# ABSTRACT

Title of dissertation: METHODOLOGY FOR EVALUATING RELIABILITY GROWTH PROGRAMS OF DISCRETE SYSTEMS.

J. Brian Hall, Ph.D., 2008

Dissertation directed by: Professor Ali Mosleh, Ph.D., Department of Mechanical Engineering

The term Reliability Growth (RG) refers to the elimination of design weaknesses inherent to intermediate prototypes of complex systems via failure mode discovery, analysis, and effective correction. A wealth of models have been developed over the years to plan, track, and project reliability improvements of developmental items whose test durations are continuous, as well as discrete. This research reveals capability gaps, and contributes new methods to 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 reduce initial probabilities (or rates) of occurrence of individual failure modes by a fractional amount, thereby increasing system reliability.

The contributions of this research are as follows. New RG management metrics are prescribed for one-shot systems under two corrective action strategies. The first is when corrective actions are delayed until the end of the current test phase. The second is when they are applied to prototypes after associated failure modes are first discovered. These management metrics estimate: initial system reliability, projected reliability (i.e., reliability after failure mode mitigation), RG potential, the expected number of failure modes observed during test, the probability of discovering new failure modes, and the portion of system unreliability associated with repeat failure modes. These management metrics give practitioners the means to address model goodness-of-fit concerns, quantify programmatic risk, assess reliability maturity, and estimate the initial, projected, and upper achievable reliability of discrete systems throughout their development programs.

Statistical procedures (i.e., classical and Bayesian) for point-estimation, confidence interval construction, and model goodness-of-fit testing are also developed. In particular, a new likelihood function and maximum likelihood procedure are derived to estimate model parameters. Limiting approximations of these parameters, as well as the management metrics, are also derived. The features of these new methods are illustrated by simple numerical example. Monte Carlo simulation is utilized to characterize model accuracy. This research is useful to program managers and practitioners working to assess the RG program and development effort of discrete systems.

**THIS PAGE IS INTENTIONALLY LEFT BLANK.**

METHODOLOGY FOR EVALUATING RELIABILITY GROWTH
PROGRAMS OF DISCRETE SYSTEMS

by

J. Brian Hall

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of:
Doctorate of Philosophy
2008

Advisory Committee:

- Professor Ali Mosleh, Committee Chair, and Advisor.
- Dr. Paul M. Ellner, Special Member, and Co-advisor.
- Associate Professor Elise Miller-Hooks, Dean's Representative.
- Professor Shapour Azarm.
- Professor Mohammad Modarres.
- Associate Professor Michel Cukier.

## - DEDICATION -

This work is dedicated to my loving wife, Alyson, who patiently tolerated me for these long years in completing my Doctorate Degree.

In that book which is
My memory …
On the first page
That is the chapter when
I first met you
Appear the words …
Here begins a new life

*– Dante Alighieri, La Vita Nuova*

Saturday, 30 June 2007, Baltimore, Maryland

THIS PAGE IS INTENTIONALLY LEFT BLANK.

# - ACKNOWLEDGEMENTS -

The author would like to extend his appreciation to his Ph.D. advisor and committee for their time, support, and assessment of this research:

- Professor Ali Mosleh, Committee Chair and Advisor.
- Paul M. Ellner, Ph.D., Special Member and Co-advisor.
- Associate Professor Elise Miller-Hooks, Dean's Representative.
- Professor Shapour Azarm.
- Professor Mohammad Modarres.
- Associate Professor Michel Cukier.

Over the last three years, several chapters of this dissertation (i.e., Chapters 4, 5, 6, and 7) were prepared and submitted to various refereed journals. The author would like to give special thanks to the following individuals for their detailed review of these papers:

- Paul M. Ellner, Ph.D., U.S. Army Materiel Systems Analysis Activity.
- Arthur Fries, Ph.D., Institute for Defense Analyses.
- Michael J. Cushing, Ph.D., U.S. Army Evaluation Center.

The author expresses grateful appreciation to his Chain-of-Command at the U.S. Army Evaluation Center (AEC) for supporting him both financially, and with a sabbatical during the Spring 2008 semester to complete his Doctorate Degree:

- SES Brian M. Simmons, Director, AEC.
- Frank J. Apicella, Technical Director, AEC.
- Colonel Roderick Burke Sr., Chief of Staff, AEC.
- Stephen P. Yuhas, Director, AEC-RAM.
- Michael J. Cushing, Ph.D., Technical Director, AEC-RAM.
- R. Kenneth Dalton, Division Chief, Ground Combat, AEC-RAM.

The author gives special thanks to his former colleagues at the U.S. Army Materiel Systems Analysis Activity (AMSAA) who gave him equal support:

- SES David J. Shaffer, Director, AMSAA.
- Colonel Philip J. DiSalvo, Deputy Director, AMSAA.
- Clarke J. Fox, Division Chief, Logistics, AMSAA.
- David E. Mortin, Ph.D., Branch Chief, Reliability, AMSAA.
- Jane G. Krolewski, Branch Chief, Mobility and Power & Energy, AMSAA.

I also thank my wife, Alyson, and parents, James and Kathy, for their love and support through the years.

THIS PAGE IS INTENTIONALLY LEFT BLANK.

## - TABLE OF CONTENTS -

**THIS PAGE IS INTENTIONALLY LEFT BLANK.**

## - LIST OF ACRONYMS -

**A**

| | | |
|---|---|---|
| ACPM | – | AMSAA Crow Projection Model |
| AEC | – | Army Evaluation Center |
| ALT | – | Accelerated Life Test |
| AMPM | – | AMSAA Maturity Projection Model |
| AMSAA | – | Army Materiel Systems Analysis Activity |
| ANSI | – | American National Standards Institute |

**C**

| | | |
|---|---|---|
| CAP | – | Corrective Action Period(s) |
| CAE | – | Continuous Analog Estimator(s) |
| CBO | – | Congressional Budget Office |
| CI | – | Confidence Interval |

**D**

| | | |
|---|---|---|
| DoD | – | Department of Defense |

**E**

| | | |
|---|---|---|
| EMATT | – | Expendable Mobile ASW Training Target |
| ESS | – | Environmental Stress Screening |

**F**

| | | |
|---|---|---|
| FD/SC | – | Failure Definition / Scoring Criterion |
| FEF | – | Fix Effectiveness Factor(s) |
| FMEA | – | Failure Mode Effects Analysis |
| FMECA | – | Failure Mode Effects and Criticality Analysis |
| FPRB | – | Failure Prevention and Review Board |
| FRACAS | – | Failure Reporting and Corrective Action System |
| FTA | – | Fault Tree Analysis |

**G**

| | | |
|---|---|---|
| GOF | – | Goodness-of-Fit |

**H**

| | | |
|---|---|---|
| HPP | – | Homogeneous Poisson Process |

**I**

| | | |
|---|---|---|
| IEC | – | International Electrotechnical Commission |

**L**

| | | |
|---|---|---|
| LCB | – | Lower Confidence Bound(s) |
| LCC | – | Life Cycle Cost |
| LS | – | Least Square |

**M**

| | | |
|---|---|---|
| MME | – | Method of Moments Estimation/Estimate(s) |
| ML | – | Maximum Likelihood |
| MLE | – | ML Estimation/Estimate(s) |
| MRBF | – | Mean Rounds Between Failure |
| MS | – | Management Strategy |
| MTBF | – | Mean Time Between Failure |
| MTTF | – | Mean Time To Failure |
| MVF | – | Mean Value Function |

**N**

| | | |
|---|---|---|
| NHPP | – | Non-Homogeneous Poisson Process |

**O**

| | | |
|---|---|---|
| OMS/MP | – | Operational Mode Summary / Mission Profile |
| O&S | – | Operations and Support |
| OSD | – | Office of the Secretary of Defense |

**P**

| | | |
|---|---|---|
| PEO | – | Program Executive Officer |
| PM | – | Program Manager |
| PM2 | – | Planning Model based on Planning Methodology |

**R**

| | | |
|---|---|---|
| RAMS | – | Reliability, Availability, and Maintainability |
| RBD | – | Reliability Block Diagram |
| RDT&E | – | Research, Development, Test and Evaluation |
| RGM | – | Reliability Growth Management |
| RGTMC | – | Reliability Growth Tracking Model Continuous |
| RGTMD | – | Reliability Growth Tracking Model Discrete |
| RPM | – | Reliability Planning Management |
| r.v. | – | Random Variable(s) |

**S**

| | | |
|---|---|---|
| SPLAN | – | System Planning Model |
| SSPLAN | – | Subsystem Planning Model |
| SSTRACK | – | Subsystem Tracking Model |

**T**

| | | |
|---|---|---|
| TAFT | – | Test, Analyze, Fix and Test |
| T&E | – | Test and Evaluation |
| TOC | – | Total Ownership Cost |
| TS | – | Test Statistic |

**V**

| | | |
|---|---|---|
| VGS | – | Visual Growth Suite |

## - LIST OF NOTATION -

| | |
|---|---|
| $k$ | – total number of potential failure modes |
| $m$ | – total number of observed failure modes |
| $T$ | – total number of trials |
| $N_i$ | – total number of failures for mode $i$ in $T$ trials |
| $P_i$ | – beta r.v. denoting the failure probability for failure mode $i$ |
| $p_i$ | – true but unknown probability of failure for mode $i$ |
| $\hat{p}_i$ | – MLE of $p_i$ |
| $\tilde{p}_i$ | – theoretical shrinkage factor estimator for $p_i$ |
| $\theta$ | – true but unknown shrinkage factor |
| $n$ | – beta parameter; pseudo number of trials |
| $x$ | – beta parameter; pseudo number of failures |
| $\breve{n}_k$ | – MME of $n$ |
| $\breve{x}_k$ | – MME of $x$ |
| $\overline{p}_u$ | – unweighted first sample moment of $P_i$ |
| $m_u^2$ | – unweighted second sample moment of $P_i$ |
| $\widehat{n}_k$ | – MLE of $n$ |
| $\widehat{x}_k$ | – MLE of $x$ |
| $d_i$ | – true but unknown FEF for mode $i$ |
| $\chi^2$ | – chi-square test statistic |
| $E_j$ | – expected chi-square GOF frequencies |
| $O_j$ | – observed chi-square GOF frequencies |
| $c$ | – number of cells in chi-square GOF table |
| $R_k\left(t \mid \vec{p}\right)$ | – conditional expected reliability on trial $t$ |
| $\mu_k\left(t \mid \vec{p}\right)$ | – conditional expected number of observed failure modes by trial $t$ |
| $h_k\left(t \mid \vec{p}\right)$ | – conditional expected probability of a new failure mode on trial $t$ |
| $\phi_k\left(t \mid \vec{p}\right)$ | – conditional expected portion of system unreliability associated with repeat failure modes observed by trial $t$ |
| $R_k\left(t\right)$ | – unconditional expectation of $R_k\left(t \mid \vec{p}\right)$ |
| $\mu_k\left(t\right)$ | – unconditional expectation of $\mu_k\left(t \mid \vec{p}\right)$ |
| $h_k\left(t\right)$ | – unconditional expectation of $h_k\left(t \mid \vec{p}\right)$ |
| $\varphi_k\left(t\right)$ | – unconditional expectation of $\phi_k\left(t \mid \vec{p}\right)$ |

THIS PAGE IS INTENTIONALLY LEFT BLANK.

## - LIST OF TABLES -

THIS PAGE IS INTENTIONALLY LEFT BLANK.

## - LIST OF FIGURES -

## - LIST OF FIGURES -

# 1. SUMMARY & CONCLUSIONS

## 1.1. Background

Per Military Handbook 189 [49], reliability growth is defined as the positive improvement in a reliability parameter over a period of time due to changes in product design or manufacturing processes. There are three major areas in the field of reliability growth including: planning, tracking, and projection. Reliability growth planning focuses on the construction of a reliability growth planning curve, which identifies the planned reliability achievement as a function of test duration, in addition to other program resources. Reliability growth tracking focuses on the analysis of a system's current demonstrated reliability. Reliability growth projection focuses on estimating system reliability following implementation of corrective actions to known failure modes. Each of these areas of reliability growth apply to complex systems whose test durations are continuous, as well as to complex systems whose test durations are discrete[1]. A great deal of research has been done over the past several decades in each of these areas. This research is summarized by literature review in Chapter 3, and reveals capability gaps in the area of discrete reliability growth projection. This area of research (i.e., discrete reliability growth projection) is the topic of this dissertation. The contributions of the dissertation are as follows.

## 1.2. Overview of Dissertation and Its Contributions

[1] For simplicity, complex systems whose test durations are discrete will be referred to as one-shot systems. One-shot systems represent (but are not limited to) items such as guns, rockets, missile systems, and torpedoes.

### 1.2.1. Chapter 1[2]

The intention of Chapter 1 is to provide a very high-level introduction of the research topic (i.e., discrete reliability growth projection), outline the organization of this document, and present a very abbreviated review of the contributions given herein. The contributions are summarized by chapter below. Chapters 2-3 give background material, and discuss prior work done in the field of reliability growth. Chapters 4-7 are (self-contained) papers on this research topic that have either been published, accepted for publication, or submitted to journals for potential publication. Chapter 8 gives Bayesian estimation procedures for the beta shape parameters[3], as well as a Monte Carlo simulation approach to construct epistemic uncertainty distributions on the management metrics derived in Chapters 5 and 6.

### 1.2.2. Chapter 2[4]

Chapter 2 provides background material on the fundamentals of reliability growth management, and areas of reliability growth obtained from the literature review given in Chapter 3. The scope of the literature review goes significantly beyond the specific research topic of discrete reliability growth projection, but was required to formulate a holistic view of the state-of-the-art in the field. Naturally, reliability growth projection models are presented in greater detail.

---

[2] References in Chapter 1 are given at the end of the document.
[3] The shape parameters *n* (pseudo trials) and *x* (pseudo failures) of the beta distribution.
[4] References in Chapter 2 are given at the end of the document.

### 1.2.3. Chapter 3[5]

Chapter 3 presents a detailed literature review of most (not all) of the work that has been done in reliability growth for complex systems. A synopses of nearly 80 papers are given, which covers 7 planning models, 25 tracking models, 6 projection models, 4 reliability growth surveys or handbooks, and 36 other papers covering theoretical results, simulation studies, real-world applications, personal-perspectives, international standards, or related statistical procedures

The literature review has answered many questions of basic interest about the field of reliability growth. For example, there are three main areas: planning, tracking, and projection. A wide array of statistical procedures (e.g., classical and Bayesian) for point-estimation, confidence interval construction, and goodness-of-fit testing are available for most of the models (not all). Models have been developed for complex systems whose test duration is continuous, as well as for complex systems whose test duration is discrete. There are at least three organizations that currently have tailored software products for reliability growth analysis.

The literature review has also revealed capability gaps mainly for one-shot systems in the areas of planning and projection, as indicated by Mortin and Ellner in [161]. With respect to the area of discrete projection, there are two types of models that depend on the type of corrective action strategy used by program management. The first type addresses the case were all corrective actions are delayed until the end of the current test phase. The second type address the more complicated case where

[5] References in Chapter 3 are given at the end of the document.

corrective actions can be applied to system prototypes after they are first discovered. The main difference between the two types of projection models are their functional forms, the data they require, and the statistical procedures involved for parameter estimation.

The genesis of discrete reliability growth projection is marked by a paper written by Corcoran, Weingarten, and Zehna in 1964 [8], which addresses the delayed case. Since then, a number of other methods have been developed. Among them include the delayed models given by Crow [54], and Ellner & Hall [162], and the non-delayed models given by Ellner [121], and Crow [157] - all of which are models for systems whose usage is measured in the continuous time domain. Hence, the need for reliability growth projection capabilities for one-shot systems. Chapters 4-8 prescribe reliability growth management metrics and associated statistical procedures that fill these capability gaps under both corrective action strategies.

### 1.2.4. Chapter 4[6,7]

Chapter 4 gives a model for estimating the true reliability growth of a complex one-shot system (as opposed to expected reliability). The model offers an alternative to the popular competing risks approach first considered by Corcoran et al. [8], and is suitable for application when one or more failure modes can be discovered in a single trial, and when catastrophic failures modes have been previously discovered and corrected. A logically derived exact expression with theoretical (i.e.,

---

[6] Chapter 4 was submitted to *IEEE Transactions on Reliability* on 18 September 2006, revised on 20 January 2007, and accepted for publication on 24 July 2007. The paper appears in the March 2008 issue of that journal (i.e., vol. 57, no. 1, pp. 174-181).

[7] References in Chapter 4 are given at the end of Chapter 4.

based on a theoretical shrinkage factor estimator), and practical estimates (i.e., based on maximum likelihood and method of moments) of reliability are developed.

A new method is given for approximating the vector of failure probabilities inherent to a complex one-shot system. This new method is based on a shrinkage factor derived herein. The benefit of this procedure is that it not only reduces error, but reduces the number of unknowns requiring estimation from $k+1$ to only three, where $k$ is the total potential number of failure modes in the system. Also, estimates of failure mode probabilities of occurrence, whether observed or unobserved during testing, are finite and positive. This in an improvement over the well-known, widely-used maximum likelihood estimate for a failure probability, which yields an estimate of zero for unobserved failure modes.

Unique limiting approximations of model equations are derived, which yield interesting simplifications. In particular, a mathematically-convenient functional form for the expected initial reliability of a one-shot system is derived. This quantity serves as an estimate of the current demonstrated reliability of a one-shot system, and offers an alternative to the typical reliability point estimate calculated as the ratio of the number of successful trials to the total number of trials.

Monte Carlo simulation is used to highlight model accuracy with respect to projection error. While all error terms are within $\pm 2.5\%$ of the true reliability, approximated normal distributions indicate that projection error is within $\pm 0.9\%$, 90% of the time. While model accuracy is generally found to be good, tailored Monte Carlo simulation studies are recommended to highlight model accuracy for specific systems of interest.

Important assumptions and limitations of this methodology is as follows:

1. The distribution of the number of failures in $T$ trials for each failure mode $i = 1,\ldots,k$ is assumed to follow a binomial random variable with probability of occurrence $p_1,\ldots,p_k$, respectively. There are two important limitations imposed by this assumption. First, the failure probabilities for each failure mode $p_1,\ldots,p_k$ must remain fixed over the entire test phase of $T$ trials. Thus, if wearout or reliability growth is encountered to a significant extent during test, this assumption will be violated and the methodology may not be suitably applied. For this reason, corrective actions are assumed to be delayed until the end of the current test phase. Second, all failure data for each failure mode $i = 1,\ldots,k$ is assumed to be generated in (and only in) the $T$ trials. Thus, lower-level subsystem test data and other data captured outside of the $T$ trials cannot be incorporated in the assessment.
2. Initial failure probabilities $p_1,\ldots,p_k$ inherent to the system are assumed to constitute a realization of a random sample from an iid beta random variable. The major limitation of this assumption is that the methodology may not be suitably applied if a given failure mode in the system fails as a consequence of failure of a different mode.
3. Because of (1) and (2) above, the system must be at a stage in development where catastrophic failure modes have been previously discovered and corrected, and are therefore not preventing the occurrence of other failure modes.

4. There must be 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.

### 1.2.5. Chapter 5[8,9]

Chapter 5 introduces a new reliability growth management metrics for one-shot systems that are applicable to the case where all corrective actions are implemented at the end of the current test phase. Associated statistical procedures for parameter estimation follow from those given in Chapter 4. The methodology consists of four primary model equations for assessing: expected reliability (i.e., initial, projected, and upper achievable limit), the expected number of failure modes observed in testing, the expected probability of discovering new failure modes, and the expected probability of observing a repeat failure mode. These metrics provide an analytical framework from which reliability practitioners can estimate reliability improvement, address goodness-of-fit concerns, quantify programmatic risk, and assess reliability maturity of one-shot systems. A numerical example is given to illustrate the value and utility of the approach. The methodology is useful to program managers and reliability practitioners developing one-shot systems under a delayed corrective action strategy. Limitations of this methodology are identical to those listed in the previous subsection (since this approach also addresses a delayed corrective action strategy).

---

[8] Chapter 5 was submitted to *Reliability Engineering & System Safety* on 26 August 2007, revised on 5 November 2007, and accepted for publication on 12 November 2007. Citation information has not yet been assigned.

[9] References in Chapter 5 are given at the end of Chapter 5.

### 1.2.6. Chapter 6[10,11]

Chapter 6 builds on the management metrics introduced in Chapter 5. It builds on these management metrics in the sense that marginal expressions are derived from the earlier model equations. These marginal expressions are in terms of the two beta shape parameters, rather than the initial failure probabilities inherent to the system (e.g., the failure probabilities $p_1, \ldots, p_k$ are integrated out of the equations). The new expressions are used for analyzing reliability growth under an arbitrary corrective action strategy (i.e., fixes can be applied to system prototypes after associated failure modes are first discovered). Thus, the system configuration need not be constant. The methodology consists of the same management metrics for identical purposes as those discussed in Chapter 5.

In addition to the management metrics, a new likelihood function and statistical estimation procedure are derived that can be used under an arbitrary corrective action strategy. The likelihood function uses the marginal beta-geometric distribution and gives consideration to Type I (i.e., time) censored data. Limiting approximations of model parameters and the management metrics are derived, yielding a number of interesting simplifications. In fact, all the model equations reduce to simple mathematically-convenient expressions in terms of only the beta shape parameter representing pseudo trials, and the initial reliability of the system. A numerical example is given to illustrate the new statistical procedures and utility of

[10] Chapter 6 was submitted to *Technometrics* on 24 March 2008 for consideration of publication.
[11] References in Chapter 6 are given at the end of Chapter 6.

the management metrics. This methodology is useful to program managers and reliability practitioners developing discrete systems under an arbitrary corrective action strategy.

Limitations (2) and (3) from Section 1.2.4 also apply to this approach. In addition, the distribution of the number of trials $t_1,\ldots,t_k$ until the first occurrence of each failure mode is assumed to constitute a realization of a random sample $T_1,\ldots,T_k$ such that $T_i \sim Geometric(p_i)$ for each failure mode $i=1,\ldots,k$. Thus, the failure probabilities $p_1,\ldots,p_k$ are implicitly assumed to remain constant until they are first discovered on trials $t_1,\ldots,t_k$. Note that following trials $t_1,\ldots,t_k$, the failure probabilities can be reduced by fractional amounts proportional to their corresponding FEF as a result of failure mode mitigation.

### 1.2.7. Chapter 7[12]

Chapter 7 develops approximate statistical procedures for goodness-of-fit testing and confidence interval construction for the expected reliability under a delayed or non-delayed corrective action strategy (given in Chapter 6). Without loss of generality, the same procedures can be applied in the delayed case by using the statistical procedures presented in Chapter 4. Two goodness-of-fit techniques are discussed. The first technique is a graphical approach that highlights the correlation between the actual cumulative number of observed failure modes versus trials, against the expected number of observed failure modes versus trials. The second technique is

[12] References in Chapter 7 are given at the end of Chapter 7.

a statistical test procedure based on a chi-squared random variable. Both techniques are based on a logically derived exact expression for the expected number of observed failure modes, also developed in Chapter 7. An important assumption associated with the proposed GOF test procedure is that the test statistic follows a chi-squared distribution with $c-2$ degrees of freedom. This assumption is investigated via Monte Carlo simulation. Provided that the expected frequencies in each class interval of the required GOF table are greater than or equal to 4, the test statistic is found to approximate well to a chi-squared distribution with $c-2$ degrees of freedom.

Using a Fisher matrix normal approximation approach, a confidence interval is constructed for the expected reliability of a one-shot system. This identical procedure can be applied to obtain an interval estimate on the other management metrics. Monte Carlo simulation is utilized to estimate the coverage probability associated with the approximation routine. In the context of this research, the coverage probability is the number of times in simulation that the confidence interval contains the true reliability out of the total number of replications. The coverage probability is found to be close to the nominal confidence level when censoring is moderate. Numerical examples are given to illustrate the proposed confidence interval and goodness-of-fit procedures. This methodology is useful to reliability practitioners who wish to address model goodness-of-fit concerns, and/or obtain a confidence interval estimate on the expected reliability of a one-shot system undergoing development.

An important assumption with this methodology is that the log-odds transform of the expected reliability is approximately normal. This assumption is investigated by Monte Carlo simulation. It was found that the aforementioned statistic was approximately normal, but not perfectly symmetric. Thus, the coverage of this approximate confidence interval routine can vary. This means that it may yield confidence intervals whose coverage is lower (e.g., 70%) than the associated nominal confidence level used (e.g., 80%). The effect that this has on the interval is that it will tend to be tighter and yield less confidence than advertised by the nominal confidence level. The volume of data censoring also has this effect on the coverage of approximate confidence interval procedures, and is researched extensively in statistical literature.

### 1.2.8. Chapter 8

Chapter 8 develops Bayesian estimation procedures for both corrective action strategies (i.e., delayed and arbitrary) that can be utilized as alternatives to the classical estimation methods developed in Chapters 4 and 6. 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 the management metrics. Another advantage is that all a priori engineering knowledge can be utilized in the assessment procedure. Analytical results (i.e., joint posterior distributions) are presented to obtain Bayes' estimates of the beta shape parameters for both corrective action strategies. A Monte Carlo approach is outlined for constructing uncertainty distributions on the management metrics. For inference on interval estimation, Bayesian probability limits are obtained in the usual manner (i.e., via desired percentiles of the uncertainty distributions). Numerical examples are given to illustrate these Bayesian procedures. In particular, Bayes' estimates of the beta shape parameters are obtained for a given sample of data. Also, Bayesian epistemic uncertainty distributions for all reliability growth management metrics are constructed via the proposed Monte Carlo approach. The limitations associated with these procedures follow directly from Section 1.2.4 under the delayed case, and Section 1.2.6 for an arbitrary correction strategy.

### 1.2.9. Future Work

Future work that could be done in the area of discrete reliability growth projection to advance the state-of-the-art further would be to develop a projection model under an arbitrary corrective action strategy that uses individual fix effectiveness factors. Individual fix effectiveness factors represent the fraction reduction for individual failure mode probabilities (or rates) of occurrence due to implementation of a corrective action. Individual fix effectiveness factors are used in this research for the delayed corrective action strategy. An average fix effectiveness factor is used for an arbitrary corrective action strategy.

A second area for future work would be to revisit the problem originally considered by Corcoran et al. in 1964 [8]. They developed the first projection model under the popular competing risks framework where at most one failure mode can be triggered on any single trial. Their reliability projection is suitable in cases where corrective actions are installed at the conclusion of a single test phase, and where the number of trial outcomes of interest is a multinomial distributed random variable. One may be able to develop an extension of this model allowing it to be applied under an arbitrary correction strategy. In addition, one may be able to develop a complete set of management metrics to address this case (i.e., competing risks), similar to those given herein. Statistical procedures, both classical and Bayesian, for point and interval estimation could also be explored.

THIS PAGE IS INTENTIONALLY LEFT BLANK.

# 2. INTRODUCTION

## 2.1. Reliability Growth Management

Early prototypes of complex systems will nearly always possess design and manufacturing deficiencies. In order to improve reliability and performance characteristics of developmental items, these initial weaknesses must be found, investigated, and eventually corrected. As a result, system prototypes are manufactured and subjected to a series of different tests, thereby exposing them to the envelop of stresses (e.g., mechanical, thermal, electrical, environmental) that the customer is likely to encounter in the operational environment. Developmental and operational testing of military equipment, for instance, reveals a wide array of problems, such as inherently incapable system designs with respect to required operational profiles; unacceptably premature component overstress and/or wear-out; quality control (e.g., variation) issues in the manufacturing process; subsystem interface problems; software failures; the presence of sneak circuits and other electronic reliability problems; and factors concerning human and operator error. Upon the discovery of each problem encountered during prototype testing, detailed reports are typically generated to document exactly what occurred, when it occurred, and what happened as a result. Test incidents are investigated to determine their root-cause, or failure mode[13]. In some cases, the root-cause of a given test incident may not be well-understood, or ever determined. In general, however, the root-causes of most test incidents are eventually identified. Once identified, engineers develop

[13] ***Failure Mode*** – the root-cause associated with the loss of a required function or component.

proposed corrective actions that, once applied to system prototypes, mitigate (or sometimes eliminate) the occurrence of the failure mode. Corrective actions can consist of, but are not limited to, engineering design modifications, alterations in manufacturing processes, or even changes in equipment operating procedures.

This process of eliminating initial design (or manufacturing) weaknesses in a system via failure mode discovery, analysis, and effective correction is generally what is meant by the term *reliability growth*. Mathematical models that are used to quantify improvements in reliability throughout development are generally referred to as *reliability growth models*. Since the 1950s (e.g., with one of the first reliability growth models given by Weiss in [1]), the genesis of three main areas of the field have emerged, where a wealth of methods have been developed to <u>plan</u>, <u>track</u>, and <u>project</u> the reliability of developmental items. Each of these three areas apply to complex systems whose test durations are measured in the continuous time domain, as well as via discrete trials (e.g., one-shot systems, such as, guns rockets, missile systems, and torpedoes). Reliability Growth Management (RGM)[14] procedures, such as those prescribed in Military Handbook 189 [49], the AMSAA Reliability Growth Guide [144], and Appendix C of the DoD's Guide for Achieving RAM [164], consist of the application of planning, tracking, and projection models with consideration to leveraging the allocation and reallocation of programmatic resources (e.g., schedule, budget, test needs) as appropriate.

There are several reasons why such management procedures are helpful. In general, these tools give program managers the means from which to make informed

[14] ***Reliability Growth Management (RGM)*** is "the systematic planning for reliability achievement as a function of time and other resources, and controlling the ongoing rate of achievement by reallocation of resources based on comparisons between planned and assessed reliability values" [49].

decisions based on quantitative assessments of various aspects of the development program. Ideally, the RGM process should begin at program initiation with the construction of a reliability growth program planning curve. Planning curves are invaluable in that they force management to specify goals for reliability achievement as a function of test time as well as other program resources. Once prototype testing begins, reliability growth tracking models give management the means to gauge the progress of the development effort. Progress is gauged by comparing quantitative assessments of system reliability against the program planning curve. A close agreement between the planned and demonstrated reliability is indicative of a successful reliability growth program (i.e., one that is progressing according to schedule). Reliability growth projection models are then used to quantify the hypothetical reliability of the next configuration of the system (i.e., the reliability that could be achieved if the system developer mitigates known failure modes with a specified level of fix effectiveness). Reliability growth projection models also give measures to quantify programmatic risk, and system maturity. The capabilities in each of these three areas of reliability growth are discussed in more detail in the following sections.

Other benefits of adopting prescribed RGM principles include the reduction of lifecycle O&S costs and programmatic risk associated with acquiring a system that does not meet its intended operational and performance requirements. Many cost studies over the years suggest that O&S costs for complex military systems can account for up to 60-84% of the total ownership cost of a weapon system. Even more interesting, the largest portion (i.e., 34%) of the DoD's budget in FY 2000 was associated with "Operations and Maintenance" costs for aging military equipment [168]. These costs were followed by "Military Personnel" (26%), "Procurement" (19%), "RDT&E" (14%), "Military Construction" (2%) and "Other" (2%). In summary, successful RGM includes the application of planning, tracking, and projection models to plan and continuously assess the progress of the development effort while reallocating programmatic resources as necessary. The ultimate goal of a reliability growth program is to develop a system whose final reliability demonstrates that which is required. Of course, a major success criterion of any reliability growth program is maturing the system (i.e., with respect to reliability achievement) within a given fixed schedule and budget.

## 2.2. Elements of Reliability Growth

In the previous section, general concepts were discussed in relation to the field of reliability growth, its areas, and associated management practices. What specifically, however, does one mean by the term reliability growth? Per Military Handbook 189 [49], reliability growth is defined as "the positive improvement of a reliability parameter over a period of time due to changes in product design or manufacturing process." While this definition has been widely adopted since 1981, one could argue that a reliability parameter, such as MTBF, and estimates thereof are merely artifacts of achieved reliability. In other words, positive improvements in reliability parameters are secondary effects of the actual concept of interest. In this research, reliability growth is considered to be the increase in the true (but unknown) reliability of a system as a result of failure mode discovery, analysis, and effective correction. Thus, as the true reliability of a system is increased, its reliability estimates (if accurate) also increase, thereby quantifying reliability growth. Naturally, reliability improvements gained during the development effort are heavily dependent upon the effectiveness of corrective actions applied to system prototypes. A corrective action is an improvement to either the hardware, software, or human factors aspects of a system. Some examples include:

1. Hardware reliability. Engineering design modifications of a system, changes to subsystem interfaces or circuit board designs, adjustments to material properties of components, and recapitalizing the facilities and equipment associated with manufacturing processes (especially packaging processes for electronic components).
2. Software reliability. Corrective actions addressing bug modes, logic errors, data quality issues, sneak logic-circuits, software design, hardware/software interfaces, and code syntax problems can be examples of software reliability improvements.
3. Human reliability. Reliability growth associated with human/operator factors could range from implementing more effective operator/maintainer training programs, modifying vehicle or equipment operating procedures, reconfiguring inspection routines/checklists, increasing the frequency of safety assessments and perhaps addressing worker supervision and motivational factors.

Regardless of the type of corrective action and area for which they are associated (i.e., hardware, software or human), a major performance indicator of the growth process entails the rate at which reliability is improved. The *rate of reliability improvement*[15] is dependent upon several factors including (but not limited to):

1. The rate at which failure modes are discovered during testing.
2. The turnaround time associated with performing root cause failure analysis.
3. The turnaround time associated with the official scoring of failure data via the system's FD/SC (which defines the customer's notion of system failure).
4. The turnaround time associated with the development of corrective actions.
5. The turnaround time associated with the corrective action review and approval process. This is typically done by a FPRB, who assigns failure mode FEF[16] based on expert engineering judgment.
6. The turnaround time associated with the implementation of approved fixes. Fixes are typically installed on system prototypes during a planned CAP but can also be applied to prototypes in a staggered, random fashion.
7. *Management strategy*[17].
8. The effectiveness of corrective actions and the overall extent to which the associated failure modes contribute to the system's failure intensity, or probability of failure.

---

[15] The ***rate of reliability improvement*** should not be confused with the growth rate parameter, $\alpha$, associated with the Duane model [7]. The rate of reliability improvement is the actual reliability improvement achieved via the *TAFT* process. The growth rate, $\alpha$, is a parameter of the Duane model that represents the slope of a linear approximation to the cumulative failure rate versus cumulative test time plotted on a log-log scale.

[16] A ***Fix Effectiveness Factor (FEF)*** is the fraction reduction in an initial mode failure rate (or failure probability) of occurrence due to implementation of corrective action. *FEFs* are commonly assigned via expert engineering judgment. Estimating demonstrated fix effectiveness is also possible. An average *FEF* of 0.80 is a common and decent level of fix effectiveness used in reliability growth projection analyses but, of course, such a quantity is highly dependent upon a given corrective action.

[17] ***Management Strategy (MS)*** is a reliability growth planning parameter that represents the portion of a system's failure intensity (or probability of failure) associated with failure modes that program management is expected to address via corrective action. A *MS* of 0.90 is commonly used in reliability growth planning but, again, it is a quantity that depends on programmatic factors (schedule, budget, management philosophy etc).

Clearly, all of the above are highly dependent on timely management commitment to provide necessary allocation and reallocation of programmatic resources, such as schedule and budget. For example, early planning in terms of schedule and budget is required to successfully plan and fund required testing, failure analysis, as well as the development and implementation of the corrective action effort.

## 2.3. Areas of Reliability Growth

### 2.3.1. Reliability Growth Planning

Reliability growth planning is an area of reliability growth that addresses program schedules, amount of testing, resources available and the realism of the test program in achieving its requirements [144]. Reliability growth planning is quantified and reflected through a reliability growth program *planning curve*[18]. Planning curves are typically in terms of MTBF expressed as a function of cumulative test duration. Reliability growth planning curves establish interim reliability goals throughout the development process and identify many important factors, such as, the number and schedule of CAP, planned MS, planned average FEF, initial MTBF, goal MTBF (i.e., system or subsystem reliability requirement). Overall, planning curves illustrate program management's planned reliability achievement as a function of test time, in addition to other resources. A logically constructed reliability growth plan is a powerful management tool for identifying

[18] A ***reliability growth planning curve*** displays the anticipated reliability growth of a system or subsystem over the course of the development program [49].

early-on the programmatic resources that will be necessary for reliability achievement. Several reliability growth planning models have been developed over the past several decades. Examples of some planning curves are presented in Chapter 3.

### 2.3.2. Reliability Growth Tracking

Reliability growth tracking [49] is an area of reliability growth that provides management the opportunity to gauge the progress of the reliability improvement effort for a system by obtaining a demonstrated numerical measure of system reliability throughout development. These reliability assessments (i.e., point and CI estimates) are compared to the system's reliability growth planning curve to determine if the actual reliability achievement is progressing according to that which was planned. Tracking models are available for both one-shot systems and systems whose test durations are continuous (e.g., measured in time or distance). Reliability growth tracking is the most well-developed area of reliability growth. Practitioners have a wealth of models from which to choose. Several tracking models are discussed in Chapter 3.

### 2.3.3. Reliability Growth Projection

Reliability growth projection is an area of reliability growth that provides an estimate of system reliability based on assessments of the effectiveness of corrective actions and failure data generated from the current and/or previous system configurations [144]. The main focus of reliability growth projection is to estimate the reliability of a future configuration of a system that would result if known failure modes are corrected via a specified fix-effectiveness. Fix-effectiveness is quantified through FEF (some historical FEF are discussed by Trapnell in his reliability growth data studies [60] and [61]). By using the FEF, the initial failure rates of occurrence associated with corrected failure modes are reduced by a fractional amount (specified by the FEF). This reduction in the initial failure rates of occurrence for corrected modes is the primary idea behind reliability growth projection models.

There are two types of projection models depending on the corrective action strategy utilized by program management. The first type address the case were all corrective actions are implemented at the end of the current test phase. In this case, all fixes are delayed, and the configuration of the system with respect to reliability remains constant. The second type of projection models address the case where fixes can be delayed or non-delayed (e.g., an arbitrary corrective action strategy). In this case, some fixes are applied during the current test phase, and some are applied at the end of testing. Therefore, from a reliability standpoint, the system configuration is (in general) dynamic (i.e., improving throughout testing as a result of design changes).

Figure 1 below displays the fundamental concepts behind reliability growth projection. The pie chart on the left represents system unreliability before correction. The pie chart on the right represents system reliability after correction. Notice that a portion of system unreliability is associated with *A-modes*[19] and a portion is

[19] An ***A-mode*** is a failure mode that will not be addressed via corrective action.

associated with *B-modes*[20] (this failure mode classification scheme was originally proposed by Crow in [54]). Failure mode mitigation is mathematically modeled as the reduction of initial failure rates (or probabilities) of occurrence associated with corrected modes by a fractional amount $(1-d_i)$. The term $d_i \in (0,1)$ is the FEF for failure mode *i*. For example, the initial failure rate $\lambda_i$, after correction, is reduced to $(1-d_i)\cdot\lambda_i$. Notice in Figure 1 that the portion of system unreliability comprised of A-modes does not change. This is because these failure modes are not addressed by corrective action. Also notice there is a portion of system unreliability comprised of B-modes that are observed during testing, and B-modes that are not observed during testing. The B-modes that are not observed also do not get corrected. Therefore, the corresponding portion of system unreliability remains unchanged. The portion of system unreliability that does get reduced, however, is the portion comprised of observed B-modes. The gray piece of the pie chart on the right in Figure 1 illustrates the portion of system unreliability eliminated by corrective action.

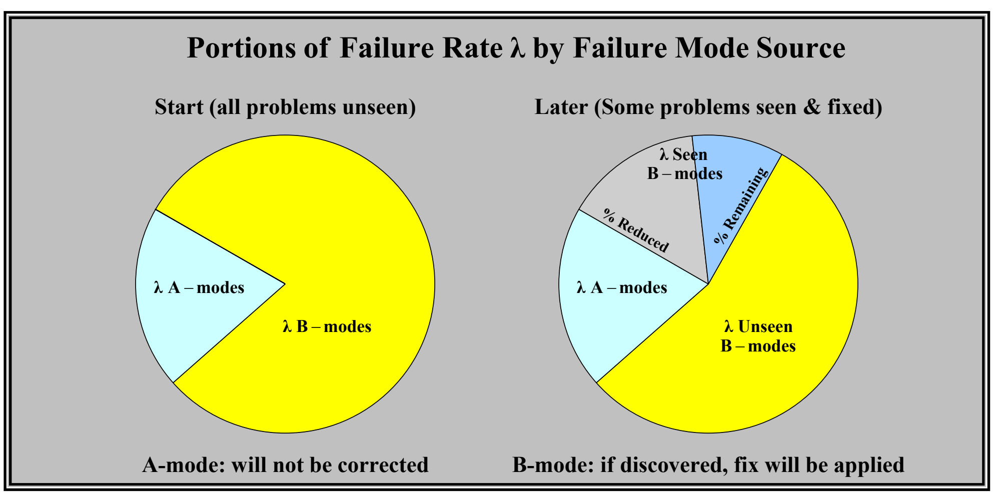


**Figure 1. Reliability Growth Projection Concepts.**

[20] A ***B-mode*** is a failure mode that program management will address via corrective action, if observed during testing.

# 3. LITERATURE REVIEW

## 3.1. Overview

Many reliability growth models have been developed over the past several decades. The purpose of these models is to help program managers and reliability practitioners address the formidable tasks of planning, tracking, and projecting the reliability improvement of a system throughout the development process. In order to summarize the current capabilities that exist, notes from the following literature review are given. This literature review briefly covers the majority of the work (not all) done in the field. Planning models, tracking models, and projection models are given in the following three sections, respectively. More comprehensive works, such as, handbooks, surveys, and guides are given in Section 3.5. A synopsis of associated theoretical results, simulation studies, real-world applications, personal-perspectives, and related statistical procedures (i.e., point-estimation, confidence interval construction, and goodness-of-fit testing) is given in Section 3.6.

## 3.2. Reliability Growth Planning

### 3.2.1. Duane's Model (1964)

In 1964 J.T. Duane [7], who at the time was an aerospace engineer with General Electric Company in Erie, PA, discovered that if changes to improve reliability are incorporated into the design of a system, then the cumulative failure rate versus cumulative test time plotted on a log-log scale exhibits a linear relationship. This relationship is sometimes referred to as the *Duane Postulate*. Duane discovered this by developing cumulative failure rate plots for a broad range of aircraft equipment, including complex hydro-mechanical devices, aircraft generators, and jet engines. Figure 2 below shows an example of a typical *Duane Plot*, where the parameter $\alpha$ represents the overall rate of reliability improvement throughout the course of the development program (for this model). The parameter $\alpha$ is commonly referred to as the growth rate and represents the negative of the slope of the logarithm of the cumulative failure rate.

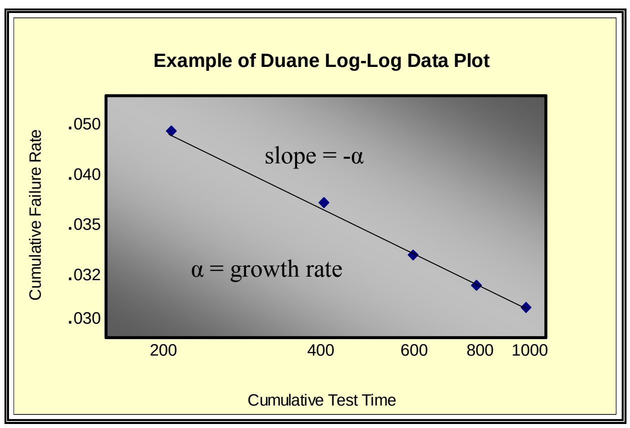


**Figure 2. Duane Reliability Growth Plot.**

Duane's original intent for the methodology was to monitor or track the reliability improvement[21] in a major subsystems for various aircrafts. This was done via the above assumed linear relationship, which approximately accounts for the overall change in the sequence of MTBF steps associated with successively redesigned

[21] Duane's original intent was to monitor reliability of a complex system undergoing design improvements, so his approach is mostly associated with reliability growth tracking. It is presented as a planning model since it can also be used as a planning tool. Moreover, the Duane postulate is used as a fundamental assumption in several other growth models. Thus, the method marks a natural starting-point from which to begin this literature review.

configurations of a system throughout the TAFT process. While Duane's original intent was to monitor reliability improvement, the model has had tantamount ramifications throughout the field of reliability growth. According to Ebeling [160] the Duane growth model is "the earliest developed and most frequently used reliability growth model." In fact, the *Duane Postulate* is utilized as a fundamental assumption in many other reliability growth models that will be discussed below. An early and detailed application of the Duane model is presented by Selby and Miller [20].

### 3.2.2. Selby-Miller RPM Model (1970)

Selby and Miller [20] present an approach to reliability planning and management of complex weapon systems, which they refer to as "Reliability Planning Management (RPM)." The basic concept behind the RPM model includes its proposed "patterned reliability growth" approach to planning. This "patterned reliability growth" methodology follows directly from Duane's postulate that the cumulative failure rate versus cumulative test duration on a log-log scale is approximately linear with slope, or growth rate, $\alpha$. While this concept is not new, the RPM model appears to be the first application of the Duane postulate for reliability growth planning (as opposed to its original intent of reliability growth monitoring).

### 3.2.3. MIL-HDBK-189 Planning Model (1982)

The purpose of the MIL-HDBK-189 model [49] is to construct a reliability growth planning curve over the developmental test program useful to program management. The planning curve serves as a baseline against which reliability assessments can be compared, and it can highlight the need to management when reallocation of resources is necessary. The model is based on the *Duane Postulate* and consists of an *idealized system reliability growth curve*[22], that portrays the profile for reliability growth throughout the developmental test period and has a constant MTBF during the initial test phase. The planning parameters that define the idealized growth curve include: (1) the initial MTBF, (2) length of the initial test phase (i.e., reliability demonstration test for the initial MTBF), (3) the final MTBF (e.g., reliability requirement, or goal), (4) the growth rate and (5) the duration of the entire growth program. Some historical data on growth rates for Army systems is discussed by Ellner and Trapnell in [89]. The model also gives a set of expected MTBF steps during each test phase in the growth program. Corrective action periods are scheduled between each of the test phases where fixes are applied to previously observed failure modes. These improvements increase system reliability iteratively and result in an increasing sequence of MTBF steps, as displayed in Figure 3.

[22] A ***reliability growth idealized curve*** is a planning curve that consists of a single smooth curve based on initial conditions, assumed growth rate and management strategy [144].

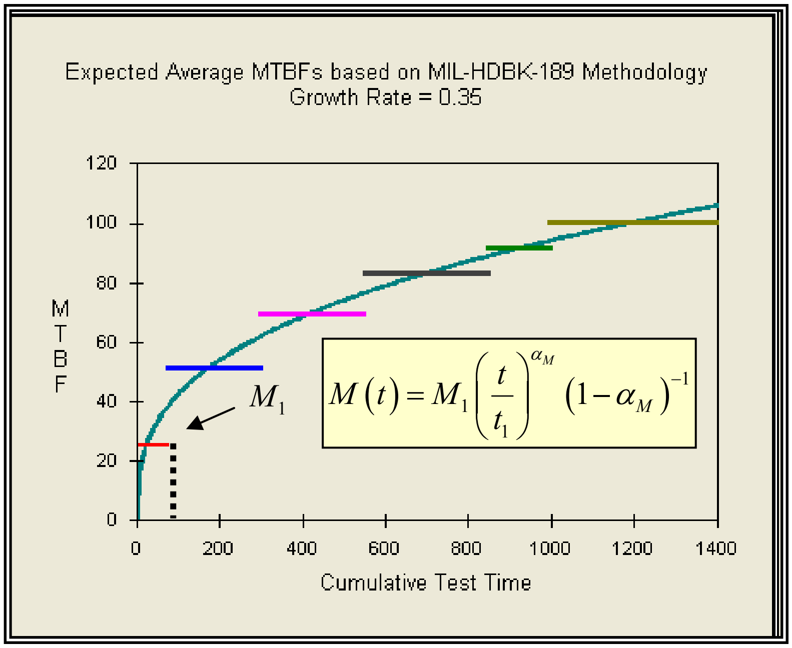


**Figure 3. MIL-HDBK-189 Planning Curve.**

Ellner and Ziad [76] later studied the statistical precision and robustness[23] of the MIL-HDBK-189 model's biased and unbiased estimators of MTBF. They conclude that the precision of the estimators strongly depend on the expected number of failures. Also robustness between the biased and unbiased estimators is approximately equivalent.

### 3.2.4. AMSAA System-Level Planning Model (1992)

The SPLAN discussed by Ellner et al. [144] is another variant of the MIL-HDBK-189 model that can be used to construct system reliability growth test plans and associated idealized system reliability growth curves. The model can also prescribe the required test duration to achieve a system reliability requirement as a point estimate. This model gives several new options for determining various

[23] ***Robustness*** refers to the effect on estimator statistical precision due to discrete configuration changes in system reliability.

planning parameters, which is convenient for conducting sensitivity analyses. For example, given any four of the five planning parameters mentioned above, SPLAN determines the value of the remaining parameter. Most often, the initial MTBF, final MTBF, growth rate, and length of the initial test phase are provided to determine the test duration required in a given development program. Figure 4 shows an example growth curve generated by SPLAN.

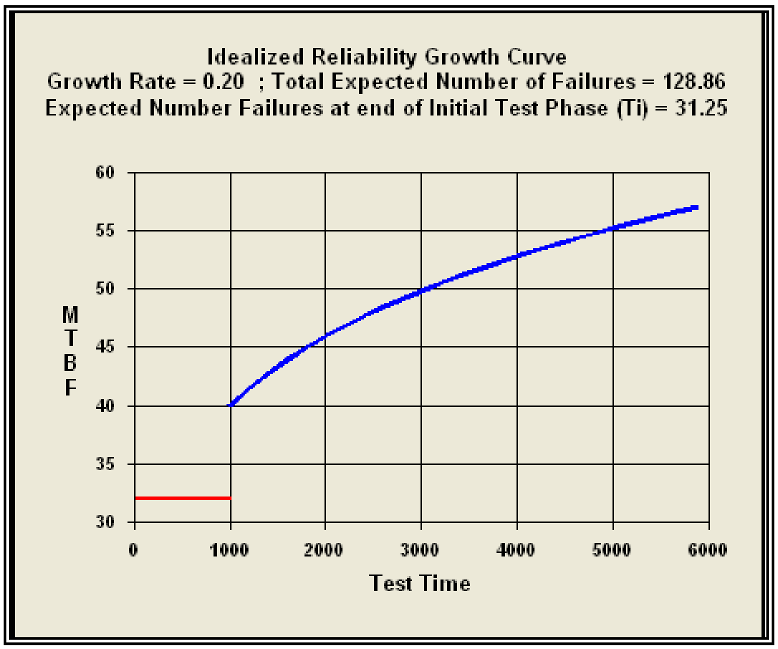


**Figure 4. SPLAN Planning Curve.**

### 3.2.5. Ellner's Subsystem Planning Model (1992)

The SSPLAN model was developed by Ellner et al. [102] and [117] to develop system or subsystem reliability growth test plans that achieve a given system-level MTBF objective with a specified level of confidence. That is, SSPLAN determines the subsystem test times and subsystem reliabilities required to demonstrate a system MTBF objective at a given level of statistical confidence.

