

Interval-valued probability in the analysis of problems containing a mixture of possibilistic, probabilistic, and interval uncertainty

Weldon A. Lodwick*, K. David Jamison

Department of Mathematical Sciences and Statistics, University of Colorado Denver, Campus Box 170, P.O. Box 173364, Denver, CO 80217-3364, USA

Available online 19 March 2008

Abstract

A simple definition of interval-valued probability measure (IVPM) is used and its implications are examined for problems in mathematical analysis. In particular, IVPs are constructed and then used to develop the extension of these measures in such a way that probability, possibility, clouds, and intervals fit within the context of IVP. With the extension principle, integration and product measures that are derived below, mathematical analysis applied to this new structures is enabled. Optimization will be the mathematical analysis used to illustrate the approaches that are developed.

© 2008 Elsevier B.V. All rights reserved.

Keywords: Possibility theory; Probability theory; Interval analysis; Interval-valued probability; Clouds

1. Introduction

This study examines a basis for linking various methods of uncertainty representation by showing that each uncertainty falls within the context of interval-valued probability (IVP). In doing so, it is seen how problems possessing mixed representations can be handled and solved. Starting with the definition of IVP as articulated by Weichselberger [13], a new method to construct interval-valued probability measures (IVPMs) and distributions is given. Moreover, in doing analysis, for example, optimization, functions of IVP distributions are required which in turn necessitates the extension principle for IVPs. How to compute the integral of IVPs and how to compute product measures of IVPs is worked out since these are key to the implementation of IVPs in mathematical analysis in general and optimization in particular.

The focus of this research, therefore, is the translation of the theory of IVPs as found in [13] into structures on which mathematical analysis can be performed. Thus, this development is distinct from [1–4] in that what is contained herein demonstrates how to do mathematical analysis with IVPs by developing the appropriate mechanisms. Implicit in our use of IVPs is the need to have a general theory which is able to account for the mixture of uncertainty that is of interest to our application. Thus, we review the basic definitions as found in [13], proving that IVPs are indeed a general enough theory to account for probabilistic, possibilistic, clouds, and interval uncertainty. Then, using IVPs so defined, a new simple way to construct IVPs is shown. Once IVPs are constructed, the extension principle for IVPs is worked out. Given the extension principle, integration and product measures are developed since

* Corresponding author.

E-mail address: Weldon.Lodwick@cudenver.edu (W.A. Lodwick).

our approach to optimization must be in a form capable of handling a mixture of uncertainty within one optimization model which requires, as will be seen, integration and product measures. Thus, what is new in this study, is the simple construction of IVPMs, the extension principle for IVPMs, IVP integration, IVP product measure, and the use of IVPs in optimization for the purpose of showing mathematical analysis on problems containing mixed uncertainty entities.

The second section of this paper defines, in a formal way, what an IVP is (as articulated by Weichselberger [13]). Weichselberger's approach begins with a set of probability measures and then defines an interval probability as a set function providing upper and lower bounds on the probabilities calculated from these measures. Weichselberger's F -probabilities are simply the tightest bounds possible for the set of interval probability measures. This definition is followed by demonstrating that various forms of uncertainty representation (probability, possibility, clouds, and intervals) all can be represented by such measures. The third section shows how IVPs may be constructed from upper and lower bounding cumulative distribution functions. Utilizing the construction, the fourth section contains the development of the extension principle for functions of uncertain variables represented by IVPs, as well as integration and product measures with respect to IVPs. These new approaches are useful and necessary in analyzing problems involving uncertainty represented by IVPs. The fifth section provides an application to a problem in optimization. The last section concludes the paper and discusses future directions.

Throughout this paper we will be primarily interested in IVP defined on the Borel sets on the real line and real-valued random variables. It is assumed that the reader is familiar with elementary measure theory [7].

2. IVPs—definitions and examples

This section begins by defining what is meant by an IVP. This generalization of a probability measure includes probability measures, possibility/necessity measures, intervals, and clouds (see [10]). The IVP set function may be thought of as a method for giving a partial representation for an unknown probability measure. Throughout, arithmetic operations involving set functions are in terms of interval arithmetic [9] and the set of all intervals contained in $[0, 1]$ is denoted, $Int_{[0,1]} \equiv \{[a, b] \mid 0 \leq a \leq b \leq 1\}$. Moreover, we use S to denote the universal set and a set of subsets of the universal set as $\mathcal{A} \subseteq S$. Here we are thinking of \mathcal{A} as being a set of subsets that has a structure imposed on it as will be seen in our development.

Definition 1 (Weichselberger [13]). Given a measurable space (S, \mathcal{A}) , an interval-valued function $i_m : \mathcal{A} \subseteq \mathcal{A} \rightarrow Int_{[0,1]}$ is called an R -probability if:

- (a) $i_m(A) = [i_m^-(A), i_m^+(A)] \subseteq [0, 1]$ with $i_m^-(A) \leq i_m^+(A)$,
- (b) \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 2 (Weichselberger [13]). 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 3 (Weichselberger [13]). An R -probability field $\mathcal{R} = (S, \mathcal{A}, i_m)$ is called an F -probability field if $\forall A \in \mathcal{A}$:

- (a) $i_m^+(A) = \sup\{\Pr(A) \mid \Pr \in \mathcal{M}(\mathcal{R})\}$,
- (b) $i_m^-(A) = \inf\{\Pr(A) \mid \Pr \in \mathcal{M}(\mathcal{R})\}$.

