Validation and Adjustment of Prior and Data for Bayesian Reliability Analysis in Engineering Design

Zhaofeng Huang
e-mail: zhaofeng@usc.edu

Yan Jin
e-mail: yjin@usc.edu

Department of Aerospace and Mechanical Engineering,
University of Southern California
Los Angels, CA 90089

Bayesian reliability analysis (BRA) technique has been actively used in reliability assessment for engineered systems. However, there are two key controversies surrounding the BRA: the reasonableness of the prior and the consistency among all data sets. These issues have been debated in Bayesian analysis for many years. As we observed, they have not been resolved satisfactorily. These controversies have seriously hindered the applications of BRA as a useful reliability analysis tool to support engineering design. In this paper, a Bayesian reliability analysis methodology with a prior and data validation and adjustment scheme (PDVAS) is developed to address these issues. As the part of the PDVAS development, a consistency measure is first defined that judges the level of consistency among all data sets including the prior. The consistency measure is then used to adjust either the prior or the data or both to the extent that the prior and the data are statistically consistent. This prior and data validation and adjustment scheme is developed for Binomial sampling with Beta prior, called Beta-Binomial Bayesian model. The properties of the scheme are presented and discussed that provides some insights of PDVAS. Various forms of the adjustment formulas are shown, and a selection framework of a specific formula, based on engineering design and analysis knowledge, is established. Several illustrative examples are presented, which show the reasonableness, effectiveness, and usefulness of PDVAS. General discussion of the scheme is offered to enhance the Bayesian reliability analysis in engineering design for reliability assessment.
[DOI: 10.1115/1.4003841]

Keywords: reliability, Bayesian reliability analysis (BRA), prior distribution, posterior distribution, prior and data validation and adjustment scheme (PDVAS)

1 Introduction

Bayesian reliability analysis (BRA) technique has been actively used in reliability assessment [1–4]. BRA follows a traditional Bayesian approach, which assumes a sampling distribution for the data being analyzed, assigns a probability distribution, called prior distribution, to the sampling distribution parameters, then collects test data to form the likelihood function, and updates the prior to the posterior distribution with the Bayes' formula that aggregates the prior with the likelihood function [3,4]. It has been long known that the reliability result obtained through BRA is very sensitive to the assignment of the prior. The criticisms on the prior assignment, and prior and data inconsistency have not stopped since the Bayesian technique was introduced, as evidenced by many literature articles [5,6]. Numerous researches have been under way, which have addressed the initial prior assignment. Several objective and noninformative prior generation methods have been developed [7,8]. Examples of the approaches include maximum entropy method [9], reference analysis technique [10], and frequentist matching method [11–13].

After an initial prior is determined, the Bayesian then collects test, experiment, or analysis data and derives the posterior distribution using the Bayesian formula. This process is called Bayesian updating. The Bayesian updating can be conducted repeatedly as multiple data sets become available. The Bayesian updating basically takes the prior and the experiment, test, or analysis data and aggregates them in a weighted average manner. All this is done procedurally, usually without considering how contradicting and inconsistent the prior is with the data as well as among data sets by the Bayesian analysis itself.

Figure 1 describes the traditional Bayesian analysis flow. After a sampling distribution is specified in step 1 and an initial prior is determined in step 2, the Bayesian then collects experiment data to form likelihood function in step 3 and derives the posterior distribution in step 4 using the Bayesian formula given by Eq. (1). If a repeated Bayesian updating is conducted, the posterior distribution derived from step 4 loops back in step 4R as a new prior input to step 3, and the Bayesian formula [Eq. (1)] is used again to aggregate this new prior with new likelihood data to arrive at an updated posterior distribution. The posterior distribution is then used as a statistical inference tool in engineering applications as indicated in step 5.

$$f_{\Theta|X}(\theta|x) = \frac{f_{\Theta,X}(x,\theta)}{f_X(x)} = \frac{f_{X|\Theta}(x|\theta) f_\Theta(\theta)}{f_X(x)} \tag{1}$$

In Eq. (1), $f_\Theta(\theta)$ is the prior density for the sampling distribution parameter Θ, $f_{X|\Theta}(x|\theta)$ is the likelihood function of the data x given a Θ value as θ, $f_{\Theta|X}(\theta|x)$ is the posterior density of the Θ, *and* $f_X(x)$ is the marginal density of the data.

If a prior distribution takes the same form of the distribution function as the posterior distribution, the prior is called conjugate prior. Mathematically, the prior, $f_\Theta(\theta)$ in a conjugate Bayesian model, has the same function form as the posterior, $f_{\Theta|X}(\theta|x)$. The

Contributed by the Design Automation Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received May 14, 2010; final manuscript received February 19, 2011; published online June 6, 2011. Assoc. Editor: Zissimos P. Mourelatos.

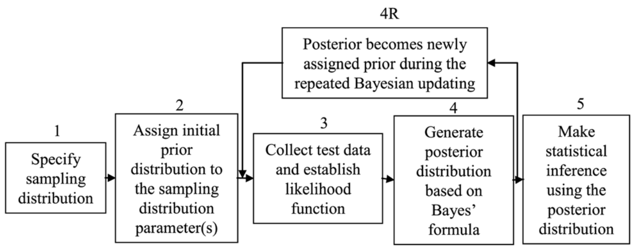

Fig. 1 Traditional Bayesian analysis flow

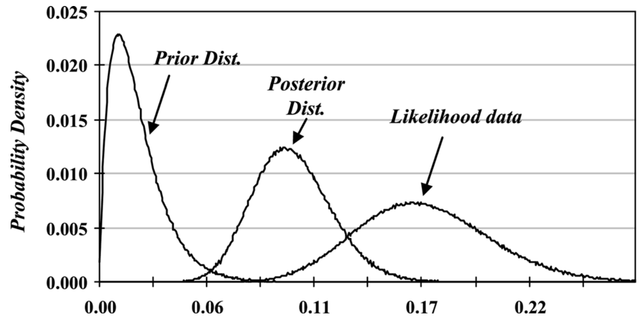

Fig. 2 Bayesian result as a weighted average of prior and data set

examples of conjugate prior include Beta prior with Binomial sampling, called Beta-Binomial Bayesian model, and normal prior for the mean with normal sampling. For more conjugate prior discussions, see Refs. [3,4]. A conjugate Bayesian model has several nice features. First, the derivation of the posterior is much easier. Usually, a closed form formula exists that aggregates the prior and the data to obtain the posterior. For example, for the Beta-Binomial model with Binomial parameter p, if we assume p is subject to Beta(α,β), and we obtain test data of F failures out of N trials, the posterior distribution of p is then given by Beta$(\alpha+F,\beta+N-F)$. Second, for a repeated Bayesian updating with k (>1) data sets available, the final Bayesian updated posterior is independent of the sequence of the updating taken on the k data sets.

During an engineering design, especially at the conceptual design stage, reliability analysis relies on various data sources, including historical failure data from similar parts or systems, expert opinions, engineering modeling and simulations, and prototype or laboratory test results. Huang and Jin [14] provided a comprehensive survey of reliability prediction data sources as potential data inputs to BRA. Grantham Lough et al. [15,16] correlated the historical failure data to the functions of system being studied during the functional design to assist risk assessment. Wang and Jin [17] developed a functional design approach, which utilized a Bayesian network technique with uniform distributions as inputs to assess an individual function's influence on the system failure probability. Wang et al. [1] applied a repeated Bayesian updating technique to a reliability assessment during a product design and development cycle using evolving, insufficient, and subjective data sets, which included a customer survey, response surface physics model results, and clinic trial test data. Huang and Jin [18] extended the traditional reliability stress–strength interference theory to the conceptual design with the combination of a team survey, historical similar function design data, physics bounds of the design to define conceptual stress, and conceptual strength for reliability quantification. However, none of the above work has addressed the data inconsistency and data contradiction issue. For some BRA applications, an obvious data inconsistency may appear between the prior and the test data or among data sets from various data sources. One extreme example is that the prior states that the Binomial sampling parameter p is equal to 0.01 with probability 1, but the experiment data show five failures out of ten trials. It is obvious that the data set "five out of ten" is very unlikely to come from Binomial sampling with $p=0.01$. A less extreme example is that the prior states that the Binomial parameter p is subject to the Beta(1,100) prior while the experiment data shows one failure out of ten trials. For data samples such as this, what Bayesian analysis produces is a weighted average of the prior and the data set as the posterior, as illustrated in Fig. 2, though the likelihood data and the prior hardly overlap as shown in the figure. A data inconsistency example is when one data set has 1 failure and 2 successes and the other has 1 failure and 50 successes. It is not likely that these two data sets are from the same sampling distribution. Therefore, it is seriously questionable whether these two data sets are combinable for the Bayesian analysis for the inference purpose. The aforementioned prior generation methods, namely maximum entropy method, reference prior analysis, and frequentist matching, only address the generation of the initial prior. As we surveyed Bayesian analysis literature, we observed that there is little active research that addresses the prior and data inconsistency, and the validation of the prior and the data during a repeated Bayesian updating process. The research in Ref. [19] is probably one of the few that discuss the data conflict.

The above discussion brings an obvious need for a prior and data validation and adjustment. If the prior is contradicting with the data, the analyst has three choices: (1) accept the prior and the data and perform Bayesian update as usual; (2) reject the prior and the data; or (3) do something about the prior and the data to continue the Bayesian process in a reasonable manner. For choice 1, the analyst may lead himself or herself to a misleading inference therefore inadequate design decision generated from the Bayesian method. For choice 2, the analyst will have no data to perform Bayesian analysis. For choice 3, the literature survey indicates there is no existing theory and method for doing so. The objective of this paper is to provide such a method to help address choice 3. In Sec. 2, we present a modified Bayesian updating process with an added step that validates the consistency among all data set and adjusts the prior and the data accordingly if inconsistencies arise. We provide the mathematical formulas for the prior and data validation and adjustment scheme (PDVAS) for the Beta-Binomial Bayesian model. In Sec. 3, we present and discuss several properties of PDVAS, which provide some insights of PDVAS. In Sec. 4, we discuss the selection of prior and data adjustment formulas based on the engineering knowledge and data available. In Sec. 5, we present various examples to illustrate the PDVAS applications. We then summarize the paper in Sec. 6, discuss the limitations of PDVAS, and future research possibilities.

