Effect of Epistemic Uncertainty Modeling Approach on Decision-Making: Example Using Equipment Performance Indicator

PSAM-11

Dana L. Kelly
Robert W. Youngblood

June 2012

This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint should not be cited or reproduced without permission of the author. This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, or any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights. The views expressed in this paper are not necessarily those of the United States Government or the sponsoring agency.
Effect of Epistemic Uncertainty Modeling Approach on Decision-Making: Example Using Equipment Performance Indicator

Dana L. Kellya, Robert W. Youngblooda*
Idaho National Laboratory, Idaho Falls, Idaho, USA

Abstract: Quantitative risk assessments are an integral part of risk-informed regulation of current and future nuclear plants in the U.S. The Bayesian approach to treatment of uncertainty, in which both stochastic and epistemic uncertainties are represented with precise probability distributions, is the standard approach to modelling uncertainties in such quantitative risk assessments. However, there are long-standing criticisms of the Bayesian approach to epistemic uncertainty from many perspectives, and a number of alternative approaches have been proposed. Among these alternatives, a promising (and rapidly developing) approach is based on the concept of imprecise probability. In this paper, we employ a performance indicator example to focus the discussion. We first give a short overview of the traditional Bayesian paradigm and review some of its controversial aspects, for example, issues with so-called noninformative prior distributions. We then discuss how the imprecise probability approach treats these issues and compare it with two other approaches: sensitivity analysis and hierarchical Bayes modelling.

Keywords: Epistemic uncertainty, Decision-making, Imprecise probability

1. INTRODUCTION

Quantitative risk assessments are an integral part of risk-informed regulation of current and future nuclear plants in the U.S. The Bayesian approach to treatment of uncertainty, in which both stochastic and epistemic uncertainties are represented with precise probability distributions, is the standard approach to modelling uncertainties in such quantitative risk assessments. However, there are long-standing criticisms of the Bayesian approach to epistemic uncertainty from many perspectives, and a number of alternative approaches have been proposed. Among these alternatives, a promising (and rapidly developing) approach is based on the concept of imprecise probability [1].

In this paper, we compare several approaches, using a performance indicator example to focus the discussion. Real applications in this area are especially challenged by sparseness of data, making the choice of prior distribution especially important. We first give a short overview of the traditional Bayesian paradigm and review some of its controversial aspects, for example, issues with so-called noninformative prior distributions, and how those distributions behave in performance indicator applications. We then discuss how the imprecise probability approach treats these issues, and compare the imprecise probability approach with two other approaches: sensitivity analysis and hierarchical Bayes modelling.

2. EXAMPLE PROBLEM

Based on current performance information, a decision-maker has to decide whether to spend additional resources on a particular piece of equipment, whose performance influences risk, measured in some appropriate metric. If the risk metric, which is a function of the component’s unreliability (on demand) $p$, is at or below a specified baseline value, then the contribution to risk from that component can be neglected, meaning no additional resources need be expended. Periodic testing provides some information about the component’s unreliability, but does not provide a precise value. The problem is that unreliability is not an observable quantity, so the decision-maker must infer unreliability from prior information and observed performance data.
Assume that the decision-maker is concerned about whether the increment to the risk metric of concern is \(> 10^{-6}\) in appropriate units.\(^1\) Assume that this value is calculated as follows:\(^2\)

\[
R = (p \mid x, n) \frac{\partial R}{\partial p}
\]  

In Eq. 1, \((p \mid x, n)\) is the posterior (current) unreliability on demand, given evidence of \(x\) failures in \(n\) trials of the component, and the partial derivative measures the sensitivity of risk \((R)\) to changes in \(p).\(^2\) We assume a value of \(5 \times 10^{-5}\) for this derivative. Thus, for \(R\) to be of concern \((> 10^{-6})\), we would need the posterior unreliability, \((p \mid x, n)\), to be \(> 0.02\). So if the decision-maker knew that the posterior unreliability were \(< 0.02\), he would not spend additional resources on the subject component.