It is interesting to note that given a measurable space (S, \mathcal{A}) and a set of probability measures P , defining $i_m^+(A) = \sup\{\Pr(A) \mid \Pr \in P\}$ and $i_m^-(A) = \inf\{\Pr(A) \mid \Pr \in P\}$ gives an F -probability and that 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 1 (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

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

and

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

then

$$i_m(A) = [i_m^-(A), i_m^+(A)]$$

defines an F -probability field $\mathcal{R} = (R, \mathcal{B}, i_m)$. To see this, simply let P be the set of all probability measures on \mathcal{B} such that $\Pr(I) = 1$.

This example also illustrates that any set A , not just an interval I , can be used to define an F -probability field.

Example 2 (A probability measure is an F -probability field). Let \Pr be a probability measure over (S, \mathcal{A}) . Define

$$i_m(A) = [\Pr(A), \Pr(A)].$$

This definition is equivalent to having total knowledge about a probability distribution over S .

The concept of a cloud was introduced by Neumaier in [10] as follows:

Definition 4. 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) $(0, 1) \subseteq \bigcup_{s \in S} c(s) \subseteq [0, 1]$.

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 $p : S \rightarrow [0, 1]$ is called a *regular possibility distribution function* if

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

Possibility distribution functions (see [12]) define a possibility measure, $Pos : S \rightarrow [0, 1]$ where

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

and its dual necessity measure

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

where $\sup\{p(x) \mid x \in \emptyset\} = 0$. An “impossibility” distribution function $n : S \rightarrow [0, 1]$ can be defined by setting

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

where $\inf\{n(x) \mid x \in \emptyset\} = 1$. In [5], it is shown that possibility distributions can be constructed which satisfy the following consistency definition.

Definition 5. Let $p : S \rightarrow [0, 1]$ 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 \forall measurable sets A , $Nec(A) \leq \Pr(X \in A) \leq Pos(A)$.

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

Proposition 1. Let \bar{p}, p be a pair of regular possibility distribution functions over set S such that $\forall s \in S \bar{p}(s) + p(s) \geq 1$. Then the mapping $c(s) = [\underline{n}(s), \bar{p}(s)]$ where $\underline{n}(s) = 1 - p(s)$ (i.e. the impossibility 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}, 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 c .

Proof. (\Rightarrow)

(1) $\bar{p}, p : S \rightarrow [0, 1]$ and $\bar{p}(s) + p(s) \geq 1$ imply property (1) of Definition 4.

(2) Since all regular possibility distributions satisfy $\sup\{p(s) \mid s \in S\} = 1$ property (2) of Definition 4 holds.

Therefore c is a cloud. Now assume consistency. Then,

$$\alpha \geq \text{Pos}\{s \mid p(s) \leq \alpha\} \geq \Pr\{s \mid p(s) \leq \alpha\} = 1 - \Pr\{s \mid p(s) > \alpha\},$$

gives the right-hand side of the inequality for part (3) of Definition 4.

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

gives the left-hand side of the inequality for part (3) of Definition 4.

(\Leftarrow) The opposite implication was proved in Section 5 of [11]. \square

Example 3 (A cloud defines an R -probability field). Let c be a cloud over the real line. Let $\text{Pos}^1, \text{Nec}^1, \text{Pos}^2, \text{Nec}^2$ be the possibility measures and their dual necessity measures relating to $\bar{p}(s)$ and $p(s)$ (where \bar{p} and p are as in Proposition 1). Define

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

Neumaier [11] proved that every cloud contains a random variable X . Since consistency requires that $\Pr(X \in A) \in i_m(A)$, the result that every cloud contains a random variable X shows consistency. Thus every cloud defines an R -probability field.

Example 4 (A possibility distribution defines an R -probability field). Let $p : S \rightarrow [0, 1]$ be a regular possibility distribution function and let Pos be the associated possibility measure and Nec the dual necessity measure. Define $i_m(A) = [\text{Nec}(A), \text{Pos}(A)]$. Defining a second possibility distribution, $\underline{p}(x) = 1 \forall x$ means that the pair p, \underline{p} define a cloud for which $i_m(A)$ defines the R -probability. Since a cloud defines an R -probability field, this means that this possibility in turn generates an R -probability.

3. Construction of IVPMs from upper and lower cumulative distribution functions

This section constructs IVPMs of an F -probability from upper and lower bounding cumulative distribution functions in a manner allowing practical computation. For example, given statistical data, a confidence interval for the underlying cumulative distribution function using the method of Kolmogorov (see [6]) can be constructed. These confidence intervals may be used to construct upper and lower cumulative distributions which in turn can be used to construct IVPMs according to the construction method that follows. However, what is developed next is a construction which includes the confidence interval approach of [6] and this approach, as will be seen, is amenable to actual use.

