

Different methods are needed to propagate ignorance and variability

Scott Ferson^a & Lev R. Ginzburg^b

^a*Applied Biomathematics, 100 North Country Road, Setauket, New York 11733, USA*

^b*Department of Ecology and Evolution, State University of New York, Stony Brook, New York 11794, USA*

There are two kinds of uncertainty. One kind arises as variability resulting from heterogeneity or stochasticity. The other arises as partial ignorance resulting from systematic measurement error or subjective (epistemic) uncertainty. As most researchers recognize, variability and ignorance should be treated separately in risk analyses. Although a second-order Monte Carlo simulation is commonly employed for this task, this approach often requires unjustified assumptions which may be inappropriate in some circumstances. We argue that the two kinds of uncertainty should be propagated through mathematical expressions with different calculation methods. Basically, interval analysis should be used to propagate ignorance, and probability theory should be used to propagate variability. We demonstrate how using an inappropriate method can yield erroneous results. We also show how ignorance and variability can be represented simultaneously and manipulated in a coherent analysis that does not confound the two forms of uncertainty and distinguishes what is known from what is assumed. © 1996 Elsevier Science Limited.

1 IGNORANCE AND VARIABILITY ARE DIFFERENT KINDS OF UNCERTAINTY

Comprehensive taxonomies of uncertainty have been offered by several authors.^{1–5} These classifications recognize many distinct kinds of uncertainty with considerable subtlety. For our purposes, it will be sufficient to follow Casti⁶ in recognizing two basic kinds of uncertainty that are fundamentally different from each other. We distinguish objective uncertainty arising from variability of the underlying stochastic system as against subjective or epistemic uncertainty resulting from our not having complete information about that system.

Table 1 gives questions exemplifying these two kinds of uncertainty in two aspects of an extinction risk analysis⁷ for an endangered species of owls. The sample questions in the first column refer to variability expressed over time and across space. This column could have been split into two columns to represent temporal and spatial variability separately if one so desired, with obvious examples for each new cell in the table. In fact, because variability can be expressed over almost any dimension, we could have multiplied the number of columns ad libitum. We have collapsed all such examples of uncertainty due to variability into one column for the sake of simplicity and to draw

attention to the fact that the same mathematical techniques are used to propagate uncertainty whether the value changes over time, across space, among individuals or on some other axis of variability.

We use the term ‘ignorance’ to denote the partial incertitude that arises because of limits on empirical study or mensurational precision. For instance, at some moment in time the number of owls present in a well defined region of forest is a particular number that is not varying. Nevertheless, this number may not be precisely known to us, just because it can be extremely difficult to tally every single bird. This uncertainty is decidedly unlike the uncertainty, say, in mortality rates arising from variability of the weather. For instance, ignorance and variability respond differently to empirical effort. Whereas ignorance can usually be reduced by additional study or by improving the techniques of measurement, variability has an objective reality that is independent of our empirical study of it. Additional effort may yield a better estimate of the magnitude of variability, but it will not tend to reduce it.

Although there may sometimes be epistemologically complex situations in which it is hard to distinguish between the two kinds of uncertainty, it is often fairly straightforward to do so. Despite this, both forms of uncertainty will be associated with many quantities of

Table 1. Uncertainty

	Variability	Ignorance
Model formulation	Do mortality mechanisms change from season to season?	Which model of density dependence should be used?
Parameter values	How does the number of owls vary in different parts of the forest?	What is the number of owls present in the forest?

interest in risk analysis problems. For example, if the future population size of owls is estimated from the mortality rate and the current population size, it is characterized by both randomness and incertitude. The relative magnitudes of ignorance and variability for any particular parameter of interest will depend on how well it has been studied and on the intrinsic stochasticity of the underlying system. In an analysis, we usually cannot control how much ignorance there is about a given parameter, nor how much it varies.

The purpose of this paper is to discuss some empirical and computational issues that arise when both kinds of uncertainty enter into a risk analysis problem. To motivate this discussion, we consider some very simplified numerical examples. Although these idealized examples are themselves unrealistic as risk analysis problems, they will help us to clarify certain important issues that commonly arise in real analyses. In the next two sections, we point out how the two kinds of uncertainty behave differently in calculations. The following sections then consider several methodological approaches to recognizing and propagating uncertainty through mathematical calculations.

2 CLASSICAL PROBABILITY THEORY INCORRECTLY PROPAGATES IGNORANCE

As risk analysts, how should we do arithmetic with numbers about whose values we are unsure? We argue that the answer depends in part on why we are unsure and whether the notion of frequency is applicable to the situation under study. Let us start this discussion with a simple question:

If the parameter A is a number somewhere between 0.2 and 0.4, and the parameter B is a number somewhere between 0.3 and 0.5, what is the value of the product AB ?

How we go about getting an answer to such a question depends on what we believe about the parameters involved and what the nature of our uncertainty about them is. For instance, if we suppose both A and B are actually fixed quantities whose values we just don't happen to know, then we might use interval analysis⁸⁻¹³ to arrive at an answer to the question. This approach asks about the possible range of the product given the stated ignorance about A and

B . In this case, the smallest possible value would be obtained when A is 0.2 and B is 0.3, yielding the product 0.06. The largest possible value would be $0.4 \times 0.5 = 0.2$. No other pair of numbers from the respective intervals yield a product outside this range. Thus, the answer is that the product is a number somewhere between 0.06 and 0.2. Figure 1(a) depicts this interval.

On the other hand, if we think the parameters A and B are varying randomly, then we might use a very different approach to arrive at an answer. Under Laplace's Principle of Insufficient Reason, the uncertainty about each parameter should be modeled with a uniform distribution. Choosing any other distributions would constitute an assertion of additional knowledge about the parameters. These distributions are then combined by the ordinary rules of probability theory.¹ In particular, if we assume they are independent, we can use probabilistic convolution or some Monte Carlo strategy to estimate the distribution of the product AB . The result is shown as a probability density function in Fig. 1(b). Figure 1(c) is the cumulated form of the same probability distribution, which is slightly more convenient for the following discussion because the ordinate always ranges between zero and one.

Although some probabilists¹⁴ have suggested probability theory is irrelevant to manipulating bounds on unknown values, others¹⁵⁻¹⁷ have insisted with some vehemence that probability theory is the only consistent calculus for propagating uncertainty of any sort. We suspect both of these views are incorrect and unnecessarily restrictive. We show how both interval analysis and probability theory will arrive at the same answer if we are careful to distinguish what is known from what is assumed.