The question of interest is: on what evidence would the decision-maker decide not to expend additional resources on this particular piece of equipment? The answer depends first of all upon the stochastic process used to model failure of the equipment on demand, which we will take to be a Bernoulli process. We assume that this process gives rise to failures, \(X\), over a specified number of demands, \(n\), and we will take \(n = 210\). Thus, \(X \sim \text{binomial}(p, 210)\).

### 2.1 Constrained Noninformative Prior

If we are working within a Bayesian framework, then the decision depends on the prior distribution of \(p\). We begin with the approximate constrained noninformative (CNI) prior [2] adopted by [3]. Reasons for adopting the CNI prior are given in [3]; basically, among the priors considered for application in [3], the CNI prior had smaller false indication probabilities in the cases examined. As described in [2], the CNI prior for Bernoulli \(p\) can be approximated well by a beta distribution with a specified mean value \((p_0)\) and first (second) parameter of 0.5 for \(p_0 < (>) 0.5\).\(^3\) The form of the beta density used is

\[
g(p) \sim p^{\alpha-1}(1-p)^{\beta-1}
\]  

(2)

In this form of the density, the first parameter is \(\alpha\) and the second parameter is \(\beta\). The mean is given by

\[
E(p) = \frac{\alpha}{\alpha + \beta}
\]  

(3)

In this example, with \(p_0 = 1.9 \times 10^{-3}\), we will have \(\alpha = 0.5\) and \(\beta = 262.66\). The posterior mean is given by

\[
E(p \mid x) = \frac{\alpha + x}{\alpha + \beta + n}
\]  

(4)

The amount of evidence needed to justify nonintervention also depends upon the decision-maker’s utility function. For this example, we adopt the linear utility function used in [4]. However, in this example, because of the small value of the derivative in Eq. 1, we can practically never have \(R > 10^{-5}\), and a value \(> 10^{-4}\) is mathematically impossible (it would require \(p > 1\)). Therefore, to keep the example as simple as possible, we restrict attention to the \(10^{-6}\) threshold alone. The utility function for this example is shown in the form of a payoff matrix in Table 1.

---

\(^1\) The threshold value of \(10^{-6}\) has been set by policy makers at the agency concerned. There are actually several thresholds, corresponding to increasing risk levels, which are color-coded. Green corresponds to \(\Delta R \leq 10^{-6}\), White corresponds to \(10^{-6} < \Delta R \leq 10^{-5}\), and Yellow to \(10^{-5} < \Delta R \leq 10^{-4}\). The highest level, Red, corresponds to \(\Delta R > 10^{-4}\).

\(^2\) In the actual NRC performance indicator program, the metric is \(\Delta R\), which is a function of \(\Delta p\). We ignore that complication for now, essentially assuming that the component’s baseline failure probability is much less than the probability that puts \(\Delta R\) over the threshold of interest.

\(^3\) The shape parameter value of 0.5 is a good approximation to the values given in [2]; the actual value depends on the mean value constraint.
Table 1. Utility function (payoff matrix) for equipment unreliability example

<table>
<thead>
<tr>
<th>Decision</th>
<th>Actual State</th>
<th>Green</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td></td>
<td>100</td>
<td>-50</td>
</tr>
<tr>
<td>White</td>
<td></td>
<td>-200</td>
<td>100</td>
</tr>
</tbody>
</table>

The optimal decision is the one that maximizes expected utility, with the expectation taken over the posterior distribution of \( p \) [5]. Thus, given \( x \) failures are observed in 210 demands, the posterior distribution of \( p \) is beta\((0.5 + x, \ 262.66 + 210 – x)\). As an example, the expected utility of a decision that the risk is in the Green region \( (R < 10^{-6}) \) is given by the following:

\[
U_{\text{Green}} = Pr(R < 10^{-6})(100) + Pr(R > 10^{-6})(-50)
\]  

(5)

Thus, with the values given in the example, the decision-maker whose utility function is described by the payoff matrix in Table 1 could see no more than 10 failures in 210 demands without deciding to expend additional resources (i.e., the optimal decision for > 10 failures is that \( R \) is White), under this particular model.

