}.$$

This equation says that the error in the estimation of $E(f)$ is on average $\sigma(f)/\sqrt{n}$ where $\sigma(f) := \sqrt{\sigma^2(f)}$. If $\sigma(f) < \infty$, the central limit theorem states that the probability that the Monte Carlo estimate of $E(f)$, $E(f, W_n)$, lies between $E(f) - a\sigma(f)/\sqrt{n}$ and $E(f) + b\sigma(f)/\sqrt{n}$ satisfies

$$\lim_{n\to\infty} \mu_C^\infty\left(\boldsymbol{\omega} \in \Omega^\infty : -a\frac{\sigma(f)}{\sqrt{n}} \leq \frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{\alpha}_i) - E(f) \leq b\frac{\sigma(f)}{\sqrt{n}}\right) = \frac{1}{\sqrt{2\pi}}\int_{-a}^{b} \exp\left(-\frac{t^2}{2}\right)\mathrm{d}t \tag{7.11}$$

In consequence, choosing n points independently and identically distributed according to the copula C, leads to an error term in Monte Carlo integration of the order $\mathcal{O}(n^{-1/2})$. Note that this error estimate is not a rigorous bound, but a probabilistic bound; thus the error term should be taken only as a rough indicator of the probable error. Also, this error estimate is independent of the dimension, and therefore it does not have the curse of dimensionality. This error term should be compared with the $\mathcal{O}(n^{-2/d})$ error bound for the Riemann sum, discussed in the previous section, when f is a continuous function whose second derivatives $\partial^2 f/\partial\alpha_i^2$ for $i = 1, \ldots, d$ are also continuous on $(0,1]^d$ and $C = \Pi$. Note that the Monte Carlo method only requires the function $|f|^2$ to be μ_C-integrable.

Note that in our particular case

$$f(\boldsymbol{\alpha}) := I[\boldsymbol{\alpha} \in F_{\mathrm{LP}}] \tag{7.12}$$

Basic Variable i	**Representation of** $(\mathscr{F}^i, P_\Gamma^i)$
1, 2, 3	$\langle$UnifCDF(-1.0, 0.5), UnifCDF(-0.5, 1.0)$\rangle$
4, 5, 6	TriangCDF(0.7, 1.0, 1.3)
7, 8, 9	TrapzPD(0.0, 0.4, 0.8, 1.2)
10, 11, 12	TriangPD(-0.25, 0.5, 1.25)

Table 7.2: Marginal basic variables considered in the numerical example of Section 7.4. Basic variables 1, 2 and 3 are described by a probability box defined by uniform CDFs, variables 4, 5 and 6 are represented by a triangular CDFs, variables 7, 8 and 9 are defined by trapezoidal possibility distributions and variables 10, 11 and 12 are defined by triangular possibility distributions.

or

$$f(\boldsymbol{\alpha}) := I[\boldsymbol{\alpha} \in F_{\mathrm{UP}}]. \tag{7.13}$$

This shows that use of simple Monte Carlo methods for the approximation of integrals (5.4) and (5.5) has as advantage over the discretization method its lack of sensibility to the dimensionality of the problem. Note that when f is a continuous function whose second derivatives $\partial^2 f/\partial \alpha_i^2$ for $i = 1,\ldots,d$ are also continuous, $C = \Pi$ and the discretization is formed by Riemann sums then, the Monte Carlo method is more efficient for dimensions $d \geq 5$.

In practice, one does not have the exact value of $\sigma(f)$, and for some sample W_n one uses the estimate

$$\hat{\sigma}(f) = \left(\frac{1}{n-1} \sum_{i=1}^{n} \big(f(\boldsymbol{\alpha}_i) - E(f, W_n) \big)^2 \right)^{1/2} \tag{7.14}$$

instead. In the particular case when f is random variable that represents a Bernoulli trial (i.e., the range of f is $\{0,1\}$), like in the case of equations (7.12) and (7.13), then equation(7.14) simplifies to

$$\hat{\sigma}(f) = \left(\frac{n}{n-1} E(f, W_n) \big(1 - E(f, W_n) \big) \right)^{1/2}. \tag{7.15}$$

7.4 Numerical example

In order to compare the proposed approach with the classical approach, namely to discretize the basic variables and thereafter form Dempster-Shafer structures, we will set up the following problem.

Consider the marginal basic variables expressed by the random sets $(\mathscr{F}^i, P_\Gamma^i)$, $i = 1,2,\ldots,12$ described in Table 7.2 and depicted in Figure 7.3. Our problem consists in assessing the lower and upper probabilities of the set $F := [0,1]^{12}$ with respect to the joint random set $(\mathscr{F}, P_\Gamma) := \times_{i=1}^{12} (\mathscr{F}^i, P_\Gamma^i)$ using the dependence relations given by the copulas:

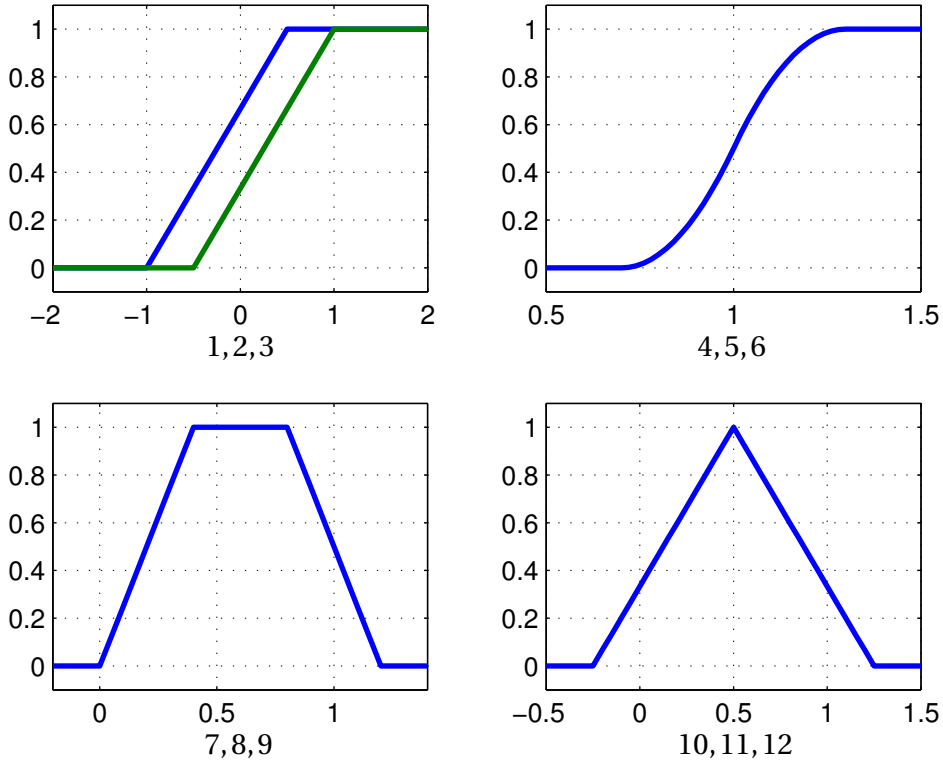


Figure 7.3: Marginal basic variables described in Table 7.2.

- Product copula: $C_1(\boldsymbol{u}) := \prod_{i=1}^{d} u_i$;
- Clayton copula: $C_2(\boldsymbol{u}) := \left(\sum_{i=1}^{d} u_i^{-\delta} - d + 1\right)^{-1/\delta}$, using $\delta = 1.0$;
- Gumbel copula: $C_3(\boldsymbol{u}) := \exp\left(-\left(\sum_{i=1}^{d}(-\ln u_i)^{\delta}\right)^{1/\delta}\right)$, using $\delta = 2.0$

In all three cases $d = 12$. For comparative purposes, the problem will be solved analytically, and with the discretization and Monte Carlo sampling methods.

7.4.1 Analytical solution

The simplicity of the problem allows us to calculate $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$ exactly, by

$$\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) = V_C(B_{\mathrm{LP}}) \tag{7.16}$$

$$\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F) = V_C(B_{\mathrm{UP}}) \tag{7.17}$$

where $B_{\mathrm{LP}} = [2/3,1]^3 \times (0,1/2]^3 \times [1/2,1]^3 \times [1/3,1]^3$ and $B_{\mathrm{UP}} = [1/3,1]^3 \times [0,1/2]^3 \times (0,1]^3 \times (0,1]^3$. Table 7.3 shows the values of $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$ for the different copulas.

Copula	$\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$	$\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$
C_1	0.00017147	0.037037
C_2	0.00020998	0.025
C_3	0.00004918	0.027383

Table 7.3: Exact values of $\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$ for the different copulas, in the numerical example considered in Section 7.4. These values were calculated using equations (7.16) and (7.17)

7.4.2 The discretization method

The lower and upper probabilities (7.16) and (7.17) were estimated by discretizing the basic variables using the outer and intermediate discretization of $(\mathscr{F},P_\Gamma)$, forming the corresponding Dempster-Shafer structure $(\mathscr{F}_n,m)$ as described in Section 7.2 and then estimating the corresponding belief $\mathrm{Bel}_{(\mathscr{F}_n,m)}(F)$ and plausibility $\mathrm{Pl}_{(\mathscr{F}_n,m)}(F)$ by equations (2.2) and (2.3) respectively. Remember that in this case the basic mass assignment of the focal element $A_{i1,\ldots,i12}$ is calculated using equation (5.1). In order to test the influence of the size and type of the discretization, each basic variable was discretized using $N = 2$, 3, 4 and 5 focal elements, as explained in Table 7.4. The results of the evaluation using Dempster-Shafer structures are shown in Table 7.5.

The estimations of the lower and upper probabilities using discretization of the basic variables requires the calculation of the V_C-volume of the corresponding B-box of every focal element. This induces in the computation additional overhead, since the copula must be evaluated $(2N)^d$-times (N^d times for each focal element and 2^d times for each evaluation of V_C, using equation (3.3), in every focal element). This operation becomes specially slow when either the size of the discretization N or the dimension d or both increase and is extremely slow when the copula does not have an explicit formula like in the case of the Gaussian or the Student's t copulas (however in this particular case, it is easy to simulate samples from those copulas). For example, in the case of the evaluation of Table 7.5 using the Clayton copula, we employed 29 seconds, 21 minutes and 10.3 hours for discretization sizes N of 2, 3 and 4 respectively using a personal computer with an AMD Processor of 2.4 GHz of velocity and 512 Mb of RAM running Fedora Core 4 Linux, C++ and the GNU Scientific Library. The estimated time for a discretization of size $N = 5$ was approximately 6 days. However, we calculated this number employing parallel computing using the LAM/MPI library and a cluster of 8 machines). Note that five is a relatively low number of discretizations to represent adequately a basic variable, compared for example with the one hundred discretizations proposed for example by Ferson et al. (2003b, p. 18) in the definition of the canonical Dempster-Shafer structure employed in the description of probability boxes.

Finally, analyzing the results in Table 7.5 one can observe that the outer discretization approach gives more conservative results than the intermediate discretization method; however, in this particular example, the results obtained

with the intermediate discretization strategy were a little bit more accurate.

7.4.3 The Monte Carlo sampling method

The probabilities (7.16) and (7.17) were also estimated using simple Monte Carlo sampling techniques, as shown in Table 7.6. Simple methods to simulate from the Clayton and the Gumbel copulas are provided in McNeil et al. (2005).

In contrast to the discretization method, the evaluation of Table 7.6 using 5000000 Monte Carlo samples and the personal computer last just 20 seconds, several orders of magnitude faster than the discretization approach. As already commented, these estimator are unbiased; note for example that even with 50000 Monte Carlo samples, the evaluation is more accurate using copulas C_2 and C_3 than the evaluation using approximately 244.1 million of focal elements ($N = 5$) with the discretization method.

In addition, Table 7.6 shows that the accuracy of the Monte Carlo simulation approach depends on the number of simulations and the exact value of the lower and upper probabilities. In a further step this value can be employed in equation (7.11) to estimate probabilistic bounds (in comparison to the hard bounds of the outer discretization method, which can be estimated at the expense of an increasingly high computational cost) on the value of the lower and upper probabilities.

These results discourage the use of the discretization of the basic variables specially in high dimensional problems and suggest instead to use the Monte Carlo approximation methods, not only because the latter is much faster but also more accurate than the former technique. In higher dimensions even more striking differences are to be expected. In low dimensions ($d \leq 4$), the discretization method can be efficient as well.

Outer discretization				
N	1, 2, 3	4, 5, 6	7, 8, 9	10, 11, 12
2	[-0.2500, 1.0000]	[1.0000, 1.3000]	[0.2000, 1.0000]	[0.1250, 0.8750]
	[-1.0000, 0.2500]	[0.7000, 1.0000]	[0.0000, 1.2000]	[-0.2500, 1.2500]
3	[0.0000, 1.0000]	[1.0551, 1.3000]	[0.2667, 0.9333]	[0.2500, 0.7500]
	[-0.5000, 0.5000]	[0.9449, 1.0551]	[0.1333, 1.0667]	[0.0000, 1.0000]
	[-1.0000, 0.0000]	[0.7000, 0.9449]	[0.0000, 1.2000]	[-0.2500, 1.2500]
4	[0.1250, 1.0000]	[1.0879, 1.3000]	[0.3000, 0.9000]	[0.3125, 0.6875]
	[-0.2500, 0.6250]	[1.0000, 1.0879]	[0.2000, 1.0000]	[0.1250, 0.8750]
	[-0.6250, 0.2500]	[0.9121, 1.0000]	[0.1000, 1.1000]	[-0.0625, 1.0625]
	[-1.0000, -0.1250]	[0.7000, 0.9121]	[0.0000, 1.2000]	[-0.2500, 1.2500]
5	[0.2000, 1.0000]	[1.1103, 1.3000]	[0.3200, 0.8800]	[0.3500, 0.6500]
	[-0.1000, 0.7000]	[1.0317, 1.1103]	[0.2400, 0.9600]	[0.2000, 0.8000]
	[-0.4000, 0.4000]	[0.9683, 1.0317]	[0.1600, 1.0400]	[0.0500, 0.9500]
	[-0.7000, 0.1000]	[0.8897, 0.9683]	[0.0800, 1.1200]	[-0.1000, 1.1000]
	[-1.0000, -0.2000]	[0.7000, 0.8897]	[0.0000, 1.2000]	[-0.2500, 1.2500]

Intermediate discretization				
N	1, 2, 3	4, 5, 6	7, 8, 9	10, 11, 12
2	[0.1250, 0.6250]	[1.0000, 1.3000]	[0.3000, 0.9000]	[0.3125, 0.6875]
	[-0.6250, -0.1250]	[0.7000, 1.0000]	[0.1000, 1.1000]	[-0.0625, 1.0625]
3	[0.2500, 0.7500]	[1.0551, 1.3000]	[0.3333, 0.8667]	[0.3750, 0.6250]
	[-0.2500, 0.2500]	[0.9449, 1.0551]	[0.2000, 1.0000]	[0.1250, 0.8750]
	[-0.7500, -0.2500]	[0.7000, 0.9449]	[0.0667, 1.1333]	[-0.1250, 1.1250]
4	[0.3125, 0.8125]	[1.0879, 1.3000]	[0.3500, 0.8500]	[0.4062, 0.5938]
	[-0.0625, 0.4375]	[1.0000, 1.0879]	[0.2500, 0.9500]	[0.2188, 0.7812]
	[-0.4375, 0.0625]	[0.9121, 1.0000]	[0.1500, 1.0500]	[0.0312, 0.9688]
	[-0.8125, -0.3125]	[0.7000, 0.9121]	[0.0500, 1.1500]	[-0.1562, 1.1562]
5	[0.3500, 0.8500]	[1.1103, 1.3000]	[0.3600, 0.8400]	[0.4250, 0.5750]
	[0.0500, 0.5500]	[1.0317, 1.1103]	[0.2800, 0.9200]	[0.2750, 0.7250]
	[-0.2500, 0.2500]	[0.9683, 1.0317]	[0.2000, 1.0000]	[0.1250, 0.8750]
	[-0.5500, -0.0500]	[0.8897, 0.9683]	[0.1200, 1.0800]	[-0.0250, 1.0250]
	[-0.8500, -0.3500]	[0.7000, 0.8897]	[0.0400, 1.1600]	[-0.1750, 1.1750]

Table 7.4: Outer and intermediate discretization of the basic variables considered in Table 7.2. Here N represents the number of discretizations of every basic variable. Note that in the case of the intermediate discretization of variables 4, 5 and 6, described by a triangular CDF, an outer discretization was employed.

Copula	n	**Intermediate discr.**		**Outer discretization**	
		$\hat{\mathrm{LP}}_{(\mathscr{F}_n,m)}(F)$	$\hat{\mathrm{UP}}_{(\mathscr{F}_n,m)}(F)$	$\hat{\mathrm{LP}}_{(\mathscr{F}_n,m)}(F)$	$\hat{\mathrm{UP}}_{(\mathscr{F}_n,m)}(F)$
C_1	2^{12}	2.4414e-3	0.125	0	1
	3^{12}	1.2042e-3	0.0877	1.5053e-5	0.2962
	4^{12}	1.0299e-3	0.1779	3.0517e-5	0.1779
	5^{12}	1.9110e-3	0.0466	7.0778e-6	0.1105
C_2	2^{12}	3.4965e-4	0.25	0	1
	3^{12}	2.3480e-5	0.0772	1.8482e-6	0.4
	4^{12}	1.0326e-4	0.1841	4.2080e-5	0.1841
	5^{12}	8.8183e-5	0.0333	2.3079e-6	0.110823
C_3	2^{12}	1.6692e-4	0.3246	0	1
	3^{12}	2.9413e-6	0.0948	4.9725e-8	0.4955
	4^{12}	1.0069e-5	0.2307	3.9306e-6	0.2307
	5^{12}	2.1120e-5	0.0357	3.3363e-8	0.1420

Table 7.5: Estimations of the lower and upper probabilities using discretization of the basic variables, in the numerical example of Section 7.4. Take into consideration that $2^{12} = 4096$, $3^{12} = 531441$, $4^{12} = 16777216$, $5^{12} = 244140625$.

Copula	n	$\hat{\mathrm{LP}}_{(\mathscr{F},P_\Gamma)}(F)$	$\hat{\mathrm{UP}}_{(\mathscr{F},P_\Gamma)}(F)$
C_1	5000	0	0.0414000
	50000	0.0001200	0.0385600
	500000	0.0001920	0.0373080
	5000000	0.0001766	0.0370564
C_2	5000	0.0002000	0.0260000
	50000	0.0001400	0.0247400
	500000	0.0001680	0.0250780
	5000000	0.0002024	0.0249836
C_3	5000	0	0.0278000
	50000	0.0000400	0.0267000
	500000	0.0000460	0.0273320
	5000000	0.0000462	0.0274142

Table 7.6: Estimations of the lower and upper probabilities using simple Monte Carlo simulations, in the numerical example of Section 7.4. Here n represents the number of simulations employed. Probabilistic intervals on the error of these estimators can be calculated by means of equations (7.15) and (7.11).

Chapter 8

Nonspecificity for infinite random sets of indexable type

8.1 Introduction

The Hartley family of measures evaluates what is called nonspecificity, which is a measure about the amount of information required to remove the epistemic uncertainty. In the discrete setting, the Hartley measure (Hartley (1928); Rényi (1970)) is employed, while for subsets of $\mathbb{R}^d$, the Hartley-like measure is used. Using the Hartley-like measure (Klir and Yuan (1995); Ramer and Padet (2001)) it is possible to define measures of nonspecificity for finite random sets defined on $\mathbb{R}^d$. The purpose of this chapter is twofold: to extend the measures of nonspecificity to infinite random sets and to give some new generalizations of concepts that were defined in the realm of finite random sets.

The plan of this chapter is the following. Section 8.2 presents a brief introduction to nonspecificity and Hartley-like measures; then in section 8.3, concepts like joint and marginal infinite random sets are introduced. In section 8.4 we generalize the concept of random set inclusion to infinite random sets and then in section 8.5, the measure of nonspecificity for infinite random sets is proposed. In section 8.6, we analyze the properties of the new measure; finally, in section 8.7, it is shown that the proposed specificity measure is the unique one that satisfies a defining set of properties.

8.2 Nonspecificity

Let us consider a universal set X and its power set $\mathscr{P}(X)$. Suppose that we have to choose an element x^* from a set of elements X. All that we are given is a class of possible choices $\mathscr{F} \subseteq \mathscr{P}(X)$, where at least one $\gamma \in \mathscr{F}$ contains x^*; here some choices γ of $\mathscr{F}$ are more probable than others. Each $\gamma \in \mathscr{F}$ is composed by some elements, however only x^* is the correct one and there is no way to make some

distinction between the elements in γ for almost all $\gamma \in \mathscr{F}$, so that one can choose x^*. The *nonspecificity* is a measure of epistemic uncertainty used in cases like the this, when we have to choose a unique element x^* from X, but we are totally indifferent about which element of the provided γ-s to choose. The nonspecificity is thus associated to the size of the elements $\gamma \in \mathscr{F}$ and may thus be measured by the average amount of information needed to completely remove the epistemic uncertainty from some set γ. The name nonspecificiy follows from the fact that full specificity will be only reached when at the most one element in γ remains for almost all possible choices γ in $\mathscr{F}$.

In contraposition, when all available information is expressed by assigning to each suboption in the universal set X a relative degree of confidence (the so called *probability*) on the assertion that the suboption is the true one, provided that the sum of all probabilities is one, then we are dealing with aleatory uncertainty. Here several alternatives are correct, however, only one of them is the true one. It is then clear that those correct claims conflict with each other. This type of uncertainty is measured for example by Shannon's entropy, which measures the average amount of uncertainty associated with the selection of a set γ from the set of weighted alternatives $\mathscr{F}$, in the case that all γ-s are singletons and $\mathscr{F}$ is finite.

In the following we will briefly discuss nonspecificity for finite sets and for convex subsets of $\mathbb{R}^d$. For details the reader is referred to Klir (2006).

8.2.1 Nonspecificity for finite sets

Consider the following problem: given a finite set E of balls, which contains a black ball, while the rest are white, we would like to measure the amount of information H required to find the black ball. Suppose that the set of balls E has $m \times n$ elements. If this set is partitioned into n sets of m balls or into m sets of n balls, the measure of nonspecificity H characterizing all those sets should satisfy

$$H(m \times n) = H(m) + H(n) \tag{8.1}$$

Also, note that the larger the set of balls E, the less specific the predictions are, and in consequence

$$H(n) \leq H(n+1) \tag{8.2}$$

where n is the cardinality of E, i.e., $n := |E|$.

Hartley (1928) proposed the formula

$$H(n) := \log_2 n$$

and Rényi (1970) showed that this is the unique expression that satisfies equations (8.1) and (8.2) up to the normalization $H(2) = 1$.

This function is known in the literature as the *Hartley measure* of uncertainty, and it measures the lack of specificity of a finite set.

8.2.2 Nonspecificity for convex and bounded subsets of $\mathbb{R}^d$

In the sequel we will use the term *isometric transformation* in $\mathbb{R}^d$ to denote a translation followed by an orthogonal transformation (rotation, reflection).

The *Hartley-like measure* was proposed by Klir and Yuan (1995) to measure nonspecificity for convex and bounded subsets of $\mathbb{R}^d$. It is characterized by the function $\mathrm{HL}:\mathscr{C}\to[0,\infty)$,

$$\mathrm{HL}(A) := \min_{t\in T}\left\{\log_2\left[\prod_{i=1}^{d}\left(1+\mu\left(A_{i_t}\right)\right)+\mu(A)-\prod_{i=1}^{d}\mu\left(A_{i_t}\right)\right]\right\} \tag{8.3}$$

where $\mathscr{C}$ represents the family of all convex and bounded subsets of $\mathbb{R}^d$, μ is the Lebesgue measure, T denotes the set of all isometric transformations and A_{i_t} denotes the i-th projection of A in the coordinate system t. Ramer and Padet (2001) considered also the possibility of extending the HL measure to nonconvex subsets of $\mathbb{R}^d$.

