A comparison of statistical approaches for assessing reliability

Jason Matthew Aughenbaugh*

Applied Research Laboratories,
University of Texas at Austin,
Austin, TX 78713, USA
E-mail: jason@arlut.utexas.edu
*Corresponding author

Jeffrey W. Herrmann

Department of Mechanical Engineering and Institute for Systems Research,
University of Maryland,
College Park, MD 20742, USA
E-mail: jwh2@umd.edu

Abstract: Reliability estimates are useful for making design decisions. We consider the case where a designer must choose between an existing component whose reliability is well-established and a new component that has an unknown reliability. This paper compares the statistical approaches for updating reliability assessments based on additional simulation or experimental data. We consider four statistical approaches for modelling the uncertainty about a new component's failure probability: a classical approach, a precise Bayesian approach, a robust Bayesian approach and an imprecise probability approach. We show that an imprecise beta model is compatible with both the robust Bayesian approach and the imprecise probability approach. The different approaches for forming and updating the designer's beliefs about the product reliability are illustrated and compared under different scenarios of available information. The goal is to gain insight into the relative strengths and weaknesses of the approaches. Examples are presented for illustrating the conclusions.

Keywords: sampling theory; robust Bayesian; imprecise probability; reliability assessment; uncertainty; information; statistical testing; Bayesian.

Reference to this paper should be made as follows: Aughenbaugh, J.M. and Herrmann, J.W. (2008) 'A comparison of statistical approaches for assessing reliability', *Int. J. Reliability and Safety*, Vol. 2, No. 4, pp.265–285.

Biographical notes: Jason Aughenbaugh is currently a Research Associate for the Applied Research Laboratories at the University of Texas at Austin. He received his MSME and PhD from the George W. Woodruff School of Mechanical Engineering at the Georgia Institute of Technology, where he was a National Science Foundation Graduate Research Fellow from 2003–2006. He also holds a BSE in Operations Research and Financial Engineering from Princeton University. He is a member of INFORMS and ASME.

Jeffrey W. Herrmann is an Associate Professor at the University of Maryland, where he holds a joint appointment with the Department of Mechanical Engineering and the Institute for Systems Research. He is the Director of the Computer Integrated Manufacturing Laboratory and Associate Director for the University of Maryland Quality Enhancement Systems and Teams (QUEST) Honours Fellows Program. He is a member of INFORMS, ASME, IIE, SME and ASEE. He earned his BS in Applied Mathematics from the Georgia Institute of Technology. As a National Science Foundation Graduate Research Fellow from 1990 to 1993, he received his PhD in Industrial and Systems Engineering from the University of Florida.

1 Introduction

Engineers make decisions based on their beliefs. These beliefs depend upon the information that the engineer has gathered, and they can change based on new information (e.g., from additional experts, experiments and simulations). Statistical reasoning includes a variety of approaches for updating beliefs based on new information. This type of reasoning is especially important when considering the reliability of a product because product failures are inherently random.

This paper considers different statistical approaches for updating beliefs about the reliability of a product and discusses the relative strengths and weaknesses of each approach. We consider a specific reliability estimation problem under a number of different scenarios of available information.

From a practical perspective, it can be argued that the following properties are desirable for a procedure of testing and analysing reliability (or more generally probability) data. First, it can be valuable to incorporate existing information that may be relevant into the analysis; otherwise, existing information is essentially wasted. Second, the procedure should allow for new information to be incorporated into the estimate. Third, for experimental design and planning, the procedure should help the engineers to determine if they need more information and if so, how much.

In Aughenbaugh and Herrmann (2007), the authors showed how each statistical approach could be applied in different information scenarios in order to help engineers understand each approach and consider the tradeoffs between them. The current paper reviews some of those results but discusses and compares the approaches differently and more completely.

This paper begins by stating the problem and introducing three scenarios that correspond to situations in which the designer initially has no information, has substantial information and has partial information. After briefly describing the statistical approaches that will be considered, the paper then discusses their performance in the no information scenario. Then, the paper considers the impact of the information scenario on the reliability estimates that each approach produces. Finally, the paper compares the approaches.

2 Reliability assessment

The performance of a product is unpredictable and often involves random behaviour. Multiple items (i.e., instances of the product being designed) will be manufactured, sold and used. Ideally, these items are identical. In practice, however, differences exist due to variations in materials and the manufacturing processes used to create them. Moreover, different items are used in different environments and in different ways.

The possibility of failures, though disagreeable to the designer, is unavoidable for the above reasons. Some items will fail, while others will not fail. Here, 'failure' is taken very generally, but we assume it is a one-time event in the lifetime of an item. In practice, the failure could be that the item's performance did not meet certain requirements during a test or that it catastrophically stops working at some point during its lifetime. We assume that each item's failure is a random Bernoulli process.

2.1 Problem statement

We consider the case where a designer must compare a new component that has an unknown reliability to an existing component whose reliability is well-established. A design team wishes to reduce the cost of a product by replacing one particularly expensive component of the system with a new, lower cost component. Cost is a typical motivation, but other issues could motivate the replacement. However, the new component must be at least as reliable as the existing component.

It will be convenient to frame things in terms of failure rates instead of reliability. The failure rate of the existing component is θ_{crit} . Let θ be the failure probability of the new component, which is the parameter of interest. In order to select the new component (as a replacement of the existing one), the designer requires that $\theta \le \theta_{crit}$. Ideally, the designer would have enough data to make a precise assessment of θ, such as '$\theta = 0.01$'. However, as discussed later in this paper, there are practical reasons why the designer cannot or is unwilling to make a precise assessment despite holding some initial beliefs about θ.

We assume that the designer has the opportunity to obtain additional data from testing in order to update his beliefs. The designer can perform n tests in which the performance of the new component – success or failure – is observed. We assume that test performance is the actual performance of interest or is an acceptable surrogate for actual performance. Let m be the number of failures observed (and consequently $n-m$ successes are observed). For convenience, we define a variable $x_i = 1$ if trial i is a failure and $x_i = 0$ otherwise. Then $m = \sum_{i=1}^{n} x_i$. We assume each test is an independent Bernoulli trial, so m is a binomial random variable.

Throughout this paper we will use an example to demonstrate the approaches. In this example, the existing component's failure probability is $\theta_{crit} = 0.05$, and the new component's true failure probability is $\theta = 0.04$, which is unknown to the designer. This paper compares the use of different statistical approaches for updating the designer's beliefs about the failure probability θ given any prior information and the set of results $X = \{x_i\}_{i=1}^{n}$.

2.2. Information scenario descriptions

The feasibility and desirability of a given approach for updating one's beliefs about an uncertain quantity depend upon the initial conditions, including the amount of data currently available and the beliefs that are currently held. In this paper, three scenarios are considered:

1. No relevant prior information is available.
2. Substantial, highly relevant data histories are available.
3. Partially relevant information or a small amount of relevant information is available.

The placement of an analysis task into one of these categories is a subjective judgement of the designer and his assessment certainly could be incorrect; in this case, it may be valuable for subsequent testing and analysis to indicate this error.

After describing the scenarios, we will briefly discuss the use of each statistical approach in each scenario.

