Bayesian Methods for Evaluating Discrete Reliability Growth

J. Brian Hall, Ph.D., U.S. Army Evaluation Center
Ali Mosleh, Ph.D., University of Maryland

Key Words: Bayesian, management metrics, posterior, projection, reliability growth

SUMMARY & CONCLUSIONS

Bayesian estimation procedures are derived herein that may be utilized to evaluate reliability growth of discrete systems, such as guns, rockets, missile systems, torpedoes, etc. One of the advantages of these Bayesian procedures is that they directly quantify the epistemic uncertainties in model parameters (i.e., the shape parameters of the beta distribution), as well as six reliability growth metrics of basic interest to program management. These metrics include: (1) the initial system reliability; (2) the projected reliability following failure mode mitigation; (3) reliability growth potential (i.e., the theoretical upper-limit on reliability achieved by finding and fixing all failure modes via a specified level of fix effectiveness); (4) the expected number of failure modes observed during testing; (5) the probability of observing a new failure mode; and (6) the fraction of the initial system probability of failure associated with failure modes for which program management is aware. These metrics give reliability practitioners the means to estimate the reliability of discrete systems undergoing development, address model goodness-of-fit concerns, quantify programmatic risk, and assess system maturity.

Analytical results are presented to obtain Bayes’ estimates of the beta shape parameters under a delayed corrective action strategy (i.e., when corrective actions are installed on system prototypes at the end of the current test phase). A Monte Carlo simulation approach is given for constructing uncertainty distributions on each of the aforementioned reliability growth management metrics. Bayesian probability limits for inference on interval estimation are obtained in the usual manner (i.e., via desired percentiles of the uncertainty distributions). These uncertainty distributions are found to be approximated very well by beta and/or Gaussian random variables. These methods are illustrated by simple numerical examples. In particular Bayes’ estimates the beta shape parameters are obtained from a small dataset, and compared against the true parameter values. Bayesian epistemic uncertainty distributions are also constructed for the reliability growth management metrics via the proposed Monte Carlo approach. This methodology is useful to program managers and reliability practitioners that wish to quantitatively assess the reliability growth program of one-shot systems developed under a delayed corrective action strategy.

ACRONYMS

AEC – Army Evaluation Center
AMSAA – Army Materiel Systems Analysis Activity
CAP – Corrective Action Period
FEF – Fix Effectiveness Factor(s)
FOT – First Occurrence Trial(s)
MME – Method of Moments Estimation/Estimate(s)
MLE – Maximum Likelihood Estimation/Estimate(s)
RG – Reliability Growth

NOTATION

\( k \) – total number of potential failure modes
\( m \) – total number of observed failure modes
\( n \) – beta shape parameter for pseudo trials
\( x \) – beta shape parameter for pseudo failures
\( T \) – total number of trials
\( N_{ij} \) – number of failures for mode \( i \) in trial \( j \)
\( N_i \) – total number of failures for mode \( i \)
\( p_i \) – prob. of occurrence of failure mode \( i \)
\( d_i \) – FEF for failure mode \( i \)
\( t_i \) – FOT of failure mode \( i \)
\( L(\hat{N} \mid x, n) \) – joint marginal likelihood function
\( \pi_0(x, n) \) – prior PDF for the beta parameters
\( \pi(x, n \mid \hat{N}) \) – posterior PDF for the beta parameters

1 INTRODUCTION

1.1 Background

The term Reliability Growth (RG) refers to the elimination of design weaknesses inherent to intermediate prototypes of complex systems via the failure mode discovery, analysis, and effective correction process [19]. A wealth of models have been published over the years to plan, track, and project reliability improvements of developmental items whose test durations are continuous, as well as discrete. The topic of this paper deals with the area of discrete RG projection. The purpose of this area of research is to quantify the reliability that could be achieved if failure modes observed during testing are corrected via a specified level of fix effectiveness. Fix Effectiveness Factors (FEF) reduce initial probabilities (or rates) of occurrence of individual failure modes by a fractional amount, thereby quantifying corrective action effectiveness and associated RG thereof. Specifically, FEF represent the fraction reduction in a failure mode initial
failure probability (or rate) of occurrence as a result of the engineering redesign / correction effort.

