Mean and Variance Bounds and Propagation for Ill-Specified Random Variables

Andrew T. Langewisch and F. Fred Choobineh, *Member, IEEE*

***Abstract*—Foundations, models, and algorithms are provided for identifying optimal mean and variance bounds of an ill-specified random variable. A random variable is ill-specified when at least one of its possible realizations and/or its respective probability mass is not restricted to a point but rather belongs to a set or an interval. We show that a nonexhaustive sensitivity-analysis approach does not always identify the optimal bounds. Also, a procedure for determining the mean and variance bounds of an arithmetic function of ill-specified random variables is presented. Estimates of pairwise correlation among the random variables can be incorporated into the function. The procedure is illustrated in the context of a case study in which exposure to contaminants through the inhalation pathway is modeled.**

***Index Terms*—Fuzzy numbers, ill-specified random variables, mean and variance bounds, risk analysis, uncertainty.**

I. Introduction

In projective modeling, the analyst must often consider the behavior of random variables that are ill-specified in values and/or probabilities. By ill-specified, we mean the outcomes that form the support for a random variable are not exactly known, but may be represented by intervals or sets, or that the probabilities associated with these outcomes are not exactly specified, but can be constrained to intervals, or both. Furthermore, the sets or intervals may be overlapping. For such random variables, the mean and variance, two prime statistics used in analysis and estimation, cannot be precisely determined. We present procedures for optimally bounding the mean and variance of an ill-specified random variable and demonstrate the use of these procedures in risk analysis, where the risk is a function of one or more ill-specified random variables.

The impetus for considering ill-specified random variables is that, in most practical situations, no information about the random variable's probability distribution is available. The analyst, lacking adequate information to faithfully model a random variable, often will assume additional knowledge, and hope to safeguard this assumption with sensitivity analysis. Our proposed definition of ill-specified random variables requires limited amounts of information that often will be available, and, in Section 3, we will show that our definition also encompasses many of the established procedures for modeling ambiguity. Furthermore, we show that nonexhaustive sensitivity analysis does not always identify the optimal bounds for the mean and variance.

Ill-specificity is a consequence of imprecision or ambiguity concerning random variables, and in situations where arithmetic functions of several ill-specified variables are of interest, researchers have examined various techniques which model the resultant distributional uncertainty and imprecision. These techniques include robust Bayesian methods (e.g., [1]), Monte Carlo simulation with best judgment to choose input distributions (e.g., [2], [3]), Monte Carlo simulation with the input distributions selected by the maximum entropy criterion (e.g., [4] and [5]), second-order Monte Carlo simulation (e.g., [6] and [7]), probability-bounds analysis (PBA; e.g., [8] and [9]), fuzzy arithmetic (e.g., [10] and [11]), and interval analysis (e.g., [12] and [13]).

In this paper, an alternative approach, here labeled interval-based mean and variance propagation analysis (IMVPA), is presented and illustrated. This approach provides a tool for identifying optimal bounds for the mean and variance of ill-specified variables, and then propagating those interval bounds through arithmetic functions. Furthermore, IMVPA can utilize estimates of correlation between pairs of ill-specified random variables. The approach is illustrated in the context of a case study in which exposure to airborne emissions via the inhalation pathway is estimated for residents living near a food-processing facility [14]. In this example, upper and lower bounds on both the mean and standard deviation (or variance) of the risk function are found to be significantly tighter than those yielded by PBA. The use of IMVPA in risk analysis can help the analyst avoid overestimating the health risks due to airborne emissions. Other applications should be readily apparent in a variety of fields where quantitative modeling of ill-specified data is useful (e.g., determining the mean and variance of electric power generation system production costs; see [15] and [16]).

II. Related Approaches

Exact analytic methods for the propagation of uncertainty are typically unwieldy and often intractable, and require well-specified probability distributions [17]. Focusing on approximations, the method of moments is applied widely to the analysis of complex models [18]. In this approach, a response variable y is analyzed as a function of n precisely-described input distributions; in contrast, we operate on and propagate uncertainty through a series of two precisely-described and/or ill-specified variables at a time, aggregating results.

Manuscript received July 25, 2002; revised January 3, 2004. This paper was recommended by Associate Editor S. Patek.

Andrew T. Langewisch is with the Department of Business Administration at Concordia University, Seward, NE 68434 USA (e-mail: alangewisch@cune.edu).

F. Fred Choobineh is with the Department of Industrial and Management Systems Engineering, University of Nebraska, Lincoln, NE 68588 USA (e-mail: fchoobineh@unl.edu).

Digital Object Identifier 10.1109/TSMCA.2004.826316

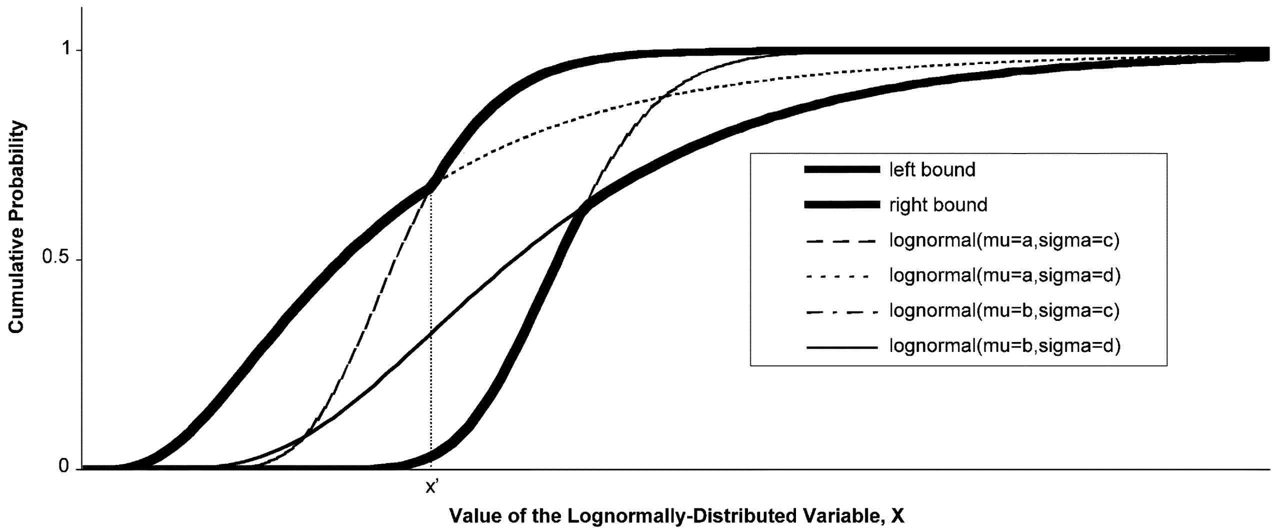


Fig. 1. PBA computes bounds on the CDF of a lognormal distribution having a mean somewhere in the interval (a, b) and standard deviation somewhere in (c, d).

Rowe [19] outlined fundamental bounds on means and variances of transformations of random variables for which moments and/or order statistics are given. Smith [20] elaborated bounds on distribution functions with given moments and shape constraints. His work generalizes that of Chebychev.

Walley [21] has developed a mathematical theory of imprecision for use in probabilistic reasoning, statistical inferences, and decision making. In the elicitation procedure, an individual's imprecise, indeterminate, and incomplete statements about his/her beliefs and judgments are modeled (in a manner that is fundamentally different from our definition of ill-specified random variables) and presented as closed convex half-spaces, and the intersection of these half-spaces identifies a class of linear previsions and a set of extreme points, one use of which is to identify upper and lower probabilities and upper and lower variances.

Saxena [22] suggests that if given upper and lower bounds on probabilities, one can maximize entropy to obtain a probability distribution. Saxena has developed a simple algorithm to facilitate determining the maximum, and he shows how the technique may be applied to investment analysis. Kmietowicz and Pearman [23] suggest that although a decision maker may not be able to specify precise probability distributions, if he or she can specify a ranking of the probabilities, i.e., $p_1 > p_2 > \cdots > p_n$, then maximum and minimum expected payoffs and variances can be derived and used to guide decision making.

Given uncertainty about a prior distribution and also about the family of sampling distributions, Lavine [24] developed a method for computing upper and lower bounds on posterior expectations over reasonable classes of sampling distributions that lie "close to" a given parametric family.

For ill-specified random variables, Langewisch and Choobineh [25] have suggested a procedure for establishing probability bounds and using those bounds for establishing stochastic dominance.

PBA computes bounds on the probability distribution of a random variable or a function of random variables when marginal distributions are known imprecisely or limited information is available about them [26], [27]. PBA is a flexible tool for propagating the effects of imprecision through calculations. In parametric cases, marginal distribution forms are known but the distributions' parameters may only be specified by intervals. For example, suppose that evidence implies that a distribution is lognormal in form, with its μ and σ known only within interval ranges. Fig. 1 illustrates probability bounds for the case $\mu \in [a,b]$ and $\sigma \in [c,d]$. The leftmost bound, for example, is comprised of two pieces: the CDF of lognormal(a,d) for $X \leq x'$, and the CDF of lognormal(a,c) otherwise. Clearly, the resulting bounds are not lognormal distributions. In nonparametric cases, the marginal distribution forms are not known but some information such as the minimum, maximum, mode, and/or percentile values are known. RAMAS Risk Calc [26] is a commercially available software package for performing PBA that uses modifications of Williamson and Downs' [8] procedure for constructing probability bounds.

