This is page i
Printer: Opaque this

OPTIMIZATION UNDER GENERALIZED UNCERTAINTY
A Unified Approach

Weldon A. Lodwick and Geraldo N. Silva

ABSTRACT

Contents

Preface

Weldon A. Lodwick
Pequistador visitante - UNESP-São José do Rio Preto
Draft 1.0, July 16, 2012
* This work has been supported by FAPESP under grant No. 2011/13985-0* This work has been supported by FAPESP under grant No. 2011/13985-0

This is page 1
Printer: Opaque this

1
Introduction

2
Generalized Uncertainty

2.1 Introduction

This chapter focuses on the various theories of generalized uncertainty that occur in data associated with optimization models. Generalized uncertainty are those uncertainties that are non-deterministic (errors) and non-probabilistic (stochastic). After we look at the theoretical aspects of generalized uncertainty, we will show how to construct distributions which will be used as inputs to optimization under generalized uncertainty models. There are two newer generalized uncertainty optimization models that are most relevant to our development, (i) interval-valued probability recourse, (ii) min/max random set optimization. Since we are interested in using the uncertainty data in optimization, our uncertainties will be quantitative, numeric. Moreover, for practical purposes, we will limit ourselves to generalized uncertainties that can be translated into interval-valued probabilities which in turn can be translated into possibility/necessity distribution pairs as we shall see.

2.1.1 Basic Definitions

Generalized uncertainty arises from incomplete or partial specification of information about the values of parameters and relationships associated in an optimization model. Many of these ideas are taken from Dubois and Prade (2009)

Definition 1 *A piece of information or data is said to be* ***incomplete*** *(imprecise, not completely specified) in a given context if it is not sufficient to allow the agent to answer a relevant question in this context.*

Example 2 *Let set* $A \equiv the\ third\ age\ of\ \ humans = \{65, ..., 90, ..., 120\}$*. When we are asking for the age of a person a and all we know is the fact that* $a \in A$*,this piece of information is an incomplete piece of information, or data since we are unable to answer the question about the age of person* a*. However, if the question is does person a get 1/2 price at the movies in Brazil or is X over 59? then the information* $Age of X \in A$ *is not incomplete.*

Definition 3 *A set used for representing incomplete information or data is called* ***disjunctive.***

Definition 4 *A set is* ***conjunctive*** *when it is defined by a collection of elements specified by listing them or giving the generating formula.*

Remark 5 *The elements of A seen as possibilistic values are mutually exclusive. That is, AgeofX is a single number in the set A. This number is different than all the other numbers. The AgeofX is real, exists, hence a piece of imprecise information or data takes the form of a disjunction of mutually exclusive values no matter the granularity.*

Definition 6 *A piece of information or data is said to be* ***uncertain*** *for an agent when the agent does not know whether or not the piece of information is true or false..*

Example 7 *In a computer program, we have the following instructions: IF x = 0 THEN y=3, OTHERWISE y =4. When the computer branches and returns y=3, do we know x=0?*

Remark 8 *According to the above definitions which come from Dubois&Prade, incomplete information has to do with the ability to answer questions in a given context given that* $a \in A$*. Uncertainty deals with our ability to tell the truth/falseness {0,1} of a given piece of information* $a \in A$*. When we cannot determine either the answer or the truth, in quantitative terms, we assign a number to it the response.*

Example 9 *"The probability that this lecture will last more than 1 hour is 0.75"* $= A, WeldonLodwick \in A$?

Remark 10 *Since we are interested in quantitative (translation to numbers) approaches, we quantify the uncertainty associated with a set A or a proposition or an event we have a measure g such that*

$$g : A \subseteq X \to [0, 1]$$

though [0,1] can be replaced by any complete order lattice. A reasonable set of assumptions on g are the following:

$$g(\varnothing) = 0, \; g(X) = 1 \tag{2.1}$$
$$A \subseteq B \Rightarrow g(A) \leq g(B). \tag{2.2}$$

The assumptions (2.1) and (2.2) define a capacity (after Choquet capacities) or a fuzzy measure. A consequence of (2.1) and (2.2) is:

$$g(A \cap B) \leq \min\{g(A), g(B)\} \tag{2.3}$$
$$g(A \cup B) \geq \max\{g(A), g(B)\}. \tag{2.4}$$

2.1.2 Theory of Generalized Uncertainty

We will consider the following distributions: (1) Possibility/Necessity, (2) P-Boxes, (3) Clouds, (4) Probability Intervals, (5) Interval-Valued Probability, (6) Belief/Plausibility, (7) Random Sets. The list is given in order of ascending generality.

<FIGURE illustrating the relationship>

Remark 11 *We will want to construct possibility and necessity pair and interval-valued pairs of distributions from special types of random sets and special types of belief/plausibility pairs. In the sequel we will see that random sets can generate interval-valued probability, some belief/plausibility pairs can be translated into interval-valued probability, all clouds can be translated into interval-valued probabilities, all probability interval can be translated into interval-valued probabilities, and all P-Boxes can be translated into interval valued probabilities. Lastly, as we will see, all interval-valued probabilities can be transformed into possibility/necessity pairs. So, our optimization under generalized uncertainty will focus on how to use interval-valued probability distributions and possibility/necessity distributions in optimization.*

Remark 12 *Since we have uncertainty in our data, we always have distinct pairs of functions describing our uncertainty. This dual set of distinct functions is not present in fuzzy membership which is unique. That is, to describe the transitional, non-Boolean set belonging, we have a unique function, not distinct dual functions as is present in generalized uncertainty distribution.*

We will need to define some structure so that the meaning of possibility and probability measures make sense. To this end we define the following concepts from measure theory.

Definition 13 *Given* $X \neq \varnothing$, *a* σ*-field defined on* X *denoted* $\mathcal{L}_X$ *is a family of subsets of* X *such that* $1)\varnothing \in \mathcal{L}_X, 2)A \in \mathcal{L}_X, 3)A_i \in \mathcal{L}_X$ *from any* *countable set (could be finite)* $\Rightarrow \{\cup_i Ai\} \in \mathcal{L}_X$. *The pair* $(X, \mathcal{L}_X)$ *is called a* ***measurable space****. Let* $(X, \mathcal{L}_X)$ *be a measurable space. By a measure* μ *on this space we mean a function*

$$\mu : \mathcal{L}_X \to \mathbb{R}^+$$

such that $\mu(\varnothing) = 0$ *and for any pair-wise disjoint set of sets* A_i, $\mu(\cup_i A_i) = \sum_{i=1} \mu(A_i)$. *The triple* $(X, \mathcal{L}_X, \mu)$ *is called a* ***measure space****.*

Remark 14 *If the mapping is* $\mu : \mathcal{L}_X \to [0,1] \subset \mathbb{R}^+$ *where* $\mu(X) = 1$, *then the measure is called a probability measure with* μ *now denoted* Pr_X *(where we drop the subscript when the context is clear) and the measure space is called a probability measure space denoted by* $(X, \mathcal{L}_X, \mathrm{Pr}_X)$.

Definition 15 *Let* $(X, \mathcal{L}_X)$ *and* $(Y, \mathcal{L}_Y)$ *be two measurable space. A function*

$$f : X \to Y$$

is said to be $(\mathcal{L}_X, \mathcal{L}_Y)$ *measurable if*

$$f^{-1}(A) = \{x \in X \mid f(x) \in A\} \in \mathcal{L}_X \text{ for each } A \in \mathcal{L}_Y.$$

Possibility/Necessity Measures and Distributions

Definition 16 *Given a measurable space, a* ***possibility measure*** *is a set-valued function*

$$\begin{aligned} Pos &: \mathcal{L}_X \to \textit{totally ordered set usually [0,1] such that} \\ Pos(\emptyset) &= 0 \\ Pos(X) &= 1 \\ Pos(A \cup B) &= \max\{Pos(A), Pos(B)\} \\ Pos(A \cap B) &= \min\{Pos(A), Pos(B)\}. \end{aligned}$$

Pos represents the state of knowledge of an agent and distinguishes (via the value) what is plausible from what is less plausible. Pos = 0 means reject as plausible while Pos =1 means completely plausible. A ***possibility distribution*** *is a function*

$$pos(x) = Pos(\{x\}).$$

If the set X *is exhaustive, then* $\exists x \in X$, *such that* $pos(x) = 1$. *If we know a possibility distribution, we can obtain a possibility measure by*

$$Pos(A) = \sup_{x \in A} pos(x).$$

Some properties associated with possibility measures are:

1. $Nec(A) = 1 - Pos(A^C)$

2. $Nec(A) + Nec(A^C) \leq 1$
3. $Pos(A) + Pos(A^C) \geq 1$
4. $Nec(A) + Pos(A^C) = 1$
5. $\min\{Nec(A), Nec(A^C) = 0$
6. $\max\{Pos(A), Pos(A^C) = 1$
7. $Nec(A) > 0 \Rightarrow Pos(A) = 1$
8. $Pos(A) < 1 \Rightarrow Nec(A) = 0$.

The necessity measure defines a necessity distribution,

$$nec(x) = Nec(\{x\})$$

and from the necessity distribution we can obtain a necessity measure

$$Nec(A) = \inf_{x \in A} nec(x).$$

General possibility theory may be derived at least in any one of the following ways:

1. Through *normalized* fuzzy sets (see [172]),
2. Axiomatically, from *fuzzy measures* g that satisfy ([35], [84])
$$g(A \cup B) = \max\{g(A), g(B)\}.$$
3. Through the belief functions of Dempster-Shafer theory, whose focal elements are normalized and nested [84],
4. By construction, though nested sets with *normalization*, for example nested $\alpha - level$ sets [66].

P-Boxes

Let Pr be a probability measure on the real line whose density (associated distribution) is denoted p. The cumulative distribution associated with Pr is

$$F_p(x) = \Pr((-\infty, x]).$$

Definition 17 *A P-Box is defined by a pair of cumulative distributions* $\underline{F}(x) \leq \overline{F}(x), x \in \mathbb{R}$. *A P-Box is used to bracket an imprecise unknown cumulative*

$$\underline{F}(x) \leq F_p(x) \leq \overline{F}(x).$$

Cloud

The idea of a cloud is to enclose uncertainty in such a way that the enclosing functions have probabilistic-like characteristics. In particular, every cloud has been shown to contain a probability distribution within it [127]. Beyond the ability to model with a mixture of uncertainty, the original impetus was to be able to model analytically when missing information, missing precision of concepts, models, and/or measurements.

Definition 18 *A **cloud** [125] over a set M is a mapping $\mathbf{x}$ that associates with each $\xi \in M$ a (non-empty, closed and bounded interval) $\mathbf{x}(\xi)$, such that,*

$$(0,1) \subseteq \bigcup_{\xi \in M} \mathbf{x}(\xi) \subseteq [0,1]. \tag{2.5}$$

*$\mathbf{x}(\xi) = [\underline{\mathbf{x}}(\xi), \overline{\mathbf{x}}(\xi)]$, is called the **level** of ξ in the cloud $\mathbf{x}$ where $\underline{\mathbf{x}}(\xi)$, and $\overline{\mathbf{x}}(\xi)$ are the **lower** and **upper level** , respectively, and $\overline{\mathbf{x}}(\xi) - \underline{\mathbf{x}}(\xi)$ is called the **width** of ξ. When the width is zero for all ξ, the cloud is called a **thin cloud**.*

When doing analysis over real numbers, a concept akin to a fuzzy number (gradual number or interval number in our previous settings) is required for clouds.

Definition 19 *[125] A real **cloudy number** is a cloud over the set $\mathbb{R}$ of real numbers. $\chi_{[a,b]}$ (χ being the characteristic function) is the cloud equivalent to an interval $[a,b]$, providing support information without additional probabilistic content. A **cloudy vector** is cloud over $\mathbb{R}^n$, where each component is a cloudy number.*

Neumaier [125] states that dependence or correlation between uncertain numbers (or the lack thereof) can be modeled by considering them jointly as components of a cloudy vector. Moreover,

> In many applications (not always, cf. Proposition 4.1, but roughly when $\overline{\mathbf{x}}(\xi) \approx 1$ for ξ near the modes of the associated distribution), the level $\mathbf{x}(\xi)$ may be interpreted as giving lower and upper bounds on the *degree of suitability* of $\xi \in M$ as a possible scenario for data modeled by the cloud $\mathbf{x}$. This degree of suitability can be given a probability interpretation by relating clouds to random variables (see (2.6) below).

We say that a random variable x with values in M belongs to a cloud $\boldsymbol{x}$ over M, and write $x \in \mathbf{x}$, if

$$\Pr(\underline{x}(x) \geq \alpha) \leq 1 - \alpha \leq \Pr(\overline{x}(x) > \alpha) \quad \forall\, \alpha \in [0,1]. \tag{2.6}$$

Pr denotes the probability of the statement given as argument, and it is required that the sets consisting of all $\xi \in M$ where $\underline{x}(x) \geq \alpha$ and $\overline{x}(x) > \alpha$

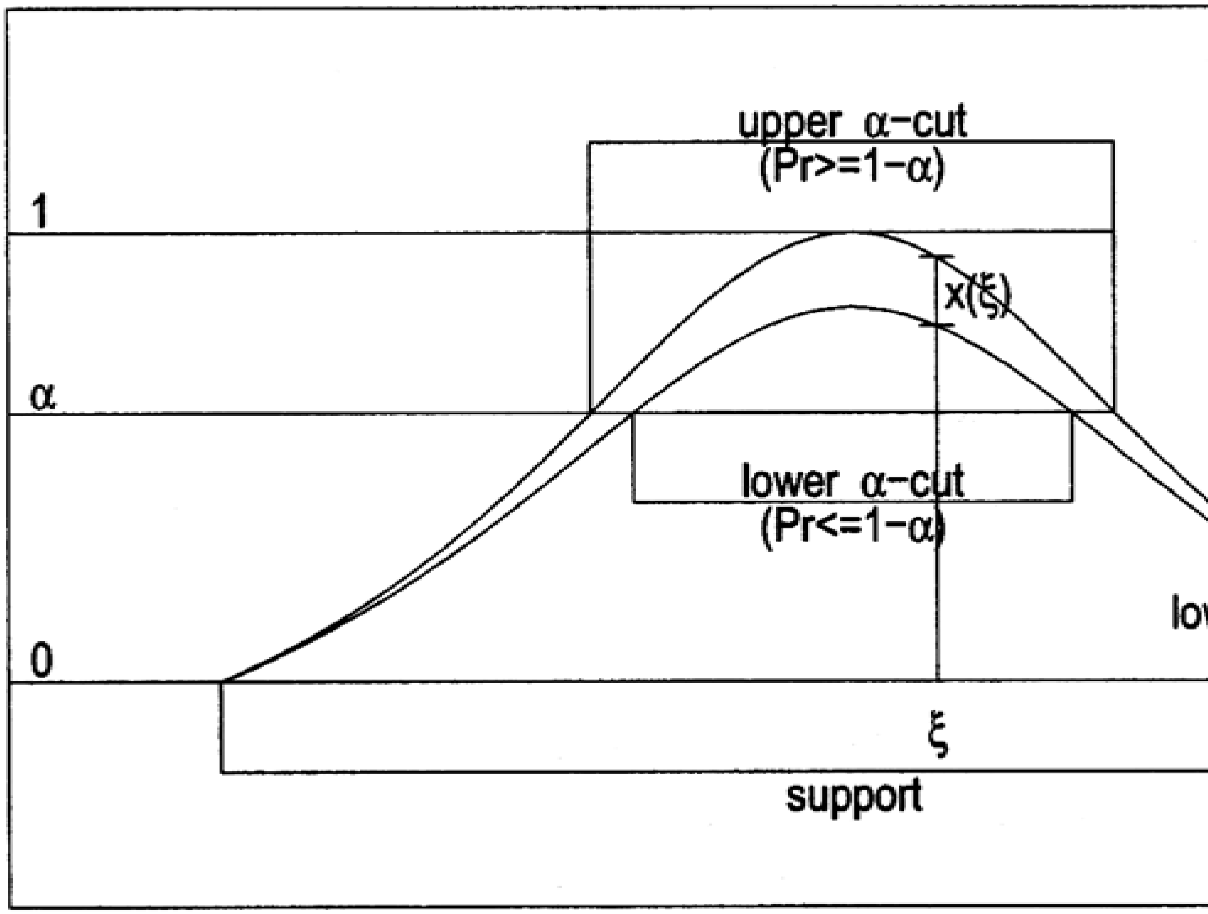

FIGURE 2.1. Cloud

are measurable in the $\sigma - algebra$ on M consisting of all sets $A \subseteq M$ for which $\Pr(x \in A)$ is defined. This approach gives clouds an underlying interpretation as the class of random variables x with $x \in \mathbf{x}$. A fuzzy cloud is one for which the level is a pair of fuzzy membership functions. The above interpretation is equivalent to the interpretation of fuzzy set membership degree as an upper bound (for a single fuzzy set considered as a cloud have the upper level being the membership function, and the $x-$axis as the lower level) for probabilities first advocated by Dubois, Moral, and Prade [40]. Since there exists at least one random variable in every cloud [127], this interpretation is meaningful.