As the development program is ongoing, program management can adopt, in general, one of two corrective action strategies to mitigate failure modes discovered during system prototype testing. The first addresses the case where all corrective actions are delayed until the conclusion of the current test phase. In this scenario, the system configuration remains static over duration of the test, and corrective actions are installed in a block of time typically referred to as a Corrective Action Period, or CAP. The second approach addresses the case where management adopts an arbitrary corrective action strategy, which results in a mixture of instantaneous fixes (e.g., installed during test), as well as delayed fixes. Thus, the system configuration from a reliability standpoint is dynamic (i.e., presumably improving) during the current test phase. Over the years, these two correction strategies have resulted in the development of two types of RG projection models. The genesis of the area of RG projection is marked by a paper written by Corcoran, Weingarten, and Zehna in 1964 [1], which addresses the delayed case for discrete systems. Since then, a number of other methods have been developed. They include the delayed models [4-5, 11-12, 16-19], and the delayed / non-delayed models given in [8], [13], and [19].

A vast amount of literature is available covering a wide variety of RG analysis capabilities. More general references for additional background material (on planning, tracking, and projection models) are as follows. The AMSAA Reliability Growth Methodology Guide given by Ellner, et al. [10], and the Fries-Sen survey of discrete RG models [9] are among the most comprehensive and detailed works on the subject. Military Handbook 189 [3], while outdated, and Appendix C of the DoD Guide for Achieving RAM [15] are also good references that outline methods available for RG analysis. Hall [19] also gives a compressive, higher-level, literature review of nearly 200 papers which cover RG planning models, tracking models, projection models, RG surveys, RG handbooks, and other works on related theoretical results, simulation studies, real-world applications, personal perspectives, international standards [14], and related statistical procedures.

1.2 Overview

This paper is organized as follows. The proposed methodology is given in Section 2, and consists of a list of assumptions and a Bayesian estimation procedure to approximate model parameters (i.e., the shape parameters of the beta distribution). A Monte Carlo simulation approach is also outlined for constructing epistemic uncertainty distributions on each the RG management metrics previously mentioned, and presented in [18]. The proposed techniques are illustrated via simple numerical example in Section 3.

2 METHODOLOGY

2.1 Assumptions

1. The distribution of the number of failures in T trials for each failure mode is binomial (i.e., the conditional distribution, \( N_i \sim \text{Binomial}(T, p_i) \)) on each failure mode \( i = 1, \ldots, k \).
2. Initial failure probabilities \( p_1, \ldots, p_k \) constitute a realization of an iid random sample \( P_1, \ldots, P_k \) such that \( P_i \sim \text{Beta}(n, x) \) for each \( i = 1, \ldots, k \). The PDF parameterization utilized herein is given as,

\[
f(p) = \begin{cases} 
\frac{\Gamma(n)}{\Gamma(x)\Gamma(n-x)} p_i^{x-1} \cdot (1-p_i)^{n-x-1} & p_i \in (0,1) \\
0 & \text{otherwise}
\end{cases}
\]  

(1)

where \( n \) represents pseudo trials, \( x \) represents pseudo failures, and \( \Gamma(x) = \int_0^\infty e^{-t} \cdot t^{x-1} dt \) denotes the Euler gamma function.

3. Corrective actions are delayed until the end of the current test phase, where a test phase is considered to consist of a sequence of \( T \) independent Bernoulli trials.

4. There is at least one repeat failure mode. If there is not at least one repeat failure mode, the moment estimators and the likelihood estimators of the beta shape parameters do not exist.

2.2 Method of Moments Estimation

The numerical integration required to obtain Bayesian estimates of the beta parameters, as defined herein, utilize the MLE of the shape parameter \( n \) (or a multiple thereof) as a general guide for the limit of integration. Similarly, the MLE use the MME as starting-values for the associated numerical (root finding) routine. Thus, the assessment procedure begins by obtaining the MME. The moment estimators for the beta shape parameters [7], per the special case considered (i.e., where all failure probabilities are estimated via the same number of trials), are given as

\[
\hat{n}_i = \frac{\bar{p}_i - m_i^2}{m_i^2 - \frac{\bar{p}_i}{T} \left( 1 - \frac{1}{T} \right) \bar{p}_i^2}
\]  

(2)

and

\[
\hat{x}_i = \bar{p}_i \cdot \bar{p}_i
\]  

(3)

where \( \bar{p}_i = \sum_{i=1}^k \hat{p}_i / k \) and \( m_i^2 = \sum_{i=1}^k \hat{p}_i^2 / k \) are the unweighted first and second sample moments, respectively.