Other work on SSPLAN includes Ellner and Mioduski's [100] operating characteristic analysis for the model. Consumer and producer's risks are expressed in terms of the model parameters. For a given confidence level, they show that these risks only depend on the expected number of failures during testing, and the ratio of the of demonstrated MTBF with confidence over the MTBF requirement. Formulas are developed for computing these risks as a function of the test duration and growth curve planning parameters.

### 3.2.6. Mioduski's Threshold Program (1992)

The Threshold Program was developed by Mioduski at AMSAA but no publication on the model is known to exist. However, the model is discussed by Broemm in [163] and is offered in AMSAA's Visual Growth Suite (VGS) [168]. The program determines at selected program milestones (e.g., thresholds), if the demonstrated reliability of a system is failing to improve as prescribed by the MIL-HDBK-189 idealized curve. It consists of a hypothesis test that compares a reliability point estimate for a system (based on actual failure data) against the theoretical threshold value consistent with the planning curve. Associated threshold values are established early in the acquisition process for program milestones or major decision-points. The test statistic in the procedure is the reliability point estimate (i.e., MTBF) computed from test data for individual system configurations. If the test statistic is inside the rejection region for the test, the program gives statistical evidence at a specified significance level that the system's reliability is not in conformance with the approved reliability growth program plan.

### 3.2.7. Ellner-Hall PM2 Model (2006)

The purpose of PM2, discussed by Ellner and Hall in [166], is to construct a reliability growth program planning curve for systems under development. Exact expressions are presented for the expected number of surfaced failure modes and system failure intensity as functions of test time. These exact expressions depend on a large number of parameters, but functional forms are derived to approximate these quantities that only depend on a small number of parameters (giving a parsimonious approximation). Simulation results are presented which show that the functional form of the derived parsimonious approximations can adequately represent the expected reliability growth associated with a variety of parent distributions for the initial failure rates inherent to the system. The main difference of this model in comparison to the other planning models is that it is independent of the NHPP assumption and utilizes parameters directly influenced by program management, such as: 1) initial MTBF; 2) MS; 3) goal MTBF; 4) average lag-time associated with fix implementation; 5) total test time; 6) average FEF; 7) the number and placement of CAP and 8) the planned monthly RAM test hours. Another benefit of PM2 is that it is the first planning model to take into consideration the lag-time due to implementation of corrective actions. An example of the type of detailed reliability growth plan that can be constructed using PM2 is shown in Figure 5 below. The vertical lines displayed in the figure correspond to four months prior to the corrective action periods (or refurbishment periods) where fixes are installed to known failure modes. The significance of this is that only failure modes discovered before the four-month lag-time are addressed in the corresponding corrective action periods. The lag-time can be due to many factors but is mainly due to the turnaround time associated with root-cause analysis and the corrective action review, approval, and implementation process.

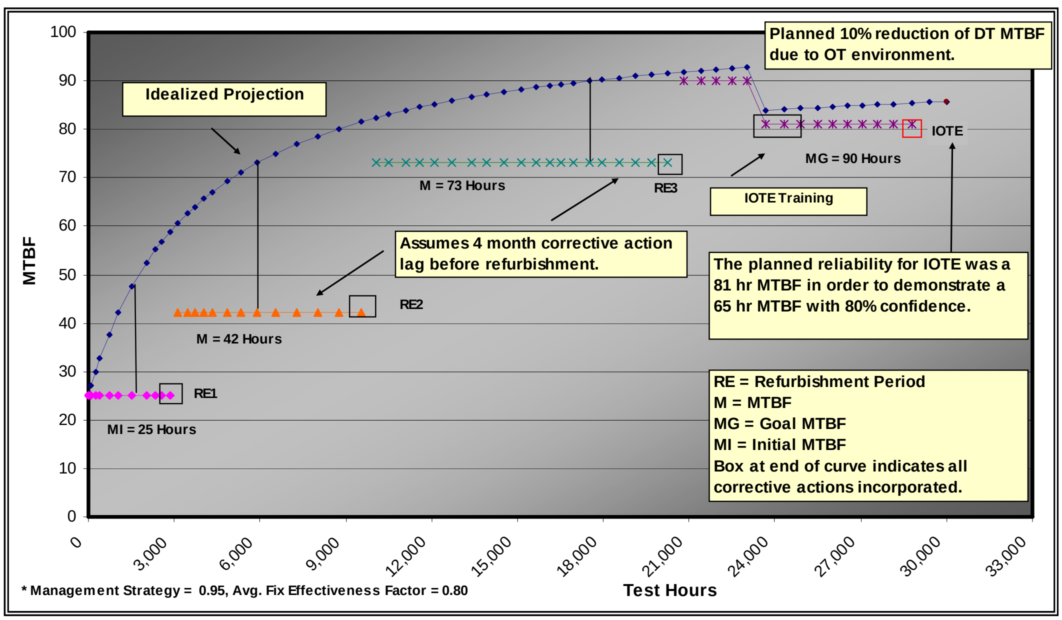


**Figure 5. PM2 Planning Curve.**

## 3.3. Reliability Growth Tracking

### 3.3.1. Weiss' Model (1956)

Weiss [1] developed methods for monitoring and extrapolating reliability growth of guided missile systems with Poisson-type failures. This paper was the earliest found on the subject. In this approach, the MTTF is believed to change over a sequence of successive trials as a result of finding and fixing failure modes in a system. MLE procedures are utilized to determine if reliability is increasing or decreasing, as well as to identify the uncertainty of the reliability estimate. The model is shown to lead to a logistic-type reliability growth curve. Expressions are given for the estimated MTTF obtained from test data, as well as its variance.

### 3.3.2. Aroef's Model (1957)

Aroef [2] developed a reliability growth tracking model for continuous systems. He assumes that the rate of reliability improvement of a system is directly proportional to the growth achieved at a given time, and inversely proportional to test duration squared. The resulting differential equation takes-on the form $df(t)/dt = \alpha \cdot f(t)/t^2$. The solution is found to be $f(t) = \theta \cdot \exp[-\alpha / t]$ where $\alpha$ is the growth rate, and $\theta$ is the upper-limit on reliability (i.e., MTBF) that can be achieved as $t \to \infty$.

### 3.3.3. Rosner's IBM Model (1961)

Rosner [3] developed what has become known as the IBM model, which is an expression for a system's failure intensity function (i.e., rate of occurrence of failure). He assumes that the rate of occurrence of failure at time *t* is proportional to the number of non-random defects remaining in the system at time *t*. The resulting differential equation is expressed as $dN(t)/dt = -b \cdot N(t)$, which has the solution $N(t) = a \cdot \exp[-b \cdot t]$. The constants $a$ and $b$ are approximated by regression. An interesting feature of the model includes its ability to estimate the required test duration for the system to be at a given "fraction corrected" (i.e., a fraction of the original failures that have been corrected). The model also estimates the number of non-random failures remaining at a given time.

### 3.3.4. Lloyd-Lipow Model (1962)

Lloyd and Lipow [4] developed a growth model to estimate the reliability of a system comprised of a single failure mode. The test program is assumed to be conducted in a series of trials. If the system fails in a given trial, a corrective action is implemented, and is mathematically modeled with a finite probability of being successful in mitigating the occurrence of the failure mode. The model has a simple exponential form given by $R_n = 1 - A \cdot \exp[-C \cdot (n-1)]$, where $R_n$ is the reliability of the system in the *n*-th trial. Model-parameters $A$ and $C$ are estimated via test data. They also present a second model, $R_k = R_\infty - \frac{\alpha}{k}$, for estimating the reliability of a system in a given stage, in this case stage *k*. MLE and LS procedures are developed for estimating the model parameters $R_\infty$ and $\alpha$. A lower-confidence limit on $R_k$ is also discussed, in addition to other potential functional forms of reliability growth models.

### 3.3.5. Chernoff-Woods Model (1962)

Chernoff and Woods [5] present several exponential regression reliability growth models. One model of interest, due to its simplicity, estimates the probability that a system will successfully operate after a given number, $r$, failures have occurred and been subsequently corrected. The model is given by the simple exponential form $P_r = 1 - \exp\left[-\left(\alpha + \beta \cdot r\right)\right]$, where $\alpha > 0$ and $\beta > 0$ are parameters estimated by a LS method. Woods later gives a review of similar models in [46].

### 3.3.6. Wolman's Model (1963)

Wolman [6] advanced the idea of AssignC failure modes (i.e., assignable cause failure modes that can be eliminated by redesign). He assumes all assignable cause failures occur with equal probability in each trial and are completely eliminated upon initial observation. Hence, reliability is improved over a sequence of trials. Wolman assesses the reliability at stage $k$ by the model $R_k = 1 - q_0 - \left(M + 1 - k\right) \cdot q$ where $q_0$ denotes the probability of a non-AssignC failure mode, $M$ is the initial number of AssignC failure modes, and $q$ is the probability of occurrence of a single AssignC failure mode. Probabilistic assessments for the model are provided via Markov chain approach. MLE procedures were later developed by Bresenham [9].

### 3.3.7. Cox-Lewis Model (1966)

Cox and Lewis [11] proposed, perhaps, one of the first NHPP models, which is sometimes referred to as the exponential-law or the log-linear model. It take-on the functional form $m(t) = \exp[\alpha \cdot t + \beta]$, where $\alpha$ and $\beta$ are parameters. The parameters are estimated from test data and GOF test procedures are developed. The model reduces to a HPP when $\alpha = 0$. Also, reliability growth is modeled when $\alpha < 0$.

### 3.3.8. Barlow-Scheuer Model (1966)

Barlow and Scheuer [12] also proposed a *k*-stage reliability growth model, where the outcome of each stage are utilized to improve the system in remaining stages. In their trinomial framework, exactly one of three outcomes can occur in a given stage: success, inherent failure, or an assignable cause failure. The reliability in the *i*-th stage is given by $r_i = 1 - q_0 - q_i$, where $q_0$ is the probability of an inherent failure, and $q_i$ is the probability of an assignable cause failure. MLE procedures are given for $q_0$ and $q_i$ under the restriction that they are non-increasing. A conservative LCB on the reliability of the system in its final configuration is also presented. Other work on this model includes the CI procedures developed by Olsen [42]. Smith [38] examined the model from a Bayesian viewpoint by imposing non-informative uniform prior distributions on each $R_k$ (i.e., the binomial probability of success in stage $k$). This formulation led to a convex combination of beta posterior distributions on $R_k$ from which interval estimates are obtained. Fard and Dietrich's simulation study [58] showed that the Barlow-Scheuer model more accurately estimated the true reliability of a system than any of the other non-Bayesian models considered. Fard and Dietrich [73] later developed a Bayesian formulation of the Barlow-Scheuer model which does not utilize the AssignC failure mode classification.

### 3.3.9. Virene's Gompertz Model (1968)

Virene [16] considered the utility of the trinomial Gompertz equation for reliability growth modeling. The reliability assessment is based on the model $R = a \cdot b^{c^t}$, where $b, c \in (0,1)$. The parameter $a$ is the upper-limit on reliability as time $t \to \infty$. He provides estimation procedures for the three model parameters as well as numerical examples.

### 3.3.10. Pollock's Model (1968)

Pollock [19] developed one of the first (if not the first) Bayesian reliability growth models. He modeled the parameters as random variables with appropriate prior distributions allowing one to project system reliability any time after initiation of the test with, or without, test data. Precision statements on the projection and estimation routines are given.

### 3.3.11. Crow's Continuous Tracking Model (1974)

In [26] Crow gives the first stochastic interpretation of the *Duane Postulate*. This is the first time the instantaneous failure rate for reliability growth (given by Duane's model) was reparameterized and recognized as being the Weibull hazard rate function for a repairable system. The model is given by $r(t) = \lambda \cdot \beta \cdot t^{\beta - 1}$, where $\lambda$ and $\beta$ are model parameters. This observation allowed the development of statistical estimation and GOF procedures for reliability growth, which were also presented for time-truncated data. The same procedures were also develop by Crow [28] shortly thereafter for failure-truncated data. Both failure and time-truncated estimation are given in [29]. CI procedures on MTBF are presented in [28]. Associated estimation procedures are based on ML, and GOF results are based on a Cramer-Von Mises test statistic. Crow gives numerical examples illustrating these procedures and a discussion of Army applications for the methodology.

These results have had a significant impact on reliability growth and repairable systems reliability modeling, as they have served as a methodological foundation for many subsequent approaches. Crow gives more comprehensive treatments to all the normal statistical procedures for the Weibull process in [32], [40], and [52]. This includes MLE procedures, hypothesis tests, and confidence bounds for model parameters (time and failure-truncated testing). Simultaneous confidence bounds on model parameters and a GOF test for the model is also given. Crow elaborates upon several applications of the methodology including: reliability growth, mission reliability, maintenance policies, industrial accidents, and applications in the medical field.

Using the stochastic interpretation of the *Duane Postulate*, the resulting model became known as the RGTMC, given in [25-27] and [30]. This model is used to assess the improvement in the reliability of a system (within a single test phase) during development for which usage is measured on a continuous scale. Applications to reliability analysis for complex, repairable systems is discussed by Crow in [32]. Four "real-world" examples are given by Crow in a much later paper [87]. Figure 6 below shows a plot of the MVF (i.e., expected number of failures) versus test time against the actual number of failures observed during testing of an Army system.

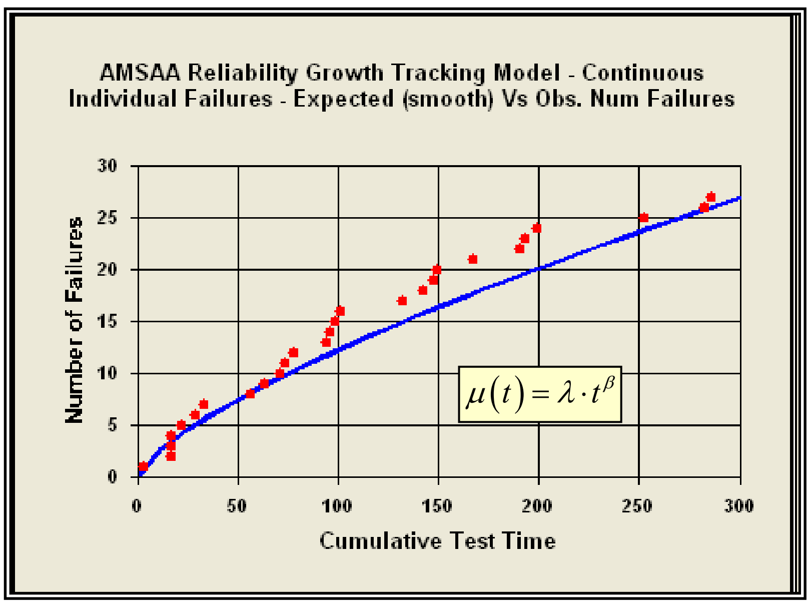


**Figure 6. RGTMC Expected No. Failures.**

Other work on this model includes Crow's MLE procedure [74] for the parameters of the RGTMC in the case where there is missing data (i.e., incomplete data). This practical reliability growth estimation procedure assumes that the actual failure history over the problem interval is unknown. Such a phenomenon occurs when failure information over a period of testing is determined to be incorrect, which leads to the reporting of either to many, or to few failures. Based on these techniques the observed number of failures over the problem interval is adjusted to "more realistically reflect the actual growth pattern." Hence, a valid reliability growth curve can then be fitted to the data and used for evaluation purposes. Two GOF procedures are developed (i.e., one based on the Cramer-Von Mises TS for individual failure time data, and the second based on a chi-squared r.v. for grouped failures). These new procedures are illustrated by several numerical applications. Years later, Crow [109] develops ML and CI procedures for failure data generated from multiple systems under test.

**3.3.12. Lewis-Shedler Model (1976)**

Lewis and Shedler [33] offer and extension of the Cox-Lewis model by developing estimation procedures for the exponential polynomial model for powers of $n = 1, \ldots, 10$. The extension addresses models of the form $m(t) = \exp\left[\alpha_0 + \alpha_1 \cdot t + \alpha_2 \cdot t^2 + \cdots + \alpha_{10} \cdot t^{10}\right]$.

**3.3.13. Singpurwalla's Model (1978)**

Using time-series[24] methods, Singpurwalla developed a discrete reliability growth model to: determine if the binomial parameter $p_i$ (i.e., the probability of success at stage $i = 1, \ldots, k$) is increasing after design modifications are applied in

[24] ***Time series*** is defined as a set of observations generated sequentially in time.

each stage. The model obtains estimates of $p_i$ at the present stage, and also forecasts $p_i$ at future stages (i.e., beyond stage $k$ ).

### 3.3.14. Crow's Discrete Tracking Model (1983)

The RGTMD was developed by Crow [55] for tracking the reliability of one-shot systems during development; such as, guns, rockets, missiles, torpedoes, mortars etc. Statistical point-estimation, CI and GOF procedures are given for both grouped data, and for data captured during a trial-by-trial basis. The model is fundamentally based on the NHPP assumption derived from the *Duane Postulate*. More specifically, the model is constructed by obtaining an equation for the probability of failure on a configuration basis, using the NHPP power-law function (sometimes referred to as the "learning curve"). This equation and a plot of the reliability growth tracking curve is shown in Figure 7 below.

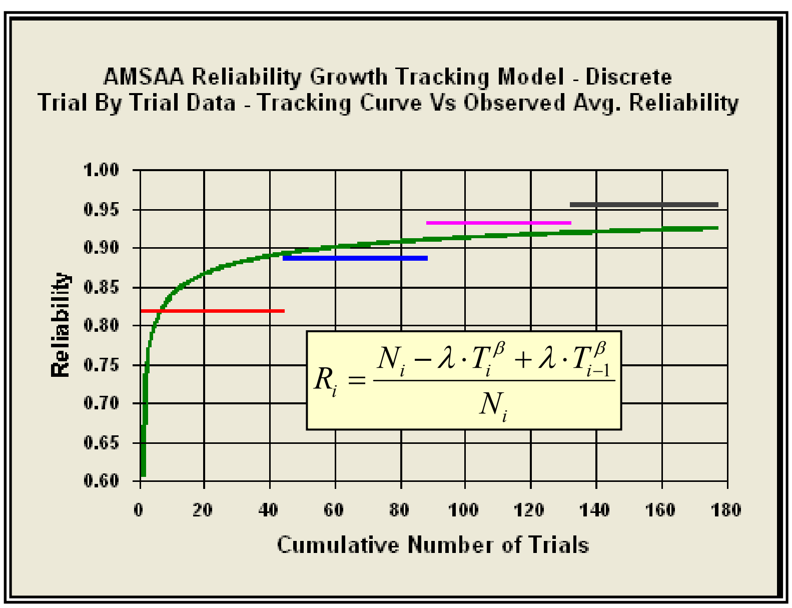


**Figure 7. RGTMD Reliability.**

Other related work on this model was done by Finkelstein [59], and Battacharyya, Fries and Johnson [84]. Finkelstein developed CAE of the model parameters for the case where only a single trial per configuration is tested. He also performed a simulation study to investigate the behavior of the CAE. He concludes that all attempts to obtain MLE of these parameters were unsuccessful and asserts the consistency of the CAE. Battacharyya, Fries and Johnson [84] generalize the CAE given by Finkelstein in the case where there is a constant pre-specified number of test trials between system configuration changes. Large-sample properties of these estimators to include consistency and normality are developed. Large-sample standard-error formulas and CI procedures are given. Finally, they provide a proof on the consistency of the CAE, which confirms the Finkelstein conjecture. Bhattacharyya & Ghosh [98] later showed that the MLE and CAE for the parameters of this model are asymptotically equivalent. Johnson [99] generalized these findings in the case where the sample sizes for each system configuration were unequal. More work was done by Hall and Wessels [145] who formulated an evolutionary programming optimization algorithm to estimate the parameters of the RGTMD [55]. A numerical example is presented, where the standard MLE of the model parameters are compared against the proposed estimates from the optimization algorithm. The estimates are nearly identical. Overall, the algorithm proves to be an effective tool for reliability growth analysis when using the RGTMD.

#### 3.3.15. Robinson-Dietrich Model (1988)

Robinson and Dietrich [78] and [85] develop a reliability growth model for monitoring the progress of the development effort at the system-level while the actual development occurs at the subsystem-level. Using the moments of the subsystem failure rate distributions as they change during testing, they show how the moments of the distribution of the system-level failure rate can be estimated. Using these moments, point-estimates and approximate CI for system reliability growth are derived. A hypothetical example is presented to illustrate some nuances of the methodology. Two additional examples are given on unspecified systems. The first system is comprised of three components, and the second system consists of eleven subsystems in a more complex structure.

### 3.3.16. Kaplan-Cunha-Dykes-Shaver Model (1990)

Kaplan et al. [90] develop a Bayesian method for assessing reliability during product development. Their "stepwise process" is implemented for analyzing failure data derived from the system and subsystem levels. Bayes' theorem is applied sequentially at each level throughout a number of test stages. The prior distribution and updating procedure at each level utilize engineering judgment to evaluate the significance of failures observed and effectiveness of corrective actions. Overall, the paper includes the development and application of a Bayesian framework for gathering, organizing and incorporating expert knowledge into reliability growth assessment. A notable feature of the approach is that assessments of reliability are derived with a concomitant measure for uncertainty.

### 3.3.17. Mazzuchi-Soyer Model (1991)

Mazzuchi and Soyer [92], [106], and [110] present a Bayesian approach for assessing reliability growth during system development. At the end of each stage of testing, failures are examined so that modification can be implemented to remove failure modes. They incorporate prior information into an ordered Dirichlet prior distribution for failure probabilities at each stage. The resulting posterior distribution of all relevant quantities is expressed as a mixture of beta, or Dirichlet, distributions. After each stage of testing, the model gives Bayes estimates of system reliability. The method is illustrated by numerical example. Overall, their approach provides a means for incorporating subjective information into reliability assessment, and provides the means for analyzing system reliability over successive stages of testing using sequential updating of a Bayesian prior distribution. These results are later extended by Erkanli, Mazzuchi and Soyer [135] who consider both the exponential and Weibull time-to-failure models.

### 3.3.18. Heimann-Clark PR-NHPP Model (1992)

Heimann and Clark [105] argue that a more accurate reliability assessments of a system can be obtained by explicitly modeling the effect of defects induced during the manufacturing process. To model this phenomenon, they develop a process-related NHPP by replacing the constant scale factor by a process age-dependent function. This function increases asymptotically over process age to a mature process scale factor value. MLE procedures are given for model parameters. The proposed PR-NHPP addresses the questions: "What will the product reliability be after a given age of the manufacturing process?" and "How much reliability growth time will be required to achieve a given product failure intensity goal?" One parameterization of the NHPP is $h(t) = \left(\frac{\beta}{\alpha}\right) \cdot \left(\frac{t}{\alpha}\right)^{\beta-1}$. The proposed modification is to replace the constant scale factor $\alpha$ in this equation by the function $\alpha(t) = \alpha \cdot (1 - \exp[-b \cdot t])$, where $t$ is the length of time that the production line has been operational, and $b$ is the shape parameter. Numerical examples are given to demonstrate the utility of the proposed PR-NHPP.

### 3.3.19. Fries' Discrete Learning-Curve Model (1993)

Fries [111] develops a learning-curve approach for discrete reliability growth analysis. This approach is particularly appropriate for destructive tests of very expensive systems. Derivations of the new model and of the RGTMD [55] are presented. Approximations of model parameters are obtained by ML procedures. Extensions of both models are discussed, which account for the distinction between *assignable*[25] and non-assignable cause failure modes. Each model accommodates for the monotonic growth in reliability during system development. The models and estimation procedures are illustrated by two numerical examples. In a later paper [116], Fries gives corrections to the likelihood equations that properly reflect the

[25] An ***assignable-cause*** failure mode is a failure mode whose root-cause is known and is therefore readily correctable.

negative binomial (geometric) behavior of the number of trials until the first observed failure.

### 3.3.20. Modified-Gompertz Model (1994)

Kececioglu, Jiang, and Vassiliou [115] observe from several datasets that reliability growth data with an S-shaped trend could not be adequately portrayed by the conventional Gompertz model [16]. They point-out that the reason is due to the model's fixed value of reliability at its inflection point. As a result, only a small fraction of reliability growth datasets following an S-shaped pattern could be fitted. Their proposed solution overcomes this shortcoming by modifying the Gompertz model to include a fourth parameter. This fourth parameter shifts the associated growth curve vertically, thus accommodating for S-shaped growth datasets. The new method is claimed to be more flexible than its predecessor for fitting data with S-shaped trends. The original Gompertz model is given by $R = a \cdot b^{c^T}$. The modification assumes the form $R = a \cdot b^{c^T} + d$. Estimation procedures are presented and consist of solving four equations for the four unknown model parameters. A detailed numerical example is given.

### 3.3.21. Ellner's Subsystem Tracking Model (1996)

The SSTRACK model was developed by Ellner [144] for assessing system level reliability from lower level subsystem testing. The motivation for this methodology was to make greater use of subsystem test data in estimating system reliability. SSTRACK takes into consideration data from both growth and non-growth subsystems. The model uses the Lindstrom-Madden method [4] for combining test data from individual subsystems. The methodology includes statistical CI and GOF procedures. Figure 8 below shows an example of approximate LCB on system MTBF computed from subsystem data as a function of the desired level of statistical confidence.

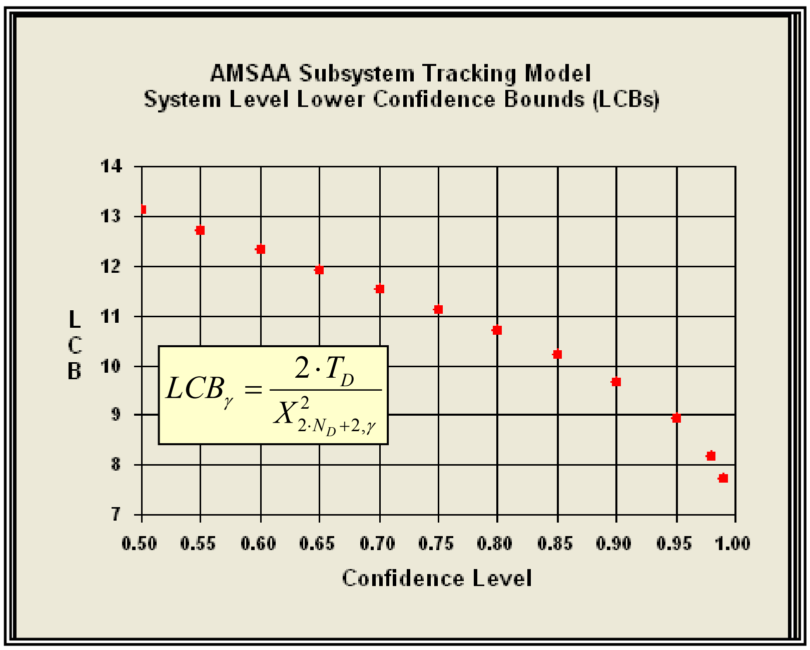


**Figure 8. SSTRACK LCB on MTBF.**

### 3.3.22. Sen's Alternative to the NHPP (1998)

Sen [136] investigates the statistical inference of current reliability of the Duane model [7]. Exact and large-sample distributional results are derived for the ML and LS estimators of the current failure intensity. The extent of misspecification of the NHPP power-law process [25] to fit failure data of a system experiencing recurrent failures is explored. Simulation results and an illustration is provided to supplement the theoretical findings and demonstrate the presented inference results. Sen concludes that the model is a suitable alternative to Crow's NHPP power-law model, in the context of analyzing recurrent failure data from systems undergoing developmental testing. Clarifications on the exact inference procedures are discussed by Sen in [140].

### 3.3.23. Donovan-Murphy Model (1999)

Donovan and Murphy [141], [143] and [146] present a new reliability growth model which is claimed to be simpler to plot and provide a better fit to data than the Duane model over the range of slopes normally observed (i.e., $\alpha \leq 0.5$). The model (for MTBF) is derived from variance stabilization transformation theory and takes-on the form $\theta(t) = \alpha \cdot \sqrt{t} + \beta$. Simulation results indicate that their model is "more effective" for growth rates less than 0.50 (which is generally the typical range for growth rates). Numerical examples are presented from two published datasets and yield findings consistent with those of the simulation results.

### 3.3.24. Pulcini's Model (2001)

Pulcini [147] presents an exponential reliability growth model, which incorporates step changes in a system's failure intensity due to engineering design improvements. He gives ML and CI procedures (exact and approximate) for obtaining estimates of the current failure intensity and lifetime expectancy. GOF procedures based on a Cramer-Von Mises TS are also developed. These statistical procedures are based on a scenario where several identical items are put on test and design modifications are introduced to all items at each failure occurrence. A numerical example is given to illustrate the inference, prediction, and test procedures using actual failure-data from a single (unspecified) military tank.

### 3.3.25. Gaver-Jacobs-Glazenbrook-Seglie Model (2003)

Gaver et al. [155] introduce probability models for sequential-stage system reliability growth. These models are appropriate in cases where a system is tested in a series of stages, whereby if a failure occurs in a given stage, later stages are not entered. System success is determined by successful operation in all test stages. At most one defect is assumed to be removed per test. Analytical procedures are developed to calculate the expected probability of field system mission success after completion of a *runs-test*[26], the distribution of the probability of system field mission success after a successful runs-test, and the expected number of individual system tests required to achieve a successful runs-test. *Seglie's stopping criterion*[27] [134] is studied quantitatively through a Bayesian model formulation which suggests the criterion provides a simple and effective test stopping-rule for a range of reasonable cost criterion.

[26] A ***runs-test*** is a sequence of tests that is conducted until a specified number of consecutive successful tests is achieved.
[27] ***Seglie's stopping criterion*** consists of stopping all testing after a successful runs-test is achieved.

## 3.4. Reliability Growth Projection

### 3.4.1. Corcoran-Weingarten-Zehna Model (1964)

Corcoran, Weingarten and Zehna [8] developed the first model for estimating reliability after corrective action. The approach was developed with consideration to estimating reliability in the final stage of development of an "expensive item." The reliability projection is suitable in cases where corrective actions are installed at the conclusion of a single test phase consisting of $N$ independent trials and where the number of trial outcomes of interest is a multinomial distributed r.v. with parameters $N$ (total number of trials), $q_0$ (unknown success probability), and $p_i$ (unknown failure probability for failure mode $i = 1,...,k$). Note that since a multinomial model is used, the equality $q_0 + \sum_{i=1}^{k} p_i = 1$ must be satisfied, which models the condition where at most one failure mode can occur per trial. In addition to deriving an exact expression for system reliability under the conditions above, Corcoran, Weingarten and Zehna presented seven different estimators and evaluated them in light of criterion typically adopted for that of point estimation (i.e., bias, consistency, conservatism, and ML). By studying these estimators they showed that an unbiased estimate of the corrected system could not be obtained. They were the first researchers to advance the idea of reducing initial failure probabilities by a fractional amount with consideration to fix effectiveness. By their model, the expected reliability (under competing risks) at the end of the current test phase is given by,

$$R\left(N \mid p_i\right) = R_I + \sum_{i=1}^{k}\left(1-d_i\right)\cdot p_i \cdot\left[1-\left(1-p_i\right)^N\right] \quad (1)$$

where $N$ is the total number of failures, $p_i$ is the failure probability of failure mode $i$, $R_I$ is the initial reliability, and $d_i$ is the FEF of failure mode $i$. Other work on this model was done by Dahiya [37] who showed that six of the seven estimators initially considered by Corcoran et al. possess the same limiting normal distributions. Thus for large samples, CI and GOF procedures follow directly. Olsen [36] showed how some of the estimators could be utilized under a multi-stage test program and developed a suitable variant of Corcoran's model in this case.

### 3.4.2. AMSAA-Crow Model (1982)

The ACPM was developed by Crow [54] and [56] for estimating system reliability at the beginning of a follow-on test phase. The model takes into consideration the reliability improvement from delayed fixes only, and is suitable for systems whose test duration is continuous. The primary framework for reducing the initial failure rates follows directly from Corcoran, Weingarten, and Zehna's model [8]. Two GOF procedures have been developed for the ACPM and are discussed by Ellner in [144]. The first procedure is based on a Cramer von Mises TS for grouped data. The second procedure, which is of the chi-squared type [69], is for individual failure time data. The ACPM is one of the first models to incorporate the important concept of *reliability growth potential*[28] [63]. No CI procedures have been reported.

[28] ***Reliability growth potential*** is the upper-limit on reliability achieved by finding and correcting all failure modes in a system with a specified fix effectiveness.

As Crow discusses in [93], this model is also an international standard adopted by IEC and ANSI [120]. The model for expected failure intensity at the end of the test phase is given by,

$$\rho(T \mid \lambda_i) = \lambda_A + \sum_{i=1}^{m}(1 - d_i) \cdot \lambda_i + \bar{d} \cdot h(T) \tag{2}$$

where $\lambda_A$ is the portion of system failure intensity associated with A-modes, $m$ is the total number of observed failure modes, $\lambda_i$ is the failure rate for failure mode $i$, $\bar{d}$ is the average FEF for observed failure modes, and $h(T) \equiv \lambda \cdot \beta \cdot T^{\beta-1}$ (i.e., the failure intensity function of the Weibull process) is the rate of occurrence of new failure modes.

### 3.4.3. Ellner-Wald AMPM Model (1995)

AMPM was developed by Ellner and Wald [121] and is the first projection model to estimate reliability under an arbitrary corrective action strategy. The benefit of this is that the system's configuration with respect to design and reliability need not be constant. The model provides estimates of: 1) the expected number of B-modes observed; 2) the percent surfaced of the B-mode initial failure intensity; 3) the rate of occurrence of new B-modes and 4) the projected system reliability. Figures 9-12 below show an example of each these model equations, comprising the robust reliability growth methodology of AMPM. The model also provides estimates of the reliability growth potential. Estimation procedures are given for both grouped data and individual failure-time data in [144]. No CI procedures have been reported. GOF procedures are discussed by Broemm in [163].

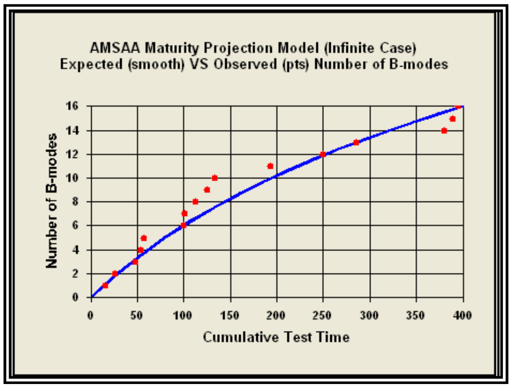


**Figure 9. Expected No. Modes.**

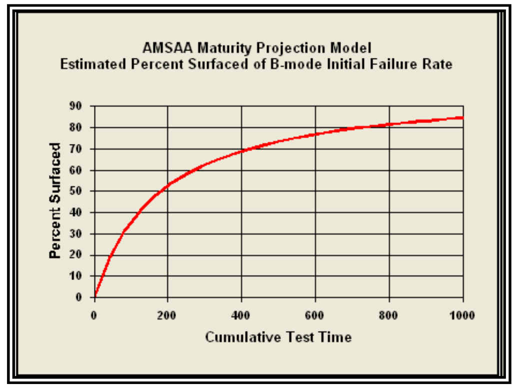


**Figure 10. Percent $\lambda_B$ Observed.**

The model equations shown in Figures 9 and 10 are given by,

$$\mu(t) = \left(\frac{\lambda_B}{\beta}\right) \cdot \ln(1 + \beta \cdot t) \tag{3}$$

and

$$\theta(t) = \frac{\beta \cdot t}{1 + \beta \cdot t} \tag{4}$$

respectively. In these expressions $\lambda_B$ is the portion of system failure intensity associated with B-modes and $\beta$ scale parameter of the gamma distribution.

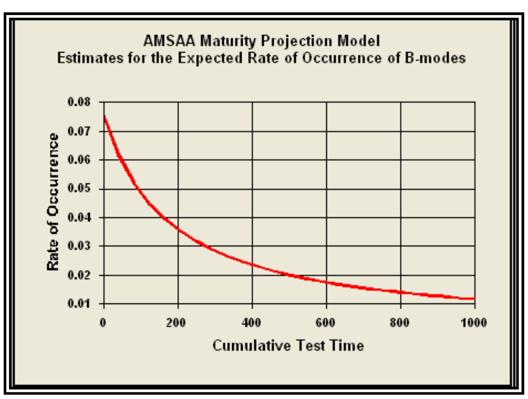


**Figure 11. ROC of New Modes.**

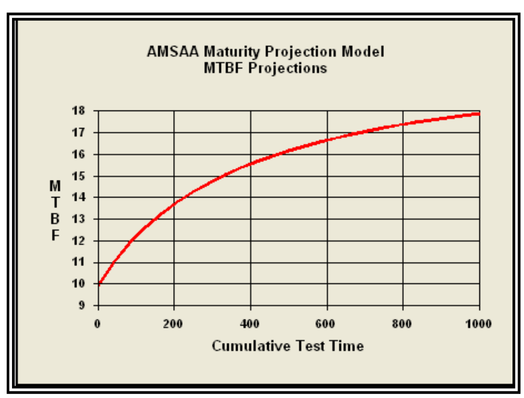


**Figure 12. Reliability Growth.**

The model equations shown in Figures 11 and 12 are given by,

$$h(t) = \frac{\lambda_B}{1 + \beta \cdot t} \tag{5}$$

(e.g., the rate of occurrence of new failure modes) and

$$\rho^{-1}(t) = \frac{1}{\lambda_A + (1 - \bar{d}) \cdot [\lambda_B - h(t)] + h(t)} \tag{6}$$

where $\lambda_A$ is the portion of system failure intensity associated with A-modes and $\bar{d}$ is the average FEF for observed failure modes.

### 3.4.4. Clark's Model (1999)

Clark [139] argues that reliability is often overlooked during early system development and many programs experience late growth programs not long before production as a result. He notes that the "popular AMSAA models" are difficult to apply in these cases since they prescribe high test durations even for aggressive growth rates. He formulates a model as an alternative for projecting reliability growth late in development, which is claimed to overcome these shortcomings. The proposed model consists of two main extensions of the ACPM [54]. The first extension includes a technical modification to allow the model to be applied in the case were fixes can be delayed or non-delayed (rather than all delayed). The second extension includes adding a term for the inherent failure rate of the system to determine how close the current reliability is to the maximum that can be achieved and decide when further growth is no longer time or cost effective. The method is illustrated via numerical example on the Airborne Warning and Control System Radar System Improvement Program. The results indicate that the model generally projected system reliability well, except when new failure modes introduced into the system by software modifications were not accounted for. Clark's model for the projected system failure intensity at time $t_f$ computed at current test time $t$ is given by,

$$\lambda_T\left(t_f,t\right)=\lambda_I+\lambda_F\left(t\right)-\bar{d}\cdot\lambda_{SF}\left(t_f,t\right)-\lambda_{VF}\left(t\right)+\lambda_U\left(t\right) \tag{7}$$

where $\lambda_I$ is the inherent failure rate, $\lambda_F\left(t\right)$ is the fixable failure rate at time $t$, $\lambda_{SF}\left(t_f,t\right)$ is the fixable failure rate at test time $t$ scheduled to be explicitly corrected by future test time $t_f$, $\lambda_{VF}\left(t\right)$ is the fixable failure rate verified to be explicitly or implicitly corrected at current test time $t$, $\lambda_U\left(t\right)$ is the unobserved failure rate at test time $t$, and $\bar{d}$ is the average FEF for observed failure modes.

### 3.4.5. Ellner-Hall AMPM-Stein Model (2004)

The AMPM-Stein model, given by Ellner and Hall [162], is used to estimate the system reliability following correction of known failures modes when fixes are delayed to the end of the test. The benefit of this approach is increased accuracy obtained by using a shrinkage factor estimator (e.g., Stein [50]) designed to minimize the expected sum of squared error. The unique feature about this estimation procedure is that all estimates of failure rates are finite and positive (whether they are observed in testing or not observed in testing). Monte Carlo simulations conducted by AMSAA [156] indicate that the accuracy in the reliability projection associated with AMPM-Stein is greater than that of the international standard adopted by IEC and ANSI [120], namely, the ACPM [54]. The model for system failure intensity is given by,

$$\rho(T) = \sum_{i \in obs} (1 - d_i) \cdot \tilde{\lambda}_i + \left(1 - \frac{m}{k}\right) \cdot (1 - \theta_S) \cdot \left(\frac{N}{T}\right) \tag{8}$$

where $d_i$ is the FEF for observed failure mode $i$, $m$ is the total number of observed failure modes, $k$ is the total potential number of failure modes in the system, and $N$ is the total number of failures observed by time $T$. The shrinkage factor estimate for individual failure mode rates of occurrence is given by,

$$\tilde{\lambda}_i \equiv \theta_S \cdot \hat{\lambda}_i + (1 - \theta_S) \cdot \frac{\sum_{j \in obs} \hat{\lambda}_j}{k} \tag{9}$$

where

$$\theta_S = \frac{Var(\lambda_i)}{\left(\frac{\sum_{i=1}^{k} \lambda_i}{k \cdot T}\right) \cdot \left(1 - \frac{1}{k}\right) + Var(\lambda_i)} \approx \frac{\beta \cdot T}{1 + \beta \cdot T} \tag{10}$$

and $\beta$ is the scale parameter of the gamma distribution and $obs$ represents the index-set of observed failure modes.

### 3.4.6. Crow-Extended Model (2005)

The purpose of the Crow-Extended model [157] is to estimate reliability in the case were corrective actions can be either delayed or non-delayed (i.e., the same as that for AMPM). This model is a trivial extension of two existing AMSAA models, namely, ACPM [54] and RGTMC [30]. Once again, this model is based on the *Duane Postulate*. Estimation procedures follow from the two existing models. Crow [157] also provides 33 metrics useful for managing a reliability growth program and introduces the notion of a further failure mode classification scheme (e.g., BD-modes[29] and BC-modes[30]). The model for system failure intensity is given by,

$$\rho = \lambda_{CA} - \lambda_{BD} + \sum_{i \in obs} \left(1 - d_i\right) \cdot \lambda_i + \bar{d} \cdot h\left(T \mid BD\right) \tag{11}$$

where $\lambda_{CA}$ (e.g., based on the tracking model [30]) is the achieved failure intensity before incorporation of BD-modes, $\lambda_{BD}$ is the constant failure intensity associated with the BD-modes, $d_i$ is the FEF for failure mode $i$, $\lambda_i$ is the failure rate of failure mode $i$, $\bar{d}$ is the average FEF of observed failure modes, and $h\left(T \mid BD\right) = \lambda \cdot \beta \cdot T^{\beta - 1}$ is the rate of occurrence of new BD-modes.

## 3.5. Reliability Growth Surveys and Handbooks

### 3.5.1. Crow's Abbreviated Literature Review (1972)

Crow [23] presents an abbreviated literature review of some reliability growth models. A limited number of numerical examples are also presented. The growth models include: Weiss [1], Lloyd-Lipow [4], Wolman [6], Duane [7], Barlow-

[29] ***BD-modes*** are "B-Delayed modes" that will not be corrected until the end of the current test phase.
[30] ***BC-modes*** modes are "B-Corrected modes" that will be corrected during testing.

Scheuer [12], Virene [16], and Pollock [19], all of which are discussed herein. A similar review of these early reliability growth models is discussed in [31] and [35].

### 3.5.2. DoD's First Military Handbook on Reliability Growth (1981)

Military Handbook 189 "Reliability Growth Management" [49] is the U.S. DoD's first handbook on reliability growth. The handbook was developed by the U.S. AMSAA with Crow as the principle author of the document. It was first published in February 1981. The handbook offers techniques to enable program managers of DoD weapon systems to plan, evaluate, and control the reliability of their systems during the development process. It also provides procuring activities, and defense contractors with an understanding of the concepts and principles of reliability growth, as well as offer guidelines and procedures to be used in managing a reliability growth program. In the main body of the handbook, two models are briefly introduced including the MIL-HDBK-189 model [49], and the RGTMC [30], both of which are discussed above. In Appendix B of the handbook, eight discrete and nine continuous reliability growth models are summarized. The discrete models include: two Lloyd-Lipow models [4], Wolman's model [6], the Barlow-Scheuer model [12], Virene's Gompertz model [16], and Singpurwalla's model [45]. The continuous reliability growth models include: Duane [7], RGTMC [30], Cox-Lewis model [11], Lewis-Shedler model [33], Rosner's model [3] and a variant thereof, a continuous Lloyd-Lipow model [4], Aroef's model [2], and an unreferenced exponential model for cumulative MTBF. In Appendix C, the RGTMC [30] and associated statistical procedures are discussed in more detail. Dr. Paul Ellner (AMSAA) is currently supervising the DoD's revision of this handbook.

### 3.5.3. Fries-Sen Survey on Discrete Reliability Growth Models (1996)

Fries and Sen [127] present a comprehensive compilation of model descriptions and characterizations, as well as discuss related statistical methodologies for parameter estimation and CI construction. The interrelationships and assumptions that underlie the various models is also presented. Their survey is extensive covering: single-stage models (e.g., Corcoran et. al [8]), multi-stage models (e.g., Lloyd-Lipow models [4] and [66, 72]), trinomial models (e.g., Wolman [6], Barlow-Scheuer [12], Weinrich-Gross [43], Mazzuchi-Soyer [110]), Bayes models (e.g., Pollock [19], Kaplan et al. [90], Jewell [65]), exponential-growth models (e.g., Lloyd-Lipow [4], Sriwastav [44]), exponential-regression models (e.g., NASA [10], Gross and Kamins [15], Virene [16], Bonis [39]), learning-curve models (e.g., Duane [7], RGTMD [55]) and several others. This survey is the most comprehensive available on the subject. Smaller-scope reviews of discrete models (more limited in detail and covered by the Fries-Sen survey) are given by Jayachadran and Moore [34], Balaban [41], Dhillon [48], Gates [68], and Woods, Drake & Chandler [75].

### 3.5.4. DoD's Guide for Achieving RAM (2005)

In August 2005, the OSD published a guide on achieving RAM [163] for DoD systems. Appendix C of the document is devoted solely to methods for reliability growth analysis. A number of associated concepts are discussed including: reliability maturity metrics for failure mode coverage and fix effectiveness, as well as some reliability growth planning, tracking, and projection models. The growth models that are discussed include: RGTMC [30], RGTMD [55], Corcoran-Weingarten-Zehna Model [8], ACPM [54], AMPM [121], Crow-Extended [157], AMPM-Stein [162], and MIL-HDBK-189 model [49].

## 3.6. Other Literature (i.e., Theoretical Results, Perspectives and Applications)

### 3.6.1. Corcoran-Read Simulation Study (1967)

Corcoran and Read [13] present a simulation study (first outlined by them in [13]) of four reliability growth models available at the time. These models include: Chernoff-Woods [5], Barlow-Scheuer [12], Wolman [6] and Lloyd-Lipow [4]. They compare the reliability estimates of these methods with three measures of effectiveness: the popular average squared-error, the average squared-error after applying an inverse sine transformation (used to stabilize the variance of success probability estimates), and a logarithmic transformation applied to the failure probabilities (i.e., the average of the absolute deviation of logarithms of the ratio of error in the failure probabilities). Based on these measures of effectiveness, they conclude that their "general ranking of the preferability" of these methods are 1, 4, 2, and 3, respectively.

### 3.6.2. Barr's Paper (1970)

Barr [21] considers a class of reliability growth models that accommodate for variations in several important factors including: the interdependencies of assignable-cause failure modes, the inclusion of an inherent failure mode, the repair policy and the distribution of initial states of the system. His paper is an exposition of several prediction models appearing in the early literature of reliability growth and identifies their general features. The methods considered include those of Lloyd-Lipow [4], Pollock [19], Weiss [1], and Wolman [6]. The overall problem Barr considered is that of predicting (before testing is undertaken) what the reliability of the system will be after a sequence of trials, and to predict the number of trials required to attain a given reliability. He divides this general class of reliability growth models in three types: single assignable-cause mode models, multiple equally likely assignable-cause mode models, and multiple assignable-cause modes not necessarily equally likely.

### 3.6.3. Read's Remark on Barlow-Scheuer Estimation Scheme (1971)

Read [22] notes that the Barlow-Scheuer estimation procedure is incomplete. He notes that this is due to not addressing the case where all trials of a stage result in only inherent failures. Read proposes a policy to handle the case which allows estimation of the trial probability of assignable-cause failures.

### 3.6.4. The AMSAA Reliability Growth Symposium (1972)

Crow [24] provides conference proceedings on, apparently, the first and last reliability growth symposium sponsored by the U.S. AMSAA. The conference took place 26-27 September 1972 and originated as an outgrowth of the recommendations of the Panel on Accelerated Development of Reliability. This panel was chaired by Jack Hope who was then serving on the White House Staff. The purpose of the symposium was to enhance the state of technical and managerial knowledge on reliability growth methodology to benefit the Army's materiel acquisition process. There were over 200 attendees and six papers given. The papers include Selby-Miller [20], Virene [16], Crow [24], Barlow-Proschan-Scheuer [24], Barlow [24], and Corcoran-Read [14].

### 3.6.5. Langberg-Proschan Theoretical Paper (1979)

Langberg and Proschan [47] present theoretical results on converting reliability growth (or decay) models involving dependent failure times into equivalent models involving only independent random variables. They consider a sequence of such conversions occurring at successive points in time where the independent random variables are becoming stochastically larger (reliability growth). Ultimately, they demonstrate that the limiting distributions in the sequence of dependent models "correctly correspond" to the limiting distributions in the sequence of independent models. No practical reliability growth model is presented, rather, associated results are mainly theoretical and focus on the technical aspects on the aforementioned conversion.