The answers from interval analysis and probability theory agree in one sense. They both say the value of the product must lie between 0.06 and 0.2. The support of the probability distribution (i.e., the range of values for which probabilities are nonzero) is identical to the result of the interval analysis. Yet the answers are clearly different. The answer given by the probabilistic approach suggests that the product is much more likely to be a value near the central tendency than to be one of the extreme values. There is a clear concentration of probability mass in the centre of the distribution. Yet there is nothing in the

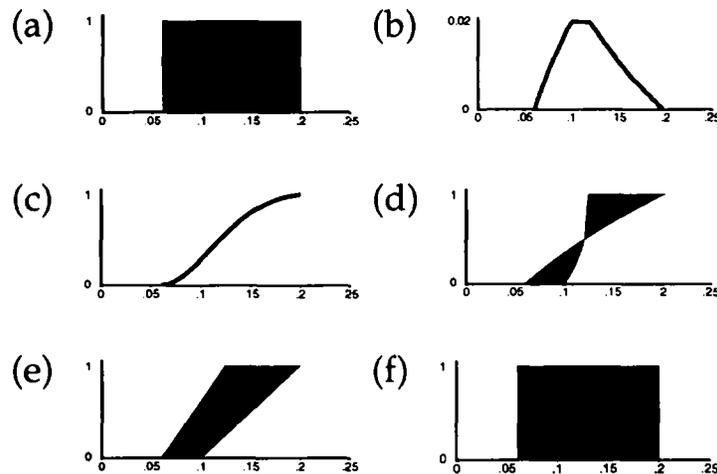


Fig. 1. (a) Depiction of the interval $[0.06, 0.2]$ which is guaranteed to contain the product of two uncertain numbers $A = [0.2, 0.4]$ and $B = [0.3, 0.5]$. Note that this interval is not the same as a uniform probability distribution. (b) The probability distribution of the product of random variables $A = \text{uniform}(0.2, 0.4)$ and $B = \text{uniform}(0.3, 0.5)$ under an assumption of independence. (c) The cumulative form of the probability distribution of this product. (d) The smallest region guaranteed to contain the cumulative distribution of the product assuming the dependence between A and B is some linear correlation between -1 and $+1$. (e) The smallest region guaranteed to enclose the cumulative distribution of the product without making any assumption about the stochastic dependence between A and B . (f) The smallest region guaranteed to enclose the cumulative probability distribution of the product of A and B given no information other than the bounds on both parameters. Note that this region is very different from a uniform distribution.

original statement of the question that obviously justifies this concentration. Let us examine just where this concentration comes from and whether it is justified.

We assumed independence between parameter A and parameter B . Perhaps this assumption accounts for the observed central concentration of probability. We can explore this idea by varying the correlation between A and B over the range of possible correlations. In general, the possible range of correlations depends on the shapes of the two univariate random distributions.¹⁸ In our case, correlations between A and B can be any value over the full theoretical range between $+1$ and -1 . Using variance reduction¹⁹ and dispersive Monte Carlo sampling,²⁰ we can calculate bounds on the cumulative distribution function that can result from the product of A and B . Figure 1(d) shows the region circumscribed by these bounds. No matter what the magnitude of the correlation between A and B , the probability distribution of their product must lie within the black region. This region is clearly much narrower than the interval depicted in Fig. 1(a), so we see that ignoring correlation cannot, by itself, have accounted for the discrepancy between the interval result and probability result.

Of course, correlation only measures linear associations between the parameters. Nothing in the original question constrained us to assume that only linear associations could exist between the two parameters. What are the bounds on the cumulative distribution function given any possible statistical

dependency between A and B ? This question is very similar to a problem first posed by Kolmogorov which has recently been solved.²¹ Applying dependency bounds analysis,^{22,23,20} we find that the distribution of the product of A and B must lie somewhere in the black region depicted in Fig. 1(e). This result shows that, no matter what statistical dependency we assume between the two parameters, or even if we do not assume anything at all about their interdependence, we still get an answer that is narrower than suggested by interval analysis and which exhibits a persistent albeit weaker concentration of probability mass in the centre of the range.

Perhaps the concentration of probability arises from our choice of uniform distributions to model the parameters. Given the paucity of information about A and B , however, it is hard to see how we could have selected any other distributions. Indeed, a long intellectual tradition in probability theory dating back to Laplace²⁴ himself demands that uniform distributions be used as the default in the absence of specific information about the frequencies of possible values. More recent theoretical arguments²⁵⁻²⁹ based on a criterion of maximizing information entropy reaffirm the selection. If we had ignored this tradition and the recent arguments, and assumed the distributions were, say, normal or lognormal, the resulting concentration of probability would have been even more pronounced. It turns out that the problem is not so much that we have assumed uniformity for the two distributions. The problem comes because we assumed particular shapes for them.

One need not assume specific shapes for the distributions if there is no basis for doing so. Recently developed methods based on interval probabilities³⁰⁻³⁶ allow us to express incertitude (ignorance) about probability itself. It is also possible to do calculations with these interval probabilities.³⁷ Figure 1(f) shows the region in which the probability distribution of the product AB must lie, given only the information about bounds on A and B . This result states that any value or distribution of values between 0.06 and 0.2 is possible and that we cannot say anything about which values are more likely than any others within this range. This interval-probabilistic approach thus agrees completely with the original result in Fig. 1(a) obtained from elementary interval analysis. They both represent an unknown quantity that may or may not be varying but whose value(s) we know to be limited to a specific range. Although the results in Fig. 1(a) and (f) were obtained by completely different computational approaches, the interpretations of the results are identical.

We see, finally, that the solutions in Fig. 1(b)–(e) are incorrect, because they assume more information than was given in the original question. In this sense, they are the result of wishful thinking, rather than a careful analysis of what is actually known. This example illustrates what may be a widespread problem with applying classical probability theory in risk analyses where the relevant empirical information is sorely incomplete (as is usually the case). Using classical probability theory to estimate even the simple product of two uncertain parameters requires several assumptions, without which no answer could be obtained. In practice, such assumptions are rarely even stated, much less tested or justified in any way. It seems obvious to us that, unless there is specific empirical information or theoretical argument to justify such assumptions, the results they produce could never be scientifically defensible. Although they may parade as precise applications of formal probability theory, the ensuing risk analyses conceal a considerable amount of unstated uncertainty that bears directly on the interpretation of the results.

The case in which some probability distributions are poorly characterized is undoubtedly common, and it may be typical in risk analysis, especially for new hazards and recently recognized environmental threats. For instance, because drilling is usually very expensive, analysis of groundwater contamination may be limited by very small sample sizes. Likewise, human effects of putative environmental hazards are typically difficult to characterize statistically because of practical or ethical constraints on data collection. Perhaps even more commonly, the correlations and statistical dependencies among variables remain unmeasured and unexplored even for long studied problems. For instance, probabilistic risk assessments

for nuclear power operations³⁸ still often neglect common-cause, common-mode and cascading failures³⁹ which introduce unknown dependencies among the variables in the analysis. Without specific analysis, there is no way to foretell whether inattention to such phenomena causes wasteful overestimation⁴⁰ or dangerous underestimation²⁰ of risks.

