DEVCOM DAC-TR-2024-044
May 2024

Bayesian Methods for Assessing Reliability Fix Effectiveness Factors

by Martin Wayne

DEVCOM DAC-TR-2024-044
May 2024

Bayesian Methods for Assessing Reliability Fix Effectiveness Factors

by Martin Wayne
DEVCOM Analysis Center

REPORT DOCUMENTATION PAGE

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**	
		START DATE	**END DATE**
May 2024	Technical Report	10/1/2023	5/1/2024

4. TITLE AND SUBTITLE
Bayesian Methods for Assessing Reliability Fix Effectiveness Factors

5a. CONTRACT NUMBER	**5b. GRANT NUMBER**	**5c. PROGRAM ELEMENT NUMBER**
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)
Martin Wayne

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Director DEVCOM Analysis Center 6896 Mauchly Street Aberdeen Proving Ground, MD 21005	**8. PERFORMING ORGANIZATION REPORT NUMBER** DEVCOM DAC-TR-2024-044

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Reliability growth testing is used to identify failure modes in the system that can be mitigated through corrective actions to reduce their rate of occurrence. Fix Effectiveness Factors (FEFs) are used to characterize the effectiveness of the corrective actions, and they are also used in reliability growth projections to understand the reliability improvement that can be expected after the corrective actions are implemented. Testing is often used to verify that the corrective actions meet the desired FEF values. This report presents a Bayesian approach that uses the data both before and after the corrective action to develop a posterior estimate of the FEF. The Bayesian prior can be informative or non-informative, and applications for designing FEF verification tests are also presented.

15. SUBJECT TERMS
reliability growth, Fix Effectiveness Factor, FEF, corrective action, Bayesian statistics

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**
a. REPORT	**b. ABSTRACT**	**C. THIS PAGE**		
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UU	26

19a. NAME OF RESPONSIBLE PERSON	**19b. PHONE NUMBER** *(Include area code)*
Martin Wayne	(410) 278-3221

STANDARD FORM 298 (REV. 5/2020)
Prescribed by ANSI Std. Z39.18

Table of Contents

List of Figures

1. INTRODUCTION

The goal of a developmental reliability growth program is to find failure modes in the system and develop corrective actions to reduce their rate of occurrence. In developmental testing it is therefore common to see numerous corrective actions or design modifications meant to mitigate failure modes that have been observed. Reliability data are also collected during these developmental test events, and the reliability of the current configuration is provided to decision-makers. Projected reliability values are also estimated when additional corrective actions are planned.

These results are obtained using reliability growth projection models, which provide an estimate of the system reliability after the corrective actions have been implemented. The projected result is obtained with the use of a Fix Effectiveness Factor (FEF), which is the fractional reduction in the rate of occurrence of a failure mode after a corrective action has been implemented. Engineering judgment is often used initially in the absence of verification test results. The assignment of the FEF through this process involves subject matter experts familiar with the design of the system or component in question, along with careful consideration of the failure mode and corrective action being implemented.

Previous work[1,2] in this area focused on estimating FEF values to provide a historical database of values. FEF estimation was performed through two different methods. The first estimated the average FEF across the entire system when a group of delayed corrective actions were implemented. The system-level average FEF was defined as

$$d_{avg} = \frac{\lambda_1 - \lambda_2}{\lambda_B - h(T)}, \tag{1}$$

where λ_1 is the overall system failure rate for the first test phase, λ_2 is the overall system failure rate for the second test phase, λ_B is the rate of occurrence of failure modes that will be fixed if they are observed, and $h(T)$ is the rate of occurrence of failure modes that have not yet been observed in testing. Values for λ_1 and λ_2 can be estimated with standard maximum likelihood estimates for the Exponential Distribution, and λ_B and $h(T)$ can be estimated using a reliability growth projection model. The U.S. Army Materiel Systems Analysis Activity (AMSAA)-Crow Projection Model[3] is mentioned as a useful parametric approach in TR-388[1]; and the AMSAA Maturity Projection Model[4] is another more recent parametric approach that is useful.

The second method involved FEF estimation for a single failure mode, with the estimate defined as

$$d = \frac{\lambda_1 - \lambda_2}{\lambda_1}. \tag{2}$$

Standard maximum likelihood estimates of the individual failure rates are used for λ_1 and λ_2 in Eq. 2. This method will provide a reasonable point estimate of the FEF in cases where there are sufficient data to obtain reasonable point estimates of the mode failure rates both before and after the corrective action has been implemented. There are no confidence interval procedures to provide a measure of uncertainty, but approximate intervals could be developed using error propagation techniques.

Estimating the FEF implicitly involves comparing two failure rates, and there are several approaches that can be found in the literature in this area. However, these methods assume that the two failure rates being compared are independent from each other. This means they are of limited value in FEF estimation because there is an inherent dependency between the two failure rates induced through the corrective action process. A common example of these approaches is the method proposed by Nelson.[5] The method works directly with the ratio of the failure rates defined as

$$\rho = \frac{\lambda_1}{\lambda_2}, \tag{3}$$

and develops a $100\gamma\%$ confidence interval on ρ. The limits of the two-sided $100\gamma\%$ confidence interval on ρ are defined as $\left(\underline{\rho}, \overline{\rho}\right)$, where $\underline{\rho}$ and $\overline{\rho}$ are defined as

$$\underline{\rho} = \left(\frac{\frac{Y}{t}}{\frac{X+1}{s}}\right)\left(\frac{1}{F\left[\frac{1+\gamma}{2}, 2X+2, 2Y\right]}\right) \tag{4}$$

$$\overline{\rho} = \left(\frac{\frac{Y+1}{t}}{\frac{X}{s}}\right)\left(\frac{1}{F\left[\frac{1+\gamma}{2}, 2Y+2, 2X\right]}\right), \tag{5}$$

where $F[\mathrm{x}, \mathrm{a}, \mathrm{b}]$ is the $100x^{th}$ percentile of the F-distribution with a degrees of freedom in the numerator and b degrees of freedom in the denominator. If the interval does not contain the value 1, the two failure rates are said to be different at the $100(1-\gamma)\%$

significance level. This approach is equivalent to providing a confidence interval on the complement of the FEF, albeit with the assumption that the two failure rates are independent.