2 A PDVAS

2.1 Motivation. As we have pointed out, the traditional Bayesian analysis has some shortcomings in dealing with data contradiction and inconsistency. We believe that one of the key applications of the Bayesian posterior is for inference modeling that predicts the sampling distribution behavior to assist design decisions. Therefore, it is very important for this inference model to be valid. Fundamentally, our motivation is how we can go beyond the simple data aggregation as the Bayesian procedure defines, evaluate data trusting and worthiness, and adjust the data when evidence of inconsistency arises with the engineering design knowledge at hand. The adjustment also needs to balance retaining the knowledge in the data and reducing the data inconsistency. This leads to the proposal of the modified Bayesian analysis flow.

2.2 Modified Bayesian Analysis Flow. Figure 3 presents our proposal of the modified Bayesian analysis process with an added prior and data validation and adjustment step, which is step 4a. All other steps are the same as the original Bayesian analysis flow as in

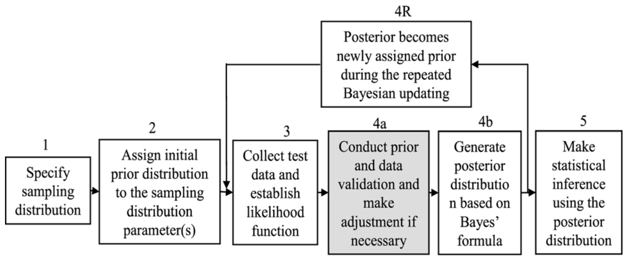

Fig. 3 Modified Bayesian analysis flow with addition of prior and data validation and adjustment step

Fig. 1. We discuss the details of the step 4a in the rest of Sec. 2, Sec. 3, and Sec. 4.

2.3 Prior and Data Validation and Adjustment Formulas. For step 4a of Fig. 3 for the Beta-Binomial model, we have a Binomial distribution as the sampling distribution with the parameter p as follows

$$p(X = x|\theta = p) = \binom{N}{x} p^x (1-p)^{(N-x)} \tag{2}$$

$x = 0, 1, 2, \ldots, N$. p is assigned a Beta distribution with Beta parameters α and β. Its density is

$$f(p) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{\mathrm{Beta}(\alpha, \beta)} \tag{3}$$

In Eq. (3), $0 < p < 1$, $\alpha > 0$ and $\beta > 0$. Beta(α, β) is a complete Beta function given by $\int_0^1 p^{\alpha-1}(1-p)^{\beta-1}dp$. The mean and the standard deviation of the Binomial sampling distribution are

$$\mu_{\mathrm{Bino}} = Np \tag{4}$$

and

$$\sqrt{Np(1-p)} \tag{5}$$

respectively.

The mean and the standard deviation of the Beta(α,β) are

$$\mu_{\mathrm{Beta}} = \frac{\alpha}{\alpha + \beta} \tag{6}$$

and

$$\sigma_{\mathrm{Beta}} = \frac{\sqrt{\alpha\beta}}{(\alpha+\beta)\sqrt{\alpha+\beta+1}} \tag{7}$$

respectively.

For a general Bayesian analysis, we have the following data: Prior distribution is given by Beta(α,β), and k data sets ($k \geq 1$) are given by (F_1,S_1), (F_2,S_2), …, (F_k,S_k). For the convenience of the notation, we name $F_0 \equiv \alpha$ and $S_0 \equiv \beta$. So we have Bayesian data sets: (F_0,S_0), (F_1,S_1), (F_2,S_2), …, (F_k,S_k). Recall F_i and S_i represent number of failures and number of successes in the ith data set, respectively. With the traditional Bayesian process, we obtain the Beta posterior, named Beta(α^*, β^*). Using Eq. (1), we get

$$\alpha^* = \sum_{i=0}^{i=k} F_i \tag{8}$$

and

$$\beta^* = \sum_{i=0}^{i=k} S_i \tag{9}$$

Remember that α^* and β^* are calculated without evaluating the data inconsistency and contradiction among (F_0,S_0), (F_1,S_1), (F_2,S_2), …, (F_k,S_k). Now the question is how we assess the data consistency and provide a measure of it. We define a consistency statistic with a probability associated with a χ^2statistic as follows

$$\chi_C^2 = \left[\sum_{i=0}^{i=k}(F_i + S_i)\right]\left(\frac{\sum_{i=0}^{i=k}\frac{F_i^2}{F_i+S_i}}{\sum_{j=0}^{j=k}F_j} + \frac{\sum_{i=0}^{i=k}\frac{S_i^2}{F_i+S_i}}{\sum_{j=0}^{j=k}S_j} - 1\right) \tag{10}$$

The χ_C^2 formula is originated from the χ^2 statistic of traditional hypothesis testing for $2 \times (k+1)$ contingency table [20,21]. χ_C^2 has a degree of freedom k. Therefore, the consistency for the data sets (F_0,S_0), (F_1,S_1), (F_2,S_2), …, (F_k,S_k) is defined as

$$\mathrm{C} \equiv \mathrm{Consistency} = P(\chi^2 > \chi_C^2) \tag{11}$$

where χ_C^2 in Eq. (10) is a χ^2 random variable with a degree of freedom k.

The motivation of the consistency formulas defined by Eqs. (10) and (11) is that $F_0 + S_0$ ($\equiv \alpha + \beta$) of the prior, per the Bayesian process, represents the prior's sample size, embedded in the prior knowledge in processing the posterior distribution [4]. F_0 ($\equiv \alpha$) and S_0 ($\equiv \beta$) approximately represent the number of failures and successes, respectively, afforded by the prior. The mean of the prior is $F_0/(F_0 + S_0)$. If the data are consistent, all the data means, namely $F_1/(F_1+S_1)$, $F_2/(F_2+S_2)$, …, $F_k/(F_k+S_k)$ should not be very far away from $F_0/(F_0+S_0)$. As matter of fact, the data sets of (F_1,S_1), (F_2,S_2), …, (F_k,S_k), according to a Bayesian assumption, all should be generated from the Binomial sampling with the parameter p, which is subject to a Beta prior with mean $F_0/(F_0+S_0)$. Therefore, when the failure fraction, $F_i/(F_i+S_i)$, $i = 1, 2, \ldots, k$, of the data sets are equal to or close to the failure fraction of the prior, $F_0/(F_0+S_0)$, we, to the maximum extent, believe that the data are consistent, and there is no contradiction between the prior and the data sets. Conversely, if one or more of the $F_i/(F_i+S_i)$ are drastically different from $F_0/(F_0+S_0)$, or some $F_i/(F_i+S_i)$ is drastically different from $F_j/(F_j+S_j)$ $(i \neq j)$, we have a reason to think that the data are not consistent and the assumption that all data sets (F_1,S_1), (F_2,S_2), …, (F_k,S_k) are generated from the same Binomial sampling with p as the parameter is not adequate. Mathematically, when $F_0/(F_0+S_0) = F_1/(F_1+S_1) = \cdots = F_k/(F_k+S_k)$, Eq. (10) leads to $\chi_C^2 = 0$. Therefore, the consistency C by Eq. (11) = 1. When some $F_i/(F_i+S_i)$ is very much different from some $F_j/(F_j+S_j)$ $(i \neq j)$, χ_C^2 in Eq. (10) becomes very big. Therefore, the consistency measure C by Eq. (11) is nearly zero.

The consistency measure C, calculated by Eqs. (10) and (11), represents the prior and data validation result. C is a value between 0 and 1. When $C = 1$ or close to 1, we believe that the data sets, (F_0,S_0), (F_1,S_1), (F_2,S_2), …, (F_k,S_k), are completely consistent or nearly consistent. Thereby we will implement the Bayesian updating unconditionally as we do in the traditional Bayesian process. When $C = 0$ or close to 0, we believe that the data sets, (F_0,S_0), (F_1,S_1), (F_2,S_2), …, (F_k,S_k), are completely inconsistent or nearly inconsistent. Thereby we will seriously challenge the assumption that all data sets, (F_1,S_1), (F_2,S_2), …, (F_k,S_k), are generated from the same Binomial sampling with p as the parameter.

Now the question is that what do we do when the consistency measure C by Eq. (11) is between 0 and 1? We propose the following data adjustment algorithm as a part of the modified Bayesian updating process. This is the substantiation of step 4a of the

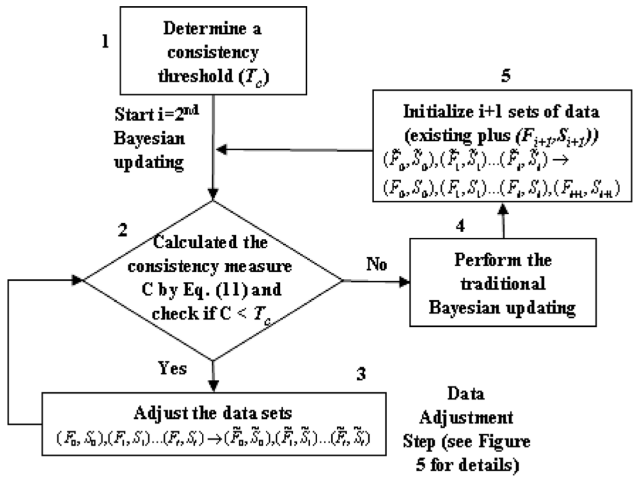

Fig. 4 Bayesian updating prior and data adjustment algorithm

process presented in Fig. 3. Figure 4 shows the algorithm and Fig. 5 presents the data adjustment step details, where $\chi^2_{C,m}$ in Fig. 5 is the χ^2 statistic by Eq. (10) using data sets (F_0,S_0), (F_1,S_1), …, (F_i,S_i), but excluding the data set (F_m,S_m). Therefore, $\chi^2_{C,m}$ has a degree of freedom of $i-1$. C_m is the consistency measure among data sets (F_0,S_0), (F_1,S_1), …, (F_i,S_i), excluding (F_m,S_m), therefore, $C_m = P(\chi^2 > \chi^2_{C,m})$. The data adjustment step (step 3) in Fig. 4 and the adjustment details (steps 3.1–3.3) in Fig. 5 basically detect the inconsistency sources and adjust the data sets to make them consistent. The targeted data set $[(F_v,S_v)$ in steps 3.2 and 3.3] for the data adjustment is the one with the biggest consistency value after it is excluded from the consistency calculation; therefore, it is identified as the source of inconsistency. It is noted in Fig. 4 that the data validation and adjustment starts from the second Bayesian updating since for the first update, we only have the initial prior and the first data set, which do not provide us any direction on how we can adjust the data. Steps 2 and 3 in Fig. 4 are in an iteration loop. Therefore, there is a convergence issue. We will discuss this in Sec. 3 . Step 3.3 is to adjust the data set (F_v,S_v), when it is found that it is the source of the inconsistency.