We are interested in mathematical analysis using IVPMs. Mathematical analysis in general and continuous optimization in particular must come to terms with complete measures since as is known, the objective function which is the measure of “best” in the context of the problem at hand, requires a complete lattice. Thus, what is shown is how to construct a complete measure for IVPMs in a simple and direct way by utilizing upper and lower cumulative distributions. This approach includes the construction of possibility and necessity measures consistent with probability [5]. The direction of what ensues parallels the development of Borel sets but tailored to our particular measure, the IVPM. What is needed is to know how the measure operates on fundamental sets of real numbers, intervals. We begin with the half open interval $(a, b]$ and show what IVPMs are on this interval, the two parts of its complement, the whole real line \mathbb{R} , and \emptyset . Next, how to use IVPMs on finite unions, intersections, symmetric differences, and complements are

added to the construction. These basic operations apply to how we measure a finite implementation of these operations to fundamental sets. The next step is to indicate how IVPM works on countably infinite unions and intersections. The underlying structure are Borel sets which are subsets of the σ -ring generated by $(a, b]$, where $a \leq b$, are real numbers.

3.1. Underlying structure

Let $F^u(x) = \Pr(X^u \leq x)$ and $F^l(x) = \Pr(X^l \leq x)$ be two cumulative distribution functions defining two random variables over the Borel sets on the real line, X^u and X^l , with the property that $F^u(x) \geq F^l(x) \forall x$. It is clear that such a pair of random variables exist since one can start with a pair of differentiable functions F_1 and F_2 that are strictly increasing in $(0, 1)$, whose range is $[0, 1]$ and attain the value of 0 and 1 for some finite value such that $F_1(x) \geq F_2(x)$. These two functions can be considered as cumulative distributions and thus it is possible to define two random variables whose probability density function is the derivative of F_1 and F_2 having the aforementioned property. Set

$$\mathcal{M}(X^u, X^l) = \{X \mid \forall x \ F^l(x) \leq \Pr(X \leq x) \leq F^u(x)\}. \quad (1)$$

Since $\mathcal{M}(X^u, X^l)$ clearly contains X^u and X^l , the structure \mathcal{M} is non-empty. If the problem has associated underlying statistical data given, construction of upper/lower cumulative distributions with the above properties is straight-forward by using [6]. The remainder of this paper restricts itself to unknown $X \in \mathcal{M}(X^u, X^l)$. In what follows, for any Borel set A , we use $\Pr(A)$ to denote $\Pr(X \in A)$.

Remark 1. It is emphasized that the underlying structure for the construction of the IVPMs as approached in this research is defined by two cumulative distributions $F^u(x) \geq F^l(x)$, $x \in \mathbb{R}$ and is precisely (1). These cumulative distributions could be derived from data (as in [6]). Note that we are assuming that $F^u(x)$ and $F^l(x)$ are known for all $x \in \mathbb{R}$.

3.2. Construction

The development is begun by constructing probability bounds for members of a *generic* family of sets. Let $a < b$ but otherwise arbitrary real numbers and let

$$\mathcal{I} = \{(a, b], (-\infty, a], (a, \infty), (-\infty, \infty), \emptyset \mid, a, b \in \mathbb{R}\}.$$

Thus, on \mathcal{I} , it is clear by definition of \mathcal{M} and \mathcal{I} that

1. For $I = (-\infty, b]$ (or $(-\infty, b)$)

$$\Pr(I) \in [F^l(b), F^u(b)]. \quad (2)$$

2. For $I = (a, \infty)$ (or $[a, \infty)$)

$$\Pr(I) \in [1 - F^u(a), 1 - F^l(a)]. \quad (3)$$

3. For $I = (a, b]$ (or (a, b) , $[a, b)$, $[a, b]$), since $I = (-\infty, b] - (-\infty, a]$ and considering minimum and maximum probabilities in each set,

$$\Pr(I) \in [\max\{F^l(b) - F^u(a), 0\}, F^u(b) - F^l(a)]. \quad (4)$$

Therefore, if the definition of F^u, F^l is extended by defining $F^u(-\infty) = F^l(-\infty) = 0$ and $F^u(\infty) = F^l(\infty) = 1$, the following general definition can be made.

Definition 6. For any $I \in \mathcal{I}$, if $I \neq \emptyset$, define

$$i_m(I) = [i_m^-(I), i_m^+(I)] = [\max\{F^l(b) - F^u(a), 0\}, F^u(b) - F^l(a)], \quad (5)$$

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

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

Remark 2. Note that with this definition $i_m((-\infty, \infty)) = [\max\{F^1(\infty) - F^u(-\infty), 0\}, F^u(\infty) - F^1(-\infty)] = [1, 1]$, which matches our intuition and thus it is easy to see that $\Pr(I) \in i_m(I) \forall I \in \mathcal{I}$.

