Theory and Decision (2005) 58: 209–238
DOI 10.1007/s11238-005-2458-y

KATHLEEN M. WHITCOMB

QUASI-BAYESIAN ANALYSIS USING IMPRECISE PROBABILITY ASSESSMENTS AND THE GENERALIZED BAYES' RULE

ABSTRACT. The generalized Bayes' rule (GBR) can be used to conduct 'quasi-Bayesian' analyses when prior beliefs are represented by imprecise probability models. We describe a procedure for deriving coherent imprecise probability models when the event space consists of a finite set of mutually exclusive and exhaustive events. The procedure is based on Walley's theory of upper and lower prevision and employs simple linear programming models. We then describe how these models can be updated using Cozman's linear programming formulation of the GBR. Examples are provided to demonstrate how the GBR can be applied in practice. These examples also illustrate the effects of prior imprecision and prior-data conflict on the precision of the posterior probability distribution.

KEY WORDS: imprecise probability, generalized Bayes' rule, second-order probability, quasi-Bayesian analysis.

1. INTRODUCTION

A requirement of any normative Bayesian procedure is that the assessed prior probability distribution is precise, or additive. This constraint can be problematic if an expert assessor does not believe that he has enough information to justify precise probability assessments or if multiple expert assessors are used. In either case it may be preferable to employ imprecise probability expressions such as interval-valued probabilities, second-order probability distributions, or belief functions (Beyth-Marom, 1982). Imprecise probability expressions reflect both first-order uncertainty—uncertainty about the occurrence or non-occurrence of events, and second-order uncertainty—uncertainty about the event probabilities themselves (Goldsmith and Sahlin, 1982).

Although normative Bayesian theory maintains that second-order uncertainty is irrelevant to decision making processes (cf. Savage, 1954), numerous experimental studies in behavioral decision theory have demonstrated that second-order uncertainty can affect a decision maker's preferences for decision alternatives (Becker and Brownson, 1964; Boiney, 1993; Curley and Yates, 1985; Einhorn and Hogarth, 1985; Ellsberg, 1961; Goldsmith and Sahlin, 1982; Yates and Zukowski, 1976). Repressing second-order uncertainty can give decision makers a false sense of control in a probabilistic sense, which can translate to overconfidence in decisions (Feagans and Biller, 1981).

While imprecise probability expressions better inform decision makers and may ultimately improve decision quality, they are incompatible with standard decision technologies. Imprecise probabilities require axiom systems more complex than for precise probability. The increased computational and cognitive demands required to implement these axiom systems have no doubt discouraged decision analysts from incorporating imprecise probability expressions into decision making processes.

In this paper, we describe and demonstrate a decision analytic procedure for incorporating second-order uncertainty into a Bayesian, or quasi-Bayesian, analysis. The procedure is based on Walley's theory of upper and lower prevision,[1] which provides a general framework for reasoning with imprecise probabilities (Walley, 1991).

The imprecise probability models examined in this study are special cases covered by the general theory of upper and lower prevision. In particular, we consider cases for which the event space consists of a finite set of mutually exclusive and exhaustive events. We demonstrate how Walley's closely related principles of avoiding sure loss, coherence, and natural extension are applied to assess and to derive coherent (internally consistent) imprecise probabilities for the special models considered in this paper.

The GBR, an inference rule derived from the principles of coherence and natural extension, is used to update coherent imprecise probabilities using a linear programming formulation of the GBR detailed in Cozman (1999). The procedure described in this paper provides decision analysts with a pragmatic approach for incorporating second-order uncertainty into an analysis. In fact, the linear programming models used to derive the imprecise prior probability models and to operationalize the GBR can be implemented using the optimization facilities available with common spreadsheet packages.

The remainder of this paper is organized as follows: Section 2 describes a linear programming model for deriving coherent imprecise probability models from two forms of imprecise probability expression—upper and lower probabilities and imprecise probability ratios. Section 3 discusses the generalized Bayes' rule and delineates a procedure for implementing the GBR using a linear programming model developed by White (1986), Snow (1991), and Cozman (1999). Section 4 describes potential applications of the GBR in reliability analysis, environmental impact assessment, and evaluation of medical trials. Section 5 discusses some issues that should be addressed before imprecise probability models are widely used by decision analysts.

2. CONSTRUCTING IMPRECISE PRIOR PROBABILITY MODELS

Alternatives to additive (precise) probability theory often employ probability intervals or convex sets of probability distributions to represent uncertainty (Fine, 1988; Gardenfors and Sahlin, 1982; Koopman, 1940; Levi, 1980; Smith, 1961). Interval probabilities and convex sets of probability distributions are closely related representations of uncertainty. A brief and non-technical discussion of these probability models and their interrelationship can be found in Cozman (1997). Walley's framework for imprecise probability requires that uncertainty be represented by a convex set of probability distributions. One way to define this convex set is to provide a set of linear constraints on event probabilities. Walley provides a general methodology for eliciting such constraints (Walley, 1991, chap. 4). A key feature of Walley's elicitation methodology is that an assessor need only provide imprecise probability assessments that collectively meet the relatively minimal requirement of "avoiding sure loss" (ASL). Coherent imprecise probability assessments can then be derived using "natural extension". The three interrelated principles—ASL, coherence, and natural extension—are briefly described below. A full description and justification for these principles is provided in Walley (1991).

Analogous to precise probability assessment, an assessor's imprecise probabilities avoid sure loss if they do not imply that he can be induced to accept a series of bets certain to result in a net loss in utility. It is relatively easy for an assessor to provide assessments that avoid sure loss. However, compliance with the ASL principle does not guarantee that the collection of imprecise probability assessments are mutually consistent, or coherent. Coherence is violated when probability assessments for some events imply a probability for some other event which differs from its assessed probability. Such inconsistency usually occurs, not because of any fundamental irrationality on the assessor's part, but because the assessor is not fully cognizant of the implications of his adjudged probabilities.

Coherence violations can be corrected using the principle of natural extension, assuming that the collection of probability assessments avoid sure loss. The principle of natural extension is fundamental to Walley's theory and is informally explained as follows: given a number of bounds on probabilities, there are any number of sets of probability measures that could attain these bounds, any such set is an "extension"; the "natural extension" is always the largest set of probability measures that satisfies the given probability constraints and is coherent. Any other such set incorporates additional information not implied by the initial constraints.

A significant advantage of Walley's approach is that the assessor is relieved of the cognitive burden of conforming to the complex coherency requirements for imprecise probability.

Empirical studies of imprecise probability assessment have demonstrated that individuals, while able to provide imprecise probability assessments that avoid sure loss, are often unable to provide assessments that satisfy coherency constraints (Benson and Whitcomb, 1993; Walley, 1991). In this section, we describe and demonstrate procedures for constructing imprecise probability models following Walley's approach.

Two modes of imprecise probability expression were adopted: (1) interval-valued event probabilities, and; (2) imprecise probability ratios. These expressions were chosen for two reasons. First, neither expression requires the assessor to have a sophisticated knowledge of probability theory in order to conform to the constraints for avoiding sure loss. Second, both form linear constraints on event probabilities and therefore are compatible with the linear programming formulations used to apply the principle of natural extension and to operationalize the GBR.

Section 2.1 describes the procedure for deriving coherent imprecise probability models from lower and upper event probabilities. The procedure for deriving imprecise probability models from imprecise probability ratios is described in Section 2.2.