Next we discuss the data adjustment formula. We first define a data adjustment as a mapping from the consistency measure to the data adjustment score, denoted as S_{DA}.

$$S_{\mathrm{DA}} = f(\mathrm{Consistency}) \tag{12}$$

The exact form of f function in Eq. (12) will be determined in Sec. 4. The S_{DA} also takes values in the range of [0, 1]. We then apply the S_{DA} to the data set (F_v,S_v) in the following discounting manner

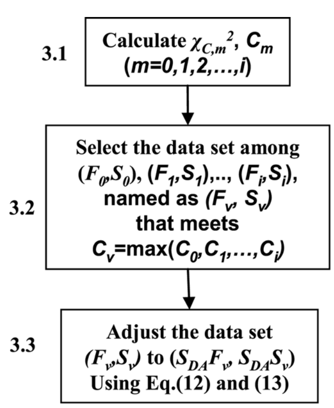

Fig. 5 Details of the data adjustment step

to obtain the discounted values of F_v and S_v, namely $\tilde{F}_v$ and $\tilde{S}_v$, respectively.

$$\tilde{F}_v = S_{DA}F_v \text{ and } \tilde{S}_v = S_{DA}S_v \tag{13}$$

So far, we have not defined the detailed f functional forms of S_{DA} yet. But we know S_{DA} should satisfy the following conditions

$$S_{\mathrm{DA}}(0) = f(0) = 0 \tag{14}$$

and

$$S_{\mathrm{DA}}(1) = f(1) = 1 \tag{15}$$

and

$$\mathrm{S}_{\mathrm{DA}}(\mathrm{consistency}_2) \geq \mathrm{S}_{\mathrm{DA}}(\mathrm{consistency}_1) \text{ when}$$

$$\mathrm{Consistency}_2 > \mathrm{Consistency}_1 (\text{monotone increase}) \tag{16}$$

Equations (13)–(16) basically state that when the consistency measure is zero, we completely ignore the data set (F_v,S_v). When the consistency measure is 1, we completely accept the data set (F_v,S_v), which is what the traditional Bayesian updating process does. The bigger the consistency measure, the less we discount the data. Recall that step 1 of Fig. 4 asks for defining a threshold value (T_c) of the consistency measure. The role and the interpretation of this threshold are similar to the concept of the significance level in a traditional statistical hypothesis testing. But the rejection criterion in hypothesis testing is a step function, that is, when the probability value (p-value) of observing a certain data sample is below the significance level, the null hypotheses will be rejected. In the context of PDVAS applications, we extend this step function to a set of smooth curves, which can incorporate engineering design and analysis knowledge for data adjustment. In this section, we present general forms of S_{DA} curves first. In Sec. 4, we will provide recommendations as to how a specific curve can be selected based on available engineering knowledge and data.

Figure 6 depicts various possibilities of S_{DA} curves with $T_c = 0.05$. For all the curves in the figure, $S_{\mathrm{DA}} = 1$ when the consistency measure ≥ 0.05. The curve A in the figure is very close to taking all S_{DA} values of 1 for any consistency measure. So if adopted, it defaults to the traditional Bayesian updating process. The curve G takes almost all S_{DA} values of 0 for the consistency

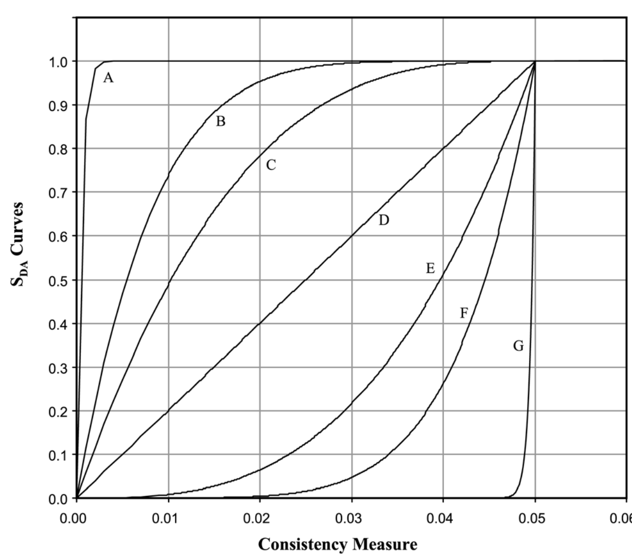

Fig. 6 Potential candidate functions for S_{DA}

measures ≤ 0.05, which leads to the nearly complete rejection of the data set (F_v,S_v), similar to the situation of a traditional hypothesis testing with a significance level of 5% of rejecting a null hypothesis. Section 4 will discuss the details how to select a T_c and S_{DA} curve.

3 Properties of PDVAS

PROPOSITION 1. *For the data set (F_v,S_v) being adjusted in Eq. (13), PDVAS does not change the mean of the data set, but increases its standard deviation σ.*

The proof of Proposition 1 is given in Appendix. Proposition 1 reflects an important PDVAS strategy, that is, to maintain the mean of the adjusted data set but increase the spread of the data distribution. In other words, PDVAS takes the face value embedded in the data regarding to the knowledge of central tendency but discounts the value of the data set by increasing its variance. The rationale of doing so is that when the prior or the data are collected, usually the uncertainty associated with the variability is much bigger than the one associated with the central tendency. Therefore, the variability of the data is more doubtful than the mean. PDVAS focuses on the adjustment of the variability to achieve the data consistency.

PROPOSITION 2. *At the ith Bayesian updating with the i sets of data (F_0,S_0), (F_1,S_1), …, (F_i,S_i) available, if $[F_v/(F_v+S_v)] < [\sum_{j=0,j\neq v}^{j=i} F_j/(\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j)]$, where $0 \leq v \leq i$, and (F_v,S_v) is one of the i data sets, we then have*

$$\begin{aligned}\frac{F_v}{F_v+S_v} &< \frac{F_v+\sum_{j=0,j\neq v}^{j=i} F_j}{F_v+S_v+\sum_{j=0,j\neq v}^{j=i} F_j+\sum_{j=0,j\neq v}^{j=i} S_j} \\ &\leq \frac{\tilde{F}_v+\sum_{j=0,j\neq v}^{j=i} F_j}{\tilde{F}_v+\tilde{S}_v+\sum_{j=0,j\neq v}^{j=i} F_j+\sum_{j=0,j\neq v}^{j=i} S_j} \\ &\equiv \frac{S_{\mathrm{DA}}F_v+\sum_{j=0,j\neq v}^{j=i} F_j}{S_{\mathrm{DA}}F_v+S_{\mathrm{DA}}S_v+\sum_{j=0,j\neq v}^{j=i} F_j+\sum_{j=0,j\neq v}^{j=i} S_j} \\ &< \frac{\sum_{j=0,j\neq v}^{j=i} F_j}{\sum_{j=0,j\neq v}^{j=i} F_j+\sum_{j=0,j\neq v}^{j=i} S_j}\end{aligned} \quad (17)$$

Note if $[F_v/(F_v+S_v)] > [\sum_{j=0,j\neq v}^{j=i} F_j/(\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j)]$, Eq. (17) still holds but with all inequality signs reversed.

The proof of Proposition 2 is given in Appendix. Proposition 2 indicates that when PDVAS detects a data inconsistency with the data set (F_v,S_v) as the source of the inconsistency, it adjusts (F_v,S_v) to reduce its weight in the posterior distribution such that the posterior mean is moving away from the mean of the data set (F_v,S_v) toward the mean of $\left(\sum_{j=0,j\neq v}^{j=i} F_j, \sum_{j=0,j\neq v}^{j=i} S_j\right)$. Under the extreme case that S_{DA} is zero or close to zero, (F_v,S_v) is completely or nearly completely ignored, and the posterior mean is the mean of $\left(\sum_{j=0,j\neq v}^{j=i} F_j, \sum_{j=0,j\neq v}^{j=i} S_j\right)$.

PROPOSITION 3. *For the data sets (F_0,S_0), (F_1,S_1), …, (F_i,S_i), there always exists a set of S_{DA} values such that steps 2 and 3 of Fig. 4 will converge.*

The proof of Proposition 3 is given in Appendix. Proposition 3 confirms that there is always a data adjustment solution for PDVAS when a data inconsistency is detected and that steps 2 and 3 of Fig. 4 will not fall into a dead loop. In reality, steps 2 and 3 often take as few as one or two iterations as the illustrative examples will show in Sec. 5.