This approach may be extended to include finite unions of elements of \mathcal{I} . For example, let

$$E = I_1 \cup I_2 = (a, b] \cup (c, d] \quad \text{with } b \leq c,$$

then consider the probabilities

$$\Pr((a, b]) + \Pr((c, d]) \tag{6}$$

and

$$1 - (\Pr((-\infty, a]) + \Pr((b, c]) + \Pr((d, \infty))) \tag{7}$$

(the probability of the sets that make up E versus one minus the probability of the intervals that make up the complement). Consider the minimum and maximum probability for each case as a function of the minimum and maximum of each set. The minimum of (6) is

$$\max(0, F^1(d) - F^u(c)) + \max(0, F^1(b) - F^u(a)) \tag{8}$$

and the maximum of (6) is

$$F^u(d) - F^1(c) + F^u(b) - F^1(a). \tag{9}$$

The minimum of (7) is

$$\begin{aligned} 1 - (F^u(\infty) - F^1(d) + F^u(c) - F^1(b) + F^u(a) - F^1(-\infty)) \\ = F^1(d) - F^u(c) + F^1(b) - F^u(a) \end{aligned} \tag{10}$$

and the maximum of (7) is

$$\begin{aligned} 1 - \max(0, F^1(\infty) - F^u(d)) + \max(0, F^1(c) - F^u(b)) + \max(0, F^1(a) - F^u(-\infty)) \\ = F^u(d) - \max(0, F^1(c) - F^u(b)) - F^1(a). \end{aligned} \tag{11}$$

This gives

$$\Pr(E) \geq \max\{F^1(d) - F^u(c) + F^1(b) - F^u(a), \max(0, F^1(d) - F^u(c)) + \max(0, F^1(b) - F^u(a))\}$$

and

$$\Pr(E) \leq \min\{F^u(d) - \max(0, F^1(c) - F^u(b)) - F^1(a), F^u(d) - F^1(c) + F^u(b) - F^1(a)\}$$

so

$$\Pr(E) \in [\max(0, F^1(d) - F^u(c)) + \max(0, F^1(b) - F^u(a)), \tag{12}$$

$$F^u(d) - \max(0, F^1(c) - F^u(b)) - F^1(a)], \tag{13}$$

where the endpoints of the interval, (12), (13), are arrived at by noting that

$$\forall x, y \quad F^1(x) - F^u(y) \leq \max(0, F^1(x) - F^u(y)).$$

Remark 3. The two extreme cases for $E = (a, b] \cup (c, d]$ (recall that we are confining ourself to the case where $b \leq c$) are as follows.

(Upper extreme) For $F^u(x) = F^1(x) = F(x) \forall x$, 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.

(Lower extreme) 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} be the algebra of sets generated by \mathcal{I} . Note that every element of \mathcal{E} has a unique representation as a union of the minimum number of elements of \mathcal{I} as a union of disjoint elements of \mathcal{I} . That there is a unique representation comes from the fact that there are finite number of endpoints of intervals in \mathcal{E} which can be ordered. Then, for the (finite number of) overlapping intervals, their union is put into the unique representation. Note also that the real line $\mathbb{R} \in \mathcal{E}$ and \mathcal{E} is closed under complements.

Assume $E = \bigcup_{k=1}^{K_E} I_k$ and $E^c = \bigcup_{j=1}^{J_E} M_j$ 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

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

This can be made more concise using the following result.

Proposition 2. If $E = \bigcup_{k=1}^K I_k$ and $E^c = \bigcup_{j=1}^J M_j$ are the unique representations of E and $E^c \in \mathcal{E}$, then $\sum_{k=1}^K i_m^-(I_k) \geq 1 - \sum_{j=1}^J i_m^+(M_j)$ and $\sum_{k=1}^K i_m^+(I_k) \geq 1 - \sum_{j=1}^J i_m^-(M_j)$.

Proof. We need only to prove $\sum_{k=1}^K i_m^-(I_k) \geq 1 - \sum_{j=1}^J i_m^+(M_j)$ since we can exchange the roles of E and E^c giving $\sum_{j=1}^J i_m^-(M_j) \geq 1 - \sum_{k=1}^K i_m^+(I_k)$ proving the second inequality.

Note $\sum_{k=1}^K i_m^-(I_k) + \sum_{j=1}^J i_m^+(M_j)$ is of the form $\sum_{k=1}^K \max(0, F^l(b_k) - F^u(a_k)) + \sum_{j=1}^J F^u(a_{j+1}) - F^l(b_j) \geq \sum_{k=1}^K (F^l(b_k) - F^u(a_k)) + \sum_{j=1}^J F^u(a_{j+1}) - F^l(b_j)$. 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. \square

Next i_m is extended to \mathcal{E} .