This report presents a Bayesian approach to estimating the FEF for a single failure mode. The posterior uses the data from both before and after the corrective action is implemented. Prior distributions are used for the failure rate and FEF. A non-informative Gamma prior is used for the failure rate and a Beta distribution is used for the FEF, which can be non-informative or informative. Methods for assigning an informative prior are discussed, and applications for the design of verification tests are presented. The methodology should help decision-makers by increasing their understanding of the impacts of any corrective actions that are made during development.

2. METHODOLOGY

Estimation of the effectiveness of a corrective action involves collecting data before and after the corrective action has been implemented. The approach presented here assumes a constant failure rate both before and after the corrective action is implemented, which is standard in reliability growth projection models.[3,4,6] The FEF, $\tilde{d}$, is defined identically to Eq. 2 as

$$\tilde{d} = \frac{\lambda_1 - \lambda_2}{\lambda_1}, \tag{6}$$

where λ_1 is the failure rate before the corrective action and λ_2 is the reduced failure rate after the corrective action has been implemented. For n_1 failures in test time T_1 before the corrective action and n_2 failures in test time T_2 after the corrective action, the likelihood is a product of Poisson probabilities given by

$$p(n_1, n_2 | \lambda_1, \lambda_2, T_1, T_2) = \frac{(\lambda_1 T_1)^{n_1}}{n_1!} \exp(-\lambda_1 T_1) \frac{(\lambda_2 T_2)^{n_2}}{n_2!} \exp(-\lambda_2 T_2). \tag{7}$$

Using the expression in Eq. 6 to express λ_2 as a function of both λ_1 and $\tilde{d}$ allows the likelihood in Eq. 7 to be reparametrized as

$$p\left(n_1, n_2 \middle| \lambda_1, \tilde{d}, T_1, T_2\right) = \frac{(\lambda_1 T_1)^{n_1}}{n_1!} \exp(-\lambda_1 T_1) \frac{\left[(1-\tilde{d})\lambda_1 T_2\right]^{n_2}}{n_2!} \exp\left[-(1-\tilde{d})\lambda_1 T_2\right]. \tag{8}$$

Assuming the corrective action does not increase the failure rate implies that $\tilde{d} \in [0,1]$. A Beta distribution is therefore a natural choice for the prior distribution on $\tilde{d}$, with a probability density defined as

$$p\left(\tilde{d} \middle| a, b\right) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \tilde{d}^{a-1} \left(1-\tilde{d}\right)^{b-1}. \tag{9}$$

This prior is not only a natural choice for the FEF, but it has an additional benefit in that it allows for an analytic solution to the posterior distribution.

For the prior distribution on the failure rate, λ_1, a Gamma distribution is a natural choice defined as

$$p(\lambda_1|\alpha,\beta) = \frac{{\lambda_1}^{\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}} \exp\left[-\frac{\lambda_1}{\beta}\right]. \tag{10}$$

The posterior distribution for $\tilde{d}$ can then be found by integrating the joint posterior with respect to the failure rate, λ_1.

$$p(\tilde{d}|n_1,n_2) = \int_0^{\infty} p(\lambda_1,\tilde{d}|n_1,n_2)d\lambda_1 \; = \int_0^{\infty} \frac{p(\tilde{d})p(\lambda_1)p(n_1,n_2|\lambda_1,\tilde{d})}{\int_0^1\int_0^{\infty} p(\tilde{d})p(\lambda_1)p(n_1,n_2|\lambda_1,\tilde{d})d\lambda_1 d\tilde{d}} d\lambda_1 \tag{11}$$

Combining Eqs. 8–10, the denominator in Eq. 11 is given by

$$\begin{aligned}&\int_0^1\int_0^{\infty} p(\tilde{d})p(\lambda_1)p(n_1,n_2|\lambda_1,\tilde{d},T_1,T_2)d\lambda_1 d\tilde{d}\\ &= \frac{{T_1}^{n_1}{T_2}^{n_2}\Gamma(a+b)\Gamma(\alpha+n_1+n_2)\Gamma(b+n_2)}{n_1!n_2!\Gamma(b)\Gamma(\alpha)\Gamma(a+b+n_2)\beta^{\alpha}\left(\frac{1}{\beta}+T_1+T_2\right)^{\alpha+n_1+n_2}} {}_2F_1\left(\alpha+n_1+n_2, a, a+b+n_2, \frac{T_2}{\frac{1}{\beta}+T_1+T_2}\right).\end{aligned} \tag{12}$$

where ${}_2F_1(a,b,c,z)$ is the hypergeometric function defined as

$${}_2F_1(a,b,c,z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 x^{b-1}(1-x)^{c-b-1}(1-zx)^{-a}dx\,. \tag{13}$$