Yet risk analyses must be performed even when empirical information is extremely sparse. The conundrum then is how should we propagate uncertainty under such conditions. Should we assume the uncertainty can always be treated like stochastic variability? We have seen that doing so can lead to unjustified results when the requisite assumptions of probability theory are untested. Huff⁴¹ summarized the danger: 'Knowing nothing about a subject is frequently healthier than knowing what is not so...'. On the other hand, relying on the calculational methods of interval analysis in all circumstances will result in unnecessarily conservative answers when those assumptions are true.

3 UNCERTAINTY DOES NOT IMPLY FLUCTUATION

Tossing a coin provides a good example of the difference between variability and ignorance in risk analyses. If each toss is an independent event and the probability of tossing heads is known to be p and the probability of tossing tails $(1-p)$, the binomial distribution gives both the expected number of heads after n tosses and the probability that the number of heads will exceed the number of tails by some quantity. We are able to perform a perfect risk analysis for this textbook case. The solid curves in Fig. 2 describe the means and 95% confidence intervals for how many more heads than tails are expected after a given number of tosses of a fair coin ($p=0.5$). Although the expected mean is consistently zero, the 95% bounds widen as the square root of the number of tosses. It is, of course, the independence assumption that makes variance grow linearly and standard deviation grow as the square root.

Now suppose that the initial estimate of p was slightly inaccurate. For instance, suppose we failed to detect that the coin was a little biased. At first, the difference between the actual and the expected tosses is negligible. Their divergence is linear with the number of tosses, however, so that the eventual excess of heads over tails is very likely to lie well outside the range predicted by the risk analysis. In this case, although the classical analysis is fine in the short run, the measurement error in the estimate of p plays a dominant role and limits our ability to make reliable risk analyses in the long run. The uncertainty due to

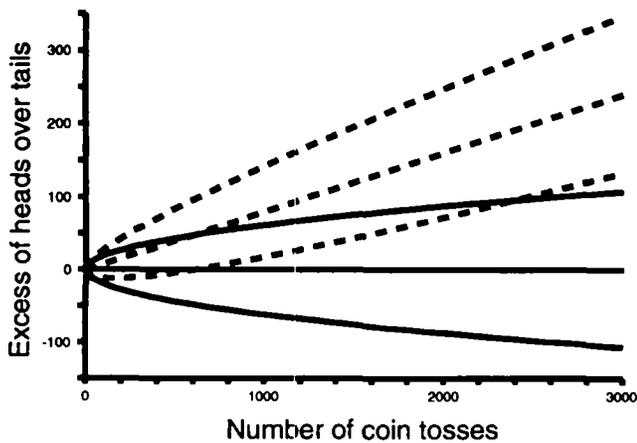


Fig. 2. Comparison of the uncertainties in coin tossing due to bias and due to fluctuation. Solid lines are for $p = 0.5$, dashed lines for $p = 0.54$. Each set shows the expected excess of heads over tails and the ranges within which the excess is likely to lie 95% of the time. The ranges expand as the square root of the number of tosses, while the expected means diverge linearly. For a small number of tosses, risk analyses for the two values of p are similar. The effect of bias begins to dominate as the number of tosses gets larger.

variability, reflected by the width of the 95% confidence bounds, grows much more slowly than the uncertainty due to ignorance about the value of p , which is reflected in the difference between the solid and dotted sets of lines. The reason for this is simply that random variations partially cancel with each other, while uncertainty accumulates linearly.

The dominance of ignorance over variation emerges in this example because the effects are iterated over time. The effects will also be iterated over the number of input parameters, so ignorance will tend to become more important as the number of inputs and calculations in a risk model increase. Therefore, like distant extrapolations, the reliability of complex risk calculations may be more controlled by ignorance than by variability. But even for a fixed problem and model, the relative importance of the two kinds of uncertainty depends on their respective magnitudes. Poor empirical study and imprecise measurements can yield a situation in which the ignorance already dominates the problem.

Variation and ignorance behave differently in a risk analysis. Random variation partially cancels itself out, but ignorance does not. This is the reason, for instance, that averaging over repeated measurements of a quantity will tend to minimize the random component of its measurement error, yet does nothing to reduce any systematic error that might be present. The coin example illustrates how cancellation can occur over time. Lucky streaks of heads are counterbalanced by subsequent bad luck with tails so that one expects to remain fairly close to even.

Partial cancellation also occurs among variables in a

risk calculation. For instance, if A and B in the previous section really were randomly varying, it would be perfectly correct to expect that values in the middle of the range $[0.06, 0.2]$ would be more likely than the end points of the range where both variables are simultaneously extreme. This expectation does not hold, however, when the uncertainty is ignorance rather than variability. We cannot assume that the systematic error in one variable will compensate for the systematic error in another variable. They might, but it would be illegitimate to assume that they must. For instance, suppose there is bias in the estimate of juvenile mortality in our extinction analysis for the endangered owl. If there is also bias in the estimate for adult mortality, is it reasonable to assume that error on the high side in one will tend to be balanced by error on the low side in the other? This would be reasonable if the errors were random, but they are systematic so we cannot expect to be so lucky. In fact, because adult and juvenile mortality are generally estimated by similar methods, one would anticipate the biases might well be in the same direction and thus reinforce rather than balance one another.

Just because we are uncertain about a quantity does not mean that it is fluctuating,⁴² and it certainly does not mean it is random. But using the uncertainty calculus of probability theory generally requires one to assume the quantities involved are random. Without this assumption, probability theory would not be of much use in formulating any estimates. Yet when we accept these assumptions without any empirical basis for doing so, we introduce error into the analysis and provide a target for criticism. Insofar as we are ignorant to some degree about the value of a quantity, we cannot discount the possibility that the uncertainty consists wholly of systematic error. The methods of probability theory cannot legitimately be used to propagate such error as though it were random.⁴³⁻⁴⁶ Only when we have some evidence or argument that the uncertainty is composed predominantly of random error, would those methods then be appropriate. In the next section we consider a way to empirically distinguish random and systematic error.

4 SYSTEMATIC AND RANDOM ERROR CAN BE DISTINGUISHED EMPIRICALLY

The theory of measurements⁴⁶ is based on three axioms about the true value of a measurand:

1. the true value exists,
2. the true value is constant, and
3. the true value cannot be found.