Proposition 3. For any $E \in \mathcal{E}$ let $E = \bigcup_{k=1}^K I_k$ and $E^c = \bigcup_{j=1}^J M_j$ be the unique representations of E and E^c in terms of elements of \mathcal{I} , respectively. If

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

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

$$i_m(E) = [\inf\{\Pr(X) \in E \mid X \in \mathcal{M}(X^u, X^l)\}, \sup\{\Pr(X) \in E \mid X \in \mathcal{M}(X^u, X^l)\}]. \quad (14)$$

Proof. Part 1: $i_m : \mathcal{E} \rightarrow \text{Int}_{[0,1]}$, is an extension. First assume $E = (a, b) \in \mathcal{I}$, then $E^c = (-\infty, a] \cup (b, \infty)$. By the definition

$$\begin{aligned} i_m(E) &= [\max\{F^l(b) - F^u(a), 0\}, 1 - (\max\{F^l(a) - F^u(-\infty), 0\} + \max\{F^l(\infty) - F^u(b), 0\})] \\ &= [\max\{F^l(b) - F^u(a), 0\}, 1 - (F^l(a) + (1 - F^u(b)))] \\ &= [\max\{F^l(b) - F^u(a), 0\}, F^u(b) - F^l(b)], \end{aligned} \quad (15)$$

which matches the definition for i_m on \mathcal{I} . The other cases for $E \in \mathcal{I}$ are similar. Thus (15) is an extension.

Part 2: $i_m : \mathcal{E} \rightarrow \text{Int}_{[0,1]}$ is well defined. The representation of any element in \mathcal{E} in terms of the minimum number of elements of \mathcal{I} is unique. Moreover, it is clear that

$$0 \leq \sum_{k=1}^K i_m^-(I_k)$$

and

$$1 - \sum_{j=1}^J i_m^-(M_j) \leq 1.$$

So we only need to show that

$$\sum_{k=1}^K i_m^-(I_k) \leq 1 - \sum_{j=1}^J i_m^-(M_j),$$

or equivalently that

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

If the endpoints of all these intervals are relabeled as $-\infty = c_1 < c_2 < \dots < c_N = \infty$ then

$$\begin{aligned} \sum_{k=1}^K i_m^-(I_k) + \sum_{j=1}^J i_m^-(M_j) &= \sum_{n=1}^{N-1} \max\{F^l(c_{n+1}) - F^u(c_n), 0\} \\ &\leq \sum_{n=1}^{N-1} \max\{F^u(c_{n+1}) - F^u(c_n), 0\} = \sum_{n=1}^{N-1} \{F^u(c_{n+1}) - F^u(c_n)\} = 1. \end{aligned} \quad (16)$$

Thus,

$$\sum_{k=1}^K i_m^-(I_k) + \sum_{j=1}^J i_m^-(M_j) \leq 1. \quad (17)$$

Part 3: Equation (14) holds. For (16) assume

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

and

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

It is next shown that

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

and there is an

$$X \in \mathcal{M}(X^u, X^l)$$

for which

$$\Pr(X \in E) = i_m^+(E).$$

Note first that

$$\sum_{j=1}^J i_m^-(M_j) = \sum_{k=1}^K \max\{F^l(a_{k+1}) - F^u(b_k), 0\} \leq \sum_{k=1}^K \max\{F(a_{k+1}) - F(b_k), 0\} = \Pr(E^c),$$

which gives both

$$\Pr(E) = 1 - \Pr(E^c) \leq i_m^+(E)$$

and by replacing E with E^c ,

$$i_m^-(E) \leq \Pr(E).$$

Next for $x \leq a_2$, set

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

and for $a_2 < x \leq b_2$ 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

$$\Pr(E^c) = \sum_{j=1}^J i_m^-(M_j)$$

or

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

The other bound is similarly derived. \square

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

$$\Pr(A) \in [\sup\{i_m^-(E) \mid E \subseteq A, E \in \mathcal{E}\}, \inf\{i_m^+(F) \mid A \subseteq F, F \in \mathcal{E}\}]. \quad (18)$$

This prompts the following:

Proposition 4. Let $i_m : \mathcal{B} \rightarrow [0, 1]$ be defined by

$$i_m(A) = [\sup\{i_m^-(E) \mid E \subseteq A, E \in \mathcal{E}\}, \inf\{i_m^+(F) \mid A \subseteq F, F \in \mathcal{E}\}]. \quad (19)$$

The i_m is an extension from \mathcal{E} to \mathcal{B} and is well defined.

Proof. The last property of Proposition 3 insures it is an extension since, for example, if $E \subseteq F$ are two elements of \mathcal{E} then $i_m^+(E) \leq i_m^+(F)$ so $\inf\{i_m^+(F) \mid E \subseteq F, F \in \mathcal{E}\} = i_m^+(E)$ similarly it ensures that $\sup\{i_m^-(F) \mid F \subseteq E, F \in \mathcal{E}\} = i_m^-(E)$.