The numerator in Eq. 11 can be solved similarly to yield

$$\begin{aligned}&\int_0^{\infty} p(\tilde{d})p(\lambda_1)p(n_1,n_2|\lambda_1,\tilde{d},T_1,T_2)d\lambda_1\\ &\qquad = \frac{{T_1}^{n_1}{T_2}^{n_2}\Gamma(a+b)\Gamma(\alpha+n_1+n_2)\tilde{d}^{a-1}(1-\tilde{d})^{b+n_2-1}}{n_1!\,n_2!\,\Gamma(a)\Gamma(b)\Gamma(\alpha)\beta^{\alpha}\left(\frac{1}{\beta}+T_1+(1-d)T_2\right)^{\alpha+n_1+n_2}}\end{aligned} \tag{14}$$

Combining Eqs. 12 and 14 provides the complete posterior for $\tilde{d}$ as

$$p(\tilde{d}|n_1,n_2,T_1,T_2) = \frac{\frac{\Gamma(a+b+n_2)}{\Gamma(a)\Gamma(b+n_2)}\tilde{d}^{a-1}(1-\tilde{d})^{b+n_2-1}\left(\frac{\frac{1}{\beta}+T_1+T_2}{\frac{1}{\beta}+T_1+(1-\tilde{d})T_2}\right)^{\alpha+n_1+n_2}}{{}_2F_1\left(\alpha+n_1+n_2,a,a+b+n_2,\frac{T_2}{\frac{1}{\beta}+T_1+T_2}\right)} \tag{15}$$

The posterior distribution in Eq. 15 resembles a Beta distribution, which is expected given the choice of the prior distribution. The left three terms in the numerator are a standard Beta distribution, but there is an additional term containing $\tilde{d}$ along with the hypergeometric function in the denominator. This viewpoint will be helpful when considering choices for the parameters of the prior distributions in Section 2.1.

The posterior mean can also be found using the posterior distribution in Eq. 15 and is given by

$$E\left[\tilde{d}\middle|n_1,n_2,T_1,T_2\right] = \left(\frac{a}{a+b+n_2}\right)\frac{{}_2F_1\left(\alpha+n_1+n_2,a+1,a+b+n_2+1,\frac{T_2}{\frac{1}{\beta}+T_1+T_2}\right)}{{}_2F_1\left(\alpha+n_1+n_2,a,a+b+n_2,\frac{T_2}{\frac{1}{\beta}+T_1+T_2}\right)}. \tag{16}$$

A corresponding 100x% lower probability bound can be found numerically by solving for the value of $\tilde{d}$ that satisfies

$$\int_0^{\tilde{d}} p(y|n_1,n_2,T_1,T_2)dy = x. \tag{17}$$

The maximum a posteriori (MAP) estimate may also be useful, as it will likely be closer to the traditional point estimate found by plugging in individual maximum likelihood estimates into Eq. 6. It can be found by setting the derivative with respect to $\tilde{d}$ of the posterior in Eq. 15 equal to zero. The result is a quadratic function that can be solved directly using the quadratic formula shown in Eq. 18, with the terms in the formula defined in Eqs. 19–21.

$$MAP\left(\tilde{d}\middle|n_1,n_2,T_1,T_2\right) = \frac{-\tilde{b} \pm \sqrt{\tilde{b}^2 - 4\tilde{a}\tilde{c}}}{2\tilde{a}} \tag{18}$$

$$\tilde{a} = \beta T_2(2-a-b+n_1+\alpha) \tag{19}$$

$$\tilde{b} = -(2-a-b-n_2)(1+\beta T_1) - \beta T_2(3-2a-b+n_1+\alpha) \tag{20}$$

$$\tilde{c} = (1-a)(1+\beta T_1+\beta T_2) \tag{21}$$

2.1 Choosing Prior Parameters

The prior distribution is intended to represent the state of knowledge before the data are observed, and this point should be considered carefully when choosing parameters for the prior distributions presented here. When assessing the FEF, minimizing the influence of the Gamma prior on the failure rate is generally desirable. The prior Gamma parameters are both additive with the failures and test times in Eq. 15, which implies that α and $\frac{1}{\beta}$ should be relatively small. Setting α = 0.0001 and β = 10,000 should sufficiently minimize their impact on the posterior.

For the prior Beta on the FEF, relevant prior information can also be used as appropriate. The type of failure mode and proposed corrective action, along with any supporting information from component testing or modeling and simulation, can all be used to inform the prior parameters. As mentioned in the introduction, relevant FEF information from various types of systems has also been documented in the past[1,2] and may be useful. Because the Beta distribution has two parameters, two prior constraints are required to define the distribution. For example, if the prior mean and variance of the FEF are determined using relevant information, the corresponding prior Beta parameters can be found by equating them to the mean and variance of the Beta distribution and solving the two equations. Alternative constraints, such as a desired percentile of the distribution, can also be used to define the distribution.