2.2.1 Scenario 1: no prior information

In this scenario, the designer has concluded that there is no available information that is relevant to the desired reliability analysis. For example, the new component may be so novel or the environment so different that the designer believes that existing data have little value in predicting the performance of the new component. Essentially, the designer needs to construct an estimate of the reliability from scratch. All inferences about the component's reliability will be made using only the data samples received from the planned experiments, specifically $X = \{x_i\}_{i=1}^n$.

2.2.2 Scenario 2: substantial prior information

In this scenario, there exists substantial information that the designer believes is relevant to the desired reliability analysis. The designer is considering testing as a way of verifying this information. For example, the new component may be a minor modification of the existing component. Alternatively, the new component may have been used in other similar settings, so its past performance is a good indication of its performance in this new setting.

2.2.3 Scenario 3: partial prior information

In this scenario, there exists some information that the designer believes is relevant to the desired reliability analysis, but the designer is considering testing in order to augment this partial information. For example, perhaps the existing data came from another setting that did not stress the component as it will be stressed in the future, or perhaps the data came from tests on an existing component with a similar, but not equivalent design. Additional experiments can be used to verify that the actual performance characteristics are similar and to refine the estimates.

3 Statistical approaches for updating reliability estimates

This paper compares four approaches for analysing data and updating beliefs about the parameter θ in the scenarios of information described in the preceding:

1. classical sampling theory
2. precise Bayesian
3. robust Bayesian
4. imprecise probability theory.

The following subsections briefly describe the use of these approaches for the reliability assessment problem. The relationship between the robust Bayesian approach and imprecise probability theory is also discussed.

3.1 Classical sampling theory approach

Classical sampling theory approaches to statistical analyses are generally emphasised in introductory texts such as Hogg and Tanis (2001). Standard in these approaches is the adoption of a frequentist interpretation of probabilities.

The classical sampling theory approach focuses on the observed data $\{x_i\}_{i=1}^{n}$. An unbiased point estimate of θ is the relative frequency of failures to trials in the sample. Specifically, one can estimate θ as $\hat{\theta} = \sum_{i=1}^{n} x_i / n = m / n$. Let z_α be the $100(1-\alpha)$ percentile of the standard normal distribution. Then, a commonly used approximate one-sided $100(1-\alpha)\%$ confidence interval gives an upper bound on θ (Hogg and Tanis, 2001):

$$\left[0, \frac{m}{n} + z_\alpha \sqrt{\frac{m}{n}\left(1-\frac{m}{n}\right)\frac{1}{n}}\right] \tag{1}$$

We use the upper bound for this problem because the designer would like the failure probability to be below the threshold. If this interval includes θ_{crit}, then one cannot reliably conclude that $\theta \le \theta_{crit}$ based on the test results. We note that other confidence intervals for estimating this parameter have been proposed. For a full discussion of them, see Brown et al. (2001).

Equation (1) is an approximation because the actual coverage probability varies with m and n. A general rule of thumb is to only use this approximation for values such that both $m \ge 5$ and $n - m \ge 5$. We will explore values on both sides of these limits for completeness. For small sample size, one may consider using the t-statistic instead of a normal, although generally when the size is small enough to justify this, the rule of thumb is not met anyway.

The classical approach is well suited to Scenario 1, in which no prior information exists, because it analyses only the observed data, that is, $X = \{x_i\}_{i=1}^{n}$. Because the prior information plays no role in the analysis, the classical approach remains the same in Scenarios 2 and 3.

3.2 Precise Bayesian approach

The Bayesian approach (e.g., Berger, 1985) provides a way to combine existing knowledge and new knowledge into a single estimate by using Bayes's theorem.

Bayes's theorem is a direct extension of the rules of conditional probability, so its validity is generally accepted (Box and Tiao, 1973). Some people use the term 'Bayesian' to describe anything that uses Bayes's theorem and this will be the usage in this paper. Others use the term to refer to any procedure that uses a subjective interpretation of probabilities (more precisely known as subjective Bayesianism or subjectivism). Under a subjective interpretation, a probability is an expression of belief based on an individual's willingness to bet (Savage, 1972; de Finetti, 1974; Lindley, 1982). Bayes's theorem provides an objective way to update any initial probability with objective data. If the prior is a subjective prior, then the posterior is necessarily subjective as well. However, one can also use 'objective Bayesian analyses' in which the prior distributions are defined without subjective input. In these cases, the analysis can be viewed as entirely objective. See Berger (2000) for further discussion and references.

One of the requirements of Bayesian analysis is a prior distribution that will be updated. Many of the strongest arguments both for and against Bayesian analysis (compared to the classical approach) involve the prior. On the one hand, if a designer has existing information about θ, it is important to consider this information in the new estimate of θ. On the other hand, if the designer has no specific information about θ, the introduction of a particular prior may distort the posterior estimate.

The objective selection of a prior distribution in the absence of relevant prior information is a topic of extensive research and debate. The approaches proposed include the use of 'non-informative priors' (Jeffreys, 1961; Zellner, 1977; Berger and Bernardo, 1992), maximum-entropy priors (Fougere, 1990) and data-dependent empirical Bayes approaches (Maritz and Lewin, 1989). Still, whether a single prior distribution can reflect all of the uncertainty is an open question to some observers.

In general, determining this posterior distribution can be computationally burdensome, as the output is a distribution over all possible values of θ. To support analytical solutions, the form of the prior is often restricted to 'conjugate' distributions with respect to the measurement model, in which case the posterior distribution that results from the update has the same type as the prior. When using non-conjugate priors, one can employ existing codes for computing the posterior. See, for instance, the Bayesian inference using Gibbs sampling (BUGS) project at http://www.mrc-bsu.cam.ac.uk/bugs/.

For the problem considered in this paper, in which the number of failures in a given number of tests is a binomial random variable, it is convenient to model the prior as a beta distribution with parameters α and β, written $Beta(\alpha,\beta)$ and having density shown in equation (2) where $B(\alpha,\beta)$ is the beta function.

$$f(\theta;\alpha,\beta)=\frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1} \tag{2}$$

Using this prior and a binomial likelihood function, the posterior will also be a beta distribution. Specifically, when one starts with the prior distribution $Beta(\alpha_0,\beta_0)$ and observes m failures out of n trials, the posterior distribution is $Beta(\alpha_0+m,\beta_0+n-m)$.

Consequently, the update can be done analytically by simple addition and subtraction, an enormous improvement in efficiency over the general case.

For our reliability assessment example, we use the priors shown in Table 1. The selections are made somewhat arbitrarily, but we have explicitly considered different priors for each scenario in order to demonstrate how prior selection can affect the results. In Scenario 1, with no prior information, a natural and common choice is the uniform [0, 1] distribution, a special case of the beta distribution $(\alpha_0 = 1$ and $\beta_0 = 1)$. The applicability of this choice will be revealed in the subsequent analysis.

In Scenario 2, we consider priors that represent the substantial, relevant information. It will be convenient to use a beta distribution. Larger values of α_0 and β_0 yield a prior distribution with less variance. We use both 'good' priors that are good estimates of the true failure probability and 'bad' priors that are not.

In Scenario 3, for the beta distribution, it is natural to choose the parameters α_0 and β_0 such that the mean $\alpha_0 / \left(\alpha_0 + \beta_0\right)$ is close to where the designer believes that the true value of θ may be. We use small absolute values to reflect the inadequate amount of information. Choosing $\alpha_0 \geq 1$ ensures that the density function has a bell-shaped curve. Lindley and Phillips (1976) discuss this process and give some relevant examples. Again, we use both 'good' priors that are good estimates of the true failure probability and 'bad' priors that are not.