Next we show that i_m is well defined. Proposition 3 shows that $\forall E \in \mathcal{E}, i_m(E) \subseteq [0, 1]$. So $0 \leq \sup\{i_m^-(E) \mid E \subseteq A\}$ and $\inf\{i_m^+(E) \mid A \subseteq E\} \leq 1$. We also have $\sup\{i_m^-(E) \mid E \subseteq A\} \leq \inf\{i_m^+(F) \mid A \subseteq F\}$. \square

Proposition 5. The function $i_m : \mathcal{B} \rightarrow \text{Int}_{[0,1]}$ defines an F -probability field on the Borel sets and

$$i_m(B) = [\inf\{\Pr(X \in B) \mid X \in \mathcal{M}(X^u, X^l)\}, \sup\{\Pr(X \in B) \mid X \in \mathcal{M}(X^u, X^l)\}], \quad (20)$$

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

Proof. Clear. \square

Thus, Eq. (20) in the above proposition indicates how to compute the IVPM of any Borel set (starting with our bounding cumulative distributions).

4. Interval-valued extension, integration, and product measure for F -probabilities

Now that the construction of IVPs has been demonstrated to be a straight-forward process, three key ideas needed for the application of IVPs to mathematical programming problems are developed from the underlying constructed IVPs. These three key ideas are: (1) the extension principle for IVPs, (2) IVPM integration, and (3) IVPM product measures.

4.1. Extension principle for IVPs

Definition 7. Let $\mathcal{R} = (S, \mathcal{A}, i_m)$ be an F -probability field and $f : S \rightarrow T$ 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

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

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

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

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

which allows for evaluation using the techniques described earlier.

4.2. Integration of IVPs

Optimization over distributions from a recourse approach as will be used here, requires a “utility” functional to be defined. The functional used here, is an integral operator and is used to map distributions into the real numbers. For example, the expectation is such an operator. To this end, integration of IVPs is defined next. It is noted that our development again parallels the development of real-valued integration over arbitrary measures. The measure in this case is the IVPM.

Definition 8. Given F -probability field $\mathcal{R} = (S, \mathcal{A}, i_m)$ with structure $\mathcal{M}(\mathcal{R})$ and an integrable function $f : S \rightarrow \mathbb{R}$, define for any $A \in \mathcal{A}$:

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

The following observations are made which are useful in calculations.

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 $A \in \mathcal{A}$ and $y \in \mathbb{R}$, then

$$\int_A f(x) di_m = yi_m(A).$$

Further, if f is a simple function taking values $\{y_k \mid k \in K\}$ on a finite set 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 $A = \bigcup_{k \in K} A_k$ then

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

where

$$i_m^+ \left(\int_A f(x) di_m \right) = \sup \left\{ \sum_{k \in K} y_k \Pr(A_k) \mid \Pr \in \mathcal{M}(\mathcal{R}) \right\} \quad (21)$$

and

$$i_m^- \left(\int_A f(x) di_m \right) = \inf \left\{ \sum_{k \in K} y_k \Pr(A_k) \mid \Pr \in \mathcal{M}(\mathcal{R}) \right\}. \quad (22)$$

Remark 4. As will be seen, when something like an expected value is used, the sup will use the lower bounding function $F^l(x)$ since its average is bigger given that it is to the right of $F^u(x)$. In like manner, the inf will be obtained by the upper bounding function $F^u(x)$. Moreover, when the sup and inf of (21) and (22) are not obtained in closed form (which will occur for simple bounding functions) these can be evaluated by solving two linear programming problems since $\Pr \in \mathcal{M}(\mathcal{R})$ implies that $\Pr(\bigcup_{l \in L \subseteq K} A_l) \in i_m(\bigcup_{l \in L \subseteq K} A_l)$ for all $L \subseteq K$.

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

$$\int_A f(x) di_m = \lim_{k \rightarrow \infty} \int_A f_k(x) di_m,$$

where

$$\lim_{k \rightarrow \infty} \int_A f_k(x) di_m = \left[\lim_{k \rightarrow \infty} i_m^- \left(\int_A f_k(x) di_m \right), \lim_{k \rightarrow \infty} i_m^+ \left(\int_A f_k(x) di_m \right) \right]$$

provided the limits exist.

Example 5. Consider the IVP constructed from the interval $[a, b]$ and look at its interval expected value,

$$\int_R x di_m. \quad (23)$$

To evaluate (23), possibility and necessity measures obtained via consistent construction of [5] will be used. That is, $Nec\{(-\infty, x]\} \leq \Pr\{(-\infty, x]\} \equiv F(x) \leq Pos\{(-\infty, x]\}$ where $F(x)$ is the cumulative distribution for any of our random variables X that belong to our structure $\mathcal{M}(\mathcal{R})$. For an interval $[a, b]$,

$$F^u(x) = Pos(x) = \begin{cases} 0 & \text{for } x < a, \\ 1 & \text{for } a \leq x, \end{cases}$$

$$F^l(x) = Nec(x) = \begin{cases} 0 & \text{for } x < b, \\ 1 & \text{for } b \leq x. \end{cases}$$

Note that when the expected value is calculated, the lower bounding cumulative distribution defines the right endpoint of the integral IVP while the upper bounding cumulative distribution defines the left endpoint. Moreover, the lower probability density function (corresponding to the upper cumulative) of the interval di_m is

$$d^-(x) = \begin{cases} 0 & \text{for } x \neq a, \\ 1 & \text{for } a = x \end{cases}$$

and the upper probability density function (corresponding to the lower cumulative) of the interval di_m is

$$d^+(x) = \begin{cases} 0 & \text{for } x \neq b, \\ 1 & \text{for } a = b. \end{cases}$$

Thus,

$$\int_R x di_m = \left[\int_{-\infty}^{\infty} x d^-(x) dx, \int_{-\infty}^{\infty} x d^+(x) dx \right] = [a, b].$$

That is, the interval expected value of an interval is the interval itself.