Probability Interval

Definition 20 *Given a discrete set* $X = \{x_1, x_2, ..., x_K\}$, *and probability interval is the set*

$$PI = \{[a_k, b_k]\}, k = 1, ..., K$$

such that

$$M_{PI} = \{\Pr \mid a_k \leq \Pr(\{x_k\}) \leq b_k,\ k = 1, ..., K\}.$$

For the continuous case

$$M_{PI} = \{F_p(x) \mid a_k \leq \Pr((-\infty, x]) \leq b_k,\ k = 1, ..., K\}.$$

Interval-Valued Probability Measures

What is presented next is an expanded version of [100] and parts may also be found in [68]. A basis for linking various methods of uncertainty representation, including clouds, is examined next. This section begins by defining what is meant by an interval-valued probability measure (IVPM). This generalization of a probability measure includes probability measures, possibility and necessity measures, intervals, and clouds [125]. The set function defining IVPM is thought of as a method for giving a partial representation for an unknown probability measure much like clouds. Throughout, arithmetic operations involving set functions are in terms of interval arithmetic [113], and $Int_{[0,1]} \equiv \{[a, b] \mid 0 \leq a \leq b \leq 1\}$ denotes an arbitrary interval within $[0, 1]$. It will be seen how problems using mixed representations can be handled and solved.

The first section defines in a formal way what is called an interval-valued probability measure as used by Weichselberger ([163], [164]). Weichselberger's definition begins with a set of probability measures (our IVPM relax "tightest") and then defines an interval probability as a set function providing lower and upper bounds on the probabilities calculated from these measures. $F-probabilities$ are simply the tightest bounds possible for the set of probability measures. This definition is followed by demonstrating that various forms of uncertainty representation (possibility, interval, cloud, and probability) all can be represented by such measures. The next section shows how interval-valued probability measures can be constructed from lower and upper bounding cumulative distribution functions. This is followed by an extension principle for a function of uncertain variables represented by interval-valued probability measures and integration with respect to interval-valued probability measures. Both of these definitions will be useful in analyzing problems involving uncertainty represented by interval-valued probability measures. An application to a problem in optimization is given.

Throughout this section we will be primarily interested in interval-valued probability defined on the Borel sets on the real line and real-valued random variables. The basic definitions from Weichselberger (with slight variation in notation) are presented next.

Definition 21 *(Weichselberger [164]) Given measurable space* $(S, \mathcal{A})$*, an interval valued function* $i_m : \mathcal{A} \to Int_{[0,1]}$ *is called an* ***R-probability*** *if:*
(a) $i_m(A) = [a^-(A), a^+(A)] \subseteq [0,1]$ *with* $a^-(A) \leq a^+(A)$
(b) There exists a probability measure $\Pr$ *on* $\mathcal{A}$ *such that*

$$\forall A \in \mathcal{A}, \Pr(A) \in i_m(A)$$

By an ***R-probability field*** *we mean the triple* $(S, \mathcal{A}, i_m)$*.*

Definition 22 *(Weichselberger [164]) Given an R-probability field* $\mathcal{R} = (S, \mathcal{A}, i_m)$ *the set*

$$\mathcal{M}(\mathcal{R}) = \{\Pr \mid \Pr \text{ is a probability measure on } \mathcal{A} \text{ such that } \forall A \in \mathcal{A}, \Pr(A) \in i_m(A)\}$$

is called the ***structure*** *of* $\mathcal{R}$*.*

Definition 23 *(Weichselberger [164]) An R-probability field* $\mathcal{R} = (S, \mathcal{A}, i_m)$ *is called an* ***F-probability field****, if* $\forall A \in \mathcal{A}$*:*
(a) $a^+(A) = \sup\{\Pr(A) \mid \Pr \in \mathcal{M}(\mathcal{R})\}$*,*
(b) $a^-(A) = \inf\{\Pr(A) \mid \Pr \in \mathcal{M}(\mathcal{R})\}$*.*

It is interesting to note that given a measurable space $(S, \mathcal{A})$ and a set of probability measures P, then defining $a^+(A) = \sup\{\Pr(A) \mid \Pr \in P\}$ and $a^-(A) = \inf\{\Pr(A) \mid \Pr \in P\}$ gives an $F-probability$, where P is a subset of the structure.

The following examples show how intervals, possibility distributions, clouds and (of course) probability measures can define $R-probability$ fields on $\mathcal{B}$, the Borel sets on the real line.

Example 24 *(An interval defines an F-probability field): Let* $I = [a, b]$ *be a non-empty interval on the real line. On the Borel sets define*

$$a^+(A) = \begin{cases} 1 & \text{if } I \cap A \neq \emptyset \\ 0 & \text{otherwise} \end{cases},$$

and

$$a^{-}(A) = \begin{cases} 1 & \text{if } I \subseteq A \\ 0 & \text{otherwise} \end{cases},$$

then $i_m(A) = [a^{-}(A), a^{+}(A)]$ *defines an F-probability field* $\mathcal{R} = (R, \mathcal{B}, i_m)$. *To see this, simply set* P *to be the set of all probability measures on* $\mathcal{B}$ *such that* $\Pr(I) = 1$.

Example 25 *(A probability measure is an F-probability field) Let* $\Pr$ *be a probability measure over* $(S, \mathcal{A})$. *Define* $i_m(A) = [\Pr(A), \Pr(A)]$ *which is equivalent to having total knowledge about a probability distribution over* S.

The concept of a cloud was introduced by Neumaier in [125], and in the context of the notation of this section, it is defined as follows.

Definition 26 *A cloud over set* S *is a mapping* c *such that:*
1) $\forall s \in S,\ c(s) = [\underline{n}(s), \bar{p}(s)]$ *with* $0 \leq \underline{n}(s) \leq \bar{p}(s) \leq 1$
2) $(0,1) \subseteq \cup_{s \in S} c(s) \subseteq [0,1]$
In addition, random variable X *taking values in* S *is said to belong to cloud* c *(written* $X \in c$*) iff*
3) $\forall \alpha \in [0,1],\ \Pr(\underline{n}(X) \geq \alpha) \leq 1 - \alpha \leq \Pr(\bar{p}(X) > \alpha)$

Clouds are closely related to possibility theory. A function $p : S \to [0,1]$ is called a regular possibility distribution function if

$$\sup\{p(x) \mid x \in S\} = 1.$$

Possibility distribution functions [160] define a possibility measure,

$$Pos : S \to [0,1],$$

where

$$Pos(A) = \sup\{p(x) \mid x \in A\},$$

and it's dual necessity measure

$$Nec(A) = 1 - Pos(A^c).$$

By convention, we define $\sup\{p(x) \mid x \in \emptyset\} = 0$. A necessity distribution function can also be defined as

$$n : S \to [0,1]$$

by setting

$$n(x) = 1 - p(x).$$

Observe that

$$Nec(A) = \inf\{n(x) \mid x \in A^c\},$$

where we define $\inf\{n(x) \mid x \in \emptyset\} = 1$. In [66] it was shown that possibility distributions could be constructed which satisfy the following consistency definition.

Definition 27 *Let $p : S \to [0,1]$ be a regular possibility distribution function with associated possibility measure Pos and necessity measure Nec. Then p is said to be* ***consistent*** *with random variable X if for all measurable sets A, $Nec\,(A) \le \Pr\,(X \in A) \le Pos\,(A)$.*

Remark 28 *Recall that a distribution acts on real numbers and measures act on sets of real numbers.*

The concept of a cloud can be stated in terms of certain pairs of consistent possibility distributions which is shown in the following proposition.

Proposition 29 *Let $\bar{p}$, $\underline{p}$ be a pair of regular possibility distribution functions over set S such that $\forall s \in S\ \bar{p}\,(s) + \underline{p}(s) \ge 1$. Then the mapping $c\,(s) = [\underline{n}\,(s)\,, \bar{p}\,(s)]$ where $\underline{n}(s) = 1 - \underline{p}(s)$ (that is, the dual necessity distribution function) is a cloud. In addition, if X is a random variable taking values in S and the possibility measures associated with $\bar{p}$, $\underline{p}$ are consistent with X then X belongs to cloud c. Conversely, every cloud defines such a pair of possibility distribution functions and their associated possibility measures are consistent with every random variable belonging to the cloud c.*

Proof:
$\Rightarrow$
1) $\bar{p}$, $\underline{\mathrm{p}}$: $S \to [0,1]$ and $\bar{p}\,(s) + \underline{\mathrm{p}}(s) \ge 1$ imply property 1) of definition 46
2) Since all regular possibility distributions satisfy $\sup\{p\,(s) \mid s \in S\} = 1$ property 2) of definition 46 holds.
Therefore c is a cloud. Now assume consistency. Then
$\alpha \ge Pos\,\{s \mid p\,(s) \le \alpha\} \ge \Pr\,\{s \mid p\,(s) \le \alpha\} = 1 - \Pr\,\{s \mid p\,(s) > \alpha\}$ gives the right-hand side of the required inequalities and

$$\begin{aligned} 1-\alpha &\ge Pos\,\{s \mid p\,(s) \le 1-\alpha\} \\ &\ge \Pr\,\{s \mid p\,(s) \le 1-\alpha\} \\ &= \Pr\,\{s \mid 1 - p\,(s) \ge \alpha\} \\ &= \Pr\,(\underline{\mathrm{n}}\,(X) \ge \alpha) \end{aligned}$$

gives the left-hand side.
$\Leftarrow$The opposite identity was proven in section 5 of [125].$\square$.

Example 30 *(A cloud defines an R-probability field) Let c be a cloud over the real line. If Pos^1,Nec^1, Pos^2, Nec^2 are the possibility measures and their dual necessity measures relating to $\bar{p}\,(s)$ and $\underline{p}(s)$, define*

$$i_m\,(A) = \left[\max\left\{Nec^1\,(A)\,, Nec^2\,(A)\right\}, \min\left\{Pos^1\,(A)\,, Pos^2\,(A)\right\}\right].$$

In [127] Neumaier proved that every cloud contains a random variable X. Consistency requires that $\Pr\,(X \in A) \in i_m\,(A)$ and thus every cloud defines an R-probability field.

Example 31 *(A possibility distribution defines an R-probability field) Let* $p : S \to [0,1]$ *be a regular possibility distribution function and let* Pos *be the associated possibility measure and* Nec *the dual necessity measure. Define* $i_m(A) = [Nec(A), Pos(A)]$*. If we define a second possibility distribution,* $\underline{p}(x) = 1 \ \forall x$ *then the pair* $p, \underline{p}$ *define a cloud for which* $i_m(A)$ *defines the R-probability.*

Dempster-Shafer Belief/Plausibility Measures and Distributions

Definition 32 *A* ***belief measure*** *is a set-valued function*

$$Bel : \mathcal{L}_X \to [0,1]$$

such that

$$\begin{aligned} Bel(\emptyset) &= 0 \\ Bel(X) &= 1 \\ A_k &\in \mathcal{L}_X, i = 1, ..., K \\ Bel(\cup_{k=1}^{K} A_k) &\geq \sum_{k=1}^{K} Bel(A_k) - \sum_{j<k} Bel(A_j \cap A_k) \\ & +... + (-1)^{K+1} Bel(\cap_{k=1}^{K} A_k). \end{aligned}$$

The last property is called super-additivity. Note that if the A_k *are mutually disjoint, then*

$$Bel(\cup_{k=1}^{K} A_k) \geq \sum_{k=1}^{K} Bel(A_k)$$

which means it is distinct from probability. A plausibility measure is a set-valued function

$$Pl : \mathcal{L}_X \to [0,1]$$

such that

$$\begin{aligned} Pl(\emptyset) &= 0 \\ Pl(X) &= 1 \\ A_k &\in \mathcal{L}_X, i = 1, ..., K \\ Pl(\cup_{k=1}^{K} A_k) &\leq \sum_{k=1}^{K} Pl(A_k) - \sum_{j<k} P(A_j \cap A_k) \\ & +... + (-1)^{K+1} Pl(\cap_{k=1}^{K} A_k). \end{aligned}$$

The last property is called sub-additivity. Note that if the A_k *are mutually disjoint, then*

$$Pl(\cup_{k=1}^{K} A_k) \leq \sum_{k=1}^{K} Pl(A_k)$$

which means that plausibility measures are distinct from probability theory. We have that

$$Pl(A) = 1 - Bel(A^C).$$

Moreover if $\mathcal{L}_X$ *consists of nested sets, then*

$$Bel(A \cup B) = \max\{Bel(A), Bel(B)\}$$

and

$$Pl(A \cap B) = \min\{Pl(A), Pl(B)\}.$$

That is, in the case of a system of nested sets, Bel and Pl are Pos and Nec respectively.