2.1. *Imprecise probability models derived from interval-valued probabilities*

Consider an event space consisting of n states of nature, $E_1,\ldots,E_n$. Lower and upper probabilities for the event space, $\underline{p}(E_j)$ and $\bar{p}(E_j)$, respectively, avoid sure loss assuming that the following constraints are satisfied (Walley, 1991, p. 198):

$$0 \leqslant \underline{p}(E_j) \leqslant \bar{p}(E_j) \leqslant 1 \quad \text{for } j = 1, \ldots, n$$

and

$$\sum_{j=1}^{n} \underline{p}(E_j) \leqslant 1 \leqslant \sum_{j=1}^{n} \bar{p}(E_j).$$

Previous studies in imprecise probability assessment revealed that assessors were almost always able to provide assessments that avoided sure loss. However, these same studies found that assessors were not always able to provide lower and upper probabilities that conformed to the following constraint for coherency (Walley, 1991, p. 198; Whitcomb, 1989, (unpublished)).

$$\bar{p}(E_j) + \sum_{i \neq j} \underline{p}(E_i) \leqslant 1 \leqslant \underline{p}(E_j) + \sum_{i \neq j} \bar{p}(E_i) \quad \text{for } j = 1, \ldots, n.$$

In order to obtain coherent lower and upper probabilities, the principle of natural extension can be applied to imprecise probability assessments that avoid sure loss. The following simple linear programming model can be used to implement the principle of natural extension:

$$\begin{aligned} \text{For } k = 1, \ldots, n \quad & \min/\max p(E_k), \\ \text{s.t.} \quad & \mathrm{L}(E_j) \leqslant p(E_j) \leqslant \mathrm{U}(E_j), \\ & p(E_j) \geq 0 \quad \text{for } j = 1, \ldots, n \text{ and} \\ & \sum_j p(E_j) = 1.0, \end{aligned}$$

where
$\mathrm{L}(E_j)$ is the lower probability assessment for E_j,
$\mathrm{U}(E_j)$ is the upper probability assessment for E_j.
As a simple example, suppose that an event space is comprised of just three mutually exclusive events E_1, E_2, and E_3. Further suppose that an assessor provides the following lower and upper probability bounds:

$$0.3 \leqslant p(E_1) \leqslant 0.5,$$
$$0.1 \leqslant p(E_2) \leqslant 0.4,$$
$$0.4 \leqslant p(E_3) \leqslant 0.7.$$

It is easy to verify that these probability assessments avoid sure loss. However, since $\bar{p}(E_2) + \sum_{j \neq 2} \underline{p}(E_j) = 1.1$ and $\bar{p}(E_3) + \sum_{j \neq 3} \underline{p}(E_j) = 1.1$, there are two coherency violations.

For example, the second violation indicates that $p(E_3)$ cannot be as high as 0.7 since $p(\text{not-}E_3)$ is constrained to be greater than or equal to 0.4—the constraints on event probabilities are inconsistent. The principle of natural extension can be implemented to derive coherent bounds on the three event probabilities by using the assessor's lower and upper bounds on

$p(E_j)$ for the $L(E_j)$ and $U(E_j)$ in the above linear programming model. The following corrected (coherent) lower and upper probability bounds were computed:

$$p(E_1) = [0.3, 0.5]; p(E_2) = [0.1, 0.3], \text{ and; } p(E_3) = [0.4, 0.6].$$

2.2. *Imprecise probability models derived from imprecise probability ratios*

Probability judgments of the form E_j is at least l and no more than u times as probable at E_i can be formulated as an imprecise probability ratio as follows:

$$l \leqslant \frac{p(E_j)}{p(E_i)} \leqslant u.$$

In order to facilitate the assessment process, Walley suggests using one of the events, $E_i = E_0$, as a reference event (Walley, 1991, p.199). Then the assessments avoid sure loss provided that $l_j \leqslant u_j$ for all j.

The expressions, $l_j \leqslant \big(p(E_j)/p(E_0)\big) \leqslant u_j$, are incorporated as linear constraints on the $p(E_j)$ in a linear program as follows:

$$\begin{aligned} \text{For } k = 1, \ldots, n \quad & \min/\max p(E_k) \\ \text{s.t.} \quad & l_j p(E_0) \leqslant p(E_j) \leqslant u_j p(E_0) \\ & p(E_j) \geqslant 0 \quad \text{for } j = 1, \ldots, n \text{ and} \\ & \sum_j p(E_j) = 1.0. \end{aligned}$$

To demonstrate, suppose again that an event space is comprised of just three outcomes E_1, E_2, and E_3 and that the following assessments have been made with respect the reference event, $E_1 = E_0$:

$$0.5\ p(E_1) \leqslant p(E_2) \leqslant 2p(E_1),$$
$$2\ p(E_1) \leqslant p(E_3) \leqslant 3p(E_1).$$

The corresponding coherent lower and upper event probabilities derived from the linear programming model are: 0.1667 $\leqslant p(E_1) \leqslant 0.2857$; $0.1111 \leqslant p(E_2) \leqslant 0.4$ and, $0.4 \leqslant p(E_3) \leqslant 0.6667$.

It is not necessary to restrict the assessment procedure to a single mode of imprecise probability expression. In fact, upper and lower probabilities are often inadequate models of uncertainty and information may be lost when used alone to model uncertainty. This is especially true when upper and lower probabilities are used to define a set of prior probability distributions in a quasi-Bayesian analysis. But it can be difficult to verify that a set of assessments comprised of different types of probability expressions avoids sure loss without using computer software written specifically for this purpose.

3. LINEAR PROGRAMMING FORMULATION FOR THE DISCRETIZED GBR

The generalized Bayes' rule is used to obtain minimum and maximum values for linear functions of posterior event probabilities. When the outcome space is finite and the prior is expressed in terms of linear constraints on event probabilities, linear programming methods can be used to operationalize the GBR. In this paper, a modified version of the White–Snow linear programming solution for the GBR was adopted. This model is described and demonstrated in this section. Additional background and details on the development of this model can be found in White (1986), Snow (1991), and Cozman (1999).

Suppose that prior uncertainty regarding events $\mathbf{E} = E_1, \ldots, E_n$ is represented by a convex set of precise probability distributions formed by a set of linear constraints on event probabilities and that observational data, x, is obtained that generates a precise likelihood function $\mathbf{C} = c(x|\mathbf{E})$ for which $c(x|E_i) > 0$ for $i = 1, \ldots, n$. The linear programming formulation of the GBR is stated initially as

$$\text{For } k = 1, \ldots, n \quad \min/\max q(E_k|x) = \frac{p(E_k) * c(x|E_k)}{\sum_{j=1}^{n} p(E_j) * c(x|E_j)},$$

$$\begin{aligned} &\text{s.t.} \quad \mathbf{Ap} \leqslant 0, \\ &p(E_j) \geqslant 0 \quad \text{for } j = 1, \ldots, n \text{ and} \\ &\sum p(E_j) = 1.0, \end{aligned}$$

where

$q(E_j|x)$ is the posterior probability for event j given observation x;

$p(E_j)$ is the prior probability for event j;

$c(x|E_j)$ is the probability of observation x given event j;

A is an $m \times n$ constraint matrix whose coefficients are denoted a_{ij}, and;

p is an n-dimensional vector whose elements are $p(E_j)$ for $j = 1, \ldots, n$.