Because the problem considered here is an estimation problem (the designer wants to know the true value of the failure probability and to compare this to the critical value), one can consider point estimates of the parameter that are based on the posterior distribution. These estimates include the mean and median of the posterior and the maximum likelihood estimate, i.e., the value of θ that maximises the density $\pi(\theta \mid Y = y)$.

Alternatively, one can express probabilities about the parameter. The probability that $\theta \leq \theta_{crit}$ is given in equation (3).

$$P[\theta \leq \theta_{crit}] = \int_0^{\theta_{crit}} \pi(\theta \mid y) \cdot d\theta \tag{3}$$

If this probability is large enough, the designer can be comfortable that the failure probability is acceptably small.

Another alternative is to use intervals to estimate the parameter. Using the posterior, a 'credible interval' (or more generally, a 'credible set') can be found such that the probability that the true value is an element of the set is equal to a desired value. It is natural to calculate the $100(1-\alpha)$ percentile of the posterior distribution, denoted $\theta_{1-\alpha}$ and form the $100(1-\alpha)\%$ credible set $\{\theta \mid 0 \leq \theta \leq \theta_{(1-\alpha)}\}$. The decision criteria then becomes determining whether this interval contains θ_{crit}. If it does, then the probability that $\theta \leq \theta_{crit}$ is less than $100(1-\alpha)\%$ and the new component is unacceptable at that level.

3.3 Robust Bayesian approach

The robust Bayesian approach, sometimes called Bayesian sensitivity analysis, addresses the problem of lack of confidence in the prior (Berger, 1984, 1985, 1993; Insua and Ruggeri, 2000). The core idea of the approach is to perform a 'what-if' analysis by changing the prior. The analyst considers several reasonable prior distributions. After additional data is collected, each candidate prior is updated, resulting in a set of posterior distributions. This set of posterior distributions yields a range of point estimates and a set of credible intervals. If there is no significant change in the conclusion across this set of posteriors, then the conclusion is 'robust' to the selection of the prior.

This analysis is not possible with a single prior. Even when using a prior with an inflated variance (a common attempt to capture lack of information in the precise Bayesian approach), the analyst is still only considering one scenario. In some cases, this proxy may work adequately, but it does not provide as rigorous an exploration of the problem as the robust Bayesian approach. This difference is discussed further in Section 6.3.

3.4 Imprecise probabilities approach

The theory of imprecise probabilities, formalised by Walley (1991), has previously been considered in risk-based engineering design decisions (Aughenbaugh and Paredis, 2006) and related methods have been considered in reliability (Coolen, 1994; Utkin, 2004a, 2004b). However, a direct comparison to other methods has not been made in the context of updating reliability estimates.

The theory of imprecise probabilities uses the same fundamental notion of rationality as de Finetti's work (1974). However, the theory allows a range of indeterminacy – prices at which a decision-maker will not enter a gamble as either a buyer or a seller. These in turn correspond to ranges of probabilities. For the problem of updating beliefs, imprecise probability theory essentially allows for prior and posterior beliefs to be expressed as sets of density functions, compared to the precise Bayesian requirement that exactly one distribution should capture an individual's beliefs.

As explained previously, the beta distribution represents a convenient prior for the example problem in this paper under the precise Bayesian approach. For the robust updating approach, it is convenient to use the imprecise beta model, described by Walley (1991) and Walley et al. (1996) and to re-parameterise the beta so that the density of θ is as given in equation (4).

$$\pi_{s,t}(\theta) \propto \theta^{st-1}(1-\theta)^{s(1-t)-1} \tag{4}$$

Compared to the standard parameterisation of $Beta(\alpha,\beta)$, this means that $\alpha = s \cdot t$ and $\beta = s \cdot (1-t)$ or equivalently that $s = \alpha + \beta$ and $t = \alpha / (\alpha + \beta)$. The convenience of this parameterisation is that t is the mean of the distribution, which has an easily grasped meaning for both the prior assessment and the posterior analysis. The model is updated as follows: if the prior parameters are s_0 and t_0, then the posterior parameters after n trials with m failures are given by $s_n = s_0 + n$ and $t_n = (s_0 t_0 + m) / (s_0 + n)$. Since $s_n = s_0 + n$, s_0 can be interpreted to be a virtual sample size of the prior information; it captures how much weight to place on the prior compared to the observed data. Selecting this parameter therefore depends on the available information and will be discussed with the different information scenarios.

Following Walley (1991), the parameters can be imprecise. That is, the priors are the set of beta distributions with $\alpha_0 = s_0 t_0$ and $\beta_0 = s_0(1 - t_0)$ such that $\underline{t}_0 \le t_0 \le \overline{t}_0$ and $\underline{s}_0 \le s_0 \le \overline{s}_0$. After the test results are observed, each prior in the set is updated as described above.

The use of the imprecise beta model obviously affects the decision process in several ways, but most revealing is how it affects point estimation. Using a precise prior input, one gets a precise posterior output and a single point estimate. In the imprecise model, there are multiple posterior distributions and consequently a range of point estimates. In particular, the interval $[\underline{t}_n, \overline{t}_n]$ is an interval estimate of θ, where the endpoints are calculated as follows:

$$\underline{t}_n = \min_{\underline{s}_0 \le s_0 \le \overline{s}_0} \{(s_0 \underline{t}_0 + m) / (s_0 + n)\} \tag{5}$$

$$\overline{t}_n = \max_{\underline{s}_0 \le s_0 \le \overline{s}_0} \{(s_0 \overline{t}_0 + m) / (s_0 + n)\} \tag{6}$$

One approach for deciding whether or not $\theta \le \theta_{crit}$ is to compare this interval to θ_{crit}. If $\overline{t}_n \le \theta_{crit}$, then the designer knows that a decision that $\theta \le \theta_{crit}$ is robust across the range of prior assumptions when using the mean estimate of θ as the decision criterion. However, as with the precise case, a designer may wish to calculate the probability that $\theta \le \theta_{crit}$ and compare this with some critical value. One can find upper and lower probabilities for this based on the set of posterior distributions. Although this varies across the set of posterior distributions, the minimal and maximal values occur using the posterior distributions that correspond to the priors at the extreme points of $[\underline{t}_0, \overline{t}_0] \times [\underline{s}_0, \overline{s}_0]$. Finally, one can determine a set of credible intervals from these posterior distributions.