Some properties are

1. $Bel(A) + Bel(A^C) \leq Bel(A \cup A^C) = Bel(X) = 1$
2. $Pl(A) + Pl(A^C) \geq Pl(A \cup A^C) = Pl(X) = 1$

Random Sets

Definition 33 *A mapping*

$$m : \mathcal{L}_X \to [0, 1]$$

such that

$$\sum_{A \in \mathcal{F}} m(A) = 1$$

generates a ***random set*** $(\mathcal{F}, m)$, $\mathcal{F}$*=*$\{A \in \mathcal{L}_X \mid m(A) > 0\}$*. The mapping m is called the* ***basic probability assignment function*** *or just assignment function.*

Remark 34

$$Bel(A) = \sum_{B \in \mathcal{F} | B \subseteq A} m(B)$$

and

$$Pl(A) = \sum_{B \in \mathcal{F} | B \cap A \neq \varnothing} m(B).$$

Moreover, if $\mathcal{F}$*=*$\{A_1, A_2, ..., A_K\}$*, a finite collection of subsets in the family* $\mathcal{F}$ *such that they are nested,* $A_1 \subseteq A_2 \subseteq ... \subseteq A_K$ *then the belief and plausibility measures generated by the random set is a possibility and necessity measure respectively. The set of all probabilities generated by a random set* $(\mathcal{F}, m)$*, for a finite family* $\mathcal{F}$*, has the form*

$$M_{RS} = \{\Pr \mid \Pr(A) = \sum_{k=1}^{K} m(A_k) \Pr{}^k(A), A \in \mathcal{L}_X\}$$

where

$$\Pr{}^k \in \{\Pr(A) = 1, A_k \subseteq A\}.$$

The representation M_{RS} of a random can also generate another belief and plausibility measure as follows (see Hung Nguyen's book).

$$\begin{aligned} Bel(A) &= \inf_{\Pr \in M_{RS}} \Pr(A), \\ Pl(A) &= \sup_{\Pr \in M_{RS}} \Pr(A). \end{aligned}$$

Moreover, if $\mathcal{F}$ is a family of singletons, then $A_x = \{x\}$ so that

$$Pl(\{x\}) = \sum_{B \in \mathcal{F} | B \cap \{x\} \neq \varnothing} m(B) = m(\{x\}).$$

That is,

$$Pl(x) = m(x).$$

Likewise,

$$Bel(\{x\}) = \sum_{B \in \mathcal{F} | B \subseteq \{x\}} m(B) = m(\{x\}).$$

That is,

$$Bel(x) = m(x).$$

Therefore in the case of singleton families,

$$Bel(x) = Pl(x) = \Pr(x).$$

Let a measure $\mu : \mathcal{L}_X \rightarrow [0,1]$ be given for a discrete space $X = \{x_1, ..., x_K\}$. Of course, in this case, $\mathcal{L}_X = P(X)$, the power set of X. On the singletons, let

$$\mu(x_k) = \alpha_k, 0 < \alpha_k \leq 1$$

since we want focal elements where $m(A) > 0$. Without loss of generality assume that $1 = \alpha_1 \geq \alpha_2 \geq ... \geq \alpha_K > 0 = \alpha_{K+1}$ by renumbering if necessary.. Let

$$\mathcal{F} = \{F(\alpha_k), k = 1, ..., K\}$$

where $F(\alpha)$ are the α-level cuts. Then μ is equivalent to the random set $(\mathcal{F}, m)$ where m is given by

$$m(F(\alpha_k)) = \alpha_k - \alpha_{k+1}, k = 1, ..., K.$$

Remark 35 *A random set can be thought of as a sets "pulled" from a given universe of sets. Suppose in a given large city we wish to find out what the distribution of education level is without asking every person in the city which would infeasible to do. We want some what to pull, from all possible unique subsets of people one of these subsets at random (according to some criteria). This is a random set*

2.2 Examples

<see my notes>..

2.3 Construction of Distributions from Generalized Uncertainty Data

2.3.1 Interval Uncertainty Data

The precursor and basis of all interval-histogram methods, the two inverse probability methods (as well as the interval convolution method above), is Moore's paper [114]. The method that R. E. Moore outlines in [114] *approximates* the cumulative distribution of a random variables and uses this to approximate of the cumulative distribution of a function of random variable (**??**) . There is no attempt to enclose the correct cumulative distribution although subsequent approaches, in particular, [11], [98], and [130], do. Given (**??**) along with the stated assumption, Moore constructs an approximation to the cumulative distribution F_Y of the function $Y = f(X)$ in the following way.

Algorithm 36 *Moore's Method [114]*
Step 1: ***Partition -*** *The support of each random variable, X_i, is subdivided into N_i subintervals, usually, of equal width.*
Step 2: ***Compute the probability of the partition*** *- The probability (histogram) for each random variable, on each of its subintervals is computed from the given probability density function.*
Step 3: ***Compute the approximate value of the function on each subinterval*** *- Use interval analysis on each of the $K = \prod_{i=1}^{n} N_i$ combinations of the subintervals. Each instantiation yields an interval where the probability assigned to the interval is the product of marginal probabilities and Moore assumes that the random variables are independent.*
Step 4: ***Order the resultant intervals and subdivide overlapping segments*** *- The intervals obtained in step 3 are ordered with respect to their left endpoints. Any overlapping intervals are subdivided.*
Step 5: ***Compute the probabilities on the overlapping segments*** *- Probabilities are assigned to the subdivided intervals in proportion to their length. That is, Moore in [114] assumes that if an interval is subdivided into say, three parts, each part of the interval receives one third of the probability associated with that interval.*
Step 6: ***Assemble into one cumulative distribution function*** *- The probability of an overlapping set of subintervals is the sum of the probabilities on each subinterval. Starting with the left-most interval, the range value of the cumulative distribution at left endpoint is zero and the right*

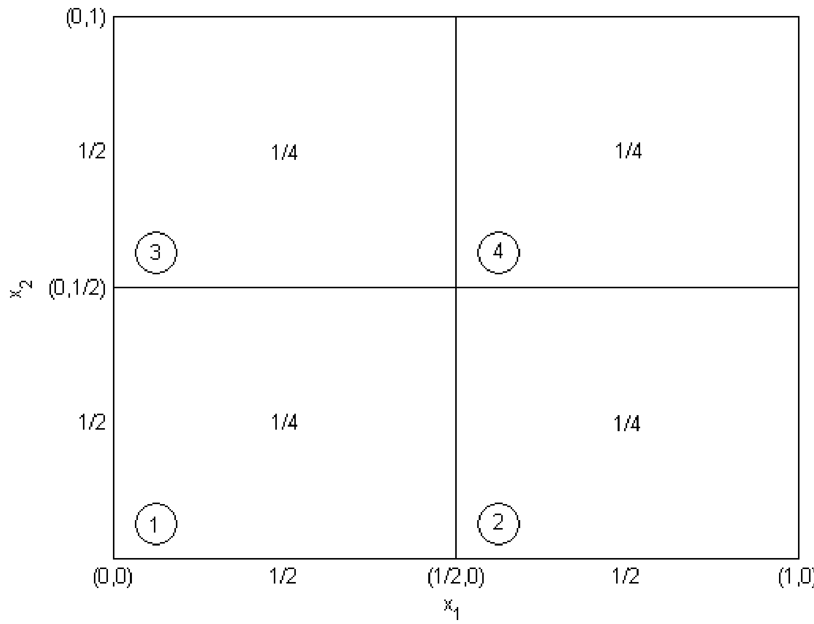

FIGURE 2.2. Partition of [0,1]×[0,1]

endpoint is probability assigned to that subinterval. Linear interpolation between the two range values is used as an approximation. The cumulative distribution at the right-endpoint of the second subinterval is the sum of the probability of the first subinterval and second subinterval. The process is continued.

Example 37 *Suppose* $Y = X_1 + X_2$, *where* $X_1 = X_2 = \begin{cases} 1 & x \in [0,1] \\ 0 & otherwise \end{cases}$, *that is,* X_1 *and* X_2 *are uniform independent random distributions on* $[0,1]$, *denoted* $U[0,1]$. *The problem is to compute, using Moore's approach, the cumulative distribution of Y.*

Step 1: **Partition -** X_1 *and* X_2 *are partitioned into two subintervals each,* $X_1 = X_{11} \cup X_{12} = [0, \frac{1}{2}] \cup [\frac{1}{2}, 1]$, *and* $X_2 = X_{21} \cup X_{22} = [0, \frac{1}{2}] \cup [\frac{1}{2}, 1]$, *resulting in a partition consisting of four boxes (1)* $[0, \frac{1}{2}] \times [0, \frac{1}{2}]$,*(2)* $[\frac{1}{2}, 1] \times [0, \frac{1}{2}]$, *(3)* $[0, \frac{1}{2}] \times [\frac{1}{2}, 1]$, *(4)* $[\frac{1}{2}, 1] \times [\frac{1}{2}, 1]$, *(see Figure 4.1) .*

Step 2: **Compute the probability of the partition** - *Since the probability on each subinterval is* $\frac{1}{2}$ *and the random variables are assumed to be independent, the probability on each box is* $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

Step 3: **Compute the approximate value of the function on each subinterval -** *Use interval arithmetic (or constraint interval arithmetic) to*

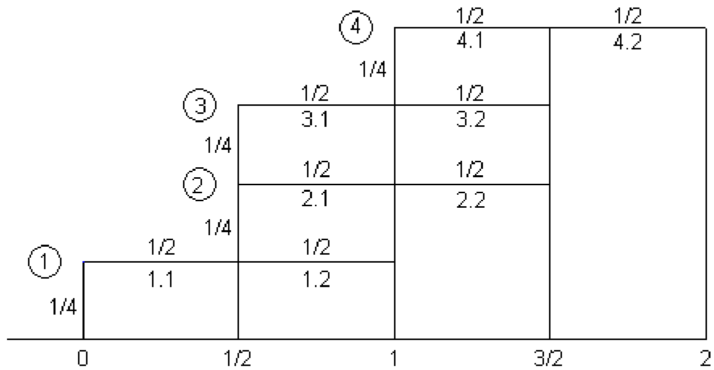

FIGURE 2.3. Resultant Intervals and Overlaps

compute the value of Y on each box. For box (1), $Y_1 = [0, \frac{1}{2}] + [0, \frac{1}{2}] = [0, 1]$, where the probability for Y_1 is $\frac{1}{4}$. For box (2), $Y_2 = [\frac{1}{2}, 1] + [0, \frac{1}{2}] = [\frac{1}{2}, \frac{3}{2}]$, where the probability for Y_2 is $\frac{1}{4}$. For box (3), $Y_3 = [0, \frac{1}{2}] + [\frac{1}{2}, 1] = [\frac{1}{2}, \frac{3}{2}]$, where the probability for Y_3 is $\frac{1}{4}$. For box (4), $Y_4 = [\frac{1}{2}, 1] + [\frac{1}{2}, 1] = [1, 2]$, where the probability for Y_4 is $\frac{1}{4}$.

Step 4: **Order the resultant intervals and subdivide overlapping segments-** *From the computations, the intervals are ordered according to their left endpoints as follows: Y_1, Y_2, Y_3, Y_4, with distinct (overlapping) subintervals of $[0, \frac{1}{2}]$ - one subinterval, $[\frac{1}{2}, 1]$ - three subintervals, $[1, \frac{3}{2}]$ - three subintervals, and $[\frac{3}{2}, 2]$ - one subinterval (see Figure 4.2).*

Step 5: **Compute the probabilities on the overlapping segments** - *Since $[0, \frac{1}{2}]$ came from subdividing $Y_1 = [0, 1]$ in half, the probability on the first subinterval is $\frac{1}{2} \times \frac{1}{4} = \frac{1}{8}$. The assumption that Moore makes is that the proportion of the division of the interval is the probability. In the same manner, there are three overlapping subintervals comprising $[\frac{1}{2}, 1]$ arising from half portions of Y_1, Y_2,* and *Y_3, each bearing probability of $\frac{1}{2} \times \frac{1}{4} = \frac{1}{8}$ so that the probability on $[\frac{1}{2}, 1]$ is $\frac{3}{8}$. In the same manner, the probability on $[1, \frac{3}{2}]$ is $\frac{3}{8}$ and on$[\frac{3}{2}, 1]$ is $\frac{1}{8}$ (see Figure 4.2).*

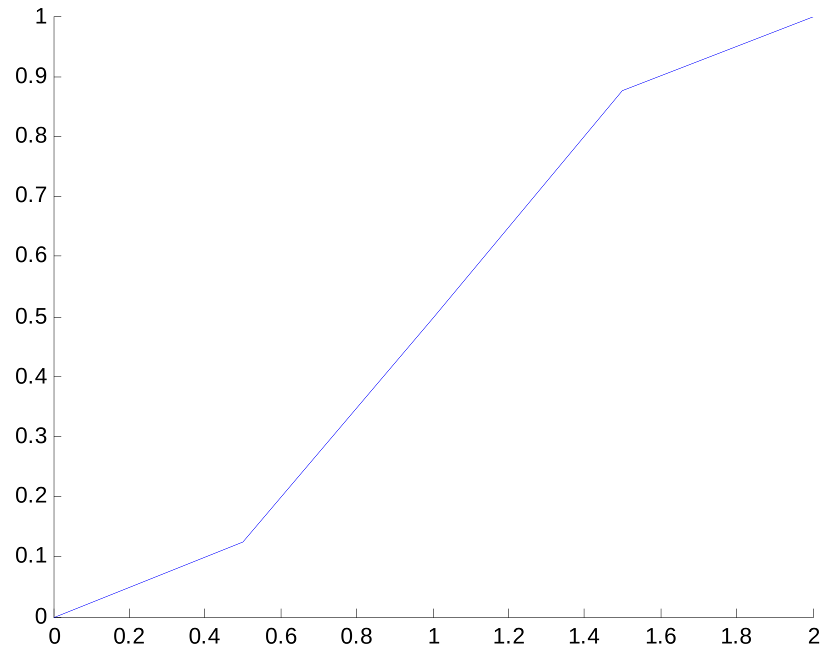

FIGURE 2.4. Assembled CDF

Step 6: **Assemble into one cumulative distribution function -** $Y(x) = \begin{cases} 0, x < 0 \\ \frac{1}{8}, x = \frac{1}{2} \\ \frac{1}{8} + \frac{3}{8} = \frac{1}{2}, x = 1 \\ \frac{4}{8} + \frac{3}{8} = \frac{7}{8}, x = \frac{3}{2} \\ \frac{1}{8} + \frac{7}{8} = 1, x = 2 \\ 1, x > 1 \end{cases}$ *with linear interpolation in between (see Figure 4.3).*

2.3.2 Construction from Kolmagorov-Smirnoff Statistics

The Kolmogorov Theorem: Given the empirical distribution function F_n for n independent distributed observations x_i where

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} C_{x_i}(x)$$

and

$$C_{x_i}(x) \begin{cases} = 1 \text{ if } x_i \leq x \\ 0 \text{ otherwise} \end{cases} .$$

Then the Kolmogorov-Smirnoff statistic states that for a given distribution F (hypothesized to be the actual underlying distribution described by the data),

$$D_n = \sup_x \|F_n(x) - F(x)\|$$

where

$$\sqrt{n} D_n \leq 1 - \alpha$$

and $1 - \alpha$ is the confidence interval and D_n was obtained by Smirnoff and is now obtained from tables.

