ILLEGIBLE DOCUMENT

THE FOLLOWING DOCUMENT(S) IS OF POOR LEGIBILITY IN THE ORIGINAL

THIS IS THE BEST COPY AVAILABLE

ANALYSIS OF COMPONENT FAILURE DATA BY

NON-CONJUGATE COMPOUND FAILURE MODELS

by

JEFFREY H. SIMMONS

B.S., Kansas State University, 1981

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Nuclear Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1983

Approved by:

J K Shultis

Major Professor

TABLE OF CONTENTS

Page

LIST OF FIGURES . . . .

LIST OF TABLES . . . .

1.0 INTRODUCTION 1

2.0 FAILURE MODELS 5
2.1 HOMOGENEOUS MODEL 6
2.2 COMPOUND MODEL 7
2.3 CLASSICAL AND BAYESIAN APPLICATION OF THE COMPOUND MODEL 9

3.0 THE PRIOR DISTRIBUTION 11
3.1 *CHOICE OF PRIOR DISTRIBUTION* 11
3.2 PROPERTIES OF VARIOUS PRIOR FAMILIES 13
3.2.1 Properties of Prior Families Selected for Analyses 13
3.3 EVALUATION OF THE MARGINAL AND POSTERIOR DISTRIBUTIONS FOR THE COMPOUND MODEL 26
3.3.1 Explicit Expresions for $I(\lambda)$ 27

4.0 ESTIMATION OF PRIOR PARAMETERS FOR THE COMPOUND MODEL 36
4.1 MATCHING MOMENTS METHODS 37
4.1.1 Prior Matching Moments Method (PMMM) 37
4.1.2 Marginal Matching Moments Method (MMMM) 39
4.1.3 Matching Moment Estimators for Particular Prior Distributions 40
4.2 MARGINAL MAXIMUM LIKELIHOOD METHOD (MMLM) 44

5.0 SAFER COMPUTER CODE 46
5.1 DESCRIPTION 46
5.2 MODIFICATIONS 47
5.2.1 Specific Changes Made in SAFER 48

6.0 ANALYSIS OF DIFFERENT PRIOR DISTRIBUTIONS IN THE COMPOUND MODEL 51
6.1 CONJUGATE ANALYSIS OF NON-CONJUGATE FAILURE DATA 52
6.1.1 Data Simulation Procedure 53
6.1.2 Results of Using a Conjugate Model to Analyze Data from Non-Conjugate Models 53
6.2 NON-CONJUGATE ANALYSIS OF DATA FROM NON-CONJUGATE MODELS 61

6.3 COMPARISON OF USING SIX DIFFERENT PRIOR DISTRIBUTIONS WITH THE SAME FAILURE DATA 70

7.0 REFERENCES 88

LIST OF FIGURES

Figure page

3.1. The gamma density function and its variation with its parameters 15

3.2. The Weibull density function and its variation with its parameters 17

3.3. The shifted lognormal density function and its variation with its parameters 19

3.4. The logbeta density function and its variation with its parameters 21

3.5. The loggamma density function and its variation with its parameters 23

3.6. The logWeibull density function and its variation with its parameters 25

6.1. Estimated prior distributions obtained by the PMMM with GOOD data, (a) linear (b) semi-log. The prior distributions were assumed to belong to the families: (1) Gamma, (2) Lognormal, (3) Weibull, (4) Logbeta, (5) Loggamma, and (6) LogWeibull 75

6.2. Estimated prior distributions obtained by the MMMM with GOOD data, (a) linear (b) semi-log. The prior distributions were assumed to belong to the families: (1) Gamma, (2) Lognormal, (3) Weibull, (4) Logbeta, (5) Loggamma, and (6) LogWeibull 76

6.3. Estimated prior distributions obtained by the MMLM with GOOD data, (a) linear (b) semi-log. The prior distributions were assumed to belong to the families: (1) Gamma, (2) Lognormal, (3) Weibull, (4) Logbeta, (5) Loggamma, and (6) LogWeibull 77

6.4. Estimated prior distributions obtained by the PMMM with BAD data, (a) linear (b) semi-log. The prior distributions were assumed to belong to the families: (1) Gamma, (2) Lognormal, (3) Weibull, (4) Logbeta, (5) Loggamma, and (6) LogWeibull 79