Bounding **Ap** by the zero vector is necessary for the methods used to solve the GBR, and results in no loss of generality (Snow, 1991). To clarify this point, note that $\sum p(E_j) = 1.0$ and let r_i denote a positive constant corresponding to the right hand side of constraint "i". Then the constraint

$$\sum_{j=1}^{n} a_{ij} p(E_j) \leqslant r_i$$

can be rewritten as

$$\sum_{j=1}^{n} \left[a_{ij} - r_i\right] p(E_j) \leqslant 0.$$

And a constraint of the form

$$\sum_{j=1}^{n} a_{ij} p(E_j) \geqslant r_i$$

can be rewritten as

$$-\sum_{j=1}^{n} \left[a_{ij} - r_i\right] p(E_j) \leqslant 0.$$

Equality constraints can be similarly expressed by a pair of zero-bounded inequalities.

In order to linearize the objective function, a change of variables must be performed. Cozman (1999) suggested the following change of variables:

$$\gamma(E_k) = \frac{p(E_k)}{\sum_{j=1}^{n} c(x|E_j) * p(E_j)}.$$

Incorporating this transformation into the previous formulation results in the following linear programming model:

$$\begin{aligned} \text{For } k = 1, \ldots, n \quad & \min/\max\, q(E_k|x) = c(x|E_k) * \gamma(E_k) \qquad (3.1)\\ \text{s.t.} \quad & \mathbf{A}\boldsymbol{\gamma} \leqslant 0,\\ & \gamma(E_j) \geqslant 0,\\ & \sum_{j=1}^{n} c(x|E_j) * \gamma(E_j) = 1.0 \quad \text{for } j = 1, \ldots, n. \end{aligned}$$

The generalized Bayes' rule follows from the principles of natural extension and coherence. The $\underline{q}(E_k|x)$ and $\overline{q}(E_k|x)$ obtained by minimizing and maximizing Equation (3.1) for each event, E_k, subject to linear constraints on the prior and coherency constraints are the *minimal* coherent lower and upper posterior probabilities. Other coherent lower and upper probabilities, $\underline{q}'(E_k|x)$ and $\overline{q}'(E_k|x)$, may exist that satisfy these constraints. But, any such $\underline{q}'(E_k|x)$ must dominate $\underline{q}(E_k|x)$ in that $\underline{q}'(E_k|x) \geqslant \underline{q}(E_j|x)$ for all E_k and $\overline{q}'(E_k|x)$ must dominate $\overline{q}(E_k|x)$ in that $\overline{q}'(E_k|x) \leqslant \overline{q}(E_k|x)$ for all E_k.

Accordingly, $\underline{q}'(E_k|x)$ and $\overline{q}'(E_k|x)$ incorporate additional information not implied by the initial constraints and thus are not the *minimal* coherent solution (Walley, 1996, p. 16).

To demonstrate the LP formulation of the GBR, we return to the example presented in Section 2.1. The prior constraints are: $0.3 \leqslant p(E_1) \leqslant 0.5; 0.1 \leqslant p(E_2) \leqslant 0.4$, and, $0.4 \leqslant p(E_3) \leqslant 0.7$. Suppose that relevant data is observed and that the normalized likelihoods are: $c(x|E_1) = 0.1, c(x|E_2) = 0.7$, and $c(x|E_3) = 0.2$. The linear program is then

$$\begin{aligned} \text{For } k = 1, 2, 3 \quad & \min/\max\, q(E_k|x) = c(x|E_k) * \gamma(E_k),\\ \text{s.t.} \quad & -0.7\gamma(E_1) \; +0.3\gamma(E_2) \; +0.3\gamma(E_3) \; \leqslant 0,\\ & +0.5\gamma(E_1) \; -0.5\gamma(E_2) \; -0.5\gamma(E_3) \; \leqslant 0,\\ & +0.1\gamma(E_1) \; -0.9\gamma(E_2) \; +0.1\gamma(E_3) \; \leqslant 0,\\ & -0.4\gamma(E_1) \; +0.6\gamma(E_2) \; -0.4\gamma(E_3) \; \leqslant 0,\\ & -0.7\gamma(E_1) \; -0.7\gamma(E_2) \; +0.3\gamma(E_3) \; \leqslant 0,\\ & +0.4\gamma(E_1) \; +0.4\gamma(E_2) \; -0.6\gamma(E_3) \; \leqslant 0,\\ & +0.1\gamma(E_1) \; +0.7\gamma(E_2) \; +0.2\gamma(E_3) \; = 1\\ & \text{and} \quad \gamma(E_1), \gamma(E_2), \gamma(E_3) \geqslant 0. \end{aligned}$$

Solution of the linear program yields the following lower and upper posterior event probabilities: $q(E_1|x) = [0.094, 0.250]$, $q(E_2|x) = [0.318, 0.656]$, and $q(E_3|x) = [0.250, 0.545]$.
If the event-space is comprised of ordered events, the analyst may want to calculate upper and lower bounds on cumulative event probabilities. This can easily be accomplished by modifying the objective function to

$$\text{For } k = 1, 2, 3 \quad \min/\max \sum_{k=1}^{k} q(E_k|x) = \sum_{k=1}^{k} c(x|E_k) * \gamma(E_k).$$

For the current example, the upper and lower bounds on cumulative event probabilities are calculated to be: $q(E_1|x) = [0.094, 0.250]$; $q(E_1 \cup E_2|x) = [0.455, 0.750]$; and $q(E_1 \cup E_2 \cup E_3|x) = [1.00, 1.00]$.
The objective function can be expressed more generally in terms of upper and lower expectations, also called upper and lower previsions. For a bounded function, $f(\mathbf{E})$, the objective function can be expressed as

$$\min/\max \sum_{k=1}^{n} f(E_k) q(E_k|x) = \sum_{k=1}^{n} f(E_k) c(x|E_k) * \gamma(E_k).$$

By appropriate selection of $f(E_k)$, any linear function of the posterior event probabilities can be optimized. Such functions include upper and lower event probabilities, cumulative event probabilities, and expected losses. To illustrate computation of upper and lower expected losses, suppose that for the current example losses associated with the occurrence E_1, E_2, and E_3 are 500, 100, and 200, respectively. The objective function is

$$\min/\max 500 \cdot 0.1 \cdot \gamma(E_1) + 100 \cdot 0.7 \cdot \gamma(E_2) + 200 \cdot 0.2 \cdot \gamma(E_3).$$

Lower and upper expected losses generated by minimizing and maximizing this function subject to the previously specified constraints are 162.50 and 240.00, respectively.

The approach for solving the GBR described in this section is quite inclusive. Although the formulation is only appropriate when the outcome space is discrete, any outcome space, if not inherently discrete, can be discretized. The LP formulation for solving the GBR can be extended to include cases where the likelihoods are interval-valued, provided that the lower bounds of the likelihoods are non-zero (Cozman, 1999). Algorithms that accomodate zero-valued likelihoods are available, but computationally more complex. These algorithms employ either sequences of linear programs (Coletti, 1994; Coletti and Scozzafava, 1999; Walley et al., 1999) or a sequence of pivoting operations (Cozman, 2002) to solve the GBR.