III. Mean and Variance Bounds

We call a real-valued variable X an *ill-specified* random variable when we do not know the precise probability measure on X, but we have enough information to constrain the possible realizations to a finite number n of points, sets or bounded intervals, i.e., $x_i \in A_i$, and we can constrain the probability mass assignments on A_i to points or bounded intervals; i.e., $m_i \in M_i, i = 1, \ldots, n$. At least one of the sets or intervals must be nondegenerate in order to have an ill-specified random variable. Furthermore, A_i's may overlap. Thus, $P[X = x] \leq \sum_{k:x \in A_k} m_k$. If there is a singleton set $A_j = \{x\}$, then there is also a lower bound $m_j \leq P[X = x]$. Since in this ill-specified space we cannot determine the precise values of the distribution's mean and variance, we propose procedures for obtaining their optimal bounds.

An unwary approach for obtaining the bounds is to use a limited or nonexhaustive-sensitivity analysis. However, this approach does not guarantee optimal bounds. For example, consider the following outcome-probability pairs $\{(11, 0.22), ([11, 14], 0.26), ([12, 15], 0.26), (15, 0.26)\}$, where all x_i's and m_i's are crisp except two x_i's which are

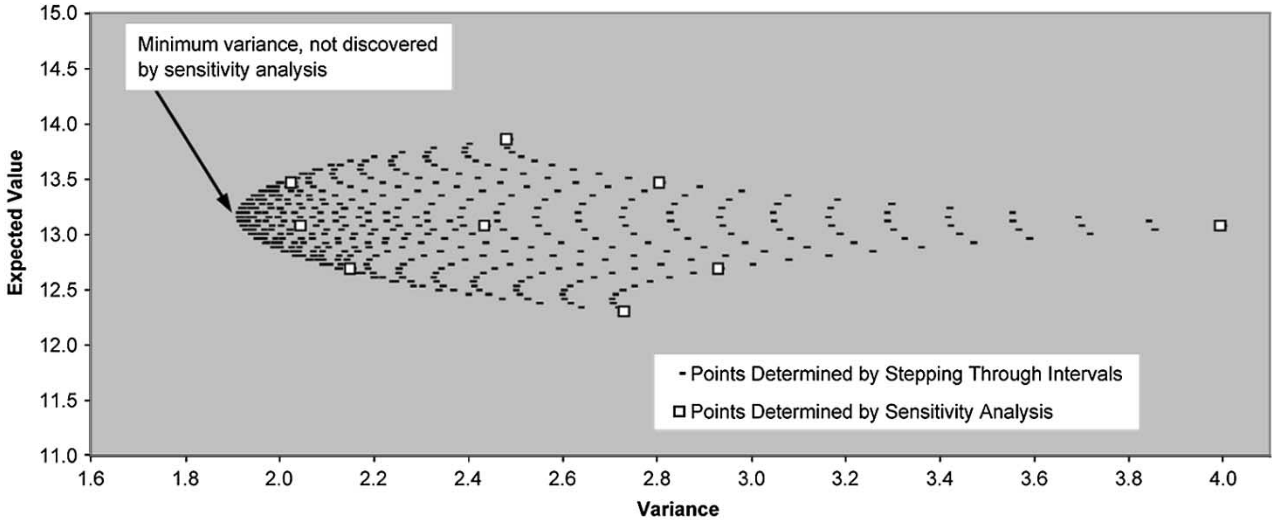


Fig. 2. Typical, irregularly shaped region of E-V points for an ill-specified random variable.

TABLE I
CLASSES OF PROBABILITY-OUTCOME ASSIGNMENTS

	Outcomes crisp	Outcomes ill-specified	
	Each A_i is a singleton	Each A_i is a finite set, at least one of which is not a singleton	Each A_i is a bounded interval, at least one of which is not degenerate
Probabilities crisp: Each m_i is given	I	IIa	IIb
Probabilities ill-specified: Each $m_i \in M_i$	III	IVa	IVb

only known to be in the given intervals. By stepping through the two intervals in small increments, the mean and variance points in Fig. 2 were generated. Now, as might be suggested by a limited or nonexhaustive sensitivity-analysis approach, we calculate the mean and variance points using the high, low, and midpoint values of the two intervals. The results are the nine highlighted points in the figure. Note the minimum variance is not found by the sensitivity-analysis approach. The smallest of the nine variance values found by sensitivity analysis is 2.02, whereas the minimum variance determined with techniques from Section III-B is 1.91.

The optimal bounds for the mean may be obtained by optimizing (separately maximizing and minimizing) the mean $= \sum_i m_i x_i$ over the space Ω comprised of the Cartesian product of A_i's mapped to the Cartesian product of M_i's. It should be noted that the mean is linear in both $\boldsymbol{m} = \{m_1, \ldots, m_n\}$ and $\boldsymbol{x} = \{x_1, \ldots, x_n\}$. Models for obtaining the optimal mean bounds are presented in Section III-A.

Similarly, the optimal bounds for variance may be obtained by optimizing variance $= \sum_i m_i x_i^2 - (\sum_i m_i x_i)^2$ over Ω. The variance is a concave function of $\boldsymbol{m}$ since the first term is linear in $\boldsymbol{m}$ and the second term is a concave function of a linear function of $\boldsymbol{m}$. In addition, for any mass assignment $\boldsymbol{m}$, variance, a nonnegative value, is a quadratic function of the form $\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{Q}\boldsymbol{x}$, where the diagonal entries of $\boldsymbol{Q}$ are $m_i - m_i^2$ and the nondiagonal entries are $-m_i m_j$. It can be shown that the variance is convex in $\boldsymbol{x}$ because $\boldsymbol{Q}$ is symmetric and positive semidefinite (see, e.g., [28]). Since variance is concave in $\boldsymbol{m}$ and convex in $\boldsymbol{x}$ and the feasible space is convex, we can in some cases, but not always, optimize variance by solving a quadratic program. Models for the quadratic cases and algorithms for the remaining cases are presented in more detail in Section III-B.

To facilitate discussion of optimization models that result from different restrictions on $\boldsymbol{x}$ and $\boldsymbol{m}$, we classify the outcome-probability space according to the conditions presented in Table I.

The Class I probability distribution represents a regular probability distribution and only for this class can precise values for the mean and variance be found. Class II considers cases where it is meaningful to assign a point probability to an ill-specified outcome. Probability mass, summing to one, is distributed over the subsets or subintervals of the outcome space. It will be delineated later in this section that the Class IIa is analogous to the problem studied by Dempster [29] and Shafer [30]. In Class III, we consider random variables with ill-specified probabilities, where probability intervals are assigned to each crisp outcome. Here, the analyst is presented with evidence suggesting $m_i \in M_i$. Clearly, for this class, we cannot expect these probability intervals to sum to one in the usual sense, but rather we can require $\sum_i \inf M_i \leq 1$ and $\sum_i \sup M_i \geq 1$. Class IV represents situations of ill-specificity in both probabilities and outcomes.

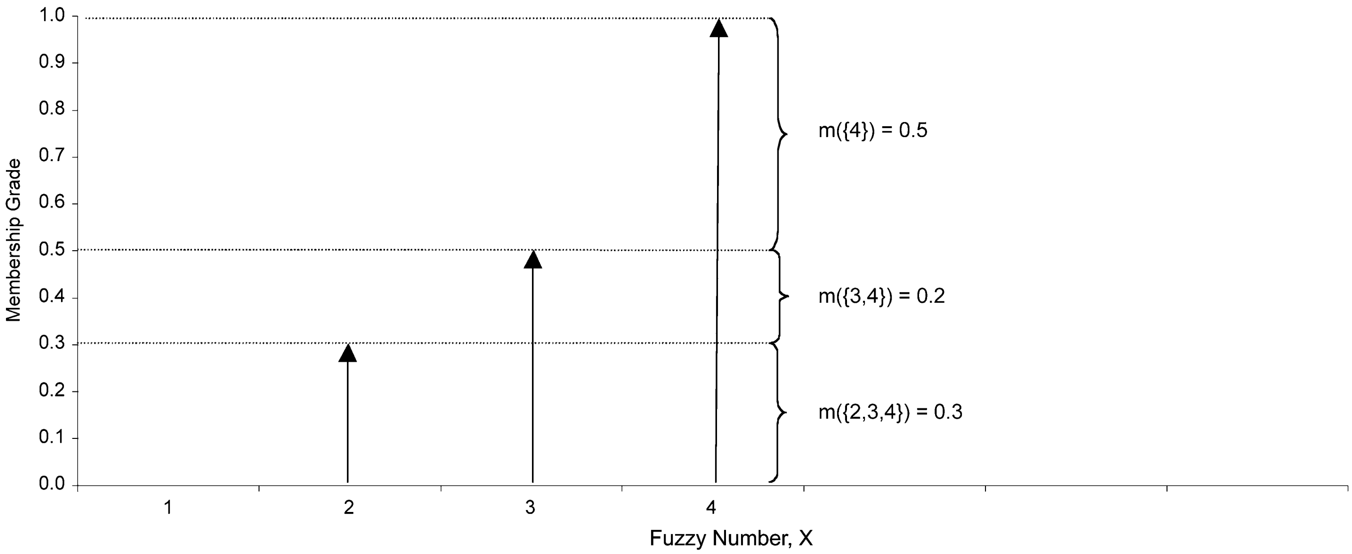


Fig. 3. Relationship between fuzzy numbers and probability mass assignments.