PROPOSITION 4. *For no-failure situation, that is, in (F_0,S_0), (F_1,S_1), …, (F_k,S_k), all F_i = 0 except F_0, χ_C^2 by Eq. (10), after substituting F_0 by $\tilde{F}_0 = S_{\mathrm{DA}}F_0$ and S_0 by $\tilde{S}_0 = S_{\mathrm{DA}}S_0$, is a monotone increase function of S_{DA}, $\chi_C^2 \geq F_0/(F_0+S_0)\sum_{i=1}^{i=k} S_i$ and $\chi_C^2 \to F_0/(F_0+S_0)\sum_{i=1}^{i=k} S_i$ when $S_{\mathrm{DA}} \to 0$.*

The proof of Proposition 4 is given in Appendix. Proposition 4 addresses a special data situation, in which there is no failure in the data. Therefore, there is no evidence of data inconsistency among (F_1,S_1), …, (F_k,S_k), since all data means $= F_j/(F_j+S_j) = 0$ ($j = 1,2,\ldots,k$). For the same reason, the failure ratio of the data cannot be compared with the prior data. If we attempt to adjust (F_0,S_0), the proposition states that the χ_C^2 has a lower bound limit. Therefore, the data adjustment does not make χ_C^2 go to zero even $S_{\mathrm{DA}} = 0$. Thereby the prior and data validation and adjustment cannot be performed meaningfully. PDVAS will not make any data adjustment under "no-failure" situation. PDVAS will default to the traditional Bayesian result under no-failure situation.

PROPOSITION 5. *The final posterior in a repeated Bayesian updating with multidata sets using PDVAS is update sequence dependent.*

Proposition 5 indicates that PDVAS loses the nice feature of the updating sequence independence possessed by the conjugate prior Bayesian process. To prove the proposition, we only need to provide a counterexample illustrating the posterior distribution from PDVAS will vary from different updating sequences in a multiple Bayesian updating process. This example will be given in Sec. 5 (example 4).

4 Selection of Data Adjustment Formulas

In Sec. 2, we discussed some general forms of the formulas being applied to data adjustment when PDVAS detects a data inconsistency. Proposition 3 in Sec. 3 indicates that we can always find a data adjustment solution to make the data consistent. However, the intention of PDVAS is to discount the data as less as possible, especially being careful not to throw out good data. Therefore, the selection of S_{DA} value and its function is essentially important. If S_{DA} is too small (close to 0), we tend to disregard good data, which defeats the original intention of the Bayesian analysis. If S_{DA} is too big (close to 1), we accept the data blindly, which may ignore possible data inconsistency and lead to a misleading inference. In this section, we first provide a family of S_{DA} functions and then recommend some criteria for selecting a specific one for general engineering applications.

4.1 A Family of S_{DA} Curves.

We select the incomplete Beta function within the range [0, T_c] to represent the S_{DA} functions. Here T_c is the threshold value defined in step 1 of Fig. 4. When the consistency measure C, calculated by Eq. (11), meets $C \geq T_c$, we accept all the data as is without any adjustment. When the C values fall within (0, T_c), we calculate S_{DA} using an incomplete Beta function as follows

$$S_{\mathrm{DA}} = f(C) = \begin{cases} 0 & \text{if } C \leq 0 \\ \frac{1}{T_c}\int_0^{C/T_c} \frac{\left(\frac{u}{T_c}\right)^{m-1}\left(1-\frac{u}{T_c}\right)^{n-1}}{\mathrm{Beta}(m,n)}du & \text{if } 0 < C < T_c \\ 1 & \text{if } T_c \leq C \end{cases} \quad (18)$$

Beta(m,n) in Eq. (18) is a complete Beta function given by $\int_0^1 u^{m-1}(1-u)^{n-1}du$. The rationale of selecting the incomplete Beta function for S_{DA} is its versatility in its shapes and the easy interpretation of the parameters m and n in Eq. (18) related to the engineering knowledge. To illustrate the merit of the Beta function, we present an example of a set of Beta curves in Fig. 7 with $T_c = 0.05$. All the curves have the saddle points (the second

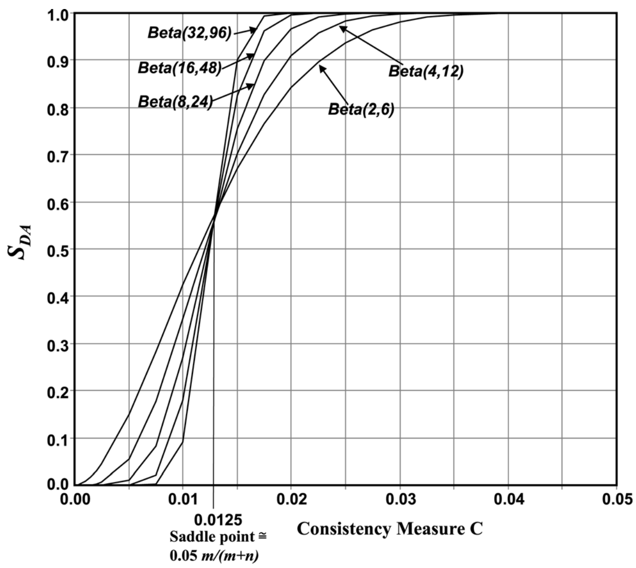

Fig. 7 A family of beta curves as potential S_{DA} functions

derivative = 0) near $m/(m+n)0.05 = 0.0125$ on the x-axis. As m and n get bigger, the curves get steeper around the saddle point, approaching the situation either PDVAS completely rejects the data or completely accepts the data depending on whether C is bigger than or smaller than 0.0125. It is recognized from a probability theory that the consistency measure C by itself is a random variable with Uniform(0,1) as its distribution, since it is a value from the accumulative probability function of the χ^2 random variable by Eq. (11). Therefore, when C falls within (0, 0.05), it uniformly takes a value between 0 and 0.05. The family of the curves in Fig. 7 basically state that there is on the average one out of four or 25% (= 0.0125/0.05) chance to significantly reject the data. At the extreme case that we are certainly one out of four times that the data are bad, we completely reject the data when $C < 0.0125$. However, in reality we make the statement "one out of four" with uncertainty, so we only reject the data partially, which is quantified by the S_{DA} values on various curves. Now the question is how we select T_c, m, and n for the S_{DA} function for a PDVAS implementation?

4.2 Selection of T_c, m, and n

4.2.1 T_c Selection. T_c represents a screening criterion, which is conceptually similar to the significance level or α value in a traditional statistical hypothesis testing. The α value is always set to be small (≤ 0.1) to avoid unacceptable false positive in the hypothesis testing. From an engineering design point view, when we use Bayesian analysis to aggregate the data to predict reliability, data often come from various sources; therefore, big uncertainty can be expected. While we want data inconsistency to be detected, we do not want to go through the data adjustment step when the evidence of data inconsistency is not strong. Therefore, we recommend using T_c value of 0.05 as a standard value for the inconsistency screening. For the case that we want to closely mimic the traditional Bayesian analysis without concerning too much about data inconsistency, we can use a value of 0.01 or smaller. The T_c value of zero defaults PDVAS to the traditional Bayesian process. Thereby, we consider the traditional Bayesian process is just a special case of PDVAS.

4.2.2 m,n Selection. As we mentioned in Sec. 4.1, the ratio of $m/(m+n)$ approximately represents our knowledge and judgment based on engineering knowledge on the possible percentage of inconsistent data. While keeping $m/(m+n)$ unchanged, the more we can pinpoint the source of data inconsistency with certainty, the bigger the m and n we can assign to. As one of our early research papers surveyed [14], reliability data for a Bayesian analysis can come from the following four sources: (1) statistical frequency method (SF); (2) similarity and comparative assessment (SCA); (3) physics based modeling and simulation (PBMS); and (4) expert elicitation (EE). Usually, SF data, which are generated from field or laboratory simulated operating environment, have the smallest modeling uncertainty for the reason that fitted statistical inference model partially addresses the sampling modeling uncertainty. Uncertainty in other three data sources (SCA, PBMS, and EE) can vary widely but data from EE probably have the biggest modeling uncertainty since they are purely based on expert judgment. With the above assessment, we use Table 1 as the selection frame work for m and n, which also serves as an illustrative example.

Column (1) in Table 1 classifies the data source categories. We use our research result in Ref. [14] to divide all possible data sources into four categories (SF, SCA, PBMS, and EE). Users of PDVAS can create their own data source categories. Column (2) is the engineering judgment of the analyst based on their knowledge on the design and analysis for what the average percentage of the inconsistent data can be. Column (3) is to assess the uncertainty of column (2), which asks approximately how many times the analyst has experienced the data inconsistency instances in the past. By the Beta function definition represented by Eq. (18), the number of inconsistency data instances equates to the parameter m value. So column (4) = column (3). n in Eq. (18) represents the estimated number of good data instances the analyst experienced in the past. So, n value in column (5) is back-calculated using Eq. (19). Column (6) completely defines the S_{DA} Beta function used in Eq. (18). Figure 8 creates the S_{DA} Beta curves based on the data from Table 1.

$$\frac{m}{m+n} = column(2) \tag{19}$$

Figure 8 indicates that we rarely reject SF data in a dramatic manner shown in curve (1). This is in line with our recognition that SF data often is the most reliable data source. Comparing curve (2) [Beta(2,2)] and curve 3 [Beta(1,1)], both of them partially discount the data in a prorated fashion. Curve (3) is strictly liner. Curve (2) discounts the data less when $C > 0.025$ and discounts more when $C < 0.025$ than curve (3) does. This is because we are more confident on Beta(2,2) than on Beta(1,1), which, by derivation, indicates we have experienced more bad data instances in Beta(2,2) situation than in Beta(1,1). For curve (4) in the figure, we discount the majority of the data (more than 50%) when C falls below 0.035 (70% of 0.05).