The set function defined in (8.3) satisfies the following set of properties:

Property HL1: Range For every $A\in\mathscr{C}$, $\mathrm{HL}(A)\in[0,\infty)$ where $\mathrm{HL}(A)=0$ if and only if $A=\{x\}$ or $A=\emptyset$ for some $x\in\mathbb{R}^d$.

Property HL2: Monotonicity If $A\subseteq B$, then $\mathrm{HL}(A)\le\mathrm{HL}(B)$ for all $A,B\in\mathscr{C}$. According to Ramer and Padet (2001), HL is strictly monotonic on convex sets, i.e., if $A\subsetneqq B$ then $\mathrm{HL}(A)<\mathrm{HL}(B)$.

Property HL3: Subadditivity For every $A\in\mathscr{C}$ we have that $\mathrm{HL}(A)\le\sum_{i=1}^{n}\mathrm{HL}(A_i)$ where A_i denotes the one-dimensional projection of A to dimension i in some coordinate system.

Property HL4: Additivity For all $A\in\mathscr{C}$ such that $A:=\times_{i=1}^{d}A_i$, then $\mathrm{HL}(A)=\sum_{i=1}^{d}\mathrm{HL}(A_i)$. A_i here has the same meaning as in Property HL3.

Property HL5: Coordinate invariance HL does not depend on isometric transformations of the coordinate system.

Property HL6: Continuity HL is a continuous set function. This criterion has not been explicitly defined by Klir and Yuan (1995), but according to the interpretation of Ramer and Padet (2001), it is intended with respect to the Hausdorff metric.

Property HL7: Normalization When $A=\left\{\times_{i=1}^{d}[a_i,a_i+1]\right\}$ for any $[a_1,a_2,\ldots,a_d]\in\mathbb{R}^d$, then $\mathrm{HL}(A)=d$, for any $d\in\mathbb{N}$.

Conditional Hartley-like measures

The conditional Hartley-like measures express the relationship between marginal and joint Hartley-like measures. Given a set $R\subseteq X\times Y$, whose projections on the sets X and Y are given by $R^X:=\mathrm{proj}_X(R)$ and $R^Y:=\mathrm{proj}_Y(R)$ where

the *projection operators* $\text{proj}_X : X \times Y \to X$ and $\text{proj}_Y : X \times Y \to Y$ stand for,

$$\text{proj}_X(\gamma) := \{x \in X : \exists y \in Y, (x,y) \in \gamma\}$$
$$\text{proj}_Y(\gamma) := \{y \in Y : \exists x \in X, (x,y) \in \gamma\}$$

given $\gamma \subseteq X \times Y$. These measures are defined by,

$$\text{HL}(R|R^Y) := \text{HL}(R) - \text{HL}(R^Y) \tag{8.4}$$
$$\text{HL}(R|R^X) := \text{HL}(R) - \text{HL}(R^X) \tag{8.5}$$

8.3 Joint and marginal infinite random sets

The definitions of joint and marginal infinite random sets follow directly from the analogous definitions for Dempster-Shafer bodies of evidence (see e.g. Klir (2006)). In this section we will consider two marginal infinite random sets. A generalization to more dimensions follows the same considerations.

Given the joint infinite random set $(\mathscr{F}^{XY}, P_\Gamma^{XY})$ on $X \times Y$, the associated *marginal random sets* $(\mathscr{F}^X, P_\Gamma^X)$ and $(\mathscr{F}^Y, P_\Gamma^Y)$, on X and Y respectively, are defined by the focal sets,

$$\mathscr{F}^X := \{\text{proj}_X(\gamma) : \gamma \in \mathscr{F}^{XY}\}$$
$$\mathscr{F}^Y := \{\text{proj}_Y(\gamma) : \gamma \in \mathscr{F}^{XY}\}$$

and the probability measures $P_\Gamma^X : \sigma_{\mathscr{F}^X} \to [0,1]$, and $P_\Gamma^Y : \sigma_{\mathscr{F}^Y} \to [0,1]$,

$$\begin{aligned} P_\Gamma^X(\mathscr{G}^X) &:= \int_{\mathscr{F}^{XY}} I[\text{proj}_X(\gamma) \in \mathscr{G}^X] \, \mathrm{d}P_\Gamma^{XY}(\gamma) \\ &= P_\Gamma^{XY}\{\gamma : \text{proj}_X(\gamma) \in \mathscr{G}^X, \gamma \in \mathscr{F}^{XY}\} \end{aligned} \tag{8.6}$$
$$\begin{aligned} P_\Gamma^Y(\mathscr{G}^Y) &:= \int_{\mathscr{F}^{XY}} I[\text{proj}_Y(\gamma) \in \mathscr{G}^Y] \, \mathrm{d}P_\Gamma^{XY}(\gamma) \\ &= P_\Gamma^{XY}\{\gamma : \text{proj}_Y(\gamma) \in \mathscr{G}^Y, \gamma \in \mathscr{F}^{XY}\} \end{aligned} \tag{8.7}$$

for all $\mathscr{G}^X \in \sigma_{\mathscr{F}^X}$ and all $\mathscr{G}^Y \in \sigma_{\mathscr{F}^Y}$ and where $\sigma_{\mathscr{F}^X}$, $\sigma_{\mathscr{F}^Y}$ are σ-algebras on $\mathscr{F}^X$ and $\mathscr{F}^Y$ respectively. Note that in the finite case $|\mathscr{F}^X| \leq |\mathscr{F}^{XY}|$ and $|\mathscr{F}^Y| \leq |\mathscr{F}^{XY}|$, and that (8.6) and (8.7) are a generalization of the definitions of marginal random sets for finite RSs, namely

$$m^X(A) = \sum_{A' \in \mathscr{F}^{XY}} I[\text{proj}_X(A') = A] m^{XY}(A')$$
$$m^Y(B) = \sum_{A' \in \mathscr{F}^{XY}} I[\text{proj}_Y(A') = B] m^{XY}(A')$$

for all $A \in \mathscr{F}^X$ and all $B \in \mathscr{F}^Y$.

When $(\mathscr{F}^{XY}, P_\Gamma^{XY})$, $(\mathscr{F}^X, P_\Gamma^X)$ and $(\mathscr{F}^Y, P_\Gamma^Y)$ are of indexable type, we can associate correspondingly the unique copulas C^{XY}, C^X and C^Y, to the probability

measures P_Γ^{XY}, P_Γ^X and P_Γ^Y. We will define the conditional copulas $C^{X|Y}$ and $C^{Y|X}$ respectively by:

$$C^{X|Y}\left(\boldsymbol{\alpha}|\boldsymbol{\beta}\right) = \begin{cases} \frac{C^{XY}(\boldsymbol{\alpha},\boldsymbol{\beta})}{C^{Y}(\boldsymbol{\beta})} & \text{for } C^{Y}\left(\boldsymbol{\beta}\right) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$C^{Y|X}\left(\boldsymbol{\beta}|\boldsymbol{\alpha}\right) = \begin{cases} \frac{C^{XY}(\boldsymbol{\alpha},\boldsymbol{\beta})}{C^{X}(\boldsymbol{\alpha})} & \text{for } C^{X}\left(\boldsymbol{\alpha}\right) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

It is important to observe the following relationships between the lower and upper probabilities of $\left(\mathscr{F}^X, P_\Gamma^X\right)$, $\left(\mathscr{F}^Y, P_\Gamma^Y\right)$ and $\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)$,

$$\mathrm{LP}_{\left(\mathscr{F}^X, P_\Gamma^X\right)}(A) = \mathrm{LP}_{\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)}(A \times Y) \tag{8.8}$$

$$\mathrm{LP}_{\left(\mathscr{F}^Y, P_\Gamma^Y\right)}(B) = \mathrm{LP}_{\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)}(X \times B) \tag{8.9}$$

$$\mathrm{UP}_{\left(\mathscr{F}^X, P_\Gamma^X\right)}(A) = \mathrm{UP}_{\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)}(A \times Y)$$

$$\mathrm{UP}_{\left(\mathscr{F}^Y, P_\Gamma^Y\right)}(B) = \mathrm{UP}_{\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)}(X \times B) \tag{8.10}$$

For example, (8.8) can be shown as follows:

$$\begin{aligned} &\mathrm{LP}_{\left(\mathscr{F}^X, P_\Gamma^X\right)}(A) = \\ &\quad = P_\Gamma^X\left\{\gamma^X : \gamma^X \subseteq A, \gamma^X \in \mathscr{F}^X, \gamma^X \neq \emptyset\right\} \\ &\quad = P_\Gamma^{XY}\left\{\gamma : \mathrm{proj}_X(\gamma) \in \left\{\gamma^X : \gamma^X \subseteq A, \gamma^X \in \mathscr{F}^X, \gamma^X \neq \emptyset\right\}, \gamma \in \mathscr{F}^{XY}\right\} \\ &\quad = P_\Gamma^{XY}\left\{\gamma : \gamma \subseteq A \times Y, \gamma \in \mathscr{F}^{XY}, \gamma \neq \emptyset\right\} \\ &\quad = \mathrm{LP}_{\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)}(A \times Y) \end{aligned}$$

The deduction of equations (8.9) to (8.10) follows the same considerations.

Non-interactivity

There is a special case of relationship between marginal random sets called *non-interactivity*. The probability measures corresponding to the marginal infinite random sets $\left(\mathscr{F}^X, P_\Gamma^X\right)$ and $\left(\mathscr{F}^Y, P_\Gamma^Y\right)$ are called *non-interactive* if and only if for all $\mathscr{G}^X \in \sigma_{\mathscr{F}^X}$, $\mathscr{G}^Y \in \sigma_{\mathscr{F}^Y}$ and $\gamma^{XY} \in \mathscr{F}^{XY}$ we have that $P_\Gamma^{XY}\left(\mathscr{G}^{XY}\right) = P_\Gamma^X\left(\mathscr{G}^X\right) P_\Gamma^Y\left(\mathscr{G}^Y\right)$ if $\mathscr{G}^{XY} = \mathscr{G}^X \times \mathscr{G}^Y$, otherwise $P_\Gamma^{XY}\left(\mathscr{G}^{XY}\right) = 0$. In section 6.1.1 it was shown that this concept is equivalent to the extension of the concept of *random set independence* (see e.g. Couso et al. (1999); Fetz and Oberguggenberger (2004)) to infinite random sets[1]. In the case of random sets of indexable type, P_Γ^{XY}, P_Γ^X and P_Γ^Y will have the associated copulas C^{XY}, C^X and C^Y correspondingly; in this case $C^{XY} = C^X C^Y$.

[1] There are two other types of independence for infinite random sets, namely *strong independence* and *fuzzy set independence*; we will not deal inhere with those cases since we will allow *all* possible probability measures residing in the marginal focal elements and also we will allow *arbitrary dependence* for the combination of those marginal probability measures. When these matters are taken into consideration, strong independence plays a main role. Also, we will not deal with fuzzy set independence since we will not restrict our analysis to consonant focal sets.

8.4 Random set inclusion for infinite random sets of indexable type

In this section, the concept of finite random set inclusion proposed by Yager (1986); Delgado and Moral (1987); Dubois and Prade (1991) (see Definition 2.2.2) is generalized to infinite random sets of indexable type.

Definition 8.4.1. *Let* $\left(\mathscr{F}^A, P_\Gamma^A\right)$ *and* $\left(\mathscr{F}^B, P_\Gamma^B\right)$ *be two infinite RSs of indexable type defined on* $X \subseteq \mathbb{R}^d$. *Then* $\left(\mathscr{F}^A, P_\Gamma^A\right)$ *is said to be* included *in* $\left(\mathscr{F}^B, P_\Gamma^B\right)$, *denoted as* $\left(\mathscr{F}^A, P_\Gamma^A\right) \subseteq \left(\mathscr{F}^B, P_\Gamma^B\right)$ *if and only if the following three conditions hold:*

1. *For all* $\gamma^A \in \mathscr{F}^A$ *there exists a set* $\gamma^B \in \mathscr{F}^B$ *such that* $\gamma^A \subseteq \gamma^B$.
2. *For all* $\gamma^B \in \mathscr{F}^B$ *there exists a set* $\gamma^A \in \mathscr{F}^A$ *such that* $\gamma^A \subseteq \gamma^B$.
3. *In the space* $(0,1]^d \times (0,1]^d$ *corresponding to the* α*-representation of* $\left(\mathscr{F}^A, P_\Gamma^A\right) \times \left(\mathscr{F}^B, P_\Gamma^B\right)$ *respectively, there exists a copula* $F : [0,1]^d \times [0,1]^d \to [0,1], \left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right) \mapsto F\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)$ *whose support is the set*

$$\left\{\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right) : \Gamma^A\left(\boldsymbol{\alpha}\right) \subseteq \Gamma^B\left(\boldsymbol{\beta}\right), \boldsymbol{\alpha}, \boldsymbol{\beta} \in [0,1]^d, \Gamma^A\left(\boldsymbol{\alpha}\right) \in \mathscr{F}^A, \Gamma^B\left(\boldsymbol{\alpha}\right) \in \mathscr{F}^B\right\}$$

such that:

$$P_\Gamma^A\left(\Gamma^A\left(G^A\right)\right) = \int_{G^A} \int_{(0,1]^d} \mathrm{d}F\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right) \tag{8.11}$$

for each $G^A \in [0,1]^d \cap \mathscr{B}^d$, $\Gamma^A\left(G^A\right) \in \sigma_{\mathscr{F}^A}$ *and*

$$P_\Gamma^B\left(\Gamma^B\left(G^B\right)\right) = \int_{(0,1]^d} \int_{G^B} \mathrm{d}F\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right) \tag{8.12}$$

for each $G^B \in [0,1]^d \cap \mathscr{B}^d$, $\Gamma^B\left(G^B\right) \in \sigma_{\mathscr{F}^B}$.

Observe that the copula F has the form,

$$F\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right) := C^{AB}\left(C^A\left(\boldsymbol{\alpha}\right), C^B\left(\boldsymbol{\beta}\right)\right) \tag{8.13}$$

where $C^A : [0,1]^d \to [0,1]$ and $C^B : [0,1]^d \to [0,1]$ are respectively the copulas associated to the RSs $\left(\mathscr{F}^A, P_\Gamma^A\right)$ and $\left(\mathscr{F}^B, P_\Gamma^B\right)$ by $P_\Gamma^A \equiv \mu_{C^A}$ and $P_\Gamma^B \equiv \mu_{C^B}$, and C^{AB} is a two-dimensional copula. Since $F\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right) = C^{AB}\left(C^A\left(\boldsymbol{\alpha}\right), C^B\left(\boldsymbol{\beta}\right)\right) = C^{A|B}\left(C^A\left(\boldsymbol{\alpha}\right) | C^B\left(\boldsymbol{\beta}\right)\right) C^B\left(\boldsymbol{\beta}\right) = C^{B|A}\left(C^B\left(\boldsymbol{\beta}\right) | C^A\left(\boldsymbol{\alpha}\right)\right) C^A\left(\boldsymbol{\alpha}\right)$, where $C^{A|B}$ and $C^{B|A}$ are the conditional copulas associated to C^{AB}, using the relations between joint, conditional and marginal CDFs, we retrieve the next basic relations from integral (8.11) (and similarly for (8.12)),

$$\begin{aligned} P_\Gamma^A\left(\Gamma^A\left(G^A\right)\right) &= \int_{G^A} \int_{(0,1]^d} \mathrm{d}C^{AB}\left(C^A\left(\boldsymbol{\alpha}\right), C^B\left(\boldsymbol{\beta}\right)\right) \\ &= \int_{G^A} \int_{(0,1]^d} \mathrm{d}C^{B|A}\left(C^B\left(\boldsymbol{\beta}\right) | C^A\left(\boldsymbol{\alpha}\right)\right) \mathrm{d}C^A\left(\boldsymbol{\alpha}\right) \\ &= \int_{G^A} \mathrm{d}C^A\left(\boldsymbol{\alpha}\right) \\ &= \mu_{C^A}\left(G^A\right) \end{aligned}$$

and $P_\Gamma^B\left(\Gamma^B\left(G^B\right)\right) = \mu_{C^B}\left(G^B\right)$ for $G^A, G^B \in [0,1]^d \cap \mathscr{B}^d$.

Examples of random set inclusion

Possibility distributions Let A and B be two possibility distributions such that $A(x) \leq B(x)$ for all $x \in X$. In this case, A is said to be a *normalized fuzzy subset* of B (see e.g. (Klir, 2006, p. 262)). If A and B are associated to the random sets $\left(\mathscr{F}^A, P_\Gamma^A\right)$ and $\left(\mathscr{F}^B, P_\Gamma^B\right)$, then $\left(\mathscr{F}^A, P_\Gamma^A\right) \subseteq \left(\mathscr{F}^B, P_\Gamma^B\right)$, or shortly $A \subseteq B$, because the three requisites for random set inclusion are satisfied, namely,

1. 2. if $\alpha^A \in (0,1]$ is the index associated to the α-cuts $A_{\alpha^A} \in \mathscr{F}^A$ and $\alpha^B \in (0,1]$ is the index associated to $B_{\alpha^B} \in \mathscr{F}^B$ then for all $\alpha^A = \alpha^B = \alpha$, such that $\alpha \in (0,1]$, we have that $A_\alpha \subseteq B_\alpha$.

3. Since $d = 1$, then the copulas C^A and C^B in (8.13) will correspond to uniform CDFs on $[0,1]$ and therefore F will be any copula with support $\text{supp}(F) := \left\{\left(\alpha^A, \alpha^B\right) : A_{\alpha^A} \subseteq B_{\alpha^B}, \alpha^A, \alpha^B \in [0,1]\right\}$.

Consider two possibility distributions that have a trapezoidal shape, $A := \text{trap}(3,4,6,9)$ and $B := \text{trap}(2,3,7,9)$. Clearly $A \subseteq B$. Figure 8.1 shows those possibility distributions and its associated set supp(F)

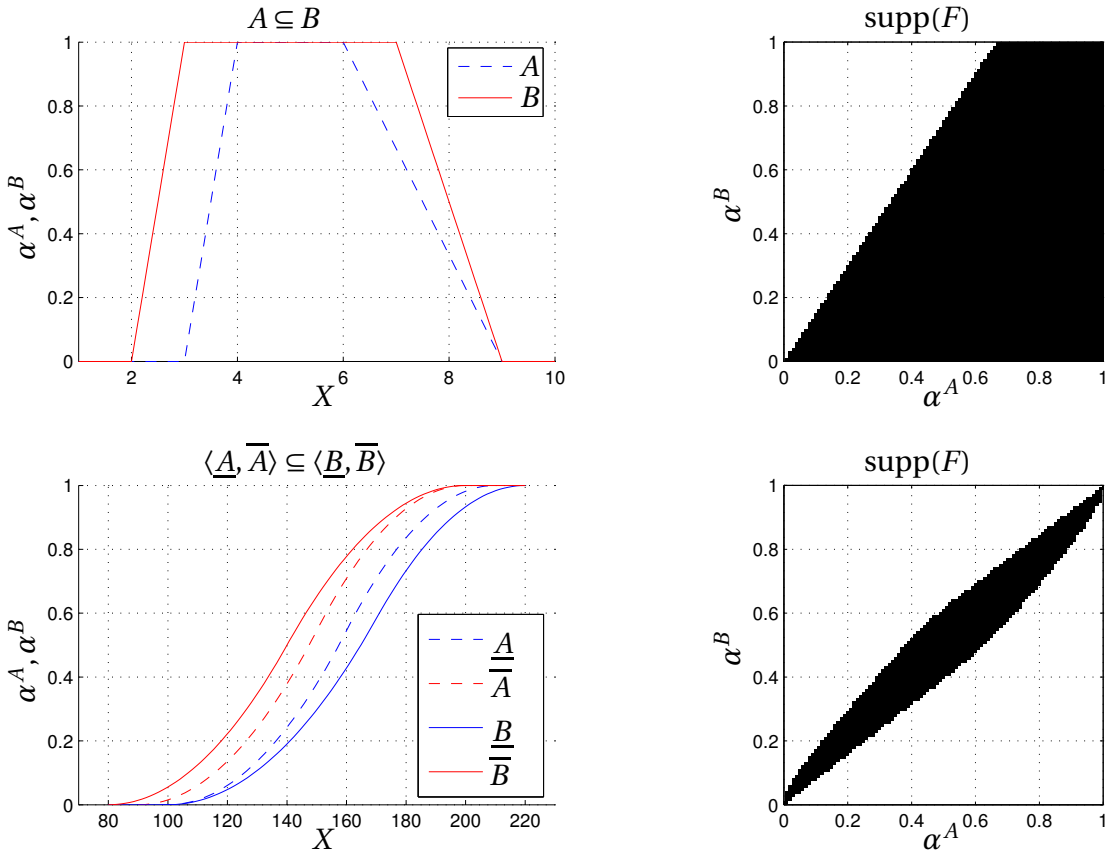


Figure 8.1: Upper row: Left: possibility distributions $A := \text{trap}(3,4,6,9)$ and $B := \text{trap}(2,3,7,9)$, here $A \subseteq B$. Right: support of the copula F. Lower row: Left: probability boxes $\langle \underline{A}, \overline{A} \rangle := \langle T(90,150,200), T(100,160,210) \rangle$ and $\langle \underline{B}, \overline{B} \rangle := \langle T(80,140,200), T(100,170,220) \rangle$, here $\langle \underline{A}, \overline{A} \rangle \subseteq \langle \underline{B}, \overline{B} \rangle$. Right: support of the copula F.

Probability boxes Let $\langle \underline{A}, \overline{A} \rangle$ and $\langle \underline{B}, \overline{B} \rangle$ two probability boxes with infinite random set representation $\left(\mathscr{F}^A, P_\Gamma^A\right)$ and $\left(\mathscr{F}^B, P_\Gamma^B\right)$, such that $\underline{B} \leq \underline{A} \leq \overline{A} \leq \overline{B}$, that is, *the probability box* $\langle \underline{A}, \overline{A} \rangle$ is *contained in the probability box* $\langle \underline{B}, \overline{B} \rangle$. We will say in this case that the probability box $\langle \underline{A}, \overline{A} \rangle$ is a subset of the probability box $\langle \underline{B}, \overline{B} \rangle$, i.e., $\langle \underline{A}, \overline{A} \rangle \subseteq \langle \underline{B}, \overline{B} \rangle$, based on the fact that $\left(\mathscr{F}^A, P_\Gamma^A\right) \subseteq \left(\mathscr{F}^B, P_\Gamma^B\right)$; the demonstration is analogous to the one of possibility distributions, with A_α replaced by $\Gamma^A(\alpha) := \left[\overline{A}^{(-1)}(\alpha), \underline{A}^{(-1)}(\alpha)\right]$ and B_α replaced by $\Gamma^B(\alpha) := \left[\overline{B}^{(-1)}(\alpha), \underline{B}^{(-1)}(\alpha)\right]$.

Figure 8.1 shows two probability boxes $\langle \underline{A}, \overline{A} \rangle := \langle T(90,150,200), T(100,160,210) \rangle$ and $\langle \underline{B}, \overline{B} \rangle := \langle T(80,140,200), T(100,170,220) \rangle$; here $T(a,b,c)$ stands for the formulation of a triangular CDF corresponding to the triangular PDF $t(a,b,c)$ with lower limit a, mode b and upper limit c. It can be shown that $\langle \underline{A}, \overline{A} \rangle \subseteq \langle \underline{B}, \overline{B} \rangle$. Figure 8.1 shows those probability boxes along with its associated set supp(F).