6.5. Estimated prior distributions obtained by the MMMM with BAD data, (a) linear (b) semi-log. The prior distributions were assumed to belong to the families: (1) Gamma, (2) Lognormal, (3) Weibull, (4) Logbeta, (5) Loggamma, and (6) LogWeibull 80

6.6. Estimated prior distributions obtained by the MMLM with BAD data, (a) linear (b) semi-log. The prior distributions were assumed to belong to the families: (1) Gamma, (2) Lognormal, (3) Weibull, (4) Logbeta, (5) Loggamma, and (6) LogWeibull 81

LIST OF TABLES

Table		page
6.1.	Results of a gamma-Poisson analysis of data from a loggamma-Poisson distribution	56
6.2.	Results of a gamma-Poisson analysis of data from a logWeibull-Poisson distribution	57
6.3.	Results of a gamma-Poisson analysis of data from a Weibull-Poisson distribution	58
6.4.	Results of a gamma-Poisson analysis of data from a lognormal-Poisson distribution	59
6.5.	Results of a gamma-Poisson analysis of data from a logbeta-Poisson distribution	60
6.6.	Results of a loggamma-Poisson analysis of data from a loggamma-Poisson distribution	62
6.7.	Results of a logWeibull-Poisson analysis of data from a logWeibull-Poisson distribution	63
6.8.	Results of a Weibull-Poisson analysis of data from a Weibull-Poisson distribution	64
6.9.	Results of a lognormal-Poisson analysis of data from a lognormal-Poisson distribution	65
6.10.	Results of a logbeta-Poisson analysis of data from a logbeta-Poisson distribution	66
6.11.	Parameter estimates for six prior models obtained by three estimation techniques for GOOD data	73
6.12.	Parameter estimates for six prior models obtained by three estimation techniques for BAD data	74
6.13.	Percentiles of six prior distributions obtained from three estimation methods with GOOD data	78
6.14.	Percentiles of six prior distributions obtained from three estimation methods with BAD data	82

THIS BOOK CONTAINS NUMEROUS PAGES WITH THE ORIGINAL PRINTING BEING SKEWED DIFFERENTLY FROM THE TOP OF THE PAGE TO THE BOTTOM.

THIS IS AS RECEIVED FROM THE CUSTOMER.

THIS BOOK CONTAINS NUMEROUS PAGES THAT HAVE INK SPLOTCHES IN THE MIDDLE OF THE TEXT. THIS IS AS RECEIVED FROM CUSTOMER.

THESE ARE THE BEST IMAGES AVAILABLE.

Chapter 1

INTRODUCTION

A safety analysis is an important aspect during the design, construction, and operation of a nuclear power plant. A good deal of effort is expended on this analysis to ensure the safe operation of the plant.

It is important to develop a failure model which has the ability to make future predictions about the failure behavior of systems under scrutiny. The nuclear industry is relatively new and the systems and components are designed to perform their task with high reliability. Therefore, a safety analysis is often based on data accumulated from components in power plants which exhibit inherently low failure rates.

When considering a safety analysis on any system, one must first determine a proper failure model to utilize in describing the failure characteristics of the particular components or systems. The better the agreement between a model and the past failure data for the system, the better one feels in his description of the stochastic nature of the failure mechanism, particularly when the development of the failure model is based on reasonable physical assumptions. Many different statistical failure models can be proposed based on different assumptions for the failure mechanism, type of failure, relation of one failure to another, the dependence of one part of the system on another, etc.

One of the most useful statistical models to describe the failure of components with low failure probabilities is the compound or

Bayesian model. Given a class of similar components, this model assumes the failure rate varies from component to component according to some unknown prior distribution, yet remains constant for each individual component in the class. The probability of observing F failures in test time T, assumed in this study to be given by the Poisson distribution, is then averaged over all possible failure rates described by the prior distribution to obtain the compound or <u>marginal</u> distribution.

Generally, the gamma distribution is chosen to represent the prior distribution in the compound model because of the analytical simplification afforded by its use when combined with the Poisson distribution. In particular, an associated distribution, the "posterior", becomes a gamma distribution (with different parameters) and for this reason the gamma distribution is called the "conjugate" distribution to the Poisson. While other prior families could, in principle, be used, the numerical evaluation of such non-conjugate compound models is exceptionally difficult. Although the conjugate gamma-Poisson model is mathematically convenient, it is unknown how the results of the failure analysis depend upon the <u>a priori</u> selection of the gamma family. To answer this question was one of the primary purposes of this study.