### 3.6.6. Jewell on Learning-Curve Models (1984)

Jewell [65] constructs a general framework for learning-curve reliability growth models with Bayesian estimation procedures for model parameters. He argues that Bayesian estimation methods must be used to incorporate engineering experience in prior estimates of the parameters of learning-curve models because ML estimators may be very inaccurate and unstable. His main conclusion is that the majority of learning-curve developmental test programs will provide insufficient data to reach the desired precision for manufacturers to make early predictions on reliability when using traditional methods. In particular, he indicates that the use of the Duane learning-curve $g(t) = k \cdot t^{\nu-1}$ leads to technical difficulties in reliability growth applications and that an exponential learning-curve $g(t) = \exp(-\nu \cdot t)$ avoids such problems.

### 3.6.7. Wong's Letter to the Editor (1988)

Wong [79] discusses the lack of organization in the vast literature on reliability growth and identifies some process that are not reflected in associated methodologies. These process include: equipment aging effects, manufacturing learning-curve (i.e., improvement in production processes over time for a single item), industry-wide part improvement (e.g., effects of industry burn-in for electronic components), and methods on the test, analyze, and fix process that only use test duration as an independent variable (e.g., Duane's model [7]). He suggests that authors of reliability growth papers should: specify what kind of reliability growth process they are modeling, and which factors in their model are held constant or randomized to smooth-out effects.

### 3.6.8. Wronka's Application of the RGTMC (1988)

Wronka [77] shows the benefits that can be obtained by conducting reliability growth tracking early in the development process. He gives an application of the RGTMC [30] for prototypes of a circuit card assembly. Results are presented for the estimation procedure of grouped data and associated GOF test.

### 3.6.9. Benton and Crow on Integrated Reliability Growth Testing (1989)

Benton and Crow [81] consider the development of reliability growth under integrated reliability growth testing. By integrated testing they refer to a development program consisting of: functional testing, environmental testing, safety testing, performance testing, mobility testing, and dedicated RAM testing. They discuss and apply the concepts of the MIL-HDBK-189 model [49], RGTMC [30] and the ACPM [54] under the framework of these types of integrated tests. They also presented results and lessons-learned on some Army programs. Years later, Crow, Franklin and Robbins [118] present a successful application of integrated reliability growth testing in the development of a large switching system. Their claimed benefits include: timely analysis of failed items, accurate problem classification, accurate laboratory failure rates, early identification of failure modes, management metrics for reporting, and reliability growth achievement using all test resources available.

### 3.6.10. Frank's Corollary of the Duane's Postulate (1989)

In [82] Frank discusses his observation that various types of avionics equipment are found to demonstrate remarkably similar gradual declines in reliability during prolonged service. He proposes a modification of Duane's learning-curve approach by extending its applicability to project a reliability profile over an equipment's planned service life. Frank claims his "revised equations" (not given) can be used to predict changes in equipment reliability, thus providing a capability to more accurately estimate life-cycle support resource requirements and costs.

### 3.6.11. Gibson-Crow Estimation Method for Fix Effectiveness (1989)

Gibson and Crow [83] develop a "practical and statistically sound" methodology for estimating the average FEF, which is a parameter utilized in some reliability growth models (e.g., Ellner's AMPM [121] and Crow's ACPM [54]). The average FEF, $\mu_D$, is basically estimated by using the ACPM in a reverse manner. In this approach, Gibson and Crow estimate the portion of the system failure intensity in a follow-on test by the common reliability point-estimate, $\lambda = f / T$ (e.g., failures over test time). This value is then equated to the reliability projection equation given by the ACPM. The equation is then algebraically manipulated to solve for the average FEF.

### 3.6.12. Woods' Study on the Effect of Discounting Failures (1990)

Woods [88] analyzes the effect of *failure discounting*[31] on the accuracy of two discrete and two continuous reliability growth models. The discrete models include: the Chernoff-Woods exponential regression model [5], and Crow's RGTMD [55]. The continuous models include: Crow's RGTMC [30] and a modification of the same model that only uses data in a given phase (i.e., not cumulative data). Woods concludes that failure discounting has a greater impact on the cumulative growth models than on the non-cumulative (i.e., there is greater bias in the cumulative models, thus yielding more optimistic reliability estimates). He also indicates that the non-cumulative growth models tracked growth patterns better than the cumulative.

[31] ***Failure discounting*** is the practice of removing fractions of the previous failures after corrective action has been taken, where no failures for the same cause reoccur in subsequent testing.

### 3.6.13. Higgins-Constantinides Application (1991)

Higgins and Constantinides [91] present and interesting reliability growth application of the U.S. Navy's EMATT system. EMATT is an open-ocean one-shot expendable target used in simulating combat missions. They faced the dilemma where no one-shot reliability growth model was suitable for their purposes. Additionally, application of established continuous growth models were deemed inappropriate[32] since the number of trials in each test phase was not relatively large nor was the reliability high. Since none of the classical growth models available at the time could provide suitable approximations, the reliability growth approach adopted for EMATT consisted of fundamentals from the Duane model [7]. Thus, they constructed a reliability growth tracking curve by plotting the cumulative reliability (i.e., cumulative successes over cumulative trials) versus cumulative trials. The results indicate a general reliability improvement trend and they note the difficulty in obtaining precise numerical reliability estimates with limited trials. In there final report [113] published two years later, they apply the RGTMD [55]. The results show that EMATT reliability grew from 0.4 to 0.8 over the several year development program, hence, demonstrating its reliability objective requirement of 0.8 (as a point-estimate).

### 3.6.14. IEC International Standards for Reliability Growth (1991)

[32] Continuous growth models were deemed inappropriate since they can only be utilized as good approximations for tracking the reliability of one-shot systems in the case where the number of trials within each test phase is relatively large and the reliability was relatively high, as stated in MIL-HDBK-189 [49].

The IEC adopted two international standards for reliability growth: IEC Standard 1014 covering "Programs for Reliability Growth," and draft IEC Standard 56 (Central Office) 150 on "Reliability Growth and Estimation Methods." IEC Standard 1014 was issued in 1989 and gives guidelines for improving the reliability and exposing the weaknesses of hardware and software items. This standard also presents basic concepts and descriptions of management, planning, testing, failure analysis, and corrective action techniques. The final draft of IEC Standard 56 (Central Office) 150 became IEC Standard 1164 [120] and was issued in 1995. This standard describes Crow's NHPP power-law reliability growth model [30] and related projection model, ACPM [54]. Step-by-step directions on their use is given. All statistical methods for the models are discussed including: MLE, CI, and GOF procedures for failure and time-truncated data. Both standards are discussed by Crow in [93].

### 3.6.15. Coolas' Application (1991)

Coolas [94] presents a dynamic reliability prediction technique for the DPS 7000, which is a mainframe computer system. Observed measures of field performance, and trends in reliability growth (due to evolving product maturity) are identified. The proposed reliability predictions are based on adjustments to component reliability and reliability growth models following from these observations. As a result of the reliability predictions and associated improvement program, more accurate spare parts provisioning and decrease in maintenance costs are claimed to have been achieved. While not specifically referenced, the Lloyd-Lipow model [4] appears to be used for component level reliability growth assessment.

### 3.6.16. Bieda's Application (1991)

Bieda [95] presents an analysis addressing product / process design concerns and validation testing issues via reliability growth testing, monitoring, and assessment. The integration of reliability growth test techniques is applied to evaluate the reliability of an unspecified electro-mechanical device. Reliability growth tracking curves are developed using Duane's model [7] and the various relationships between design iterations are identified. Product assurance analyses are performed to help identify design and process-related concerns. Point-estimates and one-sided LCB on MTBF are given using the NHPP Weibull process [26]. Results consist of successful demonstration of the relationship between failure detection and corrective action, as well as the achievement of higher reliability through reliability growth testing and use of reliability growth tracking methods.

### 3.6.17. Ellis' Robustness Study (1992)

Ellis [104] examines the robustness of techniques applied to failure time data to determine if the system failure intensity is changing over time. The techniques include: the Duane model [7], and the RGTMC [30]. Monte Carlo methods are utilized to simulate failures. MLE procedures based on time-truncated and grouped data are used to approximate model parameters and associated MTBF. Some basic advantages and disadvantages of the models are discussed. The results of the study indicate that the Duane model indicates reliability growth even in cases when the failure data are generated by an exponential distribution, and that the RGTMC is more suitable for detecting the presence of reliability growth or decay. This is largely due to the more sophisticated statistical procedures developed for the NHPP Weibull process (e.g., point-estimation, CI construction, and GOF testing).

### 3.6.18. Calabria-Guida-Pulcini Bayes Procedure for the NHPP (1992)

Calabria, Guida and Pulcini [103] develop a Bayesian estimation procedure for the parameters of the NHPP power-law process, originally developed by Crow in 1974 [26]. They provide Bayes estimates of system reliability and the failure intensity for failure-truncated testing. Their Monte Carlo simulation results show that the procedure is more accurate and efficient than that of ML, even for vague prior information. Years later, they present [128] a nonparametric Bayes-decision framework for complex repairable systems.

### 3.6.19. Meth's OSD Perspective on Reliability Growth (1992)

Meth, who at the time was Director of the Weapons Support Improvement Group of the OSD, gives a critical review [101] of reliability growth "myths and methodologies." He asserts that reliability prediction is not a reasonable application of reliability growth and that the various mathematical models may not adequately describe the reliability growth process. He conjectures that understanding of the factors for test planning has not advanced beyond the rules-of-thumb that were initially proposed by Duane [7] in 1964. Meth also challenges the reliability community to "reexamine the reliability growth concept" and how it is being applied.

### 3.6.20. Demko on Non-Linear Reliability Growth (1993)

Demko [112] identifies a shortcoming to the Duane model [7], namely, that it is insensitive to discontinuities or sudden changes in the reliability growth trend for a system. In other words, Duane's model only considers linear growth on a log-log scale and will not accurately portray non-linear growth (on the same scale). Demko proposes to utilize non-linear, piecewise regression to overcome this shortcoming. Several numerical examples and plots are given to illustrate comparisons of Duane's approach versus that of the proposed. The examples use datasets from programs that demonstrated non-linear growth patterns and show that the proposed method more accurately portrays the growth patterns.

### 3.6.21. Farquhar and Mosleh on Growth Effectiveness (1995)

Farquhar and Mosleh [124] present an approach for quantifying reliability growth effectiveness. In their approach, they develop a performance parameter, which they present from two perspectives on whether data from reliability growth testing is, or is not, available. If data is not available, a subjective assessment and characterization of attributes that are indicative of the corporate culture is used. When data is available, the parametric variable is quantified by normalizing past performance with reliability growth program goals. Five case studies were utilized to develop the performance parameter. It was then incorporated into an existing reliability growth model, known as the Tracking, Growth and Prediction model. This model was developed by P. F. Verhulst in 1845 and is based on the logistic function characterized by an S-shaped curve. They conclude that their modification to the model provides a conservative estimate of the risk involved in achieving reliability growth goals.

### 3.6.22. Demko on Reliability Growth Testing (1995)

Demko [123] argues that certain types of testing are not adequate in exposing field-related failure modes. Some of the types of testing mentioned includes: RDGT, EQT, and ESS tests. He claims that these tests yielded a high percentage of failure modes that occur only in a chamber-type environments and are not representative of failure modes that would be encountered during field use. Failure modes from over 2 million hours of field data from 13 different types of "283 Avionics Units" are compared against failure modes identified by 5 different companies who performed either RDGT (21K hours), ESS tests (28K hours) and EQT (hours not given). Several plots are given to compare the quantity of failure modes encountered in the field, and in each type of test. The results indicate that the majority of failure modes were found during field use, following ESS testing, RDGT, and EQT.

### 3.6.23. Fries-Maillart Stopping Rules (1996)

Fries and Maillart [125] present a method of when to stop testing of a one-shot system when the number of systems to be produced is predetermined and the probabilities of identifying and successfully correcting each failure mode are less than one. The stopping criterion is emphasized on maximizing the number of systems expected to perform successfully in the field after deployment of the lot. Two rules are presented. The first rule includes stopping the test when the estimated *utility*[33] (given a failure on the next trial) is less than or equal to the current utility estimate. Motivated by expected value, the second rule is to stop testing when the estimated utility after the next trial (regardless of its outcome) is less than or equal to the current utility estimate. Four discrete reliability growth models are utilized to estimate reliability improvement via Monte Carlo simulation. The models include: two Lloyd-Lipow models [4], Fries' learning-curve model [111], and Virene's Gompertz model [16]. The results indicate that both stopping rules perform well and can be practically implemented. Specific recommendations are given to implement test-stopping rules in light of several factors, such as, estimation methodology and lot size.

### 3.6.24. Ebrahimi's MLE for the NHPP (1996)

[33] ***Utility*** is defined as the number of systems expected to perform successfully in the field after deployment of the lot.

Ebrahimi [129] develops a general formulation for modeling reliability growth between design modifications. He assumes the model is either a Poisson process or the NHPP power-law process, and the times of design modifications must be known. ML estimates and CI procedures are developed in two cases, depending on the presence of constraints on the system failure intensity. His proposed MLE and CI procedures are illustrated via a limited Monte Carlo example. Corrections to several typographical and computational errors in Ebrahimi's paper is given years later (i.e., 2002) by Pulcini in [149].

### 3.6.25. Huang-McBeth-Vardeman One-Shot DT Programs (1996)

Huang, McBeth, and Vardeman [130] develop a method to efficiently conduct developmental testing of one-shot systems that are destroyed in testing upon first use. Dynamic programming is used to identify optimal test-plans that maximize the mean number of effective systems of the final design that can be purchased with the remaining budget. Several suboptimal rules are also considered and their performances are compared to that of the optimal rule.

### 3.6.26. Xie-Zhao Monitoring Approach (1996)

Xie and Zhao [131] introduce a method called the First-Model-Validation-Then-Parameter-Estimation approach. Their approach, which essentially follows directly from Duane [7], consists of model validation and parameter estimation. The model is validated by plotting the cumulative number of failures versus test duration, using linear regression to derive an equation to fit the data, and calculating the associated correlation coefficient. A subjective assessment of the magnitude of the correlation is the deciding factor on model validation. If the model is reasonable, the equation can be used for prediction purposes. The paper focuses a great deal on the need for more graphical models in reliability growth.

### 3.6.27. Seglie's OSD Perspective on Reliability Growth (1998)

Written from the perspective of the Chief Scientific Advisor to the Director of Operational Test and Evaluation, OSD, Seglie [134] argues that too many weapon system programs enter operational testing before they are ready (i.e., they have immature design margin with respect to reliability and ultimately fail their test objectives at great cost). Seglie emphasizes the false predictions that can be reported from the wealth of reliability growth models available, and notes that growth models have demonstrated a poor history of successfully predicting field reliability. He proposes that the role for reliability growth modeling should be focused on prescribing test duration required to reach a level of acceptable reliability before going into operational testing – and not on estimating system or subsystem reliability. He adds that this modest role of reliability growth methodology in developing test plans is still of great importance in determining the amount of time required for engineers to find dominant failure modes, analyze them, develop and implement corrective actions, and confirm fix-effectiveness. In his overall view, the most important attribute of reliability growth should be to provide information to help programs succeed during test and evaluation. To better do this, he suggests that growth models need to account for the effects of different environments, be system specific, and be more engineering based.

### 3.6.28. Hodge-Quigley-James-Marshal Framework (2001)

While no technical details are provided, Hodge et al. [148] discuss a modeling framework that aims to support reliability enhancement decision-making. The main objectives of their approach are to: develop a methodology to support reliability enhancement throughout the design process, and develop a model, referred to as the Reliability Enhancement Methodology Model (REMM), that facilitates the assessment of reliability throughout the product lifecycle. REMM is basically a tracking system to determine how reliability evolves throughout the design process / lifecycle by integrating statistical and engineering understanding of reliability performance. The primary outputs of REMM include point and CI estimates for: product reliability, probability of failure per unit time, and the probability of failure free periods. A list of engineering design concerns (i.e., failure modes, corrective actions) on a given product are also provided. The methodology was implemented by TRW Aeronautical Systems (Lucas Aerospace) in 2001.

### 3.6.29. Crow's Methods to Reduce LCC (2003)

Crow [150] develops methods for estimating the useful life of a fleet of repairable systems and presents a model for addressing in-service reliability growth. A minimum LCC model and associated MLE procedures are also developed. The methodology is fundamentally based on the NHPP power-law process, originally developed by Crow in 1974 [26]. A numerical example is given for 11 simulated systems to illustrate the LCC methodology. The in-service reliability growth approach follows from the ACPM [54] with a slight modification where an average FEF is utilized (i.e., rather than the individual FEF). All previously reported MLE and GOF procedures apply. A numerical example is given on simulated data for cases with and without the prevalence of wear-out.

### 3.6.30. Gurunatha-Siegel Six-Sigma Process (2003)

Gurunatha and Siegel [151] formulate 12-step six-sigma quality process of an unspecified complex commercial product developed by Xerox Corporation. As a result of implementing this process, they claim that the company achieved a reliability growth rate that exceeded that of any other program within their corporate history. The 12-step process includes: (1) material selection optimized for reliability and cost, (2) failure mode identification and physics of failure analysis, (3) total LCC calculated for each component, (4) ALT performed with lifetime predictions, (5) accuracy measurement on product and process capability, (6) identify critical parameters and their percent contribution to: survival and (7) failure, (8) Monte Carlo simulation on critical parameter uncertainty, (9) design-of-experiments on reliability characteristics, (10) subsystem design modifications to extend product life, (11) statistical process-control for process improvement and, (12) use of lessons-learned for improved reliability growth processes. No quantitative details are provided on the claimed reliability growth achievement of the product.

### 3.6.31. Yadav-Singh-Goel Approach (2003)

Yadav, Singh, and Goel [152] propose a two-stage model of system reliability growth that they develop with consideration to associated components, functions, and failure modes. The first stage consists of development of a reliability growth plan to achieve program requirements. The second stage of the framework involves a strategy to further improve the system reliability prediction following demonstration of its requirements. The prediction is decomposed by component and a prioritization index is defined to provide a rank order of components based on their potential for improving the accuracy of the system-level reliability prediction. A series system configuration is assumed, and the reliability requirement is allocated equally over all components. A gamma prior distribution is utilized under a Poisson sampling routine, which results in the typical gamma posterior for the distribution of component failure rates. Improvements in the associated system-level reliability prediction are improved by a variance reduction strategy on the component gamma posterior distributions. Methodologies for test cost estimation, and reliability improvement prioritization are given. A numerical example on a hydraulic power rack-and-pinion steering system is presented to demonstrate the proposed two-stage model.

### 3.6.32. Quigley-Walls CI Procedures (2003)

Quigley and Walls [153] develop inference properties for reliability growth analysis. They assume a Poisson prior distribution for the ultimate number of faults that would be exposed in the system if testing were to continue ad infinitum. Although, they estimate the parameters of the system failure intensity function empirically. Bias and conditions of existence of fixed-point iteration MLE procedures are investigated. The intention of the approach is to support reliability inference in situations where failure data are sparse. Their statistical CI procedures are shown to be suitable for small sample sizes and is demonstrated by numerical example.

### 3.6.33. Smith on Planning (2004)

Smith [158] describes a process for planning and estimating the cost of a reliability growth program under a Performance Based Logistics (PBL) contract. Under this type of contract, a supplier is typically responsible for structuring their reliability and support programs around a defined field availability or reliability goal. This planning process was developed with the intent to minimize cost and performance risks in the execution of a long-term PBL support contract for a complex, repairable system. The process mainly includes a detailed assessment of five areas: expected volume of field usage (e.g., flight hours, mileage), a well-defined field reliability definition, estimates of current field reliability, the goal or future requirement for field reliability, and a schedule for reliability achievement.

### 3.6.34. Krasich and Quigley on the Design Phase (2004)

Krasich and Quigley [159] discuss how there has been significantly less attention on reliability growth during the design phase (i.e., most of the literature is developed for growth during the TAFT process). They propose two models that could be utilized to assess reliability growth during design. The first model is a modification of Crow's RGTMC [30], and the second is a modification to Rosner's IBM model [3]. The data required for these models includes: the reliability requirement, a subjective assessment of the initial reliability of the system, an estimate of the number of design modifications, the mean number of faults in the initial design, and an estimate of the effectiveness in mitigating design faults. They indicate that the modified RGTMC is more appropriate when the design activities and modifications are equally spaced and well-planned. Otherwise, in more uncertain situations, the modified IBM model is deemed to be most suitable. The modified RGTMC is illustrated by numerical example to an unspecified industry example.

### 3.6.35. Mortin-Ellner Paper (2005)

Mortin and Ellner [161] address some of the advancements in reliability growth methodology (offered by AMSAA), and also highlight remaining challenges and areas in the field requiring further development. Some of the reliability growth planning models discussed include: SPLAN [144], PM2 [162], the Threshold Program [163], and SSPLAN [117]. A number of tracking models (e.g., Duane [7], RGTMC [30], RGTMD [55], and SSTRACK [126]) and projection models (e.g., AMPM [121], AMPM-Stein [162], ACPM [54]) are also discussed. One of the remaining challenges in reliability growth that they mention includes that, "more projection and tracking methodology needs to be developed for cases where the events are measured on a discrete scale rather than on a continuous basis (e.g., single-shot devices such as missiles).

### 3.6.36. Acevedo-Jackson-Kotlowitz Application (2006)

Acevedo, Jackson and Kotlowitz [165] discuss how reliability growth achievement can be realized by using a well-educated ALT program. Two product case studies are presented to show how Lucent Technologies performs ALT on critical hardware subsystems used in telecommunication systems. The hardware items studied include: an RF power amplifier module, and radio unit. ALT is used to identify product weaknesses leading to performance degradation over simulated operational lifetimes. Weaknesses are corrected through design changes prior to manufacturing and field deployment. Forecasts for the steady-state product reliability under expected field conditions are given. Crow's NHPP Weibull process [26] is used to estimate steady-state reliability under Type I censoring (time-truncated), which is consistent with the conduct of their ALT. The importance of the paper is how ALT can be incorporated into a reliability growth program, as well as how common reliability growth models can be utilized for assessment purposes when test data is obtained under accelerated conditions.

### 3.7. Conclusion

This chapter has provided a synopsis of some of the most significant research that has been done in the field of reliability growth for complex systems. The literature review has answered many questions of basic interest about the existing state-of-the-art, as well as the areas within the field. Summaries of nearly 80 papers were given, which cover 7 planning models, 25 tracking models, 6 projection models, 4 reliability growth surveys or handbooks, and 36 other papers covering theoretical results, simulation studies, real-world applications, personal-perspectives, international standards, or related statistical procedures. Thus, the literature review has identified three main areas of reliability growth including: planning, tracking, and projection. A wide array of statistical procedures (e.g., classical and Bayesian) for point-estimation, confidence interval construction, and goodness-of-fit testing are available for most of the models (not all). Models have been developed for complex systems whose test duration is continuous, as well as for complex systems whose test duration is discrete. The literature review has also revealed capability-gaps mainly for one-shot systems in the areas of planning and projection, as indicated by Mortin and Ellner in [161]. No tailored discrete reliability growth planning models were found. Also, only one discrete reliability growth projection model (i.e., Corcoran et al. [8]) is known to exist. This model is only applicable to the popular competing risks framework where at most one failure mode can be discovered in any given trial. Thus, the model cannot be applied to systems where more than one failure mode is discovered during a single trial. The U.S. Army has encountered this phenomenon, particularly with smart munitions. In some cases, up to 7 failure modes have been discovered during a single flight test.

With respect to the current research topic of discrete projection, there are two types of models that depend on the type of corrective action strategy used by program management. The first type addresses the case were all corrective actions are delayed until the end of the current test phase. The second type addresses the more complicated case where program management adopts an arbitrary corrective action strategy resulting in a mixture of delayed and/or non-delayed fixes. The main difference between the two types of projection models are their functional forms, the data they require, and their statistical procedures involved for parameter estimation. The genesis of discrete reliability growth projection is marked by a paper written by Corcoran, Weingarten, and Zehna in 1964 [8], which addresses the delayed case. Since then, a number of other methods have been developed. Among them include the delayed models given by Crow [54], and Ellner & Hall [162], and the non-delayed models given by Ellner [121], and Crow [157] - all of which are models for systems whose usage is measured in the continuous time domain. Hence, the need for reliability growth projection capabilities for one-shot systems. Chapters 4-8 prescribe reliability growth management metrics and associated statistical procedures that fill these capability gaps under both corrective action strategies.

Additional work that could be done in the area of discrete reliability growth projection includes revisiting the original problem considered by Corcoran et al. [8]. They gave the first model for estimating reliability after corrective action. Their projection is suitable in cases where: corrective actions are installed at the conclusion of a single test phase consisting of $N$ s-independent trials, and where the number of trial outcomes of interest is a multinomial distributed r.v. with parameters $N$ (total number of trials), $p_0$ (unknown success probability), and $q_i$ (unknown failure probability for failure mode $i = 1, \ldots, k$). Note that since a multinomial model is used, the equality $p_0 + \sum_{i=1}^{k} q_i = 1$ must be satisfied, which models the condition where at most one failure mode can occur on any given trial. Thus, the model was developed under the popular competing risks framework. The methodology presented herein allows zero, one, or more failure modes to be discovered during any single trial. The main difference between these two extremes is that the functional form of Corcoran's model (under competing risks) is additive, whereas the models given herein are multiplicative. The specific additional research that could be done includes developing the management metrics given in Chapters 5 and 6 under the competing risks framework. This would require investigation into both corrective action strategies previously mentioned. This may also require tailored statistical procedures (i.e., classical and Bayesian) for point estimation, interval estimation, and goodness-of-fit testing, similar to those given in Chapters 4, 6, 7, and 8.

THIS PAGE IS INTENTIONALLY LEFT BLANK.

- CHAPTER 4 -

# 4. A RELIABILITY GROWTH PROJECTION MODEL[34]


## Abstract

This paper offers several contributions to the area of discrete reliability growth projection. We present a new, logically derived model for estimating the reliability growth of complex, one-shot systems (i.e., the reliability following implementation of corrective actions to known failure modes). Multiple statistical estimation procedures are utilized to approximate this exact expression. A new estimation method is derived to approximate the vector of failure probabilities associated with a complex, one-shot system. A mathematically-convenient functional form for the *s*-expected initial reliability of a one-shot system is derived. Monte Carlo simulation results are presented to highlight model accuracy with respect to resulting estimates of reliability growth. This model is useful to program managers, and reliability practitioners who wish to assess one-shot system reliability growth.



**Keywords**: One-shot system, projection, reliability growth.


## Acronyms[35]

AEC – Army Evaluation Center

AMSAA – Army Materiel[36] Systems Analysis Activity

[34] Chapter 4 was published in the March 2008 issue of *IEEE Transaction on Reliability* (i.e., vol. 57, no. 1, pp. 174-181) as presented herein.
[35] The singular and plural of an acronym are always spelled the same.

AMPM – AMSAA Maturity Projection Model

DoD – Department of Defense

FEF – Fix Effectiveness Factor(s)

FOT – First Occurrence Time(s)

GOF – Goodness-of-Fit

MME – Method of Moments Estimation/Estimate(s)

MLE – Maximum Likelihood Estimation/Estimate(s)

## Definitions

1. Failure mode – a failure event whose occurrence is mitigated via a unique corrective action.
2. Unobserved mode – a failure mode which exhibits zero failures during testing.
3. Observed mode – a failure mode which exhibits at least one failure during testing.
4. Repeat failure mode – a failure mode which exhibits at least two failures during testing.
5. A-mode – a failure mode that will not be addressed via corrective action.
6. B-mode – a failure mode that will be addressed via corrective action, if observed.
7. FEF – fraction reduction in an initial mode failure probability due to implementation of a unique corrective action.

## Notation

[36] Materiel refers to equipment, apparatus, and supplies utilized by an organization or institution, in this case the U.S. Army.

$k$ - total number of potential failure modes.

$m$ - total number of observed failure modes.

$N_{i,j}$ - number of failures for mode $i$ in trial $j$ – zero or unity.

$N_i$ - total number of failures for mode $i$ in $T$ trials.

$p_i$ - true but unknown probability of failure for mode $i$.

$\hat{p}_i$ - MLE of $p_i$.

$\tilde{p}_i$ - theoretical shrinkage factor estimator for $p_i$.

$\theta$ - true but unknown shrinkage factor.

$n$ - beta parameter; pseudo number of trials.

$x$ - beta parameter; pseudo number of failures.

$d_i$ - true but unknown FEF for mode $i$.

$r(T)$ - true but unknown system reliability after mitigation of known failure modes.

$\tilde{r}(T)$ - theoretical approximation of $r(T)$ using $\tilde{p}_i$.

$T$ - total number of trials.

## 4.1. Introduction

### 4.1.1. Background and Motivation

There are three main areas in the field of reliability growth: planning [1], tracking [1], and projection [2]. Each of these areas apply to complex systems whose test durations are continuous, as well as to those whose test durations are discrete. While there are a number of models available in each of these areas for continuous systems, more tracking, and projection models are needed for one-shot systems, as suggested in [3]. In this paper, we present a new reliability growth projection model for one-shot systems. The model will not be suitable for application to all one-shot development programs. But it is useful in cases where one or more failure modes are, or can be, discovered in a single trial; and catastrophic failure modes have been previously discovered, and corrected. The model is unique in the area of reliability growth projection, and offers an alternative to the popular competing risks approach.

A survey of discrete reliability growth models is presented in [4], which consists of a comprehensive compilation of model descriptions, characterizations, and insights on related statistical methodologies for parameter estimation, and confidence interval construction. Of particular interest, Corcoran et al. [5] presented the first reliability growth projection model. The major limitation of the approach, however, is that it cannot be applied to one-shot systems where more than one failure mode is discovered in a given trial. This phenomenon has been encountered on a number of different DoD systems over the years, particularly with smart munitions. This is our primary motivational factor for developing the proposed method in the case considered.

A second motivational factor is associated with statistical estimation. In addition to deriving an exact expression for system reliability, Corcoran et al. [5] presented seven different estimators, and evaluated them in light of criterion typically adopted for that of point estimation including *s*-bias, consistency, conservatism, and maximum likelihood. By studying these estimators, they showed that a *s*-unbiased estimate of the corrected system could not be obtained. A natural alternative to a *s*-unbiased estimate is to utilize an estimation procedure based on a loss function to minimize error, as suggested in the concluding remarks of [5]. Years later, Stein [6] developed a statistical estimator based on such an optimality criterion; that is, based on minimizing the *s*-expected sum of squared error. After deriving the required shrinkage factor, this estimator provided good results when utilized in the development of a continuous reliability growth model, known as AMPM-Stein [7]. Simulations conducted by AMSAA [8] indicate that the accuracy in the reliability projections of AMPM-Stein are greater than that of the international standard reliability growth projection model adopted by the International Electrotechnical Commission [9]. To apply the Stein estimator in the proposed discrete setting, we derived the required shrinkage factor, which is discussed & provided below. In many respects, the presented approach serves as a discrete analogue to the continuous reliability growth projection model AMPM-Stein [7].

### 4.1.2. Overview

The methodology of our approach is presented in Section II which includes: 1) a list of model assumptions; 2) a discussion of the data required; 3) a new method for approximating the vector of failure probabilities inherent to a complex, one-shot system; 4) our exact expression for system reliability growth; 5) development of multiple estimation procedures for our model equations; and 6) a graphical method for studying GOF. To highlight model accuracy (e.g., *s*-bias, and *s*-variability), Monte Carlo simulation results are presented in Section III. Concluding remarks are given in Section IV.

## 4.2. Methodology

### 4.2.1. Model Assumptions

1. A trial results in a dichotomous success/failure outcome such that $N_{i,j} \sim Bernoulli\left(p_i\right)$ for each $i = 1,\ldots,k$, and $j = 1,\ldots,T$.
2. The distribution of the number of failures in $T$ trials for each failure mode is binomial. That is, $N_i \sim Binomial\left(T, p_i\right)$ for each $i = 1,\ldots,k$.
3. Initial failure probabilities $p_1,\ldots,p_k$ constitute a realization of a *s*-random sample $P_1,\ldots,P_k$ such that $P_i \sim Beta\left(n,x\right)$ for each $i = 1,\ldots,k$.
4. Corrective actions are delayed until the end of the current test phase, where a test phase is considered to consist of a sequence of $T$ *s*-independent Bernoulli trials.
5. One or more potential failure modes can occur in a given trial, where the occurrence of any one of which causes failure.
6. Failures associated with different failure modes arise *s*-independently of one another on each trial. As a result, the system must be at a stage in development where catastrophic failure modes have been previously discovered & corrected, and are therefore not preventing the occurrence of other failure modes.
7. 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 parameters do not exist.

#### 4.2.2. Data Required

There are two classes of projection models, and each require a unique type of data. The first class of models address the case where all fixes are delayed, as in [5], [7], [10], and the approach presented herein. The second class of projection models, as in [11], and [12], address the case where fixes can either be delayed, or non-delayed. In this case, fixes can be implemented during or following the current test phase; hence, the system configuration need not be constant. The data required for reliability growth projection consists of either: count data (i.e., the number of failures for individual failure modes), FOT data (i.e., the times or trials at which failure modes were first discovered), or a mixture of the two. While we have developed estimation procedures for both classes of projection models, we shall only present the case where all fixes are delayed in the scope of the current paper. This requires $T$, $N_i$, and $d_i$ for $i = 1, \ldots, m$. The number of trials $T$, and the count data $N_i$ for observed failure modes are obtained directly from testing. The $d_i$ can be estimated from test data, or assessed via engineering judgment. For many DoD weapon system development programs, FEF are assessed via expert engineering judgment, and assigned in failure prevention review board meetings. In our experience, the assessed FEF from such forums are those that are typically utilized in reliability growth analyses.

### 4.2.3. Estimation of Failure Probabilities

The well known, widely used MLE of a failure probability is given by

$$\hat{p}_i = \frac{N_i}{T} \tag{12}$$

The problem with this estimator is that, if there are no observed failures for failure mode *j,* then $N_j = 0$. Hence, our corresponding estimate of the failure probability is $\hat{p}_j = 0$, which results in an overly optimistic assessment. Therefore, a finite & positive estimate for each failure mode probability of occurrence is desired, whether observed during testing or not observed during testing. One way to do this is to utilize a shrinkage factor estimator [6] given by

$$\tilde{p}_i \equiv \theta \cdot \hat{p}_i + (1-\theta) \cdot \frac{\sum_{i=1}^{k} \hat{p}_i}{k} \tag{13}$$

where $\theta$ (unknown) is referred to as the shrinkage factor, and $k$ denotes the total potential number of failure modes inherent to the system. The optimal value of $\theta \in (0,1)$ can be chosen to minimize the *s*-expected sum of squared error, but it must be derived consistently with the specific case considered, and r.v. in question. The associated optimality criterion can be mathematically expressed as

$$\frac{d}{d\theta} E\left[ \sum_{i=1}^{k} (\tilde{p}_i - p_i)^2 \right] = 0 \tag{14}$$

To derive $\theta$ uniquely for our application, we have first expressed the mathematical expectation in (14) as a quadratic polynomial with respect to $\theta$ by assuming that the distribution of the number of failures in $T$ trials conditioned on a given failure mode is binomial, which gives

$$E\left[\sum_{i=1}^{k}(\tilde{p}_i - p_i)^2\right] = \theta^2 \cdot \left(\frac{p}{T} - \frac{\sum_{i=1}^{k} p_i^2}{T}\right) + 2\theta(1-\theta)\cdot\left(\frac{p}{k\cdot T} - \frac{\sum_{i=1}^{k} p_i^2}{k\cdot T}\right) + (1-\theta)^2\cdot\left(\frac{p}{k\cdot T} - \frac{\sum_{j=1}^{k} p_j^2}{k\cdot T} + \sum_{i=1}^{k} p_i^2 - \frac{p^2}{k}\right) \tag{15}$$

where $p \equiv \sum_{i=1}^{k} p_i$. Using (15), we have derived the solution to (14), which we conveniently express as

$$\theta = \frac{Var(p_i)}{\left(\frac{E(p_i)\left[1-E(p_i)\right]-Var(p_i)}{T}\right)\left(1-\frac{1}{k}\right)+Var(p_i)} \tag{16}$$

This result is significant for a number of reasons. First, we have expressed the shrinkage factor in terms of quantities that can be easily estimated; namely, the *s*-mean, and *s*-variance of the $p_i$. Second, we have reduced the number of unknowns requiring estimation from $(k+1)$ to only three. The $(k+1)$ unknowns to which we refer include the unknown failure probabilities $p_1,\ldots,p_k$, and the unknown value of $k$. Finally, estimating (or providing appropriate treatment to) these unknowns yields an approximation of the vector of failure probabilities associated with a complex, one-shot system, where each failure probability (observed or unobserved) is finite, and positive.

### 4.2.4. Reliability Growth Projection

Let $obs \equiv \{i : N_i > 0 \text{ for } i\text{=}1,\ldots,k\}$ represent the index set of failure modes observed during testing, and let $obs' \equiv \{j : N_j = 0 \text{ for } j\text{=}1,\ldots,k\}$ denote its compliment. After mitigation to (all or a portion of) failure modes observed during testing, we define the true, but unknown system reliability growth as

$$r(T) \equiv \prod_{i \in obs} \left[1-(1-d_i)\cdot p_i\right] \cdot \prod_{j \in obs'} (1-p_j) \tag{17}$$

where $d_i \in [0,1]$ represents the FEF of failure mode *i*, the true but unknown fraction reduction in initial mode failure probability *i* due to implementation of a unique corrective action. In our model, $(1-d_i)\cdot p_i$ represents the true reduction in failure probability *i* due to correction as originally developed by Corcoran et al. [5]. It will typically be the case that $d_i \in (0,1)$, as $d_i = 0$ models the condition where a given failure mode is not addressed (e.g., an A-mode), and $d_i = 1$ corresponds to complete elimination of the failure mode's probability of occurrence. We would only expect to completely eliminate a failure mode's probability of occurrence when the corrective action consists of the total removal of all components associated with the mode. Notice that our model does not require utilization of the A-mode/B-mode classification scheme proposed in [10], as A-modes need only be distinguished from B-modes via a zero FEF.

The theoretical assessment of (17) is given by

$$\tilde{r}(T) \equiv \prod_{i \in obs} \left[1-(1-d_i)\cdot \tilde{p}_i\right] \cdot \prod_{j \in obs'} (1-\tilde{p}_j) \tag{18}$$

where $\tilde{p}_i$ is expressed via (13). Note that (18) is theoretical because *k* is unknown, the $d_i$ for $i = 1,\ldots,k$ are unknown, and the $p_i$ for $i = 1,\ldots,k$ (upon which the shrinkage factor $\theta$ is based) are unknown. In the following section, we present several approximations to (18), which are derived from our estimation method for the vector of the $p_i$ in combination with classical moment-based, and likelihood-based procedures for the beta parameters. We also derive unique limiting approximations to (18).

### 4.2.5. Estimation Procedures

**4.2.5.1. Parametric Approach**: Assume that the initial mode probabilities of failure $p_1,\ldots,p_k$ constitute a realization of a s-random sample $P_1,\ldots,P_k$ from a beta distribution with the parameterization

$$f(p_i) \equiv \frac{\Gamma(n)}{\Gamma(x)\cdot\Gamma(n-x)} \cdot p_i^{x-1} \cdot (1-p_i)^{n-x-1} \tag{19}$$

for $p_i \in [0,1]$, and 0 otherwise; where *n* represents pseudo trials, *x* represents pseudo failures, and $\Gamma(x) \equiv \int_0^\infty t^{x-1} \cdot e^{-t} dt$ is the Euler gamma function. The above beta assumption not only facilitates convenient estimation of (16), but models mode-to-mode s-variability in the initial failure probabilities of occurrence. The source of such *s*-variability could result from many different factors including, but not limited to, variation in environmental conditions, manufacturing processes, operating procedures, maintenance philosophies, or a combination of the above. As indicated by Ellner & Wald [12], the approach of modeling *s*-variability in complex systems is not new. One of the earlier developments of this concept was presented by Cozzolino [13] to illustrate the effect of population failure rate heterogeneity on the population hazard function. In addition, Littlewood [14], Fakhre-Zakeri & Slud [15], and Miller [16] each, respectively, modeled initial bug rates of occurrence with exponential times to first occurrence with a gamma mixing distribution for software reliability models. In [7], and [12], mode-to-mode *s*-variability was modeled in failure rates of occurrence via a gamma r.v., where the distribution of the number of failures for each mode is assumed Poison. Sarhan et al. [17] modeled component failure probabilities as iid beta r.v. for a multi-component system in the presence of dependent masked system life test data.

Based on our beta assumption with parameterization given by (19), the associated *s*-mean, and *s*-variance are given respectively by

$$E\left(P_i\right)=\frac{x}{n}, \tag{20}$$

and

$$Var\left(P_i\right)=\frac{x\cdot\left(n-x\right)}{n^2\cdot\left(n+1\right)}. \tag{21}$$

Notice that (16) is in terms of only three unknowns; namely, the population *s*-mean of the failure probabilities, the population *s*-variance of the failure probabilities, and *k*. The first two unknowns are approximated by (20), and (21), respectively, which are in terms of the two unknown beta shape parameters. MME, and MLE procedures are utilized to approximate these parameters. The third, final unknown, *k*, is treated in two ways. First, we assume a value of *k*, which can be done in applications where the system is well understood. Second, we allow $k$ to grow without bound to study the limiting behavior of our model equations. This is suitable in cases where the number of failure modes is unknown, and the system is complex. Such a treatment of $k$ was also utilized in the development of two other reliability growth projection models [7], [12].

#### 4.2.5.2. Moment-based Estimation Procedure

Moment estimators for the beta shape parameters are given in [18]. These estimators, per the special case we consider (i.e., where all failure probabilities are estimated via the same number of trials), are given by

$$\breve{n}_k = \frac{\overline{p}_u - m_u^2}{m_u^2 - \frac{\overline{p}_u}{T} - \left(1 - \frac{1}{T}\right) \cdot \overline{p}_u^2}, \tag{22}$$

and

$$\breve{x}_k = \breve{n}_k \cdot \overline{p}_u, \tag{23}$$

where $\overline{p}_u \equiv \frac{\sum_{i=1}^{k} \hat{p}_i}{k}$, and $m_u^2 \equiv \frac{\sum_{i=1}^{k} \hat{p}_i^2}{k}$ are the unweighted first, and second sample moments, respectively. Using the above MME for the beta parameters with (16), our approximation of $\theta$ can be expressed as

$$\breve{\theta}_k = \frac{1}{\left(\frac{\breve{n}_k}{T}\right)\left(1 - \frac{1}{k}\right) + 1}. \tag{24}$$

Using (24), the moment-based shrinkage factor estimate of $p_i$ for finite $k$ is then given by

$$\breve{\tilde{p}}_{k,i} = \breve{\theta}_k \cdot \hat{p}_i + \left(1-\breve{\theta}_k\right)\left(\frac{N}{k \cdot T}\right) \tag{25}$$

where $N \equiv \sum_{i=1}^{k} N_i$ is the total number of failures observed in $T$ trials. Let the total number of observed failure modes be denoted by $m = |obs|$, which implies that there are $|obs'| = k - m$ unobserved failure modes. Then by (18), (24), and (25), the MME-based reliability growth projection for an assumed number of failure modes is given by

$$\breve{\tilde{r}}_k\left(T\right) = \prod_{i \in obs}\left[1-\left(1-d_i^*\right)\cdot \breve{\tilde{p}}_{k,i}\right]\cdot\left[1-\left(1-\breve{\theta}_k\right)\cdot\left(\frac{N}{k \cdot T}\right)\right]^{k-m}, \tag{26}$$

where $d_i^*$ estimates $d_i$.

Because the total potential number of failure modes associated with a complex system is typically large & unknown, it is desirable to study the limiting behavior of (26) as $k \to \infty$. The reliability projection under these conditions simplifies to

$$\breve{\tilde{r}}_\infty\left(T\right) \equiv \lim_{k\to\infty} \breve{\tilde{r}}_k\left(T\right) = \prod_{i \in obs}\left[1-\left(1-d_i^*\right)\cdot \breve{\tilde{p}}_{\infty,i}\right]\cdot \exp\left[-\left(1-\breve{\theta}_\infty\right)\cdot\left(\frac{N}{T}\right)\right] \tag{27}$$

where

$$\breve{\tilde{p}}_{\infty,i} = \breve{\theta}_\infty \cdot \hat{p}_i, \tag{28}$$

$$\breve{\theta}_\infty = \frac{T}{\breve{n}_\infty + T}, \tag{29}$$

and

$$\breve{n}_{\infty} \equiv \lim_{k \to \infty} \breve{n}_k = \frac{\sum_{i=1}^{m} \hat{p}_i - \sum_{i=1}^{m} \hat{p}_i^2}{\sum_{i=1}^{m} \hat{p}_i^2 - \sum_{i=1}^{m} \frac{\hat{p}_i}{T}}, \tag{30}$$

all of which are in terms of failure data that are readily available. From (23), we can see that $\breve{x}_{\infty} \equiv \lim_{k \to \infty} \breve{x}_k = 0$, which implies that the *s*-mean, and *s*-variance of the beta distribution both converge to zero as $k \to \infty$. Hence, the distribution becomes degenerate in the limit.

#### 4.2.5.3. Likelihood-based Estimation Procedure

The method of marginal maximum likelihood [18] provides estimates of the beta parameters n, and x that maximize the beta marginal likelihood function. For an assumed number of total potential failure modes, the estimates denoted by $\widehat{n}_k$, and $\widehat{x}_k$, respectively are obtained by solving the following two likelihood equations simultaneously:

$$\sum_{i=1}^{k} \sum_{j=0}^{N_i - 1} \left( \frac{1}{\widehat{x}_k + j} \right) - \sum_{i=1}^{k} \sum_{j=0}^{T - N_i - 1} \left( \frac{1}{\widehat{n}_k - \widehat{x}_k + j} \right) = 0, \tag{31}$$

and

$$\sum_{i=1}^{k} \sum_{j=0}^{T - N_i - 1} \left( \frac{1}{\widehat{n}_k - \widehat{x}_k + j} \right) - \sum_{i=1}^{k} \sum_{j=0}^{T-1} \left( \frac{1}{\widehat{n}_k + j} \right) = 0, \tag{32}$$

which are defined to be zero if $N_i = 0$. The starting values for the associated numerical routine to obtain such estimates can be chosen to be the unweighted moment estimators given by (22), and (23). Without loss of generality, the finite *k* likelihood-based estimates $\widehat{\theta}_k$, and $\widehat{\widetilde{p}}_{k,i}$ are obtained analogously to that of (24), and (25) with appropriate substitution of the MLE in place of the MME. This provides the likelihood-based estimate of system reliability growth

$$\widehat{\widetilde{r}}_k(T) = \prod_{i \in obs} \left[1 - \left(1 - d_i^*\right) \cdot \widehat{\widetilde{p}}_{k,i}\right] \cdot \left[1 - \left(1 - \widehat{\theta}_k\right) \cdot \left(\frac{N}{k \cdot T}\right)\right]^{k-m} \tag{33}$$

To estimate the limiting behavior of (33), we shall reparameterize (31) & (32), and take limits of these equations as $k \to \infty$. The true but unknown reliability of the system at the beginning of the current test phase is a realization of the product $\prod_{i=1}^{k}\left(1 - P_i\right)$, where $P_i$ is interpreted as a *s*-independent beta r.v. The mathematical expectation of this quantity with respect to the $P_i$ for $i = 1, \ldots, k$ is $\widehat{R}_k = \left(1 - \frac{\widehat{x}_k}{\widehat{n}_k}\right)^k$, which yields the useful parameterization $\widehat{x}_k = \widehat{n}_k \cdot \left(1 - \widehat{R}_k^{1/k}\right)$, where $\widehat{R}_k$ denotes an MLE of the unconditional *s*-expected initial system reliability. Notice that $\widehat{x}_k \to 0$ as $k \to \infty$. This does not come to as much of a surprise because we would expect the likelihood-based estimate of the beta parameter *x* to exhibit the same behavior as that of the moment based estimate, which also converges to zero as *k* grows without bound. By substituting this parameterization into (31), and taking the limit, we derive the following MLE-based approximation for the *s*-expected initial reliability of a complex one-shot system for $k \to \infty$:

$$\widehat{R}_\infty = \exp\left[\frac{-m}{\sum_{j=0}^{T-1}\left(\frac{\widehat{n}_\infty}{\widehat{n}_\infty + j}\right)}\right] \tag{34}$$

where $\widehat{n}_{\infty}$ denotes the limit of the MLE for the beta parameter *n* (i.e., pseudo trials). This result is significant for a number of reasons. First, we derived a new estimate for the *s*-expected initial reliability of a one-shot system, which is a basic quantity of interest to program managers, and reliability practitioners. This quantity also serves as an estimate of the current demonstrated reliability of a one-shot system. This offers an alternative to the typical reliability point estimate calculated as the ratio of the number of successful trials to the total number of trials. Second, we expressed this quantity in terms of only one unknown, which has reduced the estimation procedure to solving one equation for $\widehat{n}_{\infty}$. To derive this equation, we proceed in a similar fashion as above. Let $x_k = \frac{n_k \cdot k \cdot \left(1 - R_k^{1/k}\right)}{k} = \frac{n_k \cdot \gamma_k}{k}$, where $\gamma_k = k \cdot \left(1 - R_k^{1/k}\right)$. Note that $\gamma_k$ is finite, and positive as $k \to \infty$. By substituting this parameterization into (32), and taking the limit, the estimate $\widehat{n}_{\infty}$ for the beta parameter *n* is found such that

$$\sum_{i=1}^{m} \sum_{j=0}^{T-N_i-1} \left( \frac{1}{\widehat{n}_{\infty} + j} \right) = m \cdot \sum_{j=0}^{T-1} \left( \frac{1}{\left(\widehat{n}_{\infty} + j\right)} - \frac{1}{\left(\widehat{n}_{\infty} + j\right)^2 \cdot \sum_{i=0}^{T-1} \left( \frac{1}{\widehat{n}_{\infty} + i} \right)} \right). \quad (35)$$

Hence, the resulting limiting behavior of the likelihood-based estimate for one-shot system reliability growth is given by

$$\widehat{\widetilde{r}}_{\infty}(T) \equiv \lim_{k \to \infty} \widehat{\widetilde{r}}_{k}(T) = \prod_{i \in obs} \left[ 1 - \left(1 - d_i^*\right) \cdot \widehat{\widetilde{p}}_{\infty,i} \right] \cdot \exp\left[ -\left(1 - \widehat{\theta}_{\infty}\right) \cdot \left( \frac{N}{T} \right) \right] \quad (36)$$

where

$$\widehat{\widetilde{p}}_{\infty,i} = \widehat{\theta}_{\infty} \cdot \hat{p}_i, \quad (37)$$

$$\hat{\theta}_{\infty} = \frac{T}{\hat{n}_{\infty} + T} \tag{38}$$

and $\hat{n}_{\infty}$ is found as the solution of (35).