Dempster-Shafer structures In this case, when the RSs are finite, the concept of random set inclusion reduces to the one proposed by Yager (1986); Delgado and Moral (1987); Dubois and Prade (1991):, and already explained in Section 2.2.6. In this case, equations (2.12) and (2.13) follow directly from (8.11) and (8.12) by observing that a finite RS induces a partition of the α-space. Every entry $W\left(A_i, B_j\right)$ will correspond to the Lebesgue-Stieltjes measure of the box in the α-space associated to the focal element $A_i \times B_j$ of the RS $\left(\mathscr{F}^A, m^A\right) \times \left(\mathscr{F}^B, m^B\right)$ with respect to the related copula F.

8.5 Measures of nonspecificity for infinite random sets

In this section a measure of nonspecificity for infinite random sets is proposed, which coincides with the definitions of nonspecificity when particularized to finite random sets. Let $(\mathscr{F}, P_\Gamma)$ be an infinite random set defined on $X \subseteq \mathbb{R}^d$, the nonspecificity for $(\mathscr{F}, P_\Gamma)$ will be defined by

$$\begin{aligned} \mathrm{NL}((\mathscr{F}, P_\Gamma)) &:= \mathrm{E}_{(\mathscr{F}, P_\Gamma)}\left[\mathrm{HL}(\gamma)\right] \\ &= \int_{\mathscr{F}} \mathrm{HL}(\gamma)\, \mathrm{d}P_\Gamma(\gamma) \end{aligned} \tag{8.14}$$

where $\mathrm{HL}(\gamma)$ stands for the Hartley-like measure of the set $\gamma \in \mathscr{F}$. That is, the NL measure is the expectation of the HL measure of the focal elements within the focal set $\mathscr{F}$. In the following we will assume that all focal elements are convex and compact; if they are not, we will employ their closures. If $(\mathscr{F}, P_\Gamma)$ is of indexable type, then using the α-representation of infinite random sets, integral (8.14) is equivalent to

$$\mathrm{NL}((\mathscr{F}, P_\Gamma)) := \int_{(0,1]^d} \mathrm{HL}(\Gamma(\boldsymbol{\alpha}))\, \mathrm{d}C(\boldsymbol{\alpha}) \tag{8.15}$$

If it happens that $d = 1$, then

$$\mathrm{NL}\left((\mathscr{F}, P_\Gamma)\right) = \int_0^1 \mathrm{HL}\left(\Gamma(\alpha)\right) \mathrm{d}\alpha \tag{8.16}$$

Examples

Equation (8.15) can be particularized to the special cases when $(\mathscr{F}, P_\Gamma)$ represents possibility distributions, probability boxes or Dempster-Shafer structures, as outlined below.

Possibility distributions If $(\mathscr{F}, P_\Gamma)$ represents the unidimensional possibility distribution A, equation (8.16) turns into

$$\begin{aligned}\mathrm{NL}(A) &= \int_0^1 \mathrm{HL}(A_\alpha)\,\mathrm{d}\alpha \\ &= \int_0^1 \log_2\left(1+\mu(A_\alpha)\right)\mathrm{d}\alpha\end{aligned}$$

In the generalized information theory literature this measure is known as the UL-uncertainty (see for example (Klir, 2006, p. 206)).

Probability boxes If $(\mathscr{F}, P_\Gamma)$ is an infinite RS defined on $\mathbb{R}$ which represents the probability box $\langle \underline{F}, \overline{F} \rangle$, equation (8.16) turns into

$$\begin{aligned}\mathrm{NL}\left(\langle \underline{F}, \overline{F}\rangle\right) &= \int_0^1 \mathrm{HL}\left(\left[\overline{F}^{(-1)}(\alpha), \underline{F}^{(-1)}(\alpha)\right]\right)\mathrm{d}\alpha \\ &= \int_0^1 \log_2\left(1+\underline{F}^{(-1)}(\alpha) - \overline{F}^{(-1)}(\alpha)\right)\mathrm{d}\alpha\end{aligned} \tag{8.17}$$

Note that using (8.17), the nonspecificity of a CDF is 0. This confirms the fact that probability theory cannot measure nonspecificity, inasmuch as it requires perfect knowledge in the specification of the CDFs.

Dempster-Shafer structures If $(\mathscr{F}, P_\Gamma)$ represents the Dempster-Shafer structure $(\mathscr{F}_n, m)$ integral (8.14) turns into

$$\mathrm{NL}\left((\mathscr{F}_n, m)\right) = \sum_{i=1}^{n} \mathrm{HL}(A_i)\, m(A_i)$$

which is already recognized within generalized information theory (see e.g. (Klir and Wierman, 1998, p. 67)). See also Dubois and Prade (1985, 1987) who deal with specificity for belief functions.

8.6 Properties fulfilled by the NL measure

According to (Klir, 2006, p. 196) every measure of uncertainty is required to fulfill a set of defining properties. In the rest of this chapter it will be shown that (8.15) satisfies the following ones:

Property NL1: Range For each random set $(\mathscr{F}, P_\Gamma)$ of indexable type, whose focal elements are bounded and convex subsets of $\mathbb{R}^d$, we have that $\mathrm{NL}((\mathscr{F}, P_\Gamma))$, is nonnegative and finite; in addition $\mathrm{NL}((\mathscr{F}, P_\Gamma)) = 0$ if and only if almost all elements of $\mathscr{F}$ are singletons.

Property NL2: Monotonicity For any pair of infinite random sets of indexable type $\left(\mathscr{F}^A, P_\Gamma^A\right)$ and $\left(\mathscr{F}^B, P_\Gamma^B\right)$ such that $\left(\mathscr{F}^A, P_\Gamma^A\right) \subseteq \left(\mathscr{F}^B, P_\Gamma^B\right)$, it holds that $\mathrm{NL}\left(\left(\mathscr{F}^A, P_\Gamma^A\right)\right) \leq \mathrm{NL}\left(\left(\mathscr{F}^B, P_\Gamma^B\right)\right)$.

Property NL3: Subadditivity For any joint infinite RS $\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)$ of indexable type defined on $X \times Y$ and its associated marginals $\left(\mathscr{F}^X, P_\Gamma^X\right)$ and $\left(\mathscr{F}^Y, P_\Gamma^Y\right)$ defined on X and Y respectively, we have that,

$$\mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)\right) \leq \mathrm{NL}\left(\left(\mathscr{F}^X, P_\Gamma^X\right)\right) + \mathrm{NL}\left(\left(\mathscr{F}^Y, P_\Gamma^Y\right)\right)$$

Property NL4: Additivity For any non-interactive (independent in the sense of RS independence) infinite random sets $\left(\mathscr{F}^X, P_\Gamma^X\right)$ and $\left(\mathscr{F}^Y, P_\Gamma^Y\right)$ defined on X and Y respectively, and the associated joint infinite RS $\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)$ defined on $X \times Y$, where $\mathscr{F}^{XY} = \mathscr{F}^X \times \mathscr{F}^Y$ and $P_\Gamma^{XY}(A \times B) = P_\Gamma^X(A)\, P_\Gamma^Y(B)$ for all $A \in \sigma_{\mathscr{F}^X}$ and all $B \in \sigma_{\mathscr{F}^Y}$, we have that,

$$\mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)\right) = \mathrm{NL}\left(\left(\mathscr{F}^X, P_\Gamma^X\right)\right) + \mathrm{NL}\left(\left(\mathscr{F}^Y, P_\Gamma^Y\right)\right)$$

Property NL5: Coordinate invariance The functional NL does not change under isometric transformations of the focal elements.

Property NL7: Normalization When $\mathscr{F} = \left\{\times_{i=1}^d [a_i, a_i + 1]\right\}$ for any $[a_1, a_2, \ldots, a_d] \in \mathbb{R}^d$, then $\mathrm{NL}((\mathscr{F}, P_\Gamma)) = d$, for any $d \in \mathbb{N}$.

Property NL8: Branching Suppose we have the infinite random set of indexable type $(\mathscr{F}, P_\Gamma)$, and any partition of $\mathscr{F}$ into n sets $\mathscr{F}^i$; then to each of those sets $\mathscr{F}^i$ we will assign a probability measure $P_\Gamma\left(\mathscr{G}|\mathscr{F}^i\right) := P_\Gamma\left(\mathscr{G} \cap \mathscr{F}^i\right) / P_\Gamma\left(\mathscr{F}^i\right)$, for $\mathscr{G} \in \sigma_{\mathscr{F}^i}$, so that $(\mathscr{F}, P_\Gamma)$ is decomposed into the random sets $\left(\mathscr{F}^i, P_\Gamma\left(\cdot|\mathscr{F}^i\right)\right)$. The NL-uncertainty of $(\mathscr{F}, P_\Gamma)$ can be computed by individually calculating the uncertainty of each partition $\left(\mathscr{F}^i, P_\Gamma\left(\cdot|\mathscr{F}^i\right)\right)$ and then making a weighted addition of those uncertainties according to the importance that $\mathscr{F}^i$ has in the representation of $\mathscr{F}$, i.e.,

$$\mathrm{NL}((\mathscr{F}, P_\Gamma)) = \sum_{i=1}^{n} \mathrm{NL}\left(\left(\mathscr{F}^i, P_\Gamma\left(\cdot|\mathscr{F}^i\right)\right)\right) P_\Gamma\left(\mathscr{F}^i\right)$$

Note: If $(\mathscr{F}, P_\Gamma)$ contains a unique bounded and convex focal element, i.e., $\mathscr{F} = \{A\}$ and $P_\Gamma(A) = 1$ then Properties NLi reduce to Properties HLi, for $i = 1, \ldots, 7$.

In this and in the following section, we will show that the NL measure in fact satisfies those properties.

Finally, there is an additional property namely

Property NL6: Continuity NL is a continuous function.

which is expected to be fulfilled. However, its proof is left as an open problem.

8.6.1 Monotonicity

Random set inclusion implies monotonicity in the Hartley-like measure, i.e., if $\left(\mathscr{F}^1, P_\Gamma^1\right) \subseteq \left(\mathscr{F}^2, P_\Gamma^2\right)$ then $\mathrm{NL}\left(\left(\mathscr{F}^1, P_\Gamma^1\right)\right) \leq \mathrm{NL}\left(\left(\mathscr{F}^2, P_\Gamma^2\right)\right)$. The following theorem is an extension of a theorem proposed by Klir (2006, p. 213) for the case of finite random sets.

Theorem 8.6.1. *For any pair of infinite random sets of indexable type* $\left(\mathscr{F}^A, P_\Gamma^A\right)$ *and* $\left(\mathscr{F}^B, P_\Gamma^B\right)$ *such that* $\left(\mathscr{F}^A, P_\Gamma^A\right) \subseteq \left(\mathscr{F}^B, P_\Gamma^B\right)$*, it holds that* $\mathrm{NL}\left(\left(\mathscr{F}^A, P_\Gamma^A\right)\right) \leq \mathrm{NL}\left(\left(\mathscr{F}^B, P_\Gamma^B\right)\right)$.

Proof. Let C^A and C^B be the copulas associated to P_Γ^A and P_Γ^B respectively. Then,

$$
\begin{aligned}
&\int_{\mathscr{F}^A} \mathrm{HL}\left(\gamma^A\right) \mathrm{d}P_\Gamma^A\left(\gamma^A\right) = \\
&\quad = \int_{(0,1]^d} \mathrm{HL}\left(\Gamma^A(\boldsymbol{\alpha})\right) \mathrm{d}C^A(\boldsymbol{\alpha}) \\
&\quad = \int_{(0,1]^d} \int_{(0,1]^d} \mathrm{HL}\left(\Gamma^A(\boldsymbol{\alpha})\right) \mathrm{d}C^{B|A}\left(C^B(\boldsymbol{\beta}) | C^A(\boldsymbol{\alpha})\right) \mathrm{d}C^A(\boldsymbol{\alpha}) \\
&\quad = \int_{(0,1]^d} \int_{(0,1]^d} \mathrm{HL}\left(\Gamma^A(\boldsymbol{\alpha})\right) \mathrm{d}F(\boldsymbol{\alpha}, \boldsymbol{\beta}) \\
&\quad \leq \int_{(0,1]^d} \int_{(0,1]^d} \mathrm{HL}\left(\Gamma^B(\boldsymbol{\beta})\right) \mathrm{d}C^{A|B}\left(C^A(\boldsymbol{\alpha}) | C^B(\boldsymbol{\beta})\right) \mathrm{d}C^B(\boldsymbol{\beta}) \qquad (8.18) \\
&\quad = \int_{(0,1]^d} \mathrm{HL}\left(\Gamma^B(\boldsymbol{\beta})\right) \mathrm{d}C^B(\boldsymbol{\beta})
\end{aligned}
$$

The inequality (8.18) follows from Property HL2 and the fact that the point $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ belongs to the support of the copula F and in consequence, $\Gamma^A(\boldsymbol{\alpha}) \subseteq \Gamma^B(\boldsymbol{\beta})$. □

8.6.2 Subadditivity and additivity: relationship between the joint and the marginal generalized Hartley-like measures

The following theorem shows that the NL measure of a joint infinite random set of indexable type $\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)$ is subadditive with regard to the NL measure of

its corresponding marginal infinite random sets.

Theorem 8.6.2. *Let* $\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)$ *be a joint infinite random set of indexable type defined on* $X \times Y$, $X \subseteq \mathbb{R}^a$, $Y \subseteq \mathbb{R}^b$ *and its associated marginal infinite random sets* $\left(\mathscr{F}^X, P_\Gamma^X\right)$ *and* $\left(\mathscr{F}^Y, P_\Gamma^Y\right)$ *defined on* X *and* Y *respectively. Then:*

$$\mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)\right) \leq \mathrm{NL}\left(\left(\mathscr{F}^X, P_\Gamma^X\right)\right) + \mathrm{NL}\left(\left(\mathscr{F}^Y, P_\Gamma^Y\right)\right)$$

Proof. NOTE: sometimes we will write $C^{XY}\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)$ as $C^{XY}\left(\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)\right)$ to give a representation of C^{XY} as a multivariate CDF.

Let C^{XY}, C^X and C^Y be the copulas associated to the probability measures P_Γ^{XY}, P_Γ^X and P_Γ^Y respectively. Then

$$\begin{aligned}
\mathrm{NL}\left(\left(\mathscr{F}^X, P_\Gamma^X\right)\right) &= \int_{(0,1]^a} \mathrm{HL}\left(\Gamma^X\left(\boldsymbol{\alpha}\right)\right) \mathrm{d}C^X\left(\boldsymbol{\alpha}\right) \\
&= \int_{(0,1]^a} \int_{(0,1]^b} \mathrm{HL}\left(\Gamma^X\left(\boldsymbol{\alpha}\right)\right) \mathrm{d}C^{Y|X}\left(\boldsymbol{\beta}|\boldsymbol{\alpha}\right) \mathrm{d}C^X\left(\boldsymbol{\alpha}\right) \\
&= \int_{(0,1]^{a+b}} \mathrm{HL}\left(\mathrm{proj}_X \Gamma\left(\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)\right)\right) \mathrm{d}C^{XY}\left(\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)\right)
\end{aligned}$$

using the fact that $\Gamma^X\left(\boldsymbol{\alpha}\right) = \mathrm{proj}_X \Gamma\left(\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)\right)$. Similarly,

$$\mathrm{NL}\left(\left(\mathscr{F}^Y, P_\Gamma^Y\right)\right) = \int_{(0,1]^{a+b}} \mathrm{HL}\left(\mathrm{proj}_Y \Gamma\left(\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)\right)\right) \mathrm{d}C^{XY}\left(\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)\right) \tag{8.19}$$

Hence,

$$\begin{aligned}
&\mathrm{NL}\left(\left(\mathscr{F}^X, P_\Gamma^X\right)\right) + \mathrm{NL}\left(\left(\mathscr{F}^Y, P_\Gamma^Y\right)\right) = \\
&\quad = \int_{(0,1]^{a+b}} \mathrm{HL}\left(\mathrm{proj}_X \Gamma\left(\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)\right)\right) + \mathrm{HL}\left(\mathrm{proj}_Y \Gamma\left(\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)\right)\right) \mathrm{d}C^{XY}\left(\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)\right) \\
&\quad \geq \int_{(0,1]^{a+b}} \mathrm{HL}\left(\Gamma\left(\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)\right)\right) \mathrm{d}C^{XY}\left(\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right)\right) \\
&\quad = \mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)\right);
\end{aligned}$$

here, the last inequality follows after using the subadditivity property HL3 of the HL measure (see Section 8.2.2). □

Theorem 8.6.3. *Given the marginal infinite random sets of indexable type* $\left(\mathscr{F}^X, P_\Gamma^X\right)$ *and* $\left(\mathscr{F}^Y, P_\Gamma^Y\right)$ *which are combined under the condition of random set independence to form the joint infinite random set* $\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)$ *on* $X \times Y$, *then*

$$\mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)\right) = \mathrm{NL}\left(\left(\mathscr{F}^X, P_\Gamma^X\right)\right) + \mathrm{NL}\left(\left(\mathscr{F}^Y, P_\Gamma^Y\right)\right)$$

Proof. Since $\gamma = \gamma^X \times \gamma^Y$ and since P_Γ^X and P_Γ^Y are non-interactive, then $C^{XY}\left((\boldsymbol{\alpha},\boldsymbol{\beta})\right) = C^X(\boldsymbol{\alpha})\, C^Y(\boldsymbol{\beta})$ and using the additivity property HL4 it follows:

$$\begin{aligned}
&\mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)\right) = \\
&\quad = \int_{(0,1]^{a+b}} \mathrm{HL}\left(\Gamma\left((\boldsymbol{\alpha},\boldsymbol{\beta})\right)\right) \mathrm{d}C^{XY}\left((\boldsymbol{\alpha},\boldsymbol{\beta})\right) \\
&\quad = \int_{(0,1]^a}\int_{(0,1]^b} \mathrm{HL}\left(\Gamma^X(\boldsymbol{\alpha}) \times \Gamma^Y(\boldsymbol{\beta})\right) \mathrm{d}C^X(\boldsymbol{\alpha})\, \mathrm{d}C^Y(\boldsymbol{\beta}) \\
&\quad = \int_{(0,1]^a}\int_{(0,1]^b} \mathrm{HL}\left(\Gamma^X(\boldsymbol{\alpha})\right) + \mathrm{HL}\left(\Gamma^Y(\boldsymbol{\beta})\right) \mathrm{d}C^X(\boldsymbol{\alpha})\, \mathrm{d}C^Y(\boldsymbol{\beta}) \\
&\quad = \int_{(0,1]^a} \mathrm{HL}\left(\Gamma^X(\boldsymbol{\alpha})\right) \mathrm{d}C^X(\boldsymbol{\alpha}) + \int_{(0,1]^b} \mathrm{HL}\left(\Gamma^Y(\boldsymbol{\beta})\right) \mathrm{d}C^Y(\boldsymbol{\beta}) \\
&\quad = \mathrm{NL}\left(\left(\mathscr{F}^X, P_\Gamma^X\right)\right) + \mathrm{NL}\left(\left(\mathscr{F}^Y, P_\Gamma^Y\right)\right)
\end{aligned}$$

□

8.6.3 Coordinate invariance

The NL measure does not change under isometric transformations of the focal elements. To see this, let $\mathscr{T}$ be the set of all isometries on X. Notice that according to property HL5, if $A \subseteq X$, then $\mathrm{HL}(T(A)) = \mathrm{HL}(A)$ for all $T \in \mathscr{T}$. Now, let $(\mathscr{F}, P_\Gamma)$ be an infinite RS defined also on X. Suppose that for every focal element $\gamma \in \mathscr{F}$, we choose any isometry $T_\gamma \in \mathscr{T}$ (can be one different for every element; this is the reason of the subindex γ). In consequence, we can form a new random set $\left(\mathscr{F}', P_\Gamma'\right)$ with $\mathscr{F}' := \left\{\gamma' := T_\Gamma(\gamma) : \gamma \in \mathscr{F}\right\}$ and $P_\Gamma' := P_\Gamma$; in consequence,

$$\begin{aligned}
\mathrm{NL}\left((\mathscr{F}, P_\Gamma)\right) &= \int_{\mathscr{F}} \mathrm{HL}(\gamma)\, \mathrm{d}P_\Gamma(\gamma) \\
&= \int_{\mathscr{F}} \mathrm{HL}\left(T_\Gamma(\gamma)\right) \mathrm{d}P_\Gamma(\gamma) \\
&= \int_{\mathscr{F}'} \mathrm{HL}(\gamma')\, \mathrm{d}P_\Gamma'(\gamma') \\
&= \mathrm{NL}\left(\left(\mathscr{F}', P_\Gamma'\right)\right)
\end{aligned}$$

This shows property NL5.

8.6.4 Range

If all focal elements are bounded, then the NL measure will be nonnegative and finite. In fact, using Property HL1, the HL measure will be nonnegative and finite for all bounded focal elements; therefore, $0 \le \mathrm{HL}(\gamma) \le U$ for all γ and for some $U < \infty$. Thus, according to (8.15), $0 \le \mathrm{NL}\left((\mathscr{F}, P_\Gamma)\right) \le \int_{(0,1]^d} U\, \mathrm{d}C(\boldsymbol{\alpha}) = U$.

Suppose that almost all (with regard to P_Γ) focal elements of $(\mathscr{F}, P_\Gamma)$ are singletons; these singletons can be translated using Property NL5 of coordinate invariance to some $\{x\}$. This element will have probability one, i.e., $P_\Gamma(\{x\}) = 1$ while all others will have a null P_Γ-measure. In consequence $\mathrm{NL}\left((\mathscr{F}, P_\Gamma)\right) = \mathrm{HL}(\{x\}) = 0$.

Conversely, suppose that $\mathrm{NL}((\mathscr{F},P_\Gamma))=0$, then this implies that $\mathrm{HL}(\gamma)=0$ for almost all $\gamma\in\mathscr{F}$, since $\mathrm{NL}((\mathscr{F},P_\Gamma))\geq 0$ and since $\mathrm{HL}(\gamma)=0$ if and only if γ is a singleton or the empty set.

The justification for Property NL1 follows.

8.7 Uniqueness of the generalized NL measure for infinite random sets of indexable type

Our motivation for this section comes from the following quote from Klir (2006, p. 198): *"... the strongest justification of a functional as a meaningful measure of the amount of uncertainty of a considered type in a given uncertainty theory is obtained when we can prove that it is the only functional that satisfies the relevant axiomatic requirements and measures the amount of uncertainty in some specific measurement unit"*.

The goal of this section is to show that the generalized nonspecificity measure NL defined by equation (8.15), is the unique measure of nonspecificity for infinite random sets of indexable type that satisfies the Properties NL1 to NL5, NL7 and NL8.

The main theorem of this section follows:

Theorem 8.7.1. *Given some function* HL *that fulfills the set of Properties HL1 to HL7, there is a unique function* $\mathrm{NL}:\mathscr{M}\to[0,\infty)$ *which satisfies the Properties to NL1 to NL5, NL7 and NL8, where* $\mathscr{M}$ *is the system of all infinite random sets of indexable type with convex and bounded focal elements. Furthermore, for any* $(\mathscr{F},P_\Gamma)\in\mathscr{M}$*, with* $P_\Gamma\equiv\mu_C$*, we have*

$$\mathrm{NL}((\mathscr{F},P_\Gamma))=\int_{(0,1]^d}\mathrm{HL}(\Gamma(\boldsymbol{\alpha}))\,\mathrm{d}C(\boldsymbol{\alpha}) \tag{8.20}$$

Proof. In the following proof, $(\{A\},1)$ will represent a random set whose unique focal element is the set A.