4. ILLUSTRATIVE EXAMPLES OF THE GENERALIZED BAYES' RULE

In this section we provide examples to illustrate how the generalized Bayes' rule could be used in practical problems. The first case is set in reliability engineering and assumes that a single expert assessor provides probability constraints that can be used to derive an imprecise prior. The second and third examples are derived from an ecological impact study and a medical trials study. In both cases, the original analysis used multiple precise priors to capture the divergent beliefs of experts. We demonstrate how a quasi- Bayesian analysis of the data could have been conducted instead.

4.1. *Application to Bayesian reliability analysis*

Reliability analysis is concerned with the ability of a system to successfully perform its intended function for a given time under a specific set of conditions (Martz and Waller, 1982). Often, data for estimating reliability parameters is limited and Bayesian methods are used in order to augment the data with expert opinion in the form of a subjective probability distribution for the parameter. Previous studies have applied Walley's theory of imprecise probability to reliability engineering. Coolen (1993) and Coolen and Newby (1994) demonstrated the use of lower and upper conjugate density families to model prior beliefs about reliability parameters. Utkin and Kozine (2001) and Utkin and Gurov (2002) encode partial information about reliability parameters using linear constraints. Revised previsions for the parameter are computed using a linear programming formulation of natural extension. Their results can used for discrete and continuous lifetime distributions. Kozine and Krymsky (2003) describe a technique for obtaining more precise previsions for continuous reliability parameters by eliciting judgments on a function that dominates the probability density function of the parameter.

We present a basic example in which an expert provides imprecise probability ratios to characterize his beliefs about the mean time to failure, θ, for a machine component. Hypothetical sample results are used to update the prior using the LP formulation of the GBR described in Section 3. The data sets were chosen to reflect varying weights of evidence and degrees of prior-data conflict.

For simplicity, we assume that the mean time to failure (MTTF) of the component takes on the following values: $\theta = \{0.5, 1.0, 2.0, 3.0, 4.0, 5.0, 10.0\}$ as measured in thousands of hours.[2] Let θ_4 be the designated reference event and suppose that the engineer assesses the imprecise probability ratios as

TABLE I
Upper and lower probability, ratios and event probabilities, for prior probability model

θ_j	l_j	u_j	$\underline{p}(\theta_j)$	$\overline{p}(\theta_j)$
$\theta_1 = 0.5$	1/100	1/10	0.0025	0.0469
$\theta_2 = 1.0$	1/3	5/4	0.1030	0.4222
$\theta_3 = 2.0$	2/5	1	0.2410	0.4895
$\theta_4 = 3.0$	1	1	0.1127	0.3783
$\theta_5 = 4.0$	1/10	2/5	0.0260	0.1701
$\theta_6 = 5.0$	1/10	1/5	0.0247	0.0933
$\theta_7 = 10.0$	1/10	1/5	0.0247	0.0933
				$\Delta = 1.1590$

shown in columns (2) and (3) of Table I. Coherent lower and upper probabilities derived from these assessments using the principle of natural extension are shown in columns (4) and (5). A summary measure of imprecision, $\Delta = \sum_{j=1}^{n} \overline{p}(\theta_j) - \sum_{j=1}^{n} \underline{p}(\theta_j)$, was calculated and appears beneath columns (4) and (5) in Table I (Good, 1965).

Suppose that a life testing experiment is performed in which a certain number of new components are tested until they all fail. This is referred to as "Type II/item-censored testing without replacement" (Martz and Waller, 1982, pp. 118–121). In this situation it is common to use an exponential probability distribution to represent the distribution of failure times. The likelihood function for exponentially distributed failure times is

$$L(\theta|s, t) = \frac{1}{\theta_j^s} e^{-t/\theta},$$

where

s is a fixed number of failures; and

t is a random variable denoting the sum of the failure times for *s* units, measured in thousands of hours.

Normalized likelihoods were calculated for six different combinations of *s* and *t*. Three combinations of *s* and *t* were chosen to represent increasing amounts of data in general agreement with the prior: $(s, t) = (2, 2.4)$; $(6, 7.2)$; and $(10, 12.0)$. Three combinations were chosen to represent increasing amounts of data in conflict with the prior: $(s, t) = (2, 0.25)$; $(6, 0.75)$; and $(10, 1.2)$.

The LP formulation of the GBR was then used to update the engineer's prior beliefs using the hypothetical life testing results. The constraint set for the formulation consists of the probability ratio bounds shown in columns (2) and (3) of Table I and the requisite non-negativity and additivity constraints. Each of the six sets of normalized likelihoods, $f(s, t|\theta_j)$, was used in turn for the $c(x|\theta_j)$.

Table II contains the normalized likelihoods and lower and upper posterior probability bounds for the life testing results in general agreement with the prior. The summary measure of imprecision, Δ, declines as the amount of evidence increases.

TABLE II
Reliability example: failure-time data in agreement with prior*

θ_j	$f(2,2.4\|\theta_j)$	$\underline{p}(\theta_j\|2,2.4)$	$\overline{p}(\theta_j\|2,2.4)$	$f(6,7.2\ \|\theta_j)$	$\underline{p}(\theta_j\|6,7.2)$	$\overline{p}(\theta_j\|6,7.2)$	$f(10,12\|\theta_j)$	$\underline{p}(\theta_j\|10,12)$	$\overline{p}(\theta_j\|10,12)$
0.5	0.1042	0.0013	0.0243	0.0257	0.0002	0.0049	0.0043	0.0000	0.0008
1.0	0.2873	0.1689	0.5258	0.5373	0.3024	0.6591	0.6849	0.4263	0.7507
2.0	0.2385	0.2872	0.5682	0.3072	0.2832	0.5836	0.2698	0.2320	0.5263
3.0	0.1581	0.0860	0.3073	0.0896	0.0347	0.1544	0.0346	0.0121	0.0648
4.0	0.1086	0.0136	0.0961	0.0290	0.0027	0.0217	0.0053	0.0005	0.0041
5.0	0.0784	0.0095	0.0367	0.0109	0.0010	0.0041	0.0010	0.0001	0.0004
10.0	0.0249	0.0030	0.0118	0.0004	0.0001	0.0000	0.0001	0.0000	0.0000
		Δ =	1.00073		Δ =	0.7941		Δ =	0.6761

*Values less than 0.00005 are rounded to zero.

TABLE III
Reliability example: failure-time data in conflict with prior*

θ_j	$f(2,.25\|\theta_j)$	$\underline{p}(\theta_j\|2,.25)$	$\overline{p}(\theta_j\|2,.25)$	$f(6,.75\|\theta_j)$	$\underline{p}(\theta_j\|6,.75)$	$\overline{p}(\theta_j\|6,.75)$	$f(10,1.25\|\theta_j)$	$\underline{p}(\theta_j\|10,1.25)$	$\overline{p}(\theta_j\|10,1.25)$
0.5	0.6676	0.0179	0.3133	0.9672	0.1916	0.8943	0.9966	0.7009	0.9887
1.0	0.2143	0.3025	0.7666	0.0320	0.0986	0.7931	0.0034	0.0112	0.2986
2.0	0.0607	0.1403	0.3968	0.0007	0.0053	0.0345	0.0000	0.0000	0.0006
3.0	0.0281	0.0271	0.1656	0.0001	0.0002	0.0034	0.0000	0.0000	0.0000
4.0	0.0162	0.0038	0.0409	0.0000	0.0000	0.0003	0.0000	0.0000	0.0000
5.0	0.0105	0.0024	0.0136	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
10.0	0.0027	0.0006	0.0035	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
			Δ = 1.2057			Δ = 1.4299			Δ = 0.5758

*Values less than 0.00005 are rounded to zero.