#### 4.2.6. Goodness-of-Fit

While we have developed formal statistical GOF procedures for this model, they are not presented in the scope of the current paper. However, the GOF of the model can be graphically studied by plotting the cumulative number of observed failure modes versus trials against our estimate of the cumulative *s*-expected number of observed failure modes on trial *t* given by

$$\hat{\mu}(t) = \left( \frac{m}{\sum_{j=0}^{T-1} \left( \frac{1}{\hat{n}_{\infty} + j} \right)} \right) \cdot \left( \frac{\Gamma'(\hat{n}_{\infty} + t)}{\Gamma(\hat{n}_{\infty} + t)} - \frac{\Gamma'(\hat{n}_{\infty})}{\Gamma(\hat{n}_{\infty})} \right) \tag{39}$$

where $\hat{n}_{\infty}$ is found as the solution to (35), and $\Gamma'(x)/\Gamma(x)$ is the digamma function. Our formal statistical GOF procedures for this model will be presented in a forthcoming paper.

### 4.3. Monte Carlo Simulation

#### 4.3.1. Overview

# - CHAPTER 4 -

In previous sections, we have introduced a new model that will be helpful in estimating the demonstrated reliability, and reliability growth of one-shot systems. In light of this new model, a natural concern in its application is the accuracy associated with the resulting reliability estimates. To study model accuracy, we have developed a Monte Carlo simulation which consists of the following steps:

1. Specification of simulation inputs such as the total potential number of failure modes, and trials.
2. *S*-random generation of failure probabilities via a beta r.v.
3. *S*-random generation of failure histories via a Bernoulli r.v.
4. *S*-random generation of fix effectiveness factors via a beta r.v.
5. Estimation of the model parameters, and equations presented above.
6. Error estimation between the true, and estimated reliability growth.

Steps 1 through 6 can be viewed as simulating data analogous to that captured during a single developmental test consisting of $T$ trials for a one-shot system comprised of $k$ failure modes. These steps are replicated, which corresponds to simulating a sequence of developmental tests. Simulation inputs remain constant during each replication of the simulation. Failure probabilities, and fix effectiveness factors, however, are stochastically generated anew during each replication. After the simulation is replicated, all failure data, parameter estimates, reliability projections, and error terms are saved, and analyzed. In the next section, we present simulation results based on a given set of inputs. Simulation output consists of summary statistics, and associated relative error probability densities.

### 4.3.2. Simulation Results

#### 4.3.2.1. Summary

Via heuristics, stable simulation results are obtained at 100 replications of the simulation. The presented results are based on 300 replications with $T = 350$ trials, $k = 50$ failure modes, $\mu = 2 \cdot 10^{-3}$ for the population *s*-mean of the failure probabilities, and $\sigma^2 = 2 \cdot 10^{-4}$ for the population *s*-variance of the failure probabilities. The values of these inputs greatly reduce the volume of failures, and failure modes observed during simulation, as a conservative scenario with respect to the volume of failure data available for estimation purposes is desired. For example, only 4 of 50 failure modes were observed on average in the simulated developmental tests. In addition, only a total of 39 failures were observed on average. This is indicative of the high initial reliability of the system, as specified via the inputs above. We wish to emphasize two points. First, it is important not to confuse the difference between the number of replications, and the number of trials, *T*. Clearly, as $T \to \infty$, all failure modes will eventually be observed. However, we are simulating 350 trials per replication of a highly reliable system, and therefore we only observe about 4 of the 50 failure modes consistently on average per replication (i.e., each replication simulates 350 trials). The simulation results are stable in that a small volume of failure data are available for estimation purposes per replication, and there is not much *s*-variability in the reliability growth estimates after 100 replications. Second, a large number of trials does not imply a large volume of failure data. For example, a large number of trials is relative to the initial reliability of the system. In the presented case, 350 trials did not yield a large volume of failure data, as the true unconditional *s*-expected initial system reliability was 0.9047. The arithmetic average (over all replications) of our corresponding estimate given by (34) was 0.9029. Table 1 shows arithmetic averages of the true, and estimated reliability projections based on our approach.

| THEORETICAL | | ESTIMATED | | | |
|---|---|---|---|---|---|
| True | Stein | MME K | MME $\infty$ | MLE K | MLE $\infty$ |
| 0.9763 | 0.9740 | 0.9738 | 0.9786 | 0.9756 | 0.9784 |

**Table 1. Reliability Projections.**

The column titled True is computed via the arithmetic average of (17) over all replications. Similarly, the second column titled Stein is calculated by the arithmetic average of (18) over all replications. Both of these quantities are theoretical, as they are in terms of the true, but unknown $p_1,\ldots,p_k$, and *k*. The remaining four columns in Table 1 are estimates of the true reliability growth based on the arithmetic averages of (26), (27), (33), and (36), respectively, over all replications. The true value of $k=50$ was utilized in (26), and (33), which are shown in the third, and fifth columns, respectively. The sensitivity of not knowing *k* is given by (27), and (36), which are shown in the fourth, and sixth columns, respectively.

By addressing 4 of the 50 failure modes on average (over all replications) with a *s*-mean FEF of 0.80, the system reliability was improved from 0.9047 to 0.9763. By inspection of Table 1, the reliability projections appear quite accurate. There is, however, an element of uncertainty in studying aggregate results, as deviations in model accuracy do occur from one replication to the next. In some cases, reliability projections are conservative, whereas others are optimistic. By computing the arithmetic averages of the projections (over all replications), a portion of the error associated with the conservative estimates is canceled with that of the optimistic, thereby muting deviations in projection error that would otherwise be encountered via a single application of the model in one test phase. To address these concerns, the relative error terms obtained in each replication of the simulation are computed, and analyzed. The error analyses associated with the moment-based, and likelihood-based reliability growth estimates are presented in the following two sub-sub-sections, respectively.

#### 4.3.2.2. Accuracy of Moment-based Projections

Figure 13 displays relative error plots for the moment-based reliability growth projections using a finite, and infinite number of modes, respectively. Using (17), (26) and (27), the relative error for these projections is given respectively by

$$\breve{E}_{k,r} \equiv \frac{r(T)-\breve{\tilde{r}}_k(T)}{r(T)}, \tag{40}$$

and

$$\breve{E}_{\infty,r} \equiv \lim_{k\to\infty} \breve{E}_{k,r} = \frac{r(T)-\breve{\tilde{r}}_\infty(T)}{r(T)}. \tag{41}$$

Figure 13 displays the histograms for the relative error terms obtained from the simulation. MLE is utilized to approximate the parameters of a $s$-normal distribution, which is shown to accurately portray the probability densities of the relative error. The error densities for both the finite, and infinite $k$ reliability growth projections are similar. All error terms are within $\pm 2.5\%$ of the true reliability. Both projections possess *s*-bias with the finite *k* approach providing a slight underestimate, and the infinite *k* approach providing a slight overestimate.

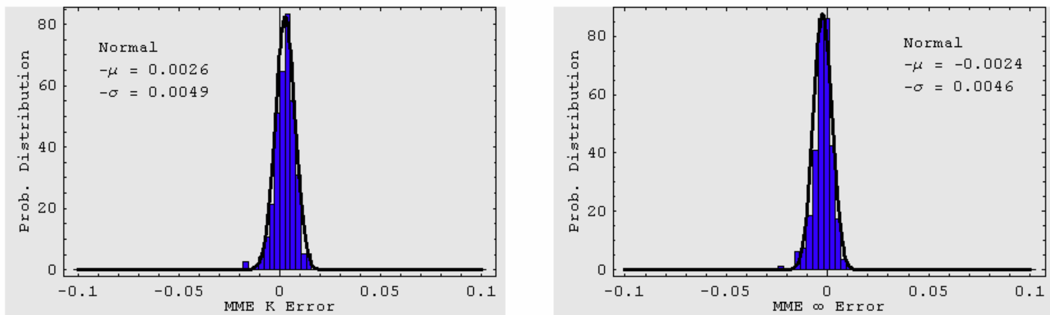


**Figure. 13. Relative Error of Moment-based Projections.**

Based on the estimated *s*-normal distribution for the finite *k* moment-based reliability growth projection $Pr\left\{\breve{E}_{k,r} < |x|\right\} = 0.90 \quad \Rightarrow |x| = 0.0091$. In other words, the projection error in (26) is within $\pm 0.0091$, 90% of the time for the simulated conditions specified above. Likewise, error in the infinite *k* moment-based reliability growth projection (27) is within $\pm 0.0085$, 90% of the time.

#### 4.3.2.3. Accuracy of Likelihood-based Projections

Using (33), and (36), the relative error in the likelihood-based projections are obtained analogously to that shown in the previous section. Without loss of generality, the error results for these projections are very similar to that of the moment-based projections. The only notable difference is that the accuracy is slightly greater using an MLE procedure. Overall, the projection error in (33), and (36) is less than $\pm 0.0076$, and $\pm 0.0081$, respectively, 90% of the time.

### 4.3.3. General Observations

The results shown in the previous sections highlight model accuracy for one set of simulation inputs. Clearly, there are infinitely many combinations of inputs under which model accuracy could be studied. Several different combinations of inputs in conjunction with their simulation output have been analyzed in an effort to generalize the conditions for which model accuracy is high (e.g., $\breve{E}_{k,r} \leq 0.10$). Based on these analyses, we observed that model accuracy is not simply a function of using (for estimation purposes) a large volume of failure data, or observing a proportional majority of failure modes in the system. Rather, model accuracy is found to be a function of obtaining good estimates for the dominant failure modes of the system. In the presented simulation results, only 4 of the 50 failure modes were observed on average, but these failure modes represented about 90% of the system unreliability. In addition, 10 failures were observed on average for each of the modes, which provided good estimates for their associated probabilities of occurrence.

Finally, with respect to the accuracy of the limiting behavior of the model, empirical evidence obtained via simulation suggests that, if *k* is sufficiently greater than *m*, the projections given by (27), and (36) will be insensitive to the value of *k*. In our experience, we have found that the condition $k \geq 5 \cdot m$ is a good rule-of-thumb for the convergence of these estimators for complex systems.

## 4.4. Air-to-Ground Missile Application

Table 2 below shows failure, and fix effectiveness data obtained from an unspecified air-to-ground missile program[37]. The second column shows the number of failures associated with $m = 7$ failure modes discovered during in $T = 27$ flight tests of the system. In discussions with design and reliability engineers working on the program, it was determined that each of these failure modes operate and fail independently of one another. The fix effectiveness data shown in column 3 of Table 2 are FEF that were assigned by a Failure Prevention and Review Board after adopting the proposed engineering design changes (e.g., corrective actions) to mitigate the occurrence of the failure modes.

| | **Failures** $N_i$ | **FEF** $d_i$ |
|---|---|---|
| **1** | 1 | 0.95 |
| **2** | 1 | 0.70 |
| **3** | 1 | 0.90 |
| **4** | 1 | 0.90 |
| **5** | 4 | 0.95 |
| **6** | 2 | 0.70 |
| **7** | 1 | 0.80 |

**Table 2. Failure and FEF Data.**

These failure data were used to calculate the estimates of the beta shape parameters $n$ and $x$, which are shown in the first two rows of Table 3, respectively. The third row shows the corresponding estimate of the shrinkage factor (16). The finite $k$ MME and MLE estimates where calculated with an assumed $k = 50$ total potential number of failure modes. The finite $k$ moment estimators are computed by Equations (22) and (23). The limiting approximation of the moment estimator for the beta parameter

[37] The system name and failure mode information cannot be provided due to propriety reasons.

$n$ is calculated by (30). The finite $k$ ML estimates are obtained as the simultaneous numerical solutions to Equations (31) and (32). The limiting ML approximation of the beta parameter $n$ is found as the solution to (35). Recall that the limiting approximation of the beta parameter $x$ (i.e., pseudo failures) converges to zero as $k \to \infty$.

| | MME K | MME ∞ | MLE K | MLE ∞ |
|---|---|---|---|---|
| $n$ | 23.310 | 19.430 | 23.960 | 19.310 |
| $x$ | 0.190 | 0 | 0.195 | 0 |
| $\theta$ | 0.5417 | 0.5815 | 0.5349 | 0.5830 |

**Table 3. Parameter Estimates.**

Using the parameter estimates given in Table 3, the projected reliabilities of the system were calculated. These estimates are shown in Table 4.

| | MME K | MME ∞ | MLE K | MLE ∞ |
|---|---|---|---|---|
| $R_P$ | 0.8218 | 0.8155 | 0.8201 | 0.8159 |

**Table 4. Demonstrated and Projected Reliability.**

These estimates were calculated from Equations (26), (27), (33), and (36), respectively. The initial reliability, computed by Equation (34) is 0.6654. By correcting the 7 failure modes with fix effectiveness specified in Table 2, system reliability is projected to be improved from about 0.67 to 0.82. Keep in mind that the actual reliability improvement depends upon the actual level fix effectiveness achieved. Thus, if the assignment of FEF were overly optimistic, these reliability assessments will also be overly optimistic. Likewise if they are overly pessimistic. The projected reliability based on the Crow-Extended model [157] is 0.85. In this example, 4 of the 7 observed failure modes were corrected during the test phase. Recall that this model assumes all the corrective actions are delayed, and that the Crow-Extended model accounts for delayed and/or non-delayed fixes. To the extent that the Crow-Extended projection is accurate, this example highlights the effect of the violation of this model's assumption regarding delayed fixes. Thus, the violation of this assumption may contribute to the difference between the reliability projections of the two models.

## 4.5. Concluding Remarks

In this paper, we have presented a new model for estimating the reliability growth of complex, one-shot systems. Our model offers an alternative to the popular competing risks approach. It is suitable for application when one or more failure modes can be discovered in a single trial, and when catastrophic failures modes have been previously discovered, and corrected. Equation (17) is our logically derived model. Our theoretical estimate of (17) is given by (18). Our practical estimates of (18) are given by (26), (27), (33), and (36).

We have developed a new method for approximating the vector of failure probabilities associated with a complex one-shot system, which is based on our derived shrinkage factor given by (16). The benefit of this procedure is that it not only reduces error, but reduces the number of unknowns requiring estimation from $k+1$ to only three. Also, estimates of mode failure probabilities, whether observed or unobserved during testing, will be finite, and positive.

We have derived unique limits of our model equations, which have yielded interesting simplifications. The limiting approximations of our model equations include (27)-(30), and (34)-(38). In particular, we derived a mathematically-convenient functional form for the *s*-expected initial system reliability of a one-shot system (34). This quantity serves as an estimate of the current demonstrated reliability of a one-shot system, and offers an alternative to the typical reliability point estimate calculated as the ratio of the number of successful trials to the total number of trials.

Finally, we have presented Monte Carlo simulation results to highlight model accuracy with respect to resulting estimates of reliability growth. While all error terms were within $\pm 2.5\%$ of their reliability estimates, the approximated s-normal distributions above indicate that the projection error is within $\pm 0.9\%$ (i.e., $\pm 0.0091$), with a probability of 0.90. While model accuracy is generally found to be good, we recommend tailored Monte Carlo simulation studies be performed to highlight model accuracy for specific systems of interest.

**Acknowledgment**

We give special thanks to both Dr. Paul M. Ellner (AMSAA), and Dr. Michael J. Cushing (AEC); Aberdeen Proving Ground, Maryland. Dr. Ellner, and Dr. Cushing provided critical comments, and assessment of this work which greatly improved the quality of the manuscript. We also extended our gratitude to Mr. Harry Calkins of Wolfram Research, Inc., Champaign, Illinois. Mr. Calkins developed a variant of *Mathematica's* "Histogram" function, which was utilized throughout this study to examine projection error.

## References

[1] *Reliability Growth Management*, Military Handbook 189, U.S. Department of Defense, Washington, DC, 1981.

[2] P. M. Ellner, W. J. Broemm, and W. J. Woodworth, “Reliability Growth Guide,” AMSAA, Aberdeen Proving Ground, MD, Tech. Rep. No. TR-652, 2000.

[3] D. E. Mortin and P.M. Ellner, “Reliability Growth for Military Vehicles – Emerging Methodology and Remaining Challenges,” *Journal of Commercial Vehicles*, vol. 114, pp.28-34, 2005.

[4] A. Fries and A. Sen, “A Survey of Discrete Reliability Growth Models,” *IEEE Trans. Reliab.*, vol. R-45, no. 4, pp. 582-604, 1996.

[5] W. J. Corcoran, H. Weingarten, and P. W. Zehna, “Estimating Reliability after Corrective Action,” *Management Science*, vol. 10, no. 4, pp. 786-798, 1964.

[6] C. M. Stein, “Estimation of the Mean of a Multivariate Normal Distribution,” *The Annals of Statistics*, vol. 9, pp. 1135-1151, 1981.

[7] P. M. Ellner and J. B. Hall, “AMSAA Maturity Projection Model based on Stein Estimation,” *Proc. RAMS*, 2005, pp. 271-277.

[8] P. M. Ellner and J. B. Hall, “AMPM-Stein and Monte Carlo Simulation Results,” AMSAA, Aberdeen Proving Ground, MD, Tech. Rep. No. TR-751, 2004.

[9] *IEC Standards for Reliability Growth – Statistical Test and Estimation Methods*, IEC Standard 61164 II, 2004.

[10] L. H. Crow, “An Improved Methodology for Reliability Growth Projection,” AMSAA, Aberdeen Proving Ground, MD, Tech. Rep. No. TR-357, 1982.

[11] L. H. Crow, “An Extended Reliability Growth Model for Managing and Assessing Corrective Actions,” *Proc. RAMS*, 2004, pp. 73-80.

[12] P. M. Ellner and L. C. Wald, “AMSAA Maturity Projection Model,” *Proc. RAMS*, 1995, pp. 174-181.

[13] J. Cozzolino, “Probabilistic Models of Decreasing Failure Rate Processes,” *Naval Research Logistics Quarterly*, vol. 15, pp. 361-374, 1968.

[14] B. Littlewood, “Stochastic Reliability-Growth: A Model for Fault-Removal in Computer-Programs and Hardware-Designs,” *IEEE Trans. Reliab.*, vol. R-30, no. 4, 1981.

[15] I. Fakhre-Zakeri and E. Slud, “Mixture Models for Reliability of Software with Imperfect Debugging: Identifiability of Parameters,” *IEEE Trans. Reliab.*, vol. R-44, no. 1, pp. 104-113, 1995.

[16] D. Miller, “Exponential Order Statistic Models of Software Reliability Growth,” *IEEE Trans. Softw. Eng*., vol. SE-12, no. 1, 1986.

[17] A. M. Sarhan, F. M. Guess, and J. S. Usher, “Estimators for Reliability Measures in Geometric Distribution Model Using Dependent Masked System Life Test Data,” *IEEE Trans. Reliab*., vol. R-56, no. 2, 2007.

[18] H. F. Martz and R. A. Waller, “Selecting Beta Prior Distributions based on Observed Binomial Sampling Data,” in *Bayesian Reliability Analysis*, 2nd ed., Malabar, FL: Krieger, 1991, pp. 312-318.

# 5. MANAGEMENT METRICS FOR DELAYED FIXES[38]

## Abstract

In this paper, we introduce a new reliability growth methodology for one-shot systems that is applicable to the case where all corrective actions are implemented at the end of the current test phase. The methodology consists of four model equations for assessing: expected reliability, the expected number of failure modes observed in testing, the expected probability of discovering new failure modes, and the expected probability of observing a repeat failure mode. These model equations provide an analytical framework for which reliability practitioners can estimate reliability improvement, address goodness-of-fit concerns, quantify programmatic risk, and assess reliability maturity of one-shot systems. A numerical example is given to illustrate the value and utility of the presented approach. This methodology is useful to program managers and reliability practitioners interested in applying the techniques above in their reliability growth program.

**Keywords:** Growth Potential, One-Shot Systems, Projection, Reliability Growth.

## Acronyms[39]

**AEC** – Army Evaluation Center

[38] Chapter 5 is accepted for publication in *Reliability Engineering & System Safety*. Citation information has not yet been assigned.

[39] The singular and plural of an acronym are always spelled the same.

**AMSAA** – Army Materiel Systems Analysis Activity

**FEF** – Fix Effectiveness Factor(s)

**GOF** – Goodness-Of-Fit

**IDA** – Institute for Defense Analysis

**MLE** – Maximum Likelihood Estimation/Estimate(s)

**MME** – Method of Moments Estimation/Estimate(s)

**TAFT** – Test, Analyze, Fix and Test

## Definitions

1. **FEF** – fraction reduction in an initial failure mode probability due to implementation of a unique corrective action.
2. **Failure mode** – the root-cause associated with the loss of a required function or component whose probability of occurrence is reduced by a specified FEF, if addressed by corrective action. Note that it may be the case that some failure modes are not observed during testing, or may not be corrected if they are observed (e.g., some failures may not be economically justifiable to correct).
3. **Unobserved failure mode** – a failure mode which exhibits zero failures during testing.
4. **Observed failure mode** – a failure mode which exhibits at least one failure during testing.
5. **Repeat failure mode** – a failure mode which exhibits at least two failures during testing.

# - CHAPTER 5 -

## Notation

$k$ - total number of potential failure modes.

$m$ - total number of observed failure modes.

$N_{i,j}$ - number of failures for failure mode $i$ in trial $j$ – zero or unity.

$N_i$ - total number of failures for failure mode $i$ in $T$ trials.

$T$ - total number of trials.

$p_i$ - true but unknown probability of failure for failure mode $i$.

$\vec{p}$ - the vector of $p_i$ for $i = 1, \ldots, k$.

$\hat{p}_i$ - MLE of $p_i$.

$\tilde{p}_{k,i}$ - theoretical shrinkage factor estimator for $p_i$.

$\theta$ - true but unknown shrinkage factor.

$n$ - beta shape parameter representing the pseudo number of trials.

$x$ - beta shape parameter representing the pseudo number of failures.

$d_i$ - true but unknown FEF for failure mode $i$.

$R_k\left(t \mid \vec{p}\right)$ - conditional expected reliability on trial $t$.

$\mu_k\left(t \mid \vec{p}\right)$ - conditional expected number of observed failure modes (i.e., unique failure modes) through trial $t$.

$h_k\left(t \mid \vec{p}\right)$ - conditional expected probability of discovering a new failure mode on trial $t$.

$\phi_k(t \mid \bar{p})$ - conditional expected probability of failure on trial $t$ due to a repeat failure mode (this quantity is independent of failure mode mitigation).

## 5.1. Introduction

### 5.1.1. Background

Many reliability growth models have been developed over the last several decades to assist reliability practitioners with the formidable task of estimating and tracking reliability improvements of a system throughout the development process. Reliability growth planning, tracking and projection are the three major areas of the field with the AMSAA Reliability Growth Methodology Guide given by Ellner et al. in [1] being among the most comprehensive works on the subject. Military Handbook 189 [2], although outdated, is also a good reference on reliability growth and is currently being updated by the U.S. Army through the AMSAA. While methodologies are available for both continuous and one-shot systems (e.g., see [3] for a survey of discrete reliability growth models), the area of discrete reliability growth projection is underdeveloped, as noted in [4]. In this paper, we present a new methodology which serves as a framework for analyzing reliability of one-shot systems undergoing development. This methodology is a discrete analogue to [5] and incorporates many of the concepts advanced in [6] and [7]. The methodology consists of four model equations designed to estimate reliability improvement of one-shot systems, address GOF concerns, and provide measures of programmatic risk and system maturity. The overall intention of these model equations are to give reliability practitioners the means to gauge the progress of the development effort of one-shot systems through the TAFT process.

Multiple statistical estimation procedures are presented in our earlier paper [8], which are utilized to approximate the model equations developed in Section II. For this reason, these estimation procedures are only briefly outlined herein, and the reader will need to reference [8] for associated details. Monte Carlo simulation results are also presented in [8] for a reliability growth model designed to estimate the true reliability of a one-shot system, as opposed the s-expected reliability (developed in Section II below). Estimates of true reliability and s-expected reliability yield similar results, but are different reliabilities and their associated expressions are constructed much differently from one another. Expressions for expected reliability are constructed by using indicator variables and mathematical s-expectations thereof, whereas the true reliability [8] is based on approximating an exact expression directly (without indicator variables and their mathematical s-expectations). Both methods are useful for estimating reliability and both methods, for the case considered (i.e., where all fixes are assumed delayed), utilize the same estimation procedures.

### 5.1.2. Overview

The methodology of our approach is presented in Section II, which contains a list of model assumptions, an abbreviated background on estimation of model parameters, and derivations of our four model equations. In Section II, we restrict our discussions to the technical aspects of the developed expressions. The value, utility, and usefulness of the methodology are discussed in greater detail and illustrated in Section III via numerical example. Concluding remarks are given in Section IV.

## 5.2. Methodology

### 5.2.1. Assumptions

1. A trial results in a dichotomous success/failure outcome such that $N_{i,j} \sim Bernoulli(p_i)$ for each failure mode $i = 1,\ldots,k$ and trial $j = 1,\ldots,T$.
2. The distribution of the number of failures in $T$ trials for each failure mode is binomial. That is, $N_i \sim Binomial(T, p_i)$ for each $i = 1,\ldots,k$.
3. Initial failure mode probabilities of occurrence $p_1,\ldots,p_k$ constitute a realization of a s-random sample $P_1,\ldots,P_k$ such that $P_i \sim Beta(n,x)$ for each $i = 1,\ldots,k$.
4. Corrective actions are delayed until the end of the current test phase, where a test phase is considered to consist of a sequence of $T$ s-independent Bernoulli trials.
5. Potential failure modes occur s-independently of one another and their occurrence is considered to constitute a failure. As a result, the system must be at a stage in development where catastrophic failure modes have been previously discovered and corrected. If catastrophic failure modes have not been previously corrected, the occurrence of other potential failure modes

may be suppressed and this assumption will therefore be violated. Hence, the system must reach a sufficient level of technological maturity in order to apply this methodology appropriately. Complications with this assumption may be avoided if the methodology is applied per phase of a mission, where the associated class of failure modes in a given phase occur s-independently.

6. 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 parameters do not exist.

### 5.2.2. Estimation

In a previous paper [8], we introduced a new method for estimating the vector of failure mode initial probabilities of occurrence associated with a one-shot system. The basis of our method is to avoid inaccuracies that arise in application of the well-known, widely-used, MLE of a failure probability given by

$$\hat{p}_i = \frac{N_i}{T} \tag{1}$$

which is zero when no failures are observed on failure mode $i$ (i.e., $\hat{p}_i = 0$ when $N_i = 0$). Our approach in avoiding this problem is to utilize the following shrinkage factor estimator [9]

$$\tilde{p}_{k,i} \equiv \theta \cdot \hat{p}_i + \left(1-\theta\right) \cdot \frac{\sum_{i=1}^{k} \hat{p}_i}{k} \tag{2}$$

where we derived the required shrinkage factor $\theta = \theta_k$ for our specific case and random variable in question. We have chosen this quantity such that

$$\frac{d}{d\theta} E\left[ \sum_{i=1}^{k} (\tilde{p}_{k,i} - p_i)^2 \right] = 0 \tag{3}$$

which minimizes the expected sum of squared-error. The resulting solution of (3) yields the optimal value of $\theta = \theta_k \in (0,1)$ that we conveniently express as

$$\theta_k = \frac{Var(p_i)}{\left( \frac{E(p_i)[1 - E(p_i)] - Var(p_i)}{T} \right)\left(1 - \frac{1}{k}\right) + Var(p_i)} \approx \frac{1}{\left(\frac{n}{T}\right)\left(1 - \frac{1}{k}\right) + 1} \tag{4}$$

This result is significant for a number of reasons but mainly because it is based on an optimality criterion to minimize error, and because (4) reduces the number of unknowns requiring estimation from $(k+1)$ to only three. The $(k+1)$ unknowns to which we refer include the unknown failure probabilities $p_1, \ldots, p_k$ and the unknown value of *k*. The three remaining unknowns include the mean and variance of the $p_i$ and *k*. By assuming that the initial failure mode probabilities of occurrence $p_1, \ldots, p_k$ constitute a realization of a s-random sample $P_1, \ldots, P_k$ from a beta distribution, we have estimated the mean and variance of the $p_i$ in (4) with the mean and variance of the beta distribution, which are only in terms of the two beta shape parameters *n*, pseudo trials, and *x*, pseudo failures. The right-hand side of (4) shows the shrinkage factor after substitution of these quantities. This facilitated estimation of our first two unknowns via the common MME and MLE procedures for a beta r.v. given in [10]. We treated the third unknown, *k*, in two ways: (1) we assumed a finite value for *k* which can be done in applications where the system is well understood, and (2) we took the limit of our model equations as $k \to \infty$, which we have found is even suitable in cases where the system is not complex, as illustrated by the numerical example below. Overall, we provided four estimation procedures in [8] which can be utilized to approximate the exact expressions herein (our notation is consistent with that of our earlier paper). Please refer to [8] for further details on statistical estimation of model parameters, as no further discussion on estimation is provided in this paper.

### 5.2.3. Reliability Growth

Let $I_i(t)$ denote the indicator function such that

$$I_i(t) \equiv \begin{cases} 1 & \text{if failure mode } i \text{ is observed on or before trial } t \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Using (5) our logically derived model of the true reliability on trial *t* (following failure mode mitigation) is given by

$$r(t \mid \bar{p}) \equiv \prod_{i=1}^{k} \left[1 - \left[1 - I_i(t-1) \cdot d_i\right] \cdot p_i\right] \tag{6}$$

When a given failure mode has been observed prior to trial *t*, the indicator function in (6) equates to unity and the initial failure probability $p_i$ is reduced by a fractional amount $(1-d_i)$. Observed failure modes that are not addressed via corrective action are simply assigned a zero FEF (i.e., $d_i = 0$). When a failure mode is not observed prior to trial *t*, the indicator function equates to zero and the original value of the failure probability $p_i$ is preserved. Assuming that trials are statistically independent and that the $p_i$ for $i=1,...,k$ are constant, the resulting mathematical expectation of (5) is

$$E\left[I_i(t)\right]=1-(1-p_i)^t \tag{7}$$

From (6) and (7), the expected reliability of the system conditioned on the vector of unknown failure probabilities $\vec{p}\equiv(p_1,p_2,\ldots,p_k)$ becomes

$$R_k(t\mid\vec{p})\equiv E\left[r(t\mid\vec{p})\right]=\prod_{i=1}^{k}\left[1-\left(1-\left[1-(1-p_i)^{t-1}\right]\cdot d_i\right)\cdot p_i\right] \tag{8}$$

Notice the initial condition of (8) equates to the initial system reliability $R_k(t=1\mid\vec{p})=\prod_{i=1}^{k}(1-p_i)$, as expected. Also notice that our model is independent of the *A-mode / B-mode*[40] classification scheme proposed in [11], as A-modes need only be distinguished from B-modes via a zero FEF (i.e., $d_i=0$ if failure mode $i$ is not observed, or is not corrected).

It is also desirable to study the limiting behavior of (8) as $k\to\infty$, since the total potential number of failure modes associated with a complex one-shot system is typically large, and since *k* is unknown. Let $obs\equiv\{i:N_i>0 \text{ for } i=1,\ldots,k\}$ represent the index set of failure modes observed during testing and $obs'\equiv\{j:N_j=0 \text{ for } j=1,\ldots,k\}$ denote its complement. Also let the total number of observed failure modes be denoted by $m=|obs|$, which implies that there are $|obs'|=k-m$ unobserved failure modes. We have found a theoretical limit of (8) by

[40] An ***A-mode*** is a failure mode that will not be addressed via corrective action, whereas a ***B-mode*** is a failure mode that will be addressed via corrective action, if observed.

expressing it in terms of the observed and unobserved failure modes and by approximating $p_i$ with an MME or MLE of $\tilde{p}_{k,i}$ for $i = 1,\ldots,k$ which yields

$$\begin{aligned}\tilde{R}_k\left(t \mid \bar{p}\right) &\equiv \prod_{i=1}^{k}\left[1-\left(1-\left[1-\left(1-\tilde{p}_{k,i}\right)^{t-1}\right]\cdot d_i\right)\cdot \tilde{p}_{k,i}\right] \\ &= \prod_{i\in obs}\left[1-\left(1-\left[1-\left(1-\tilde{p}_{k,i}\right)^{t-1}\right]\cdot d_i\right)\cdot \tilde{p}_{k,i}\right]\cdot \prod_{i\in obs'}\left[1-\left(1-\theta_k\right)\cdot \frac{\sum_{j=1}^{m}\hat{p}_i}{k}\right] \\ &= \prod_{i\in obs}\left[1-\left(1-\left[1-\left(1-\tilde{p}_{k,i}\right)^{t-1}\right]\cdot d_i\right)\cdot \tilde{p}_{k,i}\right]\cdot \left[1-\left(1-\theta_k\right)\cdot \frac{\sum_{j=1}^{m}\hat{p}_i}{k}\right]^{k-m}\end{aligned} \quad (9)$$

where $\tilde{p}_{k,i}$, given by (2), and $\theta_k$, given by (4), are approximated by MME and MLE formulas given in [8]. Note that (9) is the expected system reliability on trial $t$ if the failure modes surfaced prior to $t$ are mitigated with an effectiveness proportional to their associated FEF. Using (9), the limiting approximation can be expressed as

$$\begin{aligned}\tilde{R}_\infty\left(t \mid \bar{p}\right) \equiv \lim_{k\to\infty}\tilde{R}_k\left(t \mid \bar{p}\right) = \prod_{i\in obs}\left[1-\left(1-\left[1-\left(1-\tilde{p}_{\infty,i}\right)^{t-1}\right]\cdot d_i\right)\cdot \tilde{p}_{\infty,i}\right]\times \\ \times\exp\left[-\left(1-\theta_\infty\right)\cdot \sum_{j=1}^{m}\hat{p}_i\right]\end{aligned} \quad (10)$$

where $\tilde{p}_{\infty,i} \equiv \lim_{k\to\infty}\tilde{p}_{k,i} = \theta_\infty \cdot \hat{p}_i$, and $\theta_\infty \equiv \lim_{k\to\infty}\theta_k = \dfrac{T}{n+T}$ (see [8] for further details).

### 5.2.4. Reliability Growth Potential

Consider the theoretical upper-limit on reliability that would be achieved by finding and correcting all failure modes in a system with a specified fix effectiveness.

This theoretical upper-limit is known as the reliability growth potential and is a feature of a number of reliability growth models [5], [6] and [11]-[15]. Following from (9) (note that $m \rightarrow k$ as $t \rightarrow \infty$), the reliability growth potential given by our model becomes

$$\tilde{R}_{GP}(t \mid \bar{p}) \equiv \lim_{t \to \infty} \tilde{R}_k(t \mid \bar{p}) = \prod_{i=1}^{k} \left[ 1 - (1 - d_i) \cdot \tilde{p}_{k,i} \right] \tag{11}$$

which can be utilized as a metric for comparison against exit/entry criterion or other threshold values at select program milestones. For instance, even if current reliability growth estimates (e.g., given by (9) or (10)) are below the associated reliability requirement, the system could still have the potential (11) to demonstrate its requirement. The extent to which higher potential reliability is achieved, however, depends on finding and effectively correcting additional failure modes. On the other hand, a system could be identified as high risk of not being able to demonstrate its reliability requirement if the growth potential (11) is less then the requirement.

### 5.2.5. Number of Observed Failure Modes

Using (5), the true number of unique failure modes observed by trial *t* is

$$m(t) \equiv \sum_{i=1}^{k} I_i(t) \tag{12}$$

By using (7) and (12) we have derived the conditional expected number of unique failure modes observed by trial *t* to be

$$\mu_k(t \mid \bar{p}) \equiv E\left[m(t)\right] = \sum_{i=1}^{k} E\left[I_i(t)\right] = \sum_{i=1}^{k} \left[1 - (1 - p_i)^t\right] = k - \sum_{i=1}^{k} (1 - p_i)^t \tag{13}$$

This expression has the following convenient interpretation: the expected number of failure modes observed in $t$ trials is equal to the total potential number of failure modes in the system minus the expected number of failure modes that will not be observed in $t$ trials. For example, realize from (7) that $P\left[I_i(t)=0\right]=\left(1-p_i\right)^t$. Also notice the initial condition of (13) suggests that the expected number of failure modes on trial $t=0$ (i.e., before testing begins) is $\mu_k\left(t=0\mid\bar{p}\right)=0$, as expected.

To derive the limiting behavior of (13) as $k\to\infty$, we have expressed the sum in (13) in terms of the observed and unobserved failure modes and approximated $p_i$ with the MME and MLE of $\tilde{p}_{k,i}$. After some detailed calculation we find

$$\tilde{\mu}_\infty\left(t\mid\bar{p}\right)\equiv\lim_{k\to\infty}\tilde{\mu}_k\left(t\mid\bar{p}\right)=m-\sum_{i\in obs}\left(1-\tilde{p}_{\infty,i}\right)^t+t\cdot\left(1-\theta_\infty\right)\cdot\sum_{i\in obs}\hat{p}_i \qquad (14)$$

where $\tilde{\mu}_k\left(t\mid\bar{p}\right)$ is an MME or MLE estimate of (13) with $\tilde{p}_{\infty,i}$ approximating $p_i$. The limiting behavior of our model suggests that the expected number of observed failure modes by trial $t$ (that are unique) is equal to the expected number of known failure modes to be observed, in addition to the expected number of unknown failure modes to be observed. Notice that the initial condition $\tilde{\mu}_\infty\left(t=0\mid\bar{p}\right)=0$, as expected. Finally, one may ask the question: why are we developing this model equation when we already know how many failure modes were observed during testing? The answer is so we can construct a GOF procedure to determine if our model fits a given sample of one-shot data (i.e., to determine if our model can be suitably applied). Formal statistical GOF procedures are not presented in the scope of the current paper, but will follow in a forthcoming publication.

### 5.2.6. Probability of Discovering a New Failure Mode

Using (5) and given the $p_i$ for $i = 1,\ldots,k$, we define the exact expression for the probability of discovering a new failure mode on trial *t* as

$$h(t \mid \bar{p}) \equiv 1 - \prod_{i=1}^{k} \left[ 1 - \left[ 1 - I_i(t-1) \right] \cdot p_i \right] \qquad (15)$$

Notice from (15) that when a given failure mode has been observed prior to trial *t*, the indicator function equates to unity and its associated failure probability does not contribute to $h(t \mid \bar{p})$. Therefore, only failure modes for which program management is not yet aware (or have not observed in the associated $T$ trials) contribute to this important quantity. Using (7) the expected value of (15) becomes

$$h_k(t \mid \bar{p}) \equiv E\left[ h(t \mid \bar{p}) \right] = 1 - \prod_{i=1}^{k} \left[ 1 - (1 - p_i)^{t-1} \cdot p_i \right] \qquad (16)$$

which has the following initial condition

$$h_k(t = 1 \mid \bar{p}) = 1 - \prod_{i=1}^{k} (1 - p_i) \qquad (17)$$

In other words, our model suggests that the expected probability of discovering a new failure mode on the first trial is equivalent to the initial system probability of failure, as expected.

To derive the limiting form as $k \to \infty$, we approximate $p_i$ with an MME and MLE of $\tilde{p}_{k,i}$ and express (16) in terms of the observed and unobserved failure modes which yields

$$\tilde{h}_\infty(t \mid \bar{p}) \equiv \lim_{k\to\infty} \tilde{h}_k(t \mid \bar{p}) = 1 - \prod_{i \in obs}\left[1-\left(1-\tilde{p}_{\infty,i}\right)^{t-1}\cdot \tilde{p}_{\infty,i}\right]\cdot \exp\left[-\left(1-\theta_\infty\right)\cdot \sum_{j=1}^{m}\hat{p}_j\right] \quad (18)$$

where $\tilde{h}_k(t \mid \bar{p})$ is the an estimate of (16) with $\tilde{p}_{k,i}$ approximating $p_i$ for $i = 1,\ldots,k$.

### 5.2.7. Portion of System Unreliability Observed

Another useful metric to program management is the portion of system unreliability associated with failure modes that have already been observed during testing. Using (5) notice that the probability of observing a repeat failure mode on trial $t$ is given by

$$\psi(t \mid \bar{p}) \equiv 1 - \prod_{i=1}^{k}\left[1 - I_i(t-1)\cdot p_i\right] \quad (19)$$

In other words, if failure mode *i* is observed before trial *t*, the resulting value of the indicator function is unity and the associated failure probability, $p_i$, will contribute to $\psi(t \mid \bar{p})$. If, on the other hand, the failure mode is not observed before trial *t*, the value of the indicator will be zero and the associated failure probability will not contribute to $\psi(t \mid \bar{p})$. Using (17) and (19) we express the expected probability of failure on trial $t$ due to a repeat failure mode as a fraction of the initial system unreliability. This fraction is given by

$$\phi_k(t \mid \bar{p}) \equiv \frac{E\left[\psi(t \mid \bar{p})\right]}{h_k(t=1 \mid \bar{p})} = \frac{1-\prod_{i=1}^{k}\left[1-\left[1-\left(1-p_i\right)^{t-1}\right]\cdot p_i\right]}{1-\prod_{i=1}^{k}\left(1-p_i\right)} \quad (20)$$

Notice that the initial condition of (20) is $\phi_k\left(t=1\mid\bar{p}\right)=0$. This means that the expected probability of system failure on the first trial due to a repeat failure mode is zero, as expected. Also note that (19) and (20) are quantities that are independent of the corrective action process. To take the limit of (20) as $k\to\infty$, we proceed in a similar fashion as above by approximating $p_i$ with an MME or MLE of $\tilde{p}_{k,i}$ and expressing the equation in terms of the observed and unobserved failure modes. After simplification we obtain

$$\tilde{\phi}_\infty\left(t\mid\bar{p}\right)\equiv\lim_{k\to\infty}\tilde{\phi}_k\left(t\mid\bar{p}\right)=\frac{1-\prod_{i\in obs}\left[1-\left[1-\left(1-\tilde{p}_{\infty,i}\right)^{t-1}\right]\cdot\tilde{p}_{\infty,i}\right]}{\tilde{h}_\infty\left(t=1\mid\bar{p}\right)} \tag{21}$$

$\tilde{h}_\infty\left(t=1\mid\bar{p}\right)$ follows from (18) and $\tilde{\phi}_k\left(t\mid\bar{p}\right)$ is the estimate of (20).

## 5.3. Numerical Example

### 5.3.1. Estimation of Model Parameters

Using *Monte Carlo simulation*[41], we present the following small numerical example to illustrate the proposed methodology for a system comprised of $k=10$ failure modes which is tested for $T=50$ trials. Only 10 failure modes are simulated to minimize the volume of output presented in this example. Only 50 trials are simulated in order to minimize the volume of failure data available for estimation purposes. The second column in Table 5, titled True, represents the true failure

[41] ***Monte Carlo simulation*** was utilized to construct this example since the true failure probabilities are unknown in practice. Hence, by using Monte Carlo methods, the true values of our model equations are known and can be compared against our finite and limiting approximations.

probabilities which are unknown in practice. These failure probabilities were generated from a beta distribution with a s-mean of $\mu = 0.025$ and s-variance of $\sigma^2 = 0.0025$. Trial outcomes are generated from a Bernoulli r.v. (1 if a failure mode occurred, 0 otherwise) with parameter $p_i$ for each failure mode $i = 1, \ldots, k$ (values shown). The sum of the trial outcomes across the 50 trials are given in the third column of Table 5, titled Failures. Column four shows the FEF which are generated via a beta r.v. with a s-mean of 0.80 and a s-variance of 0.01. Unobserved failure modes are also assigned a finite and positive FEF in order to estimate the reliability growth potential given by (11). Column five shows the standard MLE of a failure probability (1) computed from failure data. The remaining columns are our approximations of the true failure probability using (2) in addition to the MME and MLE procedures presented in [10] and our derived limiting approximations in [8] for $k \to \infty$.

| | **True** $p_i$ | **Failures** $N_i$ | **FEF** $d_i$ | **Standard MLE** $\hat{p}_i$ | **MME K** $\breve{\tilde{p}}_{k,i}$ | **MME ∞** $\breve{\tilde{p}}_{\infty,i}$ | **MLE K** $\frown{\tilde{p}}_{k,i}$ | **MLE ∞** $\frown{\tilde{p}}_{\infty,i}$ |
|---|---|---|---|---|---|---|---|---|
| **1** | 0.0317 | 0 | 0.74 | 0 | 0.0051 | 0 | 0.0030 | 0 |
| **2** | 0.0539 | 2 | 0.74 | 0.0400 | 0.0366 | 0.0329 | 0.0380 | 0.0361 |
| **3** | 0.0000 | 0 | 0.88 | 0 | 0.0051 | 0 | 0.0030 | 0 |
| **4** | 0.0008 | 0 | 0.92 | 0 | 0.0051 | 0 | 0.0030 | 0 |
| **5** | 0.0109 | 0 | 0.70 | 0 | 0.0051 | 0 | 0.0030 | 0 |
| **6** | 0.1114 | 7 | 0.78 | 0.1400 | 0.1154 | 0.1152 | 0.1254 | 0.1263 |
| **7** | 0.0086 | 0 | 0.94 | 0 | 0.0051 | 0 | 0.0030 | 0 |
| **8** | 0.0140 | 3 | 0.91 | 0.0600 | 0.0524 | 0.0494 | 0.0555 | 0.0541 |
| **9** | 0.0193 | 0 | 0.62 | 0 | 0.0051 | 0 | 0.0030 | 0 |
| **10** | 0.0013 | 0 | 0.82 | 0 | 0.0051 | 0 | 0.0030 | 0 |

**Table 5. Failure Data and Estimates.**

For all tables presented herein, numbers given as 0.0000 have a small but finite and positive value, whereas numbers given as 0 are null. From Table 5 we can see that only 3 of 10 failure modes were observed in 50 trials with 2, 7, and 3 failures for failure modes 2, 6, and 8, respectively. Note that while the MME and MLE estimate of $\tilde{p}_{k,i} \rightarrow 0$ as $k \rightarrow \infty$ for $i \in obs'$, the associated vector of these estimates yields a finite and positive approximation of $\tilde{h}_{\infty}(t \mid \bar{p})$. Using the 12 failures shown in Table 5, approximations of the beta shape parameters, *n* and *x*, are calculated and shown in Table 6.

| | True | MME K | MME ∞ | MLE K | MLE ∞ |
|---|---|---|---|---|---|
| ***n*** | 8.75 | 14.99 | 10.76 | 8.03 | 5.41 |
| ***x*** | 0.22 | 0.36 | 0 | 0.19 | 0 |

**Table 6. Beta Parameters.**

The column titled true refers to the true beta parameters used to generate the $p_i$ for $i = 1,...,k$ displayed in Table 5. The columns titled MME K and MLE K refer to the well-known MME and MLE procedures in [10] utilized to estimate the beta parameters for an assumed value of $k$. The columns denoted by MME ∞ and MLE ∞ are our derived limits [8] of these estimators as $k \rightarrow \infty$. The motivation of these limits is twofold: (1) to eliminate the third and final unknown *k* and (2) to study the sensitivity of not knowing *k*.

Table 7 shows estimates of the beta mean and variance. These quantities are used to estimate $E(p_i)$ and $Var(p_i)$ in (4).

| | True | MME K | MME ∞ | MLE K | MLE ∞ |
|---|---|---|---|---|---|
| $\mu$ | 0.0250 | 0.0240 | 0 | 0.0234 | 0 |
| $\sigma^2$ | 0.0025 | 0.0015 | 0 | 0.0025 | 0 |

**Table 7. Beta Mean and Variance.**

The column titled True refers to the true mean and variance of the beta distribution. The remaining columns use the corresponding estimates from Table 6 to approximate the mean and variance of the beta distribution as given by

$$E\left(P_i\right)=\frac{x}{n} \tag{22}$$

and

$$Var\left(P_i\right)=\frac{x\cdot\left(n-x\right)}{n^2\cdot\left(n+1\right)} \tag{23}$$

Notice that the beta mean and variance both converge to zero as $k\to\infty$, which follows since $x\to 0$ as $k\to\infty$. Hence, the distribution becomes degenerate in the limit. Incidentally, this causes no inconvenience, as the important shrinkage factor approximations below (which are a function of the mean and variance) are finite and positive as $k\to\infty$.

Table 8 shows the approximations of our derived shrinkage factor (4) using the mean and variance estimates in Table 7. These approximations of $\theta$ are utilized to calculate our estimates of the failure probabilities given in Table 5, columns 6-9.

| | True | MME K | MME ∞ | MLE K | MLE ∞ |
|---|---|---|---|---|---|
| $\theta$ | 0.7394 | 0.7875 | 0.8229 | 0.8737 | 0.9023 |

**Table 8. Shrinkage Factor.**

The approximations of our model parameters above are utilized in the following sections to illustrate the proposed analytical framework.

### 5.3.2. Reliability Growth

Figure 14 below displays a plot of the conditional expected reliability growth of a one-shot system. The displayed curves can be interpreted for each $t$ as the expected true system reliability and associated estimates for trial $t$, based on the data through trial $t-1$, given the failure modes surfaced prior to $t$ have been mitigated. There are 5 series in total, all of which are very close to one another. The first series, titled Expected, represents the true reliability growth based on (8) using the true $p_i$ and $d_i$ shown in columns 2 and 4 of Table 5, respectively. The series MME K and MLE K are also computed by (8) with the true $d_i$, but are based on our corresponding MME K and MLE K estimates of $p_i$ given in Table 5. Similarly, the remaining series, MME $\infty$ and MLE $\infty$, are calculated from (10) using the true $d_i$. Error associated with fix effectiveness is not considered for a number of reasons. First, the focus of this example is to illustrate the model and highlight model accuracy with respect to our statistical estimation procedure for the $p_i$. Simulating error in the $d_i$ will incorporate another dimension of error and variability that will cloud our understanding of the accuracy per our methods of estimating failure probabilities. Second, the impetus of many applications of reliability growth projection models is to determine the reliability that could hypothetically be achieved if select failure modes are mitigated with a specified fix effectiveness. Sensitivity analyses and cost trade-off studies associated with the quantity of fixes and degree of their effectiveness are examples of such applications of reliability growth projection models. Finally, all reliability growth projection models whose purpose is estimating reliability after corrective action must use some assessment of the degree to which failures have been mitigated via fixes. Hence, all such models [5]-[8] and [11]-[13] are subject to the same error in assessed fixed effectiveness.