Once the functional form has been selected to represent the prior distribution, one must determine appropriate values for the prior parameters so that the resulting compound model agrees as closely as possible with observed failure data. Several techniques are available for the prior parameters. In this study, three methods were used to estimate values of the prior parameters: (1) matching the data moments to those of the prior distribution (PMMM), (2) matching the moments of the marginal distribution to those of the data (MMMM), and (3) the maximum likelihood method based on the marginal distribution (MMLM) where the parameter values are chosen so as to maximize the probability of observing the actual data.

This study is a summary of previous reports showing the use of non-conjugate prior distributions in compound failure models and extends the use of different prior family selections in compound failure models. In this study, simulated failure data of the (F,T) type were generated from alternative non-conjugate compound models. In particular, the prior distribution was taken as belonging to (1) the loggamma family and (2) the logWeibull family. In a previous study [1], the prior distribution was taken as belonging to (3) the gamma family, (4) the Weibull family, (5) the lognormal family, or (6) the logbeta family. The simulated failure data from non-conjugate models were analyzed by the assumed gamma-Poisson failure model as well as the correct failure models from which the data were simulated. The free parameters were estimated using the PMMM, MMMM, and MMLM. Numerical techniques for estimating the free parameters of these prior families were determined in a previous report [2] and have been extended in this study. Many comparisons are presented in this report which use both conjugate and non-conjugate models to fit data from non-conjugate distributions.

The main conclusion derived from both studies is that the simpler conjugate model can be used to describe failure data even when the data is known to come from a non-conjugate distribution. There appears to be no need to expend the large computational effort needed to use the non-conjugate compound models. The results of the failure analysis are not significantly affected by the selection of the gamma family. Of far greater importance in the safety analysis is the method used to estimate the parameters of the chosen prior family from the failure data. For some data sets the three parameter estimation techniques are found to give quite different results.

Chapter 2

FAILURE MODELS

In developing a failure model to analyze the safe operation of nuclear power plants, one needs to distinguish between components or systems which are normally active and those that are normally inactive. Normally active components are defined as components that are continually operating during the operation of the plant. These components include such items as cooling pumps, safety monitors, heat exchangers, and hundreds of others. The other class of components are those which are normally inactive and used only under specific conditions. Such components includes such items as backup cooling systems and standby diesel generators that mostly sit idle until a circumstance arises that warrants their use, at which time they are required to start without delay.

For normally inactive components, one may be concerned with predicting the probability that a component will fail to perform its designed task when needed. For this type of "failure-on-demand" situation, one may want to use failure models to estimate the probability that F failures will result from T demands of the component.

Failure models may also be used to describe normally active components. For example, a failure rate or a mean time to failure can be associated with components that are normally active. The failure rate being failure per unit time. The mean time to failure is the inverse of failure rate ($1/\lambda$) and gives the expected length of time for a component to operate before failing.

Consequently, to evaluate the safety of a nuclear power plant one would require at least two failure models; one to determine the probability of a failure-on-demand, the other to analyze the failure rate case. This study is concerned with the failure of components which are normally active, termed the "failure-rate" case. It is assumed that any component which fails can be rapidly repaired (or replaced) to its initial state and brought back into operation without any significant loss of operating time. In this study the failure-rate problem is treated by both "homogeneous" and "compound" models described below.

2.1 HOMOGENEOUS MODEL

The homogeneous model is the simpler of the two failure models used in this study. In this model the failure rate, λ, is assumed to be an unknown constant common to all components in a given class. In this study, it is assumed that a component which fails is immediately repaired and brought back into operation with its failure rate unchanged. Given that failures occur randomly and independent of one another, that is, the failure of one component does not directly or indirectly affect the failure rate of another component, the probability of obtaining F failures in time of operation, T, is assumed to be given by the Poisson distribution,

$$f(F|\lambda,T) = \frac{(\lambda T)^F}{\Gamma(F+1)} e^{-\lambda T} \qquad , F = 0,1,2,\ldots \tag{2.1}$$

where $\Gamma(F+1)$ is the gamma function.