(i) NL for sets. Consider a RS of the form $(\{A\},1)$. In this special case NL1 to NL5 reduces to HL1 to HL5, and therefore,

$$\mathrm{NL}((\{A\},1)):=\mathrm{HL}(A) \tag{8.21}$$

where HL is any function that fulfills HL1 to HL5 and HL7[2]

(ii) NL for finite random sets. Let $\left(\mathscr{F}^2,P_\Gamma^2\right)$ be a RS such that $\mathscr{F}^2=\left\{A_1^2,A_2^2\right\}$, $\sigma_{\mathscr{F}^2}=\left\{\emptyset,\mathscr{F}^2,A_1^2,A_2^2\right\}$ and $P_\Gamma^2(\emptyset)=0$, $P_\Gamma^2\left(\mathscr{F}^2\right)=1$, $P_\Gamma^2\left(A_1^2\right)=m_1$ and $P_\Gamma^2\left(A_2^2\right)=m_2$, with $m_1+m_2=1$.

[2]The Hartley-like measure (8.3) is one function that satisfies HL1 to HL7, if in addition A is convex and bounded (there is not guarantee that the Hartley-like measure is the unique function that fulfills that set of properties).

Using the branching Property NL8, it follows that,

$$\mathrm{NL}\left(\left(\mathscr{F}^2, P_\Gamma^2\right)\right) = P_\Gamma^2\left(A_1^2\right)\mathrm{NL}\left(\left(\left\{A_1^2\right\}, 1\right)\right) + P_\Gamma^2\left(A_2^2\right)\mathrm{NL}\left(\left(\left\{A_2^2\right\}, 1\right)\right) \tag{8.22}$$

Using (8.21) in (8.22) it yields

$$\mathrm{NL}\left(\left(\mathscr{F}^2, P_\Gamma^2\right)\right) = P_\Gamma^2\left(A_1^2\right)\mathrm{HL}\left(A_1^2\right) + P_\Gamma^2\left(A_2^2\right)\mathrm{HL}\left(A_2^2\right)$$

Suppose that for every finite random set with n focal elements, $\mathscr{F}^n = \left\{A_1^n, A_2^n, \ldots, A_i^n, \ldots, A_n^n\right\}$,

$$\mathrm{NL}\left(\left(\mathscr{F}^n, P_\Gamma^n\right)\right) = \sum_{i=1}^{n} P_\Gamma^n\left(A_i^n\right)\mathrm{HL}\left(A_i^n\right)$$

holds. Also, let us consider $\mathscr{F}^{n+1} = \mathscr{F}^n \cup A_{n+1}^{n+1}$ where $A_i^{n+1} = A_i^n$ for $i = 1, \ldots, n$ and $P_\Gamma^{n+1}\left(\mathscr{G}|\mathscr{F}^n\right) = P_\Gamma^n\left(\mathscr{G}\right)$ for all $\mathscr{G} \in \sigma_{\mathscr{F}^n}$. Using again the branching theorem NL8,

$$\begin{aligned}
\mathrm{NL}\left(\left(\mathscr{F}^{n+1}, P_\Gamma^{n+1}\right)\right) &= P_\Gamma^{n+1}\left(\mathscr{F}^{n+1} \setminus A_{n+1}^{n+1}\right)\mathrm{NL}\left(\left(\mathscr{F}^n, P_\Gamma^n\right)\right) \\
&\quad + P_\Gamma^{n+1}\left(A_{n+1}^{n+1}\right)\mathrm{NL}\left(\left(\left\{A_{n+1}^{n+1}\right\}, 1\right)\right) \\
&= P_\Gamma^{n+1}\left(\mathscr{F}^n\right)\mathrm{NL}\left(\left(\mathscr{F}^n, P_\Gamma^{n+1}\left(\cdot|\mathscr{F}^n\right)\right)\right) \\
&\quad + P_\Gamma^{n+1}\left(A_{n+1}^{n+1}\right)\mathrm{HL}\left(A_{n+1}^{n+1}\right)
\end{aligned}$$

P_Γ^{n+1} satisfies $P_\Gamma^{n+1}\left(\mathscr{G}\right) = P_\Gamma^{n+1}\left(\mathscr{G}|\mathscr{F}^n\right)P_\Gamma^{n+1}\left(\mathscr{F}^n\right)$ where $\mathscr{G} \in \sigma_{\mathscr{F}^n}$, $\sigma_{\mathscr{F}^n} \subseteq \sigma_{\mathscr{F}^{n+1}}$ and $P_\Gamma^{n+1}\left(A_{n+1}^{n+1}\right) = 1 - P_\Gamma^{n+1}\left(\mathscr{F}^{n+1} \setminus A_{n+1}^{n+1}\right) = 1 - P_\Gamma^{n+1}\left(\mathscr{F}^n\right)$ then

$$\begin{aligned}
\mathrm{NL}\left(\left(\mathscr{F}^{n+1}, P_\Gamma^{n+1}\right)\right) &= P_\Gamma^{n+1}\left(\mathscr{F}^n\right)\sum_{i=1}^{n} P_\Gamma^{n+1}\left(A_i^{n+1}|\mathscr{F}^n\right)\mathrm{HL}\left(A_i^{n+1}\right) \\
&\quad + P_\Gamma^{n+1}\left(A_{n+1}^{n+1}\right)\mathrm{HL}\left(A_{n+1}^{n+1}\right) \\
&= \sum_{i=1}^{n+1} P_\Gamma^{n+1}\left(A_i^{n+1}\right)\mathrm{HL}\left(A_i^{n+1}\right)
\end{aligned}$$

This shows that

$$\mathrm{NL}\left(\left(\mathscr{F}_n, m\right)\right) = \sum_{i=1}^{n} \mathrm{HL}\left(A_i\right) m\left(A_i\right)$$

is the unique measure of nonspecificity for finite random sets up to the specification of the function HL.

(iii) NL for infinite random sets. Suppose now that $\left(\mathscr{F}^1, m^1\right) \prec \left(\mathscr{F}^2, m^2\right) \prec \cdots \prec \left(\mathscr{F}^i, m^i\right) \prec \cdots$ is a sequence of every-time refining finite RSs of indexable type which converges to $\left(\mathscr{F}, P_\Gamma\right)$. Here $m^1, \ldots, m^i, \ldots$ and P_Γ are all probability measures generated by the copula C, i.e., $m^i \equiv \mu_C$ for all $i = 1, 2, \ldots$ and $P_\Gamma \equiv \mu_C$. Such a sequence can always be created.

Since HL is a bounded function and continuous in the sense of Hausdorff (by Property HL6), according to Theorem 5.6.2 we have that the limit of $\mathrm{NL}\left(\left(\mathscr{F}^i, P_\Gamma^i\right)\right)$

as $i \to \infty$ exists, is unique, and is given by the integral

$$\begin{aligned}\int_{(0,1]^d} \mathrm{HL}(\Gamma(\boldsymbol{\alpha}))\,\mathrm{d}C(\boldsymbol{\alpha}) &= \lim_{i\to\infty} \sum_{A^i \in \mathscr{F}^i} \mathrm{HL}\left(A^i\right) P_\Gamma^i\left(A^i\right) \\ &= \lim_{i\to\infty} \mathrm{NL}\left(\left(\mathscr{F}^i, P_\Gamma^i\right)\right)\end{aligned}$$

□

In consequence, (8.20) appears as the result of this limiting process and fulfills HL1 to HL5 and HL7 (which are just a particularization of NL1 to NL5 and NL7), and NL8. The other properties are shown to be satisfied in Section 8.6.

8.8 Conditional generalized Hartley-like measures

Based on the concept of conditional HL measure, according to the formulas (8.4) and (8.5), we can analogously define the conditional generalized measures $\mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right) | \left(\mathscr{F}^{Y}, P_\Gamma^{Y}\right)\right)$ and $\mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right) | \left(\mathscr{F}^{Y}, P_\Gamma^{X}\right)\right)$ by the formulas,

$$\begin{aligned}&\mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right) | \left(\mathscr{F}^{Y}, P_\Gamma^{Y}\right)\right) := \\ &\quad := \mathrm{E}_{\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)}[\mathrm{HL}\left(\gamma | \gamma^Y\right)] \\ &\quad = \mathrm{E}_{\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)}[\mathrm{HL}\left(\gamma\right)] - \mathrm{E}_{\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)}[\mathrm{HL}\left(\gamma^Y\right)] \\ &\quad = \mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)\right) - \mathrm{E}_{\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)}[\mathrm{HL}\left(\mathrm{proj}_Y \gamma\right)] \\ &\quad = \mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)\right) - \mathrm{NL}\left(\left(\mathscr{F}^{Y}, P_\Gamma^{Y}\right)\right)\end{aligned}$$

where the last equality follows from (8.19); similarly,

$$\mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right) | \left(\mathscr{F}^{Y}, P_\Gamma^{X}\right)\right) := \mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right)\right) - \mathrm{NL}\left(\left(\mathscr{F}^{X}, P_\Gamma^{X}\right)\right)$$

$\mathrm{NL}\left(\left(\mathscr{F}^{XY}, P_\Gamma^{XY}\right) | \left(\mathscr{F}^{Y}, P_\Gamma^{Y}\right)\right)$ is used to measure the uncertainty after pinching $\left(\mathscr{F}^{Y}, P_\Gamma^{Y}\right)$ towards a random set with full specify (i.e, with null nonspecificity), like a constant in the case of intervals or a CDF in the case when $(\mathscr{F}, P_\Gamma)$ represents a probability box.

Chapter 9

Sensitivity analysis using infinite random sets of indexable type

> "That process", said I, "starts upon the supposition that when you have eliminated all which is impossible, then whatever remains, however improbable, must be the truth. It may well be that several explanations remain, in which case one tries test after test until one or other of them has a convincing amount of support. ..."
>
> Sherlock Holmes in "The Blanched Soldier" by Sir Arthur Conan Doyle.

9.1 Introduction

Given a mathematical model of a system, the main aim of sensitivity analysis is to analyze the influence of the model outputs with regard to the variation in the model inputs; in fact, if small changes in an input parameter result in relatively large changes in a model output, the model is said to be *sensitive* to the parameter. This useful tool allows the analyst to recognize which input parameters affect the most the system response. Sensitivity analysis has been widely employed in different fields of research like finance, optimization, optimal design, and control systems to assess the influence of the parameters on the state of the system and to gain insight into the model behavior (see e.g. Saltelli et al. (2000) and references therein).

In particular, in reliability analysis of structural systems, sensitivity analysis is used to *a*) measure how sensitive is the probability of failure to small changes in the material, load or geometry properties of the system and to *b*) recognize which are the design variables that have more influence in that variation. In consequence, sensitivity analysis is a fundamental complement to reliability analysis, because it provides an evaluation of the robustness of the design.

In relation to random set and evidence theories, Ferson et al. (2005); Ferson and Tucker (2006b) pointed out that these theories are themselves global sensitivity analysis tools because they define families of sets that represent the uncertainty

about the available information and they propagate this uncertainty through the model to identify a family of answers which includes the correct one. In this context, sensitivity analysis methods are a useful tool to estimate how much less uncertainty in the computations we would have if additional knowledge about an input were available.

A criticism against imprecise-probabilistic-based uncertainty techniques and in particular of random and evidence theories is that the interval $[\mathrm{Bel}(F), \mathrm{Pl}(F)]$ is too wide to make useful decisions. What these criticisms would not take into account is that these theories produce such wide bounds because they take into consideration all the available information, without including additional suppositions in the calculation that will be reflected in the results. However, those criticisms raise the natural question "which is the best way to spend our resources in order to reduce the width of the interval $[\mathrm{Bel}(F), \mathrm{Pl}(F)]$?".This issue will be analyzed in the present chapter, and a solution based on techniques of sensitivity analysis will be proposed.

The plan of this chapter is as follows: first, in Section 9.2, a brief state-of-the-art review of sensitivity analysis techniques in evidence theory will be presented; in Section 9.3 a method of sensitivity analysis will be proposed based on the Hartley-like measure of nonspecificity, and which turns out to be a generalization of a method recently proposed by Ferson and Tucker (2006b); finally, a numerical example using the proposed method will be presented in Section 9.4.

9.2 Sensitivity analysis in evidence and random set theories

Even though evidence and random set theories have been present since the middle of the 1970's, only recently, Oberkampf and Helton (2002) pointed out the need of procedures to conduct sensitivity analysis. In fact, Helton et al. (2005) acknowledged that although evidence theory has become a known tool within the risk and reliability assessment community, they are unaware of any attempts to develop procedures of sensitivity analysis within it. Up to my knowledge, there are only a few contributions to the problem in consideration; in the following lines, I will list and comment on them.

In the framework of imprecise probabilities, Hall (2005, 2006) proposed a method to extend the method of variance-based sensitivity analysis (see e.g. Saltelli et al. (2000)) to imprecise probability theory. Additionally, Hall (2006) proposed two other sensitivity measures. One is based on the calculation of the partial expected value of perfect information on the credal set, which is defined as the expected gain in utility as the consequence of learning the correct value of the implied basic variables. The other method is based on the calculation of the extreme values of expectation of the Kullback-Leibler entropy on all distributions contained in the credal set.

Bae et al. (2006b) considered two methodologies for sensitivity analysis for the plausibility measure: one with respect to the basic mass assignment of each focal element (which allows to see which expert's opinion is a major uncertainty propagation source) and the other with respect to the vector of system parameters. Their approach is based on finding $\partial \mathrm{Pl}(F)/\partial a$ or $\partial \mathrm{Pl}(F)/\partial m_i$, where a is one of the several parameters that define the characterization of the basic variables (for example, a could stand for a mean or variance of a CDF or the parameters that define a possibility distribution) and m_i is the basic mass assignment of any focal element; although this method provide us information on which parameters have more influence on the value of $\mathrm{Pl}(F)$, it is not a useful tool to reduce the uncertainty $\mathrm{Pl}(F) - \mathrm{Bel}(F)$ inasmuch as it is not possible to freely push those parameters to reduce the value of $\mathrm{Pl}(F) - \mathrm{Bel}(F)$.

In the framework of probability bounds analysis (PBA, see e.g. Ferson et al. (2003b)) Ferson et al. (2005); Ferson and Tucker (2006a,b) showed that PBA is itself a global sensitivity analysis of a probabilistic calculation because it defines bounds of a CDF (by means of probability boxes) that represent the uncertainty about known input distributions and projects this uncertainty through the model to identify a neighborhood of possible answers (another probability box) in a way that guarantees the resulting bounds will enclose completely the CDF of the output. They also proposed a "meta"-sensitivity analysis to determine which variables are the ones that have the largest influence on the variability of the working probability boxes. Since in Section 9.3 a generalization of this strategy will be proposed, this method will be explained in some detail in the following lines.

The method of Ferson and Tucker

This "meta"-sensitivity analysis is based on the idea of pinching one or several input basic variables towards precise CDFs or constants with the aim of hypothetically reducing the epistemic uncertainty, the aleatory uncertainty or both; thereafter, the response of the system is computed in the form of probability boxes, and in a posterior step an indicator that measures the amount of uncertainty contained in the output probability boxes is calculated. Finally, this indicator is compared to the one calculated without pinching the input basic variables.

The goal of the pinching strategy is to assess how much less uncertainty the calculations would have if additional information about the input basic variables were available; in fact pinching a probability box to a CDF takes away the epistemic uncertainty and pinching a basic variable to a constant takes away both aleatory and epistemic uncertainty.

Ferson and Tucker proposed that after mapping the uncertainty through the system, the uncertainty in the output should be measured using some function unc(), and the measure of sensitivity

$$100\left(1 - \frac{\mathrm{unc}(R)}{\mathrm{unc}(T)}\right)\% \tag{9.1}$$

where unc(R) and unc(T) are respectively the amount of uncertainty measured with respect to the output with and without pinching of the input basic variables. The quantity (9.1) is supposed to indicate the reduction of uncertainty for that specific pinching of the input basic variables. In my opinion the indicator (9.1) should be used with caution, because it is only suitable if unc() is a homogeneous function, which may not always be the case. Ferson and Tucker (2006b) proposed

$$\begin{aligned} \text{unc}(S) &:= \left\| \overline{F} - \underline{F} \right\|_1 \\ &= \int_{-\infty}^{\infty} \left| \overline{F}(x) - \underline{F}(x) \right| \mathrm{d}x \end{aligned} \tag{9.2}$$

given the output probability box $S := \langle \underline{F}, \overline{F} \rangle$. This metric vanishes if the probability box $\langle \underline{F}, \overline{F} \rangle$ is reduced to a CDF. Note that this methodology allows to pinch several input variables at the same time, which would be useful to study interactions between groups of basic variables.

Finally, note that the results of the sensitivity analysis exercise are conditional upon the pinching used. As showed by Ferson and Tucker, different pinchings would yield different estimations of the reduction of the uncertainty. In fact, they show that taking in consideration all possible pinchings, then intervals on the indicator (9.1) would be available.

9.2.1 Some concepts

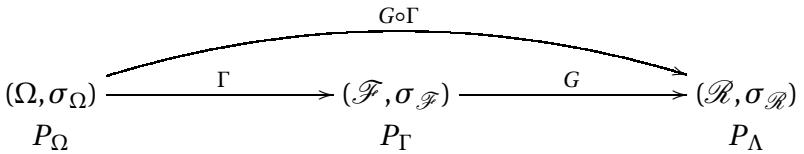
In the following we will define some concepts that will be required in the rest of this chapter.

Extension principle for infinite random sets

Let us consider two nonempty universal sets X and Y. Given the probability space $(\Omega, \sigma_\Omega, P_\Omega)$, the measurable space $(\mathscr{F}, \sigma_{\mathscr{F}})$, $\mathscr{F} \subseteq \mathscr{P}(X)$, the random set $\Gamma : \Omega \to \mathscr{F}$ and a function $g : X \to Y$. Let us define the measurable space $(\mathscr{R}, \sigma_{\mathscr{R}})$ and also the function $G : \mathscr{F} \to \mathscr{R}$ by $\gamma \mapsto \{g(x) : x \in \gamma\}$ such that G is $(\sigma_{\mathscr{F}} - \sigma_{\mathscr{R}})$-measurable. Here $\mathscr{R}$ is the set of images of all focal elements of $\mathscr{F}$ through g, i.e.,

$$\mathscr{R} := \{\lambda := g(\gamma) : \gamma \in \mathscr{F}\} \tag{9.3}$$

and $\sigma_{\mathscr{R}}$ is defined as the minimal σ-algebra generated by the system of sets $g(\sigma_{\mathscr{F}}) := \left\{ \{g(\gamma) : \gamma \in \mathscr{G}\} : \mathscr{G} \in \sigma_{\mathscr{F}} \right\}$.



The mapping $G \circ \Gamma : \Omega \to \mathscr{R}$ is also a random set since the composition of measurable functions is also measurable, and in consequence $G \circ \Gamma$ is $(\sigma_\Omega - \sigma_{\mathscr{R}})$-measurable. We already know that Γ generates a probability measure on $(\mathscr{F}, \sigma_{\mathscr{F}})$ given by $P_\Gamma := P_\Omega \circ \Gamma^{-1}$. Similarly, G can be used to generate a probability measure on $(\mathscr{R}, \sigma_{\mathscr{R}})$ given by $P_\Lambda := P_\Gamma \circ G^{-1} = P_\Omega \circ \Gamma^{-1} \circ G^{-1}$. This means that an event $\mathscr{G} \in \sigma_{\mathscr{R}}$ has the probability

$$\begin{aligned} P_\Lambda(\mathscr{G}) := P_\Gamma\{\gamma \in \mathscr{F} : G(\gamma) \in \mathscr{G}\} &= \int_{\mathscr{F}} I[G(\gamma) \in \mathscr{G}] \, \mathrm{d}P_\Gamma(\gamma) \\ = P_\Omega\{\alpha \in \Omega : G(\Gamma(\alpha)) \in \mathscr{G}\} &= \int_\Omega I[G(\Gamma(\alpha)) \in \mathscr{G}] \, \mathrm{d}P_\Omega(\alpha); \end{aligned} \tag{9.4}$$

the RSs G and $G \circ \Gamma$ will be referred to in the following also as $(\mathscr{R}, P_\Lambda)$.

We will say that the RS $(\mathscr{R}, P_\Lambda)$ was formed using the *extension principle* for random sets and we will call the RS $(\mathscr{R}, P_\Lambda)$ the *image of the RS* $(\mathscr{F}, P_\Gamma)$ *through* g. Note that if $(\mathscr{F}, P_\Gamma)$ is a random set of indexable type then $(\mathscr{R}, P_\Lambda)$ is also of indexable type.

Example: the extension principle for infinite random sets reduces to the extension principle for finite RSs defined by Dubois and Prade (1991) and explained in Section 2.2.5 when dealing with a finite number of focal elements. Note that (9.3) corresponds to (2.10) and (2.11) is a particularization of (9.4).

Pinching

Definition 9.2.1. *Suppose that we are given two random sets of indexable type* $\Gamma^A : \Omega \to \mathscr{F}^A$, *and* $\Gamma^B : \Omega \to \mathscr{F}^B$ *defined on the same probability space* $(\Omega, \sigma_\Omega, \mu_C)$ *for some copula* C *and* $\Omega := (0,1]^d$, $\sigma_\Omega := (0,1]^d \cap \mathscr{B}^d$ *and* $\mathscr{F}^A, \mathscr{F}^B \in \mathscr{P}(X)$. *The random set* Γ^A *is said to be a* pinched random set *of* Γ^B *if for all* $\boldsymbol{\alpha} \in \Omega$, *it follows that* $\Gamma^A(\boldsymbol{\alpha}) \subseteq \Gamma^B(\boldsymbol{\alpha})$.

Relationship between pinching, the interval $[\mathrm{LP}(F), \mathrm{UP}(F)]$ and the $\mathrm{NL}()$ measure

Theorem 9.2.2. *Let* $\mathscr{F}^A, \mathscr{F}^B \in \mathscr{P}(X)$. *Let* $\left(\mathscr{F}^A, P_\Gamma^A\right)$, *and* $\left(\mathscr{F}^B, P_\Gamma^B\right)$ *be defined on the same probability space* $(\Omega, \sigma_\Omega, \mu_C)$, *i.e.* $P_\Gamma^A \equiv \mu_C \equiv P_\Gamma^B$ *for some copula* C. *If* $\left(\mathscr{F}^A, P_\Gamma^A\right)$ *is a pinched random set of* $\left(\mathscr{F}^B, P_\Gamma^B\right)$ *and* $F \subseteq X$, *then*

1. $\left[\mathrm{LP}_{(\mathscr{F}^A, P_\Gamma^A)}(F), \mathrm{UP}_{(\mathscr{F}^A, P_\Gamma^A)}(F)\right] \subseteq \left[\mathrm{LP}_{(\mathscr{F}^B, P_\Gamma^B)}(F), \mathrm{UP}_{(\mathscr{F}^B, P_\Gamma^B)}(F)\right]$ *and thus* $U_{(\mathscr{F}^A, P_\Gamma^A)}(F) \leq U_{(\mathscr{F}^B, P_\Gamma^B)}(F)$;
2. $\mathrm{NL}\left(\left(\mathscr{F}^A, P_\Gamma^A\right)\right) \leq \mathrm{NL}\left(\left(\mathscr{F}^B, P_\Gamma^B\right)\right)$;
3. *if in addition* $\left(\mathscr{R}^A, P_\Lambda^A\right)$ *and* $\left(\mathscr{R}^B, P_\Lambda^B\right)$ *are respectively the images of the random sets* $\left(\mathscr{F}^A, P_\Gamma^A\right)$ *and* $\left(\mathscr{F}^B, P_\Gamma^B\right)$ *through the function* $g : X \to Y$, *then* $\left(\mathscr{R}^A, P_\Lambda^A\right)$ *will also be a pinched RS of* $\left(\mathscr{R}^B, P_\Lambda^B\right)$ *and therefore* $\left[\mathrm{LP}_{(\mathscr{R}^A, P_\Lambda^A)}(F), \mathrm{UP}_{(\mathscr{R}^A, P_\Lambda^A)}(F)\right] \subseteq \left[\mathrm{LP}_{(\mathscr{R}^B, P_\Lambda^B)}(F), \mathrm{UP}_{(\mathscr{R}^B, P_\Lambda^B)}(F)\right]$.