These axioms imply that there will always be measurement error which is the difference between

the measured value and the true value. We can never know the actual size of a measurement's error because, if we did, we would then know the true value which is unknowable. Nevertheless, we can still think about it in theoretical terms. Measurement error is often partitioned into two major components⁴³⁻⁴⁶

$$e_{\text{measurement}} = e_{\text{systematic}} + e_{\text{random}}$$

A systematic measurement error is constant or changes in a regular way over repeated measurements of a single quantity. A random measurement error, on the other hand, varies or would vary from measurement to measurement of a single quantity. There may be other components to the error of a measurement. For instance, there may occasionally be gross errors (so called because they are larger than is justifiable by measurement conditions) which can be detected and eliminated by outlier tests. And a few data sets have even been known to contain blunders (e.g., slip of the pen, etc.) which have to be eliminated by diligence and rechecking. Of course, gross errors and blunders should never be common, but, as the axioms imply, measurement error is in a strong sense inevitable. This is to say that either some systematic error or some random error, and likely both, will always be present to some extent. It is widely believed that systematic error is consistently underestimated in empirical studies.^{1,47,44}

Systematic and random errors have different consequences for measurement. When random errors are small, measurements of a single quantity under a single set of conditions will be close to one another. The measurement is said to be repeatable. When both random and systematic errors are small, measurements of a single quantity under different conditions (e.g., at different locations, by different investigators, or using different equipment) will be close to one another. In this case, the measurement is said to be reproducible. Good repeatability implies random errors are small. Good reproducibility implies that both random and systematic errors are small. From these considerations, we can deduce that systematic errors are large whenever measurements have good repeatability but poor reproducibility. This means that it is possible in principle to assess the relative magnitude of systematic error. The situation in which systematic error is fairly large is probably very common, at least in field studies collecting biological and environmental measurements.

5 SECOND-ORDER MONTE CARLO METHODS REQUIRE UNJUSTIFIED ASSUMPTIONS

A second-order Monte Carlo simulation is just a Monte Carlo simulation in which every replicate is

itself the result of a Monte Carlo simulation. (Special strategies such as Latin hypercube sampling¹ are often employed for the sake of efficiency, but the mathematical results should be the same asymptotically whatever algorithm is used.) This approach has been employed to explore the effect of an analyst's uncertainty about the parameters used to define the input distributions in risk analyses.^{48-51,4} For instance, in assessing the effects of measurement error on an extinction analysis for the endangered owl, we might select mean values for the reproduction and mortality rates from random distributions to use in a simulation that yields an estimate of the distribution of extinction risks. Repeatedly sampling the means to use in simulations yields a distribution of distributions that express uncertainty both from measurement error and from variability in the underlying population dynamics. It seems very reasonable to use a second-order method in this way to model uncertainty about the probabilities themselves. Indeed, the approach has been used for some time.^{52,53} There are, however, three practical problems one faces when employing a second-order simulation: (1) parameterization issues, (2) computational complexity, and (3) interpretational difficulty.

Parameterization of a second-order Monte Carlo simulation can require an analyst to supply considerably more specifications. Each uncertain parameter in the underlying simulation is replaced by a statistical distribution, whose form and parameters must then be specified. In principle, the cross correlations among the distributions of parameters should also be given, although independence is typically assumed in most cases. Given that the values of the second-order parameters are likely to be guesses, are not they too subject to uncertainty? Are we then facing a third-order problem? Obviously, the approach logically leads to an entire hierarchy of uncertainties,^{32,54,50} any of which has the possibility of substantially altering the final analysis. It is not clear at what point such deliberation degenerates into a merely fanciful exercise without much scientific content.

The second and third practical problems are intertwined with one another. A second-order simulation is a computational problem of roughly squared complexity compared to an ordinary Monte Carlo simulation. Each replicate in the outer loop of the simulation implies an entire Monte Carlo simulation in the inner loop. Given that at least a few hundred simulations are needed to yield a reasonably reliable picture of a statistical distribution, this translates into a substantial computational burden that might require hours or even days of time in a microcomputer implementation of a complex issue. In principle, any second-order simulation can be reexpressed as a first-order simulation^{54,55} so it should

be possible to reduce the computational burden. However, the output summaries from such a simplified simulation no longer distinguish uncertainty from the two levels and therefore mix the two kinds of uncertainty. For instance, the distribution of the extinction risk distributions would be condensed into a single distribution that does not give the probability of the owl population falling below some threshold size, but instead gives the probability that our predictions are inaccurate enough that the population would appear to fall that low, whether due to real statistical variation or due to our ignorance. Thus, reducing a second-order problem to a first-order problem confounds the two kinds of uncertainty and defeats the purpose of doing the analysis in the first place. As a result, one is faced with either a manageable simulation whose output is of tangential relevance as an objective conclusion of the risk analysis, or a much more cumbersome simulation (whose output, by the way, can itself be difficult to communicate). Although none of these three problems is insurmountable, they do often tend to complicate the practical use of second-order methods in risk analyses.

Let us now consider the intended use of second-order Monte Carlo methods for propagating parametric uncertainty. In any probabilistic risk analysis, the input probability distributions are usually selected by some sort of statistical procedure. Often, a parametric distribution of a specified shape such as the lognormal or normal is fitted to empirical data. Sometimes an input distribution's shape and parameters are simply assumed based on professional judgment. In all but a few extremely well studied cases, the parameters for input distributions are themselves the subject of some uncertainty that arises because of limited sample sizes or because the samples are collected from a statistical population other than the one specifically of interest. But what is the nature of this uncertainty about the parameters? Does the analyst believe that the parameters are in some sense drawn from a statistical distribution? Is there some variability or heterogeneity over time, across space or along some other axis so that the original underlying distribution is not stationary? Or is it really just the case that the analyst is ignorant to some degree about what may be the constant parameters for a stationary distribution? If it is the latter, then we submit that probabilistic methods, including all Monte Carlo methods, will be inappropriate for propagating the uncertainty. Exactly the same argument that applies to ignorance about the underlying values applies to second-order parameters too. In the coin example mentioned above, for instance, the uncertainty about the parameter p may well be ignorance rather than variability.

On top of the ostensible clumsiness of second-order Monte Carlo methods, we then see them to be

incomplete as well. They cannot handle ignorance about (as opposed to variability of) the distribution parameters. What computational method would be better suited to analyze the combination and interaction of the two kinds of uncertainty in a risk analysis? In the next section, we outline our suggestion.

6 PROBABILITY THEORY AND INTERVAL ANALYSIS CAN BE COMBINED

We have argued that it is important to use the methods of interval analysis for propagating ignorance and the methods of probability theory for propagating variability. Most risk analysis problems, however, must deal with both kinds of uncertainty at the same time. Is it possible to handle both variability and ignorance in a single, comprehensive analysis that is faithful to both interval analysis and probability theory? Recent developments in the theory of bounds on probabilities³⁰⁻³⁷ permit such an analysis. In this section we give a simple numerical example that combines variability with ignorance to yield a quantity that is uncertain in both senses.