If sufficient information is not available to identify two constraints for the prior, Maximum Entropy principles[7,8] can be used to arrive at an approximate Beta distribution using only the assumed average value of the FEF. Maximizing the entropy subject to the assumed mean value of the FEF parameter and a range of (0,1) results in the prior distribution for $\tilde{d}$ being a truncated Exponential distribution given by

$$p(\tilde{d}) = \frac{\mu \exp(-\mu \tilde{d})}{1 - \exp(-\mu)} \tag{22}$$

where μ is the solution to

$$\frac{1}{\mu} - \frac{\exp(-\mu)}{1 - \exp(-\mu)} = \varepsilon \tag{23}$$

for mean FEF value ε. The Beta parameters a and b can then be found by equating the means and second moments of the Truncated Exponential and Beta distributions, which results in the system of equations given by Eqs. 24 and 25.

$$\frac{a}{a+b} = \frac{1}{\mu} - \frac{exp(-\mu)}{1 - exp(-\mu)} \tag{24}$$

$$\left(\frac{a}{a+b}\right)\left(\frac{a+1}{a+b+1}\right) = \frac{exp(-\mu) + 2\left[-\frac{1}{\mu}exp(-\mu) - \frac{1}{\mu^2}exp(-\mu) + \frac{1}{\mu^2}\right]}{1 - exp(-\mu)} \tag{25}$$

However, this approach requires caution because Eq. 23 only yields a positive solution for the rate μ of the Truncated Exponential distribution when the mean FEF is less than 0.5. The symmetry of the Beta distribution allows for a workaround in cases where the mean FEF is greater than 0.5. Applying the Truncated Exponential to 1-FEF instead of the FEF itself will provide a positive solution for the rate parameter. The resulting Beta(a, b) distribution for 1-FEF induces a Beta(b, a) on the FEF with the desired mean value. The variance will also be correct because the variance of a Beta(a, b) random variable is equivalent to the variance of a Beta(b, a) random variable.

Equation 23 is also degenerate when the mean FEF value is set to 0.5; however, this problem can be overcome by perturbing the mean slightly to 0.50001. This difference is not meaningful for the prior mean FEF, and it allows for a numerical solution. The solution is a Beta distribution that is approximately Uniform, which is also the Maximum Entropy distribution for a variable with range of (0,1) with no additional constraints. This implies that a prior mean FEF of 0.5 alone is not adding much information to the prior distribution.

2.2 Alternative Prior Distribution: Truncated Normal

In some situations, it may be necessary to relax the assumption that the corrective action decreases the failure rate. This allows for the possibility of an ineffective corrective action that increases the failure rate, and the resulting FEF will be negative. The prior Beta distribution from the previously proposed model will no longer be valid, as it is bounded below by zero. Several alternative prior distributions are possible, and a Normal distribution is one potential candidate. The FEF will still have a maximum value of one in this situation; thus, a Truncated-Normal is more technically valid. The result will be the same in either case however, as the truncation involves a constant of proportionality that will not impact the posterior distribution. The joint posterior on the FEF and mode failure rate is defined as

$$p(\lambda_1, \tilde{d}|n_1, n_2, T_1, T_2) \propto exp\left(-\frac{\tilde{d}^2}{2\sigma^2}\right)(1-\tilde{d})^{n_2}\lambda_1^{\alpha+n_1+n_2-1}exp(-\lambda_1[\tau + T_1 + (1-\tilde{d})T_2]) \tag{26}$$

An analytic solution is not available in this case; therefore, Markov Chain Monte Carlo (MCMC) techniques can again be used to generate values from the posterior shown in Eq. 22. For MCMC using Gibb's sampling,[9] the full conditional posterior distributions are given by Eqs. 27 and 28.

$$p(\lambda_1|\tilde{d}, n_1, n_2, T_1, T_2) \propto \lambda_1^{\alpha+n_1+n_2-1}exp(-\lambda_1[\tau + T_1 + (1-\tilde{d})T_2]) \tag{27}$$

$$p(\tilde{d}|\lambda_1, n_1, n_2, T_1, T_2) \propto exp\left(-\frac{\tilde{d}^2}{2\sigma^2}\right)(1-\tilde{d})^{n_2}exp(-(1-\tilde{d})\lambda_1 T_2) \tag{28}$$

The conditional distribution on λ_1 in Eq. 27 is a Gamma distribution with shape parameter $\alpha + n_1 + n_2$ and rate parameter $\tau + T_1 + (1-\tilde{d})T_2$, which can be sampled directly using standard functions. The conditional posterior on $\tilde{d}$ is not a specific distribution, so alternative methods are required to generate posterior samples. Slice sampling[10] is a convenient approach that can be used to form a Slice-in-Gibb's sampler for $\tilde{d}$.

For the slice sampler for $\tilde{d}$ in Eq. 28, the log-transform of the distribution provides an easier implementation of the sampling algorithm. The log-transform of the distribution is given by

$$log[p(\tilde{d}|\lambda_1, n_1, n_2, T_1, T_2)] \propto n_2 log(1-\tilde{d}) - \frac{\tilde{d}^2}{2\sigma^2} - (1-\tilde{d})\lambda_1 T_2. \tag{29}$$

Note that this form also provides a natural transform of the support to $(-\infty, 1)$ due to the $log(1-\tilde{d})$ term. The transform of the support is beneficial since the FEF cannot be greater than one, and it provides an automatic truncation to the Normal prior. To further understand the behavior of the function, the limits for small and large values of $\tilde{d}$ can be calculated. The limit of the distribution as $\tilde{d}$ approaches 1 is $-\infty$, and the limit as $\tilde{d}$ approaches $-\infty$ is also $-\infty$. The derivative of the distribution with respect to $\tilde{d}$ is

$$\frac{dlog[p(\tilde{d}|\lambda_1, n_1, n_2, T_1, T_2)]}{d\tilde{d}} \propto -\frac{n_2}{1-\tilde{d}} - \frac{\tilde{d}}{\sigma^2} + \lambda_1 T_2, \tag{30}$$

which is positive as $\tilde{d}$ approaches $-\infty$. This implies that the log-transform is increasing from $-\infty$ to a maximum value before decreasing to $-\infty$ as $\tilde{d}$ approaches 1. These properties make it straightforward to find a horizontal slice through the function for sampling. A sample plot of the function is shown in Figure 1.