The CNI prior, being a type of maximum entropy prior [6], is intended to be minimally informative, or “gentle” in the terminology of [7]. However, as discussed in [3], the CNI prior can be relatively “rough” when updated with sparse data. As discussed by [8] and others, this is actually a problem with conjugate priors generally, as they tend to have relatively light tails. Recognizing this, [9] examined the use of a so-called “robust” prior, which in this case is a logistic-Cauchy prior. Using this prior, and proceeding otherwise analogously to the above process, with parameters determined using the approach described in [9], the decision-maker could accept at most 5 failures in 210 demands before having to conclude that more resources should be expended. In other words, the optimal decision is Green for \( x \leq 5 \), and White for \( x > 5 \).

3. ALTERNATIVE APPROACHES FOR REPRESENTING EPISTEMIC UNCERTAINTY

It has long been recognized that the choice of prior can influence a decision in the traditional Bayesian framework. In this framework, a single prior distribution is used to represent epistemic uncertainty in a quantity such as \( p \) in this example. For this reason, it is a tenet of decision-making that sensitivity analysis should be performed on the choice of prior distribution. However, it seems to be a common belief among practicing analysts that the form of the prior distribution is less important than the range, as embodied by, say, an interval from the 5th to the 95th percentile. This belief is sometimes used to pragmatically counter the argument that it is unreasonable to represent an imprecise state of knowledge with a single, precise distribution, what [1] refers to as the “dogma of ideal precision.”

We now explore two alternatives to the ideal precision of the approximate CNI prior and the heavier-tailed logistic-Cauchy prior used in this specific example. The first of these will be a type of robust Bayesian analysis, in the terminology introduced by [8]. The second will be the approach of imprecise probability described in [1], in which intervals rather than distributions are used to represent epistemic uncertainty.

3.1. Robust Bayes

For the robust Bayes approach, we will model a class of CNI priors, in which we allow the mean constraint to be uncertain. We adopt the hierarchical approach taken in [10] in which a second-stage distribution is used to represent uncertainty in the mean value constraint. In this case, we will model the class of beta prior distributions whose mean is uncertain by a factor of 3 around the value of \( \frac{1.9 \times 10^{-3}}{0.5} \), but each with \( \alpha = 0.5 \) because each member of the class is an approximate CNI prior.

We will use the OpenBUGS software [11] to carry out the Bayesian inference using Markov chain Monte Carlo (MCMC) sampling, as described in [12]. The immediate problem is how to encode the epistemic uncertainty about the mean constraint. We specify that \( \alpha = 0.5 \), as discussed above, so that the distributions will be in the class of (approximate) CNI priors. Our epistemic uncertainty is with respect to the mean value
constraint, namely that it lies in the interval \((6.3 \times 10^{-4}, 5.7 \times 10^{-3})\). If we are indifferent to subintervals in this range, we might encode this information as a uniform distribution between \(6.3 \times 10^{-4}\) and \(5.7 \times 10^{-3}\). Another approach might be to place the uniform distribution on the log scale, which would weight small values of the mean more heavily than large values; again, how to represent ignorance is a long-recognized and unsolved (some would say unsolvable) problem of standard Bayesian inference. However we encode the uncertainty in the mean value constraint, we unavoidably induce a prior distribution on \(\mu\) via the equation

\[
\beta = \frac{\alpha(1-\mu)}{\mu},
\]

where \(\mu\) is the mean value. Thus, the parameter \(\beta\) serves to index the class of CNI priors. For this example, if we encode the uncertainty in the mean with a uniform distribution, 7 failures in 210 trials will be required for the expected utility for White to exceed that for Green, making White the optimal decision. With a log-uniform prior used to represent epistemic uncertainty in the mean constraint, the results for this example are the same, although this will not be the case in general.