Many common depictions of limited knowledge, ambiguity, or imprecision concerning quantities may be classified according to Table I. A few examples are mentioned below.

The multivalued mappings and representations of Dempster-Shafer belief functions (evidence theory) can be represented by Class IIa. In this multivalued mapping, $m_i = m(A_i)$ is the probability mass attributed to a subset $A_i \in 2^X$, where 2^X is the power set of the precise finite-outcome space $\{x_1, \ldots, x_m\}$. If there are k subsets whose probability-mass assignments are nonzero, then the support $X_s = \{x_1 \in A_1, \ldots, x_k \in A_k\}$. The basic probability assignment $m(A_i)$ indicates the degree of belief that the actual outcome will be an element of A_i, with no further evidence available for establishing the likelihood of one element in the set over another in the same set.

If the analyst is presented with fuzzy numbers, we can follow the approach outlined by Klir and Yuan [10] to obtain their equivalent Dempster-Shafer belief functions. Since fuzzy set theory has a measure-theoretic counterpart in possibility theory, and since possibility theory is a special branch of Dempster-Shafer theory, we can represent numerical fuzzy sets by basic probability assignments on nested sets, i.e., $A_1 \subset A_2 \subset \cdots \subset A_k (=X_O)$, where X_O is the domain or universal set.

To illustrate both the connection with Dempster-Shafer multivalued mappings and fuzzy numbers, let B represent the concept *creditworthy*, with grades of membership $\mu_B = 0.0/1 + 0.3/2 + 0.5/3 + 1.0/4$, corresponding to a credit score $x \in X_O, X_O = \{1,2,3,4\}$. $\pi(A) = \sup_{x \in A} \mu_B(x)$ is the possibility measure of any subset A of X_O. The alpha cuts are defined as $B_\alpha = \{x \in X \,|\, \mu_B(x) \geq \alpha\}$, and will be the focal elements. There are two ways to compute the probability-mass assignments for all A. One method uses the relationship $m(A) = \sum_{A* \subseteq A} (-1)^{|A - A*|}[1 - \pi(\overline{A*})]$. For example, $m(\{3,4\}) = -(1 - \pi(\{1,2,4\})) - (1 - \pi(\{1,2,3\})) + (1 - \pi(\{1,2\})) = 0 - 0.5 + 0.7 = 0.2$. Alternatively, the probability mass assignments may be found quickly using set differences based on alpha cuts, as illustrated in Fig. 3. Specifically, let $A_0 = \emptyset, A_1 = \text{the highest alpha cut}, \ldots, A_k = \text{the lowest non-zero alpha cut}$, and $A_{k+1} = X_O$. Also let $D_i = A_i - A_{i-1}, i = 1, \ldots, k+1$. Then $m(A_i) = \pi(D_i) - \pi(D_{i+1}), i = 1, \ldots, k$. Note that while this example focuses on discrete fuzzy numbers, the same relationships hold for the intervals associated with continuous fuzzy numbers.

Interval representation of a distribution may be considered a Class IIb case. If the analyst confesses total ignorance about a particular X, except that the outcome of interest x belongs to $A_1 = [c, d]$, then $n = 1$ and we may assign $m_1 = 1$. If the analyst can also supply the median med_X, let $A_1 = [c, \mathrm{med}_X]$ and $A_2 = [\mathrm{med}_X, d]$. Then $m_1 = m_2 = 0.5$, and the support $X_s = \{x_1 \in A_1, x_2 \in A_2\}$. If additional quantiles $\{\xi_{q_1}, \ldots, \xi_{q_n}\}$ can be provided, then let $A_1 = [c, \xi_{q_1}], A_2 = (\xi_{q_1}, \xi_{q_2}], \ldots, A_{n+1} = (\xi_{q_n}, d]$ and $m_1 = q_1, m_2 = q_2 - q_1, \ldots, m_{n+1} = 1 - q_n$. If one is subjectively assessing continuous probability distributions, Clemens [31] describes a bracket median approach, which, given $a \leq X \leq b$, asks the assessor to supply a value m^* between a and b such that $P(a \leq X \leq m^*) = P(m^* \leq X \leq b)$. Class IIb techniques could be used on the set of elicited intervals.

Class III represents a common depiction of ambiguity. In eliciting subjective probabilities through a lottery process, Clemens suggests that it is important to begin with extremely wide brackets for the reference lottery and to converge on the indifference probability slowly. Franke [32] suggested that many people, when asked for a subjective probability, would be reluctant to specify a unique number and would prefer to specify an interval $[p(x) - \underline{\Delta}, p(x) + \bar{\Delta}]$ in which the "true" probability, $p(x)$, lies $(\underline{\Delta}, \bar{\Delta} \geq 0)$. These probability intervals accommodate individuals who like to think there is a true probability, even though the theory of subjective probability does not admit such a notion.

A. Optimal Mean Bounds

In Class I, the random variable's mean is unique and can be found in the usual way, i.e.,

$$E[X] = \sum_i A_i m_i. \quad (1)$$

For a Class IIa or IIb probability-outcome assignment X, bounds are determined by concentrating probability mass toward the extreme elements in the intervals or subsets. Minimum and maximum expected values, denoted by $\underline{E}[X]$ and $\overline{E}[X]$, are given by

$$\underline{E}[X] = \sum_i (\inf A_i) m_i \tag{2}$$

and

$$\overline{E}[X] = \sum_i (\sup A_i) m_i. \tag{3}$$

For Class III, $\overline{E}[X]$ is determined by the simple linear program

$$\begin{aligned} \overline{E}[X] &= \max \sum_i x_i \overline{m}_i \\ \text{s.t. } & \sum_i \overline{m}_i = 1 \\ \inf M_i &\leq \overline{m}_i \leq \sup M_i \quad \forall i. \end{aligned} \tag{4}$$

Here, $\overline{m}_i$ is a decision variable, and need not equal $\sup M_i$. $\underline{E}[X]$ is found by solving a similar minimization problem. For Class IV, we modify the optimization models of Class III, using extreme outcomes from the subsets or intervals as the objective function coefficients

$$\begin{aligned} \overline{E}[X] &= \max \sum_i (\sup A_i) \overline{m}_i \\ \text{s.t. } & \sum_i \overline{m}_i = 1, \\ \inf M_i &\leq \overline{m}_i \leq \sup M_i \quad \forall i. \end{aligned} \tag{5}$$

$\underline{E}[X]$ is found similarly.

In sum, finding expected value bounds is a straightforward process. Often the assignments can be determined by inspection. For more complex problems, the linear programs will be of assistance.

B. Optimal Variance Bounds

Before we discuss the optimization models for obtaining the optimal variance bounds, we observe that when there is imprecision concerning the outcomes, i.e., for X in Classes IIa, IIb, IVa, and IVb, the optimal solution for the maximization problem is obtained when the probability mass associated with each set or interval is focused on, or assigned to, the supremum and/or infimum of that set or interval. This is stated formally as Proposition 1 below. Utilization of this proposition yields more efficient optimization models for these classes.

Proposition 1: For an ill-specified random variable X, the maximum variance $\overline{\mathrm{Var}}[X]$ is found when all probability mass attributable to the subset or interval A_i is assigned to $\inf A_i$ or $\sup A_i$ or both for all i.

Proof: Consider an arbitrary distribution Y satisfying all the constraints of the ill-specified random variable X, consisting of known values $y_{ij} \in A_i$ and $P[Y = y_{ij}] = m_{ij}$, such that $\inf M_i \leq \sum_j m_{ij} \leq \sup M_i \; \forall i$, and $\sum_i \sum_j m_{ij} = 1$. Without loss of generality, translate Y such that $E[Y] = 0$. Translate each A_i by the same amount and direction and label it B_i.

Let Z be a second distribution with $z_{ij} = \inf B_i$ if $y_{ij} < 0, z_{ij} = \sup B_i$ if $y_{ij} \geq 0$, and $P[Z = z_{ij}] = m_{ij}$. Observe that $y_{ij}(z_{ij} - y_{ij}) \geq 0$ for all i, j. Now $\mathrm{Var}[Z] = \mathrm{Var}[Y + Z - Y] = \mathrm{Var}[Y] + \mathrm{Var}[Z - Y] + 2\mathrm{Cov}[Y, Z - Y] = \mathrm{Var}[Y] + \mathrm{Var}[Z - Y] + 2(E[Y(Z - Y)] - E[Y]E[Z - Y]) = \mathrm{Var}[Y] + \mathrm{Var}[Z - Y] + 2E[Y(Z - Y)]$. Since $y_{ij}(z_{ij} - y_{ij}) \geq 0$ for all $i, j, E[Y(Z - Y)] \geq 0$, and thus $\mathrm{Var}[Z] \geq \mathrm{Var}[Y]$. Since $\overline{\mathrm{Var}}[X]$ exists and Y was chosen arbitrarily, $\overline{\mathrm{Var}}[X]$ must be found when all probability mass attributable to the subset or interval A_i is assigned to $\inf A_i$ or $\sup A_i$ or both for all i. □

For Class I the precise value of the variance can be obtained through $\mathrm{Var}[X] = E[X - \mu]^2$. For other classes, we develop optimization models, each with a quadratic objective function and a set of linear constraints representing probability or outcome restrictions. The constraints are linear and produce a convex feasible space. In order to use a standard quadratic programming package, for a maximization problem, the objective function should be concave and, for a minimization problem, it should be convex. However, these conditions are not satisfied for all classes, and in those cases we develop special algorithms for obtaining the optimal bounds.