Table III contains the normalized likelihoods and lower and upper posterior probability bounds corresponding to the data that contradict the prior. For these cases, Δ actually increases until the weight of evidence overwhelms both first- and second-order uncertainty. Greater imprecision in the posterior versus prior probability distribution has been termed *dilation* (Siedenfeld and Wasserman, 1993). It is a significant and not uncommon drawback associated with the GBR. In practice, dilation often indicates that the prior constraints should be revisited and, if possible, assessed more precisely and/or augmented with additional constraints (Fortin et al., 2001).

In the current case, dilation is partially attributable to prior/data conflict. Figure 1, which contains graphs for the imprecise prior and the posterior distribution for $s = 6$ and $t = 750$ h,

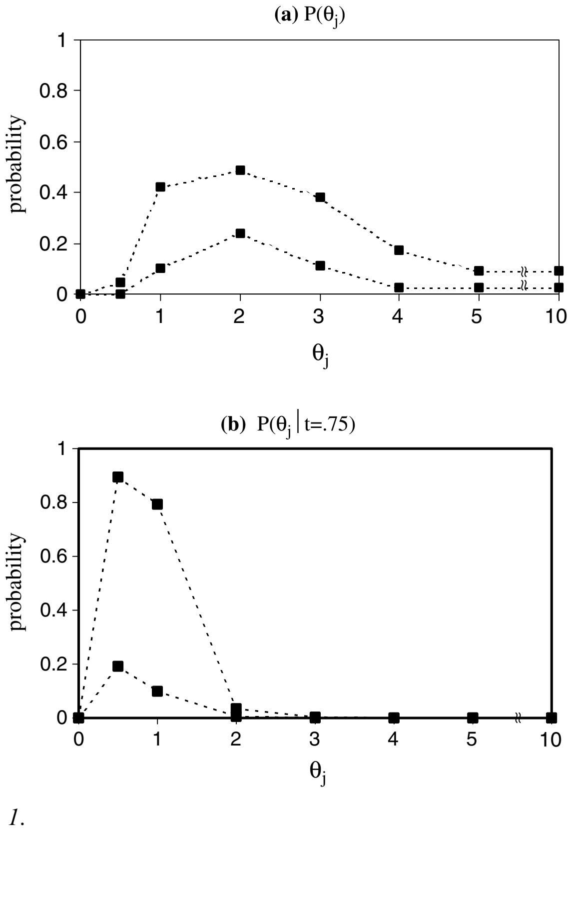

Figure 1.

illustrates this point. The likelihood of the data is allocated almost entirely to $\theta_1 = 0.5$, a parameter value assigned a low (first-order) prior probability. The remaining likelihood is allocated to $\theta_2 = 1.0$, which has a relatively high (first-order) prior probability. The imprecise posterior probability distribution reveals that the only two plausible events after the data are observed are $\theta_1 = 0.5$ and $\theta_2 = 1.0$ but that there is a great uncertainty as to how first-order uncertainty should be divided among them.

4.2. *Analyzing the effects of rain forest logging*

In environmental and ecological studies, statistical analyses are often used to make and justify controversial decisions (Wolfson et al., 1996). The example described in this section was derived from a study undertaken to assess the impact of rain forest logging on indigenous birds and small mammals in tropical Queensland, Australia (Crome et al., 1996). As in other areas involving rare, complex, or poorly understood processes, prior information about the effects of rain forest canopy reduction on indigenous species was based mainly on the opinions of various experts.

Experts were chosen to represent three groups of stakeholders: proponents of the logging industry; full-time conservationists, and; interested members of the public. The divergent beliefs of the stakeholders were reflected by constructing three separate priors—one each for the most extreme (optimistic) logger, the most extreme (pessimistic) conservationist, and the average (impartial) layperson. Each prior was derived by eliciting fractiles for the effect size, δ (the ratio of post- to pre-logging counts of a species in the logged area), and fitting a log-normal mixture distribution to the fractiles.

In our approach, an imprecise prior models the initial beliefs of the three groups of stakeholders. To construct an imprecise prior using the methods described in this paper, each stakeholder group's prior must be discretized. Ten values of δ were used and are listed in Table IV, column (1). The $p(\delta_i)$ for each precise prior were estimated from its fitted lognormal mixture

TABLE IV
Rain forest logging example*

δ_j	$\underline{p}(\delta_j)$	$\overline{p}(\delta_j)$	$f(x\|\delta_j)$	$\underline{p}(\delta_j\|x)$	$\overline{p}(\delta_j\|x)$
0.2	0.0000	0.1020	0.0000	0.0000	0.0000
0.4	0.0606	0.5102	0.0253	0.0069	0.1155
0.6	0.1212	0.2206	0.4199	0.2778	0.6296
0.8	0.1020	0.2353	0.4010	0.2394	0.5994
1.0	0.0510	0.2121	0.1265	0.0315	0.2170
1.2	0.0408	0.1667	0.0236	0.0043	0.0367
1.4	0.0204	0.1212	0.0034	0.0003	0.0038
1.6	0.0102	0.0758	0.0004	0.0000	0.0003
1.8	0.0051	0.0303	0.0000	0.0000	0.0000
2.0	0.0051	0.0303	0.0000	0.0000	0.0000
	Δ =	1.2881		Δ =	1.0421

*Values less than 0.00005 are rounded to zero.

distribution. For each δ_j, its associated probability interval was the narrowest interval that included all three values of $p(\delta_j)$. That is, $\underline{p}(\delta_j) = \min\{p_k(\delta_j)\}$ for $k = 1, 2, 3$ and $\overline{p}(\delta_j) = \max\{p_k(\delta_j)\}$ for $k = 1, 2, 3$. This combination method is known as the ‘unanimity rule’ (Coolen, 1992; Walley, 1991, p.188). The coherent probability intervals derived from the three dicretized precise priors using the unanimity rule are shown in Table IV, columns (2) and (3).

In collecting observational data, Crome et al. employed a standard design for ecological studies known as the ‘BACIP’ design (Hurlbert, 1984). In this design, count data for a particular species is obtained in a logged area and in a comparable unlogged area for several time periods before and after logging takes place. The data is analyzed using an analysis of variance model that includes time, area, and time × area effects. The logging effect, δ, is captured by the last (interaction) term. Details of the study design, sampling model, and associated likelihood function can be found in Crome et al. (1996, pp. 1107–1108).

Since actual data was not published, we computed likelihoods for data generated using the likelihood function referenced above and hypothetical values for the effect sizes.

Specifically, we generated count data for twenty periods (10 pre- and 10 post-logging) assuming a Poisson distribution with a mean of 30 for the unlogged area (pre- and post-logging), 28 for the logged area pre-logging, and 21 for the logged area post-logging. This corresponds to a 25% reduction ($\delta = 0.75$) in species count that can be attributed to logging. Normalized likelihoods for our hypothetical data are shown in column (4) of Table IV.