The unknown failure rate λ may be estimated from the given failure data. By applying the maximum likelihood method [3] to the Poisson distribution one estimates the failure rate as the ratio of the total number of failures to the sum of the testing times, i.e.;

$$\hat{\lambda} = \sum_{i=1}^{n} F_i / \sum_{i=1}^{n} T_i \ . \tag{2.2}$$

Since components in a nuclear power plant are designed to be highly reliable and therefore to have very low failure probabilities, one often has available failure data that is termed "censored" data. Censored data occurs when a group of components are tested for an extended period of time during which a certain percentage of the components do not fail. For a component which has zero failures in test time T, the homogeneous model will predict the failure rate to be zero. This is inadmissable since no machine is totally immune to failure. To obtain a better estimate of the failure rate, similar components are often lumped together and their failure rates averaged over all the components. However, it has been observed from experience that such grouped failure data may exhibit a greater variation in the observed F_i than would be expected from the homogeneous model [4-6]. Therefore, alternatives to the homogeneous model may be required to obtain more reasonable results for components with inherent low failure rates.

2.2 COMPOUND MODEL

A second and more complex model, known as the "compound model", is better equipped to treat censored data. In this model, the failure rate (λ) is assumed to vary from component to component yet remain constant for each component in the class. The failure rate λ is thus regarded as a random variable with a distribution $g(\lambda)$. This distribution is termed the "prior" distribution since it has to be determined from previous knowledge of the components in the given class. Once $g(\lambda)$ is known, the probability that a component randomly selected from the class will experience F failures in test time T is given by the "marginal" distribution

$$h(F;T) = \int_{\text{all } \lambda} f(F|\lambda;T)\, g(\lambda)\, d\lambda\,, \quad F = 0,1,2,\ldots \tag{2.3}$$

where the "conditional" distribution f is given by the Poisson distribution of Eq. (2.1). The marginal distribution can be viewed as weighing the Poisson distribution by the information about the distribution of the failure rates contained in the prior distribution. Therefore, should each component in a class under consideration have the same failure rate λ_0, the marginal distribution of Eq. (2.3) would simplify to the homogeneous model, i.e. to the Poisson distribution of Eq. (2.1). In mathematical terms, the prior distribution for this case becomes a delta function, i.e., $g(\lambda) = (\lambda - \lambda_0)$.

Also it is of practical interest to obtain a distribution which combines the prior information $g(\lambda)$ and posterior information of the data (F,T). The estimated distribution of λ for a component randomly selected from the class described by $g(\lambda)$, and for which F failures in test time T are observed, is

$$u(\lambda|F;T) = \frac{g(\lambda)f(F|\lambda;T)}{h(F;T)}\,, \quad 0 < \lambda < \infty \tag{2.4}$$

From this "posterior" distribution for an individual component, one can make a future prediction about the behavior of a component by estimating the probability interval for the failure rate λ of the component.

2.3 CLASSICAL AND BAYESIAN APPLICATION OF THE COMPOUND MODEL

In this section, interpretation of the prior distribution is discussed as to whether it is obtained through subjective judgement or not. When developing a scientific procedure for an analysis, one should try to minimize any subjective judgement by the user. If one gives the same data to different individuals, the results or answers will differ the majority of the time unless there exists a strict procedure for each to follow thereby eliminating individual judgement.

In the Bayesian application of the compound model the prior distribution represents both the physical variation of λ among components as well as the uncertainty of the analyst about what the actual values of λ are. Before the analyst observes any data, the prior distribution provides some information about the expected failure probability of the component. The prior distribution is thus an a priori probability model for the failure probability. Since the prior distribution is based upon the analyst's prior knowledge and assumptions about a components failure probability, the prior distribution represents subjective judgement or prejudice towards a component. As such, the outcome of the analysis using the compound failure model will be dependent upon the analyst's uncertainty or prejudice for a particular component.

The classical analyst, on the other hand, wishes to avoid all subjective aspects in his probabilistic models and to use only observed failure data. By eliminating subjective judgement, the analyst hopes to remove any error introduced into the model due to extraneous decision making. Nevertheless, a classical analyst can also use prior information to improve his predictions about a particular component with a low failure probability. Rather than treat each component in isolation, he may assume (subjective?) that the component of concern belongs to a class of similar components whose constant failure probabilities are distributed according to some (prior) distribution. Each component in this class is still assumed to have a constant failure probability. One may then combine the prior information (the distribution of failure probabilities in the class to which the component of concern belongs) with the observed data for the component in question. Such was the basis for the compound model developed in the previous section.