1) Class II Variance Bounds: In Class IIa, the upper bound on the variance of X is denoted $\overline{\mathrm{Var}}[X]$ and is equal to the optimal value of the nonlinear program

$$\begin{aligned} \overline{\mathrm{Var}}[X] = \max & \Bigg[\sum_i [(\inf A_i)^2 m_{i1} + (\sup A_i)^2 m_{i2}] \\ & - \left(\sum_i [(\inf A_i) m_{i1} + (\sup A_i) m_{i2}] \right)^2 \Bigg] \\ \text{s.t. } & m_{i1} + m_{i2} = m_i \quad \forall i \\ & m_{i1}, m_{i2} \geq 0 \quad \forall i \end{aligned} \tag{6}$$

where m_{i1} and m_{i2} are decision variables and denote the probability mass allocated to $\inf A_i$ and $\sup A_i$, respectively. The first $2n$ terms of the objective function are linear in $\boldsymbol{m}$ and the remaining terms may be considered a concave function $(y \to -y^2)$ of a linear function of $\boldsymbol{m}$, so the objective function is concave. Thus, the model is a standard quadratic program. It should be noted that the optimal solutions for m_{i1} and m_{i2} may both be positive. If an outcome set A_i is degenerate (i.e., $\inf A_i = \sup A_i$), we will only need one decision variable for that set; denote this as m_{i1} and set $m_{i2} = 0$.

The minimization problem for finding $\underline{\mathrm{Var}}[X]$ for X in Class IIa is not a standard quadratic program, because the objective function is not convex, and a search algorithm will be proposed for finding $\underline{\mathrm{Var}}[X]$. In the algorithm, rather than performing an exhaustive search that requires computing $\prod_i |A_i|$ possible variances, we propose a procedure that only examines a subset of these. The proposed algorithm uses the results of the following proposition, which assures its optimality.

Proposition 2: The variance of a Class IIa ill-specified random variable X is smaller when the probability mass associated with each of its sets is concentrated on one element of its respective set than when the probability mass of a set is concentrated on more than one element of its respective set.

Proof: Suppose the probability mass from at least one subset of X is assigned to two or more elements of its respective subset. We will show that there exists an alternative assignment of probability mass involving one less element of the subset that results in a smaller variance. Thus, the minimum variance cannot occur when probability mass is split among two or more elements of a subset, and so must occur when probability mass is concentrated on just one element of the subset. Denote one such subset A_k and designate a_1 and a_2 (the order to be determined below) as two elements over which probability mass is split. Excluding these two elements, list and denote the remaining elements of this subset and all other subsets as $X^* = \{x_1, \ldots, x_r\} = (A_k - \{a_1, a_2\}) \cup_{j \neq k} A_j$. Denote arbitrary probability mass assignments on these as $m(x_i)$, subject to $m(a_1) + m(a_2) + \sum m(x_i) = 1$ and $0 \leq m(x_i) \leq \sum_{A_i : x_i \in A_i} m(A_i)$. Let the central moment of X^* be denoted as $\mu = \sum x_i m(x_i)$, observing that $\sum m(x_i) < 1$. Assign a_1 and a_2, such that $|a_1 - \mu| \leq |a_2 - \mu|$. Accounting now for the effect of a_1 on the central moment, translate X, including all outcomes x_i, a_1, and a_2 to X', x_i', a_1', and a_2', such that $\mu' = \sum x_i' m(x_i') + a_1' m(a_1') = 0$. Then $(a_2')^2 > (a_1')^2$, and

$$\begin{aligned}
&\mathrm{Var}[X] = \mathrm{Var}[X'] \\
&= \sum (x_i')^2 m(x_i') + (a_1')^2 m(a_1') + (a_2')^2 m(a_2') - [a_2' m(a_2')]^2 \\
&> \sum (x_i')^2 m(x_i') + (a_1')^2 m(a_1') + (a_1')^2 (a_2') - [a_1' m(a_2')]^2
\end{aligned}$$

which is the variance if $m(a_2)$ had been combined with $m(a_1)$ and also assigned to a_1. In summary, the total variance is always smaller if the remaining assignable probability mass is concentrated closer to the central moment of the rest of the distribution, as opposed to being split between two elements. Thus, the distribution with the minimum variance will be one where the probability masses are not split among elements of a subset. □

The algorithm's general idea is as follows: Iterate once for each outcome x_i in the outcome set, concentrating unassigned probability mass from each subset onto x_i, or if x_i is not in the subset, onto the closest subset element to x_i. Compute the variance for the resulting probability distribution. Through the iterations, compare and save the minimum variance. However, we first reduce the potential number of iterations by excluding some of the smallest and largest outcomes. For any $x_i < x_{i+1} \leq \underline{E}[X]$, since we are minimizing $\mathrm{Var}[X] = E[X - \mu]^2$, we need not consider concentrating any assignable probability mass on x_i. Similarly, we need not consider concentrating any assignable probability mass on any $x_i > x_{i-1} \geq \overline{E}[X]$. Additionally, there is one subtlety. There may be neighboring elements x_i and x_k of an outcome set A_s which does not include x_j, such that $x_i < x_j < x_k$ and $x_j - x_i = x_k - x_j$. Then, when one is examining the situation where probability is concentrated toward x_j, there are two cases: we may concentrate m_s on x_i or on x_k. Pseudocode useful for programming this algorithm follows.

Algorithm for Determining the Minimum Variance of a Class IIa Ill-Specified Random Variable X

```
; Given an ordered outcome set X_O = {x_1, ..., x_m}, x_1 < ··· < x_m, and proba-
; bility masses m(A_i) distributed over the power set of X_O, let a_ij denote the jth element
; of subset A_i. Let A_1, ..., A_n denote the focal elements of X_O (i.e., m_i > 0 i =
; 1, ..., n).
;
; initialization
min_var = +∞
smallest_eligible_outcome_index = max{1, i ∋ x_i < x_{i+1} ≤ E̲[X]}
largest_eligible_outcome_index = min{m, i ∋ x_i > x_{i-1} ≥ Ē[X]}
; for breaking ties, set ε to be more than twice as small as smallest difference in distances
between any two eligible outcomes in the outcome set
ε = min{|(x_{i+2} − x_{i+1}) − (x_{i+1} − x_i)|}/3,
  i = smallest_eligible_outcome_index, ...,
  largest_eligible_outcome_index − 2
loop
k = smallest_eligible_outcome_index to largest_eligible_outcome_index
    ; with each iteration concentrate as much probability mass as possible toward the
    ; kth element of X_O.
    x* = x_k
    expectation_x = 0
    expectation_x2 = 0
    loop i = 1 to n
    ; this time through, break ties by concentrating mass toward x* − ε
    min_spread = +∞
    loop j = 1 to |A_i|
      if |a_ij − (x* − ε)| < min_spread then
        min_spread = |a_ij − (x* − ε)|
        ; a* will be the closest element in A_i to x*
        a* = a_ij
      endif
    next j
    expectation_x = expectation_x + m_i a*
    expectation_x2 = expectation_x2 + m_i (a*)^2
  next i
  test_var = expectation_x2 − (expectation_x)^2
  min_var = min{min_var, test_var}
  ; now break ties by concentrating mass toward x* + ε, except that there
  ; is no need to test both sides of the smallest and greatest eligible elements
  if k ≠ smallest_eligible_outcome_index
   and k ≠ largest_eligible_outcome_index
    expectation_x = 0
    expectation_x2 = 0
    loop i = 1 to n
      min_spread = +∞
      loop j = 1 to |A_i|
        if |a_ij − (x* + ε)| < min_spread then
          min_spread = |a_ij − (x* + ε)|
          a* = a_ij
        endif
      next j
      expectation_x = expectation_x + m_i a*
      expectation_x2 = expectation_x2 + m_i (a*)^2
      next i
      test_var = expectation_x2 − (expectation_x)^2
      min_var = min{min_var, test_var}
      endif
next k
;
; end of algorithm—the minimum variance is stored in the variable min_var
;
```

By the results of Proposition 2, the above algorithm returns the minimum variance. To demonstrate the major steps of the algorithm, suppose probability mass is distributed over the power