2.3 Maximum Likelihood Estimation

When corrective actions are delayed until the end of the current test phase, the failure probabilities \( p_i \) generating the failures \( N_i \) for \( i = 1, \ldots, k \) are not reduced by their corresponding FEF (e.g., the \( d_i \)) during test. In this case, the marginal distribution of an each observation (e.g., \( N_i \) ) [7] is,

\[
f(N_i \mid x, n) = \frac{T! \cdot \Gamma(n) \cdot \Gamma(N_i + x) \cdot \Gamma(T + n - N_i - x)}{N_i! \cdot \Gamma(T - N_i) \cdot \Gamma(T + n) \cdot \Gamma(x) \cdot \Gamma(n - x)}
\]  

(4)

The product of these terms, for each failure mode, is the joint marginal likelihood function of the entire sample,
\[ L(\vec{N} \mid x, n) = \prod_{i=1}^{k} f(N_i \mid x, n) = \prod_{i=1}^{k} \frac{T! \cdot \Gamma(n_i + x) \cdot \Gamma(T + n - N_i - x)}{N_i! \cdot (T - N_i)! \cdot \Gamma(T + x) \cdot \Gamma(n - x)} \]  

Two important assumptions associated with the derivation of (5) is that the failure probability, and number of failures observed for each failure mode (e.g., the conditional distributions) follow the beta, and binomial distributions, respectively. Thus, (4) is sometimes referred to as the marginal beta-binomial distribution, and is a popular doubly-stochastic process utilized in reliability modeling.

The method of marginal maximum likelihood provides estimates of the beta parameters \( n \) and \( x \) that maximize the beta marginal likelihood function (5). For an assumed number of total potential failure modes, \( k \), the ML estimates are obtained by numerically solving the following two likelihood equations simultaneously

\[ \sum_{i=1}^{k} \sum_{j=0}^{N_i - 1} \left( \frac{1}{\hat{n}_i - x + j} \right) - \sum_{i=1}^{k} \sum_{j=0}^{x - 1} \left( \frac{1}{\hat{n}_i - x + j} \right) = 0 \]  

and

\[ \sum_{i=1}^{k} \sum_{j=0}^{N_i - 1} \left( \frac{1}{\hat{n}_i - x + j} \right) - \sum_{i=1}^{k} \sum_{j=0}^{T - N_i + 1} \left( \frac{1}{\hat{n}_i + j} \right) = 0 \]  

Equations (6) and (7) are defined to be zero for unobserved failure modes (i.e., if \( N_i = 0 \)). The starting values for the associated numerical routine to obtain such estimates can be chosen as the unweighted moment estimators given by (2) and (3).

2.4 Bayesian Estimation

To construct a joint posterior distribution on the beta shape parameters, a natural prior selection, at least initially, would be uniform,

\[ \pi_0(x, n) = \begin{cases} \frac{1}{n} & \text{for } n \geq x \\ 0 & \text{otherwise} \end{cases} \]  

If the failure probabilities \( p_1, \ldots, p_k \) are assumed to constitute a realization of an iid sample \( p_1, \ldots, p_k \) such that \( P_i \sim \text{Beta}(n, x) \), the joint posterior based on the entire sample (for a uniform prior) is,

\[ \pi(x, n \mid \vec{N}) = \frac{\int_{x=0}^{\infty} \int_{n=0}^{\infty} L(\vec{N} \mid x, n) \cdot \pi_0(x, n) dx \, dn}{\int_{x=0}^{\infty} \int_{n=0}^{\infty} L(\vec{N} \mid x, n) \cdot \pi_0(x, n) dx \, dn} \]  

\[ = \frac{\prod_{i=1}^{k} \frac{\Gamma(N_i + x) \cdot \Gamma(T + n - N_i - x)}{N_i! \cdot (T - N_i)!}}{\int_{x=0}^{\infty} \int_{n=0}^{\infty} \frac{\Gamma(N_i + x) \cdot \Gamma(T + n - N_i - x)}{N_i! \cdot (T - N_i)!} dx \, dn} \]  