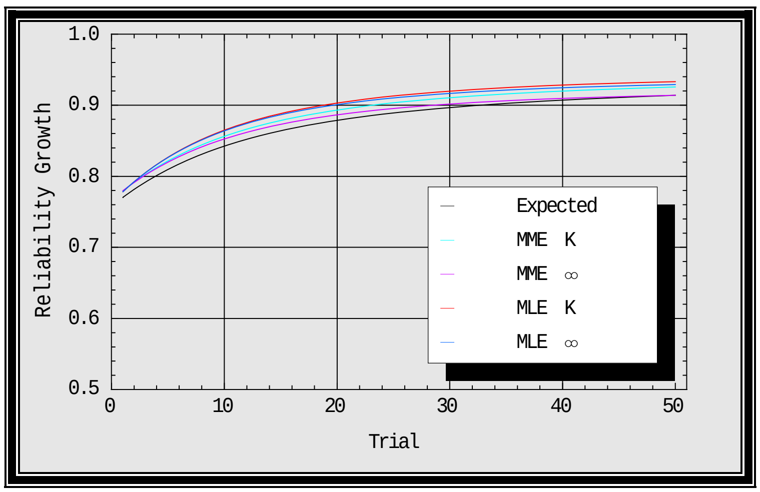


**Figure 14. Reliability Growth.**

The significance of this model equation is to estimate reliability growth (i.e., the ability to estimate reliability improvement resulting from correction to known, or observed, failure modes). In this example, notice the initial condition is the true initial reliability of the system given by $\prod_{i=1}^{k}\left(1-p_i\right)=0.77$. After correcting 3 of the 10 failure modes, with fix effectiveness as specified in Table 5, the true expected reliability has improved from 0.77 to 0.91. The reliability growth potential given by (11) is 0.94. We assume that all fixes are delayed until the end of the current test phase so the major emphasis of Figure 14 is the initial reliability in trial 1 and the final reliability after correction in trial 50. Notice that our model is insensitive to not knowing *k* given by the series displayed for MME ∞ and MLE ∞. More specifically, our assessments in these cases (i.e., where $k \to \infty$) are very accurate despite the fact that there are only a very small number of failure modes (i.e., $k = 10$). Also note model accuracy despite the few failure data (i.e., only 3 failure modes and 7 failures) available for estimation purposes. An important aspect to learn from this example is that a system need not be complex for our limiting approximations to be suitably applied, nor is a large volume of failure data required to obtain reasonably accurate results.

### 5.3.3. Number of Failure Modes

Figure 15 below displays a plot of the cumulative expected number of failure modes versus trials. The series titled Expected is given by (13) using the true value of $k$ and the true values of $p_i$ shown in column 2 of Table 5 . The series for MME K and MLE K are also given by (13) but utilize the corresponding estimates of $p_i$ which are shown in columns 6 and 8 in Table 5, respectively. The remaining series follow from (14) by estimating $p_i$ with the estimates shown in columns 7 and 9 of Table 5. In this example, the failure modes 2, 6, and 8 were discovered on trials 13, 7, and 20 respectively.

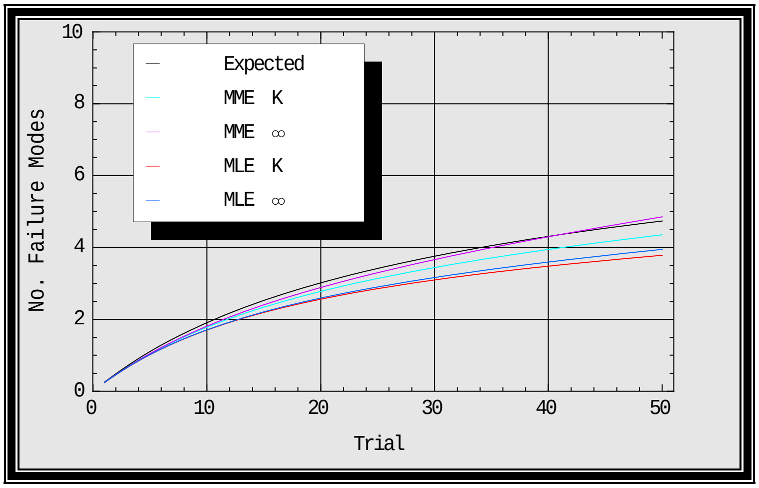


**Figure 15. Number of Failure Modes.**

The significance of this model equation is linked to GOF. Figure 15 provides graphical insight into the GOF of the presented methodology by plotting the actual number of failure modes observed, against the expected number of failure modes given by our model. The example presented in this paper is deliberately kept small (i.e., only 10 failure modes) for brevity in illustration but a more interesting plot of Figure 15 is given in the Appendix for a much larger example. Finally, while we have developed formal statistical GOF procedures for this model, they are not presented in the scope of the current paper.

### 5.3.4. Probability of Discovering a New Failure Mode

Another useful measure to program management is the probability of discovering a new failure mode on trial $t$. The significance of this model equation is associated with programmatic risk. For example, as the development effort is ongoing, it is helpful to gauge the level of maturity of the system by having a quantitative estimate for the likelihood of observing a failure mode that has not yet been uncovered. Clearly, we would like the estimate of $h_k(t \mid \bar{p}) \to 0$ as the TAFT process continues, which would indicate that program management has observed the dominant failure modes in the system. Consequently, the likelihood of the customer encountering unknown failure modes during fielding and deployment can not only be quantified by the model but mitigated through effective management and goal-setting of $h_k(t \mid \bar{p})$.

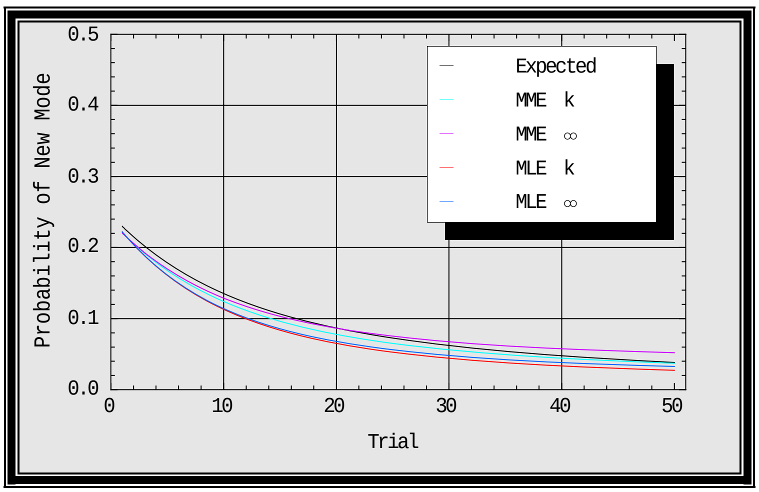


**Figure 16. Probability of Finding a New Failure Mode.**

Figure 16 shows a plot of the expected probability of discovering a new failure mode versus trials. Notice the initial condition is the initial system probability of failure (i.e., 0.23). The series titled Expected is computed via (16) using the true values for *k* and $p_i$. The series denoted by MME K and MLE K are also computed by (16) using the true value of *k*, but use our estimates of $p_i$ given in columns 6 and 8 of Table 5. The remaining series, generated from (18) with corresponding estimates from columns 7 and 9 of Table 5, show our model's insensitivity to not knowing *k*.

### 5.3.5. Portion of System Unreliability Observed

Figure 17 shows a plot of the portion of system unreliability associated with observed failure modes. Notice the initial condition is zero since failure modes have not yet been observed. The series titled Expected is computed from (20) using the true values of $k$ and $p_i$. The series denoted by MME K and MLE K are also generated by (20) with the true value of $k$ but use the corresponding estimates of $p_i$ in Table 5, columns 6 and 8, respectively. The remaining series follow from (21) with $p_i$ estimated by the corresponding estimates given in columns 7 and 9 from Table 5.

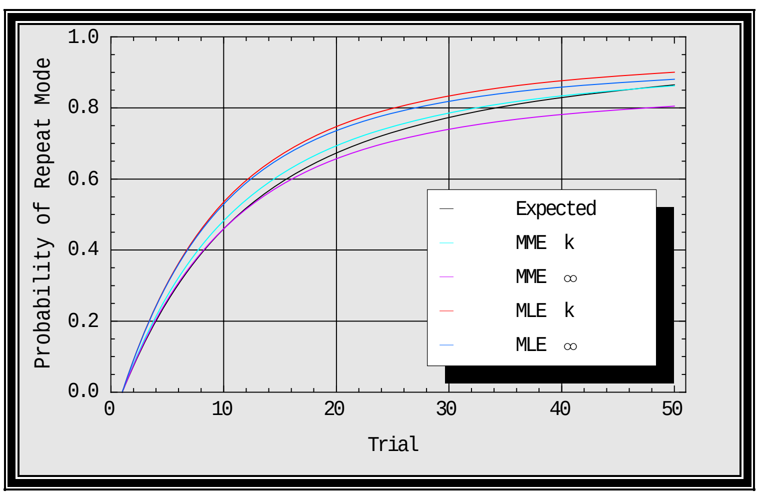


**Figure 17. Portion of System Unreliability Observed.**

This model equation is useful in serving as a system maturity metric. As suggested in [6], specifying goals for $\phi_k(t \mid \bar{p})$ at program milestones would be a good practice as part of a reliability growth program. For example, small values of $\phi_k(t \mid \bar{p})$ indicate that more testing is required to observe additional failure modes that can be corrected. Whereas, large values of $\phi_k\left(t \mid \bar{p}\right)$ may indicate further testing is no longer economically justifiable. Also notice that $\phi_k\left(t \mid \bar{p}\right)$ is independent of the corrective action process in that it does not depend on when fixes are implemented nor how effective they are. Therefore, regardless of fix effectiveness, program management can eliminate at most only a portion of $\phi_k\left(t \mid \bar{p}\right)$ from the initial system unreliability (i.e., probability of failure).

### 5.4. Concluding Remarks

In this paper we have introduced a new reliability growth methodology for one-shot systems. The methodology consists of the following four model equations:

- $R_k\left(t \mid \bar{p}\right)$, expected reliability given by (8), which estimates the reliability improvement of a one-shot system resulting from the correction of failure modes observed during testing. The theoretical upper-limit on reliability (i.e., the reliability growth potential) is given by (11).
- $\mu_k\left(t \mid \bar{p}\right)$, the expected number of failure modes to be observed in testing given by (13), which is useful for addressing model GOF concerns, as well as planning with respect to programmatic corrective action resources.
- $h_k\left(t \mid \bar{p}\right)$, the expected probability of discovering a new failure mode given by (16), which serves as a measure of programmatic risk.

- $\phi_k\left(t \mid \bar{p}\right)$, the expected portion of initial system unreliability associated with failure modes that have been observed given by (20), which serves as a system maturity metric.

These model metrics provide an analytical framework from which reliability practitioners can estimate reliability growth, address GOF concerns, quantify programmatic risk and resource needs, and assess system maturity. A numerical example was provided to illustrate the value and utility of the presented approach. Since model accuracy will vary per application, Monte-Carlo simulation studies of specific systems of interest are recommended.

## Appendix

Figure 18 below displays a plot of the actual observed number of failure modes for a one-shot system comprised of 200 failure modes. In this example, 98 failure modes were observed in a total of 200 trials. The actual observed number of failure modes are represented by black dots. The remaining series are given by our model. The series titled Expected is computed by (13) using the true values of $k$ and $p_i$. The series denoted by MLE K and MLE $\infty$ are generated via (13) and (14), respectively. Given the high correlation of our approximations in relation to the actual observed number of failure modes, Figure 18 suggests that our model reasonably fits the data and can be suitably applied.

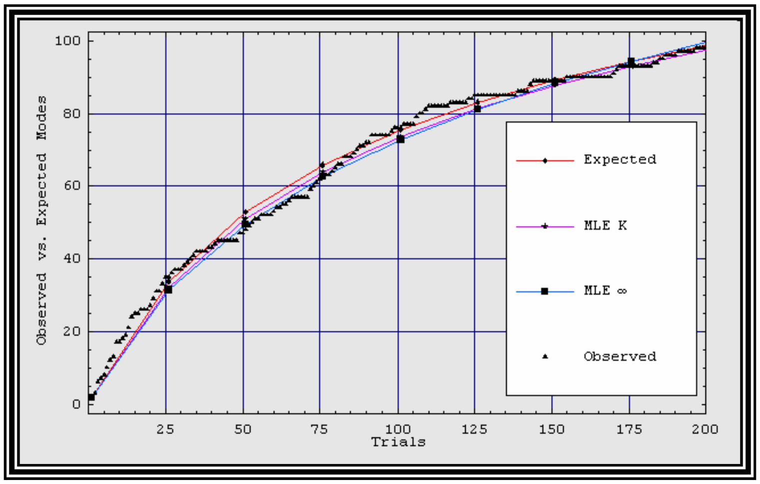


**Figure 18. Number of Failure Modes.**

## Acknowledgments

The authors are grateful to the following individuals for their support in improving this manuscript: Paul M. Ellner, Ph.D. (AMSAA), Michael J. Cushing, Ph.D. (AEC), and Arthur Fries, Ph.D. (IDA).

## References

[1] P. M. Ellner, W. J. Broemm and W. J. Woodworth, "Reliability Growth Guide," AMSAA, Aberdeen Proving Ground, MD, Tech. Rep. No. 652, 2000.

[2] *Reliability Growth Management*, U.S. Department of Defense, Washington, DC, Military Handbook 189, 1981.

[3] A. Fries and A. Sen, "A Survey of Discrete Reliability Growth Models," *IEEE Transactions on Reliability*, vol. R-45, no. 4, pp. 582-604, 1996.

[4] D. E. Mortin and P.M. Ellner, "Reliability Growth for Military Vehicles – Emerging Methodology and Remaining Challenges," *Journal of Commercial Vehicles*, vol. 114, pp.28-34, 2005.

[5] P. M. Ellner and J. B. Hall, "AMSAA Maturity Projection Model based on Stein Estimation," *IEEE Proceedings of the Reliability and Maintainability Symposium*, 2005, pp. 271-277.

[6] P. M. Ellner and L. C. Wald, "AMSAA Maturity Projection Model," *IEEE Proceedings of the Reliability and Maintainability Symposium*, 1995, pp. 174-181.

[7] W. J. Corcoran, H. Weingarten and P. W. Zehna, "Estimating Reliability after Corrective Action," *Management Science*, vol. 10, no. 4, pp. 786-798, 1964.

[8] J. B. Hall and A. Mosleh, "A Reliability Growth Projection Model for One-Shot Systems," *IEEE Transactions on Reliability*, vol. R-56, no. 4, 2007.

[9] C. M. Stein, "Estimation of the Mean of a Multivariate Normal Distribution," *The Annals of Statistics*, vol. 9, pp. 1135-1151, 1981.

[10] H. F. Martz and R. A. Waller, "Selecting Beta Prior Distributions based on Observed Binomial Sampling Data," in *Bayesian Reliability Analysis*, 2nd ed., Malabar, FL: Krieger, 1991, pp. 312-318.

[11] L. H. Crow, "An Improved Methodology for Reliability Growth Projection," *AMSAA*, Aberdeen Proving Ground, MD, Tech. Rep. No. 357, 1982.

[12] L. H. Crow, "Methods for Assessing Reliability Growth Potential," *IEEE Proceedings of the Reliability and Maintainability Symposium*, 1984, pp. 484-489.

[13] L. H. Crow, "An Extended Reliability Growth Model for Managing and Assessing Corrective Actions," *IEEE Proceedings of the Reliability and Maintainability Symposium*, 2004, pp. 73-80.

[14] P. M. Ellner and J. B. Hall, "An Approach to Reliability Growth Planning based on Failure Mode Discovery and Correction using AMSAA Projection Methodology," *IEEE Proceedings of the Reliability and Maintainability Symposium*, 2006, pp. 266-272.

[15] L. H. Crow, "Methods for Reducing the Cost to Maintain a Fleet of Repairable Systems," *IEEE Proceedings of the Reliability and Maintainability Symposium*, 2003, pp. 392-399.

THIS PAGE IS INTENTIONALLY LEFT BLANK.

# 6. MANAGEMENT METRICS FOR NON-DELAYED FIXES[42]

## Abstract

In this paper, reliability growth management metrics are prescribed for one-shot systems under an arbitrary corrective action strategy (i.e., corrective actions can be applied to prototypes at anytime after associated failure modes are first discovered). The methodology consists of four model equations for estimating reliability growth, the expected number of failure modes observed during test, the probability of discovering new failure modes, and the portion of system unreliability associated with repeat failure modes. These model equations can be utilized as management metrics to: estimate reliability (i.e., demonstrated, projected, and growth potential), address model goodness-of-fit concerns, quantify programmatic risk, and assess reliability maturity of one-shot systems undergoing development. A new likelihood function and maximum likelihood procedure is derived to estimate model parameters (i.e., the shape parameters of the beta distribution). Limiting approximations of our management metrics are also given, which are found to be simple functions in terms of only a single unknown parameter. A numerical example is given to illustrate the utility of the presented approach. This methodology is useful to program managers and reliability practitioners who wish to quantitatively assess the progress of the development effort of one-shot systems.

[42] Chapter 6 was submitted to *Technometrics* on 24 March 2008 for consideration of publication.

## - CHAPTER 6 -

**Keywords**: Discrete, Growth Potential, Management Metrics, Projection, Reliability Growth.

### Acronyms[43]

**AEC** – Army Evaluation Center

**AMSAA** – Army Materiel Systems Analysis Activity

**DoD** – Department of Defense

**FEF** – Fix Effectiveness Factor(s)

**FOT** – First Occurrence Trial

**GOF** – Goodness-of-Fit

**IDA** – Institute for Defense Analyses

**MLE** – Maximum Likelihood Estimation/Estimate(s)

**MME** – Method of Moments Estimation/Estimate(s)

**TAFT** – Test, Analyze, Fix and Test

### Definitions

1. **FEF** – fraction reduction in an initial failure mode probability due to implementation of a unique corrective action.
2. **Failure mode** – the root-cause associated with the loss of a required function or component whose probability (or rate) of occurrence is reduced by a specified FEF, if addressed by corrective action. Note that it may be the case that some

[43] The singular and plural of an acronym are always spelled the same.

failure modes are not observed during testing, or may not be corrected if they are observed (e.g., some failures may not be economically justifiable to correct).

3. **Unobserved failure mode** – a failure mode which exhibits zero failures during testing.
4. **Observed failure mode** – a failure mode which exhibits at least one failure during testing.
5. **Repeat failure mode** – a failure mode which exhibits at least two failures during testing.

## Notation

$k$ - total number of potential failure modes.

$m$ - total number of observed failure modes.

$T$ - total number of trials.

$N_i$ - total number of failures for mode $i$ in $T$ trials.

$P_i$ - random variable denoting the true but unknown probability of occurrence of failure mode $i$.

$t_i$ - the trial number when observed failure mode $i$ is first discovered.

$n$ - beta parameter representing the pseudo number of trials.

$x$ - beta parameter representing the pseudo number of failures.

$R_I$ - initial system reliability.

$R_k(t)$ - expected reliability growth on trial $t$.

$R_{k,GP}$ - expected reliability growth potential.

$\mu_k(t)$ - expected number of unique failure modes observed on or before trial $t$.

$h_k(t)$ - expected probability of discovering a new failure mode on trial $t$.

$\varphi_k(t)$ - expected probability of observing a repeat failure mode on trial $t$.

$\phi_k(t)$ - expected fraction of system unreliability on trial $t$ due to a repeat failure mode.

## 6.1. Introduction

### 6.1.1. Background

The elimination of design weaknesses inherent to intermediate prototypes of complex systems via the TAFT process is generally what is meant by the term reliability growth. Specifically, reliability growth is the improvement in the true but unknown initial reliability of a developmental item as a result of failure mode discovery, analysis, and effective correction. Corrective actions generally assume the form of fixes, adjustments, or modifications to problems found in the hardware, software, or human error aspects of a system. Some examples may include (but are not limited to) engineering redesign work of system / subsystem architectures, alterations in the material properties of components, modifications to associated manufacturing and industrial processes, elimination of electrical sneak circuits and software code syntax errors, or amendments to operating (or maintenance) procedures of a system. Since the 1950s (e.g., Weiss [1] being one of the earliest papers on the subject), the genesis of three main areas of the field have emerged where a wealth of methods have been developed to plan, track, and project the reliability of developmental items. Perhaps one of the most well-known, widely-used, and pivotal concepts in the field includes the Weibull process (i.e., the non-homogeneous Poisson process with intensity function $r(t) = \lambda \cdot \beta \cdot t^{\beta-1}$), which is accompanied by a wealth of inference procedures (e.g., see Finkelstein [3], Lee & Lee [4], Engelhardt & Bain [5], Bain & Engelhardt [6], Lee [7], and Crow [10]).

Each of these three areas of reliability growth apply to complex systems whose test durations are measured in the continuous time domain, as well as via discrete trials (e.g., one-shot systems, such as, guns rockets, missile systems, torpedoes). The AMSAA Reliability Growth Methodology Guide given by Ellner, et al. [15], and the Fries-Sen survey of discrete reliability growth models [14] are among the most comprehensive and detailed works on the subject. Military Handbook 189 [8], while outdated, and Appendix C of the DoD Guide for Achieving RAM [19] are also good references covering methods available for reliability growth analysis. Naturally, some areas are more developed than others. In particular, more work needs to be done in the area of discrete reliability growth projection, as indicated by Mortin & Ellner [20]. In this paper, we introduce a robust methodology (independent of the Weibull process) that serves as a management framework from which practitioners can gauge the progress of the development effort of one-shot systems throughout the TAFT process.

### 6.1.2. Overview

This paper is organized as follows. The methodology of our approach is presented in Section 2. This includes a list of assumptions, the technical details of our likelihood function and ML procedure for point-estimation , as well as the mathematical derivations of our management metrics. In section 3, we illustrate the presented approach by a simple numerical example. Concluding remarks are given in section 4.

## 6.2. Methodology

### 6.2.1. Model Assumptions

1. Initial failure mode probabilities of occurrence $p_1, \ldots, p_k$ constitute a realization of a s-random sample $P_1, \ldots, P_k$ such that $P_i \sim Beta(n, x)$ for each $i = 1, \ldots, k$. We shall use the following PDF parameterization,

$$f(p_i) \equiv \begin{cases} \dfrac{\Gamma(n)}{\Gamma(x) \cdot \Gamma(n-x)} \cdot p_i^{x-1} \cdot (1-p_i)^{n-x-1} & p_i \in [0,1] \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where $n$ represents pseudo trials, $x$ represents pseudo failures, and $\Gamma(x) \equiv \int_0^\infty t^{x-1} \cdot e^{-t} dt$ is the Euler gamma function. The associated s-mean, and s-variance of the $P_i$ are given respectively by,

$$E(P_i) = \frac{x}{n} \tag{2}$$

and

$$Var(P_i) = \frac{x \cdot (n-x)}{n^2 \cdot (n+1)} \tag{3}$$

2. The number of trials $t_1, \ldots, t_k$ until the first occurrence of each failure mode constitutes a realization of a random sample $T_1, \ldots, T_k$ such that $T_i \sim Geometric(p_i)$ for each $i = 1, \ldots, k$.
3. Potential failure modes occur s-independently of one another and their occurrence is considered to constitute a failure.

### 6.2.2. Estimation Procedures

#### 6.2.2.1. Likelihood Function

The area of reliability growth projection focuses on estimating the reliability that could be achieved in a system if observed failure modes are addressed via corrective action. There are two types of reliability growth projection models. The first type addresses the case where all corrective actions are delayed until the end of the current test phase, as in [2], [9], [16], [22], and [23]. In general, the functional forms of these models are expressed in terms of failure mode probabilities (or rates) of occurrence. Generally, their statistical estimation procedures only require count data (i.e., the number of failures for observed failure modes in a period of time, or number of trials) and FEF. The count data is obtained directly from testing. The FEF are typically based on expert engineering judgment and assigned by a Failure Prevention and Review Board.

The second type of reliability growth projection models (e.g., [13], [18], and the approach presented herein), address the case were corrective actions can be implemented to system prototypes anytime after associated failure modes have been found (a more complicated scenario). The statistical estimation procedures of these models are generally based on the exact trials (or times) when failure modes were first discovered. In subsections 2.3 - 2.7 below, our model equations are shown to be functions of only the two beta shape parameters (i.e., $n$ and $x$), and $k$, the total potential number of failure modes in the system. The MME and MLE procedures in Martz and Waller [12] are two well-known procedures for approximating the beta shape parameters. These procedures are very useful for the first type of projection models when all fixes are delayed. For the case we are considering, however, these procedures cannot be adequately applied (mainly because the failure probabilities that generate the count data $N_i$ for which these procedures are based does not remain fixed). For example, let $obs \equiv \{i : N_i > 0 \text{ for } i=1,\ldots,k\}$ represent the index set of failure modes observed during testing, and let $obs' \equiv \{j : N_j = 0 \text{ for } j=1,\ldots,k\}$ denote its complement. Also let the total number of observed failure modes be denoted by $m = |obs|$, which gives $|obs'| = k - m$ unobserved failure modes. Since failure modes, in the case considered, can be addressed via corrective action at anytime after they are first discovered, their associated failure probabilities $\vec{p}_{obs} \equiv \{p_i : i \in obs\}$ may not remain constant. If addressed during the test phase, they are reduced by the fractional amount specified by an assigned FEF. Hence, the MME and MLE procedures in Martz and Waller are only tailored to the case where the $p_i$ for $i = 1, \ldots, k$ do not change. As a result, we have developed a new likelihood function and associated MLE procedure to provide suitable approximations of the beta shape parameters in the case where fixes may occur during the test phase. Our likelihood function is given by,

$$L_k\left(m, \vec{t} \mid \vec{P}\right) \equiv m! \cdot \sum_{obs \in S_m} \left( \prod_{i \in obs} \left[ \left(1 - P_i\right)^{t_i - 1} \cdot P_i \right] \cdot \prod_{j \in obs'} \left(1 - P_i\right)^T \right) \tag{4}$$

where

1. $\prod_{i \in obs} \left(1 - P_i\right)^{t_i - 1} \cdot P_i$ is the joint geometric density function of a random sample of size $m$, which represents the probability that the observed failure modes first occur on trials $\vec{t} \equiv \{t_i : i \in obs\}$ (e.g., the term $\left(1 - P_i\right)^{t_i - 1} \cdot P_i$ is a geometric probability of observing failure mode $i$ on trial $t_i$).

2. $\prod_{j \in obs'} \left(1 - P_i\right)^T$ is the joint geometric reliability function of a random sample of size $k - m$, which represents the probability that the unobserved modes do not occur in $T$ trials.

3. The summation over $obs \in S_m$ represents the sum of all mutually exclusive sets comprised of exactly $m$ observed failure modes. Clearly, there are $|S_m| = \binom{k}{m}$ such sets of size $m$, and there are $m!$ ways in which the failure modes in each index set can be observed during testing.

Overall (4) represents the likelihood that the $m$ observed failure modes occur with FOT $\vec{t}$ and that the unobserved failure modes do not occur before trial $T$. By interpreting the $P_i$ in (4) as iid beta r.v., the marginal likelihood function becomes,

$$\begin{aligned} L_k\left(m,\vec{t}\,\right) &\equiv E\left[L_k\left(m,\vec{t}\mid\vec{P}\right)\right] = \\ &= m!\cdot\sum_{obs\in S_m}\left(\prod_{i\in obs} E\left[\left(1-P_i\right)^{t_i-1}\cdot P_i\right]\cdot\prod_{j\in obs'} E\left[\left(1-P_i\right)^{T}\right]\right) \\ &= m!\cdot\binom{k}{m}\cdot\prod_{i=1}^{m}\left[\frac{\Gamma\left(n-x+t_i-1\right)\cdot\Gamma\left(x+1\right)}{\mathrm{B}\left(x,n-x\right)\cdot\Gamma\left(n+t_i\right)}\right]\cdot\prod_{j=1}^{k-m}\left[\frac{\Gamma\left(n\right)\cdot\Gamma\left(n-x+T\right)}{\Gamma\left(n-x\right)\cdot\Gamma\left(n+T\right)}\right] \\ &= \frac{k!}{\left(k-m\right)!}\cdot\left[\frac{\Gamma\left(n\right)\cdot\Gamma\left(n-x+T\right)}{\Gamma\left(n-x\right)\cdot\Gamma\left(n+T\right)}\right]^{k-m}\cdot\prod_{i=1}^{m}\left[\frac{\Gamma\left(x+1\right)\cdot\Gamma\left(n-x+t_i-1\right)}{\mathrm{B}\left(x,n-x\right)\cdot\Gamma\left(n+t_i\right)}\right] \end{aligned} \tag{5}$$

where $\mathrm{B}\left(a,b\right)\equiv\frac{\Gamma\left(a\right)\cdot\Gamma\left(b\right)}{\Gamma\left(a+b\right)}=\int_0^1 t^{a-1}\cdot\left(1-t\right)^{b-1}\,dt$ is the Euler beta function. Notice that the middle product-term in (5) represents the $k-m$ Type I (i.e., time) censored observations, and that the product on the right represents the "complete observations" (i.e., the $m$ failure modes discovered on trials $t_i$ for $i=1,\ldots,m$).

### 6.2.2.2. Maximum Likelihood Estimates

#### 6.2.2.2.1. Finite $k$ Approximations

The partial derivatives of the natural logarithm of the likelihood (5) with respect to the beta shape parameters $n$ and $x$, respectively, yield the following MLE equations:

$$\frac{\partial \ln L_k\left(m,\vec{t}\right)}{\partial n}=0\Rightarrow$$
$$k\cdot\left[\psi\left(n\right)-\psi\left(n-x\right)+\psi\left(n-x+T\right)-\psi\left(n+T\right)\right]= \tag{6}$$
$$=\sum_{i=1}^{m}\left[\psi\left(n+t_i\right)-\psi\left(n-x+t_i-1\right)+\psi\left(n-x+T\right)-\psi\left(n+T\right)\right]$$

and

$$\frac{\partial \ln L_k\left(m,\vec{t}\right)}{\partial x}=0\Rightarrow$$
$$k\cdot\left[\psi\left(n-x\right)-\psi\left(n-x+T\right)\right]=\sum_{i=1}^{m}\left[\psi\left(n-x+t_i-1\right)-\frac{1}{x}-\psi\left(n-x+T\right)\right] \tag{7}$$

where $\psi\left(x\right)\equiv\frac{\Gamma'\left(x\right)}{\Gamma\left(x\right)}$ denotes the digamma function. These equations, when solved simultaneously, yield the MLE $\hat{n}$ and $\hat{x}$ that maximize the marginal likelihood function (5). Notice that these ML equations depend upon the unknown value of $k$. In applications where the system is well understood, one can assume a value of $k$ and use (6) and (7) to estimate the parameters. To avoid such an assumption, the limiting approximations of these MLE as $k\rightarrow\infty$ can be used. These limiting approximations, derived in the following section, reveal the sensitivity of not knowing $k$. They also show the limiting behavior of the assessment procedure for complex systems. Via heuristics in Monte Carlo simulation, we have found that the management metrics are not sensitive to the value of $k$ provided that it is chosen to be large (i.e., $k\geq 5\cdot m$ as a rule-of-thumb). Thus, we do not bother deriving an ML estimate for $k$. The numerical example given in Section 3 illustrates the rapid convergence of these limiting approximations (e.g., there is little difference in the magnitudes of the management metrics between assuming $k=25$ failure modes versus infinitely many failure modes).

### 6.2.2.2.2. Limiting Approximations

To derive the limiting behavior of the likelihood function, we proceed by reparameterizing (5), and taking the limit as $k \to \infty$. A logical reparameterization is obtained by using the fact that the true but unknown initial reliability of a one-shot system is a realization of the product $\prod_{i=1}^{k}(1-P_i)$ (i.e., the product of failure mode success probabilities). By interpreting the $P_i$ for $i=1,\ldots,k$ as an iid beta r.v., the resulting unconditional expectation is $R_{k,I} = \left(1-\frac{x}{n}\right)^k$. This yields the useful reparameterization of $x$ into terms of $n$ and the expected initial reliability of the system $R_{k,I}$,

$$x = n \cdot \left(1 - R_{k,I}^{1/k}\right) \tag{8}$$

In this expression, notice that $x \to 0$ as $k \to \infty$, hence our first motivation for reparameterizing. The second motivation is that the reparameterization allows us to obtain an estimate of the expected initial reliability the system, which is a quantity of basic interest to management. After reparameterizing, the limit of the natural logarithm of the likelihood (5) is,

$$\begin{aligned} \ln L_\infty\left(m,\bar{t}\right) &\equiv \lim_{k\to\infty} \ln L_k\left(m,\bar{t}\right) = \\ &= \ln R_{k,I}^{n} \cdot \left[\psi\left(n+T\right)-\psi\left(n\right)\right] + \sum_{i=1}^{m} \ln\left[\frac{\ln R_{k,I}^{-n} \cdot \Gamma\left(n+t_i-1\right)}{\Gamma\left(n+t_i\right)}\right] \end{aligned} \tag{9}$$

The first partial derivative of (9) with respect to $R_{k,I}$ is $\frac{\partial \ln L_\infty\left(m,\vec{t}\right)}{\partial R_{k,I}} = \frac{m}{\ln R_{k,I}} - n\cdot\left[\psi\left(n+T\right)-\psi\left(n\right)\right]$. Maximizing this expression yields a limiting approximation for the expected initial reliability of the system,

$$\hat{R}_{\infty,I} = \exp\left(\frac{-m}{\hat{n}_\infty\cdot\left[\psi\left(\hat{n}_\infty+T\right)-\psi\left(\hat{n}_\infty\right)\right]}\right) \tag{10}$$

In (10), $\hat{n}_\infty$ is obtained as the numerical solution to,

$$\frac{\partial \ln L_\infty\left(m,\vec{t}\right)}{\partial \hat{n}_\infty} = \sum_{i=1}^{m}\left(\frac{1}{\hat{n}_\infty+t_i-1}\right)+m\left[\frac{\psi'\left(\hat{n}_\infty+T\right)-\psi'\left(\hat{n}_\infty\right)}{\psi\left(\hat{n}_\infty+T\right)-\psi\left(\hat{n}_\infty\right)}\right]=0 \tag{11}$$

which is the partial derivative of (9) with respect to the beta parameter $n$.

There are a few notable features of these limiting approximations. First, we have derived a mathematically convenient functional form for the expected initial reliability of a one-shot system, given by (10). Second, our limiting approximations reduce the estimation procedure to solving only one equation, for one unknown, namely, the beta shape parameter $n$ (i.e., pseudo trials). Another interesting discovery is that the functional form for the initial reliability (10) derived from our likelihood function (5) is identical to that which we derived similarly in [22] from the likelihood function of the beta-binomial distribution given by Martz and Waller in [12],

$$L\left(N_i\right)=\prod_{i=1}^{k}\frac{T!\cdot\Gamma\left(n\right)\cdot\Gamma\left(N_i+x\right)\cdot\Gamma\left(T+n-N_i-x\right)}{N_i!\cdot\left(T-N_i\right)!\cdot\Gamma\left(T+n\right)\cdot\Gamma\left(x\right)\cdot\Gamma\left(n-x\right)} \tag{12}$$

Note that the two likelihood functions (5) and (12) are quite different from one another, and even require different types of data (i.e., the trial numbers of first occurrence $t_i$ for (5), and count data $N_i$ for (12). One reason why this result may not be surprising is that the initial reliability of a system is independent of the corrective action process (e.g., when fixes are applied to prototypes), which is the primary difference between the two likelihood functions.

### 6.2.3. Model Equations

#### 6.2.3.1. Overview

The methodology presented herein is comprised of four management metrics, which are derived in the following sections. These metrics build off of the methodology advanced in our earlier paper [23], which addresses the case where all corrective actions are delayed until the end of the current test phase. We now address the more complicated case where corrective actions can be installed after failure modes are first discovered. These equations are extensions of the earlier ones in the sense that they are unconditional expectations of their counterpart metrics (i.e., unconditioned on the $P_i$ for $i = 1, \ldots, k$). The resulting expressions below are found to be functions of the two beta shape parameters, rather than the vector of unknown failure probabilities inherent to the system. Equation numbers from our earlier publication [23] are given for cross-reference.

#### 6.2.3.2. Expected Reliability

Per Equation 8 in [23], the expected reliability of a one-shot system on trial $t$ conditioned on the vector of unknown failure probabilities $\vec{P} \equiv (P_1, \ldots, P_k)$ is given as,

$$R_k(t \mid \vec{P}) \equiv \prod_{i=1}^{k} \left[ 1 - \left( 1 - \left[ 1 - (1 - P_i)^{t-1} \right] \cdot d_i \right) \cdot P_i \right] \tag{13}$$

The unconditional expectation of $R_k(t \mid \vec{P})$ with respect to the $P_i$ for $i = 1, \ldots, k$ is,

$$R_k(t) \equiv E\left[ R_k(t \mid \vec{P}) \right] \approx \left[ 1 - \left( 1 - \left[ 1 - \frac{\Gamma(n - x + t - 1) \cdot \Gamma(n+1)}{\Gamma(n-x) \cdot \Gamma(n+t)} \right] \cdot \bar{d} \right) \cdot \frac{x}{n} \right]^k \tag{14}$$

where $\bar{d} \equiv \sum_{i \in obs} d_i \big/ m$ is an average FEF. This expression models the true but unknown expected reliability of a one-shot system on trial $t$, where corrective actions can be implemented at anytime after their associated failure modes are first discovered. The parameters $n$ and $x$ in (14) are estimated by the solutions which satisfy (6) and (7) simultaneously. Note that our model is independent of the *A-mode*[44] / *B-mode*[45] classification scheme proposed in [9], as A-modes need only be distinguished from B-modes via a zero FEF (i.e., $d_i = 0$ if failure mode $i$ is not observed, or is not corrected). Also notice that the initial condition of (14) equates to the expected initial reliability of the system as required,

$$R_k(t = 1) = \left( 1 - \frac{x}{n} \right)^k \tag{15}$$

It is also desirable to study the limiting behavior of (14) as $k \to \infty$, since the total potential number of failure modes inherent to a complex one-shot system is

[44] An ***A-mode*** is a failure mode that will not be addressed via corrective action.
[45] A ***B-mode*** is a failure mode that will be addressed via corrective action, if observed during testing.

typically large, and since $k$ is unknown. After reparameterizing (14) via (8), our limiting approximation simplifies too,

$$\hat{R}_{\infty}\left(t\right) \equiv \lim_{k\to\infty} \hat{R}_{k}\left(t\right) = \hat{R}_{\infty,I}^{1-\bar{d}\cdot\left(\frac{t-1}{\hat{n}_{\infty}+t-1}\right)} \tag{16}$$

where $\hat{R}_{\infty,I}$ and $\hat{n}_{\infty}$ are obtained via (10) and (11), respectively.

### 6.2.3.3. Reliability Growth Potential

Reliability growth potential [11] is a characteristic of a number of reliability growth models, such as [9], [13], [16-18], and [22-23]. It represents the theoretical upper-limit on reliability achieved by finding and effectively correcting all failure modes in a system with a specified fix effectiveness. Per Equation 11 in [23], the reliability growth potential for one-shot systems is given by,

$$R_{k,GP}\left(\vec{P}\right) \equiv \prod_{i=1}^{k}\left[1-\left(1-d_{i}\right)\cdot P_{i}\right] \tag{17}$$

The unconditional expectation of (17) with respect to the $P_i$ for $i=1,\ldots,k$ is,

$$R_{k,GP} \equiv E\left[R_{k,GP}\left(\vec{P}\right)\right] = E\left(\prod_{i=1}^{k}\left[1-\left(1-d_{i}\right)\cdot P_{i}\right]\right) \approx \left[1-\left(1-\bar{d}\right)\cdot\frac{x}{n}\right]^{k} \tag{18}$$

To estimate (18), the value of $k$ is assumed and the parameters $n$ and $x$ in approximated by the MLE obtained from (6) and (7). After reparameterizing (18) via (8) and taking the limit as $k\to\infty$, the limiting behavior simplifies to,

$$\hat{R}_{\infty,GP} \equiv \lim_{k\to\infty} \hat{R}_{k,GP} = \hat{R}_{\infty,I}^{1-\bar{d}} \tag{19}$$

where $\hat{R}_{\infty,I}$ is given by (10).

#### 6.2.3.4. Expected Number of Failure Modes

Per Equation 13 in [23], the conditional expected number of unique failure modes observed on or before trial $t$ is given by,

$$\mu_k\left(t \mid \vec{P}\right) = k - \sum_{i=1}^{k}\left(1-P_i\right)^t \tag{20}$$

The resulting unconditional expectation of (20) with respect to the $P_i$ for $i = 1,\ldots,k$ is,

$$\begin{aligned}\mu_k\left(t\right) \equiv E\left[\mu_k\left(t \mid \vec{P}\right)\right] &= k - \sum_{i=1}^{k} E\left[\left(1-P_i\right)^t\right] \\ &= k - k\cdot\left[\frac{\Gamma\left(n\right)\cdot\Gamma\left(n-x+t\right)}{\Gamma\left(n-x\right)\cdot\Gamma\left(n+t\right)}\right]\end{aligned} \tag{21}$$

These expressions have the following convenient interpretation: the expected number of unique failure modes observed in $t$ trials is equal to the total potential number of failure modes in the system minus the expected number of failure modes that will not be observed in $t$ trials. The initial condition of (21) implies that the expected number of failure modes observed on trial $t = 0$ (i.e., before testing begins) is $\mu_k\left(t = 0\right) = 0$, as expected. An estimate of (21) is obtained by using the finite $k$ MLE for the beta shape parameters $n$ and $x$.

To derive the limiting behavior of (21), we have used the reparameterization (8) and taken the limit as $k \to \infty$. After some detailed calculation we find,

$$\hat{\mu}_{\infty}(t) \equiv \lim_{k\to\infty} \hat{\mu}_k(t) = \hat{n}_{\infty} \cdot \ln \hat{R}_{\infty,I} \cdot \left[\psi(\hat{n}_{\infty}) - \psi(\hat{n}_{\infty} + t)\right]$$
$$= m \cdot \left[\frac{\psi(\hat{n}_{\infty} + t) - \psi(\hat{n}_{\infty})}{\psi(\hat{n}_{\infty} + T) - \psi(\hat{n}_{\infty})}\right] \tag{22}$$

where $\hat{n}_{\infty}$ is the numerical solution of (11). The significance of (21) and (22) is associated with model GOF. As illustrated by numerical example below, this management metric provides graphical insight into the GOF of the presented approach. This is achieved by plotting the actual cumulative number of failure modes observed during testing versus trials, against the cumulative expected number of failure modes versus trials (given by the model). Good agreement between the actual stochastic realization and our estimates indicate that the model reasonably fits the data and the associated reliability growth management metrics can be suitably applied. Finally, notice that (21) and (22) are mean-value functions that can be compared against $\mu(t) = \lambda \cdot t^{\beta}$ from a typical Weibull process approach.

### 6.2.3.5. Expected Probability of Discovering a New Failure Mode

Per Equation 16 in [23], the conditional expected probability of discovering a new failure mode on trial $t$ is given as,

$$h_k(t \mid \vec{P}) \equiv E\left[h(t \mid \vec{P})\right] = 1 - \prod_{i=1}^{k}\left[1 - (1 - P_i)^{t-1} \cdot P_i\right] \tag{23}$$

The unconditional expectation of (23) with respect to the $P_i$ for $i = 1, \ldots, k$ is,

$$h_k(t) \equiv E\left[h_k(t \mid \vec{P})\right] = 1 - \left[1 - \frac{\Gamma(n - x + t - 1) \cdot \Gamma(x + 1)}{\mathrm{B}(x, n - x) \cdot \Gamma(n + t)}\right]^k \tag{24}$$

Equation (24) is estimated by using the finite $k$ approximations for $n$ and $x$ obtained as the solution to (6) and (7). The initial condition of (24) equates to,

$$h_k(t=1)=1-\left[1-\frac{\Gamma(n-x)\cdot\Gamma(x+1)}{\mathrm{B}(x,n-x)\cdot\Gamma(n+1)}\right]^k=1-\left(1-\frac{x}{n}\right)^k=1-R_k(t=1) \qquad (25)$$

This means that the expected probability of discovering a new failure mode on the first trial is equivalent to the initial system probability of failure, as expected.

After reparameterizing via (8), the limiting approximation of (24) as $k\to\infty$ simplifies to,

$$\hat{h}_\infty(t)\equiv\lim_{k\to\infty}\hat{h}_k(t)=1-\hat{R}_{\infty,I}^{\frac{\hat{n}_\infty}{\hat{n}_\infty+t-1}} \qquad (26)$$

The estimates $\hat{R}_{\infty,I}$ and $\hat{n}_\infty$ in (26) are obtained by (10) and (11), respectively. The expressions above estimate the expected probability of discovering a new failure mode on trial $t$, and can be utilized as a measure of programmatic risk. For example, as the development effort (e.g., TAFT process) continues, we would like the estimate of $h_k(t)\to 0$, which would indicate that program management has observed the dominant failure modes in the system. Conversely, large values of $h_k(t)$ would indicate higher programmatic risk with respect to additional unseen failure modes inherent in the current design. Effective management and goal-setting of $h_k(t)$ would be a good practice to reduce the likelihood of the customer encountering unknown failure modes during fielding and deployment.

#### 6.2.3.6. Portion of System Unreliability Observed

In the absence of failure mode mitigation during test, the portion of system unreliability on trial $t$ associated with failure modes that have already been observed during testing (e.g., the probability of only observing repeat failure modes with continued testing) is given in [23] as,

$$\varphi\left(t \mid \vec{P}\right) = 1 - \prod_{i=1}^{k}\left[1 - \left[1 - \left(1 - P_i\right)^{t-1}\right] \cdot P_i\right] \tag{27}$$

The unconditional expectation of (27) with respect to the $P_i$ for $i = 1, \ldots, k$ is,

$$\varphi_k\left(t\right) \equiv E\left[\varphi\left(t \mid \vec{P}\right)\right] = 1 - \left(1 - \left[\frac{x}{n} - \frac{\Gamma\left(n - x + t - 1\right) \cdot \Gamma\left(x + 1\right)}{\mathrm{B}\left(x, n - x\right) \cdot \Gamma\left(n + t\right)}\right]\right)^k \tag{28}$$

Using (25) and (28) we express the expected probability of failure on trial $t$ due to a repeat failure mode as a fraction of initial system unreliability. This fraction in the absence of failure mode mitigation is given by,

$$\phi_k\left(t\right) \equiv \frac{\varphi_k\left(t\right)}{h_k\left(t = 1\right)} = \frac{1 - \left(1 - \left[\frac{x}{n} - \frac{\Gamma\left(n - x + t - 1\right) \cdot \Gamma\left(x + 1\right)}{\mathrm{B}\left(x, n - x\right) \cdot \Gamma\left(n + t\right)}\right]\right)^k}{1 - R_k\left(t = 1\right)} \tag{29}$$

Once again, approximations of this management metric are obtained via the MLE procedure based on (6) and (7). The initial condition of (29) is $\phi_k\left(t = 1\right) = 0$, which means that the expected probability of failure on the first trial due to a repeat failure mode is zero, as expected.

To take the limit of (29) as $k \to \infty$, we proceed in a similar fashion as above by using the reparameterization (8). After simplification we obtain,

$$\hat{\phi}_\infty\left(t\right) \equiv \lim_{k \to \infty} \hat{\phi}_k\left(t\right) = \frac{1 - \hat{R}_{\infty,I}^{\frac{t-1}{\hat{n}_\infty + t - 1}}}{1 - \hat{R}_{\infty,I}} \tag{30}$$

where $\hat{R}_{\infty,I}$ and $\hat{n}_{\infty}$ follow from (10) and (11), respectively. The value, or benefit, of these expressions is that they can be used as a system maturity metric. For instance, a good management practice would be to specify goals for $\phi_{\infty}(t)$ at important program milestones in order to track the progress of the development effort with respect to the maturing design of the system (from a reliability standpoint). Small values of $\phi_{\infty}(t)$ indicate that further testing is required to find and effectively correct additional failure modes. Conversely, large values of $\phi_{\infty}(t)$ indicate that further pursuit of the development effort to increase system reliability may not be economically justifiable (i.e., the cost may not be worth the gains that could be achieved). Finally, note that program management can eliminate at most the portion $\phi_{\infty}(t)$ from the initial system unreliability at the conclusion of trial $t$ regardless of when fixes are installed or how effective they are (i.e., since this metric is independent of the corrective action process).

## 6.3. Numerical Example

### 6.3.1. Estimation of Model Parameters

Since the true failure probabilities and values of the beta shape parameters $n$ and $x$ are unknown in practice, we have utilized Monte Carlo simulation to illustrate the proposed methodology via numerical example. In this example, the system is comprised of $k = 25$ failure modes and is tested for $T = 50$ trials. Only $m = 7$ in 25 failure modes were observed. Table 9 shows the associated failure data. The second column shows the failure mode numbers that comprise the index set of observed failure modes denoted by $obs$. The third column is the true failure mode probability of occurrence that is unknown in practice. The fourth column shows the trial number when each observed failure mode was first discovered. Our statistical estimation procedures for the beta shape parameters only require and use the FOT (i.e., the seven datapoints shown in column 4).

| | **Failure Mode** $i$ | **Failure Probability** $p_i$ | **FOT** $t_i$ |
|---|---|---|---|
| **1** | 2 | 0.1329 | 14 |
| **2** | 4 | 0.0108 | 39 |
| **3** | 6 | 0.0596 | 22 |
| **4** | 14 | 0.0171 | 33 |
| **5** | 15 | 0.0946 | 2 |
| **6** | 20 | 0.0755 | 1 |
| **7** | 23 | 0.0180 | 11 |

**Table 9. Failure Data.**

The MLE of the beta shape parameters approximated from the FOT data are shown in Table 10. The column titled True denotes the true values of beta parameters that were utilized to stochastically generate the $p_i$ for $i = 1,\ldots,k$. The column titled MLE K displays the estimates that are obtained when maximizing equations (6) and (7) simultaneously. Recall that these procedures require an assumed value of $k$, the total potential number of failure modes in the system. In this example, the true value of $k = 25$ is used. The sensitivity of not knowing $k$ is highlighted by our limiting approximations shown in the column labeled MLE ∞. In this case, the estimates are obtained as the numerical solutions that maximize (11).

| | True | MLE K | MLE ∞ |
|---|---|---|---|
| $n$ | 8.800 | 7.193 | 4.721 |
| $x$ | 0.176 | 0.152 | 0.000 |

**Table 10. Maximum Likelihood Estimates.**

Recall, per equation (8), that $x = n \cdot \left(1 - R_{k,I}^{1/k}\right) \to 0$ as $k \to \infty$, which means that the beta mean (20) and variance (21) also converge to zero. Hence, the distribution becomes degenerate in the limit. Incidentally, however, this causes no inconvenience, as the important limiting approximations of our model equations remain finite and positive.