Proof. Part 1. Here it will be proved that $\mathrm{UP}_{(\mathscr{F}^A,P_\Gamma^A)}(F) \leq \mathrm{UP}_{(\mathscr{F}^B,P_\Gamma^B)}(F)$. The demonstration that $\mathrm{LP}_{(\mathscr{F}^A,P_\Gamma^A)}(F) \geq \mathrm{LP}_{(\mathscr{F}^B,P_\Gamma^B)}(F)$ follows the same steps. By Definition 9.2.1 for all $\boldsymbol{\alpha} \in (0,1]^d$ we have that $\Gamma^A(\boldsymbol{\alpha}) \subseteq \Gamma^B(\boldsymbol{\alpha})$, and in consequence $I\left[\Gamma^B(\boldsymbol{\alpha}) \cap F \neq \emptyset\right] \geq I\left[\Gamma^A(\boldsymbol{\alpha}) \cap F \neq \emptyset\right]$; now,

$$\begin{aligned}\mathrm{UP}_{(\mathscr{F}^B,P_\Gamma^B)}(F) &= \int_{(0,1]^d} I\left[\Gamma^B(\boldsymbol{\alpha}) \cap F \neq \emptyset\right] \mathrm{d}C(\boldsymbol{\alpha}) \\ &\geq \int_{(0,1]^d} I\left[\Gamma^A(\boldsymbol{\alpha}) \cap F \neq \emptyset\right] \mathrm{d}C(\boldsymbol{\alpha}) \\ &= \mathrm{UP}_{(\mathscr{F}^A,P_\Gamma^A)}(F)\end{aligned}$$

Part 2. Since the Hartley-like measure is monotone and $\Gamma^A(\boldsymbol{\alpha}) \subseteq \Gamma^B(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in (0,1]^d$, we have that

$$\begin{aligned}\int_{(0,1]^d} \mathrm{HL}\left(\Gamma^A(\boldsymbol{\alpha})\right) \mathrm{d}C(\boldsymbol{\alpha}) &= \int_{(0,1]^d} \mathrm{HL}\left(\Gamma^A(\boldsymbol{\alpha})\right) \mathrm{d}C(\boldsymbol{\alpha}) \\ &\leq \int_{(0,1]^d} \mathrm{HL}\left(\Gamma^B(\boldsymbol{\alpha})\right) \mathrm{d}C(\boldsymbol{\alpha})\end{aligned}$$

Part 3. This follows inmediately from the fact that if $A \subseteq B$ then $g(A) \subseteq g(B)$, where $g(A)$ and $g(B)$ are the images of A and B through the function g. The rest of the theorem follows from Part 1. □

Theorem 9.2.2 shows that the epistemic uncertainty present in a basic variable defined in $\mathbb{R}$ can be reduced expressing the information by means of smaller sets to denote the focal elements; for example, by using fuzzy normalized subsets to express possibility distributions, by employing narrower probability boxes that are contained inside the original probability box, or by using narrower intervals, when the information is shaped as a Dempster-Shafer body of evidence.

9.3 A method for reduction of uncertainty based on the Hartley-like measure of nonspecificity

We will discuss how to reduce the width of the interval $[\mathrm{LP}(F), \mathrm{UP}(F)]$ under the supposition that we may be able to obtain additional information. We will see later that this method is a generalized version of the method of sensitivity analysis proposed by Ferson and Tucker (2006b).

Consider an infinite RS $(\mathscr{F}, P_\Gamma)$ defined on X. Let $U_{(\mathscr{F},P_\Gamma)}(F)$ be the width of the interval $\left[\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F), \mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)\right]$, i.e.,

$$U_{(\mathscr{F},P_\Gamma)}(F) := P_\Gamma\left\{\gamma : \gamma \cap F \neq \emptyset, \gamma \in \mathscr{F}\right\} - P_\Gamma\left\{\gamma : \gamma \subseteq F, \gamma \in \mathscr{F} \setminus \emptyset\right\} \tag{9.5}$$

Note that if in equation (9.5) $\mathscr{F}$ is a specific focal set, i.e., every focal element γ is a singleton, then $U_{(\mathscr{F},P_\Gamma)}(F) = 0$; this is in fact the case when all basic variables are random.

In relation to equation (9.5), only those focal elements γ that share points with F and $X \setminus F$ are contributing to $U_{(\mathscr{F},P_\Gamma)}(F)$. If we could reduce in some sense the size of those elements, then the value of $U_{(\mathscr{F},P_\Gamma)}(F)$ could be also reduced.

Now the question is how to measure the amount of possible alternatives residing in that focal element? One natural possibility is the Lebesgue measure of the joint focal element. It is not very useful when the joint focal element is (a) degenerate (box), because in this case the Lebesgue measure of the joint focal element will be zero. An alternative measure is the Hartley-like measure of nonspecificity, defined by equation (8.3). Using it, the nonspecificity of an infinite random set, NL(), is determined by equation (8.14). This measure is adequate to estimate the amount of possible alternatives residing in the focal set because it is zero when $\mathscr{F}$ is specific and also circumvents the problem of the Lebesgue measure when the focal elements are degenerate. In addition, the NL() measure is closely related to pinching according to Theorem 9.2.2, Part 2.

The only way to reduce the value of $U_{(\mathscr{F},P_\Gamma)}(F)$ is to obtain more information about the basic variables. Pinching therefore suggest a method to simulate the acquisition of additional information about the basic variables; in other words, the role of pinching is to artificially eliminate all epistemic uncertainty from those basic variables under the assumption that additional information were available. In fact, according to Theorem 9.2.2 (Part 3), when one or several basic variables forming the random set $(\mathscr{F},P_\Gamma)$ are pinched, in order to form the pinched random set $(\mathscr{F},P_\Gamma)_I$, and if $(\mathscr{R},P_\Lambda)_I$ and $(\mathscr{R},P_\Lambda)$ are respectively their corresponding images through some function $g : X \to Y$ then for $F \subseteq X$ we have that $\left[\mathrm{LP}_{(\mathscr{R},\Lambda)_I}(g(F)),\mathrm{UP}_{(\mathscr{R},\Lambda)_I}(g(F))\right] \subseteq \left[\mathrm{LP}_{(\mathscr{R},\Lambda)}(g(F)),\mathrm{UP}_{(\mathscr{R},\Lambda)}(g(F))\right]$.

We proceed as follows:

- first, the basic variables $I \subset \{1,\ldots,n\}$ in $(\mathscr{F},P_\Gamma)$ must be pinched towards either possibility distributions, probability boxes or intervals contained in the previous ones to form the pinched random set $(\mathscr{F},P_\Gamma)_I$;
- in a second step, a sample of n focal elements is drawn from $(\mathscr{F},P_\Gamma)_I$ to form a finite random set $(\mathscr{F}_n,m)_I$ as explained in Chapter 4.
- then, after mapping the input RS $(\mathscr{F}_n,m)_I$ through a function g that represents the mathematical model of the system, we obtain the RS $(\mathscr{R}_{n'},\rho)_I$ and then its non-specificity is calculated, i.e. $\mathrm{NL}((\mathscr{R}_{n'},\rho)_I)$.
- finally, a ranking of the images of the pinched random sets $(\mathscr{R}_{n'},\rho)_I$ according to the Hartley-like measure must be performed.

In this last step, we hope that the smaller NL() is, the largest the expected reduction of the epistemic uncertainty in the model is, and in consequence, the bigger the expected reduction of the width of the interval $\left[\mathrm{Bel}_{(\mathscr{R},P_\Lambda)}(g(F)),\mathrm{Pl}_{(\mathscr{R},P_\Lambda)}(g(F))\right]$ is; however, this is not always true, as shown by Figure 9.1. In other words, if two different RSs $(\mathscr{F},P_\Gamma)_I$ and $(\mathscr{F},P_\Gamma)_J$ are given, which result from pinching $(\mathscr{F},P_\Gamma)$ and such that $\mathrm{NL}((\mathscr{F},P_\Gamma)_I) \leq \mathrm{NL}((\mathscr{F},P_\Gamma)_J)$,

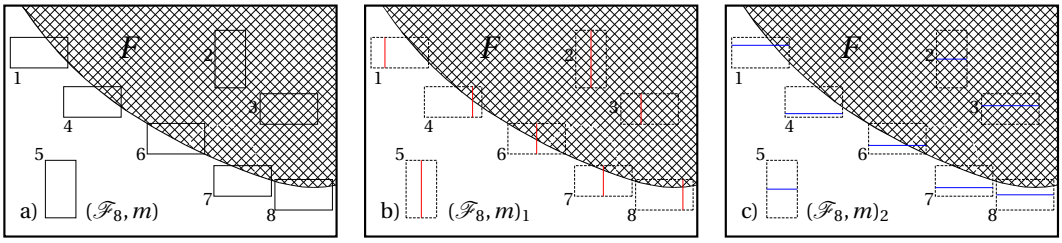


Figure 9.1: Pinching in both variables. Figure a) shows the finite RS $(\mathscr{F}_8, m)$; here every box represents one of the focal elements. Observe that $U_{(\mathscr{F}_8,m)} = 5/8$. Figures b) and c) represent respectively the pinched RSs $(\mathscr{F}_8, m)_1$ and $(\mathscr{F}_8, m)_2$; here every line represents the new focal element. Note that even though $\mathrm{NL}((\mathscr{F}_8, m)_1) \leq \mathrm{NL}((\mathscr{F}_8, m)_2)$ we have that $2/8 = U_{(\mathscr{F}_8,m)_2}(F) \leq U_{(\mathscr{F}_8,m)_1}(F) = 4/8$

.

this does not necessarily imply that $U_{(\mathscr{R},P_\Lambda)_I}(g(F)) \leq U_{(\mathscr{R},P_\Lambda)_J}(g(F))$; what a selection by the NL() measure guarantees is that the region where the true solution resides is smaller. Therefore, one must spend most of the efforts in obtaining additional information about those variables whose non-specificity NL() after pinching is small.

If we analyze this problem using an analogy to the Buffon's needle problem, it is more likely that a sample from a random set with a large area and "perimeter" will contain points of ∂F than one with small area and "perimeter". The relation between the NL() measure and the probability that a randomly sampled focal element intersects ∂F should be analyzed therefore from the viewpoint of the theory of geometric probability, which is the branch of mathematics which studies Buffon's problem.

An important feature of the proposed method is that it can cope with the influence of the scale, by taking into account the range where the basic variables usually dwell; this is done by standarizing every input variable to the same interval. Suppose for example that the basic variable x usually is found in the interval $[x_{\min}, x_{\max}]$; then a standarized variable x' can be found by the transformation

$$x' := \frac{x - x_{\min}}{x_{\max} - x_{\min}}. \tag{9.6}$$

Also the proposed method can cope with the effects of a factor while all others are varying and also is able to treat grouped factors as if they were a single factor.

This technique results to be a generalization of the one proposed by Ferson and Tucker (2006b) but, with the NL() indicator employed instead of the unc() of the associated probability box, as described in Section 9.2. In fact, the method of Ferson and Tucker requires to convert all output random sets to their associated probability boxes; then for every probability box, the value (9.2) is calculated. Now, the nonspecificity of a probability box can be calculated according to equation (8.17). Note that since $x \mapsto \log_2 x$ is a monotone increasing function in x, then the ordering of the pinchings will be the same if the integrals

$\int_0^1 1 + \underline{F}^{(-1)}(\alpha) - \overline{F}^{(-1)}(\alpha)\,\mathrm{d}\alpha$ or $\int_0^1 \underline{F}^{(-1)}(\alpha) - \overline{F}^{(-1)}(\alpha)\,\mathrm{d}\alpha$ are used. This last integral is equivalent to the area between the CDFs that define the probability box $\langle \underline{F}, \overline{F} \rangle$, and corresponds to the quantity calculated using (9.2).

Note that the indicator (9.1) should not be used in conjunction with the nonspecificity NL(), inasmuch as NL() is a nonhomogeneous function.

A final remark on how to do pinching of probability boxes is appropriate. Ferson and Tucker (2006b) already said that one should take a central estimate in some sense or nominal value used in an assessment of the probability box in consideration; for instance if $\langle \underline{F}, \overline{F} \rangle = \text{NormalCDF}(\mu = [2\ 4], \sigma = [0.5\ 1.5])$, then a natural candidate that can be used for pinching could be $\text{NormalCDF}(\mu = 3, \sigma = 1.0)$. In my opinion, if the information available is cast in a probability box without any structure, that is, that does not appear as a family of distributions with varying parameters, then the maximum entropy method (see e.g. Klir and Wierman (1998)) could be applied to extract the representative CDF; this can be approximated in the following way: consider the probability box $\langle \underline{F}, \overline{F} \rangle$. Discretize the domain of this probability box into n bins, and form the associated histograms $\{[x_{i-1}, x_i], \underline{p}_i, i = 1, \ldots, n\}$ and $\{[x_{i-1}, x_i], \overline{p}_i, i = 1, \ldots, n\}$ where $\underline{p}_i = \underline{F}(x_i) - \underline{F}(x_{i-1})$ and $\overline{p}_i = \overline{F}(x_i) - \overline{F}(x_{i-1})$. Now consider the optimization problem, $\max S = -\sum_{i=1}^n p_i \log p_i$ subject to $p_i \in [\underline{p}_i, \overline{p}_i]$ for $i = 1, \ldots, n$ and $\sum_{i=1}^n p_i = 1$. The argument that maximizes S is the histogram corresponding to the CDF that "has the maximum entropy" in the probability box $\langle \underline{F}, \overline{F} \rangle$. Figure 9.2 shows an example of the application of the technique.

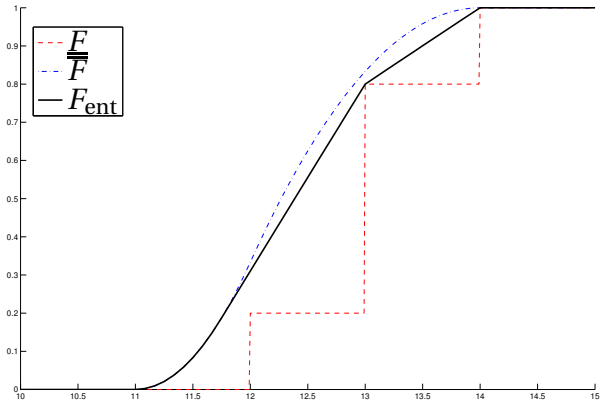


Figure 9.2: This figure shows the probability box $\langle \underline{F}, \overline{F} \rangle$ defined by $\underline{F} = 0.2H(x - 12) + 0.6H(x - 13) + 0.2H(x - 14)$, where H is the Heaviside step function and $\overline{F} = \text{CDFTrian}(x, 11, 12, 14)$. In addition it shows the CDF that has the maximum entropy of this probability box, namely F_{ent}, which was calculated using $n = 60$.

9.4 Numerical example

To test the proposed method, we will employ the example used by Ferson and Tucker (2006b) and originally considered by Hall and Lawry (2001). Let us consider a dike employed as a flood defense, and which has an associated limit state function,

$$g(\boldsymbol{x}) = \Delta D - H_s \frac{\sin\beta}{M\sqrt{s_{\text{op}}}}$$

where $\boldsymbol{x} = [\Delta, D, H_S, M, s_{\text{op}}]$, Δ is the relative density of the revetment blocks, D is the diameter of the revetment blocks, H_s is the significant wave height, β is the slope of the revetment, M is a model parameter and s_{op} is the offshore peak wave steepness. The flood defense fails when $g(\boldsymbol{x}) \leq 0$. These variables are considered to be independent.

Ferson and Tucker (2006b) considered the following parameters: $\Delta = [1.60,\ 1.65]$, $D = [0.68,\ 0.72]$ meters, $\beta = [0.309,\ 0.328]$ radians, $M = [3.0,\ 5.2]$, $H_S \sim$ WeibullCDF(scale = $[1.2,\ 1.5]$ meters, shape = $[10,\ 12]$) and $s_{\text{op}} \sim$ NormalCDF($\mu = [0.029,\ 0.041], \sigma = [0.005,\ 0.006]$).

The problem was solved using the technique proposed in Section 9.3 and for comparative purposes, using the method employed by Ferson and Tucker. In the first case, we employed a product copula to model random set independence, and performed simple Monte Carlo drawing $n = 10$ and $n = 1000$ samples from $(\mathscr{F}, P_\Gamma)$ in order to form $(\mathscr{F}_n, m)$.

In the second case, following the steps carried on by Ferson and Tucker (2006b), every basic variable was represented by a unidimensional Dempster-Shafer body of evidence. The intervals Δ, D, β and M are represented in a straightforward manner by assigning them a basic mass assignment of 1.0. The probability boxes H_S and s_{op} were discretized into 100 focal elements, each having a basic mass assignment of 0.01. In this step, we employed the averaging discretization method proposed by Tonon (2004c) which is a special case of the intermediate discretization of probability boxes, discussed in Section 7.2. Thereafter, a joint body of evidence $(\mathscr{F}_n, m)$ with $n = 10000$ focal elements was created, each one of them with a basic mass assignment of 0.0001.

The random set $(\mathscr{R}_{n'}, \rho)$ was obtained after mapping $(\mathscr{F}_n, m)$ through g using the extension principle for finite random sets and the optimization method (that is the image of $A_i \in \mathscr{F}_n$, namely $R_j \in \mathscr{R}_{n'}$, was calculated as $\left[\min_{x\in A_i} g(x), \max_{x\in A_i} g(x)\right]$, provided A_i is connected and compact). Then, the associated probability box $\langle \underline{F}, \overline{F} \rangle$ was computed, and using equation (9.2) $\text{unc}\left((\mathscr{R}_{n'}, \rho)\right) := \text{unc}(\langle \underline{F}, \overline{F} \rangle)$. Besides, $\text{NL}\left((\mathscr{R}_{n'}, \rho)\right)$ was calculated using equation (8.14). Note that for the computation of $\text{NL}\left((\mathscr{R}_{n'}, \rho)\right)$ we do not require the intermediate step of finding $\langle \underline{F}, \overline{F} \rangle$.

Then the following nominal pinchings were performed: $\Delta = 1.625$, $D = 0.70$ meters, $\beta = 0.3185$ radians, $M = 4.1$, $H_S \sim$ WeibullCDF(scale = 1.35 meters, shape = 11) and $s_{\text{op}} \sim$ NormalCDF($\mu = 0.04, \sigma =$

Table 9.1: Ranking of importance of the basic variables. Here unc() refers to the measure of uncertainty proposed by Ferson and Tucker (2006b) and NL() refers to the Hartley-like-based measure proposed here. The smaller those indicators are, the largest is the influence of the epistemic uncertainty of that variable. Here $F = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq 0\}$.

Pinched variable i	$\text{unc}((\mathscr{R}_{n'}, \rho)_i)$	$\text{NL}((\mathscr{R}_{n'}, \rho)_i)$	$\text{Bel}_{(\mathscr{F}_n, m)_i}(F)$	$\text{Pl}_{(\mathscr{F}_n, m)_i}(F)$
M	0.2665	0.3406	0	0
H_S	0.4439	0.5291	0	0.0013
D	0.5034	0.5872	0	0.0049
Δ	0.5334	0.6157	0	0.0059
β	0.5363	0.6186	0	0.0049
s_{op}	0.5476	0.6293	0	0.0032

	$\text{unc}((\mathscr{F}_n, m))$	$\text{NL}((\mathscr{F}_n, m))$	$\text{Bel}_{(\mathscr{F}_n, m)}(F)$	$\text{Pl}_{(\mathscr{F}_n, m)}(F)$
No pinching	0.5684	0.6483	0	0.0073

0.0055). Every basic variable was pinched while the others were allowed to vary. For comparison purposes with Ferson and Tucker, no standarization of the input basic variables was performed; take into consideration that a good standarization strategy would require more data, namely, the range where every basic variable dwells, as required by equation (9.6); this information is usually known by the reliability analyst.

Table 9.1 shows the NL() and the unc() measures of $(\mathscr{R}_{n'}, \rho)$ after pinching every basic variable at a time and without pinching them at all, in the case of discretizing the basic variables. Slightly different results were obtained here after repeating the calculation in comparison to those published in Ferson and Tucker (2006b). As already noted in Section 9.3, both techniques provide the same ordering of importance of the basic variables (the lower these measures are, the more important the basic variable is); therefore, the most important variable to study is M, followed in order by H_S, D, Δ, β, and s_{op}. The Table also shows the belief and plausibility of the failure region $F = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq 0\}$ with respect to the random set $(\mathscr{F}_n, m)_i$.

Table 9.2 shows the NL() and the unc() measure of $(\mathscr{F}_n, m)$ after pinching every basic variable at a time and without pinching them at all employing the proposed Monte Carlo simulation strategy. In comparison with Table 9.1, we required much less simulations ($n = 1000$) to obtain approximately the same values of $\text{NL}\left((\mathscr{R}_{n'}, \rho)\right)$ and $\text{unc}\left((\mathscr{R}_{n'}, \rho)\right)$. However, since our main goal is to make a ranking of the basic variables according to the reduction of uncertainty, the estimation of $\text{NL}\left((\mathscr{R}_{n'}, \rho)\right)$ was performed using only $n = 10$ simulations, as shown in Table 9.3. Although this computation is not as accurate as the one with $n = 1000$ simulations, it provides the same ranking. These ten simulations should be compared against the $n = 10000$ focal elements required by the method of Ferson and

Table 9.2: Ranking of importance of the basic variables using $n = 1000$ Monte Carlo simulations

Pinched variable i	$\text{unc}((\mathscr{R}_{n'}, \rho)_i)$	$\text{NL}((\mathscr{R}_{n'}, \rho)_i) \pm \hat{\sigma}(\text{HL}(\Gamma(\boldsymbol{\alpha}))_i)$	$\text{Bel}_{(\mathscr{F}_n, m)_i}(F) \pm \hat{\sigma}(I[\boldsymbol{\alpha} \in F_{\text{LP}}]_i)$	$\text{Pl}_{(\mathscr{F}_n, m)_i}(F) \pm \hat{\sigma}(I[\boldsymbol{\alpha} \in F_{\text{UP}}]_i)$
M	0.2609	0.3342±0.0212	0±0	0±0
H_S	0.4380	0.5232±0.0478	0±0	0±0
D	0.4974	0.5814±0.0540	0±0.0316	0.0010±0.0010
Δ	0.5274	0.6101±0.0530	0±0.0548	0.0030±0.0030
β	0.5304	0.6131±0.0489	0±0.0316	0.0010±0.0010
s_{op}	0.5417	0.6237±0.0469	0±0	0±0

	$\text{unc}((\mathscr{F}_n, m))$	$\text{NL}((\mathscr{F}_n, m)) \pm \hat{\sigma}(\text{HL}(\Gamma(\boldsymbol{\alpha})))$	$\text{Bel}_{(\mathscr{F}_n, m)}(F) \pm \hat{\sigma}(I[\boldsymbol{\alpha} \in F_{\text{LP}}])$	$\text{Pl}_{(\mathscr{F}_n, m)}(F) \pm \hat{\sigma}(I[\boldsymbol{\alpha} \in F_{\text{UP}}])$
No pinching	0.5624	0.6428±0.0518	0±0.0633	0.0040±0.0040

Table 9.3: Ranking of importance of the basic variables using $n = 10$ Monte Carlo simulations

Pinched variable i	$\text{unc}((\mathscr{R}_{n'}, \rho)_i)$	$\text{NL}((\mathscr{R}_{n'}, \rho)_i) \pm \hat{\sigma}(\text{HL}(\Gamma(\boldsymbol{\alpha}))_i)$
M	0.2637	0.3376±0.0136
H_S	0.4490	0.5344±0.0441
D	0.5115	0.5953±0.0476
Δ	0.5415	0.6237±0.0466
β	0.5433	0.6255±0.0427
s_{op}	0.5573	0.6384±0.0432

	$\text{unc}((\mathscr{F}_n, m))$	$\text{NL}((\mathscr{F}_n, m)) \pm \hat{\sigma}(\text{HL}(\Gamma(\boldsymbol{\alpha})))$
No pinching	0.5804	0.6591±0.0617

Tucker (2006b); more striking results are expected in higher dimensions.