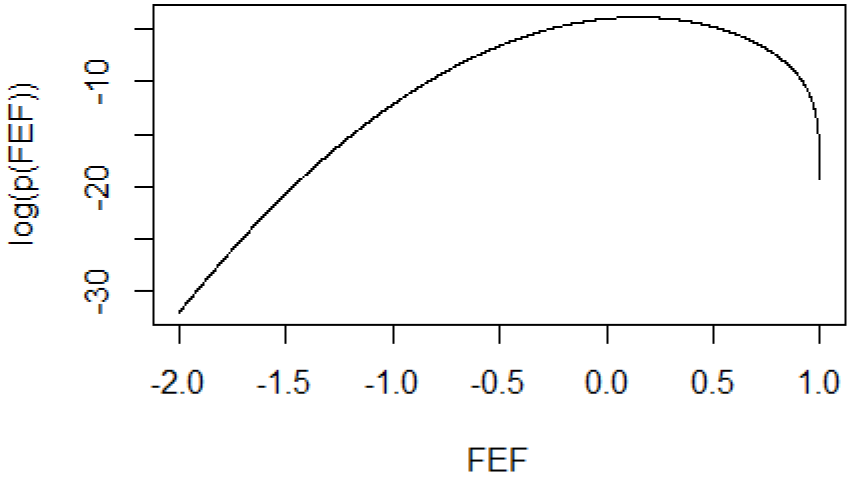

Figure 1. Example behavior of log transform of conditional distribution on FEF

3. PRIOR SENSITIVITY ANALYSIS

3.1 Beta Prior Distribution

This section provides examples of the FEF assessment using two approaches for the prior Beta distribution. This is helpful for understanding how the prior combines with various amounts of data to impact the posterior. For the example, assume the failure mode occurred four times in 500 h of testing prior to the application of the corrective action, and two times in 300 h of additional testing after the corrective action. The posterior distribution on the FEF can therefore be found by setting n_1 = 4, n_2 = 2, T_1 = 500 and T_2 = 300. The posterior distribution resulting from a non-informative Beta(1,1) prior is shown in Figure 2. The posterior mean is 0.37, and the 80% lower probability bound is 0.14. The MAP estimate using Eq. 18 is 0.17.

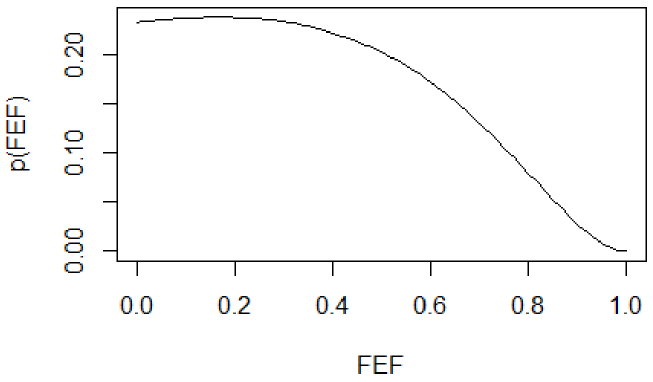

Figure 2. **Posterior on FEF assuming non-informative Beta(1,1) prior distribution**

To understand the posterior impacts from increased data after the corrective action, assume the test length and failures both increase by a factor of five so that n_2 = 10 and T_2 = 1500. The resulting failure rate is still the same, but additional data should influence the shape of the posterior. The posterior mean resulting from a uniform Beta(1,1) prior is 0.30, and the 80% lower probability bound is 0.12. The MAP estimate is still 0.17, which is to be expected.

To demonstrate the impacts of the prior parameters on the posterior distribution, Figure 3 shows the resulting posterior using the Maximum Entropy approach with a prior mean FEF of 0.7. The prior in this case is Beta(1.74,0.74), which shifts the mass of the posterior significantly to the right. The posterior mean is 0.51, and the 80% lower probability bound is 0.31. The MAP estimate is 0.58. This shows the sensitivity of the result to the prior assumptions on the FEF and demonstrates that the use of an informative prior should be underpinned by relevant supporting information.

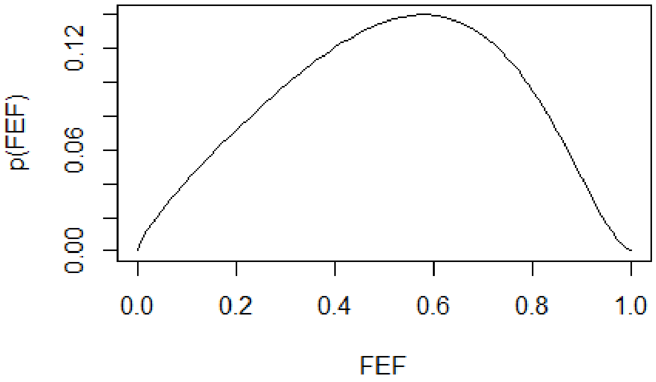

Figure 3. **Posterior on FEF assuming informative Beta(1.74,0.74) prior distribution**

3.2 Truncated-Normal Prior Distribution

For the Truncated-Normal prior distribution, again assume the failure mode occurred four times in 500 h of testing prior to the application of the corrective action, and two times in 300 h of additional testing after the corrective action. The posterior distribution on the FEF can therefore be found by setting n_1 = 4, n_2 = 2, T_1 = 500, and T_2 = 300. For the prior standard deviation, assume σ = 0.333. This implies that approximately 99% of the distribution will fall in the interval (–1,1). If the FEF is equal to –1, the failure rate after the attempted corrective action will be double the initial failure rate. FEF values below –1 will increase the failure rate even further; therefore, the choice of standard deviation is reasonably conservative.