### 6.3.2. Expected Reliability Growth

Figure 19 below displays a plot of the expected reliability of a one-shot system versus trials (that would typically be conducted in a developmental test program). The curves can be interpreted as an estimate of reliability on trial $t$ that results from implementing corrective actions to failure modes discovered in test prior to trial $t$. The series labeled Expected is generated via (14) using the true values of the beta parameters and $k$. The series labeled MLE K is also generated via (14) using the true value of $k$, but use the finite ML approximations for the beta parameters (i.e., column 3 of Table 10). The series labeled MLE ∞ is generated via (16) using the limiting ML approximations of the beta parameters (i.e., column 4 of Table 10).

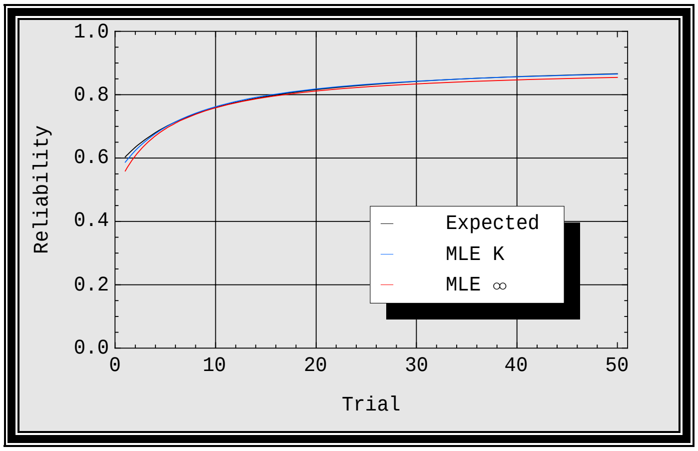


**Figure 19. Reliability Growth.**

The significance of this model equation is that it gives practitioners the capability to estimate the reliability of one-shot systems when observed failure modes are corrected at anytime during a typical developmental test program. In this example, the true initial reliability, given by (15), is $0.6035$. After correcting 7 of 25 failure modes with an average fix effectiveness of $\bar{d} = 0.80$, the true reliability has improved from 0.6035 to 0.8662. The reliability growth potential given by (18) is 0.9047. Notice that all the series are close to one another despite only having 7 datapoints from which to estimate model parameters. Also notice that our model is insensitive to not knowing the value of $k$. Thus, our limiting approximation as $k \to \infty$ is quite accurate despite the fact that the system is only comprised of a small number of failure modes (i.e., $k = 25$).

### 6.3.3. Expected Number of Failure Modes

Figure 20 displays a plot of the cumulative (observed and expected) number of failure modes versus trials. The curves can be interpreted as an estimate of the total number of unique failure modes expected to be observed on or before trial $t$. The series labeled Observed, plotted as black dots, represents the actual stochastically realized number of cumulative failure modes observed on trial $t$. The series labeled Exact is generated via (21) with the true values of the beta parameters and $k$. The series labeled MLE K is also generated by (21) with the true value of $k$, but uses the finite ML approximations for the beta parameters (e.g., column 3 of Table 10). The last series, MLE ∞, is generated via (22) using the limiting ML approximations of the beta parameters (e.g., column 4 of Table 10).

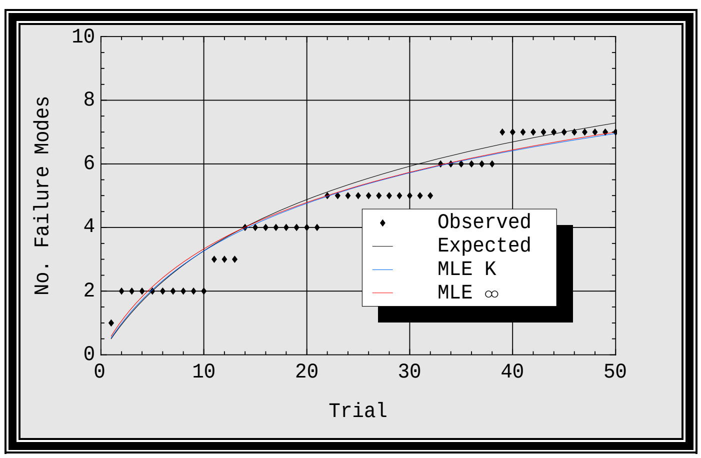


**Figure 20. Observed vs. Expected No. Failure Modes.**

One may ask the question: Why have we developed a model equation to estimate the expected number of failure modes observed during testing, when we already know how many failure modes were observed, namely $m = 7$? The answer is so we can construct a GOF procedure to determine if our model fits a given sample of data and can be suitably applied. Hence, the significance of this model equation is to give practitioners the means to assess model GOF. This is accomplished in two ways. First, a plot such as Figure 20 can provide graphical insight for such an assessment.

Good agreement between the actual observed number of failure modes versus trials, against the expected number of observed failure modes (e.g., MLE K, or MLE ∞) versus trials illustrates graphically that the model fits the data well and can be suitably applied. If, on the other hand, these two series are not in good agreement, the model may not fit the data. Due to the subjectivity of graphical assessments, we have derived a statistical GOF test procedure based on a chi-squared random variable. This test procedure is not presented in the scope of the current paper, but will follow in a later publication.

### 6.3.4. Expected Probability of a New Failure Mode

The curves in Figure 21 can be interpreted as an estimate of the expected probability on trial $t$ of observing a failure mode that has not been previously revealed. The series labeled Expected is generated via (24) using the true value of $k$ and the true values of the beta shape parameters. The series labeled MLE K is also generated via (24), and uses the finite ML approximations of the beta parameters (e.g., column 3 of Table 10) and the true value of $k$. The last series, MLE ∞, is generated via (26) using the limiting ML approximations of the beta parameters (e.g., column 4 of Table 10).

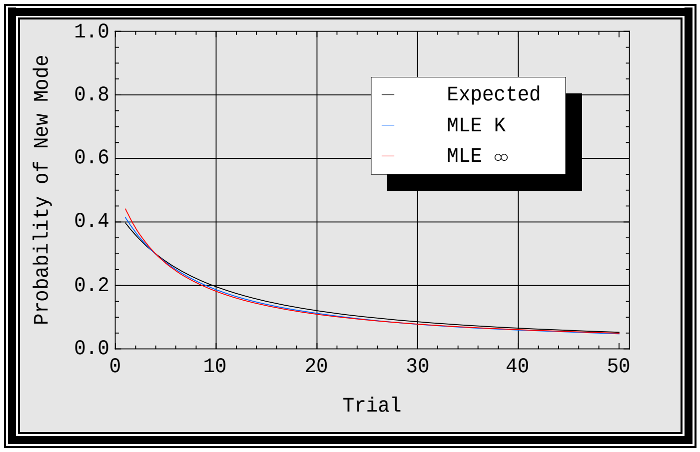


**Figure 21. Probability of a New Failure Mode.**

The significance of this model equation is that it can be utilized by management as a measure of programmatic risk. For example, as the development effort is ongoing, it is useful to gauge the likelihood of observing new problems, such as additional unknown failure modes. The prospect of finding new failure modes late in a developmental test program may translate into additional schedule and budgetary requirements associated with root-cause analysis and the correction action effort. Naturally, we would like these types of programmatic risks to diminish as the system design matures. This model equation serves as a management metric aimed at quantifying such a phenomenon.

### 6.3.5. Expected Probability of a Repeat Failure Mode

The curves in Figure 22 can be interpreted as the expected probability of observing a repeat failure mode on trial $t$ expressed as a fraction of initial system unreliability. The series labeled Expected is generated via (29) using the true value of $k$, as well as the true values of the beta shape parameters. The series labeled MLE K is also generated via (29) using the true value of $k$, but uses the finite MLE of the beta parameters (e.g., column 3 of Table 10). The series labeled MLE ∞ is generated via (30) using the limiting ML approximations (e.g., column 4 of Table 10).

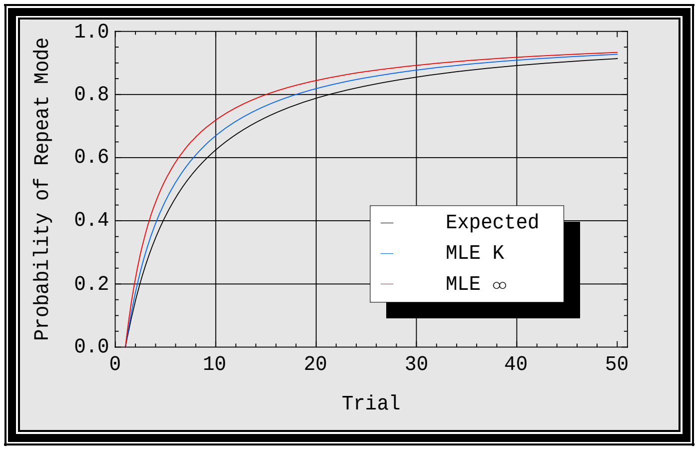


**Figure 22. Probability of a Repeat Failure Mode.**

The significance of this model equation is that it quantifies the portion of initial system unreliability associated only with known failure modes (i.e., the observed failure modes). The reason why this quantity is important is because prior to trial $t$ program management can eliminate at most the portion $\phi_k(t)$ from the initial system unreliability – regardless of when corrective actions are implemented or how effective they are. For example, the initial unreliability of the system in this example, given by (25), is $h_k(t=1)=0.3965$. On the first trial the portion of $h_k(t=1)$ associated with known failure modes is $\phi_k(t=1)=0$, as shown above. At the end of the current test phase, however, the portion of $h_k(t=1)$ associated with failure modes we know about is $\phi_k(t=50)=0.9137$. In other words, the 7 failure modes observed comprise about 91.4% of the initial probability of failure of the system. Consider trying to make a decision on whether to continue or terminate testing, especially in light of the fact that the purpose of a developmental test program is to find and eliminate design weaknesses. Given that the 7 failure modes we know about comprise 91.4% of system initial unreliability, continued testing may not be economically justifiable. The ultimate decision depends on the current reliability of the system relative to its requirement. Hence, a good management practice as part of a reliability growth program would be to monitor and specify goals for $\phi_k(t)$.

### 6.4. Ground-to-Air Missile Application

The trial numbers of first occurrence of $m = 16$ failure modes given in Table 11 were obtained in $T = 68$ flight tests of an unspecified ground-to-air missile system[46]. These failure modes were determined to have failed independently of one another during testing.

| | FOT $t_i$ | | FOT $t_i$ |
|---|---|---|---|
| **1** | 1 | **9** | 24 |
| **2** | 6 | **10** | 27 |
| **3** | 7 | **11** | 36 |
| **4** | 8 | **12** | 41 |
| **5** | 9 | **13** | 42 |
| **6** | 10 | **14** | 58 |
| **7** | 12 | **15** | 61 |
| **8** | 21 | **16** | 65 |

**Table 11. FOT.**

Using the data in Table 11, ML estimates of the beta shape parameters were calculated and are shown in Table 12. The finite $k$ estimates (given in column 2) are found as the numerical solutions that satisfy Equations (6) and (7) simultaneously for

[46] Details regarding the system and its failure mode information cannot be discussed due to proprietary reasons.

an assumed $k = 50$ total potential number of failure modes. The limiting approximation of the beta parameter $n$ (i.e., pseudo trials) given in column 3 is the solution to Equation (11). Recall via (8) that the limiting approximation of the beta parameter $x \to 0$ as $k \to \infty$.

| | MLE K | MLE ∞ |
|---|---|---|
| $n$ | 31.535 | 19.084 |
| $x$ | 0.331 | 0 |

**Table 12. Parameter Estimates.**

The ML estimates given in Table 12, were then used to generate the curves for each of the management metrics shown in Figures 23-26. Note that the blue series titled MLE K may be difficult to see in some of these plots, which is due to the close approximation between this series and the red series based on MLE ∞.

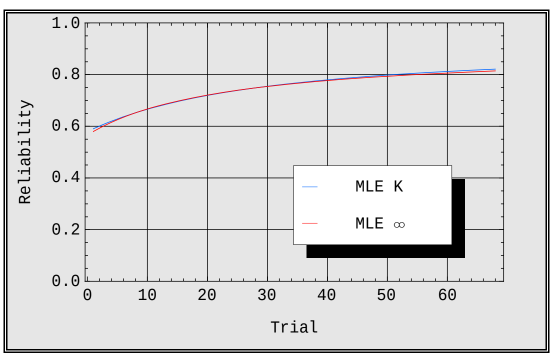


**Figure 23. Expected Reliability.**

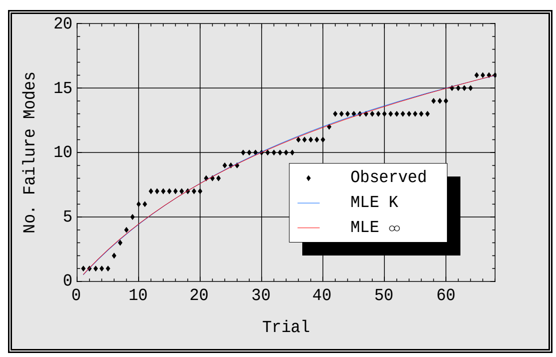


**Figure 24. Expected Failure Modes.**

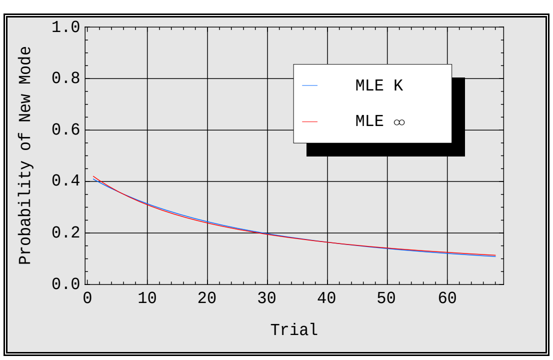


**Figure 25. Probability of New Mode.**

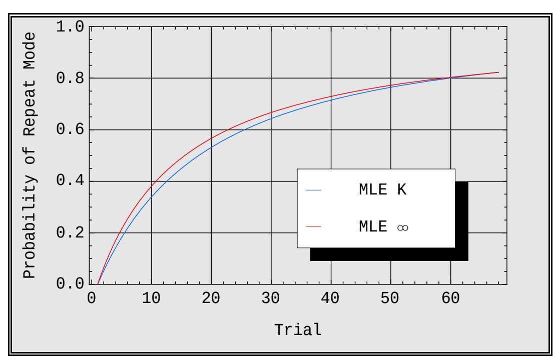


**Figure 26. Fraction Surfaced.**

Figure 23 shows that by correcting the 16 failure modes with an average FEF of $\bar{d} = 0.8$, system reliability can be increased from about 0.58 to 0.81. The reliability growth potential computed via Equation (19) is 0.90. Figure 24 shows that the model reasonably fits the data and yields and estimate of 16 failure modes by the end of the test. Figure 24 shows that the probability of discovering a new failure mode decreased from 0.42 to 0.11 by finding the 16 failure modes in these 68 trials. Figure 25 shows that the portion of the initial system probability of failure (i.e., 0.42) associated with failure modes program management found in these 68 trials has increased from 0 to 0.82. Thus, the 16 failure modes that were discovered account for 82% of the initial probability of failure of the system.

### 6.5. Concluding Remarks

In this paper, we have introduced a new methodology that serves as an analytical framework from which one-shot reliability growth programs can be assessed. The methodology consists of the following model equations that can be used as management metrics:

- Expected reliability growth of a one-shot system when corrective actions are implemented to prototypes at anytime after associated failure modes are first discovered. Our finite and limiting approximations are given by (14) and (16) , respectively.
- Expected number of failure modes observed in testing. Our finite and limiting approximations are given by (21) and (22), respectively.
- Expected probability of discovering a new failure mode. Our finite and limiting approximations are given by (24) and (26), respectively.

- Expected probability of a repeat failure mode expressed as a fraction of initial system unreliability. Our finite and limiting approximations are given by (29) and (30), respectively.

These model metrics provide the means by which reliability practitioners can estimate reliability improvement of a one-shot system, address model GOF concerns, quantify programmatic risk, and assess system maturity. To approximate these quantities, we have derived a new likelihood function (5) and associated ML procedures to estimate the shape parameters $n$ (i.e., pseudo trials) and $x$ (i.e., pseudo failures) of the beta distribution. The parameter MLE for an assumed number of total potential failure modes are obtained as the numerical solutions to (6) and (7). Since the number of total potential failure modes in a complex system is large and unknown, we have derived limiting approximations which have reduced the estimation procedure to solving only one equation (11), for one unknown. In particular, these approximations yield a mathematically convenient functional form (10) for the expected initial reliability of a one-shot system. The limiting behavior of our model equations (summarized in the Appendix) have led to interesting simplifications, and are shown to be a functions of only a single unknown, the beta shape parameter $n$.

## Appendix

| DESCRIPTION | MANAGEMENT METRICS |
|---|---|
| Initial Reliability | $\hat{R}_{\infty,I} = \exp\left(\frac{-m}{\hat{n}_\infty \cdot \left[\psi\left(\hat{n}_\infty + T\right) - \psi\left(\hat{n}_\infty\right)\right]}\right)$ |
| Projected Reliability | $\hat{R}_\infty\left(t\right) = \hat{R}_{\infty,I}^{1-\bar{d}\cdot\left(\frac{t-1}{\hat{n}_\infty + t - 1}\right)}$ |

| | |
|---|---|
| Reliability Growth Potential | $\hat{R}_{\infty,GP} = \hat{R}_{\infty,I}^{1-\bar{d}}$ |
| Number of Failure Modes | $\hat{\mu}_{\infty}(t) = \hat{n}_{\infty} \cdot \ln \hat{R}_{\infty,I} \cdot \left[\psi(\hat{n}_{\infty}) - \psi(\hat{n}_{\infty} + t)\right]$ |
| Probability of Occurrence of New Failure Modes | $\hat{h}_{\infty}(t) = 1 - \hat{R}_{\infty,I}^{\frac{\hat{n}_{\infty}}{\hat{n}_{\infty}+t-1}}$ |
| Fraction of System Initial Unreliability Associated with Repeat Failure Modes | $\hat{\phi}_{\infty}(t) = \dfrac{1 - \hat{R}_{\infty,I}^{\frac{t-1}{\hat{n}_{\infty}+t-1}}}{1 - \hat{R}_{\infty I}}$ |
| MLE Equation for the beta shape parameter $n$ | $\sum_{i=1}^{m}\left(\dfrac{1}{\hat{n}_{\infty} + t_i - 1}\right) + m\left[\dfrac{\psi'(\hat{n}_{\infty}+T) - \psi'(\hat{n}_{\infty})}{\psi(\hat{n}_{\infty}+T) - \psi(\hat{n}_{\infty})}\right] = 0$ |

**Acknowledgements**

Special thanks to Prof. Mosleh and Dr. Ellner for their advisory roles on my Ph.D. committee. The author is also grateful to Arthur Fries, Ph.D. (IDA), and Michael J. Cushing, Ph.D. (AEC) for their peer reviews which have improved initial versions of this manuscript.

## References

[1] H. K. Weiss, "Estimation of Reliability Growth in a Complex System with Poisson-Type Failure," *Operations Research*, vol. 4, no. 5, pp. 532-545, 1956.

[2] W. J. Corcoran, H. Weingarten and P. W. Zehna, "Estimating Reliability after Corrective Action," *Management Science*, vol. 10, no. 4, pp. 786-798, 1964.

[3] J. M. Finkelstein, "Confidence Bounds on the Parameters of the Weibull Process," *Technometrics*, vol. 18, no. 1, pp. 115-117, 1976.

[4] L. Lee, S. K. Lee, "Some Results on Inference for the Weibull Process," *Technometrics*, vol. 20, no. 1, pp. 41-45, 1978.

[5] M. Engelhardt, and L. J. Bain, "Prediction Intervals for the Weibull Process," *Technometrics*, vol. 20, no. 2, pp. 167-169, 1978.

[6] L. J. Bain, and M. Engelhardt, "Inferences on the Parameters and Current System Reliability for a Time Truncated Weibull Process," *Technometrics*, vol. 22, no. 3, pp. 421-426, 1980.

[7] L. Lee, "Comparing Rates of Several Independent Weibull Processes," *Technometrics*, vol. 22, no. 3, pp. 427-430, 1980.

[8] U.S. Department of Defense, "Reliability Growth Management," Washington, DC, Military Handbook 189, 1981.

[9] L. H. Crow, "An Improved Methodology for Reliability Growth Projection," AMSAA, Aberdeen Proving Ground, MD, Tech. Rep. TR-357, 1982.

[10] L. H. Crow, "Confidence Interval Procedures for the Weibull Process with Applications to Reliability Growth," *Technometrics*, vol. 24, no. 1, pp. 67-72, 1982.

[11] L. H. Crow, "Methods for Assessing Reliability Growth Potential," *IEEE Proc. RAMS*, 1984, pp. 484-489.

[12] H. F. Martz and R. A. Waller, "Selecting Beta Prior Distributions based on Observed Binomial Sampling Data," in *Bayesian Reliability Analysis*, 2nd ed., Malabar, FL: Krieger, 1991, pp. 312-318.

[13] P. M. Ellner and L. C. Wald, "AMSAA Maturity Projection Model," *IEEE Proc. RAMS*, 1995, pp. 174-181.

[14] A. Fries and A. Sen, "A Survey of Discrete Reliability Growth Models," *IEEE Trans. Reliab.*, vol. R-45, no. 4, pp. 582-604, 1996.

[15] P. M. Ellner, W. J. Broemm and W. J. Woodworth, "Reliability Growth Guide," AMSAA, Aberdeen Proving Ground, MD, Tech. Rep. TR-652, 2000.

[16] L.H. Crow, Methods for Reducing the Cost to Maintain a Fleet of Repairable Systems, *IEEE Proc. RAMS*, 2003, pp. 392-399.

[17] P. M. Ellner and J. B. Hall, "AMPM-Stein Model," AMSAA, Aberdeen Proving Ground, MD, Tech. Rep. TR-751, 2004.

[18] L. H. Crow, "An Extended Reliability Growth Model for Managing and Assessing Corrective Actions," *IEEE Proc. RAMS*, 2004, pp. 73-80.

[19] U.S. Department of Defense, "DoD Guide for Achieving Reliability, Availability and Maintainability," Washington, DC, 2005.

[20] D. E. Mortin, and P. M. Ellner, “Reliability Growth for Military Vehicles – Emerging Methodology and Remaining Challenges,” *Journal of Commercial Vehicles*, vol. 114, pp. 28-34, 2005.

[21] P.M. Ellner and J.B. Hall, “Planning Model based on Projection Methodology, AMSAA, Aberdeen Proving Ground, MD, Tech. Rep. TR-2006-9, 2006.

[22] J. B. Hall and A. M. Mosleh, “A Reliability Growth Projection Model for One-Shot Systems,” *IEEE Trans. Reliab*., vol. 57, no. 1, pp. 174-181, 2008.

[23] J. B. Hall and A. M. Mosleh, “An Analytical Framework for Reliability Growth Projection of One-Shot Systems,” *Reliability Engineering & System Safety*, doi:10.1016/j.ress.2007.11.003.

THIS PAGE IS INTENTIONALLY LEFT BLANK.

- CHAPTER 7 -

# 7. GOODNESS-OF-FIT AND CONFIDENCE INTERVAL PROCEDURES[47]


## Abstract

This paper gives Goodness-of-Fit (GOF) and Confidence Interval (CI) procedures for our reliability growth projection model published in an earlier paper. The first GOF technique is a graphical approach which compares the actual cumulative number of observed failure modes versus trials, to the expected number of observed failure modes versus trials. The second technique is a statistical GOF test procedure based on a chi-squared random variable. Both techniques are based on our exact expression for the expected number of observed failure modes, also derived herein. Maximum likelihood procedures are outlined for approximating this exact expression. A Fisher matrix normal approximation approach is employed using a log-odds transform to construct a CI estimate on expected reliability. Monte Carlo simulation is utilized to study the coverage associated with this approximate CI routine, as well as the approximating GOF test statistic. Numerical examples are given to illustrate the proposed GOF and CI procedures. This methodology is useful to practitioners who wish to address GOF concerns with our earlier reliability growth projection model, and/or obtain a CI estimate on the expected reliability of one-shot systems undergoing development.


[47] Chapter 7 will be submitted to a journal once Chapter 6 has been accepted with revisions by a journal (e.g., the material in Chapter 7 builds upon, and must refer to, the methodology given in Chapter 6).

**Keywords**: Confidence Interval, Discrete, Goodness-of-Fit, Monte Carlo, Projection, Reliability Growth.

## Acronyms[48]

**AEC** – Army Evaluation Center

**AMSAA** – Army Materiel Systems Analysis Activity

**CI** – Confidence Interval

**FEF** – Fix Effectiveness Factor(s)

**FOT** – First Occurrence Trial(s)

**GOF** – Goodness-of-Fit

**MGF** – Moment Generating Function

**ML** – Maximum Likelihood

**MLE** – ML Estimation/Estimate(s)

**MME** – Method of Moments Estimation/Estimate(s)

**NHPP** – Non-Homogeneous Poisson Process

**PLP** – Power-Law Process

## Definitions

1. **FEF** – fraction reduction in an initial failure mode probability due to implementation of a unique corrective action.
2. **Failure mode** – the root-cause associated with the loss of a required function or component whose probability, or rate, of occurrence is reduced by a specified

[48] The singular and plural of an acronym are always spelled the same.

FEF, if addressed by corrective action. Note that it may be the case that some failure modes are not observed during testing, or may not be corrected if they are observed (e.g., some failures may not be economically justifiable to correct).

3. **Unobserved failure mode** – a failure mode which exhibits zero failures during testing.
4. **Observed failure mode** – a failure mode which exhibits at least one failure during testing.
5. **Repeat failure mode** – a failure mode which exhibits at least two failures during testing.

## Notation

$k$ - total number of potential failure modes.

$m$ - total number of observed failure modes.

$n$ - beta shape parameter representing the pseudo number of trials.

$x$ - beta shape parameter representing the pseudo number of failures.

$T$ - total number of trials.

$N_{i,j}$ - number of failures for mode $i$ in trial $j$ – zero or unity.

$N_i$ - total number of failures for mode $i$ in $T$ trials.

$O_i$ - observed frequency for class interval $i$.

$E_i$ - expected frequency for class interval $i$.

$d_i$ - FEF for failure mode $i$.

$p_i$ - probability of occurrence for failure mode $i$.

$M_x^{(t)}(s)$ - $t$-th derivative of the beta MGF.

$\mu_k(t)$ - expected number of failure modes observed on or before trial $t$.

$\mu_\infty(t)$ - limiting approximation of $\mu_k(t)$ as $k \to \infty$.

$X^2$ - chi-squared test statistic.

$\alpha$ - significance level.

$c$ - number of class intervals in a chi-squared GOF table.

$X^2_{\alpha,c-2}$ - chi-squared critical point with $c-2$ degrees of freedom.

$R_k(t), R$ - expected reliability on trial $t$.

## 7.1. Introduction

### 7.1.1. Background Material

#### 7.1.1.1. Reliability Growth

In general, the term reliability growth refers to the increase in the true but unknown reliability of a developmental item that is achieved by finding, analyzing, and effectively correcting failure modes inherent to initial or intermediate system prototypes. Each of the three main areas of reliability growth (i.e., planning, tracking, and projection) apply to systems whose usage is measured in terms of discrete trials, as well as in the continuous time domain. The focus of this paper lies in the area of discrete projection. Applications involve estimating the reliability of one-shot systems[49] that could be achieved if known failure modes (discovered during testing) are mitigated via the corrective action process. The impact of corrective actions, with respect to increasing reliability, are quantified via FEF and are typically assigned by a Failure Prevention and Review Board via expert engineering judgment.

There are two types of models in the area of discrete reliability growth projection. The first type addresses the case were all corrective actions are delayed until the end of the current test phase. The second type address the more common scenario where corrective actions are applied to system prototypes anytime after associated failure modes are first discovered. The genesis of discrete reliability growth projection is marked by a paper written by Corcoran, Weingarten, and Zehna in 1964 [2], which addresses the delayed corrective action strategy. Since then, a number of other methods have been developed. They include the delayed models [10], [33], [37], and [38], and the non-delayed models [21], [32], and [39].

A vast amount of literature is available on each of the three areas of reliability growth. The most comprehensive presentation of discrete tracking and projection models is given by Fries and Sen's survey [22]. More general references include the AMSAA Reliability Growth Guide [28], Appendix C of the OSD Guide for Achieving RAM [35], and the frequently referenced Military Handbook 189 [7].

#### 7.1.1.2. GOF Methods for Reliability Growth

[49] One-shot systems such as guns, rockets, missile systems, torpedoes etc.

As in nearly any field of study, model validation is a basic necessity in reliability growth assessment. There have been several approaches published over the years for addressing GOF concerns with various reliability growth tracking and projection models. Most of these procedures have been centered around the NHPP interpretation [3] of the Duane model [1]. This is due to the wide use of the PLP[50] in reliability growth and complex, repairable systems theory. Some of the associated approaches in the literature include Crow [4], Rigdon [13], Park & Kim [16], Klefsjö & Kumar [17], Park & Seoh [20], Baker [23], Crétois & Gaudoin [24], Gaudoin [25], Crétois, El Aroui, & O. Gaudoin [26], and Gaudoin, Yang, & Xie [31]. Convenient graphical GOF explorations of the NHPP are discussed by Xie & Zhao [19], and by Donovan & Murphy [27]. Not all reliability growth models, however, are based on the PLP (e.g., the AMSAA Maturity Projection Models [21] and [33] are based on the gamma-Poisson relationship[51]), As a result, these models are accompanied by their own unique GOF procedures discussed by Broemm in [34]. The approach developed herein can be utilized for studying the conformity of our reliability growth projection models [37-39] against a given sample of discrete data. The procedures are not based on the PLP, but lend themselves to analogous graphical explorations.

#### 7.1.1.3. CI Procedures for Reliability Growth

---

[50] NHPP with failure intensity function $r(t) = \lambda \cdot \beta \cdot t^{\beta-1}$. This is also sometimes referred to as the PLP, or the Weibull process since the time to first failure follows the Weibull distribution.

[51] By gamma-Poisson relationship we are referring to the doubly-stochastic process where the distribution of the number of failures for each failure mode $i = 1, \ldots, k$ is assumed Poisson with gamma initial failure rates $\lambda_1, \ldots, \lambda_k$.

Similarly, a great deal of work has been done on CI construction for reliability growth. Finkelstein [5] developed confidence bounds on the parameters $\lambda$ and $\beta$ of the Weibull process. Crow gave small sample and asymptotic confidence intervals for system MTBF under Type I and Type II (i.e., time and failure) censoring in [3-4], [6], and [9]. These procedures were later modified in [18] to address the case were the data is generated from multiple systems operated over the same test interval $[0,T]$. Bhattacharyya, Fries, & Johnson [14] derived large-sample standard-error formulas and normal approximation CI procedures for the parameters of a discrete analogue to the PLP. Robinson & Dietrich [15] built a nonparametric-Bayes reliability growth tracking model (continuous time domain) from which Bayesian probability limits can be obtained on system failure intensity at the end of developmental testing. Ellner et al. [28] adapted the Lindström-Madden method to compute an approximate LCB on system MTBF from subsystem data (continuous time domain). Pulcini [30] developed exact and approximate CI procedures for current failure intensity, and interval prediction of current lifetime given by an exponential reliability growth model under a multiple system test program.

The above literature represents only a handful of approaches for obtaining interval estimates in reliability growth applications. Since small sample analytical results cannot always be derived, exact CI procedures may not always exist for certain models. This has led to the use of normal approximation approaches (e.g., see Nelson [8]) to be widely employed in the fields of reliability and reliability growth. The impact of data censoring on the coverage of an approximate CI procedure is particularly important. Comparisons of approximate CI procedures under Type I censoring is given by Jeng and Meeker in [29]. Hong, Meeker, and Escobar [36] also discuss how to avoid problems with normal approximation interval estimates for probabilities. Later in this paper, we use the popular Fisher matrix normal approximation approach in Nelson [8], which is widely used in reliability growth and associated commercial software products. Our procedure takes into account a number of features to include: the use of a log-odds transform to guarantee interval estimates in $\left[0,1\right]$, the use of a multi-parameter distribution, and consideration to data-censoring. The coverage probability of the approximation routine is investigated via Monte Carlo simulation.

### 7.1.2. Overview

This paper is organized as follows. The methodology of our approach is presented in Section 2 and consists of: a list of model assumptions, the derivation of our exact expression for the expected number of observed failure modes, an outline of ML procedures to estimate model parameters, graphical & statistical test procedures for assessing GOF, and CI construction on expected reliability. Section 3 illustrates these techniques via numerical example. Monte Carlo simulation results addressing the coverage probability of the CI procedures, as well as the approximating GOF test statistic are also presented. Concluding remarks are given in Section IV.

## 7.2. Methodology

### 7.2.1. Model Assumptions

1. A trial results in a dichotomous success/failure outcome such that $N_{i,j} \sim Bernoulli\left(p_i\right)$ for each $i=1,\ldots,k$ and $j=1,\ldots,T$.
2. The distribution of the number of failures in $T$ trials for each failure mode is binomial. That is, $N_i \sim Binomial\left(T, p_i\right)$ for each $i=1,\ldots,k$.
3. Initial mode failure probabilities $p_1,\ldots,p_k$ constitute a realization of a s-random sample $P_1,\ldots,P_k$ such that $P_i \sim Beta\left(n,x\right)$ for each $i=1,\ldots,k$.
4. Failures associated with different failure modes arise s-independently of one another on each trial. As a result, the system must be at a stage in development where catastrophic failure modes have been previously discovered and corrected and are therefore not preventing the occurrence of other failure modes.

### 7.2.2. Estimation Procedures

In [39] we derived a likelihood function and associated ML procedure to estimate model parameters (i.e., the shape parameters of the beta distribution) under a non-delayed corrective action strategy. These procedures are briefly outlined here. Let $obs \equiv \left\{i : N_i > 0 \text{ for } i\text{=}1,\ldots,k\right\}$ represent the index set of failure modes observed during testing, and let $obs' \equiv \left\{j : N_j = 0 \text{ for } j\text{=}1,\ldots,k\right\}$ denote its complement. Also let the total number of observed failure modes be denoted by $m = \left|obs\right|$, which gives

$|obs'| = k - m$ unobserved failure modes (i.e., right-censored, or suspended, observations). Our likelihood function is given by,

$$L_k\left(m,\vec{t} \mid n,x\right) = \frac{k!}{(k-m)!} \cdot \left[\frac{\Gamma(n)\cdot\Gamma(n-x+T)}{\Gamma(n-x)\cdot\Gamma(n+T)}\right]^{k-m} \cdot \prod_{i=1}^{m}\left[\frac{\Gamma(x+1)\cdot\Gamma(n-x+t_i-1)}{\mathrm{B}(x,n-x)\cdot\Gamma(n+t_i)}\right] \quad (1)$$

where $\mathrm{B}(a,b) \equiv \frac{\Gamma(a)\cdot\Gamma(b)}{\Gamma(a+b)} = \int_0^1 t^{a-1}\cdot(1-t)^{b-1}\,dt$ is the Euler beta function. The middle product-term in (1) represents the $k-m$ Type I (i.e., time) censored observations (i.e., the failure modes not observed on or before trial $T$). The partial derivatives of the natural logarithm of (1) with respect to the beta shape parameters $n$ and $x$ yield the following MLE equations respectively:

$$\begin{aligned}\frac{\partial \ln L_k\left(m,\vec{t}\right)}{\partial n} = 0 \Rightarrow k\cdot\left[\psi(n)-\psi(n-x)+\psi(n-x+T)-\psi(n+T)\right] &= \\ = \sum_{i=1}^{m}\left[\psi(n+t_i)-\psi(n-x+t_i-1)+\psi(n-x+T)-\psi(n+T)\right]&\end{aligned} \quad (2)$$

and

$$\begin{aligned}\frac{\partial \ln L_k\left(m,\vec{t}\right)}{\partial x} = 0 \Rightarrow k\cdot\left[\psi(n-x)-\psi(n-x+T)\right] &= \\ = \sum_{i=1}^{m}\left[\psi(n-x+t_i-1)-\frac{1}{x}-\psi(n-x+T)\right]&\end{aligned} \quad (3)$$

Equations (2) and (3), when maximized simultaneously, yield the ML estimates $\hat{n}$ and $\hat{x}$ that maximize the marginal likelihood function. Notice that these equations are a function of the unknown variable $k$, the total potential number of failure modes in the system. Thus, one must assume such a value when using this estimation procedure. To avoid this, and to assess the sensitivity of not knowing $k$ we have derived limiting approximations of the model parameters. The equations to obtain these estimates are derived by taking the limit of (1) as $k \to \infty$, followed by evaluating the partial derivatives. After some detailed analysis we have found,

$$\hat{R}_{\infty,I} = \exp\left(\frac{-m}{\hat{n}_\infty \cdot \left[\psi\left(\hat{n}_\infty + T\right) - \psi\left(\hat{n}_\infty\right)\right]}\right) \tag{4}$$

where $\hat{R}_{\infty,I}$ is a limiting approximation of initial system reliability, and where $\hat{n}_\infty$ is found as the numerical solution to

$$\sum_{i=1}^{m}\left(\frac{1}{\hat{n}_\infty + t_i - 1}\right) + m\left[\frac{\psi'\left(\hat{n}_\infty + T\right) - \psi'\left(\hat{n}_\infty\right)}{\psi\left(\hat{n}_\infty + T\right) - \psi\left(\hat{n}_\infty\right)}\right] = 0 \tag{5}$$

It can be shown that finite $k$ estimators $\hat{n}$ and $\hat{x}$ that satisfy Equations (2) and (3) converge to $\hat{n}_\infty$ and zero, respectively, as $k \to \infty$. Via heuristics in Monte Carlo simulation we have found $k \geq 5 \cdot m$ to be a good rule-of-thumb for the choice of $k$ when using the finite estimation procedure. In general, however, prior work [39] has shown that approximations of the management metrics of interest (e.g., expected reliability) are not sensitive to the value of $k$ provided it is chosen to be sufficiently large, as indicated by the rule-of-thumb criterion above. This means that the limiting approximations as $k \to \infty$ are not much different in magnitude than those based on the true value of $k$ (even for small values such as $k = 25$). Hence, the reason why we do not bother deriving an estimate of $k$. Finally, note that the use of the failure mode first occurrence data allows this assessment procedure to be used independently of the corrective action strategy. However, additional estimation procedures based on count data (i.e., number of failures for individual failure modes) [37] are available in the case where all corrective actions are delayed until the end of the current test phase. These alternate procedures can be used to estimate model parameters in a similar fashion under a delayed corrective action strategy.

### 7.2.3. GOF Procedures

#### 7.2.3.1. Expected Number of Failure Modes

The GOF of our model can be graphically studied by comparing the actual cumulative number of observed failure modes versus trials, against the expected number of failure modes given by our model. GOF can also be examined via a statistical test. Both of these approaches depend on our logically derived exact expression for the expected number of failure modes observed on or before trial $t$. To develop this expression, let $I_i(t)$ denote the indicator function such that

$$I_i(t) \equiv \begin{cases} 1 & \text{if failure mode } i \text{ is observed on or before trial } t \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

From (5), the true number of unique failure modes observed on or before trial $t$ is,

$$m(t) \equiv \sum_{i=1}^{k} I_i(t) \tag{7}$$

where $k$ is the true but unknown total potential number of failure modes in the system. Assuming that trials are statistically independent and that the $p_i$ for $i = 1,...,k$ are constant, the mathematical expectation of (5) is

$$E\left[I_i(t)\right] = 1 - (1 - p_i)^t \tag{8}$$

Let $\bar{p} \equiv (p_1, \ldots, p_k)$ denote the vector of the true (but unknown) failure probabilities inherent in the system. From (7) and (7) the conditional expected number of failure modes (i.e., conditioned on $\bar{p}$) observed on or before trial *t* is,

$$\mu_k(t \mid \bar{p}) \equiv E[m(t)] = \sum_{i=1}^{k} E[I_i(t)] = \sum_{i=1}^{k} \left[1 - (1 - p_i)^t\right] = k - \sum_{i=1}^{k} (1 - p_i)^t \tag{9}$$

This expression has the following convenient interpretation: the expected number of failure modes observed on or before trial $t$ is equivalent to the total potential number of failure modes in the system minus the expected number of failure modes that will not be observed in $t$ trials. Notice the initial condition of (13) suggests that the expected number of failure modes on trial $t = 0$ (i.e., before testing begins) is $\mu_k(t = 0 \mid \bar{p}) = 0$, as expected.

To assess the mathematical expectation of (13) with respect to the $p_i$ recall that the *t*-th moment of a beta r.v. $X$ is defined as

$$E\left[X^t\right] \equiv M_x^{(t)}(s)\Big|_{s=0} \tag{10}$$

where

$$M_x^{(t)}(s) = \sum_{i=t}^{\infty} \left[\prod_{j=0}^{i-1} \left(\frac{n - x + j}{n + j}\right)\right] \cdot \left[\frac{s^{i-t}}{(i-t)!}\right] \tag{11}$$

is the *t*-th derivative of the beta MGF. By interpreting the failure probabilities in (13) as iid beta r.v., we have derived the unconditional expected number of failure modes observed on or before trial *t*. The motivation for doing this is to express (13) in terms of the beta shape parameters, rather than the $p_i$ for $i = 1, \ldots, k$ (e.g., to obtain a marginal expression). From (10) and (11) we have found

$$E\left[\left(1-P_i\right)^t\right] \equiv M_x^{(t)}(0) = \frac{\Gamma(n)\cdot\Gamma(n-x+t)}{\Gamma(n-x)\cdot\Gamma(n+t)} \tag{12}$$

where $\Gamma(x) \equiv \int_0^\infty t^{x-1}\cdot e^{-t}dt$ denotes the Euler gamma function. From (13) and (12) the true unconditional expected number of unique failure modes observed on or before trial $t$ becomes

$$\mu_k(t) \equiv E\left[\mu_k\left(t \mid \vec{P}\right)\right] = k - \sum_{i=1}^{k} E\left[\left(1-P_i\right)^t\right] = k\cdot\left(1-\frac{\Gamma(n)\cdot\Gamma(n-x+t)}{\Gamma(n-x)\cdot\Gamma(n+t)}\right) \tag{13}$$

Notice that (13) is only a function of three unknowns, namely, $k$ and the two beta shape parameters $n$ (i.e., pseudo trials) and $x$ (i.e., pseudo failures). Since $k$ is typically large and unknown, it is desirable to derive the limiting behavior of (13) as $k \to \infty$. After reparameterizing (13) with $x = n\cdot\left(1-R_{k,I}^{1/k}\right)$ (e.g., see [37]), we find its limiting behavior to be,

$$\hat{\mu}_\infty(t) \equiv \lim_{k\to\infty} \hat{\mu}_k(t) = m\cdot\left[\frac{\psi(\hat{n}_\infty + t) - \psi(\hat{n}_\infty)}{\psi(\hat{n}_\infty + T) - \psi(\hat{n}_\infty)}\right] \tag{14}$$

where $\psi(x) \equiv \frac{\Gamma'(x)}{\Gamma(x)}$ denotes the digamma function, and $\hat{n}_\infty$ satisfies (11). Note that Equations (13) and (14) are mean-value functions that can be utilized to estimate the expected number of unique failure modes observed on or before trial $t$. Thus, they are comparable to $\mu(t) = \lambda\cdot t^\beta$ of the typical PLP approach.

#### 7.2.3.2. Hypothesis Test

To assess model GOF via a statistical test, one can use the common chi-squared approach [12] with our expected frequency function. The null hypothesis, $H_0$, is the conjecture that our model fits the data and can be suitably applied, whereas the alternative hypothesis, denoted by $H_a$, expresses the contrary. The test statistic is calculated in the usual manner,

$$X^2 \equiv \sum_{i=1}^{c} \left[ \frac{\left(O_i - \hat{E}_i\right)^2}{\hat{E}_i} \right] \tag{15}$$

where $c$ is the number of cells in the GOF table . The observed frequencies, $O_i$, represent the total number of new failure modes observed in class interval $i = 1, \ldots, c$. Using (14) (without loss of generality (13) could also be used), the expected frequencies in (15) are calculated by

$$\hat{E}_i \equiv m \cdot \left[ \frac{\widehat{\mu}_\infty\left(y_{i+1}\right) - \widehat{\mu}_\infty\left(y_i\right)}{\widehat{\mu}_\infty\left(T\right)} \right] = m \cdot \left[ \frac{\psi\left(\widehat{n}_\infty + y_{i+1}\right) - \psi\left(\widehat{n}_\infty + y_i\right)}{\psi\left(\widehat{n}_\infty + T\right) - \psi\left(\widehat{n}_\infty\right)} \right] \tag{16}$$

where $y_i$ for $i = 1, \ldots, c+1$ are the endpoints of the class intervals chosen by the practitioner (note that $y_1 = 0$). The right-hand product term in (16) represents the class probability (i.e., the probability of being in a given class interval). Thus, the condition $\frac{1}{m} \cdot \sum_{i=1}^{c} \hat{E}_i = 1$ holds. Given the high test cost of some one-shot systems, the number of observed failure modes in many applications will be limited. Thus, the total number of trials should be divided into class intervals such that the expected frequencies $E_i \geq 2$ for $i = 1, \ldots, c$ and $c \geq 3$ at a bare minimum.

The rejection region is any value of the test statistic (15) greater than the critical point. The critical point (obtained by table-lookup) is denoted by $X^2_{\alpha,c-2}$ with $\alpha$ significance level and $c-2$ degrees of freedom. A degree of freedom is lost for each estimated model parameter. In this case, the first degree of freedom is lost by using $m$, the total number of observed failure modes. This means that only $c-1$ observed cell frequencies in the GOF table are uniquely determined. The second degree of freedom is lost by estimating the beta shape parameter $n$ (e.g., pseudo trials). Hence, the critical point is based on $c-2$ degrees of freedom. Clearly, the null hypothesis, $H_0$, is rejected if $X^2 \geq X^2_{\alpha,c-2}$. In this case, there is statistical evidence at the $\alpha$ level of significance that the model does not fit the data. Failure to reject $H_0$ occurs when $X^2 < X^2_{\alpha,c-2}$, indicating that there is no statistical evidence against the model. Either way, the graphical method discussed above illustrates the associated correlation (e.g., high, or lack thereof) between (13) and/or (14) in comparison to the actual cumulative number of observed failure modes versus trials. A numerical example is given below for a dataset comprised of only a small number of (i.e., $m=7$) observed failure modes.