Table 1 m and n selection framework with an example

(1) Datacategory	(2) Estimated percentage of inconsistent data (%)	(3) Number of analyst's experiences on bad data	(4) M	(5) n	(6) S_{DA} Beta function in Eq. (18)
SF	5	1	1	19	Beta(1,19)
SCA	50	2	2	2	Beta(2,2)
PBMS	50	1	1	1	Beta(1,1)
EE	70	3	3	1.29	Beta(3,1.29)

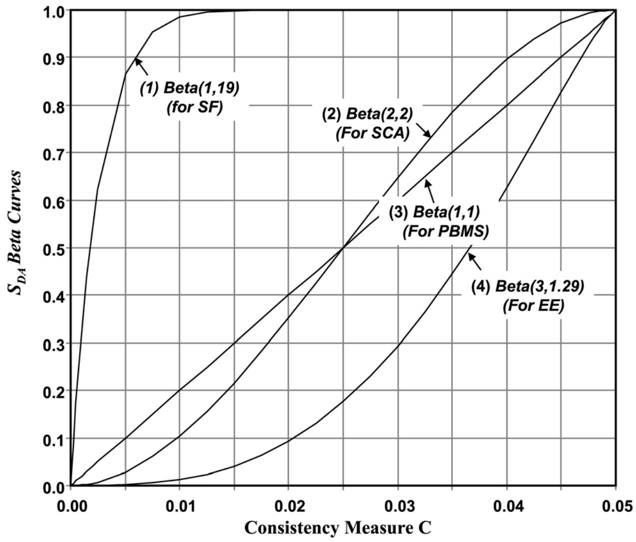

Fig. 8 S_{DA} curves for the data in Table 1

It is recognized that the data categories are created with subjectivity, the above approach is only considered to be a framework. Users of PDVAS can create their own data categories and assign m and n with their own knowledge and judgment. We recommend, for a quick analysis or for the situations that not much information is available about the data sources, the user use the linear curve (Beta(1,1)) as the S_{DA} function as Fig. 9 shows, which basically discounts the data in a linearly prorated manner when $C \leq T_c$ without assessing the fraction of possible bad data. In many of our PDVAS simulation runs, this approach is proven to be reasonable.

5 Case Examples

In this section, various examples are presented to illustrate PDVAS applications. The results from examples 1 to 4 are obtained through Monte Carlo simulations that assume certain Bayesian prior and data sampling distributions. Examples 1–3 show the effectiveness of PDVAS that detects data inconsistencies and adjusts the data accordingly. Example 4 shows the updating sequence dependency of PDVAS. Example 5 is to apply PDVAS to a rocket engine reliability analysis with various data sources.

Example 1. We assume the initial prior distribution is Beta(1,99), but the sampling distributions all come from Binomial with $p=0.10$. Intuitively, there is a data inconsistency between the prior and the data sets because the mean of prior is $1/(1+99)=0.01$ while sampling mean is $p=0.10$. We run 20 Bayesian updates following the process outlined in Fig. 3. The results are obtained through a computer simulation using the PDVAS algorithm presented in Figs. 4 and 5. We use $T_c=0.05$ and the linear curve for S_{DA} as in Fig. 9 to execute PDVAS. Figure 10 presents the mean comparison. Line (1) in the figure is the true sampling mean (normalized to $p=0.1$). Curve (2) is the Monte Carlo simulated average means from 10,000 Monte Carlo runs using PDVAS. Curve (3) is the average means of traditional Bayesian updated posteriors. This example illustrates that the PDVAS is effective in correcting the data for the case that an initial prior is too optimistic. The figure also indicates that all three curves tend to converge together eventually when i (number of updates) goes to infinity. However, PDVAS mean is much closer to the true sampling mean for the small number of updates. Therefore, PDVAS is very useful for the practical situations with small number of Bayesian updates conducted. Figure 11 presents the variance comparisons, which also indicates that PDVAS predicts the posterior variance much closer to the true sampling variance than the traditional Bayesian process does. This is because the traditional Bayesian blindly takes the initial prior Beta(1,99) as a part of the posterior updating, which significantly increases total sample size in the final posterior distribution, therefore underestimates the variance.