Under squared-error loss, the Bayes’ estimates of \( n \) and \( x \) are calculated via mathematical expectations of the posterior distribution. The Bayes’ estimates are given as,

\[ \bar{n} = E_n[\pi(x, n \mid \vec{N})] = \int_{n=0}^{\infty} \int_{x=0}^{\infty} n \cdot \pi(x, n \mid \vec{N}) dx \, dn \]  

and

\[ \bar{x} = E_x[\pi(x, n \mid \vec{N})] = \int_{n=0}^{\infty} \int_{x=0}^{\infty} x \cdot \pi(x, n \mid \vec{N}) dx \, dn \]  

respectively. In order to perform the required numerical integration to obtain these estimates, integration for the beta shape parameter \( n \) may be carried-out over the interval \( n \in (0, \hat{n}_i) \) (or a multiple thereof). Thus, the MLE \( \hat{n}_i \) is used as a general guide to define the parameter space. From a numerical implementation standpoint, the theoretical parameter space \( n \in (0, \infty) \) may not be practical, or lend itself to efficient computation.

2.5 Monte Carlo Simulation

In the preceding section, Bayesian estimates of the beta shape parameters were obtained to approximate the initial prior PDF of the failure probabilities inherent to a complex one-shot system. This PDF can be used to construct epistemic uncertainty distributions for each of the management metrics given in [18]. These metrics are summarized below with equations numbers from the earlier publication given for cross-reference. The Monte Carlo simulation approach is outlined as follows:

1. Use the Bayesian estimation procedures in the previous section to estimate the shape parameters of the beta distribution. This distribution represents an uncertainty distribution for the failure probabilities inherent to the system.

2. Stochastically generate a size \( k \) vector of failure probabilities \( p_1, \ldots, p_k \) from the distribution constructed in the previous step.

3. For each \( p_i \) for \( i = 1, \ldots, k \), simulate a failure history by generating trial outcomes (i.e., 0 or 1) from a Bernoulli distribution with parameter \( p_i \). The trial outcomes indicate either the occurrence or nonoccurrence of each failure mode in each trial \( j = 1, \ldots, T \). The count data \( N_j \) for each failure mode are obtained by summing the trial outcomes over all \( T \) trials, for each failure mode \( p_1, \ldots, p_k \). The index-set of observed failure modes \( \text{obs} \equiv \{ i : N_i > 0 \text{ for } i = 1, \ldots, k \} \) is the set of indices associated with failure modes observed during simulation. The total number of observed failure modes is denoted by \( m = |\text{obs}| \). The FOT \( t_i \) for \( i \in \text{obs} \) are obtained as the trial numbers when failure modes are first discovered.

4. Stochastically generate a size \( k \) vector of FEF \( d_1, \ldots, d_k \) from a beta distribution. The example given in the following section draws FEF via a beta distribution with a mean of 0.80 and a variance of 0.01.

5. Use the data obtained in the previous steps to calculate the management metrics of interest (summarized below).

6. Repeat steps 1-5 several times (e.g., \( r = 10,000 \) replications) to stochastically generate values of the management metrics from which their empirical distributions may be constructed.

In the example below, uncertainty distributions are constructed for all of the management metrics given in [18]. Note that the associated equations are derived from a shrinkage factor estimator discussed in [17], and [2]. The management metrics and associated model equations include:
the expected initial reliability of the system (obtained from Equation (8) in [18]),

\[ R_i(t = 1 | \hat{p}) = \prod_{i=1}^{k} (1 - p_i) \]  

(12)

the expected reliability of the corrected system on trial \( T \) (i.e., Equation (8) in [18]),

\[ R_c(t = 1 | \tilde{p}) = \prod_{i=1}^{k} \left[ 1 - \left(1 - \left[1 - (1 - p_i)^{T-1}\right] \cdot d_i \cdot p_i \right) \right] \]  

(13)

the reliability growth potential [6] (i.e., the theoretical upper-limit on reliability that is achieved if all failure modes are found and corrected via a specified level of fix effectiveness, computed by Equation (11) in [18]),

\[ R_{gr}(\hat{p}) = \prod_{i=1}^{k} \left[ 1 - (1 - d_i) \cdot p_i \right] \]  