Figure 3(a) depicts a cumulated normal distribution centred at 5.0 with unit standard deviation. For convenience, the tails have been truncated at the first and ninety-ninth percentiles. We refer to this distribution as normal(5,1). The probability distribution represents a quantity about whose value we are uncertain because of its intrinsic variability. Figure 3(b) depicts an interval between one half and one, symbolized as $[0.5,1]$. This interval represents a quantity about whose value we are uncertain because we have no more precise measurement of it. It may be varying within this range, or it may be a fixed, unvarying value somewhere within the range. We don't have any particular information one way or the other.

The remaining graphs in Fig. 3 depict the product, sum, quotient and difference of the normal(5,1) distribution and the $[0.5,1]$ interval. Let us carefully consider the product depicted in Fig. 3(c). The answer that probability bounds analysis gives us is not a single probability distribution. Rather, it is a region within which the probability distribution of the product must lie. This is to say that, whatever the true value(s) of the uncertain quantity we have represented with the interval, the distribution of the product lies somewhere within the black region of Fig. 3(c). This answer fully expresses the uncertainty induced by the two factors. Any more precise an answer would simply be underestimating the degree of uncertainty present in the calculation. For instance, if we had used a uniform distribution to represent the second factor rather than an interval, and performed the multiplication accord-

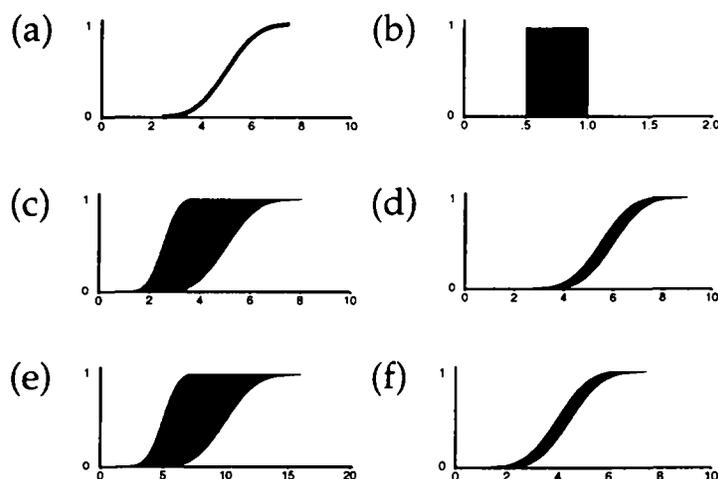


Fig. 3. (a) A cumulative normal probability distribution with mean 5.0 and standard deviation 1.0, but truncated beyond the 1st and 99th percentiles. This distribution is symbolized by the expression $\text{normal}(5,1)$. (b) An interval between one half and one, symbolized by the expression $[0.5,1]$. (c) The product of the distribution and interval, $\text{normal}(5,1) \times [0.5,1]$. The black region envelopes all the cumulative distributions that could arise as this product. (d) The sum $\text{normal}(5,1) + [0.5,1]$. (e) The quotient $\text{normal}(5,1)/[0.5,1]$. (f) The difference $\text{normal}(5,1) - [0.5,1]$.

ing to the rules of probability theory, we would have obtained a particular distribution roughly centred in the black region of Fig. 3(c). But such an answer would, however, have a wholly unjustified precision. In other words, it might be wrong, either under- or overestimating probabilities for the possible range of products. Of course it might be exactly correct by accident, but such an outcome would actually be remarkably unlikely.

The horizontal span of the probability bounds are a function of the variability in the result. The vertical breadth of the bounds is a function of our ignorance. A pure risk analysis problem with perfectly characterized probability distributions as inputs will yield a pure probability distribution as the result. Values, distributions and dependencies that are imperfectly known contribute to a widening of the bounds. The greater the ignorance, the wider the vertical distance between bounds, and the more difficult to make precise probabilistic statements about the expected frequencies of extreme events. But this is what one wants; after all, ignorance should muddle the answer to some extent. Something is obviously amiss information-theoretically if we can combine ignorance and gain more precision than we started with.

The mathematical details of the calculations used to obtain the probability bounds are elementary and have been described with numerical examples elsewhere.³⁷ High-level software⁵⁶ is available for performing arbitrary algebraic calculations with relative ease. For instance, the software allowed us to perform the multiplication simply by typing in the expression `'normal(5,1) * [0.5,1]`. The resulting pair of probability bounds can be assigned to a variable and used in subsequent calculations. Probability bounds that describe the uncertainty about the inputs in a risk

analysis problem can be derived from known constraints or developed using statistical confidence interval procedures.³³⁻³⁷

The practicality of a method for propagating uncertainty depends in part on how computationally expensive it is, although the growing availability of powerful computers is making this concern less important than it used to be. Table 2 gives comparisons for the several methods we have mentioned. If the calculation of the underlying arithmetic expression using scalar floating point numbers requires F time, Monte Carlo simulations require on the order of NF time, where N is the number of replications. The value selected for N usually ranges between 100 and 50,000, depending on the patience and zeal of the analyst. Using more Monte Carlo replicates is generally always preferable to using fewer, but special strategies such as Latin hypercube sampling¹ are used to minimize N . Interval analysis requires on the order of $4F$ time, although this could be reduced to $2F$ time if all the numbers and intermediate results are strictly positive.^{8,13} Second-order Monte Carlo methods require on the order of N^2F time because, as we mentioned above, each replicate in the outer loop implies an entire Monte Carlo simulation in the inner loop. Probability

Table 2

Analysis	Relative execution time
Deterministic point estimate	F
Interval analysis	$4F$
Monte Carlo	NF where $N \in [100,50000]$
Probability bounds	K^2F where $K \in [20,100]$
Second-order Monte Carlo	N^2F where $N \in [100,50000]$

bounds analysis requires on the order of K^2F time, where K is the number of discretization bins employed. The number of bins usually ranges between 20 and 100, depending on the finickiness of the analyst. The example calculations displayed in Fig. 3 were done using $K = 50$. This table shows that probability bounds analysis compares quite favourably with other methods of uncertainty propagation, being on a par with the computational expense of a simple Monte Carlo analysis.