### 7.2.4. CI Procedures

We shall now derive an approximate CI estimate on the expected reliability of a one-shot system [39] given by,

$$R_k(t) = \left[1 - \left(1 - \left[1 - \frac{\Gamma(n-x+t-1)\cdot\Gamma(n+1)}{\Gamma(n-x)\cdot\Gamma(n+t)}\right]\cdot\bar{d}\right)\cdot\frac{x}{n}\right]^k \tag{17}$$

where $\bar{d} \equiv \sum_{i \in obs} d_i \Big/ m$ is an average FEF. For simplicity, let the expression in (17) be denoted by $R$. To derive an approximate CI on $R$, we utilize the Fisher matrix normal approximation approach discussed by Nelson in [8]. Since $R \in (0,1)$ is not consistent with the domain of a normal r.v., we proceed by developing a CI on the monotone increasing and differentiable log-odds transform of $R$ given by,

$$h \equiv \ln\left(\frac{R}{1-R}\right) \tag{18}$$

which ensures that $h \in (-\infty, \infty)$. Thus, for large sample sizes the cumulative distribution of $\hat{h}$ (i.e. ML estimate of (18)) is approximately normal with mean $h$ and standard deviation

$$s\left(\hat{h}\right) = \sqrt{\hat{H}^T \cdot \hat{F}_0^{-1} \cdot \hat{H}} \tag{19}$$

Let the ML estimate of (17) be denoted by $\hat{R}$. In (19), $\hat{H}$ is the column vector of partial derivatives of the log-odds transform (18) w.r.t. the beta parameters given as,

$$\hat{H} \equiv \begin{bmatrix} \dfrac{\partial \hat{h}}{\partial n} \\[2ex] \dfrac{\partial \hat{h}}{\partial x} \end{bmatrix} = \begin{bmatrix} \dfrac{\hat{R}'_n}{\hat{R} \cdot \left(1-\hat{R}\right)} \\[2ex] \dfrac{\hat{R}'_x}{\hat{R} \cdot \left(1-\hat{R}\right)} \end{bmatrix} \tag{20}$$

(the partial derivatives $\hat{R}'_n$ and $\hat{R}'_x$ are given in the Appendix) and $\hat{F}_0^{-1}$ is an estimate of the inverse of the true theoretical Fisher information matrix (i.e., the true asymptotic covariance matrix of the ML estimators $\hat{n}$ and $\hat{x}$). $F_0$ is given as,

$$F_0 = \begin{bmatrix} E_0\left[\dfrac{-\partial^2 L(n,x)}{\partial n^2}\right] & E_0\left[\dfrac{-\partial^2 L(n,x)}{\partial n \cdot \partial x}\right] \\ E_0\left[\dfrac{-\partial^2 L(n,x)}{\partial x \cdot \partial n}\right] & E_0\left[\dfrac{-\partial^2 L(n,x)}{\partial x^2}\right] \end{bmatrix} \tag{21}$$

The expectations in (21) are calculated as,

$$E_0\left[\frac{-\partial^2 L(n,x)}{\partial n^2}\right] \equiv$$
$$= \sum_{i \in obs} \sum_{t_i=1}^{T} \left[\frac{-\partial^2 \ln f(t_i \mid n,x)}{\partial n^2} \cdot f(t_i \mid n,x)\right] + \sum_{j \in obs'} \left[\frac{-\partial^2 \ln R(T \mid n,x)}{\partial n^2} \cdot R(T \mid n,x)\right]$$
$$= m \cdot \sum_{t_i=1}^{T} \left[\frac{-\partial^2 \ln f(t_i \mid n,x)}{\partial n^2} \cdot f(t_i \mid n,x)\right] + (k-m) \cdot \left[\frac{-\partial^2 \ln R(T \mid n,x)}{\partial n^2} \cdot R(T \mid n,x)\right] \tag{22}$$

$$E_0\left[\frac{-\partial^2 L(n,x)}{\partial x^2}\right] \equiv$$
$$= \sum_{i \in obs} \sum_{t_i=1}^{T} \left[\frac{-\partial^2 \ln f(t_i \mid n,x)}{\partial x^2} \cdot f(t_i \mid n,x)\right] + \sum_{j \in obs'} \left[\frac{-\partial^2 \ln R(T \mid n,x)}{\partial x^2} \cdot R(T \mid n,x)\right]$$
$$= m \cdot \sum_{t_i=1}^{T} \left[\frac{-\partial^2 \ln f(t_i \mid n,x)}{\partial x^2} \cdot f(t_i \mid n,x)\right] + (k-m) \cdot \left[\frac{-\partial^2 \ln R(T \mid n,x)}{\partial x^2} \cdot R(T \mid n,x)\right] \tag{23}$$

and

$$E_0\left[\frac{-\partial^2 L(n,x)}{\partial n \partial x}\right] = E_0\left[\frac{-\partial^2 L(n,x)}{\partial x \partial n}\right] \equiv$$
$$= \sum_{i \in obs} \sum_{t_i=1}^{T} \left[\frac{-\partial^2 \ln f(t_i \mid n,x)}{\partial n \partial x} \cdot f(t_i \mid n,x)\right] + \sum_{j \in obs'} \left[\frac{-\partial^2 \ln R(T \mid n,x)}{\partial n \partial x} \cdot R(T \mid n,x)\right]$$
$$= m \cdot \sum_{t_i=1}^{T} \left[\frac{-\partial^2 \ln f(t_i \mid n,x)}{\partial n \partial x} \cdot f(t_i \mid n,x)\right] + (k-m) \cdot \left[\frac{-\partial^2 \ln R(T \mid n,x)}{\partial n \partial x} \cdot R(T \mid n,x)\right] \tag{24}$$

where,

$$f\left(t_i \mid n,x\right) \equiv \frac{\Gamma\left(x+1\right)\cdot\Gamma\left(n-x+t_i-1\right)}{\mathrm{B}\left(x,n-x\right)\cdot\Gamma\left(n+t_i\right)} \tag{25}$$

is the marginal beta-geometric density for the $m$ observed failure modes (i.e., the complete observations) and,

$$R\left(T \mid n,x\right) \equiv \frac{\Gamma\left(n\right)\cdot\Gamma\left(n-x+T\right)}{\Gamma\left(n-x\right)\cdot\Gamma\left(n+T\right)} \tag{26}$$

is the marginal beta-geometric reliability function for the $k-m$ Type I censored observations. After the detailed Fisher analysis required to obtain the estimate of the standard deviation, the desired $100\cdot\gamma\%$ CI on expected system reliability is given by,

$$\left[\underset{\sim}{R},\tilde{R}\right] = \frac{1}{1+\left(\frac{1-\hat{R}}{\hat{R}}\right)\cdot\exp\left[\pm z_\gamma\cdot s\left(\hat{h}\right)\right]} \tag{27}$$

where $\hat{R}$ is an ML estimate of (17), $z_\gamma$ is the $100\cdot\left(1-\gamma\right)/2$ percentile of the standard normal distribution, and $s\left(\hat{h}\right)$ is given by (19).

## 7.3. Numerical Examples & Simulation

### 7.3.1. GOF Procedures

#### 7.3.1.1. Graphical Method

Since the true failure probabilities, beta parameters, and other quantities of interest are unknown in practice, Monte Carlo simulation is used to generate the following numerical example. The simulation performed 50 trials of a one-shot system comprised of 25 failure modes. The probabilities of occurrence $p_1,\ldots,p_k$ for each failure mode were stochastically generated via a beta r.v. with a s-mean of $\mu = 0.0200$ and a s-variance of $\sigma^2 = 0.0020$. The number of failures for each failure mode per trial were stochastically generated from a Bernoulli r.v. with parameter $p_i$ for $i = 1,\ldots,k$. During simulation, only 7 of 25 failure modes were observed with a total of 23 failures. The statistical estimation procedures presented above (e.g., to obtain ML estimates of the beta shape parameters) only require and use the 7 FOT, denoted by $t_i$ for $i \in obs$. In this example the FOT are $\vec{t} = \{1,2,11,14,22,33,39\}$. These ML estimates in addition to their limiting approximations as $k \to \infty$ are given in Table 13. The column titled True represents the true values of the beta parameters utilized to stochastically generate the failure probabilities $p_1,\ldots,p_k$. The column labeled MLE K shows the ML estimates obtained as the solutions to Equations (2) and (3). The column titled MLE $\infty$ denotes our limiting approximations of the beta parameters. The limit of the MLE for the beta parameter $n$ is found as the solution to (11). The MLE of $x$ converges to zero as $k \to \infty$.

| | True | MLE K | MLE $\infty$ |
|---|---|---|---|
| $n$ | 8.800 | 7.193 | 4.468 |
| $x$ | 0.176 | 0.152 | 0 |

**Table 13. Beta Parameters.**

The parameter estimates shown in Table 13 are utilized to approximate Equations (13) and (14) above. These curves, shown in Figure 27, provide a graphical means from which the GOF of our model can be assessed. The actual cumulative number of observed failure modes (i.e., the stochastic realization) are represented by black dots, and correspond to the series titled Observed. The smooth black series, labeled Expected, is generated by (13) with the true values of $k$, as well as the true values of the beta parameters (e.g., column 1 of Table 13). This series represents the true expected number of failure modes observed on or before trial $t$ given by our model with no parameters estimated. The series labeled MLE K is also generated via (13) with the true value of $k$, but uses the ML estimates shown in column 2 of Table 13. The sensitivity of not knowing $k$ is illustrated by the series titled MLE ∞, which is generated by (14) with the limiting approximations of the beta parameters given in column 3 of Table 13. The high correlation of our approximations in comparison to the actual observed number of failure modes, shown in Figure 27, is a graphical indication that our model reasonably fits the data. Hence, the associated reliability growth management metrics in [39] can be suitably applied.

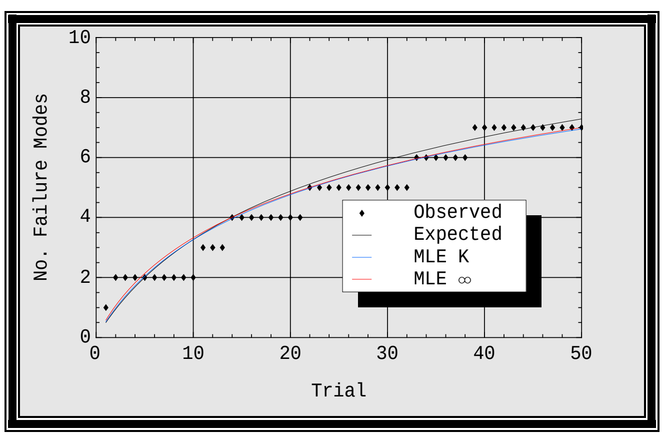


**Figure 27. Observed / Expected Failure Modes vs. Trials.**

#### 7.3.1.2. Hypothesis Test

Using the procedure outlined above, the chi-squared GOF table is constructed and shown in Table 14. The class intervals are divided such that the expected frequencies in each cell $\hat{E}_i \geq 2$ and are approximately equal in magnitude. The class intervals and associated number of trials are shown in the second and third columns. The observed frequency represents the actual number of new failure modes discovered in the each class interval. For example, there were 2 failure modes discovered in the first 5 trials. The expected frequencies are computed from our model based on (16) with $m = 7$ and $\bar{y} = \{0, 5, 15, 50\}$. The last column shows the terms of the chi-squared GOF test statistic for each class interval. The value of the test statistic is $X^2 \approx 0.0662$.

| | **Class Interval** | **No. Trials in Class** | $O_i$ **Observed Frequency** | $E_i$ **Expected Frequency** | $\frac{(O_i - E_i)^2}{E_i}$ |
|---|---|---|---|---|---|
| **1** | 1-5 | 5 | 2 | 2.1833 | 0.0154 |
| **2** | 6-15 | 11 | 2 | 2.1478 | 0.0102 |
| **3** | 16-50 | 34 | 3 | 2.6689 | 0.0411 |
| | | **Total:** | **7** | **7** | **0.0667** |

**Table 14. GOF Table.**

Using a significance level of $\alpha = 0.20$, the one-sided upper rejection region is defined as any value of the test statistic such that $X^2 \geq X^2_{0.20,1} \approx 1.6424$. Since the test statistic is not in the rejection region, we fail to reject $H_0$ (i.e., fail to reject that the model fits the data).

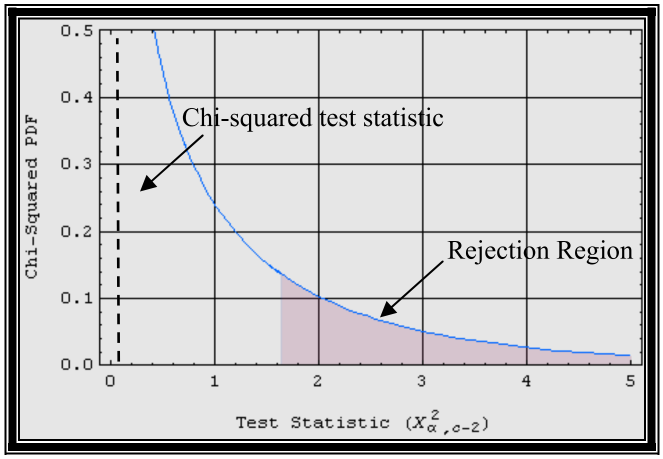


**Figure 28. Chi-Squared Test Statistic.**

### 7.3.1.3. Monte Carlo Simulation

The GOF procedure presented above assumes that the test statistic (15) follows the chi-squared distribution. As a means to investigate the assumed approximation, we have utilized Monte Carlo simulation to construct an empirical distribution for the test statistic to compare against the chi-squared PDF for the appropriate degrees of freedom. The simulation consists of the following steps:

1. Simulation inputs. Due to high costs of missile and other types of one-shot systems, the number of trials conducted and failure data obtained from developmental testing can be very limited. As a result, our study simulates a very conservative scenario with respect to the number of trials conducted and failure data available for estimation purposes. This also means that most applications of these test procedures will require 3 or 4 cells to construct the GOF table (since the expected frequencies will not be large in magnitude). Thus, this study focuses on the approximation for a chi-squared distribution based on 1 and 2 degrees of freedom. Simulation inputs include: $T = 30$ trials

(per replication), and $k = 50$ total potential number of failure modes. The mean and variance of the failure probabilities used in the study include $\mu \in (0.0165, 0.02875)$ (adjusted to increase the expected frequencies when increasing the number of cells in the GOF table) and $\sigma^2 = 0.001$, respectively. The mean and variance of the beta distribution are chosen as inputs, rather than the beta parameters, since they are more intuitive quantities to specify. An average FEF of $\overline{d} = 0.80$ was used.

2. Failure probabilities. Failure mode probabilities of occurrence $p_1, \ldots, p_k$ were stochastically generated from a beta distribution with shape parameters $n = \left(\mu - \mu^2 - \sigma^2\right) / \sigma^2$ pseudo trials, and $x = n \cdot \mu$ pseudo failures.
3. Failure histories. Failure histories for each failure mode were stochastically generated via a Bernoulli distribution with parameter $p_i$ for $i = 1, \ldots, k$. The $t_i$ for $i \in obs$ required in our estimation procedures are obtained by these failures histories as the trial numbers when each observed failure mode was first discovered.
4. Test statistic. As a bare minimum, our test procedure requires $c \geq 3$ (i.e., at least 3 cells in the GOF table to give at least one degree of freedom). Using the failure data stochastically generated from the previous step, the test statistic is calculated by a routine which takes into consideration three factors. First, the routine requires the expected frequencies $\hat{E}_i \geq 2$ for each $i = 1, \ldots, c$. Second, the cell boundaries of the class intervals are chosen such that their expected frequencies are close in magnitude to one another. For example, if

there are $m = 12$ observed failure modes and $c = 3$ cells in the GOF table, the expected frequencies would be about $E_i \approx m / c = 4$ for each $i = 1, \ldots, c$. Flexibility in the magnitudes of the expected frequencies for adjacent class intervals is incorporated as a result of the use of an optimality criterion to minimize error in the test statistic. This is achieved by selecting the cell boundaries as the trial number which yields the lowest mean squared error (in addition to satisfying the first two conditions).

The simulation, as outlined above, generates data analogous to that which is typically captured during a single developmental test for a one-shot system. The results below are based on $r = 10,000$ replications of the simulation. For Figures 29 and 30, only 12 and 20 of 50 failure modes were observed on average, respectively, from which to estimate model parameters and the test statistic. Given so few data (which is realistic in many applications) only 3 and 4 cells, respectively, were used to construct the GOF table during each replication. The associated chi-squared distributions have 1 and 2 degrees of freedom, respectively. Figure 29 and 30 show plots of the empirical distributions of the test statistic constructed via simulation, versus their actual chi-squared distributions. The close agreement between the true and empirical distributions indicate that the approximation procedure even under limited data availability for estimation purposes is quite good. Naturally, deterioration in the approximating statistic was observed as the expected frequencies in these plots get smaller than 4 and 5, respectively.

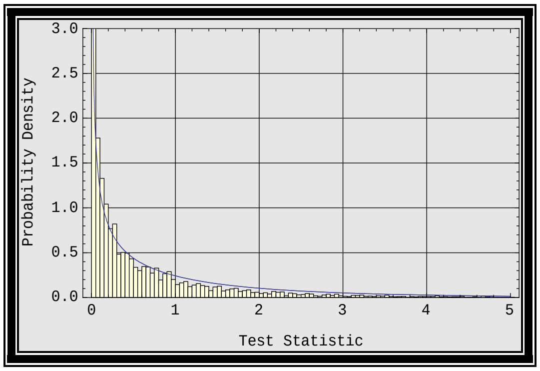

**Figure 29. Empirical Distribution, d.f. = 1.**

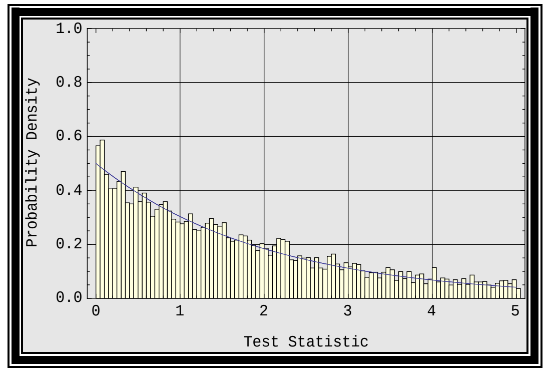

**Figure 30. Empirical Distribution, d.f. = 2.**

### 7.3.2. CI Procedures

#### 7.3.2.1. Numerical Example

Using an average FEF of $\bar{d} = 0.80$, and the $t_i$ for $i \in obs$, the ML estimate of the expected reliability via (17) is $\hat{R} = 0.8649$. Thus, the ML estimate of (18) is $\hat{h} \equiv \ln\left(\frac{\hat{R}}{1-\hat{R}}\right) = 1.8569$. The column vector of partial derivatives of the log-odds transform (18) w.r.t. the beta parameters is, $\hat{H} = [\ 0.1086, -6.4720]^T$, where the partial derivatives (i.e., given by (28) and (29) in the Appendix) evaluate to $\hat{R}'_n = 0.0127$ and $\hat{R}'_x = -0.7561$, respectively. The ML estimate of the inverse Fisher information matrix is $\hat{F}_0^{-1} = \begin{bmatrix} -227.2270 & -5.0428 \\ -5.0428 & -0.1003 \end{bmatrix}$. Given these estimates, the desired standard deviation is $s\left(\hat{h}\right) = \sqrt{\hat{H}^T \cdot \hat{F}_0^{-1} \cdot \hat{H}} = 0.4541$. From (27), an 80% CI estimate on the true but unknown expected reliability is $R \in \left(0.7816, 0.9197\right)$. The true reliability in this example is $R = 0.8662$.

#### 7.3.2.2. Monte Carlo Simulation

A natural concern in the use of an approximate CI routine is the accuracy associated with its coverage probability. In the context of this paper, the coverage probability estimates the fraction of times (out of the total numbers of replications) that the approximate CI contains the true expected reliability of a one-shot system after correction. It is well-known that CI procedures based on a normal approximation (especially when data censoring is involved) tend to yield a coverage probability less than the nominal, or advertised, confidence level (e.g., see Meeker et al. [29], and [36]). This means that an approximate CI tends to be slightly tighter than one which is exact. As a result, an approximate CI routine may advertise, for example 80% confidence, when it is really only giving say 75%.

To address these concerns we have developed a Monte Carlo simulation to estimate the coverage associated with the approximate CI procedure presented herein. The simulation consist of the previous three simulation steps in addition to the point and CI estimation of the quantities given above. Simulation results, after $r = 10,000$ replications, yield 7,920 confidence intervals that contained the true reliability (e.g., 79.2% coverage against an 80% confidence level). The histogram (scaled to a probability density) of the 10,000 estimates of (18) are shown in Figure 31.

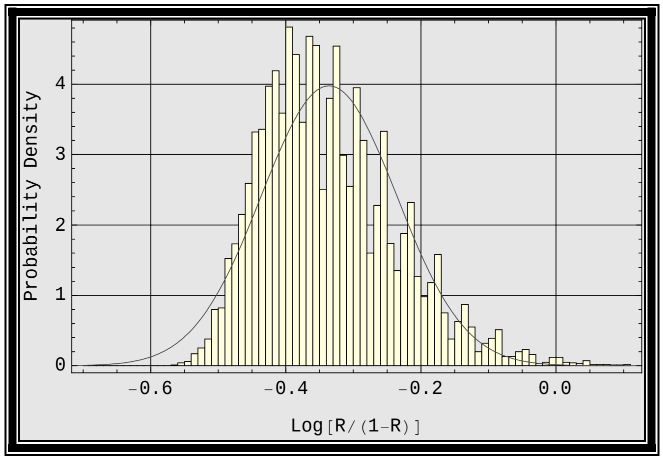


**Figure 31. Probability Density of $\hat{h}$ based on 10K Replications.**

As illustrated by Jeng and Meeker in [29], the coverage of an approximate CI routine depends on the volume of censored data involved. Using the same inputs given above with the exception of decreasing $\mu = 0.02$ (which decreases the number of observed failure modes to $m = 14$), the coverage decreases to 71.5%. Not surprisingly, 71.5% coverage is less than the nominal 80% confidence level that was used. Overall, the important point to take away from this is to be cognizant of the deterioration in the coverage probability (relative to the value of $m$) when using any approximate CI routine.

## 7.4. Concluding Remarks

In this paper, we have presented approximate GOF and CI procedures for our reliability growth projection model given in [39]. The graphical GOF approach highlights the correlation between the actual cumulative number of observed failure modes versus trials, against the expected number of observed failure modes given by our model. The second technique is a statistical GOF test procedure based on an approximating chi-squared random variable. Both of these techniques are based upon our logically derived exact expression (13) of the expected number of failure modes observed in testing. This exact expression is found to be a function of only three unknowns including $k$ (the total potential number of failure modes in the system) and the two beta shape parameters $n$ and $x$. For an assumed value of $k$, our ML estimates of these parameters are given as the solutions to (2) and (3). The sensitivity of not knowing $k$ is quantified by our limiting approximations as $k \to \infty$ (e.g., the solutions to (10) and (11)), which are used in conjunction with (14). Monte Carlo simulation results show that the approximating test statistic follows the chi-squared distribution for the appropriate degrees of freedom when the expected frequencies are sufficiently large (e.g., greater than 4). Using a Fisher matrix normal approximation approach, an approximate CI routine was developed to obtain an interval estimate on the expected reliability of a one-shot system (17). Monte Carlo simulation results indicate that the coverage of this approximate routine is largely a function of the volume of censored data involved. Coverage probabilities generally ranged between 0.70-0.80 in comparison to a nominal 80% confidence level. These results are based on a conservative volume of data (e.g., $14-33$ observed failure modes on average out of 50) available for estimation purposes. Numerical examples were presented to illustrate these techniques.

## Appendix

The partial derivatives of the expected reliability (17) with respect to the beta shape parameters $n$ (i.e., pseudo trials) and $x$ (i.e., pseudo failures) are given respectively by,

$$R_n' = k \cdot \left[ 1 - \left( 1 - \left[ 1 - \frac{n \cdot \Gamma(n-x+t-1) \cdot \Gamma(x+1)}{x \cdot \mathrm{B}(x, n-x) \cdot \Gamma(n+t)} \right] \cdot \bar{d} \right) \cdot \frac{x}{n} \right]^{k-1} \times$$
$$\times \left[ \frac{x}{n^2} - \left( \frac{x}{n^2} + \frac{x \cdot \Gamma(n) \cdot \Gamma(n-x+t-1) \cdot \left[ \psi(n) - \psi(n+t) - \psi(n-x) + \psi(n-x+t-1) \right]}{\Gamma(n+t) \cdot \Gamma(n-x)} \right) \cdot \bar{d} \right] \tag{28}$$

and

$$R_x' = k \cdot \left[ 1 - \left( 1 - \left[ 1 - \frac{n \cdot \Gamma(n-x+t-1) \cdot \Gamma(x+1)}{x \cdot \mathrm{B}(x, n-x) \cdot \Gamma(n+t)} \right] \cdot \bar{d} \right) \cdot \frac{x}{n} \right]^{k-1} \times$$
$$\times \left[ \left( \frac{1}{n} - \frac{\Gamma(n) \cdot \Gamma(n-x+t-1) \cdot \left[ 1 + x \cdot \psi(n-x) - x \cdot \psi(n-x+t-1) \right]}{\Gamma(n+t) \cdot \Gamma(n-x)} \right) \cdot \bar{d} - \frac{1}{n} \right] \tag{29}$$

The ML estimates, $\hat{R}_n'$ and $\hat{R}_x'$, for these expressions are obtained by substitution of the ML estimates $\hat{n}$ and $\hat{x}$ in place of the true but unknown beta parameters $n$ and $x$, respectively.

## Acknowledgements

The authors are grateful to Paul M. Ellner, Ph.D. (AMSAA), Michael J. Cushing, Ph.D. (AEC) and Arthur Fries, Ph.D. (IDA) for their review and critical comments on initial drafts of this manuscript.

## References

[1] J.T. Duane, "Learning curve approach to reliability monitoring," *IEEE Trans. Aerospace*, vol. AS-2, pp. 563-566, 1964.

[2] W.J. Corcoran, H. Weingarten and P.W. Zehna, "Estimating Reliability after Corrective Action," *Management Science*, vol. 10, no. 4, pp. 786-798, 1964.

[3] L.H. Crow, "Reliability analysis for complex, repairable systems," in *Reliability and Biometry*, eds. F. Proschan and R. J. Serfling, Philadelphia: SIAM, 1974, pp. 379-410.

[4] L.H. Crow, "Tracking Reliability Growth," AMSAA, Aberdeen Proving Ground, MD, Technical Report R-30, 1974.

[5] J.M. Finkelstein, "Confidence Bounds on the Parameters of the Weibull Process," *Technometrics*, vol. 18, no. 1, pp. 115-117, 1976.

[6] L.H. Crow, "Confidence Interval Procedures for Reliability Growth Analysis," AMSAA, Aberdeen Proving Ground, MD, Technical Report 197, 1977.

[7] Department of Defense, "Reliability Growth Management," Washington, DC, Military Handbook 189, 1981.

[8] W. Nelson, "Multi-parameter Distribution and Multiply Censored Data" in *Applied Life Data Analysis*, NY: John Wiley & Sons, 1982, pp. 375-384.

[9] L.H. Crow, "Confidence Interval Procedures for the Weibull Process with Applications to Reliability Growth," *Technometrics*, vol. 24, no. 1, pp. 67-72, 1982.

[10] L.H. Crow, "An Improved Methodology for Reliability Growth Projection," AMSAA, Aberdeen Proving Ground, MD, Technical Report 357, 1982.

[11] L. H. Crow, "AMSAA Discrete Reliability Growth Model," AMSAA, Aberdeen Proving Ground, MD, Office Note 1-83, 1983.

[12] R.B. D'Agostino and M.A. Stephens, "Tests of Chi-Squared Type," in *Goodness-of-Fit Techniques*, New York, NY: Marcel Dekker, 1986, pp. 93-93.

[13] S.E. Rigdon, "Testing goodness-of-fit for the power law process," *Communications in Statistics: Theory and Methods*, vol. 18, pp. 4665-4676, 1989.

[14] G.K. Bhattacharyya, A. Fries, and R.A. Johnson, "Properties of Continuous Analog Estimators for a Discrete Reliability-Growth Model," *IEEE Trans. Reliab.*, vol. 38, no. 3, pp. 373-378, 1989.

[15] D. Robinson, and D. Dietrich, "A Nonparametric-Bayes Reliability-Growth Model," *IEEE Trans. Reliab.*, vol. 38, no. 5, pp. 591-598, 1989.

[16] W.J. Park and Y.G. Kim, "Goodness-of-fit test for the power-law process," *IEEE Trans. Reliab.*, vol. 41, pp. 107-111, 1992.

[17] B. Klefsjö and U. Kumar, "Goodness-of-fit test for the power-law process based on the TTT-plot," *IEEE Trans. Reliab.*, vol. 41, pp. 593-598, 1992.

[18] L.H. Crow, "Confidence Intervals on the Reliability of Repairable Systems," *IEEE Proc. RAMS*, 1993, pp. 126-134.

[19] M. Xie and M. Zhao, "On some reliability growth models with graphical interpretations," *Microelectronics and Reliability*, vol. 33, pp. 149-167, 1993.

[20] W.J. Park and M. Seoh, "More goodness-of-fit tests for the power law process," *IEEE Trans. Reliab.*, vol. 43, pp. 275-278, 1994.

[21] P.M. Ellner and L.C. Wald, "AMSAA Maturity Projection Model," *IEEE Proc. RAMS*, 1995, pp. 174-181.

[22] A. Fries and A. Sen, "A Survey of Discrete Reliability Growth Models," *IEEE Trans. Reliab.*, vol. R-45, no. 4, pp. 582-604, 1996.

[23] R.D. Baker, "Some new tests of the power law process," *Technometrics*, vol. 38, pp. 256-265, 1996.

[24] E. Crétois and O. Gaudoin, "New results on goodness-of-fit tests for the power-law process and application to software reliability," *Int. J. Reliab. Qual. Saf. Eng.*, vol. 5, pp. 249-267, 1998.

[25] O. Gaudoin, "CPIT goodness-of-fit test for the power-law process," *Communications in Statistics: Theory and Methods*, vol. 27, pp. 165-180, 1998.

[26] E. Crétois, M.A. El Aroui and O. Gaudoin, "U-plot method for testing the goodness-of-fit of the power-law process," *Communications in Statistics: Theory and Methods*, vol. 28, pp. 1731-1747, 1999.

[27] J. Donovan and E. Murphy, "Reliability growth – A new graphical model," *Qual. Reliab. Eng. Int.*, vol. 15, 1999, pp. 167-174.

[28] P.M. Ellner, W.J. Broemm, and W.J. Woodworth, "Reliability Growth Guide," AMSAA, Aberdeen Proving Ground, MD, Technical Report 652, 2000.

[29] S.L. Jeng, and W.Q. Meeker, "Comparisons of Approximate Confidence Interval Procedures for Type I Censored Data," *Technometrics*, vol. 42, no. 2, pp. 135-148, 2000.

[30] G. Pulcini, "An Exponential Reliability Growth Model in Multicopy Testing Program," *IEEE Trans. Reliab.*, vol. 50, no. 4, pp. 365-373, 2001.

[31] O. Gaudoin, B. Yang, and M. Xie, "A Simple Goodness-of-Fit Test for the Power-Law Process, based on the Duane Plot, *IEEE Trans. Reliab.*, vol. 52, num. 1, pp. 69-74, 2003.

[32] L.H. Crow, "An Extended Reliability Growth Model for Managing and Assessing Corrective Actions," *IEEE Proc. RAMS*, 2004, pp. 73-80.

[33] P.M. Ellner and J.B. Hall, "AMSAA Maturity Projection Model based on Stein Estimation," AMSAA, Aberdeen Proving Ground, MD, Technical Report 751, 2004.

[34] W.J. Broemm, "User Guide for the Visual Growth Suite of Reliability Growth Models," AMSAA, Aberdeen Proving Ground, MD, Technical Report 764, 2005.

[35] Department of Defense, "DoD Guide for Achieving Reliability, Availability and Maintainability," Washington, DC, 2005.

[36] Y. Hong, W.Q. Meeker, and L.A. Escobar, "Avoiding Problems with Normal Approximation Confidence Intervals for Probabilities," *Technometrics*, vol. 50, no. 1, pp. 64-68, 2008.

[37] J.B. Hall and A. Mosleh, "A Reliability Growth Projection Model for One-Shot Systems," *IEEE Trans. Reliab*., vol. 57, no. 1, pp. 174-181, 2008.

[38] J.B. Hall and A. Mosleh, "An Analytical Framework for Reliability Growth Projection of One-Shot Systems," *Reliab Eng & System Safety*, doi: 10.1016/j.ress.2007.11.003.

[39] J.B. Hall, A. Mosleh, and P.M. Ellner, Reliability Growth Management and Methodology of Missile Systems," (to be submitted for publication).

THIS PAGE IS INTENTIONALLY LEFT BLANK.

# 8. BAYESIAN ESTIMATION PROCEDURES


## Abstract

The purpose of this chapter is to develop Bayesian procedures that can be utilized as alternatives to the classical estimation methods developed in Chapters 4 and 6. 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), and the management metrics previously discussed. Another advantage is that all a priori engineering knowledge can be utilized in the assessment procedure. Analytical results are presented to obtain Bayes' estimates of the beta shape parameters for both corrective action strategies. A Monte Carlo approach is outlined for constructing uncertainty distributions on the management metrics. For inference on interval estimation, Bayesian probability limits are obtained in the usual manner (i.e., via desired percentiles of the uncertainty distributions). Numerical examples are given to illustrate these Bayesian procedures. In particular, Bayes' estimates of the beta shape parameters are obtained for a given sample of data. Also, Bayesian epistemic uncertainty distributions for all reliability growth management metrics are constructed via the proposed Monte Carlo approach.


**Keywords**: Bayesian, beta distribution, shape parameters.

# - CHAPTER 8 -

## Acronyms

**FEF** – Fix Effectiveness Factor(s)

**FOT** – First Occurrence Trial(s)

**MME** – Method of Moments Estimation/Estimate(s)

**MLE** – Maximum Likelihood Estimation/Estimate(s)

## Notation

$k$ - total number of potential failure modes.

$m$ - total number of observed failure modes.

$n$ - beta shape parameter representing the pseudo number of trials.

$x$ - beta shape parameter representing the pseudo number of failures.

$T$ - total number of trials.

$N_{i,j}$ - number of failures for mode $i$ in trial $j$ – zero or unity.

$N_i$ - total number of failures for mode $i$ in $T$ trials.

$t_i$ - trial number of first occurrence of failure mode $i$.

$p_i$ - probability of occurrence for failure mode $i$.

## 8.1. Background

The foundation of Bayesian statistical inference rests upon the notion of subjective probability, which contrasts with the well-known classical frequency interpretation[52]. The Bayesian approach is achieved via the construction of a posterior distribution of belief for a given parameter. Such a distribution follows directly from Bayes' Theorem [97] given by,

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)} \tag{1}$$

where

1. $P(A \mid B)$ is the posterior distribution.
2. $P(B \mid A)$ is the likelihood function.
3. $P(A)$ is the prior distribution.
4. $P(B)$ is referred to the normalization, or correction, factor.

In (1) the prior, $P(A)$, expresses the state of knowledge, or ignorance, about event *A* without sample data or previous experience. The likelihood function, $P(B \mid A)$, expresses the state of knowledge about event *B* given evidence *A*. The normalization factor, $P(B)$, ensures that the posterior distribution is a *PDF* and can be expressed as,

$$P(B) = P(B \mid A) \cdot P(A) + P(B \mid \bar{A}) \cdot P(\bar{A}) \tag{2}$$

where $\bar{A}$ denotes the compliment of the event *A*. Finally, the posterior distribution, $P(A \mid B)$, expresses the state of knowledge about *A* given some type of known information, data or evidence, *B*.

## 8.2. Likelihood Functions

[52] The frequency interpretation refers to the concept of probability as defined by the limiting frequency of repeatable events.

### 8.2.1. Delayed Strategy

Estimation procedures (e.g., MME and MLE) were developed in Chapter 4 to estimate the beta shape parameters under a delayed corrective action strategy. 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 $d_i$ for observed failure modes) during test. In this case, the marginal distribution of an individual observation (e.g., $N_i$) given by Martz and Waller [97] is,

$$f\left(N_i \mid x,n\right) = \frac{T!\cdot\Gamma\left(n\right)\cdot\Gamma\left(N_i + x\right)\cdot\Gamma\left(T+n-N_i-x\right)}{N_i!\cdot\left(T-N_i\right)!\cdot\Gamma\left(T+n\right)\cdot\Gamma\left(x\right)\cdot\Gamma\left(n-x\right)} \tag{3}$$

where $n$ represents pseudo trials, $x$ represents pseudo failures, and $\Gamma\left(x\right) \equiv \int_0^\infty t^{x-1}\cdot e^{-t}dt$ is the Euler gamma function. The product of these terms for each failure mode is the (joint) marginal likelihood function of the entire sample,

$$L\left(\vec{N} \mid x,n\right) = \prod_{i=1}^{k} f\left(N_i \mid x,n\right) = \prod_{i=1}^{k} \frac{T!\cdot\Gamma\left(n\right)\cdot\Gamma\left(N_i + x\right)\cdot\Gamma\left(T+n-N_i-x\right)}{N_i!\cdot\left(T-N_i\right)!\cdot\Gamma\left(T+n\right)\cdot\Gamma\left(x\right)\cdot\Gamma\left(n-x\right)} \tag{4}$$

Two important assumptions associated with the derivation of (4) is that the failure probability, and number of failures observed <u>for each failure mode</u> (e.g., the conditional distributions) follow the beta, and binomial distributions, respectively. Thus, (3) is sometimes refereed to as the marginal beta-binomial distribution, and is a popular doubly stochastic process.

### 8.2.2. Arbitrary Strategy

If program management adopts an arbitrary corrective action strategy where, failure probabilities for observed failure modes (e.g., the $p_i$ for $i \in obs \equiv \{i : N_i > 0 \text{ for } i=1,\ldots,k\}$) can be reduced by the fractional amounts specified by their corresponding FEF. Under this corrective action strategy, the marginal distribution for a single observation $N_i$ is not, in general, given by (3). The reason why is because this strategy does not operate under binomial sampling since the otherwise binomial parameter $p_i$ may not remain constant over the entire $T$ trials for each $i = 1,\ldots,k$, as assumed in the derivation of (3). Thus, to estimate the beta shape parameters under this corrective action strategy a likelihood function is needed that is consistent with the manner in which failure modes are mitigated. Such a likelihood function is presented in detail in Chapter 6 and is based on the trial numbers when failure modes are first discovered. The (joint) marginal likelihood function for the sample is given by,

$$L(\vec{t} \mid n, x) = m! \cdot \binom{k}{m} \cdot \left[ \frac{\Gamma(n) \cdot \Gamma(n - x + T)}{\Gamma(n - x) \cdot \Gamma(n + T)} \right]^{k-m} \cdot \prod_{i \in obs} \left[ \frac{\Gamma(x+1) \cdot \Gamma(n - x + t_i - 1)}{\Gamma(n + t_i) \cdot \mathrm{B}(x, n - x)} \right] \quad (5)$$

where $\mathrm{B}(a,b) \equiv \frac{\Gamma(a) \cdot \Gamma(b)}{\Gamma(a+b)} = \int_0^1 t^{a-1} \cdot (1-t)^{b-1}\, dt$ is the Euler beta function, $m$ is the total number of observed failure modes, and $\vec{t}$ is the vector of first occurrence trial numbers $t_i$ for $i \in obs$. Notice that the middle product-term in (5) represents the $k - m$ Type I (i.e., time) censored observations which result from the unobserved failure modes. Thus, data censoring is accounted for.

## 8.3. Estimation of Beta Shape Parameters

### 8.3.1. Delayed Strategy

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

$$\pi_0(x,n) = c \in \mathbb{R}^+ \tag{6}$$

If the failure probabilities $p_1,\ldots,p_k$ are assumed to constitute a realization of an *i.i.d.* sample $P_1,\ldots,P_k$ such that $P_i \sim Beta(n,x)$, the joint posterior based on the entire sample for a uniform prior is,

$$\begin{aligned}\pi(x,n \mid \vec{N}) &= \frac{L(\vec{N} \mid x,n)\cdot \pi_0(x,n)}{\int_{n=0}^{\infty}\int_{x=0}^{n} L(\vec{N} \mid x,n)\cdot \pi_0(x,n)\,dxdn} \\ &= \frac{\prod_{i=1}^{k}\left[\frac{T!\cdot\Gamma(n)\cdot\Gamma(N_i+x)\cdot\Gamma(T+n-N_i-x)}{N_i!\cdot(T-N_i)!\cdot\Gamma(T+n)\cdot\Gamma(x)\cdot\Gamma(n-x)}\right]}{\int_{n=0}^{\infty}\int_{x=0}^{n}\prod_{i=1}^{k}\left[\frac{T!\cdot\Gamma(n)\cdot\Gamma(N_i+x)\cdot\Gamma(T+n-N_i-x)}{N_i!\cdot(T-N_i)!\cdot\Gamma(T+n)\cdot\Gamma(x)\cdot\Gamma(n-x)}\right]dxdn}\end{aligned} \tag{7}$$

Under squared-error loss, the Bayes' estimates of $n$ and $x$ are calculated as mathematical expectations of the posterior distribution. The Bayes' estimates are given by,

$$\bar{n} \equiv E_n\left[\pi(x,n \mid \vec{N})\right] \equiv \int_{n=0}^{\infty}\int_{x=0}^{n} n\cdot\pi(x,n \mid \vec{N})\,dxdn \tag{8}$$

and

$$\bar{x} \equiv E_x\left[\pi(x,n \mid \vec{N})\right] \equiv \int_{n=0}^{\infty}\int_{x=0}^{n} x\cdot\pi(x,n \mid \vec{N})\,dxdn \tag{9}$$

respectively.

### 8.3.2. Arbitrary Strategy

For a delayed or non-delayed corrective action strategy, the posterior distribution is constructed similarly using (5),

$$\pi\left(x,n\,|\,\vec{t}\right)=\frac{L\left(\vec{t}\,|\,x,n\right)\cdot\pi_0\left(x,n\right)}{\int_{n=0}^{\infty}\int_{x=0}^{n}L\left(\vec{t}\,|\,x,n\right)\cdot\pi_0\left(x,n\right)dxdn}$$
$$=\frac{\left[\frac{\Gamma(n)\cdot\Gamma(n-x+T)}{\Gamma(n-x)\cdot\Gamma(n+T)}\right]^{k-m}\cdot\prod_{i\in obs}\left[\frac{\Gamma(x+1)\cdot\Gamma(n-x+t_i-1)}{\Gamma(n+t_i)\cdot\mathrm{B}(x,n-x)}\right]}{\int_{n=0}^{\infty}\int_{x=0}^{n}\left(\left[\frac{\Gamma(n)\cdot\Gamma(n-x+T)}{\Gamma(n-x)\cdot\Gamma(n+T)}\right]^{k-m}\cdot\prod_{i\in obs}\left[\frac{\Gamma(x+1)\cdot\Gamma(n-x+t_i-1)}{\Gamma(n+t_i)\cdot\mathrm{B}(x,n-x)}\right]\right)dxdn} \tag{10}$$

The resulting Bayes' estimates of the beta shape parameters based on the entire sample are,

$$\overline{n}\equiv E_n\left[\pi\left(x,n\,|\,\vec{t}\right)\right]\equiv\int_{n=0}^{\infty}\int_{x=0}^{n}n\cdot\pi\left(x,n\,|\,\vec{t}\right)dxdn \tag{11}$$

and

$$\overline{x}\equiv E_x\left[\pi\left(x,n\,|\,\vec{t}\right)\right]\equiv\int_{n=0}^{\infty}\int_{x=0}^{n}x\cdot\pi\left(x,n\,|\,\vec{t}\right)dxdn \tag{12}$$

### 8.3.3. Numerical Example

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 that was generated is shown in Table 15. 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 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.

| | **Failure Mode** | **Probability** $p_i$ | **Failures** $N_i$ | **FOT** $t_i$ | **FEF** $d_i$ |
|---|---|---|---|---|---|
| **1** | 10 | 0.0211 | 2 | 11 | 0.8874 |
| **2** | 12 | 0.0172 | 1 | 27 | 0.8010 |
| **3** | 13 | 0.1016 | 10 | 7 | 0.8613 |
| **4** | 14 | 0.0195 | 1 | 30 | 0.8958 |
| **5** | 17 | 0.0286 | 2 | 17 | 0.5812 |
| **6** | 19 | 0.1187 | 7 | 4 | 0.6189 |
| **7** | 20 | 0.0164 | 1 | 32 | 0.9222 |

**Table 15. Failure Data.**

For a delayed corrective action strategy, estimates of the beta shape parameters are obtained using MME, MLE, as well as the Bayesian estimation procedures given above. These parameter estimates, which are based on the count data shown in column 4 of Table 15, are given in columns 2-5 of Table 16, 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 Equations (22) and (23), respectively, from Chapter 4. The MLE estimates, shown in column 3, are obtained as the solutions to Equations (31) and (32) from Chapter 4 when solved simultaneously. The Bayesian estimates of $n$ and $x$ are obtained by Equations (8) and (9) above, respectively.

For an arbitrary corrective action strategy, estimates of the beta parameters are obtained via MLE and Bayesian estimation procedures. These estimates are shown in columns 6 and 7 in Table 16. The MLE estimates are obtained as the solutions to Equations (6) and (7) in Chapter 6. The Bayes estimates are given by Equations (11) and (12) above. 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 must 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. The best results were discovered to result when integrating over the parameter space $n \in (0,1.5 \cdot \widehat{n}_k)$, where $\widehat{n}_k$ is the finite $k$ MLE estimate for the beta parameter $n$ given in column 4 of Table 16. For some stochastically generated datasets, this MLE can be lower than the true value of $n$. This is the reason why the integration was performed using a multiple of 1.5 beyond $\widehat{n}_k$. This endeavors to ensure the entire volume of under the parameter space is accounted for numerically.

| | | DELAYED | | | ARBITRARY | |
|---|---|---|---|---|---|---|
| | **True** | **MME** | **MLE** | **Bayes** | **MLE** | **Bayes** |
| $n$ | 8.750 | 9.650 | 9.206 | 8.797 | 32.320 | 9.168 |
| $x$ | 0.219 | 0.232 | 0.220 | 0.241 | 0.458 | 0.245 |

**Table 16. Beta Parameter Estimates.**

Notice that the MLE obtained under an arbitrary corrective action strategy is larger then the Bayes estimate. There are a few points to be aware of when studying these estimates. First, the estimates under an arbitrary corrective action strategy are not directly comparable to the true parameters, or the parameter approximations given in the delayed case. The reason why is because the distribution is changing as a result of the failure probabilities being reduced via the corrective action effort. Second, what would seem like large error in the magnitudes of the beta parameters, does not translate into large error for approximating the distribution, or the management metrics. Finally, it should be noted that these estimates were generated from a single dataset (not estimated over several stochastic realizations). Thus, the departures between the MLE and Bayes estimates are only an artifact of the outcome of a single dataset, and not stable results that would be achieved by replicating this process thousands of times.

Figures 32 and 33 show the beta PDF approximations using the parameter estimates given in Table 16 for the delayed and arbitrary corrective action strategies, respectively. Not surprisingly, the PDF approximations based on classical versus Bayesian methods are very close to one another. For this particular stochastic realization, the Bayesian estimation procedure based on FOT (i.e., Figure 33) is found to more accurately approximate the true beta PDF in comparison to MLE.

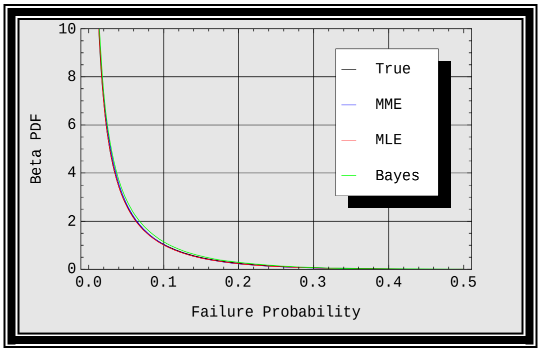


Figure 32. Beta PDF – Delayed.

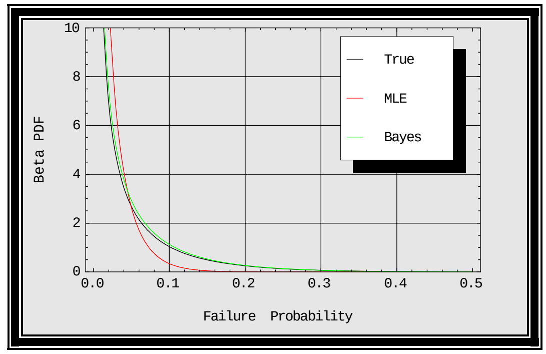


Figure 33. Beta PDF – Arbitrary.

## 8.4. Management Metrics

### 8.4.1. Simulation Approaches

In the preceding section, Bayesian estimates of the beta shape parameters were obtained to approximate the initial prior density function of the failure probabilities inherent to a complex one-shot system. This density can now be used to construct an initial prior density for each of the management metrics (e.g., a prior for the reliability of the corrected system). Two Monte Carlo approaches, outlined below, are utilized to construct the desired empirical distributions. The first simulation approach addresses the case where the $P_1,\ldots,P_k$ are considered to be a random sample from an *i.i.d.* beta r.v.. The second simulation approach address the case where the $P_i$ for $i=1,\ldots,k$ are still independent beta random variables, but not necessarily identically distributed. Without loss of generality, a numerical example is given to illustrate the first simulation approach. The simulation steps are as follows:

1. Use the Bayesian estimation procedures in the preceding 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 $N_i$ 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 is the set of indices associated with failure modes observed during simulation. The FOT $t_i$ for $i \in obs$ 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 below calculates the beta parameters based on a mean of 0.80 and variance of 0.01, which yield values of FEF typically assigned by a FPRB. Note that these FEF remain fixed (i.e., are not generated anew) during each replication.
5. Use the data obtained in the previous steps to calculate the management metric of interest. In the example below, uncertainty distributions are

constructed for all of the management metrics. They include: the expected initial reliability of the system (Equation (8) in Chapter 5),

$$R_k\left(t=1 \mid \vec{p}\right)=\prod_{i=1}^{k}\left(1-p_i\right) \tag{13}$$

the expected reliability of the corrected system in trial $T$ (i.e., the final reliability, also computed via Equation (8) in Chapter 5),

$$R_k\left(T \mid \vec{p}\right)=\prod_{i=1}^{k}\left[1-\left(1-\left[1-\left(1-p_i\right)^{T-1}\right]\cdot d_i\right)\cdot p_i\right] \tag{14}$$

the reliability growth potential (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 Chapter 5),

$$R_{GP}\left(t \mid \vec{p}\right)=\prod_{i=1}^{k}\left[1-\left(1-d_i\right)\cdot p_i\right] \tag{15}$$

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 Chapter 5),

$$\mu_k\left(T \mid \vec{p}\right)=k-\sum_{i=1}^{k}\left(1-p_i\right)^{T} \tag{16}$$

the expected probability of observing a new failure mode on trial $T$ (Equation (16) in Chapter 5),

$$h_k\left(T \mid \vec{p}\right)=1-\prod_{i=1}^{k}\left[1-\left(1-p_i\right)^{T-1}\cdot p_i\right] \tag{17}$$

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 Chapter 5),

$$\phi_k\left(T \mid \vec{p}\right) = \frac{1-\prod_{i=1}^{k}\left(1-\left[1-\left(1-p_i\right)^{T-1}\right]\cdot p_i\right)}{1-\prod_{i=1}^{k}\left(1-p_i\right)} \tag{18}$$

6. Repeat steps 1-3 and step 5 several times (e.g., $r = 10,000$). This generates values of the management metrics from which their empirical distributions are constructed.

After the simulation is replicated several times, construct a histogram scaled to an area of unity using the data generated in step 6. Beta and normal PDF approximations to the histogram can be obtained by estimating distribution parameters as either the mean and variance of the data, or functions thereof. The Bayesian point-estimate and probability limits for the expected reliability of the corrected system are obtained in the usual manner (e.g., as the mean and desired percentiles of the distribution, respectively).