Example 6. Consider the IVPM of a trapezoidal possibility interval $a/b/c/d$, core $[b, c]$, support $[a, d]$ and compute $\int_R x \, di_m$. Again, a construction using possibility and necessity measures obtained via consistent construction of [5] is used. For a trapezoidal possibility interval $a/b/c/d$

$$F^u(x) = Pos(x) = \begin{cases} 0 & \text{for } x < a, \\ \frac{1}{b-a}x - \frac{a}{b-a} & \text{for } a \leq x \leq b, \\ 1 & \text{for } b \leq x. \end{cases}$$

Note that

$$dF^u(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b, \\ 0 & \text{otherwise,} \end{cases}$$

so that (again, the upper cumulative distribution produces the lower integral):

$$\int_{-\infty}^{\infty} x \, d^-(x) \, dx = \int_a^b x \frac{1}{b-a} \, dx = \frac{1}{2} \frac{1}{b-a} \{b^2 - a^2\} = \frac{1}{2}(a+b).$$

$$F^l(x) = Nec(x) = \begin{cases} 0 & \text{for } x < c, \\ \frac{1}{d-c}x - \frac{c}{d-c} & \text{for } c \leq x \leq d, \\ 1 & \text{for } d \leq x. \end{cases}$$

Note that

$$dF^l(x) = \begin{cases} \frac{1}{d-c} & \text{for } c \leq x \leq d, \\ 0 & \text{otherwise,} \end{cases}$$

so that (again, the lower cumulative distribution produces the upper integral):

$$\int_{-\infty}^{\infty} x \, d^-(x) \, dx = \int_c^d x \frac{1}{d-c} \, dx = \frac{1}{2} \frac{1}{d-c} \{d^2 - c^2\} = \frac{1}{2}(c+d).$$

Thus,

$$\int_R x \, di_m = \left[\int_{-\infty}^{\infty} x \, d^-(x) \, dx, \int_{-\infty}^{\infty} x \, d^+(x) \, dx \right] = \left[\frac{a+b}{2}, \frac{c+d}{2} \right].$$

For triangular possibility intervals $a/b/c$,

$$\int_R x \, di_m = \left[\int_{-\infty}^{\infty} x \, d^-(x) \, dx, \int_{-\infty}^{\infty} x \, d^+(x) \, dx \right] = \left[\frac{a+b}{2}, \frac{b+c}{2} \right]. \quad (24)$$

4.3. Product measures of IVPs

The combination of IVPMs when the variables are independent is developed next. The situation when dependencies may be involved is not addressed.

Definition 9. Let $\mathcal{R} = (S, \mathcal{A}, i_X)$ and $\mathcal{Q} = (T, \mathcal{B}, i_Y)$ be F -probability fields representing uncertain independent 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

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

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

where $(S \times T, \mathcal{A} \times \mathcal{B})$ is the usual product of σ -algebra of sets and $\Pr_X \times \Pr_Y$ is the product measure on $S \times T$.

It is clear from this definition that

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

for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Thus, if there are several uncertain independent parameters in a problem with the uncertainty characterized by IVPs, the IVPM for the product space can be obtained and set E can be formed by a disjoint union of sets of the form $A \times B$ and multiplication by the respective IVPs for each of the sets and summing.

5. Application to optimization

An example application of these concepts is the following recourse problem. Suppose the optimization problem is to maximize $f(\vec{x}, \vec{a})$ subject to $g(\vec{x}, \vec{b}) = 0$ (where \vec{a} and \vec{b} are parameters). This section gives a more robust solution to problems of optimization under uncertainty. That is, instead of obtaining a (single) “best” solution, the IVP approach uses bounds thereby incorporating the variability of the parameters into its calculation where the variable includes the uncertainty distributions captured by IVPs as pointed out in the second section. Assume \vec{a} and \vec{b} are vectors of independent uncertain parameters, each with an associated IVPM. Assume the constraint may be violated at a cost $\vec{p} > 0$ per unit so that the problem becomes one to maximize

$$h(\vec{x}, \vec{a}, \vec{b}) = f(\vec{x}, \vec{a}) - \vec{p}|g(\vec{x}, \vec{b})|.$$

Given the independence assumption, form an IVPM for the product space $i_{\vec{a} \times \vec{b}}$ for the joint distribution. Next, calculate the interval-valued expected value with respect to this IVPM. Then the interval-valued expected value is