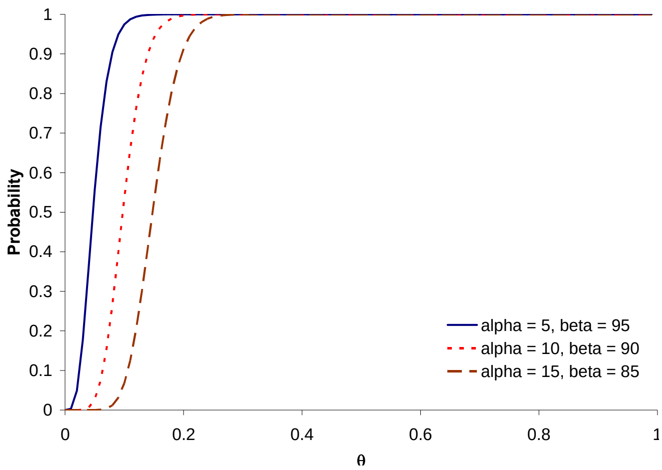
3.5 Connection between robust Bayesian and imprecise probability approaches

Although the motivations differ between the robust Bayesian approach and the imprecise beta model, the computational approaches both involve the application of Bayes's theorem to a set of priors, thus resulting in a set of posterior distributions. For example, consider a statistical description of experiments on a similar component as the baseline, which is well modelled by the beta prior $\alpha_0 = 10$ and $\beta_0 = 90$. However, the engineers believe that the new component actually could have a smaller mean failure rate, and thus the baseline distribution could be shifted to the left to form the prior, yielding $\alpha_0 = 5$ and $\beta_0 = 95$. Conversely, there may also be a reason to believe that the new design has a higher failure rate, and thus it would be prudent to shift the distribution to the right to form the prior, yielding $\alpha_0 = 15$ and $\beta_0 = 85$. Figure 1 shows these three probability distribution functions. The extreme distributions form the boundaries of a box of distribution functions, which is called a p-box (Ferson et al., 2002).

This representation is consistent with both a set of prior distributions and a statement of imprecise probabilities, although the formulation of the priors may differ (Pericchi and

Walley, 1991). All of the beta distributions bounded by this p-box could be considered priors for robust Bayesian analysis. Using the parameters s_0 and t_0 instead of α_0 and β_0, we see that $s_0 = 100$ for all of these distributions, but t_0 ranges from 0.05 to 0.15. Thus, this set of distributions is also consistent with an imprecise beta model describing imprecise probabilities. Consequently, in this reliability assessment problem, it is consistent to use the imprecise beta model to represent both imprecise probability distributions and a robust Bayesian set of prior distributions and to apply Bayes's theorem to update the model to determine the set of posterior distributions. However, not all methods for imprecise probabilities are compatible with robust Bayesian techniques.

Figure 1 P-box of different beta distributions (see online version for colours)



3.6 *Prior selection for the imprecise beta model*

In Scenario 1, when there is a complete lack of prior information, the appropriate starting point is a 'vacuous prior' (so-called because it contains the least information). In the imprecise beta model, this means setting $\underline{t}_0 = 0$ and $\overline{t}_0 = 1$, all that is said is that the probability is somewhere between 0 and 1, which is the least specific statement that is possible. We must also choose a 'learning parameter' s_0. Because it reflects how much 'importance' to assign to the prior data and there is no prior data, one should select a small learning parameter. However, if it is too small, the posterior may react too quickly to the data. As described by Walley (1991), $s_0 = 2$ has a number of good properties, but one can also allow for a range, such as $\underline{s}_0 = 0$ and $\overline{s}_0 = 2$, that contains several of the different precise Bayesian priors proposed in the literature.

Note that $t_0 = 0.5$ and $s_0 = 2$ corresponds to the beta distribution with $\alpha_0 = 1$ and $\beta_0 = 1$. Thus, this vacuous set of priors includes the single uniform prior considered in the precise Bayesian approach with no information. This highlights the difference between a uniform (precise) prior and a vacuous prior. A vacuous prior is a set of prior distributions that happens to contain the uniform prior; therefore, the vacuous prior is more general than the uniform prior.

In Scenario 2, given substantial prior information, the designer can choose $\overline{t}_0$ and $\underline{t}_0$ to be upper and lower bounds on θ. If the designer is unsure of how much 'importance' to assign to the prior data, then the designer could specify upper and lower learning parameters $\overline{s}_0$ and $\underline{s}_0$. In the case of substantial information, we would expect $\overline{t}_0$ and $\underline{t}_0$ to be close to each other and $\overline{s}_0$ and $\underline{s}_0$ to be large.

In Scenario 3, with less information than in Scenario 2, the range of probabilities is wider and the values of s_0 are smaller, indicating a smaller 'pseudo sample size' for the prior. In general, the range for t_0 will be larger since there is more uncertainty in the estimates when there is less information.

Table 1 lists the priors that we chose, using the above guidelines, for our reliability assessment example. The robust Bayesian priors are selected to be generally consistent with the precise priors in that the Bayesian priors contain the precise priors as specific cases.

Table 1 Bayesian priors for $\theta = 0.04$ case

Approach	*No prior info.*	*Partial prior info. (good)*	*Partial prior info. (bad)*	*Substantial prior info. (good)*	*Substantial prior info. (bad)*
Precise Bayesian	$\alpha = 1$ $\beta = 1$	$\alpha = 1.1$ $\beta = 25$	$\alpha = 1.8$ $\beta = 22$	$\alpha = 4$ $\beta = 96$	$\alpha = 6$ $\beta = 94$
Robust Bayesian	$\underline{t}_0 = 0$ $\overline{t}_0 = 1$ $\underline{s}_0 = 0$ $\overline{s}_0 = 2$	$\underline{t}_0 = 0.00$ $\overline{t}_0 = 0.05$ $\underline{s}_0 = 20$ $\overline{s}_0 = 30$	$\underline{t}_0 = 0.05$ $\overline{t}_0 = 0.10$ $\underline{s}_0 = 20$ $\overline{s}_0 = 30$	$\underline{t}_0 = 0.035$ $\overline{t}_0 = 0.045$ $\underline{s}_0 = 80$ $\overline{s}_0 = 120$	$\underline{t}_0 = 0.055$ $\overline{t}_0 = 0.065$ $\underline{s}_0 = 80$ $\overline{s}_0 = 120$

4 Analysis of no information scenario

In this section we will compare the statistical approaches under Scenario 1, the no information scenario. Their performance in Scenarios 2 and 3 is considered in Section 5.

To illustrate the differences, consider the example presented earlier and the following choices for the statistical approaches: the classical approach uses the 95% confidence interval of equation (1), the Bayesian prior is the uniform $(\alpha_0 = 1, \beta_0 = 1)$, and the imprecise beta prior is the vacuous prior $(\underline{t}_0 = 0, \overline{t}_0 = 1)$ with a learning parameter of $\underline{s}_0 = 0$ and $\overline{s}_0 = 2$.

Table 2 describes the conclusions for some of the possible outcomes of running $n = 100$ trials. Note that for $m < 5$, $n = 100$ the rule of thumb for equation (1) is not met. As the number of observed failures increases, the confidence interval on θ widens and the posterior beta distributions shift, which increases the 95% credible intervals for both the precise Bayesian approach and the imprecise beta model.