Since in Tables 9.2 and 9.3, NL, Bel and Pl were calculated using Monte Carlo simulation, we can also calculate some estimator of the deviation of those indicators, using the equation (7.14).

9.5 How to recalculate the belief and plausibility after obtaining new information

After performing the sensitivity study, the analyst will take decisions on how to invest the available resources to obtain new information and in this way reduce the width of the interval $\left[\text{LP}_{(\mathscr{F}, P_\Gamma)}(F), \text{UP}_{(\mathscr{F}, P_\Gamma)}(F)\right]$. If it were possible to obtain

new information for some basic variables that is consistent with the one contained in $(\mathscr{F}, P_\Gamma)$, we could form the new joint RS $(\mathscr{F}', P'_\Gamma)$ that is basically a pinched RS of the initial RS $(\mathscr{F}, P_\Gamma)$; according to Theorem 9.2.2 we have that $[\mathrm{LP}_{(\mathscr{F}', P'_\Gamma)}(F), \mathrm{UP}_{(\mathscr{F}', P'_\Gamma)}(F)] \subseteq [\mathrm{LP}_{(\mathscr{F}, P_\Gamma)}(F), \mathrm{UP}_{(\mathscr{F}, P_\Gamma)}(F)]$.

Suppose that we sampled from $(\mathscr{F}', P'_\Gamma)$, by means of n Monte Carlo simulations, the finite random set $(\mathscr{F}'_n, m')$; then the lower and upper probabilities $\mathrm{LP}_{(\mathscr{F}', P'_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F}', P'_\Gamma)}(F)$, could be estimated by means of the belief and plausibility $\mathrm{Bel}_{(\mathscr{F}'_n, m')}(F)$ and $\mathrm{Pl}_{(\mathscr{F}'_n, m')}(F)$ of F with respect to the the finite random set $(\mathscr{F}'_n, m')$. However, we could use the following procedure in order to save lots of computational effort.

1. When sampling the random set $(\mathscr{F}, P_\Gamma)$ to form $(\mathscr{F}_n, m)$, we keep in some database the samples $\boldsymbol{\alpha} \in (0,1]^d$ which belong to $F_{\mathrm{UP}} \setminus F_{\mathrm{LP}}$, where F_{UP} and F_{LP} are the sets in the α-space associated to the random set $(\mathscr{F}, P_\Gamma)$ and the set $F \subseteq X$.

2. For every one of the points $\boldsymbol{\alpha} \in F_{\mathrm{UP}} \setminus F_{\mathrm{LP}}$, check if they belong to the regions F'_{UP} and F'_{LP} associated to the random set $(\mathscr{F}', P'_\Gamma)$ and the set F. This is done by evaluating again the conditions $\Gamma(\boldsymbol{\alpha}) \cap F \neq \emptyset$ and $\Gamma(\boldsymbol{\alpha}) \subseteq F$ respectively.

3. The updated belief and plausibility given by,

$$\mathrm{Bel}_{(\mathscr{F}'_n, m')}(F) := \mathrm{Bel}_{(\mathscr{F}_n, m)}(F) + \frac{\left|\{\boldsymbol{\alpha} \in F_{\mathrm{UP}} \setminus F_{\mathrm{LP}} : \boldsymbol{\alpha} \in F'_{\mathrm{LP}}\}\right|}{n}$$
$$\mathrm{Pl}_{(\mathscr{F}'_n, m')}(F) := \mathrm{Pl}_{(\mathscr{F}_n, m)}(F) - \frac{\left|\{\boldsymbol{\alpha} \in F_{\mathrm{UP}} \setminus F_{\mathrm{LP}} : \boldsymbol{\alpha} \in F'_{\mathrm{UP}}\}\right|}{n}$$

   are just the approximations of $\mathrm{LP}_{(\mathscr{F}', P'_\Gamma)}(F)$ and $\mathrm{UP}_{(\mathscr{F}', P'_\Gamma)}(F)$ that we are looking for.

This method of updating the belief and plausibility is much more efficient than calculating $\mathrm{Bel}_{(\mathscr{F}'_n, m')}(F)$ and $\mathrm{Pl}_{(\mathscr{F}'_n, m')}(F)$ again from scratch since usually $\left|\{\boldsymbol{\alpha} \in F_{\mathrm{UP}} \setminus F_{\mathrm{LP}} : \boldsymbol{\alpha} \in F'_{\mathrm{UP}}\}\right| \ll n$.

The method is based on the following theorem:

Theorem 9.5.1. *Let be $\mathscr{F}^A, \mathscr{F}^B \in \mathscr{P}(X)$. Let $(\mathscr{F}^A, P^A_\Gamma)$, and $(\mathscr{F}^B, P^B_\Gamma)$ be defined on the same probability space $(\Omega, \sigma_\Omega, \mu_C)$, i.e. $P^A_\Gamma \equiv \mu_C \equiv P^B_\Gamma$ for some copula C. If $(\mathscr{F}^A, P^A_\Gamma)$ is a pinched random set of $(\mathscr{F}^B, P^B_\Gamma)$ and $F \subseteq X$, then $F^B_{\mathrm{LP}} \subseteq F^A_{\mathrm{LP}}$ and $F^B_{\mathrm{UP}} \supseteq F^A_{\mathrm{UP}}$.*

Proof. Using Definition 9.2.1, $\Gamma^A(\boldsymbol{\alpha}) \subseteq \Gamma^B(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in (0,1]^d$. This implies

$$I\left[\Gamma^B(\boldsymbol{\alpha}) \cap F \neq \emptyset\right] \geq I\left[\Gamma^A(\boldsymbol{\alpha}) \cap F \neq \emptyset\right] \tag{9.7}$$
$$I\left[\Gamma^B(\boldsymbol{\alpha}) \subseteq F\right] \leq I\left[\Gamma^A(\boldsymbol{\alpha}) \subseteq F\right] \tag{9.8}$$

Also, since $(\mathscr{F}^A, P^A_\Gamma)$ and $(\mathscr{F}^B, P^B_\Gamma)$ are random sets of indexable type, then related to F there exist in the α-space of those random sets the sets F^A_{LP}, F^A_{UP}, F^B_{LP} and F^B_{UP}.

But then inequalities (9.7) and (9.8) can be rewritten as

$$I\left[\boldsymbol{\alpha} \in F^B_{\mathrm{UP}}\right] \geq I\left[\boldsymbol{\alpha} \in F^A_{\mathrm{UP}}\right]$$
$$I\left[\left[\boldsymbol{\alpha} \in F^B_{\mathrm{LP}}\right] \leq I\left[\boldsymbol{\alpha} \in F^A_{\mathrm{LP}}\right]\right.$$

for all $\boldsymbol{\alpha} \in (0,1]^d$. This implies $F^B_{\mathrm{LP}} \subseteq F^A_{\mathrm{LP}}$ and $F^B_{\mathrm{UP}} \supseteq F^A_{\mathrm{UP}}$. □

Part III

Summary of results, conclusions and open problems

Chapter 10

Final considerations, summary and future work

> I keep the subject of my inquiry constantly before me, and wait till the first dawning opens gradually, by little and little, into a full and clear light.
>
> Sir Isaac Newton

Evidence and random set theories provide the engineer with a tool that allows her/him to express her/his opinions intuitively and realistically for given partial information without making additional assumptions and in consequence is much more reliable in terms of the predicted uncertainty of the ocurrence of an event. The main aim for the risk analyst is to produce the best (tightest) possible intervals, that can be justified on the available information and well grounded assumptions. The bounds provided by random set theory are best possible in the sense that they cannot be any narrower and still contain all outcomes that agree with the given information and well founded assumptions. However, these bounds might be too wide to make useful decisions; this should, however, not be taken as an argument against random set theory. What it shows is the danger, even in simple problems, of making wishful assumptions in order to obtain a unique value of the probability of occurrence of the event of consideration.

In the following I will list the main contributions of this research:

Infinite random sets

The main contribution of this investigation is the extension of the method of finite random sets to infinite random sets in the framework of reliability and uncertainty analysis. This method is suitable for the calculation of the bounds of the probability of events when there is either epistemic or aleatory uncertainty in the definition of the basic variables. It allows to model, under a unified framework, the available information about the basic variables using probability boxes, possibility and probability distribution functions and Dempster-Shafer bodies of

evidence (families of intervals provided by experts). Since it takes in consideration all possible variation due to the uncertainty in the representation of the basic variables employed in the calculation of the probability of failure, it gives an interval as an answer, not a unique value of $P_{\boldsymbol{X}}(F)$. In addition techniques for sampling from a probability box, random set and possibility distributions were proposed; it was also shown that each of those cases is just a particularization of an infinite random set.

The proposed method introduces in addition, a new geometrical interpretation of the space of basic variables, called here the α-space. The lower and upper probability bounds are given by integrals (4.5) and (4.6), which are defined in that space. Since the evaluation of those integrals is analytically difficult or even impossible, a direct Monte Carlo sampling strategy was proposed to estimate such integrals. It was shown that those estimators are unbiased.

In contrast to Dempster-Shafer evidence theory, using an infinite number of focal elements we do not require anymore to perform discretizations of the basic variables. Using the proposed Monte Carlo approach, the precision of the results depend however, on the number of simulations performed, independent of the dimensionality of the problem.

Convergence considerations

A series of results regarding the convergence of finite random sets to infinite random sets was obtained; they follow the intuitive fact that if a finer discretization of the basic variables is performed, then more accurate results are obtained. The limit of these discretizations are infinite random sets.

A particular class of random sets was defined, namely random sets of indexable type. Probability boxes, possibility and probability distribution functions and Dempster-Shafer bodies of evidence were shown to be examples of this kind of random sets.

The results found were demonstrated for the special case that the refinements of random sets are by elimination. Further research must be carried out also in the case that the refinements are done by addition or by a combination of refinements by addition and elimination. Also, additional work is required in finding under which conditions $\mu_C\left(F_{\mathrm{LP}} \setminus \mathrm{supp}(\varphi^{\mathrm{Bel}})\right) = 0$ and $\mu_C\left(\mathrm{supp}(\varphi^{\mathrm{Pl}}) \setminus F_{\mathrm{UP}}\right) = 0$ in the fourth item of Theorem 5.6.3, when C has a singular component.

Dependence considerations

Some results relating infinite random sets and the theory of coherent upper and lower probabilities and upper and lower expectations were found. In fact, when random set theory is employed using only basic variables that are random, the proposed methodology is a particularization of the above mentioned theories.

It is left as an open problem to describe under which conditions the solution of the optimization problems (6.8) and (6.9) for $k = 1, 2, \ldots$ converge respectively to the solution of the optimization problems (6.10) and (6.11), as $k \to \infty$.

Discretizations and approximations of solutions

We characterized the different possible types of discretizations of basic variables, concluding that the outer discretization is the most conservative of the types of discretizations for a basic variable, in view of the fact that it does not entail loss of information. However, the results obtained with the outer discretization are too conservative, and therefore intermediate discretizations may be preferred as a way of discretizing the basic variables in Dempster-Shafer theory.

Also it was shown that the classical discretization of infinite random sets is equivalent to an approximation by Riemann-Stieltjes sums of the integrals (5.4) and (5.5). In addition, it was shown that more efficient algorithms like Monte Carlo methods exist for the evaluation of the lower and upper probability measures. This is a strong result, that would discourage in the future the use of the classical discretization of the basic variables employed for example in Demspter-Shafer evidence theory, and that suggest the use of infinite random sets of indexable type as a method for the faster evaluation of the lower and upper probability of events, specially in high dimensional problems. In low dimensions ($d \leq 4$), the discretization method can be efficient as well.

Monte Carlo simulation is the most basic algorithm for integration of functions in high dimensions. This is a current topic of research. We may benefit of any additional development in the field of high dimensional integration that outperforms Monte Carlo simulation. For example, it is known that variance reduction techniques, like stratified or importance sampling, or even adaptive Monte Carlo methods like the VEGAS algorithm, outperform simple Monte Carlo integration of Riemann integrals. Additional research is required on understanding how to effectively incorporate such methodology in the evaluation of the integrals (4.5) and (4.6). Using other Monte Carlo sampling methods, the computational cost required for estimating the lower and upper probability of the failure region could notably decrease.

It is left as an open problem to include, in the introduced methodology, fuzzy random variables. The reader is referred to Baudrit et al. (2008) for some ideas in this direction.

Nonspecificity

A Hartley-like based measure of nonspecificity for infinite random sets was proposed. This is a measure of lack of information, that is, a measure of epistemic uncertainty. In addition, some concepts related to random set theory like joint and marginal infinite random sets and random set inclusion were extended to

infinite random sets of indexable type. A mathematical derivation of the formulas for nonspecificity of possibility distributions and probability boxes was performed. It was also shown that the proposed measure corresponds to the one defined in the realm of Dempster-Shafer evidence theory when dealing with finite random sets. It was shown that the proposed measure is the only one that fulfills a set of characterizing properties.

Further work is required on extending the measures of aleatory uncertainty for infinite random sets, and in proving property NL6 (see Section 8.6). Note also that the treatment of random sets of indexable type is restricted to convex and compact focal elements. It should be studied whether it can be extended to more general types of focal elements.

Additional work is required to define a nonspecificity measure for sets of probability measures which are defined on some universal set $X \subseteq \mathbb{R}^d$. Since random sets of indexable are a particularization of sets of probability measures (see Appendix B) that measure must coincide with the one inhere defined when the particularization is done . Note that the definition of a nonspecificity measure when X is finite has already been analyzed by Abellán and Moral (2000) and further commented in Klir (2006).

Sensitivity analysis

We analyzed a relationship between sensitivity analysis and the task of reducing the width of the interval $[\mathrm{LP}(F), \mathrm{UP}(F)]$. The proposed method resulted to be a generalization of a method developed by Ferson and Tucker (2006b). The proposed technique can cope with the influence of scale dimension and shape of input PDFs, can evaluate the effects of a factor, while all others are varying as well and is able to treat grouped factors as if they were a single factor. Its most important drawback is that selecting a single representative pinching can be very misleading, especially if we want guaranteed results since these techniques erase uncertainty rather than propagate it; that is, this method depends strongly on the pinching employed. In this sense very well grounded assumptions must be made to justify the pinching strategy used, and in this way we can have some idea of the importance of those variables. The analysis reveals which basic variables should be the focus of a thoughtful investigation in order to reduce the epistemic uncertainty, which will be reflected in the width of the lower probability-upper probability interval.

In comparison to the method described in Ferson and Tucker, the proposed approach eliminates the intermediate step of converting the output body of evidence to a probability box in order to compute the indicator unc(), avoiding in this way unnecessary intermediate operations; also the proposed method is based in particular on the application of the Hartley-like measure, which is based on the theory of generalized information theory, therefore it appears to be a more natural measure of sensitivity to epistemic uncertainty. It was seen in the numerical example that it is sufficient to apply the extension principle for RSs to a small

number of focal elements in order to estimate the ranking of the basic variables; the proposed method allows the analyst to sample until some requirement in the variance of NL() or unc() is fulfilled, requiring in this way a smaller number of evaluations of the image of a focal element than in the method of Ferson and Tucker; in fact, in the example of Section 9.4, our method required 10 focal element simulations, while the method of Ferson and Tucker required 10000, to obtain the same ranking of the basic variables; more striking results are expected in higher dimensions.

Another technique was proposed for updating the lower and upper probability after obtaining new information that saves lots of the computational effort in computing again these numbers.

Further research is required on how to make the method robust for taking into account all possible pinchings, on possibility distributions, probability boxes and Dempster-Shafer bodies of evidence without increasing notably the computational overhead.

Other contributions

Another contribution of this thesis is the demonstration, in Section 3.6, of the fact that the set of all copulas $\mathscr{C}$ is compact under the metric d defined by equation (3.13).

10.1 Some other directions for future research

In addition to the open problems described before in this Chapter and also those listed in Sections 2.8, the following is a list of ideas that could be employed as a source of inspiration for new research.

The definition of infinite random sets requires the definition of suitable rules of combination of information, like for example, Dempster's rule of combination extended to infinite random sets.

It is required to develop techniques to measure the amount of uncertainty present when there is none or limited information about the dependence between the basic variables. This measure of uncertainty could then be applied to create a method of reduction of the width of the interval $[\mathrm{LP}(F), \mathrm{UP}(F)]$ that also takes into consideration the dependence information available.

There seems to exist an intrinsic relation between pinched random sets and random set inclusion for infinite random sets; this issue deserves further study.

Finally, Appendix B discusses some ideas that may be useful to understand random sets as probability measures. They may be the clue to extend the principles of strong and fuzzy set independence (see e.g. Fetz and Oberguggenberger (2004)) to infinite random sets.

Part IV

Appendix

Appendix A

Literature review

Since their appearance in the 1970s, evidence theory has been used in several areas of research like cartography, classification, decision making, failure diagnosis, robotics, signal processing and risk and reliability analysis. Sentz and Ferson (2002) provide a comprehensive review of the application of Dempster-Shafer evidence theory in these fields. In particular, this chapter will contain a literature review on the use of Dempster-Shafer evidence theory and finite random set theory, in the fields of risk, reliability and uncertainty analysis.

A.1 General remarks

In the last years, one of the leading research groups in the field of uncertainty analysis has been the Epistemic Uncertainty Project (http://www.sandia.gov/epistemic/) at Sandia National Laboratories (NM, USA). During 2003 and 2004, the group developed a sequence of challenge problems, containing varying mixtures of epistemic and aleatory uncertainty in the input parameters of a mechanical system. Researchers worldwide were asked to solve these problems, furnishing in this way a common framework for experimentation, discussion, and comparison of multiple approaches used in the assessment and propagation of uncertainty in real systems. The results of this benchmark were published in a special issue of Reliability Engineering & System Safety, edited by Helton and Oberkampf (2004). Some other outstanding contributions of the group are: *a*) Sentz and Ferson (2002) a comprehensive survey on a number of possible combination rules for Dempster-Shafer bodies of evidence; *b*) Ferson et al. (2003b) a complete review of probability boxes, and aggregation methods for handling agreement and conflict when multiple Dempster-Shafer bodies of evidence and probability boxes are obtained from different sources; *c*) Ferson et al. (2004) which reviews several of the most popular and dangerous myths

among risk analysts about dependence in probabilistic models, as well as summarizes methods to incorporate information (or lack of information) about intervariable dependence into risk assessments that use evidence theory or probability bounds analysis to address epistemic and aleatory uncertainty; this report also reviews techniques for simulating correlated variates for a given correlation specification and dependence model, computation of bounds on CDFs under a specified dependence model, formulation of parametric and empirical dependence models, and bounding approaches that can be used when information about the intervariable dependence is incomplete; *d*) Ferson et al. (2003a), which documents CONSTRUCTOR, a software that facilitates the collation of empirical information and expert judgment for the specification of CDFs, probability boxes, and Dempster-Shafer structures from data, qualitative shape information, constraints on moments, order statistics, densities, and coverage probabilities about uncertain unidimensional quantities, which may be real-valued, integer-valued or logical values; *e*) Ferson and Tucker (2006b) which deals with the problem of sensitivity analysis when the basic variables of the model are specified by quantities that can involve both aleatory and epistemic uncertainty and the method of calculation is Dempster-Shafer evidence theory or probability bounds analysis.

Among the papers contained in the special issue edited by Helton and Oberkampf (2004), the following papers are relevant Tonon (2004c), Hall and Lawry (2004) and Fetz and Oberguggenberger (2004).

Tonon (2004c) solves the first benchmark problem. He proposes two methods to discretize the probability boxes when the information given is composed by upper and lower envelope CDFs. He also remarks that Dempster's rule of combination can lead to a loss of information when the sources of information are not coherent. He used two grids: one fine and one rough. Tonon found that the calculation of those grids give almost the same results, except in the tails of the upper and lower CDFs of the response, where considerable differences where obtained. In this sense, he says that adaptive techniques to select the intervals are required because when a regular grid is used a poor estimation of the tails of the PDFs is done.

Hall and Lawry (2004) proposed a method for generating a random set with the corresponding belief and plausibility measures that are an outer approximation to coherent lower and upper probability distributions. This paper presents an alternative method to the discretization of probability boxes to the one proposed by Ferson et al. (2003b).

Fetz and Oberguggenberger (2004) deal with the specification of uncertainties, represented by either intervals, random sets, families of probability distributions, or fuzzy sets, and thereafter, they consistently reduce them to sets of probability measures, providing in this way a unifying interpretation for all these models and allowing the propagation through functional evaluation; thereafter, the paper deals with the question of dependence and independence of basic variables, represented by sets of probability measures.

Some other articles that deserve consideration in the area are:

Joslyn (2003) which present a useful method to construct random intervals by eliciting from experts in addition to the main interval $[a,b]$, information on how high and how low could be a and b themselves.

Joslyn and Ferson (2004) presented the relationship between random intervals, probability boxes and traces, which stands as the fuzzy set representation of a probability box or random interval.

Joslyn and Kreinovich (2005) prove that the estimates of belief and plausibility of the response of a system converge to the actual belief and plausibility bounds, and provides the requirements under which these convergence holds.

Bernardini and Tonon (2004) suggest a rule for the aggregation of different sources of information in the context of RS theory.

Tonon and Chen (2005) analyze the relation between random set inclusion and non-interactivity and stochastic independence. They show that inclusion is not necessarily preserved by non-interactivity and that it is preserved by stochastic independence.

Kriegler and Held (2003, 2005) apply DS theory in the analysis of future climate change. They show that the lower envelope of a set of probabilities bounded by CDFs is a belief function.

Finally, closely related to the theory of random sets is the theory of probability bounds analysis (PBA), proposed by Williamson (1989), Williamson and Downs (1990) and further developed by Ferson et al. (2003b). An important contribution in PBA is Berleant and Zhang (2004a,b); Zhang and Berleant (2005), where probability boxes are obtained given the marginal CDFs and partial dependence information among the random variables.

A.2 Use of random set theory in the uncertainty analysis in civil engineering

The following is intended to be, up to the author's best knowledge, a complete overview of the topic in the area of civil and mechanical engineering.

According to Tonon et al. (1998), the first application of random sets for structural engineering was presented by Bernardini and Tonon (1996), Tonon et al. (1996), Bernardini (1997) and Tonon and Bernardini (1998) following the ideas and derivations of Dubois and Prade (1991). Ever since, most of the works on this topic in the area of civil engineering have been carried out among others by Bernardini and Tonon, Bae and Grandhi, Fetz, Fellin and Oberguggenberger, Hall and Larry and Peschl and Schweiger and their coworkers. In the following, some of their works in the field are listed and some comments about the contributions and the new concept by them introduced is done.