The theorem is used to test whether or not a cumulative distribution that we think will describe our data meets a confidence interval criteria given by Kolmogorov's Theorem as computed via Smirnoff's tables. What we will do is the opposite. Given a data set, we want to obtain upper and lower bounds on the unknown cumulative distribution postulated to have generated our given data set so the confidence interval on the unknown distribution is what we specify.

Theorem 38 *A* $100(1-\alpha)\%$ *confidence band for the unknown cumulative distribution* F *is given by* $\underline{F}(x)$ *and* $\overline{F}(x)$ *where* d *is selected such that* $P(D_n \geq d) = \alpha$ *and*

$$\begin{aligned} \underline{F}(x) &= \begin{cases} 0 \text{ if } F_n(x) - d < 0 \\ F_n(x) - d \text{ if } F_n(x) - d \geq 0 \end{cases} , \\ \overline{F}(x) &= \begin{cases} F_n(x) + d \leq 0 \text{ if } F_n(x) + d \leq 1 \\ 1 \text{ if } F_n(x) + d > 1 \end{cases} . \end{aligned}$$

There are tables that given an α give you the d.

Example 39 *(from the internet) Suppose we have a random sample of 10 professors at UNESP who are asked how many minutes they spend each weekday looking at email and we are given the following data: 108, 112, 117, 130, 111, 131, 113, 113, 105, 128. We first order these data points* *and form* $F_n(x_k) = \frac{k}{n}$ *which in our case is* $F_{10}(x_k) = \frac{k}{10}$.

k	x_k	$F_{10}(x_k) = \frac{k}{10}$
1	105	0.1
2	108	0.2
3	111	0.3
4	112	0.4
5	113	0.5
6	113	0.6
7	117	0.7
8	128	0.8
9	130	0.9
10	131	1.0

Now, if we want a 95% confidence interval,

$$\sqrt{n}D_n \leq (1 - 0.95) = 0.05$$

$\Rightarrow$

$$D_{10} \leq \frac{0.05}{\sqrt{10}} = 1.58113... \times 10^{-2}$$

and looking 0.0158113... up in the tables, we get $d \approx 0.459$ *or rounding to two digits, we use* 0.46 *and we calculate according to the above theorem the upper and lower cumulative probabilities to be* $\underline{F}(x) = F_{10}(x) - d = F_{10}(x) - 0.46$ *and* $\overline{F}(x) = F_{10}(x) + d = F_{10}(x) + 0.46$ *but with the restrictions. We get the following.*

k	x_k	$F_{10}(x)$	$F_{10}(x) - d$	$\underline{F}(x)$	$F_{10}(x) + d$	$\overline{F}(x)$
1	105	0.1	$0.1 - 0.46 = -0.36$	0	$0.1 + 0.46 = 0.56$	0.56
2	108	0.2	$0.2 - 0.46 = -0.26$	0	$0.2 + 0.46 = 0.66$	0.66
3	111	0.3	$0.3 - 0.46 = -0.16$	0	$0.3 + 0.46 = 0.76$	0.76
4	112	0.4	$0.4 - 0.46 = -0.06$	0	$0.4 + 0.46 = 0.86$	0.86
5	113	0.5	$0.5 - 0.46 = 0.04$	0.04	$0.5 + 0.46 = 0.96$	0.96
6	113	0.6	$0.6 - 0.46 = 0.14$	0.14	$0.6 + 0.46 = 1.06$	1
7	117	0.7	$0.7 - 0.46 = 0.24$	0.24	$0.7 + 0.46 = 1.16$	1
8	128	0.8	$0.8 - 0.46 = 0.34$	0.34	$0.8 + 0.46 = 1.26$	1
9	130	0.9	$0.9 - 0.46 = 0.44$	0.44	$0.9 + 0.46 = 1.36$	1
10	131	1.0	$1.0 - 0.46 = 0.54$	0.54	$1.0 + 0.46 = 1.46$	1

2.3.3 Cloud Uncertainty Data

Computation with clouds involves arithmetic on the (vertical) interval $\mathbf{x}(\zeta)$ for each ξ that defines the lower and upper level of the cloud. Efficient methods to compute with clouds is still an open question. However, once a cloud is constructed, methods that have been discussed can be used.

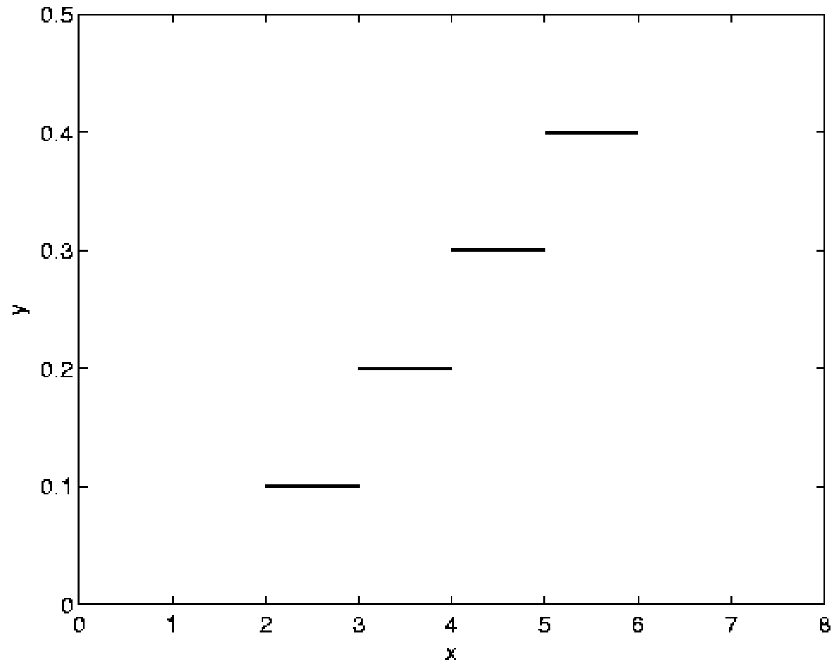

FIGURE 2.5. Probability Density - Histogram-Based Random Variable

Example 40 *[79] Suppose a histogram-based random variable x is given as follows.*

Intervals X_i		$\Pr(x \in X_i)$	*Walley [158] uses a cumulative α_i*	$\mathbf{x}(\xi) := [\alpha_{i-1}, \alpha_i]$
X_1	$(-\infty, 2)$	0.0	0.0	$[0.0, 0.0]$
X_2	$[2, 3)$	0.1	0.1	$[0.0, 0.1]$
X_3	$[3, 4)$	0.2	0.3	$[0.1, 0.3]$
X_4	$[4, 5)$	0.3	0.6	$[0.3, 0.6]$
X_5	$[5, 6]$	0.4	1.0	$[0.6, 1.0]$
X_6	$(6, \infty)$	0.0	1.0	$[1.0, 1.0]$.

The original density is not crisp. The density of Figure 4.14 assumes uniform distribution within each interval. If the X_i are chosen in the order shown in Figure 4.14, then we have a discrete cloud which looks like Figure 4.15.

The dashed lines of Figure 4.15 form $\underline{\mathbf{x}}(\xi)$, the lower level, and the solid line of Figure 4.15 form $\overline{\mathbf{x}}(\xi)$, the upper level. There is nothing to prevent the cloud from being formed using a different order of the X_i. In fact, there may some advantage in choosing the intervals in a different order so that

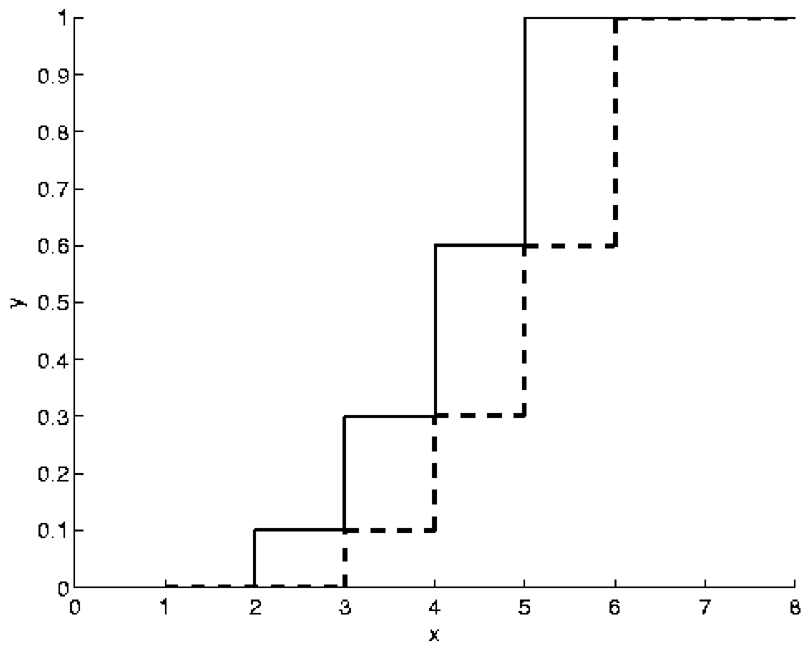

FIGURE 2.6. Discrete Cloud Constructed from Probability Density in Order Given

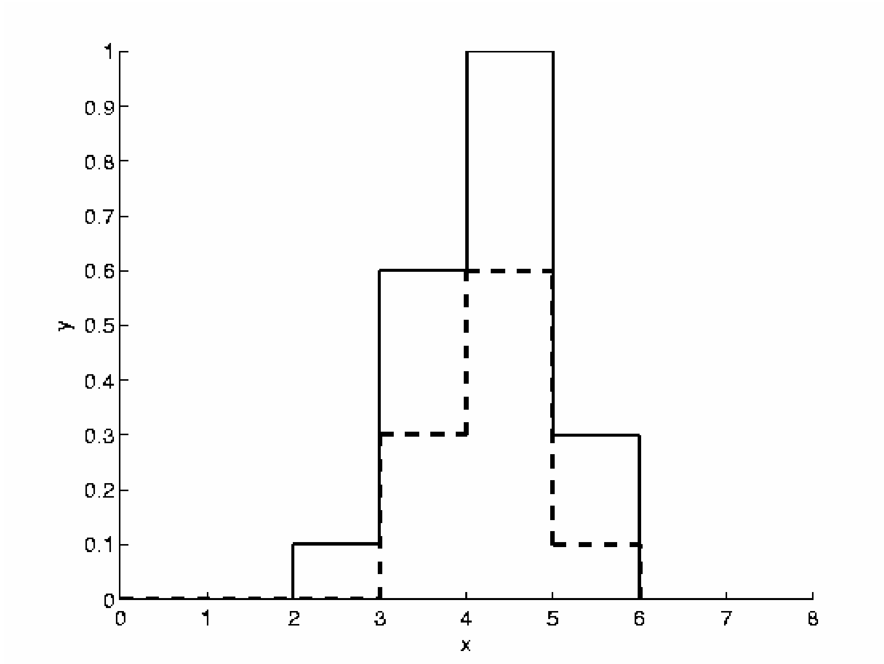

FIGURE 2.7. Discrete Cloud Constructed from Density in Different Order Than That Given

it results in a "bell-shaped" cloud as can be seen in Figure 4.16.

2.3.4 Interval-valued Probability Measures

What is presented next is an expanded version of [100] and parts may also be found in [68]. A basis for linking various methods of uncertainty representation, including clouds, is examined next. This section begins by defining what is meant by an interval-valued probability measure (IVPM). This generalization of a probability measure includes probability measures, possibility and necessity measures, intervals, and clouds [125]. The set function defining IVPM is thought of as a method for giving a partial representation for an unknown probability measure much like clouds. Throughout, arithmetic operations involving set functions are in terms of interval arithmetic [113], and $Int_{[0,1]} \equiv \{[a,b] \mid 0 \le a \le b \le 1\}$ denotes an arbitrary interval within $[0,1]$. It will be seen how problems using mixed representations can be handled and solved.

The first section defines in a formal way what is called an interval-valued probability measure as used by Weichselberger ([163], [164]). Weichselberger's definition begins with a set of probability measures (our IVPM relax "tightest") and then defines an interval probability as a set function providing lower and upper bounds on the probabilities calculated from these measures. $F-probabilities$ are simply the tightest bounds possible for the set of probability measures. This definition is followed by demonstrating that various forms of uncertainty representation (possibility, interval, cloud, and probability) all can be represented by such measures. The next section shows how interval-valued probability measures can be constructed from lower and upper bounding cumulative distribution functions. This is followed by an extension principle for a function of uncertain variables represented by interval-valued probability measures and integration with respect to interval-valued probability measures. Both of these definitions will be useful in analyzing problems involving uncertainty represented by interval-valued probability measures. An application to a problem in optimization is given.

Throughout this section we will be primarily interested in interval-valued probability defined on the Borel sets on the real line and real-valued random variables. The basic definitions from Weichselberger (with slight variation in notation) are presented next.

Definition 41 *(Weichselberger [164]) Given measurable space $(S,\mathcal{A})$, an interval valued function $i_m : \mathcal{A} \to Int_{[0,1]}$ is called an **R-probability** if:*
(a) $i_m(A) = [a^-(A), a^+(A)] \subseteq [0,1]$ with $a^-(A) \leq a^+(A)$
(b) There exists a probability measure $\Pr$ on $\mathcal{A}$ such that

$$\forall A \in \mathcal{A}, \Pr(A) \in i_m(A)$$

*By an **R-probability field** we mean the triple $(S,\mathcal{A},i_m)$.*

Definition 42 *(Weichselberger [164]) Given an R-probability field $\mathcal{R} = (S,\mathcal{A},i_m)$ the set*

$$\mathcal{M}(\mathcal{R}) = \{\Pr \mid \Pr \text{ is a probability measure on } \mathcal{A} \text{ such that } \forall A \in \mathcal{A}, \Pr(A) \in i_m(A)\}$$

*is called the **structure** of $\mathcal{R}$.*

Definition 43 *(Weichselberger [164]) An R-probability field $\mathcal{R} = (S,\mathcal{A},i_m)$ is called an **F-probability field**, if $\forall A \in \mathcal{A}$:*
(a) $a^+(A) = \sup\{\Pr(A) \mid \Pr \in \mathcal{M}(\mathcal{R})\}$,
(b) $a^-(A) = \inf\ \{\Pr(A) \mid \Pr \in \mathcal{M}(\mathcal{R})\}$.

It is interesting to note that given a measurable space $(S,\mathcal{A})$ and a set of probability measures P, then defining $a^+(A) = \sup\{\Pr(A) \mid \Pr \in P\}$ and $a^-(A) = \inf\{\Pr(A) \mid \Pr \in P\}$ gives an $F-probability$, where P is a subset of the structure.

The following examples show how intervals, possibility distributions, clouds and (of course) probability measures can define $R-probability$ fields on $\mathcal{B}$, the Borel sets on the real line.