Some theoreticians⁵⁷ have argued that it is reasonable to use Monte Carlo methods to take account of the fact that there is disagreement about the appropriate risk model to use, and even to average the results from different models, weighted by the respective evidence each one has supporting its claim as the truth. Others^{1,58,59} disagree with this approach, suggesting it is nonsensical to average the results of mutually exclusive theories about nature. Their position leaves unanswered the question of how uncertainty about model form ought to be propagated through an analysis and suggests that multiple analyses are required which grow combinatorially in complexity as the number of disputes about model form increases. It seems to us that not knowing which model is correct means we cannot be sure which outcome will follow. This is ignorance, not variability, and should be treated as incertitude in a risk analysis.

Figure 4 illustrates how the predictions of two theories with respect to some variable are combined as the envelope of their cumulative distribution functions. The region depicted in Fig. 4(b) can be used as an input for probability bounds analysis. We see then that an approach based on probability bounds analysis is flexible enough to incorporate uncertainty arising from indecision or controversy about the mathematical form of the risk model, including which functions are used, the values of parameters, the shapes of input distributions, and the correlation and dependency structure among variables. Although the resulting probability bounds are wider than they would be if we could agree on the correct form of the model, they will enclose the true result so long as the correct model is among those considered. By using probability bounds analysis the computational complexity does not grow when there are more questions about model form.

Although the probability bounds approach is flexible, it is not very delicate. By lumping all sources of incertitude together, one cannot later resolve what component of the uncertainty in the output is the result of what source of uncertainty, unless dual with-and-without computations are made. The approach may also be unwieldy in analysing the weight of evidence⁵⁷ for one theory or model choice over another. One might think that a second-order Monte

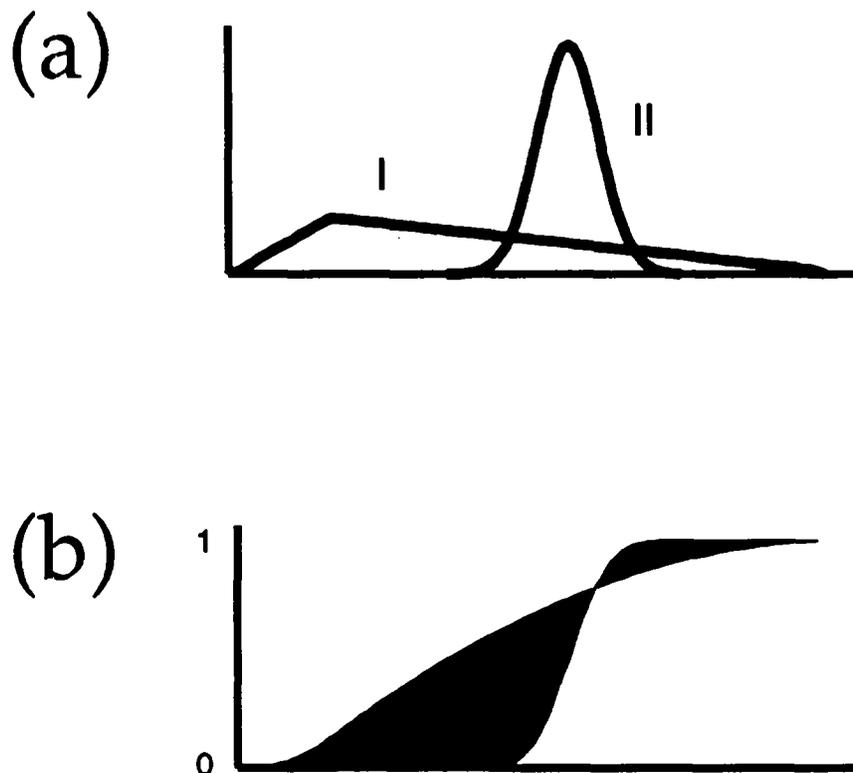


Fig. 4. (a) (Hypothetical) probability distributions for a parameter from competing theories I and II. (b) Probability bounds representing uncertainty in the parameter resulting from controversy about whether theory I or II is correct and the variability each theory predicts for the parameter. The region is the envelope of the respective cumulated forms of the two distributions.

Carlo simulation would be better suited for the task, but Loui³² has argued that a probability bounds approach is not improved on by higher-order Monte Carlo methods in that it answers all the practical questions they can answer. This is not to say that evidence for one theory over another should be ignored. However, to make use of the information in a probability bounds approach would require a series of nested calculations at different levels of confidence to incorporate the evidence comprehensively. A hybrid arithmetic⁶⁰ that generalizes probability bounds analysis in the same way that fuzzy arithmetic⁶¹ generalizes interval analysis would be needed to do this.

7 RISK ANALYSIS IS NOT PROBABILITY THEORY

As we somewhat simplistically depict in Fig. 5, risk analysis is a new field of inquiry that is distinct from probability theory. The two disciplines address different problems. Since the time of Laplace, probability theory has concentrated on making good estimates for single quantities whose values might be varying or might be uncertain. The typical problem in probability theory posits the existence of an answer that is a single real number. Risk analysis, on the other hand, focuses on fundamentally different questions. In risk analysis, we are usually dealing with populations and various potential magnitudes of risk. We are generally interested in the entire distribution rather than a single number. Whereas probability theory uses a distribution to characterize the uncertainty about a single scalar number, risk analysis requires some comparable device to express the uncertainty about a distribution. Second-order Monte Carlo methods seek to provide this needed perspective, as does probability bounds analysis.

We have argued that different computational methods need to be used for variability and ignorance,

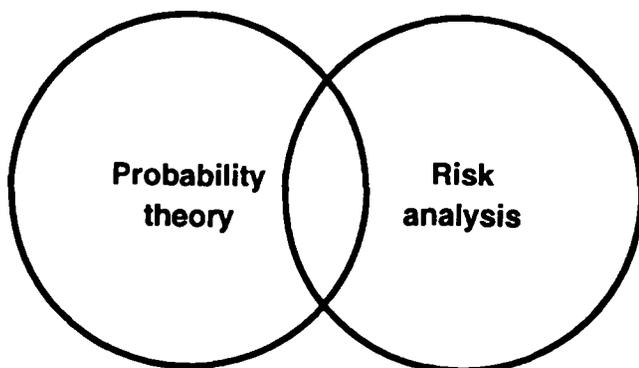


Fig. 5. Qualitative comparison of the subject matter of risk analysis and probability theory.

but we are not the first to suggest this. In decision theory,^{62,63} of which risk analysis is surely a subdiscipline, ignorance and variability have long been recognized as deserving different analytical treatments. Knight⁶⁴ himself differentiated between decisions under risk and decisions under uncertainty which correspond to situations in which probabilities are either well known or completely unknown. It has been argued that maximum entropy criteria can be used to reduce a decision-under-uncertainty problem to a decision-under-risk problem so long as one supposes that the available information is sufficient to uniquely determine the probabilities.²⁸ We argue that such a supposition is wishful thinking that cannot be justified in real-life problems. Instead, one should just concede that empirical knowledge is limited and honestly observe this fact when making calculations. Recent methodological research in probability intervals^{31,32,62} has shown how formal decisions may be made in circumstances in which both ignorance and variability play important roles. This approach has the comprehensiveness and objectivity required of a method used for making decisions affecting the public good.