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

To optimize over such a value requires an ordering of intervals. One such ordering is 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 $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ and $v : \text{Int}_R \rightarrow \mathbb{R}^2$ by setting, for any interval $I = [a, b]$, $v(I) = ((a + b)/2, b - a)$. Thus v gives the midpoint and width of an interval. In this case u would measure the preference for one interval over another considering both its midpoint and width (a risk measure). Using this measure, the optimization problem becomes (leaving off the vector notation which will be assumed):

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

Example 7. Consider the problem

$$\begin{aligned} \max f(x, a) &= 8x_1 + 7x_2 \\ \text{subject to :} \\ g_1(x, b) &= 3x_1 + [1, 3]x_2 + 4 = 0 \\ g_2(x, b) &= \tilde{2}x_1 + 5x_2 + 1 = 0 \\ \vec{x} &\in [0, 2], \end{aligned}$$

where $\tilde{2} = 1/2/3$, that is, $\tilde{2}$ is a triangular possibilistic interval with support $[1, 3]$ and modal value at 2. For $\vec{p} = (1, 1)^T$,

$$h(x, a, b) = 5x_1 - \tilde{2}x_1 + [3, 5]x_2 - 6$$

so that

$$\begin{aligned}\int_R h(x, a, b) di_{a \times b} &= 5x_1 + \left[\int_0^1 (\alpha - 3) d\alpha, \int_0^1 (-1 - \alpha) d\alpha \right] x_1 + [3, 5]x_2 - 6 \\ &= 5x_1 + \left[-\frac{5}{2}, -\frac{3}{2} \right] x_1 + [3, 5]x_2 - 5.\end{aligned}$$

Note that $[-\frac{5}{2}, -\frac{3}{2}]$ is precisely the expected value of the triangular fuzzy interval $\tilde{2}$ whose calculation is (24). Since the constant -5 will not affect the optimization, it will be removed (then added at the end), so that

$$v \left(\int_R h(x, a, b) di_{a \times b} \right) = v \left(\left[\frac{5}{2}, \frac{7}{2} \right] x_1 + [3, 5]x_2 \right) = (3, 1)x_1 + (4, 2)x_2.$$

Let $u(\vec{y}) = \sum_{i=1}^n y_i$ which for the context of this problem yields

$$\max_x u \left(v \left(\int_R h(x, a, b) di_{a \times b} \right) \right) = \max_{\vec{x} \in [0, 2]} (4x_1 + 6x_2) - 5 = 20 - 5 = 15.$$

6. Conclusion

The definition of an interval-valued probability measure provides a formal setting in which various representations of uncertainty (for example intervals, probability, possibility, and clouds) can be combined. This allows solution methods for problems containing mixed representations. Future research will focus on the theory and applications of such measures to problems in optimization in which uncertainty cannot be fully captured by probability alone.

References

- [1] S. Destercke, D. Dubois, A unified view of some representations of imprecise probabilities, in: J. Lawry, E. Miranda, A. Bugarin, S. Li (Eds.), *Internat. Conf. on Soft Methods in Probability and Statistics (SMPS)*, Advances in Soft Computing, Bristol, Springer, 2006, pp. 249–257.
- [2] S. Destercke, D. Dubois, E. Chojnacki, Relating practical representations of imprecise probabilities, in: *5th Symp. Imprecise Probabilities and their Applications*, Prague, Czech Republic, 2007.
- [3] D. Dubois, Possibility theory and statistical reasoning, *Comput. Statist. Data Anal.* 51 (2006) 47–69.
- [4] S. Ferson, V. Kreinovich, L.R. Ginzburg, K. Sentz, D.S. Myers, Constructing probability boxes and Dempster–Shafer structures, Sandia National Laboratories, Technical Report SAND2002-4015, Albuquerque, New Mexico, 2003.
- [5] K.D. Jamison, W.A. Lodwick, The construction of consistent possibility and necessity measures, *Fuzzy Sets and Systems* 132 (2002) 1–10.
- [6] A.N. Kolmogorov, Confidence limits for an unknown distribution function, *Ann. Math. Statist.* 12 (1941) 461–463.
- [7] A.N. Kolmogorov, S.V. Fomin, *Introductory Real Analysis*, Dover Publications Inc., New York, 1975.
- [8] W.A. Lodwick, K.D. Jamison, Estimating and validating the cumulative distribution of a function of random variables: toward the development of distribution arithmetic, *Reliable Comput.* 9 (2003) 127–141.
- [9] R.E. Moore, *Methods and Applications of Interval Analysis*, SIAM, Philadelphia, 1979.
- [10] A. Neumaier, Clouds, fuzzy sets and probability intervals, *Reliable Comput.* 10 (4) (2004) 249–272.
- [11] A. Neumaier, On the structure of clouds, 2004 (<http://www.mat.univie.ac.at/~neum/papers.html>).
- [12] Z. Wang, G.J. Klir, *Fuzzy Measure Theory*, Plenum Press, New York, 1992.
- [13] K. Weichselberger, The theory of interval-probability as a unifying concept for uncertainty, *Internat. J. Approx. Reason.* 24 (2000) 149–170.