Example 44 *(An interval defines an F-probability field): Let* $I = [a, b]$ *be a non-empty interval on the real line. On the Borel sets define*

$$a^{+}(A) = \begin{cases} 1 & \text{if } I \cap A \neq \emptyset \\ 0 & \text{otherwise} \end{cases},$$

and

$$a^{-}(A) = \begin{cases} 1 & \text{if } I \subseteq A \\ 0 & \text{otherwise} \end{cases},$$

then $i_m(A) = [a^{-}(A), a^{+}(A)]$ *defines an F-probability field* $\mathcal{R} = (R, \mathcal{B}, i_m)$*. To see this, simply set* P *to be the set of all probability measures on* $\mathcal{B}$ *such that* $\Pr(I) = 1$*.*

Example 45 *(A probability measure is an F-probability field) Let* $\Pr$ *be a probability measure over* $(S, \mathcal{A})$*. Define* $i_m(A) = [\Pr(A), \Pr(A)]$ *which is equivalent to having total knowledge about a probability distribution over* S*.*

The concept of a cloud was introduced by Neumaier in [125], and in the context of the notation of this section, it is defined as follows.

Definition 46 *A cloud over set* S *is a mapping* c *such that:*
1) $\forall s \in S,\ c(s) = [\underline{n}(s), \bar{p}(s)]$ *with* $0 \leq \underline{n}(s) \leq \bar{p}(s) \leq 1$
2) $(0, 1) \subseteq \cup_{s \in S} c(s) \subseteq [0, 1]$
In addition, random variable X *taking values in* S *is said to belong to cloud* c *(written* $X \in c$*) iff*
3) $\forall \alpha \in [0, 1],\ \Pr(\underline{n}(X) \geq \alpha) \leq 1 - \alpha \leq \Pr(\bar{p}(X) > \alpha)$

Clouds are closely related to possibility theory. A function $p : S \to [0, 1]$ is called a regular possibility distribution function if

$$\sup\{p(x) \mid x \in S\} = 1.$$

Possibility distribution functions [160] define a possibility measure,

$$Pos : S \to [0, 1],$$

where

$$Pos(A) = \sup\{p(x) \mid x \in A\},$$

and it's dual necessity measure

$$Nec(A) = 1 - Pos(A^c).$$

By convention, we define $\sup\{p(x) \mid x \in \emptyset\} = 0$. A necessity distribution function can also be defined as

$$n : S \to [0, 1]$$

by setting

$$n(x) = 1 - p(x).$$

Observe that

$$Nec(A) = \inf\{n(x) \mid x \in A^c\},$$

where we define $\inf\{n(x) \mid x \in \emptyset\} = 1$. In [66] it was shown that possibility distributions could be constructed which satisfy the following consistency definition.

Definition 47 *Let $p : S \to [0,1]$ be a regular possibility distribution function with associated possibility measure Pos and necessity measure Nec. Then p is said to be **consistent** with random variable X if for all measurable sets A, $Nec(A) \leq \Pr(X \in A) \leq Pos(A)$.*

Remark 48 *Recall that a distribution acts on real numbers and measures act on sets of real numbers.*

The concept of a cloud can be stated in terms of certain pairs of consistent possibility distributions which is shown in the following proposition.

Proposition 49 *Let $\bar{p}$, $\underline{p}$ be a pair of regular possibility distribution functions over set S such that $\forall s \in S$ $\bar{p}(s) + \underline{p}(s) \geq 1$. Then the mapping $c(s) = [\underline{n}(s), \bar{p}(s)]$ where $\underline{n}(s) = 1 - \underline{p}(s)$ (that is, the dual necessity distribution function) is a cloud. In addition, if X is a random variable taking values in S and the possibility measures associated with $\bar{p}$, $\underline{p}$ are consistent with X then X belongs to cloud c. Conversely, every cloud defines such a pair of possibility distribution functions and their associated possibility measures are consistent with every random variable belonging to the cloud c.*

Proof:
$\Rightarrow$
1) $\bar{p}$, $\underline{p}$: $S \to [0,1]$ and $\bar{p}(s) + \underline{p}(s) \geq 1$ imply property 1) of definition 46
2) Since all regular possibility distributions satisfy $\sup\{p(s) \mid s \in S\} = 1$ property 2) of definition 46 holds.
Therefore c is a cloud. Now assume consistency. Then
$\alpha \geq Pos\{s \mid p(s) \leq \alpha\} \geq \Pr\{s \mid p(s) \leq \alpha\} = 1 - \Pr\{s \mid p(s) > \alpha\}$ gives the right-hand side of the required inequalities and

$$\begin{aligned} 1 - \alpha &\geq Pos\{s \mid p(s) \leq 1 - \alpha\} \\ &\geq \Pr\{s \mid p(s) \leq 1 - \alpha\} \\ &= \Pr\{s \mid 1 - p(s) \geq \alpha\} \\ &= \Pr(\underline{n}(X) \geq \alpha) \end{aligned}$$

gives the left-hand side.
$\Leftarrow$The opposite identity was proven in section 5 of [125].$\square$.

Example 50 *(A cloud defines an R-probability field) Let c be a cloud over the real line. If Pos^1,Nec^1, Pos^2, Nec^2 are the possibility measures and their dual necessity measures relating to $\bar{p}\,(s)$ and $\underline{p}(s)$, define*

$$i_m\,(A) = \left[\max\left\{Nec^1\,(A)\,, Nec^2\,(A)\right\}, \min\left\{Pos^1\,(A)\,, Pos^2\,(A)\right\}\right].$$

In [127] Neumaier proved that every cloud contains a random variable X. Consistency requires that $\Pr(X \in A) \in i_m\,(A)$ and thus every cloud defines an R-probability field.

Example 51 *(A possibility distribution defines an R-probability field) Let $p : S \to [0, 1]$ be a regular possibility distribution function and let Pos be the associated possibility measure and Nec the dual necessity measure. Define $i_m\,(A) = [Nec\,(A)\,, Pos\,(A)]$. If we define a second possibility distribution, $\underline{p}(x) = 1\ \forall x$ then the pair p,$\underline{p}$ define a cloud for which $i_m\,(A)$ defines the R-probability.*

2.3.5 Construction of Interval-Valued Probability Measures from Kolmagorov-Smirnoff Statistics

In this section an $F-$*probability* is constructed from lower and upper bounding cumulative distribution functions in a manner allowing practical computation. For example, given statistical data we can construct a confidence interval for the underlying cumulative distribution function using the Kolmogorov estimation of confidence limits (see [87]). Then using this confidence interval we can use the following development to construct an interval-valued probability measure. Although a simple definition could be used by simply setting the interval equal to the lower and upper bound over all probability measures contained in the bound, it is not clear how to use this definition in practice. The development that follows is more amenable to actual use.

Let

$$F^u\,(x) = \Pr\left(X^u \leq x\right),$$

and

$$F^l\,(x) = \Pr\left(X^l \leq x\right),$$

be two cumulative distribution functions for random variables over the Borel sets on the real line, X^u and X^l, with the property that $F^u\,(x) \geq F^l\,(x)\ \forall x$. Set

$$\mathcal{M}\left(X^u, X^l\right) = \left\{X \mid \forall x\ F^u\,(x) \geq \Pr\left(X \leq x\right) \geq F^l\,(x)\right\},$$

which clearly contains X^u and X^l. We will think in terms of an unknown $X \in \mathcal{M}\left(X^u, X^l\right)$. For any Borel set A, let $\Pr(A) = \Pr(X \in A)$.

We begin by developing probability bounds for members of the family of sets

$$\mathcal{I} = \{(a, b]\,, (-\infty, a]\,, (a, \infty)\,, (-\infty, \infty)\,, \emptyset \mid a < b\}$$

For $I = (-\infty, b]$ it is clear by definition that

$$\Pr(I) \in \left[F^l(b), F^u(b)\right].$$

For $I = (a, \infty)$, let

$$\Pr(I) \in \left[1 - F^u(a), 1 - F^l(a)\right].$$

For $I = (a, b]$, since $I = (-\infty, b] - (-\infty, a]$, and considering minimum and maximum probabilities in each set, let

$$\Pr(I) \in \left[\max\left\{F^l(b) - F^u(a), 0\right\}, F^u(b) - F^l(a)\right].$$

Therefore, if we extend the definition of F^u, and F^l by defining

$$F^u(-\infty) = F^l(-\infty) = 0,$$

and

$$F^u(\infty) = F^l(\infty) = 1,$$

we can make the following general definition.

Definition 52 *For any $I \in \mathcal{I}$, if $I \neq \emptyset$, define*

$$i_m(I) = \left[a^-(I), a^+(I)\right] = \left[\max\left\{F^l(b) - F^u(a), 0\right\}, F^u(b) - F^l(a)\right]$$

where a and b are the left and right endpoints of I. Otherwise set

$$i_m(\emptyset) = [0, 0].$$

Remark 53 *Note that with this definition*

$$\begin{aligned} i_m((-\infty, \infty)) &= \left[\max\left\{F^l(\infty) - F^u(-\infty), 0\right\}, F^u(\infty) - F^l(-\infty)\right] \\ &= [1, 1], \end{aligned}$$

which matches our intuition and thus, it is easy to see that $\Pr(I) \in i_m(I)$ *$\forall I \in \mathcal{I}$.*

We can extend this to include finite unions of elements of $\mathcal{I}$. For example, if

$$E = I_1 \cup I_2 = (a, b] \cup (c, d]$$

with $b < c$, then we consider the probabilities,

$$\Pr((a, b]) + \Pr((c, d]),$$

$$1 - \left(\Pr((-\infty, a]) + \Pr((b, c]) + \Pr((d, \infty))\right),$$

(the probability of the sets that make up E versus one less the probability of the intervals that make up the complement), and consider the minimum and maximum probability for each case as a function of the minimum and maximum of each set. The *minimum* for the first sum is

$$\max\left(0, F^l(d) - F^u(c)\right) + \max\left(0, F^l(b) - F^u(a)\right),$$

and the *maximum* is

$$F^u(d) - F^l(c) + F^u(b) - F^l(a).$$

The minimum for the second is

$$\begin{aligned} & 1 - \left(F^u(\infty) - F^l(d) + F^u(c) - F^l(b) + F^u(a) - F^l(-\infty)\right) \\ = \; & F^l(d) - F^u(c) + F^l(b) - F^u(a) \end{aligned}$$

and the maximum is

$$\begin{aligned} & 1 - \left(\max\left(0, F^l(\infty) - F^u(d)\right) + \max\left(0, F^l(c) - F^u(b)\right) + \max\left(0, F^l(a) - F^u(-\infty)\right)\right) \\ = \; & F^u(d) - \max\left(0, F^l(c) - F^u(b)\right) - F^l(a). \end{aligned}$$

This gives

$$\Pr(E) \geq \max\left\{\begin{array}{l} F^l(d) - F^u(c) + F^l(b) - F^u(a) \\ \max\left(0, F^l(d) - F^u(c)\right) + \max\left(0, F^l(b) - F^u(a)\right) \end{array}\right., \text{ and}$$

$$\Pr(E) \leq \min\left\{\begin{array}{l} F^u(d) - \max\left(0, F^l(c) - F^u(b)\right) - F^l(a) \\ F^u(d) - F^l(c) + F^u(b) - F^l(a) \end{array}\right.,$$

so

$$\begin{aligned} \Pr(E) \in \; & \\ & [\max\left(0, F^l(d) - F^u(c)\right) + \max\left(0, F^l(b) - F^u(a)\right), \\ & F^u(d) - \max\left(0, F^l(c) - F^u(b)\right) - F^l(a)]. \end{aligned}$$

The final line is arrived at by noting that

$$\forall x, y F^l(x) - F^u(y) \leq \max\left(0, F^l(x) - F^u(y)\right).$$

Remark 54 *Note the two extreme cases for $E = (a, b] \cup (c, d]$. For $F^u(x) = F^l(x) = F(x)\ \forall x$, then, as expected,*

$$\Pr(E) = F(d) - F(c) + F(b) - F(a) = \Pr((a, b]) + \Pr((c, d]),$$

that is, it is the probability measure. Moreover, for $F^l(x) = 0\ \forall x$,

$$\Pr(E) \in [0, F^u(d)],$$

that is, it is a possibility measure for the possibility distribution function $F^u(x)$.

Let

$$\mathcal{E} = \left\{ \cup_{k=1}^{K} I_k \mid I_k \in \mathcal{I} \right\}.$$

That is, $\mathcal{E}$ is the algebra of sets generated by I. Note that every element of E has a unique representation as a union of the minimum number of elements of $\mathcal{I}$ (or, stated differently, as a union of disconnected elements of $\mathcal{I}$). Note also that $R \in \mathcal{E}$ and $\mathcal{E}$ is closed under complements.

Assume $E = \cup_{k=1}^{K} I_k$ and $E^c = \cup_{j=1}^{J} M_j$ are the unique representations of E and E^c in $\mathcal{E}$ in terms of elements of $\mathcal{I}$. Then, considering minimum and maximum possible probabilities of each interval, it is clear that

$$\Pr(E) \in [\max\left(\Sigma_{k=1}^{K} a^{-}(I_k), 1 - \Sigma_{j=1}^{J} a^{+}(M_j)\right), \min\left(\Sigma_{k=1}^{K} a^{+}(I_k), 1 - \Sigma_{j=1}^{J} a^{-}(M_j)\right)].$$

This can be made more concise using the following result.

Proposition 55 *If $E = \cup_{k=1}^{K} I_k$ and $E^c = \cup_{j=1}^{J} M_j$ are the unique representations of E and $E^c \in \mathcal{E}$, then $\Sigma_{k=1}^{K} a^{-}(I_k) \geq 1 - \Sigma_{j=1}^{J} a^{+}(M_j)$, and $\Sigma_{k=1}^{K} a^{+}(I_k) \geq 1 - \Sigma_{j=1}^{J} a^{-}(M_j)$.*

Proof:
We need only prove

$$\Sigma_{k=1}^{K} a^{-}(I_k) \geq 1 - \Sigma_{j=1}^{J} a^{+}(M_j),$$

since we can exchange the roles of E and E^c, giving

$$\Sigma_{j=1}^{J} a^{-}(M_j) \geq 1 - \Sigma_{k=1}^{K} a^{+}(I_k),$$

thereby proving the second inequality. Note $\Sigma_{k=1}^{K} a^{-}(I_k) + \Sigma_{j=1}^{J} a^{+}(M_j)$ is of the form

$$\begin{aligned} & \Sigma_{k=1}^{K} \max\left(0, F^{l}(b_k) - F^{u}(a_k)\right) + \Sigma_{j=1}^{J} F^{u}(a_{j+1}) - F^{l}(b_j) \\ \geq \quad & \Sigma_{k=1}^{K} \left(F^{l}(b_k) - F^{u}(a_k)\right) + \Sigma_{j=1}^{J} F^{u}(a_{j+1}) - F^{l}(b_j). \end{aligned}$$

Since the union of the disjoint intervals yields all of the real line, we have either $F^{u}(\infty)$ or $F^{l}(\infty)$ less either $F^{u}(-\infty)$ or $F^{l}(-\infty)$ which is one regardless.□

Next i_m is extended to $\mathcal{E}$.