Table 2 Conclusions based on 100 trials

# of failures	*Classical approach*	*Precise Bayesian*	*Imprecise beta*
m	*95% confidence interval*	*95% credible interval*	*Range of 95% credible intervals*
0	[0, 0.000]	[0, 0.029]	[0, 0] to [0, 0.046]
1	[0, 0.026]	[0, 0.046]	[0, 0.029] to [0, 0.061]
2	[0, 0.043]	[0, 0.061]	[0, 0.046] to [0, 0.075]
3	[0, 0.058]	[0, 0.075]	[0, 0.061] to [0, 0.088]
4	[0, 0.072]	[0, 0.088]	[0, 0.075] to [0, 0.101]
5	[0, 0.086]	[0, 0.101]	[0, 0.088] to [0, 0.114]
6	[0, 0.099]	[0, 0.114]	[0, 0.101] to [0, 0.126]
7	[0, 0.112]	[0, 0.126]	[0, 0.114] to [0, 0.138]
8	[0, 0.125]	[0, 0.138]	[0, 0.126] to [0, 0.150]
9	[0, 0.137]	[0, 0.150]	[0, 0.138] to [0, 0.162]
10	[0, 0.149]	[0, 0.162]	[0, 0.150] to [0, 0.174]

The imprecise beta approach covers all possible low-evidence beta priors with means that range from 0 to 1. Each prior leads to a single posterior and consequently the imprecise beta approach gives a range, or interval, of point estimates in addition to credible intervals. Note that the larger imprecise beta credible intervals contain the precise Bayesian credible intervals, as they should because the set of priors includes the uniform prior used in the precise Bayesian approach.

A comparison between the Bayesian credible intervals and the classical confidence intervals is not exact because the meaning of the two types of intervals is different. However, it is common in both approaches to use a 95%-level for decision making since these intervals form reasonable bounds on a point estimate of the reliability. Consequently, an informal comparison is informative. As shown in Table 2, the classical confidence intervals are just slightly smaller than the smallest credible intervals for the imprecise beta approach and are considerably smaller than the credible intervals for the precise Bayesian approach. In general, one expects more information to yield smaller intervals.

As we will see also in other results later in this paper, the quality of the information matters as well. The smallest imprecise beta credible intervals much more closely match the classical confidence intervals because the set of priors contains those with near-zero values of α and β (which corresponds to a near-zero learning parameter). These near-zero priors have little influence on the posterior distribution and are therefore arguably less informative than the uniform prior distribution, in which $\alpha_0 = \beta_0 = 1$ (Zhu and Lu, 2004). When the prior's parameter values are near zero, the mean of the posterior is very close to m/n and the credible interval is close to the classical confidence interval. Other priors have larger parameter values (up to $\alpha_0 + \beta_0 = 2,$ in which case the learning parameter equals 2). Some of these have means close to $m/n,$ but many others have means that are far away from $m/n.$ The combination of the larger learning parameter and large mean leads to posteriors whose means are not as close to m/n and credible intervals that are not as close to the classical confidence interval. The precise Bayesian credible interval is not as close to the classical confidence interval for the same reason: it has $\alpha_0 + \beta_0 = 2$ and its mean equals 0.5. The posterior mean is nearly 0.01 greater than m / n.

Alternatively, consider how the conclusions would change based on the number of trials, for a fixed relative frequency of observed failures to trials. For instance, if *m/n* = 0.04, then the conclusions change based on the number of trials as shown in Table 3. The confidence intervals decrease in width as the number of trials increases because the large number of trials reduces the sample variance. Similarly, the credible intervals decrease in width because the variance of the posterior distribution decreases. This also increases the probability that $\theta \le \theta_{crit}$. The probability from the precise Bayesian approach is in the middle of the range of upper and lower probabilities that result from the imprecise beta model. Note also that this range significantly decreases as the number of trials increases.

Tables 2 and 3 illustrate how the results of the imprecise beta approach reflect the amount of information available, while the precise Bayesian approach conceals it. If we compare the results for 6 failures out of 100 trials from Table 2 and the results for 2 failures out of 50 trials from Table 3, we see that the credible intervals for the precise Bayesian approach are nearly the same, despite the difference in observed failure rate (6% versus 4%) and the different amounts of data. However, the ranges of credible intervals for the imprecise beta approach are quite different and reflect the underlying amount of information. In the first case (where there is more information), the difference of the upper bounds is 0.025. In the second case (where there is less information), the difference of the upper bounds is 0.055.

Table 3 Conclusions with observed failures at 4%

Number of trials	*Classical approach*	*Precise Bayesian*	*Imprecise beta*
n	*95% confidence interval*	*95% credible interval*	*Range of 95% credible intervals*
25	[0, 0.104]	[0, 0.170]	[0, 0.109] to [0, 0.223]
50	[0, 0.086]	[0, 0.118]	[0, 0.090] to [0, 0.145]
100	[0, 0.072]	[0, 0.088]	[0, 0.075] to [0, 0.101]
200	[0, 0.063]	[0, 0.071]	[0, 0.064] to [0, 0.077]
500	[0, 0.054]	[0, 0.057]	[0, 0.055] to [0, 0.060]
1000	[0, 0.050]	[0, 0.052]	[0, 0.051] to [0, 0.053]
2000	[0, 0.047]	[0, 0.048]	[0, 0.047] to [0, 0.048]

5 Estimate comparison

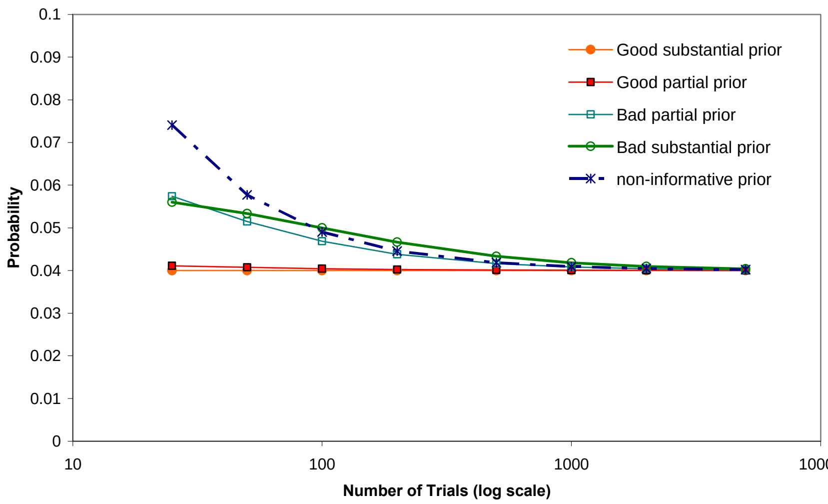
In this section we will consider how the available information affects the results of the precise Bayesian approach and imprecise beta model. We will consider both the quality of the prior distribution (is it close to θ or not) and the magnitude of the prior (does it represent no information, partial information, or substantial information). We will consider the case in which the number of observed failures equals the expected number. To perform this comparison, we will use the priors listed in Table 1.

5.1 Point estimates for θ

All three methods can be used to create point estimates for θ. While these point estimates may be complemented by confidence or credible intervals (as discussed in Section 4 for the case of no prior information), it is informative to consider how the point estimates change under the different information scenarios. We assume that the number of observed failures equals the expected number; that is $m = \theta n$, with $\theta = 0.04$, for all n. In practice, the observed frequency will differ from the true long range frequency. The average performance in the context of decision making is addressed in other articles (Aughenbaugh and Herrmann, 2008, 2009).