If the $P_1, \ldots, P_k$ are not identically distributed an uncertainty distribution can be constructed for each observed failure mode by using in (7), $f\left(N_i \mid x, n\right)$ given by (3) in place of $L\left(\vec{N} \mid x, n\right)$. For a uniform prior distribution, this gives a different joint distribution of the beta parameters $n_i$ and $x_i$ for each $p_i$ for $i = 1, \ldots, k$,

$$\begin{aligned}\pi\left(x_i, n_i \mid N_i\right) &= \frac{f\left(N_i \mid x_i, n_i\right)\cdot\pi_0\left(x_i, n_i\right)}{\int_{n_i=0}^{\infty}\int_{x_i=0}^{n_i} f\left(N_i \mid x_i, n_i\right)\cdot\pi_0\left(x_i, n_i\right)dx_i dn_i} \\ &= \frac{\left[\dfrac{\Gamma\left(n_i\right)\cdot\Gamma\left(N_i + x_i\right)\cdot\Gamma\left(T + n_i - N_i - x_i\right)}{\Gamma\left(T + n_i\right)\cdot\Gamma\left(x_i\right)\cdot\Gamma\left(n_i - x_i\right)}\right]}{\int_{n_i=0}^{\infty}\int_{x_i=0}^{n_i}\left[\dfrac{\Gamma\left(n_i\right)\cdot\Gamma\left(N_i + x_i\right)\cdot\Gamma\left(T + n_i - N_i - x_i\right)}{\Gamma\left(T + n_i\right)\cdot\Gamma\left(x_i\right)\cdot\Gamma\left(n_i - x_i\right)}\right]dx_i dn_i}\end{aligned} \tag{19}$$

with Bayesian parameter estimates

$$\bar{n}_i \equiv E_{n_i}\left[\pi\left(x_i, n_i \mid N_i\right)\right] \equiv \int_{n_i=0}^{\infty}\int_{x_i=0}^{n_i} n_i \cdot \pi\left(x_i, n_i \mid N_i\right) dx_i dn_i \tag{20}$$

and

$$\bar{x}_i \equiv E_{x_i}\left[\pi\left(x_i, n_i \mid N_i\right)\right] \equiv \int_{n_i=0}^{\infty}\int_{x_i=0}^{n_i} x_i \cdot \pi\left(x_i, n_i \mid N_i\right) dx_i dn_i \tag{21}$$

Note that the distribution and associated Bayesian parameter estimates will be identical for failure modes $i$ and $j$ when $N_i = N_j$. Using and, failure probabilities $p_1, \ldots, p_k$ are stochastically generated from a beta distribution with Bayesian shape parameter estimates $\bar{n}_i$ (pseudo trials) and $\bar{x}_i$ (pseudo failures). Let $P_i \sim beta\left(\bar{n}_i, \bar{x}_i\right)$ denote this distribution, and let $g\left(t \mid \bar{p}\right)$ represent one of the model metrics given in Chapter 5 (e.g., the expected reliability of the corrected system on trial $t$). Using a single realization for each $p_i$, one can calculate a single estimate of $g\left(t \mid \bar{p}\right)$. Once again, an empirical distribution for the model metric is constructed by replicating this process (e.g., say $r = 10,000$ times). The Bayesian point-estimate and probability limits of the expected reliability of the corrected system are obtained in the usual manner. Note that these procedures take into account Type 1 (i.e., time) censoring, as well as complete data. For example, right-censored, or suspended, observations occur when $N_i = 0$, which yields the reliability function of the marginal beta-binomial distribution[53] given by $P\left(N_i = 0\right) = \dfrac{\Gamma\left(n_i\right) \cdot \Gamma\left(n_i - x_i + T\right)}{\Gamma\left(n_i + T\right) \cdot \Gamma\left(n_i - x_i\right)}$.

[53] The reliability functions for the beta-binomial and beta-geometric distributions are identical. For example, if $X \sim bin\left(T, p_i\right)$, then $R_X \equiv P\left(X = 0\right) = \left(1 - p_i\right)^T$. Also, if $Y \sim geo\left(p_i\right)$, then $R_Y \equiv P\left(Y > T\right) = 1 - P\left(Y \le T\right) = 1 - \sum_{t=1}^{T}\left(1 - p_i\right)^{t-1} \cdot p_i = \left(1 - p_i\right)^T$. $\therefore R_X = R_Y$.

### 8.4.2. Numerical Example

Continuing from the example presented in Section 8.3.3, 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 16. Using these failure probabilities, failure histories for each failure mode where generated from a Bernoulli random variable. The index-set of observed failure modes were defined uniquely in 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 the management metrics were calculated. This process was replicated $r = 10,000$ times during simulation. Figures 34-39 below show the empirical densities for each of the management metrics that were constructed. Table 17 summarizes the true values of the management metrics, as well as their associated point and interval estimates. The point and interval estimates are obtained as the mean and as the $10^{th}$ and $90^{th}$ percentiles of the uncertainty distributions, respectively.

Using the same inputs given above, the approximate confidence interval on projected reliability given in Chapter 7 is $R \in (0.73,\ 0.91)$ with a point estimate of $\hat{R} = 0.8436$. While the results obtained from the classical and Bayesian approaches are nearly identical, one should note that the Bayesian point and interval estimates are much more stable. The reason why is because the Bayesian estimates are based on an uncertainty distribution (e.g., shown in Figure 35) constructed from 10,000 stochastically generated datasets. The classical confidence interval procedure is only structured for one dataset, hence, it is only based on a single stochastic realization. As mentioned, the same simulation inputs were used to generate the data for these two approaches.

| Management Metric | Distribution | Parameters | True & Point Estimate | 80% Interval Estimate |
|---|---|---|---|---|
| Initial Reliability | Beta | $n = 11.62$<br>$x = 6.62$ | $R_I = 0.6918$<br>$\hat{R}_I = 0.5699$ | $R_I \in (0.38, 0.75)$ |
| Projected Reliability | Beta | $n = 45.25$<br>$x = 37.74$ | $R = 0.8766$<br>$\hat{R} = 0.8341$ | $R \in (0.76, 0.90)$ |
| Growth Potential | Beta | $n = 32.43$<br>$x = 28.81$ | $R_{GP} = 0.9182$<br>$\hat{R}_{GP} = 0.8883$ | $R_{GP} \in (0.81, 0.95)$ |
| Expected Failure Modes | Normal | $\mu = 7.58$<br>$\sigma = 1.71$ | $m = 7$<br>$\hat{\mu} = 6.38$ | $\mu \in (5.39, 9.78)$ |
| Probability of New Mode | Beta | $n = 374.90$<br>$x = 17.12$ | $h = 0.0542$<br>$\hat{h} = 0.0510$ | $h \in (0.04, 0.07)$ |
| Fraction of P[F] Observed | Beta | $n = 17.10$<br>$x = 15.54$ | $\phi = 0.8697$<br>$\hat{\phi} = 0.9085$ | $\phi \in (0.82, 0.98)$ |

**Table 17. Management Metrics.**

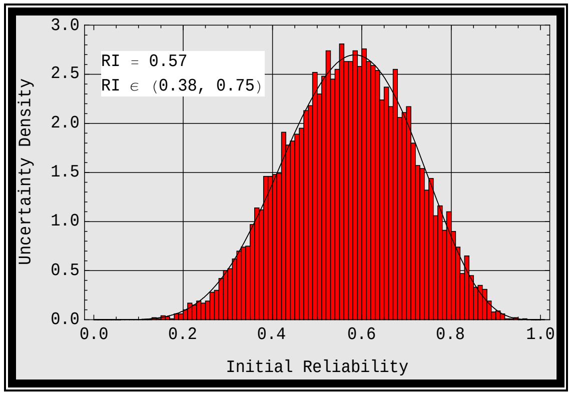


**Figure 34. Initial Reliability.**

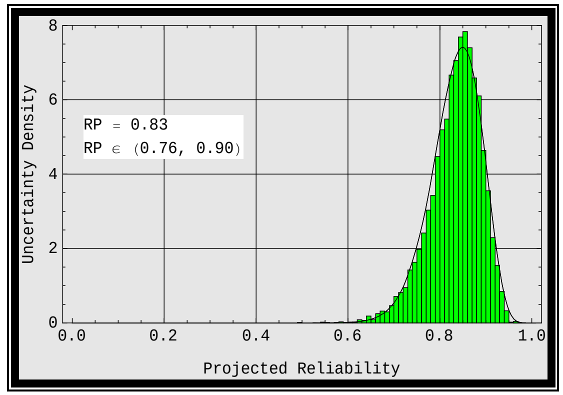


**Figure 35. Projected Reliability.**

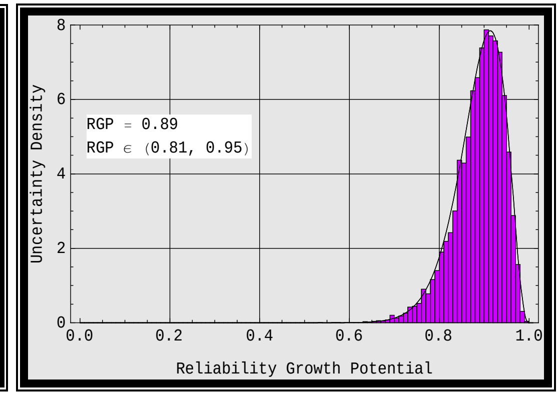


**Figure 36. Reliability Growth Potential.**

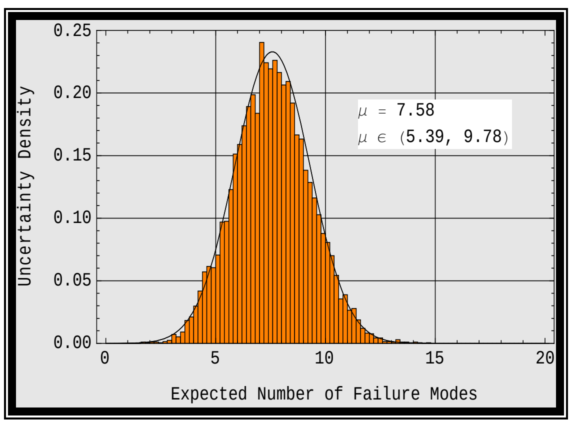


**Figure 37. Expected Number of Failure Modes.**

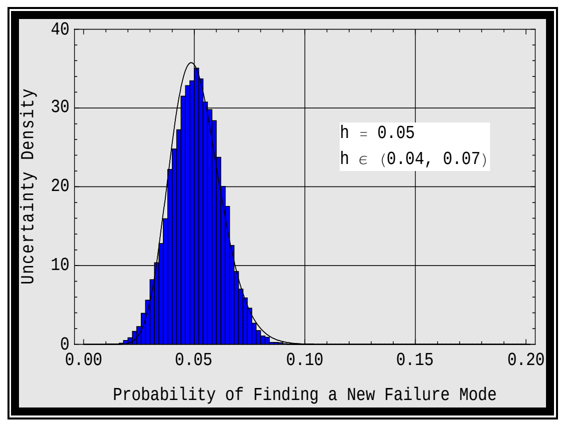


**Figure 38. Probability of New Failure Mode.**

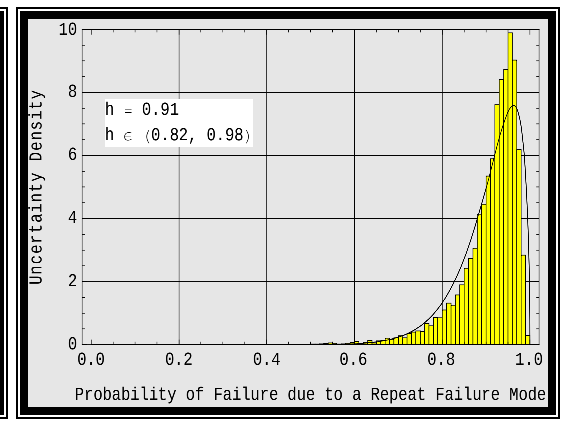


**Figure 39. P[F] due to a Repeat Failure Mode.**

| | MEMBER | TITLE | ADDRESS | PHONE | E-MAIL |
|---|---|---|---|---|---|
| 1 | **Committee Chair**<br><br>Dr. Ali Mosleh | Professor of Mechanical Engineering | University of Maryland, College Park<br>Department of Mechanical Engineering<br>2181 Glenn L. Martin Hall, Room 0151-F<br>College Park, Maryland 20742 | 301-405-5215 | mosleh@umd.edu |
| 2 | **Regular Member**<br><br>Dr. Mohammad Modarres | Professor of Reliability Engineering | University of Maryland, College Park<br>Department of Mechanical Engineering<br>2181 Glenn L. Martin Hall, Room 0151-C<br>College Park, Maryland 20742 | 301-405-5226 | modarres@umd.edu |
| 3 | **Regular Member**<br><br>Dr. Shapour Azarm | Professor of Mechanical Engineering | University of Maryland, College Park<br>Department of Mechanical Engineering<br>2181 Glenn L. Martin Hall, Room 2155<br>College Park, Maryland 20742 | 301-405-5250 | azarm@umd.edu |
| 4 | **Regular Member**<br><br>Dr. Michel Cukier | Associate Professor of Reliability Engineering | University of Maryland, College Park<br>Department of Mechanical Engineering<br>2181 Glenn L. Martin Hall, Room 0151-E<br>College Park, Maryland 20742 | 301-314-2804 | mcukier@umd.edu |
| 5 | **Dean's Representative**<br><br>Dr. Elise Miller-Hooks | Associate Professor of Civil and Environmental Engineering | University of Maryland, College Park<br>Department of Civil and Environmental Engineering<br>1173 Glenn L. Martin Hall, Room 1182<br>College Park, Maryland 20742 | 301-405-2046 | elisemh@umd.edu |
| 6 | **Special Member**<br><br>Dr. Paul Ellner | Technical Expert in Reliability Growth | Director, U.S. AMSAA<br>Logistics Analysis Division<br>392 Hopkins Road (ATTN: AMSRD-AMS-LA)<br>APG, Maryland, 21005 | 410-278-9314 | paul.m.ellner@us.army.mil |

**STUDENT BIOGRAPHY:**

**J. Brian Hall** is a team leader in the Ground Combat Division of the Reliability and Maintainability Directorate; U.S. Army Evaluation Center, Aberdeen Proving Ground, Maryland. In this position, Mr. Hall performs, and oversees others in performing, independent reliability evaluations of Army, and DoD systems. Before this position, Mr. Hall worked as a Project Leader, and Operations Research Analyst with the U.S. AMSAA. At AMSAA, Mr. Hall worked in logistics for nearly eight years providing support in the area of reliability, and reliability growth to various DoD organizations and their supporting defense contractors on associated weapon system development programs. He was also detailed on a six month assignment working on the staff of the Commanding General (General Benjamin S. Griffin), U.S. Army Materiel Command; Fort Belvoir, Virginia, where he led various operations research projects. Mr. Hall earned a BS in Mathematics from Loyola College, and a MS in Mathematics from The Johns Hopkins University; Baltimore, Maryland. He is a graduate of the U.S. Army Logistics Management College at Fort Lee, Virginia, as well as the U.S. Army Management Staff College at Fort Belvoir, Virginia. He is a certified member of the U.S. Army Acquisition Corps, and is currently pursuing a Ph.D. in Engineering Reliability at the University of Maryland, College Park. He is recently married to the former Alyson Taylor Moore.

**CONTACT INFORMATION:**


| | ADDRESS | PHONE* | E-MAIL |
|---|---|---|---|
| **Work** | Director; U.S. Army Evaluation Center<br>ATTN: TEAE-REG<br>4120 Susquehanna Avenue<br>Aberdeen Proving Ground, Maryland 21005 | Phone<br>410-306-1026<br><br>FAX<br>410-306-0476 | brian.hall@us.army.mil |
| **Home** | 611 Lochern Terrace<br>Bel Air, Maryland 21015 | Phone<br>410-420-6937<br><br>FAX<br>410-420-9854 | jbrianhall@msn.com |
| **Mobile** | N/A | 443-643-6968 | brian.hall@us.army.mil |

## A

- **A-mode** – a failure mode that will not be addressed via corrective action.
- **Assignable-cause failure mode** – a failure mode whose root-cause has been identified.

## B

- **B-mode** – a failure mode that will be addressed via corrective action, if observed.
- **BC-mode** – a B-mode that is addressed via corrective action before the conclusion of the current test phase.
- **BD-mode** – a B-mode that is addressed via corrective action at the conclusion of the current test phase.

## C

- **Corrective action** – any physical action taken to permanently mitigate the occurrence of a failure mode. Tactical fixes are not considered to be a corrective action.

## D

- **Developmental test** – a test of initial or intermediate prototypes of a system (e.g., typically possessing an immature design) which is focused on exposing design weaknesses (i.e., failure modes) that can be analyzed and effectively corrected.

## F

- **Failure discounting** – the practice of removing fractions of previous failures associated with a given failure mode after corrective action has been implemented.
- **Failure mode** – the root-cause associated with the loss of a required function or component whose probability (or rate) of occurrence is reduced by a specified FEF, if addressed by corrective action. Note that it may be the case that some failure modes are not observed during testing, or may not be corrected if they are observed (e.g., some failures may not be economically justifiable to correct).
- **Fix Effectiveness Factor (FEF)** – the fraction reduction in an initial mode failure probability (or rate) due to implementation of a unique corrective action.
- **First Occurrence Trial/Time (FOT)** – the trial number (or exact time) when an individual failure mode was first discovered during testing.

## G

- **Growth Rate** – the growth rate, typically denoted by $\alpha$, is a reliability growth planning parameter that represents the negative of the cumulative failure rate versus cumulative test time on a log-log scale (i.e., the slope of the Duane plot). The growth rate should not be confused with the general rate of improvement in the reliability of a developmental item. It is a specific reliability growth planning parameter associated with a single model, namely, the Duane model.

## I

- **Idealized Curve** – a reliability growth planning curve that is based on initial conditions (e.g., initial MTBF, length of the initial test phase), and other planning parameters, such as, an assumed growth rate and management strategy.

## M

- **Management Strategy (MS)** – a reliability growth planning parameter that represents the portion of a system's failure intensity (or probability of failure) associated with failure modes that program management is planning to address via corrective action.

## N

- **Non-homogeneous Poisson Process (NHPP)** – a non-stationary Poisson process (e.g., a Poisson process with an increasing or decreasing intensity function, such as the power-law process).

## O

- **Observed mode** – a failure mode which exhibits at least one failure during testing.
- **One-shot system** – a system whose usage is measured in terms of discrete trials, or demands, such as, guns, rockets, missile systems, and torpedoes.

## P

- **Planning curve** – a smooth-curve representation (i.e., given by the Duane model) of the anticipated reliability growth of a system over the course of its planned test program.
- **Poisson Process** – a stationary counting process of discrete events, say $N(t)$, in an interval of time that is independent of the number of events that have previously occurred. A Poisson process must have an initial condition of $N(t=0)=0$, and events must be orderly in the sense that the occurrence of two or more events in an small interval of time is impossible.
- **Power-law process (PLP)** – a non-homogeneous Poisson Process with intensity function $r(t) \equiv \lambda \cdot \beta \cdot t^{\beta-1}$. A PLP is also referred to as a Weibull process because its time to first failure follows the Weibull distribution.

## R

- **Reliability growth** – the increase in the true (unknown) reliability of a system as a result of failure mode discovery, analysis, and effective correction.
- **Reliability growth management** – the systematic planning for reliability achievement by controlling the ongoing rate of achievement by the allocation and reallocation of program resources based on comparisons between planned and

demonstrated reliability values.

- **Reliability growth planning** – an area of reliability growth that addresses program schedules, amount of testing, resources available, and the realism of the test program in achieving its requirements. Reliability growth planning is portrayed and quantified through a reliability growth planning curve.
- **Reliability growth potential** – the theoretical upper-limit on system reliability achieved by finding and correcting <u>all failure modes</u> with a specified level of fix effectiveness.
- **Reliability growth projection** – an area of reliability growth that focuses on quantifying the reliability that could be achieved if observed failure modes inherent to the system are mitigated by a specified level of fix effectiveness.
- **Reliability growth tracking** – an area of reliability growth that provides management the opportunity to gauge the progress of the development effort by quantifying the demonstrated reliability of a system throughout its test program.
- **Repeat failure mode** – a failure mode which exhibits at least two failures during testing. Repeat failure modes are particularly important under a delayed corrective action strategy. The reason why is because the moment estimators and the likelihood estimators of the beta (or gamma) parameters do not exist unless there is at least one repeat failure mode.
- **Runs-test** – a sequence of tests that are conducted until a specified number of consecutive successful trials are achieved.

## S

- **Seglie's stopping criterion** – a stopping criterion for a developmental test that consists of stopping all trials after a successful runs-test is achieved.

## T

- **Tactical fix** – a physical action that temporarily mitigates the occurrence of a failure mode during test (e.g., a tactical fix is not a permanent design change).

## U

- **Unobserved mode** – a failure mode which exhibits zero failures during testing.
- **Utility** – the number of systems expected to perform successfully in the field after deployment of a single lot. The size, or number of units, in a lot varies by system.

## W

- **Weibull process** – a non-homogeneous Poisson process with a power-law mean value function. A Weibull process is also referred to as a power-law process.

THIS PAGE IS INTENTIONALLY LEFT BLANK.

# - REFERENCES -


[1] H. K. Weiss, "Estimation of Reliability Growth in a Complex System with Poisson-Type Failure," *Operations Research*, vol. 4, num. 5, pp. 532-545,1956.

[2] M. Aroef, "Study of Learning Curves of Industrial Manual Operations," Unpublished Master's Thesis, *Cornell University*, 1957.

[3] N. Rosner, "System Analysis - Nonlinear Estimation Techniques," *Proceeding of the National Symposium on Reliability and Quality Control*, 1961, pp. 203-207.

[4] D. K. Lloyd and M. Lipow, "Reliability Growth Models," *Reliability: Management, Methods, and Mathematics, Prentice-Hall Space Technology Series*, Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1962, pp. 330-348.

[5] H. Chernoff and W. M. Woods, "Reliability growth models-Analysis and applications," *CEIR*, Los Altos, CA, Field Memo, Feb. 1962.

[6] W. Wolman, "Problems in System Reliability Analysis," *Statistical Theory of Reliability*, Marvin Zelen Ed., Madison, Wisconsin: University of Wisconsin Press, 1963, pp. 149-160.

[7] J. T. Duane, "Learning Curve Approach To Reliability Monitoring," *IEEE Transactions on Aerospace*, vol. 2, pp. 563-566, 1964.

[8] W. J. Corcoran, H. Weingarten and P. W. Zehna, "Estimating Reliability after Corrective Action," *Management Science*, vol. 10, num. 4, pp. 786-798, 1964.

[9] J. E. Bresenham, "Reliability growth models," Dept of Statistics, *Stanford University*, TR 74, Contract Number NONR225 53, AD-605 993, 1964.

[10] NASA, "Reliability growth of US rockets," *Space Technology Labs Inc*, vol. II, Contract NAS 8-11037, 1964. (SECRET).

[11] D. R. Cox and P. A. Lewis, *The Statistical Analysis of Series of Events*. New York: NY: John Wiley and Sons, 1966.

[12] R. E. Barlow and E. M. Scheuer, "Reliability Growth during a Development Testing Program," *Technometrics*, vol. 8, num. 1, pp. 53-60, 1966.

[13] W. J. Corcoran and R. R. Read, "Modeling the Growth of Reliability - Extended Model," *United Technology Center*, Sunnydale, CA, UTC 2140-ITR, 1966.

[14] W.J. Corcoran and R.R. Read, "Comparison of some reliability growth estimation and prediction schemes," *United Technology Center*, Sunnydale, CA, UTC 2140-ITR, 1967.

[15] A. J. Gross, M. Kamins, “Reliability assessment in the presence of reliability growth,” *RAND Corporation*, Memo RM-5346-PR, Sep. 1967. Also *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1968, pp. 406-416.

[16] E. P. Virene, "Reliability Growth and the Upper Reliability Limit, How to Use Available Data to Calculate," *Boeing Company*, Seattle, Washington, Report No. D2-24142-1, , AD-804 150, pp. 33, 1966. Also "Reliability Growth and Its Upper Limit," *IEEE Proceedings of the Annual Symposium on Reliability*, 1968, pp. 265-270.

[17] J. Cozzolino, “Probabilistic Models of Decreasing Failure Rate Processes,” *Naval Research Logistics Quarterly*, vol. 15, 1968, pp. 361-374.

[18] E. O. Codier, "Reliability Growth in Real Life," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1968, pp. 458-469.

[19] S. M. Pollock, “A Bayesian reliability growth model,” *IEEE Trans. Reliability*, vol. R-17, num. 4, pp. 187-198, Dec. 1968.

[20] J. D. Selby and S. G. Miller, "Reliability Planning and Management (RPM)," *Proceedings ASQC/SRE Seminar*, Milwaukee, WS: ASQC, 1970.

[21] D. R. Barr, “A class of general reliability growth models,” *Operations Research*, vol. 18, num. 1, pp. 52-65, 1970.

[22] R. R. Read, “A remark on the Barlow-Scheuer reliability growth estimation scheme,” *Technometrics*, vol. 13, num. 1, pp. 199-200, 1971.

[23] L. H. Crow, “Reliability Growth Modeling,” *AMSAA*, Aberdeen Proving Ground, MD, Tech. Rep. TR-55, 1972.

[24] L. H. Crow, “AMSAA Reliability Growth Symposium,” *AMSAA*, Aberdeen Proving Ground, MD, Interim Note 22, 1972.

[25] L. H. Crow, "Reliability Analysis for Complex, Repairable Systems in Reliability and Biometry," *SIAM*, Proschan and R. J. Serfling Eds., Philadelphia, PA, 1974, pp. 379-410.

[26] L. H. Crow, “Reliability Analysis for Complex, Repairable Systems,” *AMSAA*, Aberdeen Proving Ground, MD, Interim Note 26, 1974.

[27] L. H. Crow, “On Reliability Growth Modeling,” *AMSAA*, Aberdeen Proving Ground, MD, Interim Note 27, 1974.

[28] L. H. Crow, “On Reliability Growth Tracking,” *AMSAA*, Aberdeen Proving Ground, MD, Interim Note R-30, 1974.

[29] L. H. Crow, “Reliability Growth Estimation from Failure and Time Truncated Testing,” *AMSAA*, Aberdeen Proving Ground, MD, Interim Note R-31, 1974.

[30] L. H. Crow, "On Tracking Reliability Growth," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1975, pp. 438-443.

[31] Rome Air Development Center, "Reliability Growth Study," *Air Force Systems Command*, Griffiss Air Force Base, NY, RADC-TR-75-253, 1975.

[32] L. H. Crow, “Reliability Analysis of Complex Repairable Systems,” *AMSAA*, Aberdeen Proving Ground, MD, Tech. Rep. 138, 1975.

[33] P. Lewis and G. Shedler, "Statistical Analysis of Non-Stationary Series of Events," *IBM Journal of Research and Development*, vol. 20, pp. 465-482, 1976.

[34] T. Jayachadran and L. R. Moore, “A comparison of reliability growth models,” *IEEE Trans. Reliability*, vol. R-25, pp. 49-51, 1976.

[35] H. Braun and J. M. Paine, “A comparative study of models for reliability growth,” Dept. of Statistics, *Princeton University*, Tech. Rep. 126, ser 2, Jul. 1977.

[36] D. E. Olsen, “Estimating reliability growth,” *IEEE Trans. Reliability*, vol. R-26, pp. 50-53, Apr. 1977.

[37] R. C. Dahiya, “Estimating reliability after corrective action,” *IEEE Trans. Reliability*, vol. R-26, pp. 348-351, Dec. 1977.

[38] A. F. M. Smith, “A Bayesian note on reliability growth during a developmental testing program,” *IEEE Trans. Reliability*, vol. R-26 pp. 346-347, Dec. 1977.

[39] A. J. Bonis, “Reliability growth curves for one-shot devices,” *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1977, pp. 181-185.

[40] L. H. Crow, “Confidence Interval Procedures for Reliability Growth Analysis,” *AMSAA*, Aberdeen Proving Ground, MD, Tech. Rep. 197, 1977.

[41] H. S. Balaban, “Reliability growth models,” *J. Environ. Sciences*, vol. 21, pp. 11-18, 1978.

[42] D.E. Olsen, "A Confidence Interval for the Barlow-Scheuer Reliability Growth Model," *IEEE Trans. Reliability*, vol. R-27, pp. 308-310, Dec. 1978.

[43] M. C. Weinrich and A. J. Gross, "The Barlow-Scheuer Reliability Growth Models from a Bayesian Viewpoint," *Technometrics*, vol. 20, num. 3, pp. 249-254, 1978.

[44] G. L. Sriwastav, "A Reliability Growth Model," *IEEE Trans. Reliability*, vol. R-27, pp. 306-307, Dec. 1978.

[45] N. D. Singpurwalla, "Estimating reliability growth (or deterioration) using time series analysis," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1978, pp. 111-114.

[46] W. M. Woods, "Reliability growth models," *Avionics Reliability, Its Techniques and Related Disciplines*, num. 261M. C. Jacobsen, ED, NATO, Seine; Advisory Group for Aerospace Research & Development Conf. Proc, 1979.

[47] N. A. Langberg and F. Proschan, "A Reliability Growth Model Involving Dependent Components," *The Annals of Probability*, vol. 7, num. 6, pp. 1082-1087, 1979.

[48] B. S. Dhillon, "Reliability growth: A survey," *Microelectronics & Reliability*, vol. 20, pp. 743-751, 1980.

[49] U.S. Department of Defense, Reliability Growth Management, "Military Handbook 189," *Naval Publications and Forms Center*, Philadelphia, 1981.

[50] C. M. Stein, "Estimation of the Mean of a Multivariate Normal Distribution," *The Annals of Statistics*, vol. 9, pp. 1135-1151, 1981.

[51] B. Littlewood, "Stochastic Reliability-Growth: A Model for Fault-Removal in Computer-Programs and Hardware-Designs," *IEEE Trans. Reliability*, vol. R-30, num. 4, 1981.

[52] W. Nelson, *Applied Life Data Analysis.* New York, NY: John Wiley & Sons, 1982.

[53] L. H. Crow, "Confidence Interval Procedures for the Weibull Process with Applications to Reliability Growth," *Technometrics*, vol. 24, num.1, pp. 67-72, 1982.

[54] L. H. Crow, "An Improved Methodology for Reliability Growth Projection," *AMSAA*, Aberdeen Proving Ground, MD, Tech. Rep. TR-357, 1982.

[55] L. H. Crow, "AMSAA Discrete Reliability Growth Model," *AMSAA*, Aberdeen Proving Ground, MD, Office Note 1-83, 1983.

[56] L. H. Crow, "Reliability Growth Projection from Delayed Fixes," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1983, pp. 84-89.

[57] L. H. Crow, "Methods for Reliability Growth Assessment During Development," *In Electronic Systems Effectiveness and Life Cycle Testing*, J. K. Skwirzynski Ed.: Springer-Verlag, 1983.

[58] N. S. Fard and D. L. Dietrich, "Comparison of Attribute Reliability Growth Models," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1983, pp. 24-29.

[59] J. M. Finkelstein, "A logarithmic reliability-growth model for single mission systems," *IEEE Trans. Reliability*, vol. R-32, pp. 508-511, Dec. 1983.

[60] P. B. Trapnell, "Study on Reliability Fix Effectiveness Factors for Army Systems," *AMSAA*, Aberdeen Proving Ground, MD, Tech. Rep. 388, 1984.

[61] P. B. Trapnell, "Update of Reliability Fix Effectiveness Factors for Army Systems," *AMSAA*, Aberdeen Proving Ground, MD, Interim Note R-86, 1984.

[62] S. Jiang, D. B. Kececioglu and P. Vassiliou, "Modified Gompertz," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1994.

[63] L. H. Crow, "Methods for Assessing Reliability Growth Potential," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1984, pp. 484-489.

[64] "Reliability growth testing effectiveness," Rome Air Development Center Report, RADC-TR-84-20, Jan. 1984.

[65] W. S. Jewell, "A General Framework for Learning Curve Reliability Growth Models," *Operations Research*, vol. 32, num. 3, pp. 547-558, 1984.

[66] D. K. Lloyd, "Forecasting Reliability Growth," *IEEE Quality and Reliability Engineering International*, pp. 19-23, 1986.

[67] D. Miller, "Exponential Order Statistic Models of Software Reliability Growth," *IEEE Trans. Software Engineering*, vol. SE-12, num. 1, 1986.

[68] R. K. Gates, G. J. Gibson, K.K. McLain, "Reliability growth prediction," Rome Air Development Center Report, RADC-TR-86-148, Sept. 1986.

[69] R. B. D'Agostino and M. A. Stephens, "Tests of Chi-Squared Type," in *Goodness-of-Fit Techniques*, New York, NY: Marcel Dekker, 1986, pp. 93-93.

[70] W. P. Clay, "Maximum Likelihood Estimator for the Discrete Application of the AMSAA Growth Model," *AMSAA*, Aberdeen Proving Ground, MD, Tech. Rep. 433, 1987.

[71] Department of Defense, Reliability Test Methods, Plans, and Environments for Engineering Development, Qualification, and Production "Military Handbook 781," *Naval Publications and Forms Center, Philadelphia*, 1987.

[72] D. K. Lloyd, "Forecasting reliability growth," *Proceedings of the 33rd Annual Technical Meeting*, 1987, pp. 171-175.

[73] N. S. Fard, D. L. Dietrich, "A Bayes reliability growth model for a developmental testing program," *IEEE Trans. Reliability*, vol. R-36, pp. 568-572, Dec. 1987.

[74] L. H. Crow, "Reliability Growth Estimation With Missing Data - II," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1988, pp. 248-253.

[75] W. M. Woods, J. E. Drake, J.D. Chandler Jr., "Analysis and evaluation of discrete reliability growth models with and without failure discounting," *U.S. Naval Postgraduate School*, Tech. Rep. NPS55-88-013, Dec. 1988.

[76] P. M. Ellner and T. Ziad, "Statistical Precision and Robustness of the AMSAA Continuous Reliability Growth Estimators," *AMSAA*, Aberdeen Proving Ground, MD, Tech. Rep. 453, 1988. Also *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1990, pp. 210-214.

[77] J. C. Wronka, "Tracking of Reliability Growth in Early Development," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1988, pp. 238-242.

[78] D. Robinson and Duane Dietrich, "A System-Level Reliability Growth Model," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1988, pp. 243-247.

[79] K. L. Wong, "Reliability Growth - Myth or Mess," *IEEE Trans. Reliability*, vol. 37, num. 2, p. 209, 1988.

[80] H. A. Malec, "Reliability Growth from the Customer's Perspective," *IEEE Journal on Selected Areas of Communications*, vol. 6, num. 8, pp. 1287-1293, 1988.

[81] A. W. Benton, "Integrated Reliability Growth Testing," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1989, pp. 160-166.

[82] D. G. Frank, "A corollary to: Duane's postulate on reliability growth," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1989, pp. 167-170.

[83] G. J. Gibson and L. H. Crow, "Reliability Fix Effectiveness Factor Estimation," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1989, pp. 171-177.

[84] G. K. Bhattacharyya, A. Fries, R. A. Johnson, "Properties of continuous analog estimators for a discrete reliability-growth model," *IEEE Trans. Reliability*, vol. 38, pp. 373-378, Aug. 1989.

[85] D. Robinson and Duane Dietrich, "A Nonparametric-Bayes Reliability Growth Model," *IEEE Trans. Reliability*, vol. 38, num. 5, pp. 591-598, 1989.

[86] S. E. Rigdon, "Testing goodness-of-fit for the power law process," *Communications in Statistics: Theory and Methods*, vol. 18, pp. 4665-4676, 1989.

[87] L. H. Crow, "Evaluating the Reliability of Repairable Systems," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1990, pp. 275-279.

[88] W. M. Woods, "The effect of discounting failures and weighting data on the accuracy of some reliability growth models," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1990, pp. 200-204.

[89] P. M. Ellner and B. Trapnell, "AMSAA Reliability Growth Data Study," *AMSAA*, Aberdeen Proving Ground, MD, Interim Note R-184, 1990.

[90] S. Kaplan, G. D. M. Cunha, A. A. Dykes and D. Shaver, "A Bayesian Methodology for Assessing Reliability during Product Development," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1990, pp. 205-209.

[91] P. E. Higgins and A. Constantinides, "Reliability Growth for Expendable Mobile ASW Training Target (EMATT): A 1-Shot Device," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1991, pp. 497-501.

[92] T. A. Mazzuchi and R. Soyer, "A Bayesian Attribute Reliability Growth Model," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1991, pp. 322-325.

[93] L. H. Crow, "New International Standards on Reliability Growth," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1991, pp. 478-480.

[94] G. Coolas, "Dynamic Reliability Prediction: How to Adjust Modeling and Reliability Growth?," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1991, pp. 301-306.

[95] J. Bieda, "Reliability Growth Test Management in a Product Assurance Environment within the Automotive Component Industry," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1991, pp. 317-321.

[96] Kececioglu, Dimitri, B., "Reliability Growth," *Reliability Engineering Handbook*, 4$^{th}$ ed., vol. 2, Englewood Cliffs, New Jersey: Prentice-Hall, 1991, pp. 415-418.

[97] H. F. Martz and R. A. Waller, "Selecting Beta Prior Distributions based on Observed Binomial Sampling Data," *Bayesian Reliability Analysis*, 2nd ed., Malabar, FL: Krieger, 1991, pp. 312-318.

[98] G. K. Bhattacharyya, J. K. Ghosh, "Estimation parameters in a discrete reliability growth model," *Journal for Statistics Planning and Inference*, vol. 29, pp. 43-53, 1991.

[99] R. A. Johnson, "Confidence estimation with discrete reliability growth models," *Journal for Statistics Planning and Inference*, vol. 29, pp. 87-97, 1991.

[100] P. M. Ellner and R. Mioduski, "Operating Characteristic Analysis for Reliability Growth Programs," *AMSAA*, Aberdeen Proving Ground, MD, Tech. Rep. 524, 1992.

[101] M. A. Meth, "Reliability-growth myths and methodologies: A critical view," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1992, pp. 337-342.

[102] P. M. Ellner, D. E. Mortin and D. M. Querido, "Developing a Subsystem Reliability Growth Program using the Subsystem Planning Model," *AMSAA*, Aberdeen Proving Ground, MD, Division Note R-24, 1992.

[103] R. Calabria, M Guida and G. Pulcini, "A Bayes Procedure for Estimation of Current System Reliability," *IEEE Trans. Reliability*, vol. 41, num. 4, pp. 616-620, 1992.

[104] K. E. Ellis, "Robustness of Reliability Growth Analysis Techniques," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1992, pp. 303-315.

[105] D. I. Heimann and W. D. Clark, "Process-Related Reliability Growth Modeling - How and Why," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1992, pp. 316-321.

[106] T. A. Mazzuchi and R. Soyer, "Reliability Assessment and Prediction during Product Development," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1992, pp. 468-474.

[107] W. J. Park and Y. G. Kim, "Goodness-of-fit test for the power-law process," *IEEE Trans. Reliability*, vol. 41, pp. 107-111, 1992.

[108] B. Klefsjö and U. Kumar, "Goodness-of-Fit Test for the Power-Law Process based on the TTT-Plot," *IEEE Trans. Reliability*, vol. 41, pp. 593-598, 1992.

[109] L. H. Crow, "Confidence Intervals on the Reliability of Repairable Systems," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1993, pp. 126-134.

[110] T. A. Mazzuchi, R. Soyer, "A Bayes Method for Assessing Product Reliability during Development Testing," *IEEE Trans. Reliability*, vol. 42, pp. 503-510, Sep. 1993.

[111] A. Fries, "Discrete Reliability-Growth Models on Learning-Curve Property," *IEEE Trans. Reliability*, vol. 42, pp. 303-306, Jun. 1993.

[112] E. Demko, "One Non-linear Reliability Growth," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1993, pp. 411-416.

[113] P. E. Higgins, "Final Reliability Growth Results for EMATT, A 1-Shot Device," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1993, pp. 399-403.

[114] M. Xie and M. Zhao, "On Some Reliability Growth Models with Graphical Interpretations," *Microelectronics and Reliability*, vol. 33, pp. 149-167, 1993.

[115] D. Kececioglu, S. Jiang, P. Vassiliou, "The Modified-Gompertz Reliability Growth Model," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1994, pp. 160-165.

[116] A. Fries, "Correction- Discrete Reliability Growth Models based on a Learning-Curve Property," *IEEE Trans. Reliability*, vol. 43, p.475, Sep. 1994.

[117] P. M. Ellner, M. J. McCarthy, D. E. Mortin and D. M. Querido, "Developing a Subsystem Reliability Growth Program using SSPLAN," *AMSAA*, Aberdeen Proving Ground, MD, Tech. Rep. 555, 1994.

[118] L. H. Crow, P. H. Franklin and N. B. Robbins, "Principles of Successful Reliability Growth Applications," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1994, pp. 157-159.

[119] W. J. Park and M. Seoh, “More Goodness-of-Fit Tests for the Power Law Process,” *IEEE Trans. Reliability*, vol. 43, pp. 275-278, 1994.

[120] International Electrotechnical Commission “Reliability Growth - Statistical test and estimation methods,” *IEC International Standard*, IEC 1164, 1995.

[121] P. M. Ellner and L. C. Wald, “AMSAA Maturity Projection Model,” *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1995, pp. 174-181.

[122] I. Fakhre-Zakeri and E. Slud, “Mixture Models for Reliability of Software with Imperfect Debugging: Identifiability of Parameters,” *IEEE Trans. Reliability*, vol. R-44, num. 1, pp. 104-113, 1995.

[123] E. Demko, "On Reliability Growth Testing," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1995, pp. 162-165.

[124] K. J. Farquhar and A. Mosleh, "An Approach to Quantifying Reliability Growth Effectiveness," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1995, pp. 166-173.

[125] A. Fries and L. M. Maillart, "TAAF (Test-Analyze-And-Fix) Stopping Rules for Maximizing the Operational Utility of One-Shot Systems," *Institute of Defense Analysis*, Arlington, VA, IDA Document D-1880, 1996.

[126] W. J. Broemm, “Subsystem Tracking Model (SSTRACK),” *AMSAA*, Aberdeen Proving Ground, MD, Briefing Chart Rep., 1996.

[127] A. Fries and A. Sen, “A Survey of Discrete Reliability Growth Models,” *IEEE Trans. Reliability*, vol. R-45, num. 4, pp. 582-604, 1996.

[128] R. Calabria, M. Guida and G. Pulcini, "A Reliability Growth Model in a Bayes-Decision Framework," *IEEE Trans. Reliability*, vol. 45, num. 3, pp. 505-510, 1996.

[129] N. Ebrahimi, "How to Model Reliability Growth when Times of Design Modifications are Known," *IEEE Trans. Reliability*, vol. 45, num. 1, pp. 54-58, 1996.

[130] M. Y. Huang, D. McBeth and S. B. Vardeman, "Developmental Test Programs for 1 Shot Systems: 2-State Reliability and Binary Development-Test Results," *IEEE Trans. Reliability*, vol. 45, num. 3, pp. 379-385, 1996.

[131] M. Xie and M. Zhao, "Reliability Growth Plot - An Underutilized Tool in Reliability Analysis," *Microelectronics Reliability*, vol. 36, num. 6, pp. 797-805, 1996.

[132] R. D. Baker, “Some new tests of the power law process,” *Technometrics*, vol. 38, pp. 256-265, 1996.

[133] D. P. Gaver, and P. A. Jacobs, "Testing or fault-finding for reliability growth: A missile destructive-test example," *Naval Research Logistics*, vol. 44, pp. 623-637, 1997.

[134] E. A. Seglie, "Reliability Growth Programs from the Operational Test and Evaluation Perspective," *Reliability Growth Modeling: Objectives, Expectations, and Approaches*, K. J. Farquar and A. Mosleh Eds., College Park, MD: The Center for Reliability Engineering, University of Maryland, 1998.

[135] A. Erkanli, T. A. Mazzuchi and R. Soyer, "Bayesian Computations for a Class of Reliability Growth Models," *Technometrics*, vol. 40, num. 1, pp. 14-23, 1998.

[136] A. Sen, "Estimation of Current Reliability in a Duane-based Reliability Growth Model," *Technometrics*, vol. 40, num. 4, pp. 334-344, 1998.

[137] E. Crétois and O. Gaudoin, “New Results on Goodness-of-Fit Tests for the Power-Law Process and Application to Software Reliability,” *International Journal of Reliability, Quality & Safety Engineering*, vol. 5, pp. 249-267, 1998.

[138] O. Gaudoin, “CPIT Goodness-of-Fit Test for the Power-Law Process,” *Communications in Statistics: Theory and Methods*, vol. 27, pp. 165-180, 1998.

[139] J. A. Clark, "Modeling Reliability Growth Late in Development," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 1999, pp. 201-207.

[140] A. Sen, "Correction to: Estimation of Current Reliability in a Duane-based Reliability Growth Model," *Technometrics*, vol. 41, num. 4, p. 385, 1999.

[141] J. Donovan and D. Murphy, "Reliability Growth - A New Graphical Example," *Quality and Reliability Engineering International*, vol. 15, pp. 167-174, 1999.

[142] E. Crétois, M. A. El Aroui and O. Gaudoin, “U-Plot Method for Testing the Goodness-of-Fit of the Power-Law Process,” *Communications in Statistics: Theory and Methods*, vol. 28, pp. 1731-1747, 1999.

[143] J. Donovan and D. Murphy, "A New Reliability Growth Model: Its Mathematical Comparison to the Duane Model," *Microelectronics Reliability*, vol. 40, pp. 533-539, 2000.

[144] P. M. Ellner, W. J. Broemm and W. J. Woodworth, "Reliability Growth Guide," *AMSAA*, Aberdeen Proving Ground, MD, Tech. Rep. TR-652, 2000.

[145] J. D. Hall and W. R. Wessels, "An Automated Solver to Determine ML Estimates for the AMSAA Discrete Reliability Growth Model," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 2001, pp. 290-295.

[146] J. Donovan and D. Murphy, "Improvements in Reliability Growth Modeling," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 2001, pp. 296-301.

[147] G. Pulcini, "An Exponential Reliability Growth Model in Multi-copy Testing Program," *IEEE Trans. Reliability*, vol. 50, num. 4, pp. 365-373, 2001.

[148] R. Hodge, J. Quigley, I. James and J. Marshall, "Integrating Reliability Improvement Modeling into Practice-Challenges and Pitfalls," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 2002, pp. 158-164.

[149] G. Pulcini, "Correction to: How to Model Reliability Growth when Times of Design Modifications are Known," *IEEE Trans. Reliability*, vol. 51, num. 2, pp. 252-253, 2002.

[150] L. H. Crow, "Methods for Reducing the Cost to Maintain a Fleet of Repairable Systems," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 2003, pp. 392-399.

[151] T. Gurunatha and R. P. Siegel, "Applying Quality Tools to Reliability: A 12-Step Six-Sigma Process to Accelerate Reliability Growth in Product Design," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 2003, pp. 562-567.

[152] O. P. Yadav, N Singh and P. S. Goel, "A Practical Approach to System Reliability Growth Modeling and Improvement," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 2003, pp. 351-359.

[153] J. Quigley and L. Walls, "Confidence Intervals for Reliability Growth Models with Small Sample Sizes," *IEEE Trans. Reliability*, vol. 52, num. 2, pp. 257-262, 2003.

[154] O. Gaudoin, B. Yang and M. Xie, "A Simple Goodness-of-Fit Test for the Power-Law Process, based on the Duane Plot," *IEEE Trans. Reliability*, vol. 52, num. 1, pp. 69-74, 2003.

[155] D. P. Gaver, P. A. Jacobs, K. D. Glazenbrook and E. A. Seglie, "Probability Models for Sequential-Stage System Reliability Growth via Failure Mode Removal," *International Journal of Reliability, Quality & Safety Engineering*, vol. 10, issue 1, pp. 15-40, 2003.

[156] P. M. Ellner and J. B. Hall, “AMPM-Stein,” *AMSAA*, Aberdeen Proving Ground, MD, Tech. Rep. 751, 2004.

[157] L. H. Crow, "An Extended Reliability Growth Model for Managing and Assessing Corrective Actions," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 2004, pp. 73-80.

[158] T. C. Smith, "Reliability Growth Planning under Performance based Logistics," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 2004, pp. 418-423.

[159] M. Krasich and J. Quigley, "Modeling Reliability Growth in the Product Design Process," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 2004, pp. 424-430.

[160] C. E. Ebeling, *An Introduction to Reliability and Maintainability Engineering*. Illinois: Waveland Press Inc., 2005.

[161] D. E. Mortin and P.M. Ellner, “Reliability Growth for Military Vehicles – Emerging Methodology and Remaining Challenges,” *Journal of Commercial Vehicles*, vol. 114, pp.28-34, 2005.

[162] P. M. Ellner and J. B. Hall, “AMSAA Maturity Projection Model based on Stein Estimation,” *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 2005, pp. 271-277.

[163] W. J. Broemm, "User Guide for the Visual Growth Suite of Reliability Growth Models," AMSAA, Aberdeen Proving Ground, MD, Tech. Rep. 764, 2005.

[164] Department of Defense, “DoD Guide for Achieving Reliability, Availability and Maintainability,” Washington, DC, 2005.

[165] P. E. Acevedo, D. S. Jackson and R. W. Kotlowitz, "Reliability Growth and Forecasting for Critical Hardware through Accelerated Life Testing," *Bell Labs Technical Journal*, vol. 11, num. 3, pp. 121-135, 2006.

[166] P. M. Ellner and J. B. Hall, "An Approach to Reliability Growth Planning based on Failure Mode Discovery and Correction using AMSAA Projection Methodology," *IEEE Proceedings of the Annual Reliability and Maintainability Symposium*, 2005, pp. 271-277.

[167] A. M. Sarhan, F. M. Guess and J. S. Usher, "Estimators for Reliability Measures in Geometric Distribution Model Using Dependent Masked System Life Test Data," *IEEE Trans. Reliability*, vol. R-56, num. 2, 2007.

[168] "The Effects of Aging on Costs of Operating and Maintaining Military Equipment, *U.S. Congressional Budget Office*, 2001.

[169] *IEC Standards for Reliability Growth-Statistical Test and Estimation Methods*, IEC Standard 61164 II, 2004.

[170] S.L. Jeng, and W.Q. Meeker, "Comparisons of Approximate Confidence Interval Procedures for Type I Censored Data," *Technometrics*, vol. 42, no. 2, pp. 135-148, 2000.

[171] Y. Hong, W.Q. Meeker, and L.A. Escobar, "Avoiding Problems with Normal Approximation Confidence Intervals for Probabilities," *Technometrics*, vol. 50, no. 1, pp. 64-68, 2008.

[172] L. Lee, S. K. Lee, "Some Results on Inference for the Weibull Process," *Technometrics*, vol. 20, no. 1, pp. 41-45, 1978.

[173] M. Engelhardt, and L. J. Bain, "Prediction Intervals for the Weibull Process," *Technometrics*, vol. 20, no. 2, pp. 167-169, 1978.

[174] L. J. Bain, and M. Engelhardt, "Inferences on the Parameters and Current System Reliability for a Time Truncated Weibull Process," *Technometrics*, vol. 22, no. 3, pp. 421-426, 1980.

[175] L. Lee, "Comparing Rates of Several Independent Weibull Processes," *Technometrics*, vol. 22, no. 3, pp. 427-430, 1980.