Proposition 56 *For any $E \in \mathcal{E}$, let $E = \cup_{k=1}^{K} I_k$, and $E^c = \cup_{j=1}^{J} M_j$ be the unique representations of E and E^c in terms of elements of $\mathcal{I}$, respectively. If*

$$i_m(E) = \left[\Sigma_{k=1}^{K} a^{-}(I_k), 1 - \Sigma_{j=1}^{J} a^{-}(M_j)\right],$$

then $i_m : \mathcal{E} \rightarrow Int_{[0,1]}$ is an extension of $\mathcal{I}$ to $\mathcal{E}$ and is well-defined. In addition,

$$i_m(E) = \left[\inf\left\{\Pr(X) \in E \mid X \in \mathcal{M}\left(X^u, X^l\right)\right\}, \sup\left\{\Pr(X) \in E \mid X \in \mathcal{M}\left(X^u, X^l\right)\right\}\right].$$

Proof:
First assume $E = (a, b] \in \mathcal{I}$, then $E^c = (-\infty, a] \cup (b, \infty)$, so by the definition,
$i_m(E) = \left[\max\left\{F^l(b) - F^u(a), 0\right\}, 1 - \left(\max\left\{F^l(a) - F^u(-\infty), 0\right\} + \max\left\{F^l(\infty) - F^u(b), 0\right\}\right)\right]$
which matches the definition for i_m on $\mathcal{I}$. The other cases for $E \in \mathcal{I}$ are similar. Thus it is an extension. It is easy to show that it is well-defined, since the representation of any element in $\mathcal{E}$ in terms of the minimum number of elements of $\mathcal{I}$ is unique. In addition it is clear that

$$0 \leq \Sigma_{k=1}^{K} a^{-}(I_k),$$

and

$$1 - \Sigma_{j=1}^{J} a^{-}(M_j) \leq 1.$$

So we only need to show that

$$\Sigma_{k=1}^{K} a^{-}(I_k) \leq 1 - \Sigma_{j=1}^{J} a^{-}(M_j).$$

That is,

$$\Sigma_{k=1}^{K} a^{-}(I_k) + \Sigma_{j=1}^{J} a^{-}(M_j) \leq 1.$$

If we relabel the endpoints of all these intervals as $-\infty = c_1 < c_2 ... < c_N = \infty$, then

$$\begin{aligned}
\Sigma_{k=1}^{K} a^{-}(I_k) + \Sigma_{j=1}^{J} a^{-}(M_j) &= \Sigma_{n=1}^{N-1} \max\left\{F^l(c_{n+1}) - F^u(c_n), 0\right\} \\
&\leq \Sigma_{n=1}^{N-1} \max\left\{F^u(c_{n+1}) - F^u(c_n), 0\right\} = \Sigma_{n=1}^{N-1} F^u(c_{n+1}) - F^u(c_n) \\
&= 1.
\end{aligned}$$

Thus

$$\Sigma_{k=1}^{K} a^{-}(I_k) + \Sigma_{j=1}^{J} a^{-}(M_j) \leq 1.$$

For the last equation assume

$$E = \cup_{k=1}^{K} I_k = (-\infty, b_1] \cup (a_2, b_2] \cup ... \cup (a_K, b_K],$$

and

$$E^c = \cup_{j=1}^{J} M_j = \cup (b_1, a_2] \cup ... \cup (b_K, \infty).$$

We will show that

$$X \in \mathcal{M}\left(X^u, X^l\right) \Rightarrow \Pr(X \in E) \in i_m(E),$$

and there is an $X \in \mathcal{M}\left(X^u, X^l\right)$ for which $\Pr(X \in E) = a^{+}(E)$. Note first that

$$\begin{aligned}
\Sigma_{j=1}^{J} a^{-}(M_j) &= \Sigma_{k=1}^{K} \max\left\{F^l(a_{k+1}) - F^u(b_k), 0\right\} \\
&\leq \Sigma_{k=1}^{K} \max\left\{F(a_{k+1}) - F(b_k), 0\right\} \\
&= \Pr(E^c)
\end{aligned}$$

which gives both

$$\Pr(E) = 1 - \Pr(E^c) \le a^+(E),$$

and by replacing E with

$$E^c a^-(E) \le \Pr(E).$$

Next for $x \le a_2$ set

$$F(x) = \min\left(F^l(b_1), F^u(x)\right),$$

and for $a_2 < x \le b_2$ set

$$F(x) = \min\left(F^l(b_2), \left(\frac{x - a_2}{b_2 - a_2}\right) F^u(x) + \left(\frac{b_2 - x}{b_2 - a_2}\right) F^u(x)\right).$$

Continuing in this way gives a cumulative distribution function for which

$$\Pr(E^c) = \Sigma_{j=1}^{J} a^-(M_j),$$

and

$$\Pr(E) = 1 - \Sigma_{j=1}^{J} a^-(M_j).$$

The other bound is similarly derived.□

The family of sets, $\mathcal{E}$,is a ring of sets generating the Borel sets $\mathcal{B}$. For an arbitrary Borel set S, then it is clear that

$$\Pr(S) \in \left[\sup\left\{a^-(E) \mid E \subseteq S, E \in \mathcal{E}\right\}, \inf\left\{a^+(F) \mid S \subseteq F, F \in \mathcal{E}\right\}\right]$$

This prompts the following:

Proposition 57 *Let $i_m :.\mathcal{B} \to [0,1]$ be defined by*

$$i_m(A) = \left[\sup\left\{a^-(E) \mid E \subseteq A, E \in \mathcal{E}\right\}, \inf\left\{a^+(F) \mid A \subseteq F, F \in \mathcal{E}\right\}\right]$$

Then i_m is an extension from $\mathcal{E}$ to $\mathcal{B}$, and it is well-defined.

Proof:
The last property of proposition 56 insures it is an extension since, for example, if $E \subseteq F$ are two elements of $\mathcal{E}$ then $a^+(E) \le a^+(F)$ so

$$\inf\left\{a^+(F) \mid E \subseteq F, F \in \mathcal{E}\right\} = a^+(E).$$

Similarly it ensures that

$$\sup\left\{a^-(F) \mid F \subseteq E, F \in \mathcal{E}\right\} = a^-(E).$$

Next we show that i_m is well-defined. Proposition 56 shows that

$$\forall E \in \mathcal{E}, i_m(E) \subseteq [0,1].$$

Thus,

$$0 \leq \sup \left\{ a^{-}(E) \mid E \subseteq S \right\},$$

and

$$\inf \left\{ a^{+}(E) \mid S \subseteq E \right\} \leq 1.$$

We also have

$$\sup \left\{ a^{-}(E) \mid E \subseteq S \right\} \leq \inf \left\{ a^{+}(F) \mid S \subseteq F \right\}.$$

□

Proposition 58 *The function $i_m : \mathcal{B} \to Int_{[o,1]}$ defines an F-probability field on the Borel sets and*

$$i_m(B) = \left[\inf \left\{ \Pr(X \in B) \mid X \in \mathcal{M}\left(X^u, X^l\right) \right\}, \sup \left\{ \Pr(X \in B) \mid X \in \mathcal{M}\left(X^u, X^l\right) \right\} \right],$$

that is, $\mathcal{M}\left(X^u, X^l\right)$ defines a structure.

Proof: Clear.

2.3.6 Interval-Valued Integration, Extension and Independence of $F-probabilities$

In this section we define three key concepts needed for the application of IVPMs to mathematical programing problems, integration, extension, and independence.

Definition 59 *Given F-probability field $\mathcal{R} = (S, \mathcal{A}, i_m)$ and an integrable function $f : S \to R$ we define:*

$$\int_A f(x)\, di_m = \left[\inf_{p \in \mathcal{M}(\mathcal{R})} \int_A f(x)\, dp, \sup_{p \in \mathcal{M}(\mathcal{R})} \int_A f(x)\, dp \right].$$

We make the following observations which are useful in actual evaluation.

It is easy to see that if f is an $\mathcal{A}-$measurable simple function such that $f(x) = \begin{cases} y & x \in A \\ 0 & x \notin A \end{cases}$, with $A \in \mathcal{A}$, then

$$\int_A f(x)\, di_m = y i_m(A).$$

Further, if f is a simple function taking values $\{y_k \mid k \in K\}$ on an at most countable collection of disjoint measurable sets $\{A_k \mid k \in K\}$ that is, $f(x) = \begin{cases} y_k & x \in A_k \\ 0 & x \notin A \end{cases}$, where $A = \cup_{k \in K} A_k$, then

$$\int_A f(x)\, di_m = \left[a^{-}\left(\int_A f(x)\, di_m \right), a^{+}\left(\int_A f(x)\, di_m \right) \right],$$

where

$$a^{+}\left(\int_{A} f(x)\, di_m\right) = \sup\left\{\Sigma_{k\in K} y_k \Pr(A_k) \mid \Pr \in \mathcal{M}(\mathcal{R})\right\},$$

and

$$a^{-}\left(\int_{A} f(x)\, di_m\right) = \inf\left\{\Sigma_{k\in K} y_k \Pr(A_k) \mid \Pr \in \mathcal{M}(\mathcal{R})\right\}. \tag{2.7}$$

Note that these can be evaluated by solving two linear programing problems, since $\Pr \in \mathcal{M}(\mathcal{R})$ implies that $\Sigma_{k\in K} \Pr(A) = 1$, and $\Pr(\cup_{l\in L\subset K} A_l) \in i_m(\cup_{l\in L\subset K} A_l)$ so the problem may be tractable. In general, if f is an integrable function, and $\{f_k\}$ is a sequence of simple functions converging uniformly to f, then the integral with respect to f can be determined by noting that

$$\int_{A} f(x)\, di_m = \lim_{k\to\infty} \int_{A} f_k(x)\, di_m,$$

where

$$\lim_{k\to\infty} \int_{A} f_k(x)\, di_m = \left[\lim_{k\to\infty} a^{-}\left(\int_{A} f_k(x)\, di_m\right), \lim_{k\to\infty} a^{+}\left(\int_{A} f_k(x)\, di_m\right)\right],$$

provided the limits exist.

Example 60 *Consider the IVPM constructed from the interval* $[a,b]$. *Then* $\int_R x di_m = [a,b]$, *that is, the interval-valued expected value is the interval itself.*

Definition 61 *Let* $\mathcal{R} = (S, \mathcal{A}, i_m)$ *be an F-probability field and* $f : S \to T$ *a measurable function from measurable space* $(S, \mathcal{A})$ *to measurable space* $(T, \mathcal{B})$. *Then the F-probability* $(T, \mathcal{B}, l_m)$ *defined by*

$$l_m(B) = \left[\inf\left\{\Pr\left(f^{-1}(B)\right) \mid \Pr \in \mathcal{M}(\mathcal{R})\right\}, \sup\left\{\Pr\left(f^{-1}(B)\right) \mid \Pr \in \mathcal{M}(\mathcal{R})\right\}\right] \tag{2.8}$$

is called the extension of the R-probability field to $(T, \mathcal{B})$.

That this defines an $F-probability$ field is clear from our earlier observation. In addition, it's easy to see that this definition is equivalent to setting

$$l_m(A) = i_m\left(f^{-1}(A)\right),$$

which allows for evaluation using the techniques described earlier.

We now address the combination of IVPMs when the variables are independent. We do not address the situation when dependencies may be involved. Given measurable spaces $(S, \mathcal{A})$ and $(T, \mathcal{B})$ and the product space$(S \times T, \mathcal{A} \times \mathcal{B})$, assume $i_{X\times Y}$ is an IVPM on $\mathcal{A} \times \mathcal{B}$. Call i_X and i_Y defined by $i_X(A) = i_{X\times Y}(AxT)$ and $i_Y(B) = i_{X\times Y}(S \times B)$ the marginals of $i_{X\times Y}$. The marginals, i_X and i_Y, are IVPMs.

Definition 62 *Call the marginal IVPMs independent if and only if* $i_{X\times Y}(A\times B) = i_X(A)\, i_Y(B)\ \forall A, B \subseteq S$.

Definition 63 *Let* $\mathcal{R} = (S, \mathcal{A}, i_X)$ *and* $\mathcal{Q} = (T, \mathcal{B}, l_Y)$ *be F-probability fields representing uncertain random variables* X *and* Y. *We define the F-probability field* $(S\times T, \mathcal{A}\times\mathcal{B}, i_{X\times Y})$ *by defining*

$$i^{+}_{X\times Y}(A\times B) = \sup\left\{ \Pr_X(B) \Pr_Y(B) \mid \Pr_X \in \mathcal{M}(\mathcal{R}), \Pr_Y \in \mathcal{M}(\mathcal{Q}) \right\},$$

$$i^{-}_{X\times Y}(A\times B) = \inf\left\{ \Pr_X(B) \Pr_Y(B) \mid \Pr_X \in \mathcal{M}(\mathcal{R}), \Pr_Y \in \mathcal{M}(\mathcal{Q}) \right\},$$

where $(S\times T, \mathcal{A}\times\mathcal{B})$ *is the usual product of sigma algebra of sets*

It is clear from this definition that

$$i_{X\times Y}(AxB) \equiv i_X(A)\, i_Y(B)$$

for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Thus, if we have several uncertain parameters in a problem with the uncertainty characterized by IVPMs, and all are independent, we can form an IVPM for the product space by multiplication and use this IVPM.

2.3.7 Application to Optimization

An example application of IVPM is an optimization problem, the recourse problem. Suppose we wish to optimize $f(x,a)$ subject to $g(x,b) = 0$. Assume a and b are vectors of independent uncertain parameters, each with an associated IVPM. Assume the constraint can be violated at a known cost c so that the problem is to solve:

$$h(x,a,b) = \max\left(f(x,a) - c(g(x,b))\right).$$

Using the independence, form an IVPM for the product space, i_{axb}, for the joint distribution. Then calculate the interval-valued expected value with respect to this IVPM. The resulting interval-valued expected value is

$$\int_R h(x,a,b)\, di_{a\times b}.$$

To optimize over such a value requires an ordering of intervals. One such ordering is to use the midpoint of the interval on the principle that in the absence of additional data, the midpoint is the best estimate for the true value. Another possible ordering is to use risk/return multi-objective decision making. For example, determine functions $u : \mathbb{R}^2 \to \mathbb{R}$ and $v : Int_{\mathbb{R}} \to \mathbb{R}^2$ by setting, for any interval $I = [a,b]$, $v(I) = \left(\frac{a+b}{2}, b-a\right)$. Thus, v gives the midpoint and width of an interval. Then u would measure the decision makes preference for one interval over another considering both its midpoint and width (a risk measure). The optimization problem becomes

$$\max_x u\left(v\left(\int_R h\left(x,a,b\right)di_{a\times b}\right)\right).$$

2.4 Extension Principles for Generalized Distribution

One extension principle associated with general distributions is presented above (2.8), perhaps the only one that is able to deal with the complete set of uncertainty that is of interest to this monograph. The current research seems to lack a generalized extension principle except what is presented [68] and which is given here. The reason for this is not only the complexity of the problem but a lack of a general theory that captures the broad spectrum of uncertainty distribution, which IVPMs of [164] and [68] do.