(14)

the expected number of unique failure modes observed on or before trial \( T \) (i.e., observed before the end of the test phase, computed via Equation (13) in [18]),

\[ \mu_T(T | \tilde{p}) = k - \sum_{i=1}^{k} (1 - p_i)^T \]  

(15)

the expected probability of observing a new failure mode on trial \( T \) (Equation (16) in [18]),

\[ h_T(T | \tilde{p}) = 1 - \prod_{i=1}^{k} \left[ 1 - (1 - p_i)^{T-1} \cdot p_i \right] \]  

(16)

the expected probability of failure on trial \( T \) due to a repeat failure mode, expressed as a fraction of the initial system unreliability (Equation (20) in [18]),

\[ \phi_T(T | \tilde{p}) = \frac{1 - \prod_{i=1}^{k} \left[ 1 - (1 - p_i)^{T-1} \cdot p_i \right]}{1 - \prod_{i=1}^{k} (1 - p_i)} \]  

(17)

After the simulation is replicated, one may use the data generated on each metric to construct a histogram scaled to area unity (e.g., empirical distribution). Beta and normal PDF approximations to the histogram can be obtained by estimating distribution parameters from the sample mean and sample variance of the data, or functions thereof. For example, the beta parameter estimates can be calculated as,

\[ \hat{n} = \frac{\bar{x} - \bar{x}^2 - s^2}{s^2} \]  

(18)

and

\[ \hat{x} = \hat{n} \cdot \bar{x} \]  

(19)

where \( \bar{x} \) and \( s \) represent the mean and standard deviation, respectively, of the data generated for a given metric. The Bayesian point-estimates and probability limits are obtained in the usual manner (e.g., as the mean and desired percentiles of the distribution, respectively).

3 NUMERICAL EXAMPLE

3.1 Bayesian Parameter Estimates

Since the true values of the beta shape parameters are unknown in practice, Monte Carlo simulation is used to stochastically generate the data from which the proposed Bayes estimates are obtained and compared against the true parameters. In this example, the system is comprised of \( k = 20 \) failure modes and the system is tested for \( T = 50 \) trials. The mean and variance of the beta distribution used to generate the failure mode probabilities of occurrence are \( \mu = 0.025 \) and \( \sigma^2 = 0.0025 \), respectively. During simulation, \( m = 7 \) (of 20) failure modes were observed with a combined total of \( N = 24 \) failures. The failure data shown in Table 1. The second column shows the failure mode numbers comprising the index-set of observed failure modes. The third column gives the true probability of failure (stochastically generated via a beta r.v.) for each of these observed failure modes. The number of failures, trials of first occurrence, and individual FEF are shown in columns 4-6, respectively.

<table>
<thead>
<tr>
<th>Failure Mode</th>
<th>Probability</th>
<th>Failures</th>
<th>FOT</th>
<th>FEF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0211</td>
<td>2</td>
<td>11</td>
<td>0.8874</td>
</tr>
<tr>
<td>2</td>
<td>0.0172</td>
<td>1</td>
<td>27</td>
<td>0.8010</td>
</tr>
<tr>
<td>3</td>
<td>0.1016</td>
<td>10</td>
<td>7</td>
<td>0.8613</td>
</tr>
<tr>
<td>4</td>
<td>0.0195</td>
<td>1</td>
<td>30</td>
<td>0.8958</td>
</tr>
<tr>
<td>5</td>
<td>0.0286</td>
<td>2</td>
<td>17</td>
<td>0.5812</td>
</tr>
<tr>
<td>6</td>
<td>0.1187</td>
<td>7</td>
<td>4</td>
<td>0.6189</td>
</tr>
<tr>
<td>7</td>
<td>0.0164</td>
<td>1</td>
<td>32</td>
<td>0.9222</td>
</tr>
</tbody>
</table>

Table 1 - Failure Data

For a delayed corrective action strategy, estimates of the beta shape parameters are obtained using the MME, MLE, as well as the Bayesian procedures given above. These parameter estimates (which are based on the count data shown in column 4 of Table 1) are given in columns 3-5 of Table 2, respectively. The true values of the parameters are given in column 2. The MME estimates for \( n \) and \( x \), given in column 3, are obtained via (2) and (3). The MLE estimates, shown in column 4, are obtained as the numerical solutions to Equations (6) and (7) when solved simultaneously. The Bayesian estimates of \( n \) and \( x \) are obtained by Equations (10) and (11) respectively.