In the original analysis, derivation of the three precise posterior distributions was straightforward since the prior probability distributions were natural conjugates for the likelihood function. In the re-analysis, the imprecise posterior probability intervals are obtained using the prior probability intervals as constraints, the normalized likelihoods, and the linear programming formulation of the GBR. The posterior probability intervals for each δ_j appear in columns (5) and (6) of Table IV.

Examination of Table IV reveals that there are fewer plausible events indicated by the imprecise posterior than by the imprecise prior. Also, there is less imprecision associated with the posterior ($\Delta = 1.0421$) compared to the prior ($\Delta = 1.2881$). Accordingly, for this example both first- and second-order uncertainty have been reduced after updating with the GBR.

4.3. *Evaluating the efficacy of medical trials*

In the medical field, prior distributions are constructed using information from clinical databases, expert opinion, medical literature, and other sources (Spiegelhalter et al., 1993). As in the previous example, selection of an appropriate prior probability distribution can be controversial. This is especially true when evaluating the efficacy of new medical therapies. Walley et al. (1996) used an imprecise beta distribution (Walley, 1991, section 5.3) to model prior beliefs for the difference in survival probabilities for standard and new therapies. We present a simple application in which the percentage improvement in survival rates for a new versus standard therapy is modeled in terms of upper and lower event probabilities. Our example is derived from a statistical analysis of the CHART (continuous hyperfractionated accelerated radiotherapy) trials.

Parmar et al. (1994) demonstrated the use of Bayesian methods to evaluate a new therapy called 'CHART'. Nine clinicians who were to participate in the trials assessed probabilities for selected values of **x**, the change in the proportion of patients who were disease free after two years using CHART versus standard therapies. Despite large differences in clinicians' prior probability distributions for **x**, simple averaging across priors was employed to derive a precise prior. A normal probability distribution was then fit to the average prior after transforming **x** to a log hazard ratio, ln(**h**), scale.[3] Because this distribution had a mean survival rate equivalent to a 12% improvement in the two-year survival rate using CHART (ln(**h**) ~ $N(0.351, 0.0625)$) it was termed an 'enthusiastic' prior. A 'skeptical' prior was then constructed by shifting the enthusiastic prior downward so that the mean survival rate translated to a 0% improvement in survival rate (ln(**h**) ~ N(0.0, 0.0625)).

An imprecise prior for the nine clinicians could have been constructed using the unanimity rule. However, since the clinicians' priors varied greatly, the unanimity rule resulted in a highly imprecise prior ($\Delta = 2.803$). This level of imprecision often results in posterior probabilities so imprecise that they have no practical value. Consequently, we decided to employ a less conservative combination rule. This rule is detailed in Moral and Selgado (2001) and is as follows:

For $k = 1,\ldots,m$ probability distributions and a constant, c, between 0 and 1, inclusive:

(1) calculate $\mathbf{p} = (1/m)\sum_{k=1}^{m}\mathbf{p}_k$, the 'average' probability distribution;
(2) calculate $\mathbf{p}_k^c = c\mathbf{p} + (1-c)\cdot\mathbf{p}_k$ for $k = 1,\ldots,m$;
(3) use the convex set determined by points $\mathbf{p}_k^c$ to derive upper and lower bounds, $\overline{p}(\delta_j)$ and $\underline{p}(\delta_j)$, for $j = 1,\ldots,n$.

For $c = 0$, the above rule coincides with the unanimity rule. For $c = 1.0$, it is equivalent to simple averaging. We com-

TABLE V
Clinical trials example

X	ln(h)	$\underline{p}(h)$	$\overline{p}(h)$	$f(h_d\|h)$	$\underline{p}(h\|h_d)$	$\overline{p}(h\|h_d)$
−0.0750	−0.2057	0.0125	0.0625	0.2471	0.0195	0.1816
−0.0250	−0.0691	0.0250	0.1729	0.2499	0.0425	0.4110
0.0250	0.0701	0.0969	0.3469	0.2007	0.1369	0.6125
0.0750	0.2145	0.0969	0.2969	0.1369	0.0825	0.4363
0.1250	0.3667	0.1094	0.4094	0.0830	0.0555	0.3948
0.1750	0.5300	0.0969	0.3469	0.0460	0.0273	0.2077
0.2250	0.7088	0.0500	0.2000	0.0243	0.0074	0.0631
0.2750	0.9094	0.0125	0.0625	0.0120	0.0009	0.0098
		$\Delta = 1.3979$			$\Delta = 1.9443$	

promised by choosing $c = 0.5$. This value avoids the problems associated with a prohibitively imprecise prior and still reflects disagreement in expert opinion. The ln(**h**) values and lower and upper probabilities derived from the clinicians' precise priors using $c = 0.5$ appears in columns (2)–(4) of Table V. Column (1) lists **x** values equivalent to the ln(**h**) values in column (2).

Parmar et al. (1994) used hypothetical interim results to demonstrate how a Bayesian analysis could be conducted for the CHART trials. They supposed that of the 492 patients entered in the trials an equal number had been assigned to either therapy, that 37 events had been observed for CHART and 33 for the standard therapy, and that the estimated log hazard ratio, ln($\mathbf{h}_d$), was −0.114. Again assuming that the log hazard ratio follows a normal distribution, the observed data is normal with mean ln($\mathbf{h}_d$) and variance $4/N_d$. N_d is the sum of the events (recurrences of illness) for the standard and new therapies. The derived enthusiastic and skeptical posterior probability distributions are then ln($\mathbf{h}_e$) ~ N(.108, .0299) and ln($\mathbf{h}_s$) ~ N(−0.0597, 0.0299), respectively. Additional details regarding the construction of the prior, data, and posterior probability distributions for the log hazard ratio can be found in Fayers et al. (1997).

As in the previous example, derivation of the precise posterior probability distributions was easy since the priors were natural conjugate densities for the likelihood function. To update our imprecise prior, it was again necessary to explicitly calculate likelihoods. The likelihood function for the observed log hazard ratio is expressed as follows (Fayers et al., 1997)

$$\mathrm{L}(h|h_d) = \frac{1}{\sqrt{8\pi/N_d}} \mathrm{e}^{-(1/2)\left[(\ln(h_d)-\ln(h))(4/N_d)^{-0.5}\right]^2},$$

where

ln(**h**) is the expected (mean) log hazard ratio;

ln($\mathbf{h}_d$) the observed log hazard ratio, and

$4/N_d$ the variance of the observed log hazard ratio.

The normalized likelihoods for the data and imprecise posterior probability distribution are shown in Table V, columns (5)–(7). In this case, the posterior probability distribution is more imprecise than the prior ($\Delta = 1.9443$ versus $\Delta = 1.3979$). For this example, dilation can be attributed to high imprecision of the prior and a relatively diffuse likelihood function.

In examples 4.2 and 4.3, the original analysis used multiple precise priors to bound the range of expert opinion. In effect, these priors defined a convex set of priors. However, the multiple precise posterior distributions derived using Bayes' rule do not define the largest set of probability measures consistent with the convex set of priors and the data. In our reanalysis, we employed the GBR to update bounds on the prior. Consequently, the resultant posterior bounds were minimally coherent. But, because we defined the prior bounds solely in terms of upper and lower event probabilities, our imprecise posterior distributions tended to be excessively wide. In practice, additional constraints on the prior should be elicited to avoid unnecessary imprecision.