TABLE II
SEEKING MINIMUM VARIANCE FOR A CLASS IIA PROBLEM

		probability mass concentrated towards					
		1		$2-\varepsilon$		$2+\varepsilon$	
Subset A_i	m_i	$E[X]$	$E[X^2]$	$E[X]$	$E[X^2]$	$E[X]$	$E[X^2]$
{1}	0.20	0.20	0.20	0.20	0.20	0.20	0.20
{2,3}	0.15	0.30	0.60	0.30	0.60	0.30	0.60
{1,3}	0.10	0.10	0.10	0.10	0.10	0.30	0.90
{1,3,8}	0.05	0.05	0.05	0.05	0.05	0.15	0.45
{3,8}	0.30	0.90	2.70	0.90	2.70	0.90	2.70
{1,2,3,8}	0.20	0.20	0.20	0.40	0.80	0.40	0.80
Totals		1.75	3.85	1.95	4.45	2.25	5.65
Variance		0.7875		0.6475		0.5875	

		probability mass concentrated towards					
		$3-\varepsilon$		$3+\varepsilon$		8	
Subset A_i	m_i	$E[X]$	$E[X^2]$	$E[X]$	$E[X^2]$	$E[X]$	$E[X^2]$
{1}	0.20	0.20	0.20	0.20	0.20	0.20	0.20
{2,3}	0.15	0.45	1.35	0.45	1.35	0.45	1.35
{1,3}	0.10	0.30	0.90	0.30	0.90	0.30	0.90
{1,3,8}	0.05	0.15	0.45	0.15	0.45	0.40	3.20
{3,8}	0.30	0.90	2.70	0.90	2.70	2.40	19.20
{1,2,3,8}	0.20	0.60	1.80	0.60	1.80	1.60	12.80
Totals		2.60	7.40	2.60	7.40	5.35	37.65
Variance		0.6400		0.6400		9.0275	

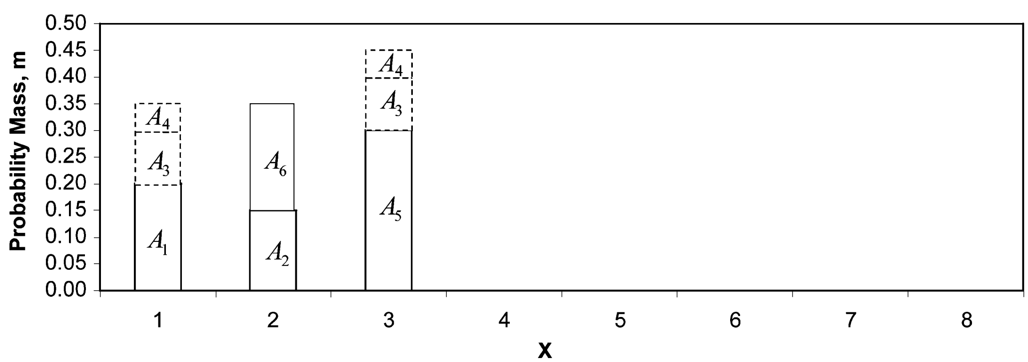


Fig. 4. Allocation of probability mass in the Class IIa variance-minimizing search for the example in Table II. To find the minimum variance, alternative allocations of probability mass must be tested. Here, probability mass is concentrated toward $X = 2$. Subsets 3 and 4 allow assignment to $X = 1$ or $X = 3$, which are equally close. Thus, two cases are generated. Concentrating mass toward $X = 2 - \epsilon$ leads to the probability mass of subsets 3 and 4 being assigned to $X = 1$, with an associated variance of 0.6475. Concentrating mass toward $X = 2 + \epsilon$ leads to probability mass being assigned to $X = 3$, with an associated variance of 0.5875.

set of $X_O = \{1, 2, 3, 8\}$ as shown in the first two columns of Table II. $\underline{E}[X] = 1.75$ and $\overline{E}[X] = 5.35$ so we cannot exclude any of the smallest or largest outcomes from consideration. The remaining columns illustrate the results of the algorithm. Fig. 4 illustrates the iteration where probability mass is being concentrated toward $X = 2$. The minimum variance is 0.5875, found when probability mass is concentrated toward $2 + \varepsilon$, as summarized in Table III. In this example, only six of the 96 extreme points needed to be examined.

For Class IIb the variance maximization model is identical to that of Class IIa, other than substituting *interval* A_i for *subset*

TABLE III
THE VARIANCE-MINIMIZING ASSIGNMENTS FOR THE CLASS IIA PROBLEM IN TABLE II

Subset A_i	$a^* \in A_i$	m_i
{1}	{1}	0.20
{2,3}	{2}	0.15
{1,3}	{3}	0.10
{1,3,8}	{3}	0.05
{3,8}	{3}	0.30
{1,2,3,8}	{2}	0.20

A_i. To obtain the minimum variance, since variance as a function of $\boldsymbol{x}$ is convex, $\underline{\mathrm{Var}}[X]$ is given by the quadratic program

$$\underline{\mathrm{Var}}[X] = \min \sum_{i=1}^{n} m_i x_i^2 - \left(\sum_{i=1}^{n} m_i x_i\right)^2 \quad \text{s.t. } \inf A_i \le x_i \le \sup A_i \quad \forall i \tag{7}$$

which is a standard quadratic program.

2) Class III Variance Bounds: In Class III, we have $\mathrm{Var}[X]$ as a function of $\boldsymbol{m}$, so here the function is concave. $\overline{\mathrm{Var}}[X]$ is given by the standard quadratic program

$$\begin{aligned}\overline{\mathrm{Var}}[X] &= \max \sum_{i=1}^{n} x_i^2 m_i - \left(\sum_{i=1}^{n} x_i m_i\right)^2 \\ \text{s.t. } &\sum_{i=1}^{n} m_i = 1, \\ &\inf M_i \le m_i \le \sup M_i \quad \forall i.\end{aligned} \tag{8}$$

To determine $\underline{\mathrm{Var}}[X]$, since we are minimizing a concave function, we will have to check the extreme feasible points of the convex outcome space. Because of the equality constraint, we can choose to let the value of one of the n variables be determined by values assigned to the other $n-1$ variables. For each of the remaining $n-1$ variables, there are two endpoints to be examined. Conceptually, for each interval there are two inequalities with two associated slack variables. A potential extreme point is generated by forcing one of the two slack variables for each interval to be zero. Thus, there are $n2^{n-1}$ potential extreme points. A potential extreme point may not satisfy all remaining constraints, but these can be checked quite quickly, and a potential extreme point is simply discarded if it doesn't satisfy all interval constraints. Variances are computed for all feasible extreme points and the minimum is saved.

3) Class IV Variance Bounds: For X in Class IVa and IVb the variance maximization model is similar to that of Class IIa; only the constraints differ. Therefore, the maximum variance model is

$$\begin{aligned}\overline{\mathrm{Var}}[X] = \max &\left[\sum_i [(\inf A_i)^2 m_{i1} + (\sup A_i)^2 m_{i2}] \right.\\ &\left. - \left(\sum_i [(\inf A_i) m_{i1} + (\sup A_i) m_{i2}]\right)^2\right] \\ \text{s.t. } &\inf M_i \le m_{i1} + m_{i2} \le \sup M_i \quad \forall i \\ &\sum_i m_{i1} + m_{i2} = 1, \quad m_{i1}, m_{i2} \ge 0 \quad \forall i.\end{aligned} \tag{9}$$

The minimum variance problem of Class IVa can be formulated as the following nonlinear program:

$$\begin{aligned}\underline{\mathrm{Var}}[X] = \min &\left[\sum_{i=1}^{n}\sum_{j=1}^{|A_i|} a_{ij}^2 m_{ij} - \left(\sum_{i=1}^{n}\sum_{j=1}^{|A_i|} a_{ij} m_{ij}\right)^2\right] \\ \text{s.t. } &\inf M_i \le \sum_j m_{ij} \le \sup M_i \quad \forall i \\ &\sum_i \sum_j m_{ij} = 1, \quad m_{ij} \ge 0 \quad \forall i, j\end{aligned} \tag{10}$$

where a_{ij} denotes the jth element of the outcome set A_i. Similar to the Class IIa case, the objective function is concave, so the nonlinear program (10) is not a standard quadratic program. Therefore, we apply a combination of Class IIa and III procedures to find the minimum variance. Following the algorithm suggested for Class IIa, we will examine the smaller set of distributions generated by focusing as much probability mass as possible toward one element of X_O. For each element, a Class III subproblem is generated. These Class III subproblems are solved by the previously described procedure.

The last bound of possible interest here is $\underline{\mathrm{Var}}[X]$ for X in Class IVb. This variance function is, in general, neither convex nor concave. A proof is given below. To proceed then, one approach is to replace the outcome intervals $A_i = [a_{i1}, a_{i2}]$ with sets $A_i' = \{a_{i1}', a_{i2}', \ldots, a_{ir}'\}$, where $a_{i1}' = a_{i1}, a_{ir}' = a_{i2}$, and $a_{i1}' < a_{i2}' < \cdots < a_{ir}'$. These replacements yield a Class IVa problem. Depending on the coarseness of the replacement sets, the resulting minimum may indeed be the global minimum, or may be an approximation of indeterminate quality.

Proposition 3: The variance function for a Class IVb ill-specified random variable X is, in general, neither convex nor concave.