<table>
<thead>
<tr>
<th>True</th>
<th>MME</th>
<th>MLE</th>
<th>Bayes</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>8.750</td>
<td>9.650</td>
<td>9.206</td>
</tr>
<tr>
<td>( x )</td>
<td>0.219</td>
<td>0.232</td>
<td>0.220</td>
</tr>
</tbody>
</table>

Table 2 - Parameter Estimates

Notice that the Bayesian estimation procedure requires the beta shape parameter \( n \) (i.e., pseudo trials) to be theoretically integrated over the entire parameter space \( n \in (0, \infty) \). The practical parameter space (i.e., as opposed to the theoretical) from which the numerical integration can be carried-out includes realistic values for the parameter. For example, if \( n \) represents the true value of the parameter, the parameter space should cover the interval \((0,n)\). Since the true value of the parameter is unknown, one can use either the MME, MLE, or a multiple thereof. These results were obtained by integrating over the parameter space \( n \in (0,1.5 \cdot \hat{n}_k) \), where \( \hat{n}_k \) is the finite \( k \) MLE estimate for the beta parameter \( n \) given in column 4 of Table 2. For some
stochastically generated datasets, this MLE can be lower than the true value of \( n \). This is the reason why it may be desirable to perform the numerical integration over a multiple beyond \( \hat{n} \). This endeavors to ensure the entire volume under the surface (9) is accounted for numerically.

Figure 1 shows the beta PDF approximations using the parameter estimates given in Table 2. Not surprisingly, the PDF approximations based on classical versus Bayesian methods are very close to one another.

![Beta PDF Approximations](image)

**Figure 1 - Beta PDF Approximations**

### 3.2 Uncertainty Distributions

Continuing from the example presented in Section 3.1, a size \( k = 20 \) realization of failure probabilities \( p_1, \ldots, p_k \) were stochastically drawn from a beta distribution whose Bayes’ estimates are shown in column 5 of Table 2. Using these failure probabilities, failure histories for each failure mode were generated from a Bernoulli random variable. The index-set of observed failure modes were defined uniquely for each replication. A size \( k = 20 \) realization of FEF were then generated from a beta distribution. Only failure probabilities associated with the observed failure modes are reduced by their corresponding FEF during each replication. Unobserved failure modes are assigned a zero FEF. Using these data, estimates of each of the management metrics were

![Initial Reliability](image)

**Figure 2 – Initial Reliability**

![Projected Reliability](image)

**Figure 3 – Projected Reliability**

![Reliability Growth Potential](image)

**Figure 4 – Reliability Growth Potential**

![Expected Number of Failure Modes](image)

**Figure 5 – Expected Number of Failure Modes**

![Probability of New Failure Mode](image)

**Figure 6 – Probability of New Failure Mode**
calculated. This process was replicated $r=10,000$ times during simulation. Figures 2-7 show the empirical densities and corresponding PDF approximations for each of the management metrics that were constructed. Each of the empirical densities were found to be closely represented by a beta r.v., with exception to Figure 5, which was closely represented by a Gaussian r.v. The point and interval estimates shown were obtained as the mean, and as the 10th and 90th percentiles of the uncertainty distributions, respectively. Thus, these Bayesian probability limits are analogous to the notion of an 80% confidence interval.

![Figure 7 – Probability of Failure due to a Repeat Mode](image-url)

REFERENCES


BIOGRAPHIES

J. Brian Hall, Ph.D.
U.S. Army Evaluation Center
4120 Susquehanna Avenue, ATTN: TEAE-REG
Aberdeen Proving Ground, Maryland 21005, USA
e-mail: brian.hall@us.army.mil

A recent biography is given in [19].

Ali Mosleh, Ph.D.
University of Maryland, College Park
Department of Mechanical Engineering
2181 Glenn L. Martin Hall, Room 0151-F
College Park, Maryland 20742, USA
e-mail: mosleh@umd.edu

A recent biography is given in [18].