In Tonon et al. (1998) it is shown how belief and plausibility concepts can be used to bound the probability of failure in the design of engineering structures and to mix the subjective opinion of experts. Tonon and Bernardini (1998) propose a reliability based optimization method of structures whose parameters are assigned as fuzzy numbers or fuzzy relations. Tonon et al. (2000a) propose a method to transform a general random set into a fuzzy set; this is advantageous in order to speed up the computations, at the expense of loss of information. Tonon et al. (2000b) propose and explore some procedures leading to the prediction of the behavior of rock engineering works when uncertainty affects basic parameters of the rock mass. The proposed procedure is founded on the theory of random sets. This paper also presents a technique to bracket the results of the estimation of the probability of failure based on Monte Carlo simulations. Tonon (2004a,b) use random sets to bound the results of Monte Carlo simulation of a traditional reliability analysis of structural systems, in the sense of Ditlevsen and Madsen (1996). This approach can be useful because a bound can be enough to make a decisions. In case the bound is required to be tighter (even locally, when the whole CDFs are calculated), the method can successively refine it to minimize the degree of uncertainty. Also, given the bounds on the CDF of the system response y, Tonon proposes an approximation of the CDF of y within those bounds. This method can also deal with cases in which the input information is expressed as an interval, a fuzzy or random set, as well as any combination of them. However this algorithm is extremely sensitive to the dimension of the problem, since the required number of calls of the system solver are $(n+1)^d$ where n is the number of discretization intervals per variable and d is the number of random variables. Tonon et al. (2004) present a classical application of the theory of random sets to analyze the reliability bounds of an aircraft wing, subject to situations of little information about the uncertainties in loads and materials. Tonon and Pettit (2004, 2005) propose a method to take into consideration the dependence information inside the joint focal elements. This method, in particular, assumes that there is a correlation index between the unidimensional marginal focal elements. Bernardini (2005) analyses the vulnerability and the expected seismic damage to residential buildings using random sets. Other contributions by Tonon, Bernardini and coworkers in the topic are: Tonon and Bernardini (1999) and Tonon et al. (2001).

Bae et al. (2003, 2004a,b) estimated a bound on the probability of failure of a fighter aircraft wing subject to static loads and modelled through the finite element method. This bound is calculated as $[\mathrm{Bel}(F), \mathrm{Pl}(F)]$, where F is the failure region. They employed a quadratic response surface to mitigate the computational effort of the optimization method. They called their approach the *cost-effective method*. Also, Bae et al. (2004b) solved the reliability analysis problem of surpassing of the maximum allowed displacement of the one node of a truss; using incomplete information shaped in a random set, they employed possibility, probability and evidence theory to estimate the reliability of the truss. In the first case, the input random set was transformed to a consonant one using a method proposed by Tonon et al. (2000a) and in consequence interpreted as a fuzzy set. Using Zadeh's extension principle, the fuzzy response (displacement) of the node

was estimated and then the possibility of failure. The possibility approach gave the most conservative result. In addition, the authors observed that the inclusion techniques lead to the loss of the physical meaning of the information. In the probabilistic case, since the PDF of the uncertain variables was unknown, the uniform distribution was used in accordance with Laplace's principle of insufficient reason, i.e. it was assumed that each interval had a uniform PDF. Monte Carlo simulation was employed to calculate the probability of failure, which was obtained as a single number, leading to a possible underestimation of it, since this measure is extremely sensitive to the changes of the PDF. Bae et al. (2006a) propose an approximation of the plausibility function (called by them the plausibility decision function) to enable its continuous modeling. This approximation is used as an upper bound of the probability of failure of the system, which is employed as a constraint in the optimization problem. Also, to speed up the calculations, the cost-effective method is employed. Another publication of Bae and coworkers in the topic is Bae et al. (2006b).

Oberguggenberger and Fellin (2002, 2004) showed that the probability of failure of a shallow foundation may fluctuate even by orders of magnitude (they were seen to range between 10^{-11} and 10^{-3}) when different likely PDFs associated to the soil parameters, and estimated using a relatively large number of soil tests, were considered. Using the evidence theory approach, no additional assumptions were made. The interval of failure obtained contained that one calculated through the probabilistic approach, as expected. Oberguggenberger and Fellin suggest a method for modeling probability boxes from a set of data. The method is based on the use of Tchebycheff's inequality in the finite (Oberguggenberger and Fellin (2005)) and in the infinite random set case (Oberguggenberger and Fellin (2006)).

Rubio et al. (2004) analyzed the stability of an earth slope. They used random set theory to combine different sources of information, like intervals, sets of intervals, PDFs and upper and lower bounds on CDFs which expresses the knowledge about the different parameters of the soil. In Hall et al. (2004), they analyzed the same problem, using random sets of probability measures.

Schweiger and Peschl (2004), Peschl and Schweiger (2004), Peschl (2004) introduced a general framework for dealing with uncertainties in the design and construction process of geotechnical engineering using finite random sets to describe the variability of material parameters and geometrical data.

Appendix B

Some ideas on how to understand random sets as sets of probability measures

In this Appendix, we discuss some ideas that deserve further research.

A finite random set can be understood a set of probability measures that results as the weighted addition of the sets of probability measures that are defined in each of the focal elements of $\mathscr{F}$. A proof of this result in the finite random set case can be found in Fetz and Oberguggenberger (2004). In the following it is shown an idea of how to extend this result to the infinite case.

Following Fetz and Oberguggenberger (2004), let $X \subseteq \mathbb{R}^d$ be our universal space, and let (X, σ_X) be a measure space, where $\sigma_X = \mathscr{B}^d \cap X$. Let $\mathscr{P}$ be the set of all probability measures that can be defined in (X, σ_X). In particular, $\mathscr{P}$ contains all δ-Dirac measures on X. Let F be an element of σ_X.

Sets

The probability that an element x of a set $\gamma \in \sigma_X$ belongs to F can be, on the one hand, bounded from below and above respectively by

$$\underline{P}(x \in F) = I\left[\gamma \subseteq F\right]$$

and

$$\overline{P}(x \in F) = I\left[\gamma \cap F \neq \emptyset\right]$$

On the other hand, if we consider some set P_γ of probability measures defined on γ, the corresponding lower and upper probabilities are defined by P_γ are

$$\underline{\underline{P}}(x \in F) = \inf_{P \in \mathscr{P}_\gamma} P(\gamma \cap F)$$

and

$$\overline{\overline{P}}(x \in F) = \sup_{P \in \mathscr{P}_\gamma} P(\gamma \cap F)$$

Note that

$$\underline{P}(x \in F) \leq \underline{\underline{P}}(x \in F) \tag{B.1}$$

and

$$\overline{P}(x \in F) \geq \overline{\overline{P}}(x \in F). \tag{B.2}$$

The equality in (B.1) and (B.2) holds when P_γ is the set of all probability measures defined on γ, i.e., $P_\gamma := \{P : P(\gamma) = 1, P \in \mathscr{P}\}$ (in particular, this set contains all δ-Dirac measures on γ). In fact, if $\gamma \subseteq F$ then all $P \in P_\gamma$ assign 1 to the set $\gamma \cap F$; on the contrary, if $\gamma \not\subseteq F$, making P to be any δ-Dirac measure concentrated on some point in $\gamma \setminus F$, will assign a null probability to the set $\gamma \cap F$. Similar considerations can be made to show (B.2).

Random sets

Let $(\mathscr{F}, P_\Gamma)$ be a random set defined on X; here $\gamma \in \mathscr{F}$ is such that $\gamma \in \sigma_X$. The probability that an element x that belongs to the support of the focal set $\mathscr{F}$, i.e., $x \in \cup_{\gamma \in \mathscr{F}} \gamma$ belongs to $F \in \sigma_X$, is bounded by below and above by equations (4.1) and (4.3), which are rewritten here by convenience:

$$\mathrm{LP}_{(\mathscr{F}, P_\Gamma)}(F) = \int_{\mathscr{F}} I\left[\gamma \subseteq F\right] \mathrm{d}P_\Gamma(\gamma), \tag{B.3}$$

$$\mathrm{UP}_{(\mathscr{F}, P_\Gamma)}(F) = \int_{\mathscr{F}} I\left[\gamma \cap F \neq \emptyset\right] \mathrm{d}P_\Gamma(\gamma) \tag{B.4}$$

provided $I\left[\gamma \subseteq F\right]$ and $I\left[\gamma \cap F \neq \emptyset\right]$ are P_Γ-measurable functions of γ.

If we consider that for every focal element γ, there exists some set P_γ of probability measures defined on γ, then corresponding lower and upper probabilities are defined by those P_γ's are

$$\underline{\underline{P}}(x \in F) = \int_{\mathscr{F}} \inf_{P \in \mathscr{P}_\gamma} P(\gamma \cap F) \, \mathrm{d}P_\Gamma(\gamma), \tag{B.5}$$

$$\overline{\overline{P}}(x \in F) = \int_{\mathscr{F}} \sup_{P \in \mathscr{P}_\gamma} P(\gamma \cap F) \, \mathrm{d}P_\Gamma(\gamma) \tag{B.6}$$

The measurability of the integrands must be required here. Using relations (B.1) and (B.2) in (B.3) and (B.4) respectively, we have

$$\mathrm{LP}_{(\mathscr{F}, P_\Gamma)}(F) \leq \underline{\underline{P}}(x \in F) \tag{B.7}$$

and

$$\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F) \geq \overline{\overline{P}}(x \in F). \tag{B.8}$$

Again, the equality in (B.7) and (B.8) holds when for each focal elements γ, we have that P_γ is the set of all probability measures defined on γ, i.e., $P_\gamma := \{P : P(\gamma) = 1, P \in \mathscr{P}\}$.

Now, let us interpret $(\mathscr{F}, P_\Gamma)$ as the weighted set of probability measures $\mathscr{P}_\mathscr{F}$ with elements of the form

$$p = \int_\mathscr{F} p_\gamma \, \mathrm{d}P_\Gamma(\gamma)$$

(considering again the question of the measurability of the integrand) where $p_\gamma \in P_\gamma$. It appears an open question to show the relationship between the lower and upper probabilities

$$\underline{P}(x \in F) = \inf_{p \in \mathscr{P}_\mathscr{F}} p(F)$$

and

$$\overline{P}(x \in F) = \sup_{p \in \mathscr{P}_\mathscr{F}} p(F)$$

and equations to (B.5) and (B.6).

Copulas and t-norms

When we are dealing with random sets, we already know that copulas are the functions that couple the marginal basic variables. For example, if $X_1, X_2, \ldots, X_n$ are random variables with CDFs $F_{X_1}, F_{X_2}, \ldots, F_{X_n}$, the joint CDF $F_{X_1,X_2,\ldots,X_n}$ that relates those random variables is defined using the copula C by

$$F_{X_1,X_2,\ldots,X_n} = C\left(F_{X_1}, F_{X_2}, \ldots, F_{X_n}\right)$$

Now, it is a well know fact from possibility theory that when $A_1, A_2, \ldots, A_n$ are possibility distributions, the joint possibility distribution $A_{1,2,\ldots,n}$ that relates those possibility distributions is defined using the t-norm T by,

$$A_{1,2,\ldots,n} = T\left(A_1, A_2, \ldots, A_n\right)$$

Some t-norms are copulas and many copulas are special cases of t-norms; Klement et al. (2000) discuss this issue. Therefore, we have to study the relations between these functions and how they are related with our problem. There may be for example possibility distributions that cannot be modelled as a joint random set, inasmuch as the t-norm that joins the marginals is not a copula; the natural question that arises is which other information is required in order to model that

joint possibility distribution as a random set. Maybe we require to model in addition which are the probability measures residing in each of the focal elements, i.e., the sets P_γ.

Some other fact that deserves attention is that the t-norm T is the function in charge of defining how is the shape of the α-cuts of the joint possibility distribution $A_{1,2,\ldots,n}$. Figure B.1, shows an example where the same marginal possibility distributions generate different joint possibility distributions. How is this information modelled in random set theory? This has obviously some relation with the sets of probability measures P_γ; in other words, the t-norms seem to model not only the dependence relation between but also and within the focal elements. This open question may be the link to model for example joint probability boxes, and to define strong and fuzzy independence.

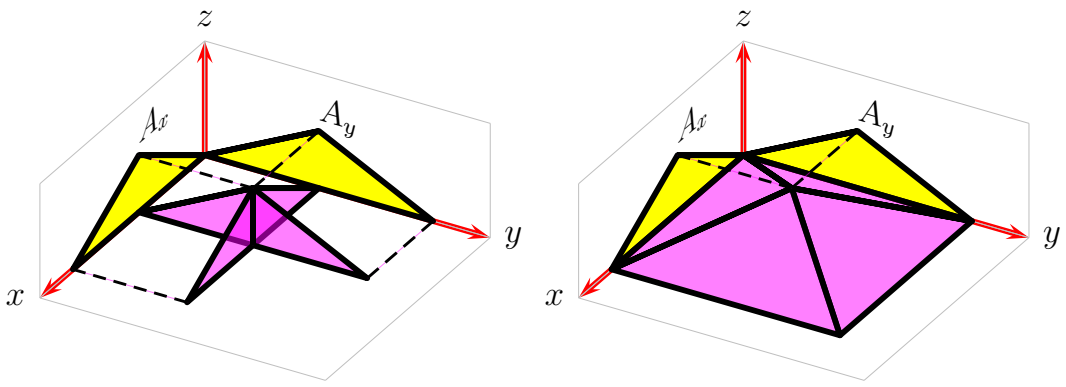


Figure B.1: Joint possibility distributions with the same marginal distributions.

Bibliography

Abellán, J. and Moral, S. (2000). A non-specificity measure for convex sets of probability distributions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 8(3):357–367. See also Corrigendum in Volume 14(3) of the same Journal, page 467, 2005. (Cited on page 132)

Anderson, E. J. and Nash, P. (1987). *Linear Programming in Infinite Dimensional Spaces: theory and applications*. Series in Discrete Mathematics and Optimization. Wiley-Interscience. (Cited on page 78)

Bae, H.-R., Grandhi, R. V., and Canfield, R. A. (2003). Uncertainty quantification of structural response using evidence theory. *AIAA Journal*, 41(10):2062–2068. (Cited on pages 2, 3, 56, and 140)

Bae, H.-R., Grandhi, R. V., and Canfield, R. A. (2004a). An approximation approach for uncertainty quantification using evidence theory. *Reliability Engineering and System Safety*, 86(3):215–225. (Cited on pages 17, 56, and 140)

Bae, H.-R., Grandhi, R. V., and Canfield, R. A. (2004b). Epistemic uncertainty quantification techniques including evidence theory for large-scale structures. *Computers and Structures*, 82:1101–1112. (Cited on pages 56 and 140)

Bae, H.-R., Grandhi, R. V., and Canfield, R. A. (2006a). Reliability based optimization of engineering structures under imprecise information. *International Journal of Materials and Product Technology*, 25(1,2,3):112–126. (Cited on page 141)

Bae, H.-R., Grandhi, R. V., and Canfield, R. A. (2006b). Sensitivity analysis of structural response uncertainty propagation using evidence theory. *Journal of Structural and Multidisciplinary Optimization*, 31(4):270–279. (Cited on pages 114 and 141)

Baudrit, C., Dubois, D., and Guyonnet, D. (2006). Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment. *IEEE Transactions on Fuzzy Systems*, 14(5):593–608. (Cited on page 84)

Baudrit, C., Dubois, D., and Perrot, N. (2008). Representing parametric probabilistic models tainted with imprecision. *Fuzzy Sets and Systems*, 159(15):1913–1928. (Not cited)

Bellman, R. (1961). *Adaptive Control Processes: A Guided Tour*. Princeton University Press, Princeton, NJ. (Cited on page 16)

Berleant, D. and Cheng, H. (1998). A software tool for automatically verified operations on intervals and probability distributions. *Reliable Computing*, 4:71–82. (Cited on page 20)

Berleant, D. and Goodman-Strauss, C. (1998). Bounding the results of arithmetic operations on random variables of unknown dependency using intervals. *Reliable Computing*, 4:147–165. (Cited on page 20)

Berleant, D. and Zhang, J. (2004a). Representation and problem solving with distribution envelope determination (DEnv). *Reliability Engineering and System Safety*, 1–3:153–168. Available at http://class.ee.iastate.edu/berleant/home/me/cv/papers/challengeProblems.pdf. (Cited on pages 20 and 139)

Berleant, D. and Zhang, J. (2004b). Using Pearson correlation to improve envelopes around the distributions of functions. *Reliable Computing*, 10:139–161. Available at http://class.ee.iastate.edu/berleant/home/me/cv/papers/correlationpaper.pdf. (Cited on pages 20 and 139)

Bernardini, A. (1997). A fuzzy set approach to the response evaluation of uncertain mechanical systems. In Soffara, D. H., editor, *Proceedings of the Third International Conference on Stochastic Structural Dynamics*, volume 12, pages 113–147, San Juan, Puerto Rico. (Cited on page 139)

Bernardini, A. (2005). Random sets of the expected seismic damage to building types. In Augusti, G., Schuëller, G. I., and Ciampoli, M., editors, *Proceedings of the Ninth International Conference on Structural Safety and Reliability ICOSSAR'05*, pages 1733–1740, June 19–23, 2005, Rome. Millpress, Rotterdam. (Cited on page 140)

Bernardini, A. and Tonon, F. (1996). A combined fuzzy and random set approach to the multiobjective optimization of uncertain systems. In Frangopol, D. M. and Grigoriu, M. D., editors, *Proceedings of the Seventh Speciality ASCE Conference on Advances in Probabilistic Mechanics and Structural Reliability*, pages 314–317, August 7–9, 1996, Worcester, Massachusetts. ASCE, New York. (Cited on page 139)

Bernardini, A. and Tonon, F. (2004). Aggregation of evidence from random and fuzzy sets. *ZAMM (Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik)*, 84:700–709. (Cited on page 139)

Bertoluzza, C., Gil, M. A., and Ralescu, D. A., editors (2002). *Statistical modeling, analysis and management of fuzzy data*, volume 87 of *Studies in fuzziness and soft computing*. Physica Verlag, Heidelberg; New York. (Cited on page 12)

Blockley, D. I. (1980). *The nature of structural design and safety*. Ellis Horwood, Chichester. (Cited on page 1)

Blockley, D. I. (1999). Risk based structural reliability methods in context. *Structural safety*, 21:335–348. (Cited on page 1)

Couso, I., Moral, S., and Walley, P. (1999). Examples of independence for imprecise probabilities. In De Cooman, G., Cozman, F. G., Moral, S., and Walley, P., editors, *Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications, ISIPTA'99*, pages 121–130. Available at http://decsai.ugr.es/~smc/isipta99/proc/068.html. (Cited on pages 18, 76, and 101)

De Cooman, G., Troffaes, M., and Miranda, E. (2005). n-monotone lower previsions and lower integrals. In *Proceedings of the Fourth International Symposium on Imprecise Probabilities and Their Applications, ISIPTA'05*, Pittsburgh, Pennsylvania. Available at http://www.sipta.org/isipta05/proceedings/017.html. (Cited on page 48)

Delgado, M. and Moral, S. (1987). A definition of inclusion for evidences. *Fuzzy Mathematics (Huazhong, China)*, 6(4):81–87. (Cited on pages 17, 102, and 104)

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339. (Cited on pages 3, 13, 26, and 27)

Dennenberg, D. (1994). *Non-additive Measure and Integral*. Kluwer Academic Publishers, Dordrecht, Boston. (Cited on page 11)

Ditlevsen, O. and Madsen, H. O. (1996). *Structural Reliability Methods*. John Wiley and Sons, New York. 384 p. (Cited on pages 1 and 140)

Dodagoudar, G. R. and Venkatachalam, G. (2000). Reliability analysis of slopes using fuzzy set theory. *Computers and Geotechnics*, 27(2):101–115. (Cited on page 27)

Dong, W. and Shah, H. (1987). Vertex method for computing functions on fuzzy variables. *Fuzzy Sets and Systems*, 24:65–78. (Cited on page 16)

Dubois, D. and Prade, H. (1985). A note on measures of specificity for fuzzy sets. *International Journal of General Systems*, 10(4):279–283. (Cited on page 105)

Dubois, D. and Prade, H. (1987). The principle of minimum specificity as a basis for evidential reasoning. In Bouchon, B. and Yager, R., editors, *Uncertainty in Knowledge-Based Systems*, volume 286 of *LNCS*, pages 75–84. Springer Verlag. (Cited on page 105)

Dubois, D. and Prade, H. (1988). *Possibility Theory*. Plenum Press, New York. (Cited on page 21)

Dubois, D. and Prade, H. (1991). Random sets and fuzzy interval analysis. *Fuzzy Sets and Systems*, 42(1):87–101. (Cited on pages 9, 15, 17, 18, 22, 102, 104, 117, and 139)

Embrechts, P., McNeil, A., and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In Dempster, M., editor, *Risk Management: Value at Risk and Beyond*, pages 176–223. Cambridge University Press, Cambridge, UK. (Cited on page 30)

Ferson, S. and Hajagos, J. (2002). Don't open that envelope: solutions to the Sandia problems using probability boxes. Poster presented at Sandia National Laboratory. Epistemic Uncertainty Workshop. Albuquerque. August 6–7. Webpage http://www.sandia.gov/epistemic/eup_workshop1.htm. Available at http://www.sandia.gov/epistemic/Papers/ferson.pdf. (Cited on page 22)

Ferson, S., Hajagos, J., Myers, D. S., and Tucker, W. T. (2003a). Constructor: synthesizing information about uncertainty variables. Report SAND2005-3769, Sandia National Laboratories, Albuquerque, NM. Available at http://www.ramas.com/construct.pdf. (Cited on page 138)

Ferson, S., Hajagos, J., and Tucker, W. T. (2005). Probability bounds analysis is a global sensitivity analysis. In Hanson and Hemez (2005), pages 211–220. Available at http://library.lanl.gov/ccw/samo2004. (Cited on pages 113 and 115)

Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D. S., and Sentz, K. (2003b). Constructing probability boxes and Dempster-Shafer structures. Report SAND2002-4015, Sandia National Laboratories, Albuquerque, NM. Available at http://www.ramas.com/unabridged.zip. (Cited on pages 22, 44, 45, 52, 92, 115, 137, 138, and 139)

Ferson, S., Nelsen, R. B., Hajagos, J., Berleant, D. J., Zhang, J., Tucker, W. T., Ginzburg, L. R., and Oberkampf, W. L. (2004). Dependence in probabilistic modelling, Dempster-Shafer theory and probability bounds analysis. Report SAND2004-3072, Sandia National Laboratories, Albuquerque, NM. Available at http://www.ramas.com/depend.zip. (Cited on pages 63 and 137)

Ferson, S. and Tucker, W. T. (2006a). Sensitivity analysis using probability bounding. *Reliability Engineering and System Safety*, 91(10–11):1435–1442. (Cited on page 115)

Ferson, S. and Tucker, W. T. (2006b). Sensitivity in risk analysis with uncertain numbers. Report SAND2006-2801, Sandia National Laboratories, Albuquerque, NM. Available at http://www.ramas.com/sensanal.pdf. (Cited on pages 113, 114, 115, 116, 118, 120, 121, 122, 123, 132, and 138)

Fetz, T. and Oberguggenberger, M. (2004). Propagation of uncertainty through multivariate functions in the framework of sets of probability measures. *Reliability Engineering and System Safety*, 85(1–3):73–87. (Cited on pages 15, 18, 19, 20, 76, 101, 133, 138, and 143)

Fishman, G. S. (1996). *Monte Carlo - Concepts, Algorithms and Applications*. Springer-Verlag, New York. (Cited on page 88)

Goodman, I. R. and Nguyen, H. T. (2002). Fuzziness and randomness. In Bertoluzza, C., Gil, M. A., and Ralescu, D. A., editors, *Statistical modeling, analysis and management of fuzzy data*, volume 87 of *Studies in fuzziness and soft computing*, pages 3–21. Physica Verlag, Heidelberg; New York. (Cited on pages 14 and 21)

Hall, J. and Lawry, J. (2001). Imprecise probabilities of engineering system failure from random and fuzzy set reliability analysis. In De Cooman, G., Fine, T., and Seidenfeld, T., editors, *Proceedings of the Second International Symposium on Imprecise Probabilities and their Applications, ISIPTA'01*, pages 195–204, Maastricht. Cornell University, June 27-29, 2001. (Cited on page 122)

Hall, J. W. (2005). Sensitivity indices for imprecise probability distributions. In Hanson and Hemez (2005), pages 54–60. Available at http://library.lanl.gov/ccw/samo2004. (Cited on page 114)

Hall, J. W. (2006). Uncertainty-based sensitivity indices for imprecise probability distributions. *Reliability Engineering and System Safety*, 91(10–11):1443–1451. (Cited on page 114)

Hall, J. W. and Lawry, J. (2003). Fuzzy label methods for constructing imprecise limit state functions. *Structural Safety*, 25(4):317–341. (Cited on page 27)

Hall, J. W. and Lawry, J. (2004). Generation, combination and extension of random set approximations to coherent lower and upper probabilities. *Reliability Engineering and System Safety*, 85(1–3):89–101. (Cited on pages 22, 67, 81, and 138)

Hall, J. W., Rubio, E., and Anderson, M. G. (2004). Random sets of probability measures in slope hydrology and stability analysis. *ZAMM (Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik)*, 84(10–11):710–720. (Cited on page 141)

Hanson, K. M. and Hemez, F. M., editors (2005). Santa Fe, New Mexico. Los Alamos National Laboratory. Available at http://library.lanl.gov/ccw/samo2004. (Cited on pages 150, 151, and 152)

Hartley, R. V. L. (1928). Transmission of information. *The Bell System Technical Journal*, 7(1):535–563. (Cited on pages 97 and 98)

Helton, J. C. (1997). Uncertainty and sensitivity analysis in the presence of stochastic and subjective uncertainty. *Journal of Statistical Computation and Simulation*, 57:3–76. (Cited on page 2)

Helton, J. C. and Burmaster, D. E., editors (1996). *Treatment of Aleatory and Epistemic Uncertainty in Performance Assessments for Complex Systems*, volume 54. Reliability Engineering and System Safety. (Cited on page 2)

Helton, J. C., Johnson, J. D., and Oberkampf, W. L. (2004). Sensitivity analysis in conjunction with evidence theory representations of epistemic uncertainty.