The posterior mean for this case is 0.003, and the 80% lower probability bound is –0.254. This shows that the relaxed assumption of the prior distribution has a significant impact on the posterior result. If the data after the corrective action is increased by a factor of five so that n_2 = 10 and T_2 = 1500, the posterior mean increases slightly to 0.004, and the 80% lower probability bound becomes –0.234. The change is not meaningful in practice, but it does illustrate that more data will generally offer a reduction in the uncertainty in the posterior distribution.

4. TEST PLANNING TO VERIFY FIX EFFECTIVENESS

Figure 4 shows the impacts of testing after the corrective action on the posterior distribution for the FEF. The assumed prior is a Beta(1,1), which is a uniform distribution, with n_1 = 4 and T_1 = 500. The horizontal black line indicates that the posterior is still a uniform distribution when no additional data are available after the corrective action.

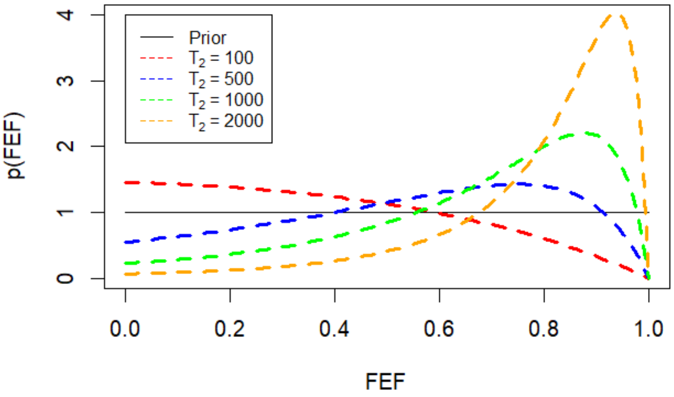

Figure 4. Effects of testing after corrective action on FEF posterior assuming Uniform prior distribution

Figure 4 demonstrates that the posterior shape will change significantly when holding the number of failures after the corrective action constant and varying the amount of testing. If the intent of the testing after the corrective action is to verify the achieved FEF, there are two approaches that can be used. The first uses the posterior distribution directly. The second approach uses the assurance testing framework discussed by Meeker & Escobar[11] and Hamada et al.[12]

4.1 Using Posterior Percentiles

For a desired posterior probability and allowable number of failures after the corrective action, the required amount of testing can be determined. This can be performed using numerical integration on the posterior distribution in Eq. 15 within a simple bisection search algorithm. For the desired FEF, desired number of failures c, and desired posterior percentile γ, the verification test length T_2 can be found by solving Eq. 31 numerically.

$$\gamma = \int_0^{FEF} \frac{\frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b+c)} x^{a-1}(1-x)^{b+c-1}\left(\frac{\frac{1}{\beta}+T_1+T_2}{\frac{1}{\beta}+T_1+(1-d)T_2}\right)^{\alpha+n_1+c}}{{}_2F_1\left(\alpha+n_1+c,a,a+b+c,\frac{T_2}{\frac{1}{\beta}+T_1+T_2}\right)} dx \qquad (31)$$

4.2 Using Assurance Testing

Assurance testing treats the verification test as a pass/fail event. The plan is developed by determining the maximum number of allowable failures that can be observed in the assurance (i.e., verification) test. Two posterior risks are considered. The posterior consumer risk is defined as the posterior probability that the FEF value is below the desired threshold when the test is passed. The posterior producer risk is the opposite risk and is defined as the posterior probability that the FEF value is above the desired threshold when the test is failed. The mathematical definitions of the posterior consumer and posterior producer risks are defined in Eqs. 32 and 33, respectively.

$$p\left(\tilde{d} \leq FEF \mid n_2 \leq c\right) = \int_{0}^{FEF} p\left(\tilde{d} \middle| n_2 \leq c\right) d\tilde{d} = \int_{0}^{FEF} \frac{p\left(n_2 \leq c \middle| \tilde{d}\right) p\left(\tilde{d}\right)}{\int_0^1 p\left(n_2 \leq c \middle| \tilde{d}\right) p\left(\tilde{d}\right) d\tilde{d}} d\tilde{d} \tag{32}$$

$$p\left(\tilde{d} > FEF \mid n_2 > c\right) = \int_{FEF}^{1} p\left(\tilde{d} \middle| n_2 > c\right) dd = \int_{FEF}^{1} \frac{p\left(n_2 > c \middle| \tilde{d}\right) p\left(\tilde{d}\right)}{\int_0^1 p\left(n_2 > c \middle| \tilde{d}\right) p\left(\tilde{d}\right) d\tilde{d}} d\tilde{d} \tag{33}$$

The risk definitions contain two probability distributions. The first is the prior distribution for the FEF $\tilde{d}$, and the second is the likelihood of observing n_2 given a value of $\tilde{d}$. The prior distribution for $\tilde{d}$ is the same as that defined in Eq. 9, which is the Beta distribution defined by

$$p\left(\tilde{d} \middle| a, b\right) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \tilde{d}^{a-1} \left(1 - \tilde{d}\right)^{b-1}. \tag{34}$$