5. DISCUSSION

Numerous studies in behavioral decision theory have demonstrated the importance of second-order uncertainty for decision making. The rapid growth of interest in imprecise probability has led to the development of imprecise probability models with many potential applications. The generalized Bayes' rule, formulated as a mathematical programming model, is an important example. However, the GBR has scarcely been used by decision analysts.

Several factors have impeded adoption of the GBR by practitioners of decision analysis. Certainly, the amount of cognitive and computational effort required to construct imprecise versus precise probability models for prior beliefs is an obstacle. We have addressed this issue by demonstrating that Walley's approach to imprecise probability elicitation can be used in conjunction with simple linear programming models to construct coherent imprecise prior probability models. Conducting a Bayesian analysis with imprecise prior probability models is another impediment. Closed-form solutions to the GBR exist for special types of imprecise probability models, such as upper and lower natural conjugate densities (Walley, 1991, p. 205). Linear programming formulations of the GBR are more general, but require more computational effort. In this paper, we showed that Cozman's linear programming formulation of the GBR can be used in conjunction with imprecise prior probability models that are specified by linear constraints on event probabilities. This methodology is unwieldy for event spaces with very large n, but is quite manageable for event spaces that are not overly fine. Linear programming formulations of the GBR are easily implemented using the optimization modules available with common spreadsheet software or any standard linear programming software.

Other issues need to be addressed before imprecise probabilities are routinely used for decision making. Notably, more attention should be devoted to the development of general assessment strategies for imprecise probabilities. Accurate representation of the imprecise prior probability model is critical in a quasi-Bayesian analysis, since probability expressions that are too imprecise or highly contradictory to the data can generate posterior probability models that are practically vacuous—i.e., so vague as to be useless for decision making. While this effect is unavoidable if the assessor's knowledge of the events is limited or inaccurate, it should not occur because the assessment procedure failed to adequately assist the assessor in capturing his beliefs. To date, most research in imprecise probability assessment has been directed at theoretical and computational aspects of assessment rather than the development of practical assessment procedures. The elicitation protocol delineated in Lins et al. (2001) is a useful example of the latter. Interactive computer programs for checking the consistency of probability constraints such as the one developed by Dickey (2003) and the expert systems for assessment suggested by Walley (1996) are also critical if assessment procedures for imprecise probability are to be successful and widely used.

The development of efficient algorithms for propagating imprecise probabilities in influence diagrams, or Bayesian networks, is also an important concern. Breese and Fertig (1991) and Tessem (1992) describe approximative algorithms for solving such problems. Andersen and Hooker (1994) and Fagiuoli and Zaffalon (1998) use non-linear programming methods to formulate efficient algorithms for solving Bayesian networks with imprecise probabilities. But a general effective solution method for large-scale systems has yet to be formulated.

Comparing decision alternatives can also be problematic when probabilities are imprecise. Lower and upper bounds on expectations can be computed if event outcomes are precise. But the ranges of expected values for different decision alternatives can overlap. When this occurs, the expected utility criterion cannot be used to determine the optimal course of action. A variety of strategies for choosing decision alternatives have been proposed for such cases. These include: choosing the alternative that minimizes the maximum possible expected loss (Berger, 1985; Gardenfors and Sahlin, 1982); using secondary measures that enforce 'security' (Levi, 1980); and reporting all admissible plans, but leaving the actual choice of alternatives unspecified (Fertig and Breese, 1990). Recently, Schervish et al.

(2003) examine extensions of expected utility theory to convex sets of probabilities using Γ-minimax (Berger, 1985), E-admissibility (Levi, 1980), and maximality (Walley, 1991).

The potential for indeterminacy may discourage decision makers from including second-order uncertainty in a decision analysis. However, it should be remembered that indeterminacy reflects the limits of knowledge available for making decisions. Forcing assessors to provide precise probabilities in order to generate a unique ordering of decision alternatives conceals but does not eliminate second-order uncertainty and any resultant indeterminacy. Such forced uncertainty absorption can harm decision quality by giving decision makers a false sense of confidence in their decisions (Feagans and Billess, 1981).

NOTES

1. In Walley's theory, a coherent lower prevision is equal to an affinely superadditive lower expectation. The term 'prevision' conveys the idea that an expectation is a subjective guess about future events (Cozman, 1997).
2. Alternatively, a continuous parameter can be discretized by partitioning the parameter space into a finite number of intervals. Calculation of the likelihood function becomes more complicated in this case, since the likelihood for each interval of parameter values is a definite integral of the likelihood function over that interval.
3. The log hazard ratio scale is defined as follows: $\ln(h) = \ln(\ln(P_1)/\ln(P_2)) \cdot P_1$ is the survival rate (proportion) in group 1 and P_2 is the survival rate (proportion) in group 2.

REFERENCES

Andersen, K.A. and Hooker, J.N. (1994), Bayesian logic, *Decision Support Systems* 11, 191–210.

Becker, R.W. and Brownson, F.O. (1964), What price ambiguity? or the role of ambiguity in decision making, *Journal of Political Economy* 72, 62–73.

Benson, P.G. and Whitcomb, K.M. (1993), The effectiveness of imprecise probability forecasts, *Journal of Forecasting* 12, 139–159.

Berger, J.O. (1985), *Statistical Decision Theory and Bayesian Analysis*, 2nd Edition. New York: Springer-Verlag.

Beyth-Marom, R. (1982), How probable is probable? a numerical translation of verbal probability expressions, *Journal of Forecasting* 1, 257–269.

Boiney, L.G. (1993), The effects of skewed probability on decision making under ambiguity, *Organizational Behavior and Human Decision Processes* 56, 134–148.

Breese, J.S. and Fertig, K.W. (1991), Decision making with interval influence diagrams. In Bonissone, P.P., Henrion, M., Kanal, L.N., and Lemmer, J.F. (eds.), Uncertainty in Artificial Intelligence 6, Elsevier Science Publishers, pp. 467–478.

Coletti, G. (1994) Coherent Numerical and Ordinal Probabilistic Assessments, *IEEE Transactions on Systems, Man, and Cybernetics* **24**, 1747–1753.

Coletti, G. and Scozzafava, R. (1999), Coherent upper and lower bayesian updating, in G. DeCooman, F.G. Cozman, S. Moral and Walley, P. (eds.), *Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications*, pp. 101–110, The Imprecise Probabilities Project.

Coolen, F.P.A. (1992), Elicitation of expert knowledge and assessment of imprecise prior densities for lifetime distributions, Memorandum COSOR 92–12, Eindhoven University of Technology, Department of Mathematics and Computing Science, Eindhoven, The Netherlands.

Coolen, F.P.A. (1993), Imprecise conjugate prior densities for the one-parameter exponential family of distributions, *Statistics and Probability Letters* 16, 337–342.

Coolen, F.P.A. and Newby, M.J. (1994), Bayesian reliability analysis with imprecise prior probabilities, *Reliability Engineering and System Safety* 43, 75–85.

Cozman, F.G. (1997), An informal introduction to quasi-bayesian theory for AI. Working paper CMU-RI-TR 97–24, Pittsburgh, PA: Carnegie Mellon University.

Cozman, F.G. (1999), Computing posterior upper expectations, in G. DeCooman, F.G. Cozman, S. Moral and Walley, P. (eds.), *Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications*, The Imprecise Probabilities Project, pp. 131–140.