8 CONCLUSIONS

Using idealized numerical examples we have illustrated an important principle: measurement error should not blithely be treated as though it were random error. Although some of it may well be random, it is illegitimate to assume that it is always random and doing so will lead to unjustifiably precise answers. While empirical effort can in principle reveal the relative magnitudes of the systematic and random components in measurement error, enough information can never be collected to make such a determination perfectly. More generally, uncertainty should not always be treated as though it were variability. Doing so will lead a risk analysis to incorrect results. Although variability in parameters will partially cancel itself out when the parameters are combined in a risk analysis, ignorance need not do so because it can contain systematic error. One cannot assume ignorance about one thing is cancelled out by ignorance about another. It might actually be the case that errors balance one another to some extent, but assuming they always will is merely wishful thinking. The more complex the mathematical expression used in a risk analysis is in terms of the number of parameters or time steps, the more important this argument will be.

Of course we do not suggest that all uncertainty should necessarily be treated as though it were ignorance. Doing so would neglect available information and abdicate the responsibility to obtain an

answer that is as precise as is justified. Analysts in individual circumstances are free to specify what they know and admit what they do not know. This process may, of course, involve assumptions one way or the other such as the analyst may consider reasonable. We expect that conscientious deliberation will in most cases recognize the presence of uncertainty that should be considered ignorance rather than variability, although the relative magnitudes of the two kinds of uncertainty will vary widely among circumstances.

We conclude that not only should ignorance be distinguished from variability in risk analyses, but the two forms of uncertainty should be propagated with different analytical methods. A method like interval analysis is appropriate for propagating ignorance through mathematical calculations. Although this method is general enough to be applied to both ignorance and variability, in the case of the latter it will inefficiently yield results with overestimated (i.e., hyperconservative) dispersions. Likewise, when the dependence between inputs can be specified, interval analysis cannot make use of the information to narrow the results. Probability theory provides the methods appropriate for propagating random variability with known dependencies through calculations. Although appropriate for variability, probability theory in its classical form cannot be used to propagate real ignorance however. Although several theoreticians have asserted there is no distinction to be made between ignorance and probability, we follow the tradition in decision theory which says there is. Because risk analysis is generally concerned with entire distributions of risks and the reliability of these distributions, as a purely practical matter, variability and ignorance should be treated in a way that does not confuse the two forms of uncertainty. Because both are present in almost all practical situations, an approach (such as probability bounds analysis) that is faithful to both interval analysis and probability theory should be used to propagate uncertainty in risk analysis.

REFERENCES

1. Morgan, M.G. & Henrion, M., *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press, Cambridge, England, 1990.
2. Finkel, A.M., *Confronting Uncertainty in Risk Management: A Guide for Decision-Makers*, Centre for Risk Management, Resources for the Future, Washington, 1990.
3. Faber, M., Manstetten, R. & Proops, J., *Toward an open future: ignorance, novelty and evolution. Ecosystem Health: New Goals for Environmental Management*, (eds R. Costanza, B.G. Norton & D.D. Haskell) Island Press, Washington, 1992.
4. Hoffman, F.O. & Hammonds, J.S., Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk Anal.*, **14** (1994) 707-712.
5. Rowe, W.D., Understanding uncertainty. *Risk Anal.*, **14** (1994) 743-750.
6. Casti, J.L., *Searching for Certainty*, William Morrow, New York, 1990.
7. Burgman, M., Ferson, S. & Akçakaya, H.R., *Risk Assessment in Conservation Biology*, Chapman & Hall, London, 1993.
8. Dwyer, P., *Linear Computations*, John Wiley, New York, 1951.
9. Moore, R.E., *Interval analysis*, Prentice-Hall, Englewood Cliffs, New Jersey, 1966.
10. Moore, R.E., *Methods and Applications of Interval Analysis. SIAM Studies on Applied Mathematics*, Vol. 2, Philadelphia, 1979.
11. Kulisch, U.W. & Miranker, W.L., *Computer Arithmetic in Theory and Practice*, Academic Press, New York, 1981.
12. Alefeld, G. & Herzberger, J., *Introduction to Interval Computations*, Academic Press, New York, 1983.
13. Neumaier, A., *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, England, 1990.
14. Neapolitan, R.E., A survey of uncertain and approximate inference. In *Fuzzy Logic for the Management of Uncertainty*, (eds L. Zadeh & J. Kacprzyk), John Wiley and Sons, New York, 1992, pp. 55-82.
15. Cox, R.T., Probability, frequency and reasonable expectation. *Am. J. Physics*, **14** (1946) 1-13.
16. Cheeseman, P., An inquiry into computer understanding (with responses). *Computational Intelligence*, **4** (1988) 0.
17. Lindley, D.V., Scoring rules and the inevitability of probability. *Int. Stat. Rev.*, **50** (1982) 1-26.
18. Whitt, W., Bivariate distributions with given marginals. *Annals of Stat.*, **4** (1976) 1280-1289.
19. Bratley, P., Fox, B.L. & Schrage, L.E., *A Guide to Simulation*, Springer-Verlag, New York, 1983.
20. Ferson, S., Naive Monte Carlo methods yield dangerous underestimates of tail probabilities. In *Proc. High Consequence Operations Safety Symposium*, Sandia National Laboratories, SAND94-2364, (ed. J.A. Cooper), 1994, pp.507-514.
21. Frank, M.J., Nelson, R.B. & Schweizer, B., Best-possible bounds for the distribution of a sum—a problem of Kolmogorov. *Probability Theory and Related Fields*, **74** (1987) 199-211.
22. Williamson, R.C. & Downs, T., Probabilistic arithmetic I: numerical methods for calculating convolutions and dependency bounds. *Int. J. Approximate Reasoning*, **4** (1990) 89-158.
23. Ferson, S. & Long, T.F., Conservative uncertainty propagation in environmental risk assessments. In *Environmental Toxicology and Risk Assessment - Third Volume*, ASTM STP 1218, (eds J.S. Hughes, G.R. Biddinger & E. Mones) American Society for Testing and Materials, Philadelphia, 1995, pp.97-110.
24. Laplace, P.-S. marquis de, *Théorie analytique des probabilités*, J. Gabay, Paris, 1812.
25. Jaynes, E.T., Information theory and statistical mechanics. *Physical Rev.*, **106** (1957) 620-630.
26. Levine, R.D. & Tribus, M., *The Maximum Entropy Formalism*, MIT Press, Cambridge, 1979.
27. Jaynes, E.T., Where do we stand on maximum entropy? In *The Maximum Entropy Formalism*, (eds R.D. Levine & M. Tribus) MIT Press, Cambridge, Massachusetts, 1979.