The likelihood of observing n_2 given a value of $\tilde{d}$ can be defined as

$$p\left(n_2 \mid \tilde{d}\right) = \int_0^{\infty} p(n_2 \mid \tilde{d}, \lambda)\, p(\lambda)\, d\lambda\,, \tag{35}$$

where $p(\lambda)$ is the distribution on the failure rate λ prior to the verification test. Using the observed failure data of n_1 failures in test length T_1 with the prior Gamma distribution defined in Eq. 10, the posterior after the original test will be a Gamma distribution defined as

$$p(\lambda|\alpha,\beta,n_1,T_1) = \frac{\lambda^{\alpha+n_1-1}\left(\frac{1}{\beta}+T_1\right)^{\alpha+n_1}}{\Gamma(\alpha+n_1)} \exp\left[-\left(\frac{1}{\beta}+T_1\right)\lambda\right]. \tag{36}$$

The likelihood $p(n_2 \mid \tilde{d},\lambda)$ is then defined as

$$p(n_2 \mid \tilde{d},\lambda) = \frac{\left[(1-\tilde{d})\lambda T_2\right]^{n_2}}{n_2!} \exp\left[-(1-\tilde{d})\lambda T_2\right]. \tag{37}$$

Using the distributions defined in Eqs. 36 and 37 in Eq. 35 yields

$$p\left(n_2 \mid \tilde{d}\right) = \frac{\left[(1-\tilde{d})T_2\right]^{n_2}\left(\frac{1}{\beta}+T_1\right)^{\alpha+n_1}\Gamma(\alpha+n_1+n_2)}{n_2!\,\Gamma(\alpha+n_1)\left[\frac{1}{\beta}+T_1+(1-\tilde{d})T_2\right]^{\alpha+n_1+n_2}} \tag{38}$$

This likelihood terms in the risk definitions are then found through summation with respect to the c allowable failures.

$$p\left(n_2 \leq c|\tilde{d}\right) = \sum_{i=0}^{c} \frac{\left[(1-\tilde{d})T_2\right]^{i}\left(\frac{1}{\beta}+T_1\right)^{\alpha+n_1}\Gamma(\alpha+n_1+i)}{i!\,\Gamma(\alpha+n_1)\left[\frac{1}{\beta}+T_1+(1-\tilde{d})T_2\right]^{\alpha+n_1+i}} \tag{39}$$

$$p\left(n_2 > c|\tilde{d}\right) = 1-\sum_{i=0}^{c} \frac{\left[(1-\tilde{d})T_2\right]^{i}\left(\frac{1}{\beta}+T_1\right)^{\alpha+n_1}\Gamma(\alpha+n_1+i)}{i!\,\Gamma(\alpha+n_1)\left[\frac{1}{\beta}+T_1+(1-\tilde{d})T_2\right]^{\alpha+n_1+i}} \tag{40}$$

The posterior risks in Eqs. 32 and 33 can then be calculated using the distributions defined in Eqs. 34, 39, and 40. Further simplification of the risks is not necessary, because numerical integration can be used to evaluate the integrals in the definitions.

$$
p\left(\tilde{d} \le FEF \mid n_2 \le c\right) = \frac{\int_0^{FEF} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \tilde{d}^{a-1}\left(1-\tilde{d}\right)^{b-1} \sum_{i=0}^{c} \frac{\left[(1-\tilde{d})T_2\right]^i \left(\frac{1}{\beta}+T_1\right)^{\alpha+n_1} \Gamma(\alpha+n_1+i)}{i!\,\Gamma(\alpha+n_1)\left[\frac{1}{\beta}+T_1+(1-\tilde{d})T_2\right]^{\alpha+n_1+i}}\, d\tilde{d}}{\int_0^{1} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \tilde{d}^{a-1}\left(1-\tilde{d}\right)^{b-1} \sum_{i=0}^{c} \frac{\left[(1-\tilde{d})T_2\right]^i \left(\frac{1}{\beta}+T_1\right)^{\alpha+n_1} \Gamma(\alpha+n_1+i)}{i!\,\Gamma(\alpha+n_1)\left[\frac{1}{\beta}+T_1+(1-\tilde{d})T_2\right]^{\alpha+n_1+i}}\, d\tilde{d}} \tag{41}
$$

$$
p\left(\tilde{d} > FEF \mid n_2 > c\right) = \frac{\int_{FEF}^{1} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \tilde{d}^{a-1}\left(1-\tilde{d}\right)^{b-1} \left[1 - \sum_{i=0}^{c} \frac{\left[(1-\tilde{d})T_2\right]^i \left(\frac{1}{\beta}+T_1\right)^{\alpha+n_1} \Gamma(\alpha+n_1+i)}{i!\,\Gamma(\alpha+n_1)\left[\frac{1}{\beta}+T_1+(1-\tilde{d})T_2\right]^{\alpha+n_1+i}}\right] d\tilde{d}}{1 - \int_0^{1} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \tilde{d}^{a-1}\left(1-\tilde{d}\right)^{b-1} \sum_{i=0}^{c} \frac{\left[(1-\tilde{d})T_2\right]^i \left(\frac{1}{\beta}+T_1\right)^{\alpha+n_1} \Gamma(\alpha+n_1+i)}{i!\,\Gamma(\alpha+n_1)\left[\frac{1}{\beta}+T_1+(1-\tilde{d})T_2\right]^{\alpha+n_1+i}}\, d\tilde{d}} \tag{42}
$$

There are multiple ways to find the verification test length and allowable failures to meet desired posterior consumer and producer risk thresholds. The simplest is to use a two-step process:

1. For the initial minimum number of allowable failures, solve for the test length that meets the posterior consumer risk threshold.
2. Calculate the posterior producer risk for the test length and allowable failures from step 1. If the risk meets the desired threshold, use the test length and allowable failures. Otherwise, increase the number of allowable failures by 1 and return to step 1.