References

[1] O. Aberth (1978), "A method for exact computation with rational numbers," *Journal of Computational and Applied Mathematics,* **4**, pp. 285-288,

[2] O. Aberth (1988). *Precise Numerical Analysis.* Wm. C. Brown Publishers, Dubuque, Iowa.

[3] G. Alefeld and J. Herzberger (1983). *Introduction to Interval Computations.* Academic Press, New York.

[4] G. Alefeld (1990), "Enclosure methods," in *Computer Arithmetic and Self-Validating Numerical Methods,* C. Ullrich (editor), Academic Press, Boston, pp. 55-72.

[5] M. Anile, S. Deodato and G. Privitera (1995), "Implementing fuzzy arithmetic," *Fuzzy Sets and Systems*, **72:2**, pp. 239-250.

[6] Archimedes of Siracusa, "On the measurement of the circle," In: Thomas L. Heath (ed.), *The Works of Archimedes*, Cambridge University Press, Cambridge, 1897; Dover edition, 1953.

[7] O. Artbauer (1988), "Application of interval, statistical, and fuzzy methods to the evaluation of measurements," *Metrologia*, **25**, pp. 81-86.

[8] J-P. Audin and H. Frankkowska (1990). *Set-Valued Analysis.* Birkhäuser, Boston.

[9] C. Baudrit, D. Dubois, and H. Fargier (2005), "Propagation of uncertainty involving imprecision and randomness," *ISIPTA 2005*, pp. 31-40.

[10] R.E. Bellman and L. A. Zadeh (1970) "Decision-making in a fuzzy environment," *Management Science*, Serial B 17, 141-164.

[11] D. Berleant (1993), "Automatically verified reasoning with both intervals and probability density functions," *Interval Computations,* **2**, pp. 48-70.

[12] D. Berleant and C. Goodman-Strauss (1998), "Bounding the results of arithmetic operations on random variables of unknown dependency using intervals," *Reliable Computing,* **4**, pp. 147-165.

[13] D. Berleant and J. Zhang, (2004), "Representation and problem solving with distribution envelope determination (DEnv)," *Reliability Engineering and Systems Safety,* **85:**(1-3), pp. 153-168.

[14] K. Böhmer, P. Hemker and H. J. Stetter (1984), "The defect correction approach," *Computing Supplement,* **5**, pp. 1-32.

[15] J. Bondia, A. Sala, and M. Sáinz (2006), "Modal fuzzy quantities and applications to control," in K. Demirli and A. Akgunduz (editors), *2006 Conference of the North American Fuzzy Information Processing Society, June 3-6, 2006, Montréal, Canada*.

[16] F. Bornemann, D. Laurie, S. Wagon, and J. Waldvogel (2004). *The SIAM 100-Digit Challenge: A Study in High-Accuracy Numerical Computing.* SIAM, Philadelphia, PA.

[17] J. J. Buckley (1989), "A generalized extension principle," *Fuzzy Sets and Systems*, **33**, pp.241-242.

[18] G. F. Corliss (2004), "Tutorial on validated scientific computing using interval analysis," *PARA'04 Workshop on State-of-the-Art Computing*, Technical University of Denmark, June 20-23, 2004 (see http://www.eng.mu.edu/corlissg/PARA04/READ_ME.html).

[19] J. C. Burkill (1924), "Functions of intervals," *Proceedings of the London Mathematical Society,* **22,** pp. 375-446.

[20] D. Daney, Y. Papegay and A. Neumaier (2004), "Interval methods for certification of the kinematic calibration of parallel robots," *in: Proc. 2004 IEEE Int. Conf. Robotics Automation*, New Orleans, LA, April 2004, pp. 191-198 .

[21] B. Davies (2005), "Whither mathematics?," *Notices of the AMS,* **52**:11, December 2005, pp. 1350-1356.

[22] A. Deif (1986). *Sensitivity Analysis in Linear Systems,* Springer-Verlag, New York.

[23] A. P. Demster (1967), "Upper and lower probabilities induced by multivalued mapping," *Annals of Mathematical Statistics,* **38**, pp. 325-339.

[24] M. A. H. Dempster (1969), "Distributions in interval and linear programming," in *Topics in Interval Analysis* (edited by E. R. Hansen), Oxford Press, pp. 107-127.

[25] M. A. H. Dempster (1974), "An application of quantile arithmetic to the distribution problem in stochastic linear programming," *Bulletin of the Institute of Mathematics and its Applications,* **10**, pp. 186-194.

[26] D. Dubois and H. Prade (1977), "Le flou, mécédonka?", *Tech. Rep. C.E.R.T.-D.E.R.A.,* Toulouse, Avril 1977.

[27] D. Dubois and H. Prade (1978), "Fuzzy algebra, analysis, logics," *Technical Report, Purdue University, N$^{\underline{o}}$ TR-EE 78-13*, March 1978.

[28] D. Dubois and H. Prade (1978), "Operations on fuzzy numbers," *International Journal of Systems Science,* **9**:6, pp. 613-626.

[29] D. Dubois and H. Prade (1979), "Fuzzy real algebra: some results," *Fuzzy Sets and Systems*, **2:**4, pp. 327-348.

[30] D. Dubois and H. Prade (1980). *Fuzzy Sets and Systems: Theory and Applications.* Academic Press, New York.

[31] D. Dubois and H. Prade (1981), "Additions of interactive fuzzy numbers," *IEEE Trans. on Automatic Control,* **26**:4, pp. 926-936.

[32] D. Dubois and H. Prade (1987), "Fuzzy numbers: An overview," Tech. Rep. n^{o}219, (L.S.I., Univ. Paul Sabatier, Toulouse, France), also in Chapter 1 of *Analysis of Fuzzy Information, Volume 1, Mathematics and Logic*, James C. Bezdek editor, CRC Press, Inc., 1987, pp.3-39.

[33] D. Dubois and H. Prade (1987), "Evidence theory and interval analysis," *Second IFSA Congress, Tokyo, July 20-25, 1987,* pp. 502-505.

[34] D. Dubois and H. Prade (1987), *Special Issue on Fuzzy Numbers: Fuzzy Sets and System,* **24**:3.

[35] D. Dubois and H. Prade (1988). *Possibility Theory an Approach to Computerized Processing of Uncertainty.* Plenum Press, New York.

[36] D. Dubois and H. Prade (1991), "Random sets and fuzzy interval analysis," *Fuzzy Sets and Systems*, **42**, pp.87-101.

[37] D. Dubois and H. Prade (2000), Editors. *Fundamentals of Fuzzy Sets.* Kluwer Academic Press, 2000.

[38] D. Dubois and H. Prade (2005), "Fuzzy elements in a fuzzy set," in *Proceedings of the 10th International Fuzzy System Association (IFSA) Congress, Beijing* 2005, pp. 55-60.

[39] D. Dubois, E. Kerre, R. Mesiar and H. Prade (2000), Chapter 10: "Fuzzy Interval Analysis," in D. Dubois and H. Prade (2000), Editors. *Fundamentals of Fuzzy Sets.* Kluwer Academic Press, 2000.

[40] D. Dubois, S. Moral and H. Prade (1997), "Semantics for possibility theory based on likelihoods," *J. Math. Anal. Appl.*, **205**, pp. 359-380.

[41] P. S. Dwayer (1951). *Linear Computations.* Wiley, New York.

[42] J. S. Ely (1993), "The VPI software package for variable precision interval arithmetic," *Interval Computation,* **2:**2, pp. 135-153.

[43] J. S. Ely and G. R. Baker (1993), "High-precision calculations of vortex sheet motion," *Journal of Computational Physics* **111**, pp. 275-281.

[44] S. Ferson (1995), "Quality assurance for Monte Carlo risk assessment," *IEEE Proceedings of ISUMA-NAFIPS'95.*, pp. 14-19.

[45] S. Ferson (1996), "Probability bounds analysis software," *Computing in Environmental Resource Management: Proceedings of a Speciality Conference,* Research Triangle Park, NC, December 2-4, 1996, Pittsburgh, PA, Air and Waste Management Association, pp. 669-677.

[46] S. Ferson, L. Ginzburg, and R. Akçakaya, "Whereof one cannot speak: when input distributions are unknown," (see http://www.ramas.com).

[47] S. Ferson (2002). *RAMAS Risk Calc 4.0 Software : Risk Assessment with Uncertain Numbers*. Lewis Publishers, Boca Raton, Florida.

[48] S. Ferson, V. Kreinovich L. R. Ginzburg, K. Sentz, and D. S. Myers (2003). *Constructing Probability Boxes and Dempster-Shafer Structures*. Sandia National Laboratories, Technical Report SAND2002-4015, Albuquerque, New Mexico.

[49] S. Ferson and V. Kreinovich, "Combining interval and probabilistic uncertainty: Foundations, algorithms, challenges - an overview," *World-Wide Web,* 2005.

[50] P. C. Fischer, "Automatic Propagated and Round-off Error Analysis", *Proceedings of the 13th national meeting of the Association for Computing Machinery*, June 1958 (http://portal.acm.org/citation.cfm?id=610971).

[51] J. Fortin, D. Dubois, H. Fargier, "Gradual numbers and their application to fuzzy interval analysis," (to appear in *IEEE Transactions on Fuzzy Systems).*

[52] M. J. Frank, R. B. Nelsen, and B. Schweizer (1987), "Best-possible bounds for the distribution of a sum - A problem of Kolmogorov," *Probability Theory and Related Fields,* **74**, pp. 199-211.

[53] R. Fullér and T. Keresztfalvi (1990), "On generalization of Nguyen's theorem," *Fuzzy Sets and Systems*, **41**, pp. 371-374.

[54] E. Gardeñes, H. Mielgo, and A. Trepat (1985), "Modal intervals: Reasons and ground semantics," in K. Nickel (editor), *Interval mathematics 1985 : Proceedings of the international symposium, Freiburg i. Br., Federal Republic of Germany, September 23-26, 1985,* pp. 27-35.

[55] K. G. Guderley and C. L. Keller (1972), "A basic theorem in the computation of ellipsoidal error bounds," *Numerische Mathematik,* **19**(3), pp. 218-229.

[56] T. C. Hales (2000), "Cannonballs and honeycombs," *Notices of the American Mathematical Society* **47,** pp. 440-449.

[57] E. R. Hansen (1969) editor. *Topics in Interval Analysis*, Oxford Press.

[58] E. R. Hansen (1975), "A generalized interval arithmetic" in K. Nickel (editor). *Interval Mathematics.* Lecture Notes in Computer Science 29, Springer Verlag, New York, pp. 7-18.

[59] E. R. Hansen (1978), "Interval forms of Newton's method," *Computing* **20**, pp.153-163.

[60] E. R. Hansen (1980), "Global optimisation using interval analysis - the multidimentional case," *Numerische Mathematik,* **34**, pp. 247-270.

[61] E. R. Hansen (1992). *Global Optimization Using Interval Arithmetic.* Marcel Dekker, Inc., New York.

[62] E. R. Hansen (2001), "Publications related to early interval work of R. E. Moore", August 13, 2001, (see http://interval.louisiana.edu/Moores_early_papers/bibliography.html).

[63] M. Hanss (2005). *Applied Fuzzy Arithmetic,* Springer Verlag, Berlin.

[64] T. Hickey, Q. Ju and M. H. van Emden (2001), "Interval arithmetic: From principles to implementation," *Journal of the ACM,* **48**:5, pp. 1038-1068, September 2001.

[65] K. Jahn (1980), "The importance of 3-valued notions for interval mathematics," in *Interval Mathematics 1980* (edited by Karl E. Nickel), Academic Press, New York, pp. 75-98.

[66] K. D. Jamison and W. A. Lodwick (2002) "The construction of consistent possibility and necessity measures", *Fuzzy Sets and Systems* **132:1**, pp. 1-10.

[67] K. D. Jamison and W. A. Lodwick, "Interval-valued probability measures," *UCD/CCM Report No. 213,* March 2004.

[68] K. D. Jamison and W. A. Lodwick, "Interval-valued probability in the analysis of problems that contain a mixture of fuzzy, possibilistic and interval uncertainty" (submitted to *International Journal of Approximate Reasoning*)

[69] K. D. Jamison, W. A. Lodwick and M. Kawai, "A simple closed form estimation for the cumulative distribution function of a monotone function of random variables," *UCD/CCM Report No. 187,* May 2002.

[70] L. Jaulin (2001), "Path planning using intervals and graphs," *Reliable Computing* **7**(1), pp. 1-15.

[71] L. Jaulin, M. Kieffer, O. Didrit and E. Walter (2001). *Applied Interval Analysis.* Springer, New York.

[72] W. M. Kahan (1968), "Circumscribing an ellipsoid about the intersection of two ellipsoids," *Canadian Mathematical Bulliten* **11:3**, pp. 437-441.

[73] W. M. Kahan (1968). *A More Complete Interval Arithmetic.* Lecture Notes for a Summer Course at University of Michigan.

[74] S. Kaplan (1981), "On the method of discrete probability distributions in risk and reliability calculations - Applications to seismic risk assessment," *Journal of Risk*, **1**:3, pp. 189-196.

[75] E. Kaucher (1973), "Über metrische und algebraische Eigenschaften eiginger beim numerischen Rechnen auftretender Räume," Ph.D. Thesis, Univ. Karlsruhe.

[76] E. Kaucher (1980), "Interval analysis in the extended space $I\mathbb{R}$," *Computing, Suppl.* **2**, pp. 33-49.

[77] E. Kaucher and S. M. Rump, "E-methods for fixed point equations f(x)=x," *Computing* **28**, pp. 31-42.

[78] A. Kaufmann and M. M. Gupta. *Introduction to Fuzzy Arithmetic - Theory and Applications,* Van Nostrand Reinhold, 1985.

[79] M. Kawai (2006), "Discrete clouds," (seminar presentation April 28, 2006, Department of Mathematical Sciences of University of Colorado at Denver).

[80] U. Kaymak and J. M. Sousa, "Weighting of Constraints in Fuzzy Optimization," *Constraints*, **8**, 2003, pp. 61-78.

[81] R. B. Kearfott (1996). *Rigorous Global Search: Continuous Problem.* Kluwer Academic Publishers, Boston.

[82] R. B. Kearfott (1996), "Interval computations: Introduction, uses, and resources," *Euromath Bulletin* **2**(1), pp. 95-112.

[83] D. Klaua (1969), "Partielle Mengen und Zhlen," *Mtber. Dt. Akad. Wiss.* **11**, pp. 585-599.

[84] G. J. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic.* Prentice Hall, Upper Saddle River, NJ, 1995.

[85] G. J. Klir (1997), "Fuzzy arithmetic with requisite constraints," *Fuzzy Sets and Systems,* **91**:2, pp. 165-175.

[86] G. J. Klir (1998), "Chapter 1: The Role of Constrained Fuzzy Arithmetic in Engineering," in Ayyub, Bilal and Gupta, Madan (editors), *Uncertainty Analysis in Engineering and Sciences: Fuzzy Logic, Statistics, and Neural Network Approach*, Kluwer Academic Publishers.