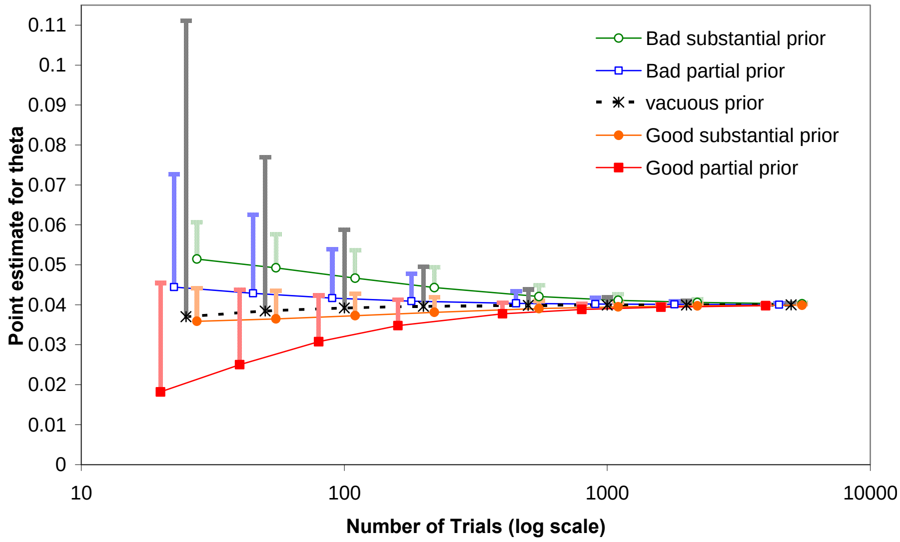
In the classical approach, the point estimate is independent of the initial information scenario and is given by $\hat{\theta} = m / n$ which is constant by assumption. For the other two approaches, the results change as the new information updates the prior assumptions. The results for the precise Bayesian approach are shown in Figure 2 for each of the prior scenarios defined in Table 1. Notice that the 'bad partial prior' information scenario converges more quickly to the true probability than the non-informative prior does, but that the bad substantial prior scenario converges more slowly. This reveals an interesting trade-off between the quantity and quality of information. A little information that is in the neighbourhood of the truth is valuable, but a lot of such close-but-inaccurate information can be detrimental.

Figure 2 For the precise beta model, the evolution of point estimates for θ (see online version for colours)



The imprecise beta model can be used to determine upper and lower bounds on point estimates of θ, as shown in Figure 3 for each of the prior scenarios defined in Table 1. To improve the readability of the chart, the lower bounds are shown with lines, the upper bounds are shown using the error bars and the abscissa for each point for an informative prior is shifted to the right or left of the point for the non-informative prior. In all cases, the lower and upper bounds converge as the number of trials increases. Additionally, it should be noted that initially the bad substantial prior result is a subset of the bad partial prior result, which in turn is a subset of the vacuous prior result. As n increases, the vacuous prior and partial prior results adapt faster to the observed data and converge more quickly to the truth than the substantial prior case does.

Figure 3 For the imprecise beta model, the range of point estimates for θ (see online version for colours)



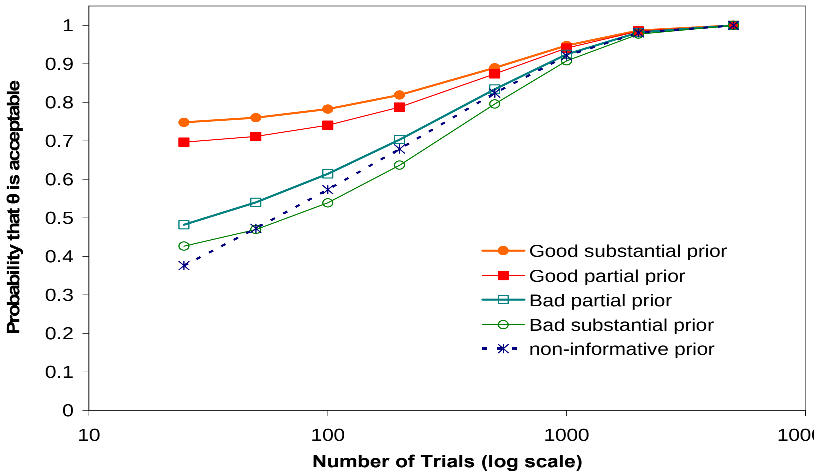
Note: The lower bounds are shown with the solid lines and the upper bounds are shown with the error bars. The abscissa of each series is shifted slightly to improve display.

5.2 Probability of an acceptable design

The probability that $\theta \leq \theta_{\text{crit}}$ is a key output of both the precise Bayesian approach and the imprecise beta model approach, but this estimate is not applicable in the classical approach. As in the previous section, we assume that the number of observed failures equals the expected number; that is, $m = \theta n$, with $\theta = 0.04$.

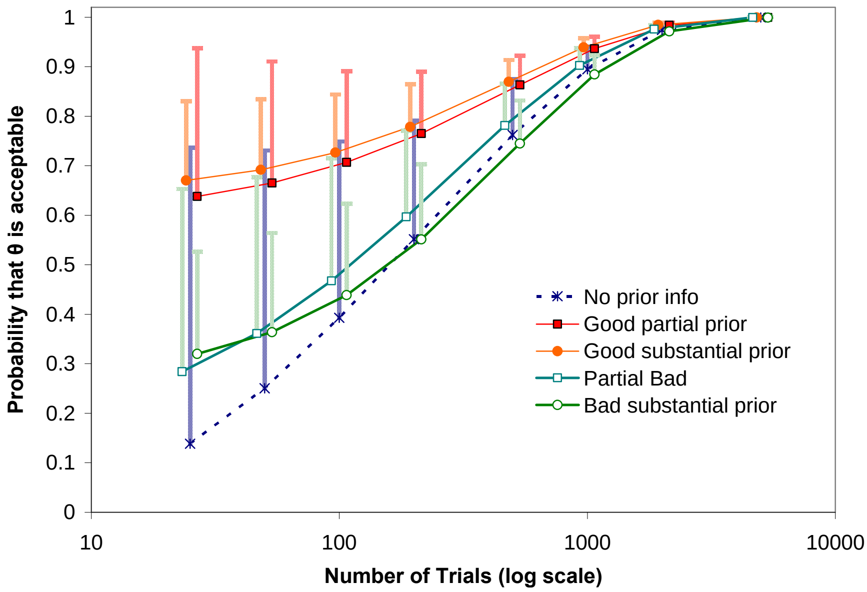
Figure 4 shows the change in this probability using the precise Bayesian approach for the five priors listed in Table 1. The plot reveals that, as expected, more good information leads to a better assessment. However, it also reveals that having a small amount of bad prior information is useful (compared to the non-informative prior), but a lot of bad information can be detrimental, as was the case with the point estimates. A 'bad' estimate that is near the truth but has little weight essentially provides a good starting point for fast updates based on the observed data, especially compared to the uniform prior, which essentially estimates that θ is just as likely to be near 0.96 (or any other value) as near the true value of 0.04. A bad estimate that is given a lot of weight will not allow the posterior to adapt quickly enough to the observed data, as illustrated by the non-informative curve crossing the bad substantial prior curve around 90 trials.

Figure 4 For the precise Bayesian approach, the posterior probability of that θ is acceptable (see online version for colours)



The results for the imprecise beta model are shown in Figure 5 for the five priors listed in Table 1. As before, the lower bounds are shown with the solid lines, the upper bounds are shown with the error bars and the abscissa of each series is shifted slightly to improve display. In this figure, it is clear not only how the probability converges to one (as it did in the precise case, too), but also how the available information narrows in on the appropriate value.