Cozman, F.G. (2002), Algorithms for conditioning on events of zero lower probability. *Proceedings of the Fifteenth International Florida Artificial Intelligence Society*, Florida: Pensacola, pp. 248–252.

Crome, F.H.J., Thomas, M.R., and Moore, L.A. (1996), A novel approach to assessing impacts of rain forest logging, *Ecological Applications* 6, 104–123.

Curley, S.P. and Yates, J.F. (1985), The center and range of probability intervals as factors affecting ambiguity preferences, *Organizational Behavior and Human Decision Processes* 36, 273–287.

Dickey, J. (2003), Convenient interactive computing for coherent imprecise prevision assessments, in J.M. Bernard, T. Seidenfeld and M. Zaffalon (eds.), *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications,* pp. 218–230, Carleton Scientific.

Einhorn, H.J. and Hogarth, R.M. (1985), Ambiguity and uncertainty in probabilistic inference, *Psychological Review* 92, 433–461.

Ellsberg, D. (1961), Risk, uncertainty, and the savage axioms, *Quarterly Journal of Economics* 75, 643–669.

Fagiuoli, E. and Zaffalon, M. (1998), 2U: An exact interval propagatgion algorithm for polytrees with binary variables, *Artificial Intelligence* 106, 77–107.

Fayers, P.M., Ashby, D. and Parmar, M.K.B. (1997), Tutorial in biostatistics: bayesian data monitoring in clinical trials, *Statistics in Medicine* 16, 1413–1430.

Feagans, T.B. and Biller, W.F. (1981), Assessing the health risks associated with air quality standards, *The Environmental Professional* 3, 235–247.

Fertig, K.W. and Breese, J.S. (1990), Interval influence diagrams, in M. Henrion, R.D. Shachter, L.N. Kanal, and Lemmer, J.F. (eds.), *Uncertainty in Artificial Intelligence 5,* Amsterdam: North-Holland, Elsevier Science Publishers, pp. 149–161.

Fine, T.L. (1988), Lower probability models for uncertainty and nondeterministic processes, *Journal of Statistical Planning and Inference* 20, 389–411.

Fortin, V., Parent, E. and Bobée, B. (2001), Posterior previsions for the parameter of a binomial model via natural extension of a finite number of judgments, in G. DeCooman, T. Fine and Seidenfeld, T. (eds.), *Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications*, Shaker Publishing.

Gardenfors, P. and Sahlin, N.E. (1982), Unreliable probabilities, risk taking, and uncertainty, *Synthese* 53, 361–386.

Goldsmith, R.W. and Sahlin, N.E. (1982), The Role of second-order probabilities in decision making, in P.C. Humphreys, O. Svenson, Vari, A. (eds.), *Analysing and Aiding Decision Processes,* Amsterdam: North Holland, pp. 319–329.

Good, I.J. (1965), *The Estimation of Probabilities: An Essay on Modern Bayesian Methods*, Cambridge, Massachusetts: MIT Press.

Hurlbert, S.H. (1984), Pseudoreplication and the design of ecological field experiments, *Ecological Monographs* 54, 187–211.

Koopman, B.O. (1940), The axioms and algebra of intuitive probability, *Annals of Mathematics* 41, 269–292.

Kozine, I. and Krymsky, V. (2003), Reducing uncertainty by imprecise judgments on probability distribution: application to system reliability. in J.M. Bernard, T. Seidenfeld and Zaffalon, M. (eds.), *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*, pp. 335–344, Carleton Scientific.

Levi, I. (1980), *The Enterprise of Knowledge,* Cambridge, Massachusetts: MIT Press.

Lins, G.C.N. and Campello de Souza, F.M. (2001), A protocol for the elicitation of prior distributions, in G. DeCooman, T. Fine and

Seidenfeld, T. (eds.), *Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications*, Shaker Publishing.

Martz, H.F. and Waller, R.A. (1982). *Bayesian Reliability Analysis*, New York: John Wiley and Sons.

Moral, S. and del Sagnado, J. (1997), Aggregation of imprecise probabilities, in B. Bouchon-Meunier (ed.), *Aggregation and Fusion of Imperfect Infor*mation, Heidelberg, Germany: Physica-Verlag, pp. 162–188.

Parmar, M.K.B., Spiegelhalter D.J. and Freedman, L.S. (1994), The CHART trials: bayesian design and monitoring in practice, *Statistics in Medicine* 13, 1297–1312.

Savage, L.J. (1954), *The Foundations of Statistics*, New York: John Wiley and Sons.

Schervish, M.J., Seidenfeld, T., Kadane, J.B. and Levi, I. (2003), Extensions of expected utility and some limitations of pairwise comparisons, in J.M. Bernard, T. Seidenfeld and Zaffalon, M. (eds.), *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications,* pp. 496–510, Carleton Scientific.

Seidenfeld, T. and Wasserman, L. (1993), Dilation for sets of probabilities, *Annals of Statistics* 21, 1139–1154.

Smith, C.A.B. (1961), Consistency in statistical inference and decision, *Journal of the Royal Statistical Society B* 23, 1–25.

Snow, P. (1991), Improved posterior probability estimates from prior and conditional linear constraint systems, *IEEE Transactions on Systems, Man and Cybernetics*, Part A, 21, 464–469.

Spiegelhalter, D.J., Freedman, L.S. and Parmar, M.K.B. (1993), Applying bayesian thinking in drug development and clinical trials, *Statistics in Medicine* 12, 1501–11.

Tessem, B. (1992), Interval probability propagation, *International Journal of Approximate Reasoning* 7, 95–120.

Utkin, L.V. and Kozine I.O. (2001), Computing the reliability of complex systems. in G. DeCooman, T. Fine and Seidenfeld, T. (eds.), *Proceedings of the Second Internation Symposium on Imprecise Probabilities and Their Applications*, Shaker Publishing.

Utkin, L.V. and Gurov, S.V. (2002), Imprecise reliability measures for some new lifetime distribution classes, *Journal of Statistical Planning and Inference* 105, 215–232.

Walley, P. (1991), *Statistical Reasoning with Imprecise Probabilities*, London: Chapman and Hall.

Walley, P. (1996), Measures of uncertainty in expert systems, *Artificial Intelligence*, 83, 1–58.

Walley, P., Gurrin, L. and Burton, P. (1996), Analysis of clinical data using imprecise prior probabilities, *The Statistician* 45, 457–485.

Walley, P. Pelessoni, R. and Vicig, P. (1999), Direct algorithms for checking coherence and making inferences from confiditional probability assessments. Quad. n. 6/99 del Dipartimento di Matematica Applicata ‘‘Bruno de Finetti’’, University of Trieste.

White III, C.C. (1986), A posteriori representations based on linear inequality descriptions of a priori and conditional distributions, *IEEE Transactions on Systems, Man and Cybernetics*, SMC-16, 570–573.

Wolfson, L.J., Kadane, J.B. and Small, M.J. (1996), Bayesian environmental policy decisions: two case studies, *Ecological Applications* 6, 1056–1066.

Yates, J.F. and Zukowski, L.G. (1976), Characterization of ambiguity in decision making, *Behavioral Science* 21, 19–25.

Address for correspondence: Kathleen M. Whitcomb, Moore School of Business, University of South Carolina, Columbia, SC 29208, U.S.A. e-mail: whitcomb@moore.sc.edu