In Hanson, K. M. and Hemez, F. M., editors, *Proceedings of the 4th International Conference on Sensitivity Analysis of Model Output (SAMO 2004)*, pages 410–420, Santa Fe, New Mexico. Los Alamos National Laboratory, The Research Library. Available at http://library.lanl.gov/ccw/samo2004/. (Cited on page 9)

Helton, J. C., Johnson, J. D., and Oberkampf, W. L. (2005). Sensitivity analysis in conjunction with evidence theory representations of epistemic uncertainty. In Hanson and Hemez (2005), pages 410–420. Available at http://library.lanl.gov/ccw/samo2004. (Cited on page 114)

Helton, J. C. and Oberkampf, W. (2004). Special volume on alternative representations of epistemic uncertainty. *Reliability Engineering and System Safety*, 85(1–3). (July-Sept. 2004), pp. 1-369. (Cited on pages 137 and 138)

Himmelberg, C. J. (1975). Measurable relations. *Fundamenta Mathematicae*, 87:53–72. (Cited on page 48)

Hoeffding, W. (1940). Masstabinvariante korrelationstheorie. *Schriften des Matematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin*, 5, Heft 3:179–Ű233. (Cited on page 32)

Hurtado, J. E. and Alvarez, D. A. (2001). Neural-network-based reliability analysis: a comparative study. *Computer Methods in Applied Mechanics and Engineering*, 191(1–2):113–132. (Cited on page 17)

Joslyn, C. (2003). Multi-interval elicitation of random intervals for engineering reliability analysis. In *Proceedings of the 4th International Symposium on Uncertainty Modelling and Analysis, ISUMA'03*, page 168, Washington, DC, USA. IEEE Computer Society. (Cited on page 139)

Joslyn, C. and Booker, J. M. (2004). Generalized information theory for engineering modeling and simulation. In Nikolaidis, E. and Ghiocel, D., editors, *Engineering Design Reliability Handbook*, pages 9:1–40. CRC Press. (Cited on pages 3 and 9)

Joslyn, C. and Ferson, S. (2004). Approximate representations of random intervals for hybrid uncertain quantification in engineering modeling. In Hanson, K. M. and Hemez, F. M., editors, *Proceedings of the 4th International Conference on Sensitivity Analysis of Model Output (SAMO 2004)*, pages 453–469, Santa Fe, New Mexico. Los Alamos National Laboratory, The Research Library. Available at http://library.lanl.gov/ccw/samo2004/. (Cited on pages 22, 23, 44, and 139)

Joslyn, C. and Kreinovich, V. (2005). Convergence properties of an interval probabilistic approach to system reliability estimation. *International Journal of General Systems*, 34(4):465–482. (Cited on pages 16 and 139)

Kendall, D. G. (1974). Foundations of a theory of random sets. In Harding, E. F. and Kendall, D. G., editors, *Stochastic geometry*, pages 322–376, London. Wiley. (Cited on page 3)

Klement, E. P., Mesiar, R., and Pap, E. (2000). *Triangular norms*. Kluwer. (Cited on page 145)

Klir, G. J. (1989). Is there more to uncertainty than some probability theorists might have us believe? *Internaltional Journal of General Systems*, 15:347–378. (Cited on page 10)

Klir, G. J. (1995). Principles of uncertainty: what are they? why do we need them? *Fuzzy Sets and Systems*, 74:15–31. (Cited on page 9)

Klir, G. J. (2006). *Uncertainty and Information : Foundations of Generalized Information Theory*. John Wiley and Sons, New Jersey. (Cited on pages 14, 21, 98, 100, 103, 105, 106, 107, 110, and 132)

Klir, G. J. and Folger, T. A. (1988). *Fuzzy Sets, Uncertainty and Information*. Prentice-Hall, Englewood Cliffs, New Jersey. (Cited on page 21)

Klir, G. J. and Wierman, M. J. (1998). *Uncertainty-Based Information: Elements of Generalized Information Theory*, volume 15 of *Studies in Fuzziness and Soft Computing)*. Physica-Verlag, Heidelberg, Germany. (Cited on pages 2, 9, 105, and 121)

Klir, G. J. and Yuan, B. (1995). On nonspecificity of fuzzy sets with continuous membership functions. In *Proc. 1995 Intern. Conf. on Systems, Man, and Cybernetics*, pages 25–29, Vancouver. (Cited on pages 97 and 99)

Kohlas, J. and PA, P. M. (1995). *A mathematical theory of hints. An approach to the Dempster-Shafer theory of evidence*, volume 425 of *Lecture notes in economics and mathematical systems*. Springer-Verlag, Berlin. (Cited on page 4)

Kolmogorov, A. N. and Fomin, S. V. (1970). *Introductory Real Analysis*. Dover Publications, New York. (Cited on pages 36 and 64)

Kriegler, E. and Held, H. (2003). Climate projections for the 21st century using random sets. In Bernard, J. M., Seidenfeld, T., and Zaffalon, M., editors, *Proceedings of the Third International Symposium on Imprecise Probabilities and their Applications, ISIPTA'03*, Lugano, Switzerland; July 14-17 2003. Available at http://www.carleton-scientific.com/isipta/PDF/026.pdf. (Cited on page 139)

Kriegler, E. and Held, H. (2005). Utilizing belief functions for the estimation of future climate change. *International Journal of Approximate Reasoning*, 39:185–209. (Cited on page 139)

Krommer, A. R. and Ueberhuber, C. W. (1998). *Computational integration*. SIAM, Philadelphia. (Cited on page 87)

Kyburg, H. E. (1987). Bayesian and non-bayesian evidential updating. *Artificial Intelligence*, 31:271–293. (Cited on page 26)

Loève, M. (1977). *Probability theory I*. Springer Verlag, Berlin, 4th edition. (Cited on pages 50 and 64)

Matheron, G. (1975). *Random sets and integral geometry*. Wiley, New York. (Cited on page 3)

McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative risk management: concepts, techniques and tools*. Princeton Series in Finance. Princeton University Press, Princeton, NJ. (Cited on page 93)

Melchers, R. E. (1999). *Structural Reliability Analysis and Prediction*. John Wiley and Sons, New York, 2nd edition. 456 p. (Cited on page 1)

Miranda, E., Couso, I., and Gil, P. (2005a). Random intervals as a model for imprecise information. *Fuzzy sets and systems*, 154(3):386–412. (Cited on pages 14 and 65)

Miranda, E., Couso, I., and Gil, P. (2005b). Random sets as imprecise random variables. *Journal of Mathematical Analysis and Applications*, 307(1):32–47. (Cited on page 14)

Möller, B. and Beer, M. (2004). *Fuzzy Randomness - Uncertainty in Civil Engineering and Computational Mechanics*. Springer, Berlin Heidelberg New York. (Cited on pages 4 and 26)

Nelsen, R. B. (1999). *An Introduction to Copulas*, volume 139 of *Lectures Notes in Statistics*. Springer Verlag, New York. (Cited on pages 29, 32, 34, 35, 49, 73, and 88)

Nguyen, H. (1978). On random sets and belief functions. *Journal of Mathematical Analysis and Applications*. (Cited on page 48)

Nguyen, H. T. (2006). *An introduction to random sets*. Chapman and Hall CRC, Boca Raton, FL. (Cited on page 68)

Nguyen, H. T. and Walker, E. A. (1996). *A first course in fuzzy logic*. CRC Press, Boca Raton. (Cited on page 21)

Niederreiter, H. (1992). *Random number generation and quasi-Monte Carlo methods*. Number 63 in CBMS-NSF regional conference series in applied mathematics. Society for Industrial and Applied Mathematics, Philadelphia, Pensilvania. (Cited on pages 87 and 89)

Oberguggenberger, M. and Fellin, W. (2002). From probability to fuzzy sets: the struggle for meaning in geotechnical risk assessment. In Pötter, R., Klapperich, H., and Schweiger, H. F., editors, *Probabilistics in geotechnics: technical and economic risk estimation*, pages 29–38, Essen. Verlag Glückauf GmbH. (Cited on pages 2 and 141)

Oberguggenberger, M. and Fellin, W. (2004). The fuzziness and sensitivity of failure probabilities. In Fellin, W., Lessmann, H., Oberguggenberger, M., and Vieider, R., editors, *Analyzing Uncertainty in Civil Engineering*, pages 33–48, Berlin. Springer-Verlag. (Cited on pages 2 and 141)

Oberguggenberger, M. and Fellin, W. (2005). Assessing the sensitivity of failure probabilities: a random set approach. In Augusti, G., Schuëller, G. I., and Ciampoli, M., editors, *Proceedings of the Ninth International Conference on Structural Safety and Reliability ICOSSAR'05*, pages 1733–1740, June 19–23, 2005, Rome. Millpress, Rotterdam. (Cited on pages 23 and 141)

Oberguggenberger, M. and Fellin, W. (2006). Reliability bounds through random sets: nonparametric methods and geotechnical applications. *Leopold-Franzens Universität Innsbruck*. In preparation. (Cited on pages 23 and 141)

Oberkampf, W. L. and Helton, J. C. (2002). Investigation of evidence theory for engineering applications. In *Proceedings of the fourth AIAA Non-Deterministic Approaches Forum*, number Paper No. AIAA-2002-1569, Denver, Colorado. American Institute of Aeronautics and Astronautics. (Cited on page 114)

Oberkampf, W. L., Helton, J. C., Joslyn, C. A., Wojtkiewicz, S. F., and Ferson, S. (2004). Challenge problems: Uncertainty in system response given uncertain parameters. *Reliability Engineering and System Safety*, 85(1–3):11–20. (Cited on page 51)

Oberkampf, W. L., Helton, J. C., and Sentz, K. (2001). Mathematical representation of uncertainty. In *Non-deterministic approached Forum 2001*, AIAA-2001-1645, Seattle, WA. American Institute of Aeronautics and Astronautics. April 16-19. (Cited on pages 5 and 9)

Pan, Y. and Klir, G. J. (1997). Bayesian inference based on interval-valued prior distributions and likelihoods. *Journal of Intelligent and Fuzzy Systems*, 5(3):193–203. (Cited on page 26)

Papadrakakis, M., Papadopoulos, V., and Lagaros, N. D. (1996). Structural reliability analysis of elastic-plastic structures using neural networks and Monte Carlo simulation. *Computer Methods in Applied Mechanics and Engineering*, 136:145–163. (Cited on page 17)

Papoulis, A. and Pillai, S. U. (2002). *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, New York, 4th edition. (Cited on pages 3 and 5)

Peschl, G. and Schweiger, H. (2004). Application of the random set finite element method (RS-FEM) in geotechnics. In Pande, G. and Pietruszczak, S., editors, *Proceedings of the 9th International Symposium on Numerical Models in Geomechanics*, pages 249–255, Balkema, Leiden. (Cited on page 141)

Peschl, G. M. (2004). *Reliability Analyses in Geotechnics with the Random Set Finite Element Method*. PhD dissertation, Technische Universität Graz, Graz, Austria. (Cited on page 141)

Rackwitz, R. (2000). Reviewing probabilistic soil modelling. *Computers and Geotechnics*, 26:199–223. (Cited on page 2)

Ramer, A. and Padet, C. (2001). Nonspecificity in $\mathbb{R}^n$. *International Journal of General Systems*, 30(6):661–680. (Cited on pages 97 and 99)

Rényi, A. (1970). *Probability theory*. North-Holland, Amsterdam. Chapter IX: Introduction to Information Theory, p. 540–616. (Cited on pages 97 and 98)

Rubinstein, R. and Kroese, D. (2007). *Simulation and the Monte Carlo Method*. John Wiley & Sons, New York, 2 edition. (Cited on page 88)

Rubinstein, R. Y. (1981). *Simulation and the Monte Carlo Method*. John Wiley & Sons, Inc., New York, NY, USA. (Cited on page 45)

Rubio, E., Hall, J. W., and Anderson, M. G. (2004). Uncertainty analyis in a slope hydrology and stability model using probabilistic and imprecise information. *Computers and Geotechnics*, 31:529–536. (Cited on page 141)

Rudin, W. (1973). *Functional Analysis*. Mc. Graw Hill, New Delhi. (Cited on page 36)

Saltelli, A., Chan, K., and Scott, E. (2000). *Sensitivity Analysis*. John Wiley and Sons, Chichister. (Cited on pages 113 and 114)

Schweiger, H. and Peschl, G. (2004). Numerical analysis of deep excavations utilizing random set theory. In Brinkgreve, R., Schad, H., Schweiger, H., and Willand, E., editors, *Proceedings of the Symposium on Geotechnical Innovations*, pages 277–294, June 25, 2004, Stuttgart. Velag Glückauf Essen. (Cited on page 141)

Sentz, K. and Ferson, S. (2002). Combination of evidence in Dempster-Shafer theory. Report SAND2002-0835, Sandia National Laboratories, Albuquerque, NM. (Cited on pages 9, 24, 25, 27, 56, and 137)

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton University Press, Princeton, NJ. (Cited on pages 3, 14, 25, and 56)

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statis. Univ. Paris*, 8:229–231. (Cited on pages 32 and 33)

Sklar, A. (1996). Random variables, distribution functions, and copulas - a personal look backward and forward. In Rüschendorf, L., Schweizer, B., and Taylor, M., editors, *Distributions with fixed marginals and related topics*, pages 1–14. Institute of Mathematical Statistics, Hayward, CA. (Cited on page 33)

Smets, P. and Kennes, R. (1994). The transferable belief model. *Artificial Intelligence*, 66:191Ű–234. (Cited on page 4)

Sugeno, M. (1974). *Theory of fuzzy integrals and its applications*. PhD dissertation, Tokyo Institute of Technology, Tokyo, Japan. (Cited on page 4)

Tonon, F. (2004a). Efficient calculation of CDF and reliability bounds using random set theory. In Wojtkiewicz, S., Ghanem, R., and Red-Horse, J., editors, *Proceedings of the 9th ASCE Joint Specialty Conference on Probabilistic Mechanics and Structural Reliability, PMC04*, July 26-28, 2004, Albuquerque, NM. ASCE, Reston, VA. Paper No. 07-104. (Cited on page 140)

Tonon, F. (2004b). On the use of random set theory to bracket the results of Monte Carlo simulations. *Reliable Computing*, 10:107–137. (Cited on page 140)

Tonon, F. (2004c). Using random set theory to propagate epistemic uncertainty through a mechanical system. *Reliability Engineering and System Safety*, 85(1–3):169–181. (Cited on pages 22, 26, 52, 53, 67, 81, 84, 122, and 138)

Tonon, F. (2007). Some properties of a random set approximation to upper and lower distribution functions. *International Journal of Approximate Reasoning*, 48(1):174–184. (Cited on page 86)

Tonon, F., Bae, H.-R., Grandhi, R. V., and Pettit, C. (2004). Using random set theory to calculate reliability bounds in situations of little information: an application to a finite element intermediate complexity wing model. In *Proceedings of the 45th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference and 6th AIAA Non-Deterministic Approaches Forum*, AIAA-2004-1583, April 18-21, 2004, Palm Springs, California. American Institute for Aeronautics and Astronautics. (Cited on page 140)

Tonon, F. and Bernardini, A. (1998). A random set approach to the optimization of certain structures. *Computers and structures*, 68:583–600. (Cited on pages 139 and 140)

Tonon, F. and Bernardini, A. (1999). Multiobjective optimization of uncertain structures through fuzzy set and random set theory. *Computer-Aided Civil and Infrastructure Engineering*, 14:119–140. (Cited on page 140)

Tonon, F., Bernardini, A., and Elishakoff, I. (1998). Concept of random sets as applied to the design of structures and analysis of expert opinions for aircraft crash. *Chaos, Solitrons & Fractals*, 10(11):1855–1868. (Cited on pages 139 and 140)

Tonon, F., Bernardini, A., and Mammino, A. (2000a). Determination of parameters range in rock engineering by means of random set theory. *Reliability Engineering and System Safety*, 70:241–261. (Cited on page 140)

Tonon, F., Bernardini, A., and Mammino, A. (2000b). Reliability of rock mass response by means of random set theory. *Reliability Engineering and System Safety*, 70:263–282. (Cited on page 140)

Tonon, F., Bernardini, A., and Mammino, A. (2001). Hazard bounding in rock engineering and ground-structure interaction by means of random set theory. In Corotis, R. B., Schuëller, G. I., and Shinozuka, M., editors, *Proceedings of*

the Eighth International Conference on Structural Safety and Reliability ICOSSAR'01, June 17–22, 2001, Newport Beach, California, USA. (Cited on page 140)

Tonon, F. and Chen, S. (2005). Inclusion properties for random relations under the hypothesis of stochastic independence and non-interactivity. *International Journal of General Systems*, 34(5):615–624. (Cited on page 139)

Tonon, F., Mammino, A., and Bernardini, A. (1996). A random set approach to the uncertainties in rock engineering and tunnel design. In Barla, G., editor, *Proceedings of the ISRM International Symposium on Prediction and Performance in Rock Mechanics and Rock Engineering EUROCK '96*, volume 2, pages 861–868, September 2-5, 1996, Torino, Italy. Balkema, Rotterdam. (Cited on page 139)

Tonon, F. and Pettit, C. L. (2004). Structural reliability application of correlation in random-set valued variables. In Wojtkiewicz, S., Ghanem, R., and Red-Horse, J., editors, *Proceedings of the 9th ASCE Joint Specialty Conference on Probabilistic Mechanics and Structural Reliability, PMC04*, July 26-28, 2004, Albuquerque, NM. ASCE, Reston, VA. Paper No. 08-201. (Cited on page 140)

Tonon, F. and Pettit, C. L. (2005). Toward a definition and understanding of correlation for variables constrained by random relations. In Augusti, G., Schuëller, G. I., and Ciampoli, M., editors, *Proceedings of the Ninth International Conference on Structural Safety and Reliability ICOSSAR'05*, pages 1741–1745, June 19–23, 2005, Rome. Millpress, Rotterdam. (Cited on page 140)

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London. (Cited on pages 4, 26, 79, and 86)

Walley, P. (1996). Measures of uncertainty in expert systems. *Artificial Intelligence*, 83(1):1–58. (Cited on page 26)

Walley, P. (2000). Towards a unified theory of imprecise probabilities. *International Journal of Approximate Reasoning*, 24(2–3):125–148. (Cited on page 26)

Wang, Z. and Klir, G. J. (1992). *Fuzzy Measure Theory*. Plenum Press, New York. (Cited on page 11)

Weichselberger, K. (2000). The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, 24:149–170. (Cited on page 26)

Weichselberger, K. (2001). *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung. Volume I - Intervallwahrscheinlichkeit als umfassendes Konzept*. Physika-Verlag, Heidelberg. (Cited on page 26)

Weichselberger, K. and Pohlmann, S. (1990). *A methodology for uncertainty in knowledge-based systems*. Springer Verlag, New York. (Cited on page 26)

Williamson, R. C. (1989). *Probabilistic Arithmetic*. PhD dissertation, University of Queensland, Australia. Available at http://users.rsise.anu.edu.au/~williams/papers/thesis300dpi.pdf. (Cited on pages 35 and 139)

Williamson, R. C. and Downs, T. (1990). Probabilistic arithmetic I: numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning*, 4:89–158. (Cited on page 139)

Wilson, N. (2000). Algorithms for Dempster-Shafer theory. In Kohlas, J. and Moral, S., editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, volume 5: Algorithms for Uncertainty and Defeasible Reasoning, pages 421–476. Kluwer, Dordrecht. Available at http://4c.ucc.ie/web/upload/publications/inBook/Handbook00wilson.pdf. (Cited on page 81)

Wolkenhauer, O. (2001). *Data Engineering*. John Wiley and Sons, New York. (Cited on page 12)

Yager, R. R. (1986). The entailment principle for dempster-shafer granules. *International Journal of Intelligent Systems*, 1:247–262. (Cited on pages 17, 102, and 104)

Zadeh, L. (1984). Review of Shafer's "A Mathematical Theory of Evidence". *Artificial Intelligence Magazine*, 5:81–83. (Cited on page 25)

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8:338–353. (Cited on page 3)

Zadeh, L. A. (1978). Fuzzy sets as the basis for a theory of possibility. *Fuzzy sets and systems*, 1:3–28. (Cited on page 4)

Zhang, J. and Berleant, D. (2005). Arithmetic on random variables: squeezing the envelopes with new joint distribution constraints. In Cozman, F. G., Nau, R., and Seidenfeld, T., editors, *Proceedings of the Fourth International Symposium on Imprecise Probabilities and their Applications, ISIPTA'05*, Carnegie Mellon University, Pittsburgh, PA, USA; July 20-23 2005. Available at http://www.sipta.org/isipta05/proceedings/papers/s013.pdf. (Cited on page 139)

Index

Curriculum vitae

Diego Andrés Alvarez Marín

He was born in September 1977 in Manizales (Caldas), Colombia.

Studies

1984-1994	Primary and secondary education. Liceo Arquidiocesano de Nuestra Señora at Manizales
1995-2000	Bachelor degree on Civil Engineering. National University of Colombia at Manizales
2001-2003	Master degree on Industrial Automatization. National University of Colombia at Manizales
2004-2007	Doctoral studies in the Faculty of Civil Engineering at the Leopold-Franzens University, Innsbruck, Austria.

Previous research

Application of machine learning techniques in the evaluation of the probability of failure of structural systems.

Application of stochastic structural control techniques to bridges subjected to wind turbulences.

Honors and awards

2004–2007	Programme Alßan scholarship, European Union programme of high level scholarships for Latin America
2003	Thesis of master degree: *Stochastic structural control of a bridge subjected to wind-induced vibrations using separated surfaces*. Marking: *meritorious* (this marking is similar to *cum laude*)
II-2002 – 2003	COLCIENCIAS "Young Researcher" (COLCIENCIAS is the Colombian research promoter institute)
I-2002	Scholarship for graduate studies at the National University of Colombia granted by the same university
2000	Final undergraduate work: *Structural reliability assessment using artificial neural networks*. Marking: *meritorious* (similar to *cum laude*)
II-1999	Prize for outstanding academic performance, National University of Colombia
1994	First place on the National University's admission test for Civil Engineering
1994	ICFES exam: 372/400; this is the Colombia's state examination for university admission. He was between the ten best results in Caldas that year.

Publications

He has five publications in International Journals and five publications in Conference Proceedings.

February, 2007