3.2. Imprecise Probability

This is the generalized probabilistic framework developed by Peter Walley for relaxing the requirement of traditional Bayesian inference that epistemic uncertainty be encoded via a single probability distribution. The details of this approach can be found in [1]. For the example at hand, we will use the lower envelope theorem from [1] to derive lower and upper bound probabilities on \(\Delta R\) (and thus on the expected utility) from the class of beta distributions with \(\alpha = 0.5\). We encode our epistemic uncertainty about the mean constraint as lower and upper previsions; nothing is implied about subintervals, and so no distribution is used, only the lower and upper previsions for the mean value. Stating that the mean value is in the interval \((6.3 \times 10^{-4}, 5.7 \times 10^{-3})\) produces lower and upper (cumulative) prior distributions for \(p\) as shown in Figure 1. The lower distribution is beta(0.5, 87.2) and the upper distribution is beta(0.5, 789).\(^4\) As an aside, the lower curve corresponds to the belief function, and the upper curve corresponds to the plausibility function, as discussed in [13]. Ferson et al. [14] refers to these distribution function bounds as a \(p\)-box.

![Figure 1. Lower and upper prior cumulative distribution functions for \(p\), derived from epistemic uncertainty on the mean constraint of the CNI prior.](image)

One can also reparameterize the beta distribution as follows:

\(^4\)Although the enveloping distributions are beta distributions that approximate CNI priors, this should not be taken to imply that there are distributions between the bounds, and that these are also CNI priors; the imprecise probability approach does not use distributions between the bounds.
In this parameterization, \( t \) is the mean value, and \( s \) is a learning parameter that measures how much the mean is influenced by the observed data. Small values of \( s \) correspond to large influence by the observed data, and vice versa. This can be seen by examining the expression for the posterior mean:

\[
E(p | x) = \frac{st + x}{s + n}
\]

Thus, the posterior mean is a weighted average of the prior mean, \( t \), and the maximum likelihood estimate (MLE) of \( p \), which is \( x/n \). The weights are proportional to \( s \) and \( n \), respectively. Thus large values of \( s \) weight the prior mean more heavily than small values. This parameterization is related to the more standard one by \( \alpha = st \) and \( \beta = s(l - t) \). With \( \alpha = 0.5 \) and \( t \) in the interval \( (6.3 \times 10^{-4}, 5.7 \times 10^{-3}) \), \( s \) will be in the interval \( (87.7, 789) \).

The posterior results conditional upon observing \( x \) are obtained by updating the parameters of the lower and upper distributions in Figure 1. This is straightforward since each is a beta distribution, conjugate to the binomial likelihood. Now, however, there is not a single precise value of expected utility. Rather, there is an uncertainty interval for expected utility, whose bounds are the lower and upper previsions for utility, as discussed in [1]. We will choose the optimal decision based on maximality, as described in [1]. In this approach, a decision is treated as a gamble with an uncertain payoff. The optimal decision is to choose an interval that is not dominated by any other interval. In our example, we consider the lower previsions of the gambles (Green – White) and (White – Green) and choose as our optimal decision the interval whose lower prevision is > 0. For up to 6 failures in 210 demands, the optimal decision under this criterion is Green, as the lower prevision of (Green – White) is > 0, and that for (White – Green) is < 0. However, for \( 7 \leq x \leq 21 \), there is no optimal decision, as neither gamble dominates the other. For \( x > 21 \), the lower prevision of (White – Green) > 0 and that of (Green – White) < 0, and so the optimal decision is now White. Thus, until there are 22 failures in 210 demands, the optimal decision is indeterminate. The results are summarized in Table 2.