Proof by Example: Suppose we are given two outcome intervals, $x_1 \in A_1$ and $x_2 \in A_2$, corresponding to two probability mass intervals $m_1 \in M_1$ and $m_2 \in M_2$, and we have $m_1 + m_2 = 1$. Then, $\mathrm{Var}[X] = x_1^2 m_1 + x_2^2 m_2 - (x_1 m_1 + x_2 m_2)^2$. The diagonal entries of the Hessian for $\mathrm{Var}[X]$, with respect to x_1, x_2, m_1, and m_2, are

$$\begin{bmatrix} 2m_1 - 2m_1^2 & \cdots & \cdots & \cdots \\ \cdots & 2m_2 - 2m_2^2 & \cdots & \cdots \\ \cdots & \cdots & -2x_1^2 & \cdots \\ \cdots & \cdots & \cdots & -2x_2^2 \end{bmatrix}.$$

If none of the variables are zero, the first principal minors are positive, positive, negative and negative. Thus, the function cannot be concave or convex. □

IV. Mean and Variance Bounds for a Function of Ill-Specified Random Variables

Often in risk modeling or economic analysis, the variable of interest is an arithmetic function of one or more ill-specified random variables. Having found optimal bounds for the mean and variance of each ill-specified random variable, we need to be able to propagate those bounds through the model. By combining variables two at a time and determining the mean and variance bounds of the arithmetic combination, we can work up to the complete model.

Exact and approximate functions for the mean and variance of certain arithmetic combinations of two random variables [33] are summarized in Tables IV and V. While useful for many cases, two issues must be addressed. First, approximations for the mean and variance of the quotient of two random variables can be quite poor, and lead us away from our focus on bounds. Second, determining the variance of the product of correlated random variables using the derived formula involves higher-order moments and knowledge of the

TABLE IV
IDENTIFYING BOUNDS ON THE MEAN OF ARITHMETIC FUNCTIONS OF IMPRECISE VARIABLES

Operation	Function *f* for determining mean	Possible critical points to examine in addition to extreme points
$Y = cX_i$	$\mu_Y = c\mu_i$	None
$Y = X_1 \pm X_2$	$\mu_Y = \mu_1 \pm \mu_2$	None
$Y = X_1 X_2$	$\mu_Y = \mu_1\mu_2 + \rho_{12}\sigma_1\sigma_2$	$\mu_1 = 0, \mu_2 = 0, \sigma_1 = 0,$ $\sigma_2 = 0, \rho_{12} = 0$
$Y = X_1 / X_2$	An approximation is $\mu_Y \cong \frac{\mu_1}{\mu_2} - \frac{1}{\mu_2^2}\rho_{12}\sigma_1\sigma_2 + \frac{\mu_1}{\mu_2^3}\sigma_2^2$	

TABLE V
IDENTIFYING BOUNDS ON THE VARIANCE OF ARITHMETIC FUNCTIONS OF IMPRECISE VARIABLES

Operation	Function *f* for determining variance	Possible critical points to examine in addition to extreme points
$Y = cX_i$	$\sigma_Y^2 = c^2\sigma_i^2$	None
$Y = X_1 + X_2$	$\sigma_Y^2 = \sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2$	$\sigma_1 = -\rho_{12}\sigma_2,\ \sigma_2 = -\rho_{12}\sigma_1,$ $\sigma_1 = 0, \sigma_2 = 0$
$Y = X_1 - X_2$	$\sigma_Y^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2$	$\sigma_1 = \rho_{12}\sigma_2,\ \sigma_2 = \rho_{12}\sigma_1,$ $\sigma_1 = 0, \sigma_2 = 0$
$Y = X_1 X_2$	If X_1 and X_2 are independent, $\sigma_Y^2 = \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2$ Otherwise, $\sigma_Y^2 = \mu_2^2\sigma_1^2 + \mu_1^2\sigma_2^2$ $+ 2\mu_1\mu_2 \operatorname{cov}[X_1, X_2] - (\operatorname{cov}[X_1, X_2])^2$ $+ E[(X_1 - \mu_1)^2 (X_2 - \mu_2)^2]$ $+ 2\mu_2 E[(X_1 - \mu_1)^2 (X_2 - \mu_2)]$ $+ 2\mu_1 E[(X_1 - \mu_1)(X_2 - \mu_2)^2]$	$\mu_1 = 0, \mu_2 = 0,$ $\sigma_1 = 0, \sigma_2 = 0$
$Y = X_1 / X_2$	An approximation is $\sigma_Y^2 \cong \frac{\sigma_1^2}{\mu_2^2} + \left(\frac{\mu_1}{\mu_2^2}\right)^2 \sigma_2^2 - 2\rho_{12}\left(\frac{\mu_1}{\mu_2^3}\right)\sigma_1\sigma_2$	

joint distribution of X_1 and X_2. In these cases, we suggest developing nonlinear optimization models to find the bounds on the mean and variance of the function. First, if necessary, slice the distribution into outcome intervals, assigning probability mass to each interval. Then denote the probability mass assignments for the ith outcome interval in X_1 as $m(A_{1,i})$ and the probability mass assignments for the jth outcome interval in X_2 as $m(A_{2,j})$. If X_1 and X_2 are independent, set $m(\{A_{1,i}, A_{2,j}\}) = m(A_{1,i})m(A_{2,j}) \forall i, j$. If X_1 and X_2 are dependent but we lack any information about the joint distribution, we incur additional decision variables and constraints: $\sum_i m(\{A_{1,i}, A_{2,j}\}) = m(A_{1,i})\ \forall i$, and $\sum_j m(\{A_{1,i}, A_{2,j}\}) = m(A_{2,j})\ \forall j$. For example, the max variance model for the quotient X_1/X_2, with no knowledge of the dependency relationship between the variables, may be stated as

$$\begin{aligned} \max z = & \sum_{i,j} (x_{1,i}/x_{2,j})^2 m(\{A_{1,i}, A_{2,j}\}) \\ & - \left(\sum_{i,j} (x_{1,i}/x_{2,j}) m(\{A_{1,i}, A_{2,j}\}) \right)^2 \\ \text{s.t. } & x_{1,i} \in A_{1,i} \quad \forall i \\ & x_{2,j} \in A_{2,j} \quad \forall j \\ & \sum_i m(\{A_{1,i}, A_{2,j}\}) = m(A_{1,i}) \quad \forall i \text{ and} \\ & \sum_j m(\{A_{1i}, A_{2j}\}) = m(A_{2j}) \quad \forall j. \end{aligned} \tag{11}$$

If there is additional uncertainty regarding probability masses, as in Class IV, we add constraints $m(A_{1,i}) \in M(A_{1,i})\ \forall i, m(A_{2,j}) \in M(A_{2,j})\ \forall j, \sum_i m(A_{1,i}) = 1$ and $\sum_j m(A_{2,j}) = 1$. As such a function in general lacks convexity or concavity, finding the global optima may be challenging. It would require an exhaustive search, evaluating the function at all combinations of interval endpoints, plus evaluation at all critical points where a partial derivative equals zero. Further research may be helpful here.

Returning to the cases in Tables IV and V, where an exact function is available and appropriate, let $\underline{\mu}_i, \overline{\mu}_i, \underline{\sigma}_i^2$, and $\overline{\sigma}_i^2$ denote the bounds on the mean and variance of X_i such that $\mu_i \in [\underline{\mu}_i, \overline{\mu}_i], \sigma_i \in [\underline{\sigma}_i, \overline{\sigma}_i], i = 1, 2$. The correlation coefficient $\rho_{12} \in [-1, 1]$, but additional information may allow one to specify a tighter interval $[\underline{\rho}_{12}, \overline{\rho}_{12}]$ on ρ_{12}. The various functions $f(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho_{12})$ are usually not strictly concave or convex, as may be seen by examining the principal minors (see, for example, [34]). Determining bounds on $f(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho_{12})$, then, requires one at most to evaluate f at the 32 extreme points $\{\{\underline{\mu}_1, \overline{\mu}_1\} \times \{\underline{\mu}_2, \overline{\mu}_2\} \times \{\underline{\sigma}_1, \overline{\sigma}_1\} \times \{\underline{\sigma}_2, \overline{\sigma}_2\} \times \{\underline{\rho}_{12}, \overline{\rho}_{12}\}\}$, as well as at the critical points, which are found by solving the set of partial derivatives equated to zero. The critical points may not all be permissible, given the intervals.

V. Example: Application in Environmental Risk Analysis

To illustrate the use of IMVPA technique in risk analysis, we consider the model developed by Copeland *et al.* [14] for determining the risk of exposure to contaminants via the inhalation pathway. This exposure risk, denoted *Dose-inh* (mg/kg·day), was modeled by

$$Dose\text{−}inh = GLC^{*}BIOinh^{*}ED^{*}EF/AT^{*}(RR/ABW). \tag{12}$$

Table VI summarizes characteristics of variables used by Copeland *et al.* in their model.

The exposure duration variable ED represents an empirical example of a random variable with an imprecise probability distribution. While it may not be unreasonable to assume that the distribution of, say, 23 to 33 year olds is uniform, we lack enough evidence to do so. To alleviate this ignorance one could try to locate the original works by the Census Bureau and the EPA, or one could assume a specific distribution over each interval, or one could proceed, assuming ignorance within each interval. We assume ignorance and treat ED as a Class IIb variable. The techniques of Section 3 bound the ED mean in the interval [8.72, 17.64] and its standard deviation in the interval [8.19, 18.79].

An obvious dependency exists between the respiration rate (RR) and the average body weight (ABW). For illustration purposes, we assume this dependency relationship is positive but, otherwise, unknown. Table VII also shows the bounds for the exposure risk *Dose-inh* with and without dependency between RR and ABW.