Figure 5 For the imprecise beta approach, the posterior probability that θ is acceptable (see online version for colours)



Note: As before, the ranges are shown from lower (solid) to upper (error bars), abscissa offset slightly.

For example, with 25 trials starting with no prior information, the precise approach suggests that $P[\theta \le \theta_{\text{crit}}] = 0.38.$ Thus, based on this evidence, it is more likely that $\theta > \theta_{\text{crit}},$ meaning the new design is inferior to the old. In comparison, the imprecise beta approach yields the interval estimate $0.14 \le P[\theta \le \theta_{\text{crit}}] \le 0.74.$ This suggests that, given the available information, the likelihood that the new product is an improvement ranges from highly unlikely to fairly likely. This imprecision explicitly reflects the underlying scarcity of information.

6 Discussion of approaches

This section will summarise the above results and make additional observations in order to compare the statistical approaches.

6.1 Adopting a Bayesian viewpoint

The debate between Bayesians and non-Bayesians has been long, contentious and without universal resolution (Rothenberg, 1974; Zellner, 1974; Clarotti, 1993). However, there are certain properties of Bayesian analysis that are often attractive in practice.

First, Bayesian analyses obey the 'likelihood principle', which basically states that all relevant information from an experiment is contained in the likelihood function and that two likelihoods contain the same information if they are proportional to each other [for example, see (Berger, 1985)]. This implies that inferences should be made based on the data that was actually observed, not on hypothetical data that may have been observed but was not.

Lindley and Phillips (1976) give the following example of the likelihood principle as it applies to Bayesian and sampling theory approaches, showing that sampling theory approaches depend on the stopping rule used in the experiment. Consider experiment A in which n trials are performed and the number of failures m are counted. In experiment B, trials are performed until m failures are observed, which happens to take n trials in this example. Even though the same result of m failures out of n trials was observed in each experiment, a sampling theory approach can lead to different conclusions in a hypothesis test. This occurs because experiment A involves a binomial distribution and experiment B involves a negative binomial distribution. However, the likelihood functions differ only by a scalar constant and consequently the Bayesian inference will be the same in experiments A and B, just as the actually observed data is the same in both experiments.

Second, a Bayesian analysis can take advantage of existing information. Thus, in some cases the Bayesian approach can incorporate more information than a classical approach. Of course, the impact of prior beliefs decreases as more data is obtained. Third, Bayesian approaches enable analysts to make direct probability statements about hypotheses and model parameters, statements that cannot be made in classical approaches.

The classical and Bayesian approaches are compatible under certain scenarios. For example, if the Bayesian prior is a constant non-informative one and the analyst uses the mode of the posterior distribution as the point estimate, then this is exactly the same as the maximum likelihood estimate in the classical approach (Hogg and Tanis, 2001).

The precise Bayesian approach can involve more computational effort than the classical approach, especially when the use of conjugate priors cannot be justified. It also requires the prior information to be elicited and formulated correctly, which is not always a straightforward or inexpensive task. A robust Bayesian approach reduces the necessity of complete elicitation while still incorporating information that is easily obtained or available.

The classical approach requires less computational effort but does not take into account any information that might be available before testing. This feature has the additional advantage of being insensitive to the quality of the information. If one ignores the information, it doesn't matter if it is wrong.

6.2 Selecting a prior distribution

The Bayesian approaches can use prior information to help the designer make an accurate estimate with less additional information. However, they are greatly affected by the quality and amount of existing information. In general, it appears that using a prior based on partial information is less 'risky' than using a prior that represents substantial information. Using a prior based on substantial information can lead to poor results if it is a 'bad' prior because overcoming the incorrect information will take a great amount of new information. Of course, if the prior is 'good', then good decisions can be made quickly. Therefore, using partial information may be a good compromise. A prior (or set of priors) based on such partial information provides an effective head start compared to the non-informative prior (or vacuous priors) with minimal investment in information gathering, but, if the prior information is incorrect, new information can overcome the prior more easily than it would if the prior was based on substantial information. Even when substantial information is available, a prior can be constructed that discounts this information, for example by choosing smaller parameters of the beta distribution.

Selecting the form of the prior distribution is also an important part of the Bayesian approach. The precise Bayesian can be more difficult to calculate when conjugate priors are not available. Although a particular conjugate prior may be useful for computational convenience, it may not represent the prior information perfectly. When there is little prior information available, additional challenges occur. Using a single precise prior in this case seems to conceal the lack of information, as shown by the comparison to the robust Bayesian approach.

6.3 Specifying a set of prior distributions

There are at least two arguments as to why a designer should not consider a single prior distribution – one practical and one philosophical. First, because eliciting and assessing an individual's beliefs is a resource intensive process, it will often be impractical to fully characterise them (Savage, 1971; Weber, 1987; Walley, 1991; Groen and Mosleh, 2005). Consequently, only a partial (i.e., imprecise) characterisation will be available. This is the view that advocates of the robust Bayesian or Bayesian sensitivity analysis approaches hold.

The second argument asserts that true beliefs need not be precise. A person may be unwilling to commit to either side of a gamble when he has no information about it. It is possible that one would be satisfied with some price, but this is not a condition for rationality. This view is the one that advocates of imprecise probabilities generally hold (Pericchi and Walley, 1991; Walley, 1991).

The use of upper and lower probabilities has a number of advantages in this situation (Walley, 1996; Walley et al., 1996). Most relevant to the decision being discussed here, the results reflect the amount of information on which the conclusions are drawn; one can distinguish between probabilities based on a large amount of data and those based on ignorance. Moreover, the approach is able to model prior ignorance in a very intuitive way and it includes common Bayesian models for prior ignorance as special cases.

For example, in the partial information scenario, the precise Bayesian approach chooses one prior for the new component's failure probability, an approach that is consistent with the premise that an individual's beliefs can always be modelled by exactly one prior distribution. If the designer is unsure about the failure probability, a precise Bayesian analyst may address this lack of confidence in the estimate by increasing the variance of the prior model, thus reflecting more uncertainty of some kind. Taken to the extreme, a complete lack of information generally leads to a uniform distribution. Unfortunately, the use of a uniform distribution confounds two cases: first, that nothing is known; second, that all failure probabilities between 0 and 1 are equally likely, which is actually substantial information.

In the context of a large engineering project in which there are many individuals, this is an important distinction. For example, one engineer's complete lack of knowledge about some aspect of the system may be offset by another engineer's expertise or by additional experimentation. However, if substantial analysis has already led to the conclusion that certain outcomes are equally likely, then it would be inefficient to expend additional resources examining those probabilities. In a precise Bayesian approach, both scenarios would be represented by a uniform prior distribution, even though the underlying scenarios are inherently different. The robust Bayesian approach allows one to consider the different scenarios independently rather than aggregating them together. This affords the design team the opportunity to make different, more appropriate decisions about information management under the two scenarios.

Using the robust Bayesian approach or imprecise probabilities can require more effort than the precise Bayesian approach, especially when conjugate models such as the imprecise beta (or imprecise Dirichlet) models do not apply. Also, as multiple sources of uncertainty are propagated, it may be that the resulting intervals are quite large, leading to considerable regions of imprecision, as our results showed.