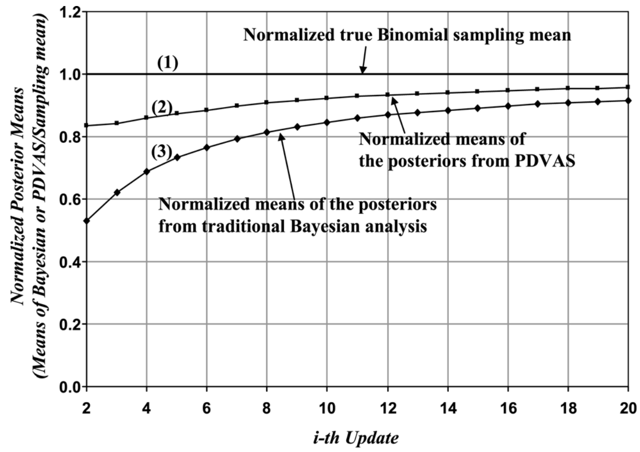

Fig. 10 Mean comparisons of PDVAS and traditional Bayesian for example 1

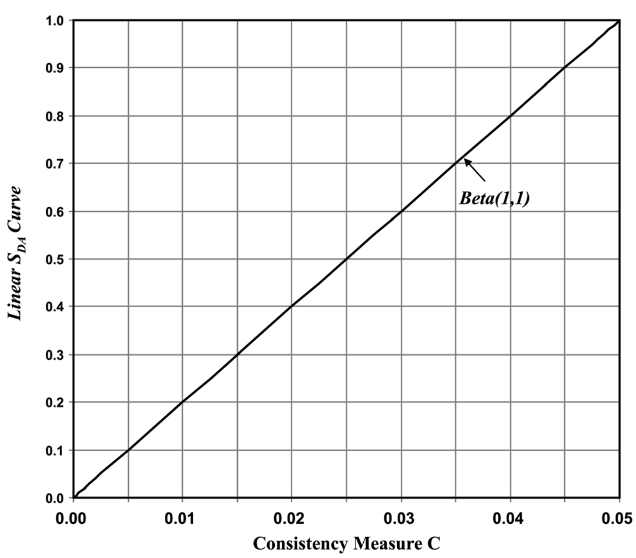

Fig. 9 Linear S_{DA} curve

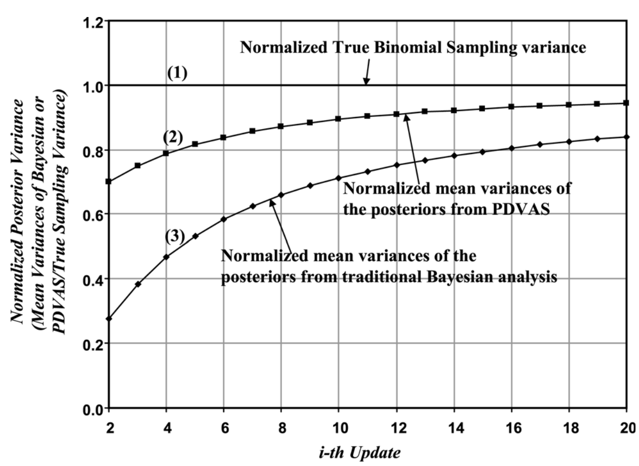

Fig. 11 Variance comparisons of PDVAS and traditional Bayesian for example 1

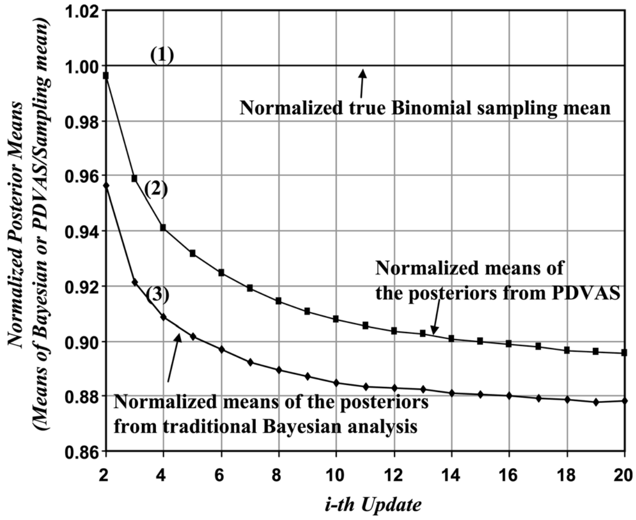

Fig. 12 Mean comparisons of PDVAS and traditional Bayesian for example 2

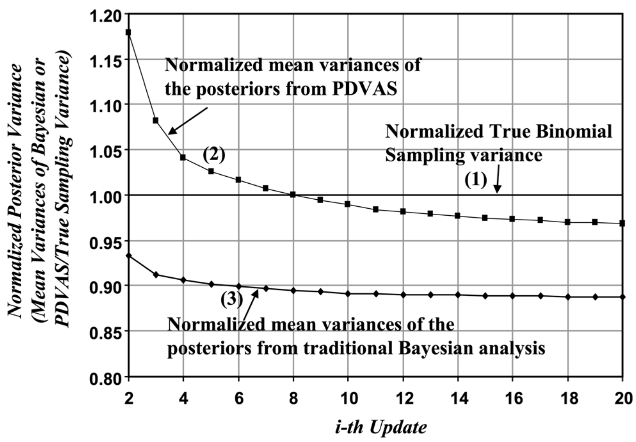

Fig. 13 Variance comparisons of PDVAS and traditional Bayesian for example 2

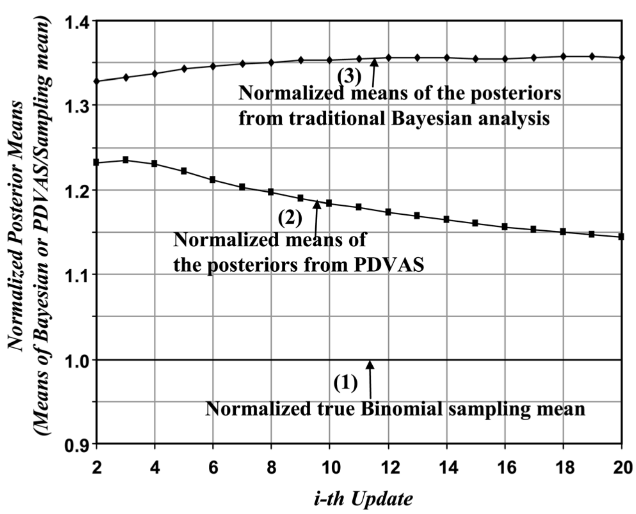

Fig. 14 Mean comparisons of PDVAS and traditional Bayesian for example 3

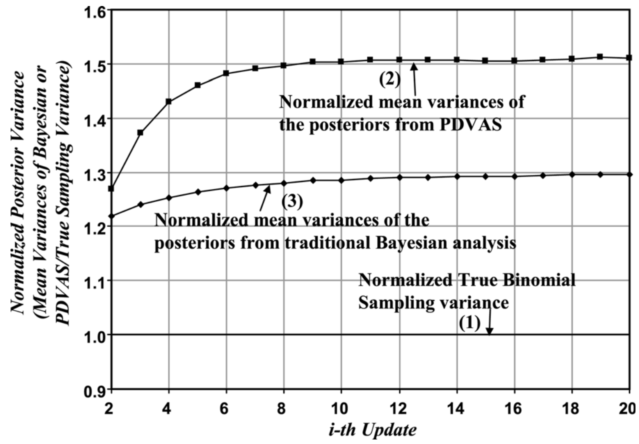

Fig. 15 Variance comparisons of PDVAS and traditional Bayesian for example 3

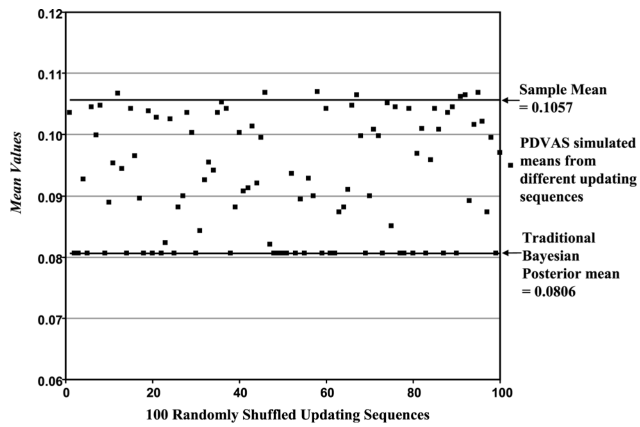

Fig. 16 Means of PDVAS posteriors from 100 different update sequences

Example 2. Similar to example 1, we assume the sampling distribution is a Binomial with $p=0.10$. Initial prior is the noninformative prior Beta(0.5,0.5). However, in this example, we insert a data inconsistency anomaly by randomly generating 10% of the data sampling with Beta(3,97) as the sampling distribution. Notice Beta(3,97) has a much smaller mean $[3/(3+97)=0.03]$ than the Binomial sampling mean $p=0.10$. We want to test, in a simulated practical application, whether PDVAS can detect the data inconsistency as a multiple Bayesian updating is executed to incorporate newly obtained data. Again we use $T_c=0.05$ and the linear S_{DA} curve. Figures 12 and 13 present the mean and variance

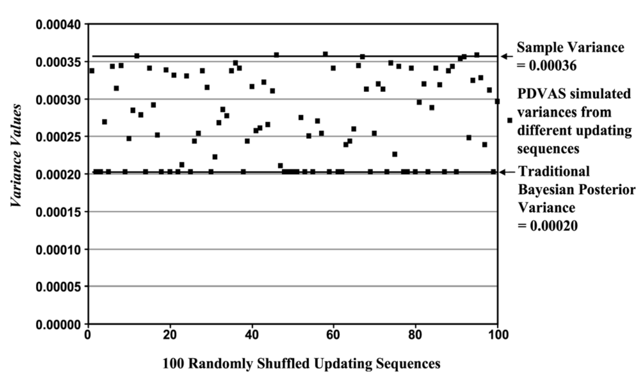

Fig. 17 Variances of PDVAS posteriors from 100 different update sequences

Table 2 Rocket engine reliability data sets in example 5

Design stage	Data category	Number of failures	Number of successes
Concept exploration	Demonstrated reliability from heritage engine A	0	69
	Demonstrated reliability from heritage engine B	0	13
Conceptual design	Combination of SCA and PBMS	Predicted 1 failure per 1000 engine hot fires	
Embodiment design	Laboratory test result	1	4
Development	Subscale development test results	2	18
	Full scale development test results	3	147
Certification	Certification test results	0	120

comparisons, respectively. The results from the figures show that PDVAS is superior to the tradition Bayesian since both means and variances of PDVAS posteriors are closer to the true sampling mean and variance than the tradition Bayesian posteriors after three updates. It is interesting to notice that both PDVAS and traditional Bayesian means and variances diverge from the true sampling mean and variance. This is because there is always 10% of the sampling data with Beta(3,97) as their distribution with a smaller mean and variance than the sampling mean and variance. PDVAS for this case, even though better than the traditional Bayesian process, is not detecting all data inconsistency, which is intentional in the design of PDVAS that is to avoid overcorrecting of the data.

Example 3. All simulation set up in this example is the same as in example 2 except the difference in 10% data inconsistency anomaly. Instead of inserting 10% of sampling data with Beta(3,97), we insert 10% of sampling data with Beta(30,70). Mean comparison in Fig. 14 shows again the superiority of PDVAS over the tradition Bayesian. Variance comparison in Fig. 15 shows PDVAS is worse (bigger variances which are further away from the true sampling variance than the traditional Bayesian variances). A detailed examination of the simulation data indicates that the PDVAS algorithm defined in Figs. 3 and 4 does not differentiate the bad data from the good well, increasing the variance due to some of the good data being thrown out. Further research is needed to improve the algorithm to deal with this problem, which will be discussed in Sec. 6.

Example 4. This example is to show that PDVAS is update sequence dependent and to provide some ideas for how much PDVAS posterior means and variances can vary from different update sequences. The property of the update sequence independence is possessed as a nice feature by the traditional Bayesian process with conjugate priors. It is lost in PDVAS traded with the data inconsistency check and data adjustment with the intention to produce more valid posterior. In this example, we use the noninformative prior Beta(0.5,0.5) and assume the following ten data sets are available for Bayesian updating: (2,12), (3,47),(3,38), (7,45), (2,9),(6,57), (1,6), (2,13), (2,10), and (1,99). The first nine data sets are randomly drawn from the Binomial sampling with $p = 0.10$. The last data set represents a data inconsistency source, which is from Beta(1,99). We randomly shuffled the 10 data sets 100 times and applied PDVAS to each of these 100 shuffles. Figure 16 shows that all 100 PDVAS posterior means (symbolized by dark squares) are closer to the sampling mean (the upper line) than to the traditional Bayesian posterior mean (the lower line), which is a fixed value due to the update sequence independence. Similar phenomenon is observed in Fig. 17 for the simulated variances. The noticeable scatter of the means and variances in PDVAS due to update sequence variations brings an open research question whether it is necessary or possible to further define the PDVAS algorithm to produce an optimized but unique posterior with some predefined optimization criteria. We will discuss this in Sec. 6.

Example 5. This example is to apply PDVAS to a rocket engine reliability analysis. Table 2 presents the data that are assumed to be chronologically obtained as the design maturity evolves. The Bayesian update is performed repeatedly as the new data sets become available to support on-going design decisions.

Table 3 Traditional Bayesian and PDVAS comparisons in example 5

Design stage	Data category	Number of failures	Number of successes	Traditional Bayesian posterior	PDVAS posterior
Initial prior	Noninformative Beta(0.5,0.5)	0.5	0.5		
Concept exploration	Demonstrated reliability from heritage engine A	0	69		
	Demonstrated reliability from heritage engine B	0	13	Beta(0.5,82.5) Mean = 0.0060 Sd = 0.0084	Beta(0.5,82.5) Mean = 0.0060 Sd = 0.0084
Conceptual design	Combination of SCA and PBMS	1	999	Beta(1.5,1081.5) Mean = 0.0014 Sd = 0.0011	Beta(1.0,1081.0) Mean = 0.0009 Sd = 0.0009
Embodiment design	Laboratory test result	1	4	Beta(2.5,1085.5) Mean = 0.0023 Sd = 0.0015	Beta(1.0,1081.0) Mean = 0.0009 Sd = 0.0009
Development	Subscale development test results	2	18	Beta(4.5,1103.5) Mean = 0.0041 Sd = 0.0019	Beta(1.01,1081.09) Mean = 0.0009 Sd = 0.0009
	Full scale development test results	3	147	Beta(7.5,1250.5) Mean = 0.0060 Sd = 0.0022	Beta(2.29,1144.42) Mean = 0.0020 Sd = 0.0013
Certification	Certification test results	0	120	Beta(7.5,1370.5) Mean = 0.0054 Sd = 0.0020	Beta(2.29,1264.42) Mean = 0.0018 Sd = 0.0012

For the SCA and PBMS combined data set, we interpret as 1 failure in 1000 trials so $(F,S)=(1{,}999)$. If we apply the traditional Bayesian analysis to the above data with a noninformative priorBeta(0.5,0.5) as the initial prior, we get the final posterior $\text{Beta}(0.5+1+1+2+3{,}0.5+69+13+999+4+18+147+120) \equiv \text{Beta}(7.5{,}1370.5)$. This posterior has a mean of $7.5/(7.5+1370.5)=0.0054$, a standard deviation of $\sqrt{7.5\text{x}1370.5/[(7.5+1370.5)^2(7.5+1370.5+1)]}=0.0020$, and a COV (coefficient variation, standard deviation/mean) of $0.0020/0.0054=37\%$. Now we apply PDVAS to the same data. Table 3 summarizes the traditional Bayesian and PDVAS results in the chronologically manner. The final PDVAS posterior has a mean of $2.29/(2.29+1264.42)=0.0018$, a standard deviation of $\sqrt{2.29\text{x}1264.42/[(2.29+1264.42)^2(2.29+1264.42+1)]}=0.0012$ and a COV of $0.0012/0.0018=67\%$. Comparing with the traditional Bayesian, PDVAS posterior has a smaller mean (0.0018 versus 0.0054) but bigger COV (67% versus 37%). The mean reduction is because PDVAS discounts several inconsistent data sets, which have significant failure probability. The COV increase is because PDVAS discounts the total data sample size (notice for $\text{Beta}(\alpha,\beta)$, $\text{COV} = \sqrt{\beta}/\left(\sqrt{\alpha}\sqrt{\alpha+\beta+1}\right)$). All these results are in line with the engineering design and development practice that addresses failures when they occur so the failure probability is reduced but at the same time, redesign of the product or implementation of some corrective actions may introduce new uncertainty and unknowns, which leads to a bigger COV.