4.3 Verification Testing Guidance

There is a direct connection between the approaches in Sections 4.1 and 4.2. The consumer risk definition in Eq. 41 will be equivalent to the posterior percentile defined in Eq. 31 when the number of allowable failures c is set to zero. If the number of allowable failures is greater than zero, the posterior percentile for the desired FEF with c failures will be larger than the chosen consumer risk threshold. This relationship can be used to provide practical guidance for developing a verification test.

The posterior risks can be used to control the probabilities of making the incorrect decision from the test. The risks are defined to cover a range of possible outcomes, which is useful to understand for planning the test. However, the test will have a single outcome. It is generally desirable that the outcome meets a certain posterior percentile, meaning that the posterior probability of the FEF being below a given value is sufficient. Using the posterior risks alone will not meet this condition.

To demonstrate this issue, assume there are six occurrences of a failure mode in 500 h of testing prior to the corrective action being implemented with a desired FEF of 0.70. For desired posterior consumer and producer risks of 0.20, using a Beta(1,1) prior for the FEF results in a verification test of 915 h with one allowable failure. The posterior consumer risk is 0.20, and the posterior producer risk is 0.16. But using Eq. 31, the posterior percentile for the FEF of 0.70 is 0.29. This means that the test can be passed successfully with one failure using risks of 0.20, but the posterior probability of the FEF being less than 0.70 is 0.29. Understanding these differences in the risks and posterior percentiles is important, as the resulting posterior probability may not be sufficient.

An alternative approach combining the two methods can be used to provide more practical verification tests in this setting. Step 1 from Section 4.2 can be altered slightly to use the posterior percentile directly, which guarantees that the posterior results will meet desired thresholds and the risks will also be adequately controlled. The altered steps are as follows:

1. For the initial minimum number of allowable failures, solve for the test length that meets the desired posterior percentile threshold.
2. Calculate the posterior producer risk for the test length and allowable failures from step 1. If the risk does not meet the desired threshold, increase the number of allowable failures by 1 and return to step 1. Otherwise, use the test length and allowable failures and calculate the posterior consumer risk.

Using this approach on the previous example results in a verification test that is slightly longer at 1,191 h with one allowable failure. The posterior percentile with one failure is 0.20, and it decreases to 0.07 if zero failures are observed. The posterior consumer risk is 0.13, and the posterior producer risk is 0.19. This approach yields a slightly longer verification test, but it has the added benefit of controlling the risks while also keeping posterior probabilities at desired levels.

5. CONCLUSIONS

The methodology presented here provides a way to assess the effectiveness of corrective actions in a Bayesian framework. The posterior uses the data from both before and after the corrective action is implemented. Prior distributions are used for the failure rate and FEF. A non-informative Gamma prior is used for the failure rate, and a Beta distribution is used for the FEF, which can be non-informative or informative. The flexibility of the Beta prior allows for any relevant information about the corrective action to be included in the assessment.

An additional benefit of the methodology is that it can be used to plan for efficient tests to verify the FEF after the corrective action has been implemented. Posterior risks are developed in an assurance testing framework, which again allows for any relevant prior information to be used. The test design approach using both the posterior percentile and the posterior risks provides a method for obtaining desired posterior results from the test while also controlling the probability of making an incorrect decision with the posterior risks. The verification testing provides a useful tool to help decision makers make efficient choices regarding testing, and it also helps to better inform them on the effectiveness of the corrective actions that have been implemented.

6. REFERENCES

1. Trapnell, B. (1983). *Reliability fix effectiveness factors for Army systems* (TR-388). U.S. Army Materiel Systems Analysis Activity.

2. Trapnell, B. (1984). *Update on reliability fix effectiveness factors* (Interim Note R-86). U.S. Army Materiel Systems Analysis Activity.

3. Crow, L. (1982). *An improved methodology for reliability growth projection* (TR-357). U.S. Army Materiel Systems Analysis Activity.

4. Ellner, P., & Wald, L. (1995). *AMSAA maturity projection model*. Proceedings of the Annual Reliability and Maintainability Symposium.

5. Nelson, W. (1982) *Applied life data analysis.* John Wiley & Sons.

6. Wayne, M., & Modarres, M. (2015). A Bayesian model for complex system reliability growth under arbitrary corrective actions. *IEEE Transactions on Reliability*, 64–1.

7. Sivia, D. (2006). *Data analysis – a Bayesian tutorial*. Oxford University Press.

8. Jaynes, E. (2003). *Probability theory - the logic of science*. Cambridge University Press.

9. Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6), 721–741.

10. Neal, R. (2003). Slice sampling. *Annals of Statistics*, 31–3.

11. Meeker, W., & Escobar, L. (2004). Reliability: The other dimension of quality. *Quality Technology and Qualitative Management*,*1*, 1–25.

12. Hamada, M., Wilson, A., Shane Reese, C., & Martz, H. (2010). *Bayesian reliability*. Springer Series in Statistics.

LIST OF ACRONYMS

AMSAA	U.S. Army Materiel Systems Analysis Activity
DAC	DEVCOM Analysis Center
DEVCOM	U.S. Army Combat Capabilities Development Command
FEF	Fix Effectiveness Factor
MAP	maximum a posteriori
MCMC	Markov Chain Monte Carlo