The imprecise beta model allows the designer to use ranges to express his beliefs about the value of the parameter (via a range of means) and the importance of the prior data (via a range of learning parameters). Such caution about the prior information reduces the precision of the estimate, but it also improves its accuracy. Essentially, the analyst can say less about the parameters, but what the analyst does say is more likely to be true. Thus, this model reflects both the quality and quantity of information available, including indirectly revealing any conflict between the prior information and the new data.

7 Summary

A variety of statistical approaches are available for updating assessments of the reliability of a component based on test results. This paper has compared the classical sampling theory approach, the precise Bayesian approach and an imprecise beta model, which is compatible with both a robust Bayesian approach and a particular imprecise probabilities approach for this problem. We have focused on exploring how each method behaves as more information is collected and using those results to illustrate the strengths and weaknesses of each approach. Direct quantitative comparisons and general judgements of absolute superiority are not possible due to the inherently different assumptions and structures underlying each method. These conclusions have been illustrated using a specific example decision. A designer can use the presented methods to gain insight into the tradeoffs that exist in a specific domain.

Acknowledgements

This work was supported in part by the Applied Research Laboratories at the University of Texas at Austin IR&D grant 07–09.

References

Aughenbaugh, J.M. and Herrmann, J.W. (2007) 'Updating uncertainty assessments: a comparison of statistical approaches', *2007 ASME International Design Engineering Technical Conferences,* Las Vegas, NV.

Aughenbaugh, J.M. and Herrmann, J.W. (2008) 'A comparison of information management using imprecise probabilities and precise Bayesian updating of reliability estimates', *NSF Workshop on Reliable Engineering Computing,* Savannah, GA.

Aughenbaugh, J.M. and Herrmann, J.W. (2009) 'Reliability-based decision-making: a comparison of statistical approaches', *Journal of Statistical Theory and Practice* (to appear).

Aughenbaugh, J.M. and Paredis, C.J.J. (2006) 'The value of using imprecise probabilities in engineering design', *Journal of Mechanical Design*, Vol. 128, No. 4, pp.969–979.

Berger, J.O. (1984) 'The robust Bayesian viewpoint', in J.B. Kadane (Ed.): *Robustness of Bayesian Analysis*, North-Holland, New York, pp.63–124.

Berger, J.O. (1985) *Statistical Decision Theory and Bayesian Analysis*, 2nd ed., Springer.

Berger, J.O. (1993) *An Overview of Robust Bayesian Analysis*, Report Num. Technical Report #93–53c, Purdue University, West Lafayette, IN.

Berger, J.O. (2000) 'Bayesian analysis: a look at today and thoughts of tomorrow', *Journal of the American Statistical Association*, Vol. 95, No. 452, pp.1269–1276.

Berger, J.O. and Bernardo, J.M. (1992) 'On the development of the reference prior method', in J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith (Eds.): *Bayesian Statistics 4*, Oxford University Press, Oxford.

Box, G.E.P. and Tiao, G.C. (1973) *Bayesian Inference in Statistical Analysis,* Addison-Wesley, Reading, Massachusetts.

Brown, L.D., Cai, T.T. and DasGupta, A. (2001) 'Interval estimation for a binomial proportion', *Statistical Science*, Vol. 16, No. 2, pp.101–133.

Clarotti, C.A. (1993) 'Decision making: drawbacks of the classical statistics approach', in R.E. Barlow, C.A. Clarotti and F. Spizzichino (Eds.): *Reliability and Decision Making*, Chapman & Hall, New York, pp.1–12.

Coolen, F.P.A. (1994) 'On Bernoulli experiments with imprecise prior probabilities', *The Statistician*, Vol. 43, No. 1, pp.155–167.

de Finetti, B. (1974) *Theory of Probability Volume 1: A Critical Introductory Treatment*, Wiley, New York.

Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D.S. and Sentz, K. (2002) *Constructing Probability Boxes and Dempster-Shafer Structures,* Report Num. SAND2002-4015, Sandia National Laboratories, Albuquerque, New Mexico.

Fougere, P. (1990) *Maximum Entropy and Bayesian Methods*, Kluwer Academic Publishers, Dordrecht.

Groen, F.J. and Mosleh, A. (2005) 'Foundations of probabilistic inference with uncertain evidence' *International Journal of Approximate Reasoning*, Vol. 39, No. 1, pp.49–83.

Hogg, R.V. and Tanis, E.A. (2001) *Probability and Statistical Inference,* 6th ed., Prentice Hall, Upper Saddle River, NJ.

Insua, D.R. and Ruggeri, F. (2000) *Robust Bayesian Analysis,* Springer, New York.

Jeffreys, H. (1961) *Theory of Probability*, 3rd ed., Oxford University Press, London.

Lindley, D.V. (1982) 'Subjectivist view of decision making', *European Journal of Operational Research*, Vol. 9, No. 3, pp.213.

Lindley, D.V. and Phillips, L.D. (1976) 'Inference for a Bernoulli process (a Bayesian view)', *American Statistician*, Vol. 30, pp.112–119.

Maritz, J.S. and Lewin, T. (1989) *Empirical Bayes Methods*, 2nd ed., Chapman and Hall, London.

Pericchi, L.R. and Walley, P. (1991) 'Robust Bayesian credible intervals and prior ignorance', *International Statistical Review/Revue Internationale de Statistique*, Vol. 59, No. 1, pp.1–23.

Rothenberg, T.J. (1974) 'The Bayesian approach and alternatives in econometrics-II', in S.E. Fienberg and A. Zellner (Eds.): *Studies in Bayesian Econometrics and Statistics: In Honor of Leonard J. Savage*, North-Holland Publishing Company and American Elsevier Publishing Company, New York, pp.55–67.

Savage, L.J. (1971) 'Elicitation of personal probabilities and expectations', *Journal of the American Statistical Association*, Vol. 66, pp.783–801.

Savage, L.J. (1972) *The Foundation of Statistics*, Dover Publications, New York.

Utkin, L.V. (2004a) 'Interval reliability of typical systems with partially known probabilities', *European Journal of Operational Research*, Vol. 153, No. 3, pp.790–802.

Utkin, L.V. (2004b) 'Reliability models of M-out-of-N systems under incomplete information', *Computers and Operations Research*, Vol. 31, No. 10, pp.1681–1702.

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

Walley, P. (1996) 'Inferences from multinomial data: learning about a bag of marbles', *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol. 58, No. 1, pp.3–57.

Walley, P., Gurrin, L. and Burton, P. (1996) 'Analysis of clinical data using imprecise prior probabilities', *The Statistician*, Vol. 45, No. 4, pp.457–485.

Weber, M. (1987) 'Decision making with incomplete information', *European Journal of Operational Research*, Vol. 28, No. 1, pp.44.

Zellner, A. (1974) 'The Bayesian approach and alternatives in econometrics-I', in S.E. Fienberg and A. Zellner (Eds.): *Studies in Bayesian Econometrics and Statistics: in Honor of Leonard J. Savage*, North-Holland Publishing Company and American Elsevier Publishing Company, New York, pp.39–54.

Zellner, A. (1977) 'Maximal data information prior distributions', in A. Aykac and C. Brumat (Eds.): *New Methods in the Applications of Bayesian Methods*, North-Holland, Amsterdam.

Zhu, M. and Lu, A.Y. (2004) 'The counter-intuitive non-informative prior for the Bernoulli family', *Journal of Statistics Education*, Vol. 12.