6 Summary and Concluding Remarks

A Bayesian PDVAS was developed for the Beta-Binomial Bayesian model to address the two controversial issues surrounding the Bayesian reliability analysis, which are the reasonableness of the prior and data consistency. PDVAS attempts to balance retaining the knowledge in the data and reducing the data inconsistency. PDVAS is also devised so that the traditional Bayesian becomes a special case of it and is a default position if we do not have enough knowledge to adjust the data. Several properties of PDVAS were presented that provide insights about PDVAS data adjustment strategy, PDVAS posterior's convergence, and update sequence dependency. The PDVAS adjustment formulas were related to the reliability data categories from engineering design. A detailed data adjustment selection framework was provided to assist PDVAS implementation. Several illustrative examples show the adequateness, effectiveness, and usefulness of PDVAS for the presented application instances. Designers can use PDVAS for data validation and adjustment, especially if suspecting data inconsistency exists. With PDVAS, Bayesian reliability analysis will be more valid with less data inconsistency and contradiction to better support engineering design decisions.

There are some limitations for PDVAS. One limitation is that PDVAS aims at detecting and reducing data noise and disturbance with the assumption that correct sampling data are from a single sampling distribution. PDVAS is not for data comprising that aggregates different sampling distributions in a weighted average manner. Another limitation is the proper balance of the risk of rejecting good data and the risk of failing to detect data inconsistency. As one illustrative example (example 2 of Sec. 5) shows, PDVAS may not be aggressive enough to detect the data inconsistency. The mean of the PDVAS derived posterior in that example is not converging to the true sampling mean, though it is still closer to the true sampling mean, therefore, better than the mean of the traditional Bayesian. Another example (example 4 in Sec. 5) shows, under certain data situations, PDVAS may throw away good data, which leads to overestimating the sampling variance. The results of these two cases indicate that a proper balance of the risk of rejecting good data and the risk of failing to detect data inconsistency may not be easily achieved. As discussed early, PDVAS is an update sequence dependent; therefore, PDVAS applications on conjugated prior data sets need to be cautioned for possible noticeably different posteriors due to different update sequences.

There are further research opportunities for PDVAS. One immediate interest is to extend the PDVAS approach and formulas developed in this paper for the Beta-Binomial model to general prior-sampling distribution situations. Quick examinations of the PDVAS algorithm presented by Figs. 4 and 5 in Sec. 2 and five propositions presented in Sec. 3 indicate that they all can be easily generalized to other prior-sampling distribution situations. The PDVAS consistency measure (Eq. (11) in Sec. 2) and the data adjustment formulas (Eqs. (18) and (19) in Sec. 4) will have to be developed for a specific prior and sampling distribution pair of interest.

Other research opportunities are as follows. Some details of the PDVAS algorithm need to be refined to accommodate more versatile data situations for balancing the two risks (the risk of rejecting good data and the risk of failing to detect data inconsistency). The selection criteria of the PDVAS screen threshold (T_c) needs to be more rigorously established, which links to a data validity measure and data acceptance and rejection risks. Update sequence dependency of PDVAS is unavoidable but the optimization of PDVAS posteriors is worthwhile to explore that can lead to uniqueness of the posterior. The PDVAS data categorization can be more closely defined with engineering design and analysis data as inputs. Finally, PDVAS may provide another criterion to assess the adequacy of the initial prior assignment, which has been an active research area for years in Bayesian analysis.

Appendix: The Proofs of the Propositions

PROPOSITION 1. *For the data set (F_v,S_v) being adjusted in Eq. (13), PDVAS does not change the mean of the data set but increases its standard deviation σ.*

Proof. The mean of the data set (F_v,S_v) is $F_v/(F_v+S_v)$. The mean of the adjusted data set $(\tilde{F}_v,\tilde{S}_v)$ is $\tilde{F}_v/\left(\tilde{F}_v+\tilde{S}_v\right) \equiv S_{\text{DA}}F_v/(S_{\text{DA}}F_v+S_{\text{DA}}S_v)=F_v/(F_v+S_v)$ per Eq. (13). For the σ (standard deviation) of $(\tilde{F}_v,\tilde{S}_v)$, denoted as $\tilde{\sigma}_{(\tilde{F}_v\tilde{S}_v)}$, we use σ formula of the Beta distribution

$$\tilde{\sigma}_{\text{Beta}(\tilde{F}_v,\tilde{S}_v)} = \frac{\sqrt{\tilde{F}_v\tilde{S}_v}}{(\tilde{F}_v+\tilde{S}_v)\sqrt{\tilde{F}_v+\tilde{S}_v+1}} = \frac{\sqrt{S_{\text{DA}}F_vS_{\text{DA}}S_v}}{(S_{\text{DA}}F_v+S_{\text{DA}}S_v)\sqrt{S_{\text{DA}}F_v+S_{\text{DA}}S_v+1}}$$
$$= \frac{\sqrt{F_vS_v}}{(F_v+S_v)\sqrt{S_{\text{DA}}(F_v+S_v)+1}} > \frac{\sqrt{F_vS_v}}{(F_v+S_v)\sqrt{(F_v+S_v)+1}}$$

$\equiv \sigma_{(F_v,S_v)} \equiv$ standard deviation of the unadjusted data set (F_v,S_v) for any $S_{\text{DA}}<1$.

PROPOSITION 2. *At the ith Bayesian updating with the i sets of data (F_0,S_0), (F_1,S_1), …, (F_i,S_i) available, if $[F_v/(F_v+S_v)] < \left[\sum_{j=0,j\neq v}^{j=i}F_j\Big/\left(\sum_{j=0,j\neq v}^{j=i}F_j+\sum_{j=0,j\neq v}^{j=i}S_j\right)\right]$, where $0\le v\le i$, and (F^v,S^v) is one of the i data sets, we then have*

$$\frac{F_v}{F_v+S_v} < \frac{F_v+\sum_{j=0,j\neq v}^{j=i}F_j}{F_v+S_v+\sum_{j=0,j\neq v}^{j=i}F_j+\sum_{j=0,j\neq v}^{j=i}S_j}$$

$$\le \frac{\tilde{F}_v+\sum_{j=0,j\neq v}^{j=i}F_j}{\tilde{F}_v+\tilde{S}_v+\sum_{j=0,j\neq v}^{j=i}F_j+\sum_{j=0,j\neq v}^{j=i}S_j}$$

$$\equiv \frac{S_{\text{DA}}F_v+\sum_{j=0,j\neq v}^{j=i}F_j}{S_{\text{DA}}F_v+S_{\text{DA}}S_v+\sum_{j=0,j\neq v}^{j=i}F_j+\sum_{j=0,j\neq v}^{j=i}S_j} < \frac{\sum_{j=0,j\neq v}^{j=i}F_j}{\sum_{j=0,j\neq v}^{j=i}F_j+\sum_{j=0,j\neq v}^{j=i}S_j}$$

Proof. First, we show

$$\frac{F_v}{F_v+S_v} < \frac{F_v + \sum_{j=0,j\neq v}^{j=i} F_j}{F_v+S_v+\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j} < \frac{\sum_{j=0,j\neq v}^{j=i} F_j}{\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j}. \quad \text{(A1)}$$

From $[F_v/(F_v+S_v)] < [\sum_{j=0,j\neq v}^{j=i} F_j/(\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j)]$, we have $[F_v(\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j)] < [(F_v+S_v)\sum_{j=0,j\neq v}^{j=i} F_j]$ Adding $F_v(F_v+S_v)$ on both sides, we get

$$F_v(F_v+S_v) + F_v\left(\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j\right) < F_v(F_v+S_v) + (F_v+S_v)\sum_{j=0,j\neq v}^{j=i} F_j.$$

So we have $[F_v(F_v+S_v+\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j)] < [(F_v+S_v)(F_v+\sum_{j=0,j\neq v}^{j=i} F_j)]$.
Therefore, we have

$$\frac{F_v}{F_v+S_v} < \frac{F_v + \sum_{j=0,j\neq v}^{j=i} F_j}{F_v+S_v+\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j}$$

Similarly, we can show the right side inequality of (A1).
For proving

$$\frac{F_v}{F_v+S_v} < \frac{S_{DA}F_v + \sum_{j=0,j\neq v}^{j=i} F_j}{S_{DA}F_v+S_{DA}S_v+\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j} < \frac{\sum_{j=0,j\neq v}^{j=i} F_j}{\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j} \quad \text{(A2)}$$

we have

$$\frac{S_{DA}F_v}{S_{DA}F_v+S_{DA}S_v} \equiv \frac{F_v}{F_v+S_v} < \frac{\sum_{j=0,j\neq v}^{j=i} F_j}{\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j}$$

So, similar to the proof of (A1), we have

$$\frac{F_v}{F_v+S_v} \equiv \frac{S_{DA}F_v}{S_{DA}F_v+S_{DA}S_v} < \frac{S_{DA}F_v + \sum_{j=0,j\neq v}^{j=i} F_j}{S_{DA}F_v+S_{DA}S_v+\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j} < \frac{\sum_{j=0,j\neq v}^{j=i} F_j}{\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j}.$$