28. Grandy, W.T. Jr. & Schick, L.H., *Maximum Entropy and Bayesian Methods*, Kluwer Academic Publishers, Dordrecht, 1991.
29. Lee, R.C. & Wright, W.E., Development of human exposure-factor distributions using maximum-entropy inference. *J. Exposure Anal. & Environ. Epidemiology*, **4** (1994) 329–341.
30. Walley, P. & Fine, T.L., Towards a frequentist theory of upper and lower probability. *Annals of Stat.*, **10** (1982) 741–761.
31. Wolfenson, M. & Fine, T.L., Bayes-like decision making with upper and lower probabilities. *J. Am. Stat. Assoc.*, **77** (1982) 80–88.
32. Loui, R.P., Interval-based decisions for reasoning systems. In *Uncertainty in Artificial Intelligence*, (eds L.N. Kanal & J.F. Lemmer) Elsevier Science Publishers, Amsterdam, 1986, pp.459-472.
33. Grosz, B.N., An inequality paradigm for probabilistic reasoning. In *Uncertainty in Artificial Intelligence*, (eds L.N. Kanal & J.F. Lemmer) Elsevier Science Publishers, Amsterdam, 1986, pp.259-275.
34. Pearl, J., On probability intervals. *Int. J. Approximate Reasoning*, **2** (1988) 211–216.
35. van der Gaag, L.C., Computing probability intervals under independency constraints. *Uncertainty in Artificial Intelligence*, **6** (1991) 457–466.
36. Tessem, B., Interval probability propagation. *Int. J. Approx. Reas.*, **7** (1992) 95–120.
37. Ferson, S., Ginzburg, L. & Akçakaya, H.R. Whereof one cannot speak: when input distributions are unknown. *Risk Analysis* (to appear).
38. Hickman, J.W., PRA procedures guide: a guide to the performance of probabilistic risk assessments for nuclear power plants. *NUREG/CR-2300-V1 and -V2*, National Technical Information Service, Washington, 1983.
39. Smith, A.M. & Watson, I.A., Common cause failure—a dilemma in perspective. *Reliab. Engng.*, **1** (1980) 127–142.
40. Calabrese, E.J. & Gilbert, C.E., Lack of total independence of uncertainty factors (UFs): implications for the size of the total uncertainty factor. *Regulatory Toxicology and Pharmacology*, **17** (1993) 44–51.
41. Huff, D., *How to Lie With Statistics*, Norton, New York, 1954.
42. Gull, S.F., Some misconceptions about entropy. In *Maximum Entropy in Action*, (eds B. Buck & V.A. Macauley), Clarendon Press, Oxford, 1991.
43. Beers, Y., *Introduction to the Theory of Error*, Addison-Wesley, New York, 1957.
44. Wilson, E.B. Jr., *An Introduction to Scientific Research*, McGraw-Hill, New York, 1952.
45. Taylor, J.R., *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, University Science Books, Mill Valley, California, 1982.
46. Rabinovich, S., *Measurement Errors: Theory and Practice*, American Institute of Physics, New York, 1993.
47. Shlyakhter, A., Shlyakhter, I., Broido, C. & Wilson, R., Estimating uncertainty in physical measurements, observational and environmental studies: Lessons from trends in nuclear data. In *Proc. Second Int. Symp. Uncertainty Modeling and Analysis*, IEEE Computer Society Press, Los Alamitos, California, 1993, pp.310-317.
48. Bogen, K.T. & Spear, R.C., Integrating uncertainty and interindividual variability in environmental risk assessments. *Risk Anal.*, **7** (1987) 427–436.
49. McKone, T.E., Uncertainty and variability in human exposures to soil contaminants through home-grown food: a Monte Carlo assessment. *Risk Anal.*, **14** (1994) 449–463.
50. Hattis, D. & Burmaster, D.E., Assessment of variability and uncertainty distributions for practical risk analyses. *Risk Anal.*, **14** (1994) 713–730.
51. Carrington, C.D. & Bolger, M., *Risk Assessment for Contaminants in Food*, Centre for Food Safety and Applied Nutrition, US Food and Drug Administration, Washington, 1995.
52. Fisher, R.A., The underworld of probability. *Sankhya*, **18** (1957) 201–210.
53. Good, I.J., *The Estimation of Probabilities*, MIT Press, Cambridge, Massachusetts, 1965.
54. Kyburg, H.E. Jr., Higher order probabilities. In *Uncertainty in Artificial Intelligence 3*, (eds L.N. Kanal, T.S. Levitt & J.F. Lemmer) Elsevier Science Publishers, Amsterdam, 1989, pp.15-22.
55. Cheeseman, P., In defense of probability. In *Proc. Ninth Int. Joint Conf. Artificial Intelligence*, Los Angeles, 1985, pp.1002-1009.
56. Kuhn, R. & Ferson, S., *Risk Calc*, Applied Biomathematics, Setauket, New York, 1994.
57. Holland, C.D. & Sielken, R.L. Jr., *Quantitative Cancer Modeling and Risk Assessment*, Prentice Hall, Englewood Cliffs, New Jersey, 1994.
58. Committee on Risk Assessment of Hazardous Air Pollutants, Science and Judgment in Risk Assessment. National Research Council, Washington, 1994.
59. Finkel, A.M., A second opinion on an environmental misdiagnosis: The risky prescriptions of Breaking the Vicious Circle. *New York Univ. Env. Law J.*, **3** (1995) 295–381.
60. Ferson, S. & Ginzburg, L., Hybrid arithmetic. In *Proc. 1995 Joint ISUMA Third Int. Symp. Uncertainty Modeling and Anal. and NAFIPS Ann. Conf. North Am. Fuzzy Information Processing Soc.*, IEEE Computer Society Press, Los Alamitos, California, 1995.
61. Kaufmann, A. & Gupta, M.M., *Introduction to Fuzzy Arithmetic: Theory and Applications*, Van Nostrand Reinhold, New York, 1985.
62. Kmietowicz, Z.W. & Pearman, A.D., *Decision Theory and Incomplete Knowledge*, Gower Publishing Company, Hampshire, England, 1981.
63. Finkel, A.M., Stepping out of your own shadow: a didactic example of how facing uncertainty can improve decision-making. *Risk Anal.*, **14** (1994) 751–761.
64. Knight, F.H., *Risk, Uncertainty and Profit*, Houghton Mifflin, Boston, 1921.