To determine the mean and variance bounds on RR/ABW, we slice the normal distributions into 100 intervals each, choosing interval endpoints, such that each interval has probability mass 0.01. The theoretically infinite distributions are arbitrarily truncated at the 0.1 percentile and the 99.9 percentile. We do this for compatibility comparisons with the Risk Calc implementation of PBA. Letting RR take the role of X_1 and ABW take the role of X_2, we have, for example, $A_{1,1} = [\Phi^{-1}_{20,2^2}(0.001), \Phi^{-1}_{20,2^2}(0.01)] = [13.8, 15.3]$, $A_{1,2} = [\Phi^{-1}_{20,2^2}(0.01), \Phi^{-1}_{20,2^2}(0.02)] = [15.3, 15.9], \ldots$, $A_{1,99} = [\Phi^{-1}_{20,2^2}(0.98), \Phi^{-1}_{20,2^2}(0.99)] = [24.1, 24.7]$, $A_{1,100} = [\Phi^{-1}_{20,2^2}(0.99), \Phi^{-1}_{20,2^2}(0.999)] = [24.7, 26.2]$. The bounds on the tabulated RR distribution are $\mu_1 \in [19.94, 20.06]$ and $\sigma_1^2 \in [3.65, 4.42]$. As the specified mean is 20 and the variance 4, we are introducing some imprecision here. The imprecision can be reduced by using more intervals. On the other hand, we are confident that we are maintaining bounds. In the case where we assume independence, $m(\{A_{1,i}, A_{2,j}\}) = m(A_{1,i})m(A_{2,j}) = 0.01 \cdot 0.01 = 0.0001\ i = 1, \ldots, 100, j = 1, \ldots, 100$. The bounds on the mean of RR/ABW are given by $\max/\min \sum_{i,j}(x_{1,i}/x_{2,j})0.0001$, s.t. $x_{1,i} \in A_{1,i}, x_{2,j} \in A_{2,j}, i = 1, \ldots, 100, j = 1, \ldots, 100$, or $[0.323, 0.331]$. The bounds on the variance are given by $\max/\min \sum_{i,j}(x_{1,i}/x_{2,j})^2 0.0001 - (\sum_{i,j}(x_{1,i}/x_{2,j})0.0001)^2$ s.t. $x_{1,i} \in A_{1,i}, x_{2,j} \in A_{2,j}, i = 1, \ldots, 100, j = 1, \ldots, 100$, or $[0.006\,09, 0.009\,32]$. In the case where we assume dependency, some structure can be added by considering just the case of maximal positive dependency, where $m(\{A_{1,i}, A_{2,j}\}) = 0.01$ if $i = j$ and 0 otherwise. That is, if independence implies the probability mass for a given RR outcome is uniformly spread amongst the j outcomes of ABW, then maximal dependence implies the probability mass for a given RR outcome is concentrated on just one outcome of ABW, sensibly here being that outcome which shows perfect positive correlation. Under these conditions, the mean of RR/ABW is bounded by [0.316, 0.324] and the variance is bounded by [0.001 08, 0.002 64]. With both of these cases, bounds on the quotient are combined with the bounds on the other factors of *Dose-inh* using the independent product function. Results are summarized in Table VII.

Copeland *et al.* approached the problem of estimating *Dose-inh* by using Monte Carlo simulation. We replicate their analysis by using the Crystal Ball [35] software package using 10 000 trials. Since generating a random sample from the probability distribution of ED requires specification of a probability distribution for each interval of that distribution, we follow the assumption of Copeland *et al.* and use a uniform distribution for each interval. In addition, we determine mean and standard deviation bounds using Risk Calc. The Risk Calc model is listed in Fig. 5. Table VIII shows the results obtained for the exposure risk using the three methods.

One may observe from Table VIII that both IMVPA and PBA intervals contain the simulation results. Under the assumption of independence among variables, the mean interval identified by IMVPA is 61% as wide as that identified by PBA. Moreover,

TABLE VI
INPUT DISTRIBUTIONS FOR VARIABLES OF RISK EXPOSURE TO CONTAMINANTS (ADAPTED FROM [14])

Parameters		Distribution		
Symbol	Description	Type	Relevant Descriptors	Notes
GLC	Concentration, measured 2 meters above the ground (mg/m^3)	Point estimate	1.99E-11	The point estimate is the highest (worst case) reading found in the vicinity.
BIOinh	Fraction of particulates retained in lung	Uniform	$[0.46, 1.0]$ $\mu = 0.73,\ \sigma = 0.16$	The authors' discussion suggests a uniform distribution over $[0.46, 0.59]$ may have been more realistic, arguing that a large portion of the mass of suspended particulates is exhaled and/or swallowed, but the reported model used the wider interval to ensure that the regulatory guideline value would be included.
ED	Exposure duration $(years)$	Custom	Years / Interval Probability 0-1 7.5% 1-3 16.9% 3-13 40.2% 13-18 11.0% 18-23 7.9% 23-33 9.5% > 33 7.0%	This distribution was based on a U.S. Census Bureau survey, as evaluated by the EPA. No discussion of the distribution within each of these year intervals was given. Copeland et al assume uniform distributions within each interval. However, the PBA and IMVPA analyses assume ignorance within each interval.
EF	Exposure frequency (fraction of a year)	Uniform	$[0.58, 1.0]$ $\mu = 0.79, \sigma = 0.12$	Specified by regulatory guidelines.
AT	Averaging time $(years)$	Static value	70	Specified by regulatory guidelines.
RR	Respiration rate (m^3/day)	Normal	$\mu = 20, \sigma = 2$	The distribution for this factor has not been found to be critical to modeling uncertainty and hence is typically replaced by a point estimate. We use a normal distribution here to illustrate the capability of IMVPA to model correlated parameters.
ABW	Average body weight (kg)	Normal	$\mu = 64.2, \sigma = 13.19$	Specified by regulatory guidelines.

the upper bound on standard deviation identified by IMVPA is 52% smaller than that identified by PBA. When a positive correlation between the respiration rate and the average body weight is assumed, these percentages are 45% and 58%, respectively. A brief explanation for these differences is as follows. In Fig. 1, for example, the maximum variance, using Risk Calc, would be computed using the points on the lower part of the left bound and the upper part of the right bound. These points do not follow a single realizable lognormal CDF. With IMVPA, the optimally-determined bounds would correspond to a possible realization of a distribution. Thus, while PBA has the advantage of maintaining a set of bounds at all percentiles, its mean and variance functions operate simplistically with the bounds.

VI. SUMMARY AND CONCLUSION

A general category of ill-specified random variables has been identified that encompasses a variety of imprecise representations, including fuzzy numbers, percentile estimates, and Dempster-Shafer multivalued mappings. Within the general category, classes are defined in terms of ill-specificity concerning outcomes (set-valued or interval), probabilities, or both. Algorithms and models have been developed to determine the optimal bounds for the mean and variance of ill-specified random variables in each of these classes.

Uncertainty can be propagated through a function of ill-specified and/or precise random variables using the proposed

TABLE VII
IMVPA CALCULATION RESULTS FOR EXAMPLE

Ill-Specified Random Variable or Arithmetic Expression	$\underline{\mu}$	$\overline{\mu}$	$\underline{\sigma}$	$\overline{\sigma}$
ED	8.72	17.64	8.19	18.79
Intermediate results, assuming independence among all variables				
*GLC * BIOinh*	1.45E-11	1.45E-11	3.10E-12	3.10E-12
*(GLC * BIOinh) * ED*	1.27E-10	2.56E-10	1.25E-10	2.85E-10
*((GLC * BIOinh) * ED) * EF*	1.00E-10	2.02E-10	1.01E-10	2.30E-10
*(((GLC * BIOinh) * ED) * EF) / AT*	1.43E-12	2.89E-12	1.44E-12	3.28E-12
RR / ABW	3.23E-01	3.31E-01	7.80E-02	9.66E-02
*((((GLC * BIOinh) * ED) * EF) / AT)* (RR / ABW) = Dose-inh*	4.62E-13	9.58E-13	4.91E-13	1.17E-12
Intermediate results, assuming a positive dependence between RR and ABW				
RR / ABW	3.16E-01	3.24E-01	3.28E-02	5.13E-02
*((((GLC * BIOinh) * ED) * EF) / AT)* (RR / ABW) = Dose-inh*	4.51E-13	9.36E-13	4.60E-13	1.08E-12

```
tOp = .999                                  // truncate infinite distributions at the 99.9 percentile
bOt = .001                                  // truncate infinite distributions at the 0.001 percentile
GLC = 1.99e-11 mg m{-3}                     // concentration of contaminant
BIOinh = uniform(0.46,1)                    // fraction of particulates retained in lung
ED = @(0,0) (0,0.075) (1,0.075) (1,0.244) (3,0.244) (3,0.646) (13,0.646) (13,0.756) (18,0.756) (18,0.835)
(23,0.835) (23,0.93) (33, 0.93) (33,1) (1,0) (1,0.075) (3,0.075) (3,0.244) (13,0.244) (13,0.646) (18,0.646)
(18,0.756) (23,0.756) (23,0.835) (33,0.835) (33,0.93) (70,0.93) (70,1)@ years
                                            // exposure duration
EF = uniform(0.58, 1)                       // exposure frequency (fraction of a year)
ABW = normal(64.2 kg, 13.19 kg)             // average body weight
AT = 70 years                               // averaging time
RR = normal(20 m{3} day{-1}, 2 m{3} day{-1}) // respiration rate
// exposure to contaminants via inhalation assuming independence among all variables
x1 = RR |/| ABW
Idose = GLC |*| BIOinh |*| ED |*| EF |/| AT |*| x1
mean(Idose)                                 // return an interval bounding the mean
stddev(Idose)                               // return an interval bounding the std. deviation
// exposure to contaminants via inhalation assuming dependence between RR and ABW
x2 = RR / ABW
dose =  GLC |*| BIOinh |*| ED |*| EF |/| AT |*| x2
mean(dose)
stddev(dose)
```

Fig. 5. PBA model.