Lastly, to prove

$$\frac{F_v + \sum_{j=0,j\neq v}^{j=i} F_j}{F_v+S_v+\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j} \le \frac{S_{DA}F_v + \sum_{j=0,j\neq v}^{j=i} F_j}{S_{DA}F_v+S_{DA}S_v+\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j} \quad \text{(A3)}$$

equivalently, we prove

$$\left(F_v + \sum_{j=0,j\neq v}^{j=i} F_j\right)\left(S_{DA}F_v+S_{DA}S_v+\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j\right) - \left(F_v+S_v+\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j\right)\left(S_{DA}F_v + \sum_{j=0,j\neq v}^{j=i} F_j\right) \le 0 \quad \text{(A4)}$$

Reorganizing the left-hand side (LHS) of (A4), we get

$$\text{LHS} = (S_{DA}-1)\left[(F_v+S_v)\sum_{j=0,j\neq v}^{j=i} F_j - F_v(\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j)\right]$$

Since $[F_v/(F_v+S_v)] < [\sum_{j=0,j\neq v}^{j=i} F_j/.(\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j)]$

$$(F_v+S_v)\sum_{j=0,j\neq v}^{j=i} F_j - F_v\left(\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j\right) > 0$$

Since $S_{DA} \le 1$, $S_{DA}-1 \le 0$. Therefore,

$$(S_{DA}-1)\left((F_v+S_v)\sum_{j=0,j\neq v}^{j=i} F_j - F_v(\sum_{j=0,j\neq v}^{j=i} F_j + \sum_{j=0,j\neq v}^{j=i} S_j)\right) \le 0$$

This proved (A3).

PROPOSITION 3. *There always exists a set of* S_{DA} *values such that steps 2 and 3 of Fig. 4 will converge.*

Proof. At the *i*th Bayesian update, the consistency statistic, calculated by Eq. (10), is given by

$$\chi_C^2 = \left[\sum_{j=0}^{j=i}(F_j+S_j)\right]\left(\frac{\sum_{j=0}^{j=i}\frac{F_j^2}{F_j+S_j}}{\sum_{j=0}^{j=i}F_j} + \frac{\sum_{j=0}^{j=i}\frac{S_j^2}{F_j+S_j}}{\sum_{j=0}^{j=i}S_j} - 1\right)$$

This is a χ^2 statistic with the degree of freedom of i. Under the extreme case, we can have all data sets adjusted with the same data adjustment score S_{DA}. Then the adjusted χ_C^2 value, named as $\tilde{\chi}_C^2$, becomes

$$\tilde{\chi}_C^2 = \left[\sum_{j=0}^{j=i}(S_{DA}F_j+S_{DA}S_j)\right]\left(\frac{\sum_{j=0}^{j=i}\frac{(S_{DA}F_j)^2}{S_{DA}F_j+S_{DA}S_j}}{\sum_{j=0}^{j=i}S_{DA}F_j} + \frac{\sum_{j=0}^{j=i}\frac{(S_{DA}S_j)^2}{S_{DA}F_j+S_{DA}S_j}}{\sum_{j=0}^{j=i}S_{DA}S_j} - 1\right)$$

$$= S_{DA}[\sum_{j=0}^{j=i}(F_j+S_j)]\left(\frac{\sum_{j=0}^{j=i}\frac{F_j^2}{F_j+S_j}}{\sum_{j=0}^{j=i}F_j} + \frac{\sum_{j=0}^{j=i}\frac{S_j^2}{F_j+S_j}}{\sum_{j=0}^{j=i}S_j} - 1\right)$$

Therefore, we can always pick an S_{DA} that is small enough such that $C = P(\chi^2 > \tilde{\chi}_C^2) \geq T_c$.

PROPOSITION 4. *For the no-failure situation, that is, in* (F_0,S_0), (F_1,S_1), …, $(F_k S_k)$, *all* $F_i = 0$ *except* F_0, χ_C^2 *by Eq. (10), after substituting* F_0 *by* $\tilde{F}_0 = S_{DA}F_0$ *and* S_0 *by* $\tilde{S}_0 = S_{DA}S_0$, *is a monotone increase function of* S_{DA}, $\chi_C^2 \geq \{[F_0/(F_0+S_0)]\sum_{i=1}^{i=k} S_i\}$ *and* $\chi_C^2 \to \{[F_0/(F_0+S_0)]\sum_{i=1}^{i=k} S_i\}$ *when* $S_{DA} \to 0$.

Proof. After substituting F_0 by $\tilde{F}_0 = S_{DA}F_0$, S_0 by $\tilde{S}_0 = S_{DA}S_0$, and F_i by 0 for $i \geq 1$, Eq. (10) becomes

$$\chi_C^2 = \left[S_{DA}F_0 + S_{DA}S_0 + \sum_{i=1}^{i=k} S_i\right]\left(\frac{F_0\sum_{i=1}^{i=k} S_i}{(F_0+S_0)(S_{DA}S_0 + \sum_{i=1}^{i=k} S_i)}\right)$$

$$= \frac{F_0 + S_0 + \frac{\sum_{i=1}^{i=k} S_i}{S_{DA}}}{S_0 + \frac{\sum_{i=1}^{i=k} S_i}{S_{DA}}}\frac{F_0}{F_0+S_0}\sum_{i=1}^{i=k} S_i$$

It can easily be proven that $[F_0 + S_0 + (\sum_{i=1}^{i=k} S_i/S_{DA})]/[S_0 + (\sum_{i=1}^{i=k} S_i/S_{DA})]$ is a monotone increase function of S_{DA} from calculus. It is also easily seen that when $S_{DA} \to 0$, $[F_0 + S_0 + (\sum_{i=1}^{i=k} S_i/S_{DA})]/[S_0 + (\sum_{i=1}^{i=k} S_i/S_{DA})] \to 1$, which leads to $\chi_C^2 \to (F_0/F_0+S_0)\sum_{i=1}^{i=k} S_i$. Since χ_C^2 is a monotone increase function of S_{DA}, therefore $\chi_C^2 \geq (F_0/(F_0+S_0))\sum_{i=1}^{i=k} S_i$.

References

[1] Wang, P., Youn, B. D, Xi, Z., and Kloess, A., 2009, "Bayesian Reliability Analysis With Evolving, Insufficient and Subjective Data Sets," ASME J. Mech. Des., **131**, p. 111008.

[2] Coolen, F. P. A., and Newby, M. J., 1994, "Bayesian Reliability Analysis With Imprecise Prior Probabilities," Reliab. Eng. Syst. Saf., **43**, pp. 75–85.

[3] Martz, H. F., and Waller, R. A., 1982, *Bayesian Reliability Analysis*, John Wiley & Sons, New York.

[4] Hamada, M. S., Wilson, A. G., Reese, C. S., and Martz, H. F., 2008, *Bayesian Reliability*, Springer, New York.

[5] Gelman, A., 2008, "Objections to Bayesian Statistics," Bayesian Anal., **3**, pp. 445–450.

[6] Efron, B., 1986, "Why Isn't Everyone a Bayesian?" Am. Stat., **40**(1), pp. 1–5.

[7] Jeffreys, H., 1946, "An Invariant Form for the Prior Probability in Estimation Problems," Proc. R. Soc. London, **186**, pp. 453–461.

[8] Kass, R. E., and Wasserman, L., 1996, "The Selection of Prior Distributions by Formal Rules," J. Am. Stat. Assoc., **91**(435), pp. 1343–1370.

[9] Jaynes, E. T., 1982, "On the Rationale of Maximum Entropy Methods," Proc. IEEE, **70**, pp. 939–952.

[10] Bernardo, J. M., 1979, "Reference Posterior Distribution for Bayesian Inference," J. R. Stat. Soc. B, **41**(2), pp. 113–147.

[11] Welch, B., and Peers, H. W., 1963, "On Formulas for Confidence Points on Integrals of Weighted Likelihood," J. R. Stat. Soc. B, **25**, pp. 318–329.

[12] Datta, G. S., 1996, "On Priors Providing Frequentist Validity of Bayesian Inference for Multiple Parametric Functions," Biometrika, **83**(2), pp. 287–298.

[13] Fraser, D. A. S., and Reid, N., 2002, "Strong Matching of Frequentist and Bayesian Parametric Inference," J. Stat. Plann. Inference, **103**, pp. 263–285.

[14] Huang, Z., and Jin, Y., 2009, "Reliability Prediction Methods: A Survey and Selection for Mechanical Design-for-Reliability," ASME IDETC 2009-87103, San Diego, CA.

[15] Grantham Lough, K., Stone, R. B., and Tumer, I., 2005, "Function Based Risk Assessment: Mapping Function to Likelihood," DETC 2005-85053, Long Beach, CA.

[16] Grantham Lough, K., Stone, R. B., and Tumer, I., 2006, "Prescribing and Implementing the Risk in Early Design (RED) Method," DETC 2006-99374, Philadelphia, PA.

[17] Wang, K., and Jin, Y., 2002, "An Analytical Approach to Functional Design," DETC 2002-34084, Montreal, Canada.

[18] Huang, Z., and Jin, Y., 2009, "Extension of Stress and Strength Interference Theory for Conceptual Design-for-Reliability," ASME J. Mech. Des., **131**, p. 071001.

[19] Evan, M., and Moshonov, H., 2006, "Checking for Prior-Data Conflict," Bayesian Anal., **4**, pp. 893–914

[20] Pearson, E. S., 1947, "The Choice of Statistical Tests Illustrated on the Interpretation of Data Classed in a 2×2 Table," Biometrika, **34**, pp. 139–167.

[21] Plackett, R. L., 1983, "Karl Pearson and the Chi-Square Test," Int. Statist. Rev., **51**(1), pp. 59–72.