<table>
<thead>
<tr>
<th>No. failures</th>
<th>CNI</th>
<th>Logistic-Cauchy</th>
<th>Robust Bayes</th>
<th>Imprecise Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Green</td>
<td>White</td>
<td>Green</td>
<td>White</td>
</tr>
<tr>
<td>5</td>
<td>90.8</td>
<td>-181.6</td>
<td>22.5</td>
<td>-45.0</td>
</tr>
<tr>
<td>6</td>
<td>81.4</td>
<td>-162.8</td>
<td>-4.8</td>
<td>9.7</td>
</tr>
<tr>
<td>7</td>
<td>67.6</td>
<td>-135.1</td>
<td>-25.5</td>
<td>51.0</td>
</tr>
<tr>
<td>8</td>
<td>50.1</td>
<td>-100.2</td>
<td>-37.1</td>
<td>74.3</td>
</tr>
<tr>
<td>9</td>
<td>30.6</td>
<td>-61.1</td>
<td>-44.6</td>
<td>89.1</td>
</tr>
<tr>
<td>10</td>
<td>11.1</td>
<td>-22.2</td>
<td>-47.7</td>
<td>95.5</td>
</tr>
<tr>
<td>11</td>
<td>-6.4</td>
<td>12.7</td>
<td>-49.2</td>
<td>98.4</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>21</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>22</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>

From a traditional Bayesian perspective, the indeterminate results can be interpreted as arising from using a probability distribution to represent the epistemic uncertainty in the mean value constraint, but one which is chosen from the class of all possible distributions whose support is the interval \( (6.3 \times 10^{-4}, 5.7 \times 10^{-3}) \). Some of these distributions will be symmetric, and others will be asymmetric, weighting small and large values of the mean unequally. Some of these would lead to the optimal decision being Green, others to White. Without being able to select a single precise distribution from the possible distributions over this interval, which the robust Bayes approach requires the analyst to do, it is not always possible to reach an optimal decision.
From the perspective of sensitivity analysis, one could simply use expected utility to find the optimal decision at the limits of the constraint. At the lower constraint of $6.3 \times 10^{-4}$, 22 failures in 210 demands are required for White to be the optimal decision. At the upper constraint of $5.7 \times 10^{-3}$, 7 failures are required for White to be the optimal decision. Thus, for intermediate values of the constraint, the optimal decision could be Green or White, and we would have to specify a constraint value in order to determine the optimal decision. Without knowing the actual constraint value, we cannot specify the optimal decision, which is the result that imprecise probability conveys.

4. CONCLUSION

This paper has compared different approaches to a particular example in performance assessment. It was already well established that the choice of prior distribution can have a significant effect on a decision analysis, and this is true in the present example. Here, a distribution based on the maximum entropy principle of [6] leads to quite a different decision than a distribution that encodes similar information but which has much heavier distributional tails. The robust Bayes approach suggested by [8] has also been explored, via a hierarchical approach that uses a class of beta prior distributions. The results are similar to those using a single robust prior, as described in [9]. However, the results with this hierarchical robust Bayes approach can be sensitive to parameterization and choice of prior distributions.

Finally, the imprecise probability approach of [1] was examined, using lower and upper distribution functions from within the beta distribution family as a pragmatic means of forming the natural extension. Unlike the other approaches examined above, this approach does not always point to an optimal decision. This is not a “flaw” of the imprecise probability approach. Quite the contrary: one interpretation of this result, and the one put forth in [1], is that imprecise probability is a more honest representation of the true state of knowledge, when only interval-valued knowledge is at hand. Thus, from a decision-making perspective, even the robust Bayes approach, in which epistemic uncertainty in the mean value constraint of the CNI prior is modeled with a second-stage distribution, may be skewing the results in comparison to the imprecise probability approach, in which intervals are used to represent epistemic uncertainty.

Finally, we note that the imprecise probability approach of [1] contains, as special cases, Dempster-Shafer belief functions and the so-called p-boxes of [14]. Note that a p-box was used in the example above; it is what [1] refers to as upper and lower distribution functions. In [14] Ferson discusses at some length the relationship between p-boxes and belief functions. Imprecise probability also encompasses the possibility functions of Dubois, as discussed by [15]. More on these relations can be found in [1], [14], and [15].

Acknowledgements

This paper was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights.

This manuscript has been authored by Battelle Energy Alliance, LLC under Contract No. DE-AC07-05ID14517 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

References