TABLE VIII
COMPARISON OF INHALATION DOSE (MG/KG-DAY) ESTIMATES IDENTIFIED BY INTERVAL MEAN-VARIANCE PROPAGATION ANALYSIS, PBA, AND SIMULATION

	$\underline{\mu}$	$\overline{\mu}$	$\underline{\sigma}$	$\overline{\sigma}$
A. Assuming independence among all variables				
IMVPA	4.62E-13	9.58E-13	4.91E-13	1.17E-12
PBA	4.08E-13	1.22E-12	4.13E-13	2.44E-12
Simulation	7.14E-13		8.52E-13	
B. Assuming a dependence between RR and ABW				
IMVPA	4.51E-13	9.36E-13	4.60E-13	1.08E-12
PBA	3.32E-13	1.42E-12	3.13E-13	2.56E-12
Simulation	6.98E-13		8.00E-13	

IMVPA method. Indeed, in many analyses of uncertainty, a risk variable is modeled as an arithmetic function of several variables, some of which may be ill-specified. The value of IMVPA is in providing an analytical bound for the first and second moments of a risk variable. Bounded moments extend and complement traditional probabilistic analyzes, capturing both the variability associated with a precisely-specified distribution and the uncertainty associated with incomplete knowledge about distribution parameters or outcome/probability mappings. When risk assessments are made, and resources are committed to address these risks, it is important not to overstate the risks. IMVPA complements sensitivity analysis (which may not discover the mean/variance bounds), simulation (which requires precise input probability distributions and generates point estimates, not bounds), and PBA (which maintains CDF bounds at all percentiles but computes mean/variance bounds that may be too wide and not realizable) with mean/variance bounds determined by optimization methods.

REFERENCES

[1] J. O. Berger, "Robust Bayesian analysis: Sensitivity to the prior," *J. Statis. Planning Inference*, vol. 25, pp. 303–328, 1990.

[2] T. E. McKone and P. B. Ryan, "Human exposures to chemicals through food chains: An uncertainty analysis," *Environ. Sci. Technol.*, vol. 23, pp. 1154–1163, 1989.

[3] G. W. Suter II, *Ecological Risk Assessment*. Boca Raton, FL: Lewis Publishers, 1993.

[4] E. T. Jaynes, *The Maximum Entropy Formalism*, R. D. Levine and M. Tribus, Eds. Cambridge, MA: MIT Press, 1979.

[5] R. C. Lee and W. E. Wright, "Development of human exposure-factor distributions using maximum-entropy inference," *J. Exposure Anal. Environ. Epidemiol.*, vol. 4, pp. 329–341, 1994.

[6] I. J. Good, *The Estimation of Probabilities*. Cambridge, MA: MIT Press, 1965.

[7] J. C. Helton, "Treatment of uncertainty in performance assessments for complex systems," *Risk Anal.*, vol. 14, pp. 483–511, 1994.

[8] R. C. Williamson and T. Downs, "Probabilistic arithmetic. I. Numerical methods for calculating convolutions and dependency bounds," *Int. J. Approx. Reason.*, vol. 4, pp. 89–158, 1990.

[9] S. Ferson, L. Ginzburg, and R. Akcakaya, "Whereof one cannot speak: When input distributions are unknown," *Appl. Biomed.*, working paper, Setauket, NY.

[10] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[11] J. A. Cooper, "Fuzzy algebra uncertainty analysis for abnormal environmental safety assessment," *J. Intell. Fuzzy Syst.*, vol. 2, pp. 337–345, 1994.

[12] R. E. Moore, *Interval Anal.*. Englewood Cliffs, NJ: Prentice-Hall, 1966.

[13] F. F. Choobineh and A. Behrens, "Use of intervals and possibility distributions in economic analysis," *J. Oper. Res. Soc.*, vol. 43, no. 9, pp. 907–918, 1992.

[14] T. L. Copeland, A. M. Holbrow, J. M. Otani, K. T. Connor, and D. J. Paustenbach, "Use of probabilistic methods to understand the conservatism in California's approach to assessing health risks posed by air contaminants," *J. Air Waste Manage. Assoc.*, vol. 44, pp. 1399–1413, 1994.

[15] F. Shi, M. Mazumdar, and J. Bloom, "Asymptotic mean and variance of electric power generation system production costs via recursive computation of the fundamental matrix of a Markov chain," *Oper. Res.*, vol. 47, no. 5, pp. 703–712, 1999.

[16] S. Ryan and M. Mazumdar, "Chronological influences on the variance of electrical power production costs," *Oper. Res.*, vol. 40, no. S2, pp. S284–S292, 1992.

[17] M. D. Springer, *The Algebra of Random Variables*. New York: Wiley, 1979.

[18] M. G. Morgan and M. Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge, MA: Cambridge Univ. Press, 1990.

[19] N. C. Rowe, "Absolute bounds on the mean and standard deviation of transformed data for constant-sign derivative transformations," *SIAM J. Sci. Stat. Comput.*, vol. 9, pp. 1098–1113, 1988.

[20] J. E. Smith, "Generalized Chebychev inequalities: Theory and applications in decision analysis," *Operations Research*, vol. 43, no. 5, pp. 807–825, 1995.

[21] P. Walley, "Statistical reasoning with imprecise probabilities," in *Monographs on Statistics and Applied Probability*. London, UK: Chapman and Hall, 1991, vol. 42.

[22] U. Saxena, "Investment analysis under uncertainty," *Eng. Econ.*, vol. 29, no. 1, pp. 33–40, 1983.

[23] Z. W. Kmietowicz and A. D. Pearman, *Decision Theory and Incomplete Knowledge*, Aldershot, UK: Gower Publishing, 1981.

[24] M. Lavine, "Sensitivity in Bayesian statistics: The prior and the likelihood," *J. Amer. Stat. Assoc.*, vol. 86, pp. 396–399, 1991.

[25] A. T. Langewisch and F. F. Choobineh, "Stochastic dominance tests for ranking alternatives under ambiguity," *Eur. J. Oper. Res.*, vol. 95, pp. 139–154, 1996.

[26] S. Ferson, *RAMAS Risk Calc 4.0 Software: Risk Assessment with Uncertain Numbers*. Boca Raton, FL: Lewis Publishers, 2002.

[27] W. T. Tucker and S. Ferson, *Probability Bounds Analysis in Environmental Risk Assessments*. Setauket, NY: Applied Biomathematics, 2003.

[28] H. A. Taha, *Operations Research*, 5th ed. New York: Macmillan, 1992.

[29] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *Annals Math. Stat.*, vol. 38, pp. 325–339, 1967.

[30] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton Univ. Press, 1976.

[31] R. T. Clemens, *Making Hard Decisions*, 2nd ed. Belmont, CA: Duxbury Press, 1996.

[32] G. Franke, "Expected utility with ambiguous probabilities and 'irrational' parameters," *Theory and Decision*, vol. 9, pp. 267–283, 1978.

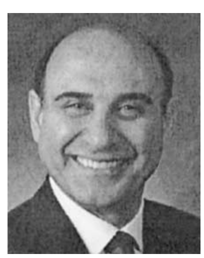
[33] A. M. Mood, F. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics*, 3rd ed. Boston, MA: McGraw-Hill, 1974.

[34] W. L. Winston, *Operations Research: Applications and Algorithms*, 3rd ed. Belmont, CA: Duxbury Press, 1994.

[35] *Crystal Ball v. 4.0*. Denver, CO: Decisioneering, 1996.

Andrew Langewisch received the B.A. degree in mathematics from Concordia University, Seward, NE, in 1982, the M.B.A. degree from the University of Michigan, Ann Arbor, in 1985, and the Ph.D. degree in industrial and management systems engineering from the University of Nebraska, Lincoln, in 1998.

He represented the University of Nebraska at the INFORMS Doctoral Colloquium, 1998. He has taught full-time in the Department of Business Administration, Concordia University since 1985. His research interests are concentrated on representing uncertainty in business and engineering problems.



F. Fred Choobineh (M'89) received the B.S.E.E., M.S.I.E., and Ph.D. degrees from Iowa State University, Ames, in 1972, 1976, and 1979, respectively.

He is a Professor of Industrial and Management Systems Engineering at the University of Nebraska, Lincoln, where he also holds a courtesy appointment as a Professor of Management. His research interests are in design and control of manufacturing systems and use of approximate reasoning techniques in decision making.

Prof. Choobineh is a Fellow of the Institute of Industrial Engineers (IIE) and a Member of INFORMS.