[87] A. N. Kolmogorov (1941), "Confidence limits for an unknown distribution function," *Annuals of Mathematics Statistics*, **12**, pp. 461-463.

[88] A. N. Kolmogorov and S. V. Fomin (1975). *Introductory Real Analysis,* Dover Publications, Inc., New York.

[89] P. Korenerup and D. W. Matula (1983), "Finite precision rational arithmetic: An arithmetic unit," *IEEE Transactions on Computers,* **32**:4, pp. 378-388.

[90] U. Kulisch (1971), "An axiomatic approach to rounded computations," *Numerische Mathematik.***18**, pp. 1-17.

[91] U. Kulisch and W. L. Miranker. *Computer Arithmetic in Theory and Practice.* Academic Press, New York, 1981.

[92] U. Kulisch and W. L. Miranker (1986), "The arithmetic of the digital computer: A new approach," *SIAM Review* **28:1**, pp. 1-40.

[93] T. Y. Lin (2005), "A function theoretic view of fuzzy sets: New extension principle," in D. Filev and H. Ying (editors), *Proceedings of NAFIPS05* .

[94] W. A. Lodwick (1989), "Constraint propagation, relational arithmetic in AI systems and mathematical programs," *Annals of Operations Research* **21**, pp. 143-148.

[95] W. A. Lodwick (1990), "Analysis of structure in fuzzy linear programs," *Fuzzy Sets and Systems* **38**, pp. 15-26.

[96] W. A. Lodwick (1999), "Constrained interval arithmetic," *CCM Report* **138**.

[97] W. A. Lodwick and K. D. Jamison, editors (2003). Special Issue on the Interfaces Between Fuzzy Set Theory and Interval Analysis, *Fuzzy Sets and Systems* , **135:1**, April 2003.

[98] W. A. Lodwick and K. D. Jamison, "Estimating and validating the cumulative distribution of a function of random variable: Toward the development of distribution arithmetic," *Reliable Computing* **9**:127-141, 2003.

[99] W. A. Lodwick, editor (2004). Special Issue on Linkages Between Interval Analysis and Fuzzy Set Theory, *Reliable Computing*

[100] W. A. Lodwick and K. D. Jamison, "Interval-valued probability in the analysis of problems that contain a mixture of fuzzy, possibilistic and interval uncertainty," in K. Demirli and A. Akgunduz (editors), *2006 Conference of the North American Fuzzy Information Processing Society, June 3-6, 2006, Montréal, Canada.*

[101] K. Makino and M. Berz (2003), "Taylor models and other validated functional inclusion methods," *International Journal of Pure and Applied Mathematics* **4**:4, pp. 379-456.

[102] K. Makino and M. Berz (2005), "Verified global optimization with Taylor model based range bounders," *Transactions on Computers* **11**:4 (November 2005), pp. 1611-1618.

[103] G. Mayer (1995), "Epsilon-inflation in verification algorithms," *Journal of Computational and Applied Mathematics* **60**, pp. 147-169.

[104] G. Mayer (1995), "Success in epsilon-inflation," in G. Alefeld and B. Lang (editors), *Scientific Computing and Validated Numerics:*

Proceedings of the International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics SCAN-95 held in Wuppertal, Germany, September 26-29, 1995, Akademie Verlag, Berlin, May 1996, pp. 98-104.

[105] M. Mizumoto and K. Tanaka (1976), "The four operations of arithmetic on fuzzy numbers," *Syst. Compt. Control,* **7**:5, pp. 73-81.

[106] A. S. Moore, "Interval risk analysis of real estate investment: A non-Monte Carlo approach," *Freiburger Intervall-Berichte* 85/3 (1985), pp. 23-49.

[107] R. E. Moore (1959), "Automatic error analysis in digital computation," *Technical Report LMSD-48421,* Lockheed Missile and Space Division, Sunnyvale, CA, (see http://interval.louisiana.edu/Moores_early_papers/bibliography.html).

[108] R. E. Moore and C. T. Yang (1959), "Interval analysis I," *Technical Report Space Div. Report LMSD285875, Lockheed Missiles and Space Co.*, 1959.

[109] R. E. Moore, W. Strother and C. T. Yang (1960), "Interval integrals", *Technical Report Space Div. Report LMSD703073, Lockheed Missiles and Space Co.*, 1960.

[110] R. E. Moore (1962), "Interval arithmetic and automatic error analysis in digital computing," Ph.D. Thesis, Stanford University, Stanford, California, Published as Applied Mathematics and Statistics Laboratories Technical Report No. 25.November 15, 1962 (see http://interval.louisiana.edu/Moores_early_papers/bibliography.html).

[111] R. E. Moore (1965), "The automatic analysis and control of error in digital computing based on the use of interval numbers," In L. B. Rall, (editor), *Error in Digital Computation, Vol. I*, Chapter 2, pages 61-130. John Wiley and Sons, Inc., New York.

[112] R. E. Moore (1966). *Interval Analysis.* Prentice-Hall, Englewood Cliffs N. J.

[113] R. E. Moore (1979). *Methods and Applications of Interval Analysis.* SIAM, Philadelphia, PA.

[114] R. E. Moore (1984), *Risk Analysis without Monte Carlo methods,* Freiburger Intervall-Berichte **1** pp. 1-48.

[115] R. E. Moore (1985). *Computational Functional Analysis.* Ellis Horwood, Chichester, England.

[116] R. E. Moore (1992), "Computing to arbitrary accuracy," in C. Bresinski and U. Kulisch (editors). *Computational and Applied Mathematics I: Algorithms and Theory* pp. 327-336, North-Holland, Amsterdam.

[117] R. E. Moore (1999), "The dawning," *Reliable Computing*, **5**:4, November 1999, pp.423-424.

[118] R. E. Moore and W. A. Lodwick, (2003), "Interval Analysis and Fuzzy Set Theory," *Fuzzy Sets and Systems*, **135**:1, pp. 5-9.

[119] S. Nahmias (1978), "Fuzzy variable," *Fuzzy Sets and Systems*, **1**, pp. 97-110.

[120] R. B. Nelsen (1995), "Copulas, characterization, correlation, and counterexamples," *Mathematics Magazine,* **68**:3, pp. 193-198.

[121] C. V. Negoita and D. A. Ralescu (1975). *Applications of Fuzzy Sets to Systems Analysis.* Birkhauser, Boston, MA.

[122] A. Neumaier (1990). *Interval Methods for Systems of Equations.* Cambridge Press, Cambridge.

[123] A. Neumaier (1993), "The wrapping effect, ellipsoid arithmetic, stability and confidence regions," *Computing Supplementum,* **9**, pp. 175-190.

[124] A. Neumaier (2003), "Fuzzy modeling in terms of surprise," *Fuzzy Sets and Systems,* **135**:1**,** pp. 21-38.

[125] A. Neumaier (2004), "Clouds, fuzzy sets and probability intervals," *Reliable Computing* **10** , pp. 249-272.

[126] A. Neumaier (2004), "Complete search in continuous global optimization and constraint satisfaction," in A. Iserles (editor). *Acta Numerica 2004.* Cambridge University Press, pp. 271-369

[127] A. Neumaier (2005), "Structure of clouds," (submitted - downloadable http://www.mat.univie.ac.at/~neum/papers.html).

[128] H. T. Nguyen (1978), "A note on the extension principle for fuzzy sets," *Journal of Mathematical Analysis and Applications*, **64**, pp. 369-380.

[129] K. Nickel (1969), "Triplex-Algol and its applications," in *Topics in Interval Analysis* (edited by E. R. Hansen), Oxford Press, pp. 10-24.

[130] G. Olsen, "The inverse probability method: An interval algorithm for measuring uncertainty," Masters Thesis, University of Colorado at Denver, Department of Mathematics, 2005.

[131] H.-J. Ortolf (1969), "Eine Verallgemeinerung der Intervallarithmetik," *Geselschaft fuer Mathematik und Datenverarbeitung,* Bonn **11**, pp. 1-71.

[132] G. M. Phillips (1981), "Archimedes the numerical analyst," *American Mathematical Monthly,* pp. 165-169.

[133] E. D. Popova (1998), (see http://www.math.bas.bg/~epopova/directed.html).

[134] J. D. Pryce and G. F. Corliss (2006), "Interval arithmetic with containment sets," (submitted for publication - under review).

[135] J. Ramik (1986), "Extension principle in fuzzy optimization," *Fuzzy Sets and Systems*, **19**, pp. 29-35.

[136] Ratschek, H. and J. Rokne (1988). *New Computer Methods for Global Optimization.* Ellis Horwood, Chichester, England.

[137] N. Revol and F. Rouillier (2005), "Motivations for an arbitrary precision interval arithmetic and the MPFI library," *Reliable Computing*, **11**:4, pp. 275–290.

[138] B. Russell (1924), "Vagueness," *Australian Journal of Philosophy,* **1**, pp. 84-92.

[139] M. A. Sáinz (2001), "Modal intervals," *Reliable Computing* **7**:2, pp. 77-111.

[140] M. J. Schulte and E. E. Swartzlander Jr.(200), "A family of variable-precision interval processors," *IEEE Transactions on Computers,* **49**:5, pp. 387-397.

[141] G. Shafer (1976). *A Mathematical Theory of Evidence.* Princeton University Press, Princeton, NJ.

[142] G. Shafer (1987), "Belief functions and possibility measures," in Chapter 3 of *Analysis of Fuzzy Information, Volume 1, Mathematics and Logic*, James C. Bezdek editor, CRC Press, Inc., pp. 51-84.

[143] J. R. Shewchuk (2002), "Delaunay refinement algorithms for triangular mesh generation," *Computational Geometry: Theory and Applications* **22**:1-3, pp. 21-74.

[144] S. Smale (1998), "Mathematical problems for the next century," *Mathematical Intelligencer* **20**:2, pp. 7-15.

[145] Stolfi, J., Andrade, M. V. A., Comba, J. L. D., Van Iwaarden, R. "Affine arithmetic: A correlation-sensitive variant of interval arithmetic," (see http://www.dcc.unicamp.br/~stolfi/EXPORT/projects/affine-arith)

[146] M. D. Springer (1979). *The Algebra of Random Variables.* Wiley & Sons, New York.

[147] H. J. Stetter (1978), "The defect correction principle and discretization methods," *Numerische Mathematiche,* **29**, pp. 425-443.

[148] W. Strother, "Continuity for multi-valued functions and some applications to topology," (doctoral dissertation) Tulane University, 1952.

[149] W. Strother (1955), "Fixed points, fixed sets, and m-retracts," *Duke Mathematical Journal* **22**:4, pp. 551-556

[150] W. Strother (1958), "Continuous multi-valued functions," *Boletim da Sociedade de Matematica de São Paulo*, Vol. 10, 1958, pp. 87-120.

[151] T. Sunaga (1958), "Theory of an interval algebra and its application to numerical analysis," *RAAG Memoirs*, **2**, pp. 547-564 (can be downloaded from http://www.cs.utep.edu/interval-comp/early.html).

[152] H. Tanaka, T. Okuda, and K. Asai (1974), "On fuzzy mathematical programming," *Journal of Cybernetics,* **3**:4, pp. 37-46.

[153] F. Tonon (2004), "On the use of random set theory to bracket the results of Monte Carlo simulations," *Reliable Computing* **10**, pp.107-137.

[154] F. Tonon (2004), "Aggregation of evidence from random and fuzzy sets," *Zaam Z. Agnew. Math. Mech.* **84**:10-11, pp.700-709.

[155] W. Tucker (2002), "A rigorous ODE solver and Smale's 14^{th} problem," *Foundation of Computational Mathematics* **2**, pp. 53-117.

[156] J. A. Tupper (1996), "Graphing equations with generalized interval arithmetic," Ph.D. Thesis, University of Toronto.

[157] P. Walley (1991). *Statistical Reasoning with Imprecise Probabilities,* Chapman & Hall, London.

[158] P. Walley (1999), "Towards a unified theory of imprecise probability," in Gert De Cooman, Fabio Gagliardi Cozman, Serafín Moral, Peter Walley (editors), *Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications,* held at the Conference Center "Het Pand" of the Universiteit Gent, Ghent, Belgium, 29 June - 2 July 1999 ISIPTA 1999.

[159] G. W. Walster (1998), "The extended real interval system," (personal copy from the author).

[160] Z. Wang and G. J. Klir, *Fuzzy Measure Theory*, Plenum Press, New York, 1992.

[161] M. Warmus (1956), "Calculus of approximations," *Bulliten de l'Académié Polonaise de Sciences* Cl. III, No. 4, pp. 253-259 (can be downloaded from http://www.cs.utep.edu/interval-comp/early.html).

[162] M. Warmus (1961), "Approximations and inequalities in the calculus of approximations. Classification of approximate numbers," *Bulliten de l'Académié Polonaise de Sciences* Cl. IX, No. 4, pp. 241-245 (can be downloaded from http://www.cs.utep.edu/interval-comp/early.html)

[163] K. Weichselberger (1996), "Interval probability on finite sample spaces," in H. Rieder (editor). *Robust Statistics, Data Analysis and Computer-Intensive Methods,* Springer, New York, pp. 391-409.

[164] K. Weichselberger (2000), "The theory of interval-probability as a unifying concept for uncertainty," *International Journal of Approximate Reasoning* **24,** pp. 149-170.

[165] R. C. Williamson and T. Downs (1990), "Probabilistic arithmetic I: Numerical methods for calculating convolutions and dependency bounds," *International Journal of Approximate Reasoning* **4**, pp. 89-158.

[166] R. C. Williams (1990), "Interval arithmetic and probability arithmetic," in C. Ullrich (editor), *Contributions to Computer Arithmetic and Self-Validating Numerical Methods,* J. C. Baltzer AG, IMACS, pp. 67-80.

[167] R. R. Yager (1981), "A procedure for ordering fuzzy subsets of the unit interval," *Information Sciences*, **24**, pp. 143-161.

[168] R. R. Yager (1986), "A characterization of the extension principle," *Fuzzy Sets and Systems*, **18**, pp. 205-217.

[169] R. C. Young (1931), "The algebra of many-valued quantities," *Mathematiche Annalen* Band **104**, pp. 260-290.

[170] L. A. Zadeh (1965), "Fuzzy Sets," *Information and Control*, **8**, pp. 338-353.

[171] L. A. Zadeh (1968), "Probability measures of fuzzy events," *Journal of Mathematical Analysis and Applications.*, **23**, pp. 421-427.

[172] L. A. Zadeh (1975), "The concept of a linguistic variable and its application to approximate reasoning," *Information Sciences,* Part I: **8**, pp. 199-249, Part II: **8**, pp. 301-357, Part III: **9**, pp. 43-80.

[173] L. A. Zadeh (1978), "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets and Systems*, **1**, pp. 3-28, 1978.

[174] L. A. Zadeh (2005), "Toward a generalized theory of uncertainty (GTU) - An outline," *Information Science*, **172**:1-2, pp. 1-40.