In both the classical compound model analysis and the Bayesian analysis, information about the distribution of the failure probability concerning the component must be obtained. In the Bayesian analysis this information is obtained from one's own experience and personal judgement thereby introducing subjective judgement and possible error into the analysis. In a classical compound model, this information is given either a priori or must be established from data observed for the components in the class.

The development of the failure probability model for a particular component is identical for both analyses. In summary, the difference between the Bayesian approach and the compound model of classical analysis is the interpretation attached to the prior distribution and the methods used to estimate it.

Chapter 3

THE PRIOR DISTRIBUTION

3.1 CHOICE OF PRIOR DISTRIBUTION

Before the compound model discussed in the previous section can be used to analyze failure data from a given class of components and to make predictions about any one of the components, one needs to predetermine or estimate an appropriate prior distribution. In principle, any non-negative function with unit area could be used. However, there are criteria which the prior distribution should satisfy. In practice, one often choses a function that simplifies the marginal or posterior distributions such that they can be evaluated analytically. Other requirements of the functions used to represent the prior distribution are that it should have several free parameters whose values determine the shape and range of the distribution. The number of free parameters in the function should be kept to a minimum so as to minimize the associated problem of obtaining appropriate parameter values. By varying the values of the parameters, the prior distribution should have the ability to assume a wide variety of shapes in order to model as many component classes as possible.

Usually the gamma family is chosen as the prior distribution in the application of the compound model of Eq. (2.3). This decision is, more often than not, based on analytical convenience since the gamma family permits the marginal distribution to be evaluated explicitly and has the property of yielding another member of the gamma family for the posterior distribution. The gamma family is thus called the "conjugate" prior distribution due to this pairing of the prior and posterior distributions.

The two free parameters in the gamma family allow the functional form to assume many different shapes. In the past, it has often been assumed that the gamma family has sufficient flexibility to model the actual distribution of λ among the components in any given class. One may question this assumption based on the limitations forced on the compound model by the functional form of the gamma distribution, and wonder to what degree the results are affected by the choice of a gamma prior distribution.

In this study, alternatives to the gamma distribution are investigated for use as a prior distribution in the compound model. The non-conjugate distributions, loggamma and logWeibull, are presented in this study as possible alternatives for use as prior distributions. In order to obtain a more complete analysis, three other non-conjugate distributions, previously investigated [1], are also included. These are the lognormal, Weibull, and logbeta distributions.

3.2 PROPERTIES OF VARIOUS PRIOR FAMILIES

In this section, six different functions are presented as possible prior distributions for failure-rate analyses. A brief summary of some of the properties of each function, and figures are presented showing the possible shapes each function may assume.

3.2.1 Properties of Prior Families Selected for Analyses

GAMMA DISTRIBUTION [2]

[conjugate distribution to the conditional Poisson Distribution]

(a) Density Function:

$$g(\lambda;\alpha,\beta,\theta) = \frac{(\lambda-\theta)^{\alpha-1}\beta^{\alpha}}{\Gamma(\alpha)} e^{-\beta(\lambda-\theta)} \qquad (3.1)$$

(b) Parameters and Range:

(i) Parameters: shape $\alpha>0$; scale $\beta>0$; shift $\theta\geq0$

(ii) Range: $0 \leq \lambda_{min} = \theta \leq \lambda \leq \infty$

(c) Shape of Density Function: (see Fig. 3.1)

(i) if $\alpha>1$, $g(\lambda)$ has a mode at $\lambda_{max} = \theta + (\alpha-1)/\beta$

(ii) if $\alpha<1$, $g(\lambda)$ becomes unbounded at $\lambda_{min} = \theta$, and $g(\lambda)$ decreases monotonically.

(d) Cumulative Distribution:

$$g(\lambda;\alpha,\beta,\theta) = P(\alpha,\beta(\lambda-\theta)) \qquad (3.2)$$

where the "normalized incomplete gamma function" P is defined as

$$P(\alpha,x) \equiv \frac{1}{\Gamma(\alpha)}\int_0^x e^{-t}\, t^{\alpha-1}\, dt \qquad (3.3)$$

(e) Moments about Zero:

Use the recursion relation

$$\omega_k' \equiv E[\lambda^k] = \frac{1}{\beta^k} \prod_{i=0}^{k-1} (\alpha+i) - \sum_{i=1}^{k} (-1)^i \binom{k}{i} \theta^i \omega_{k-i}', \quad k=1,2,\ldots \qquad (3.4)$$

with $\omega_0' = 1$.

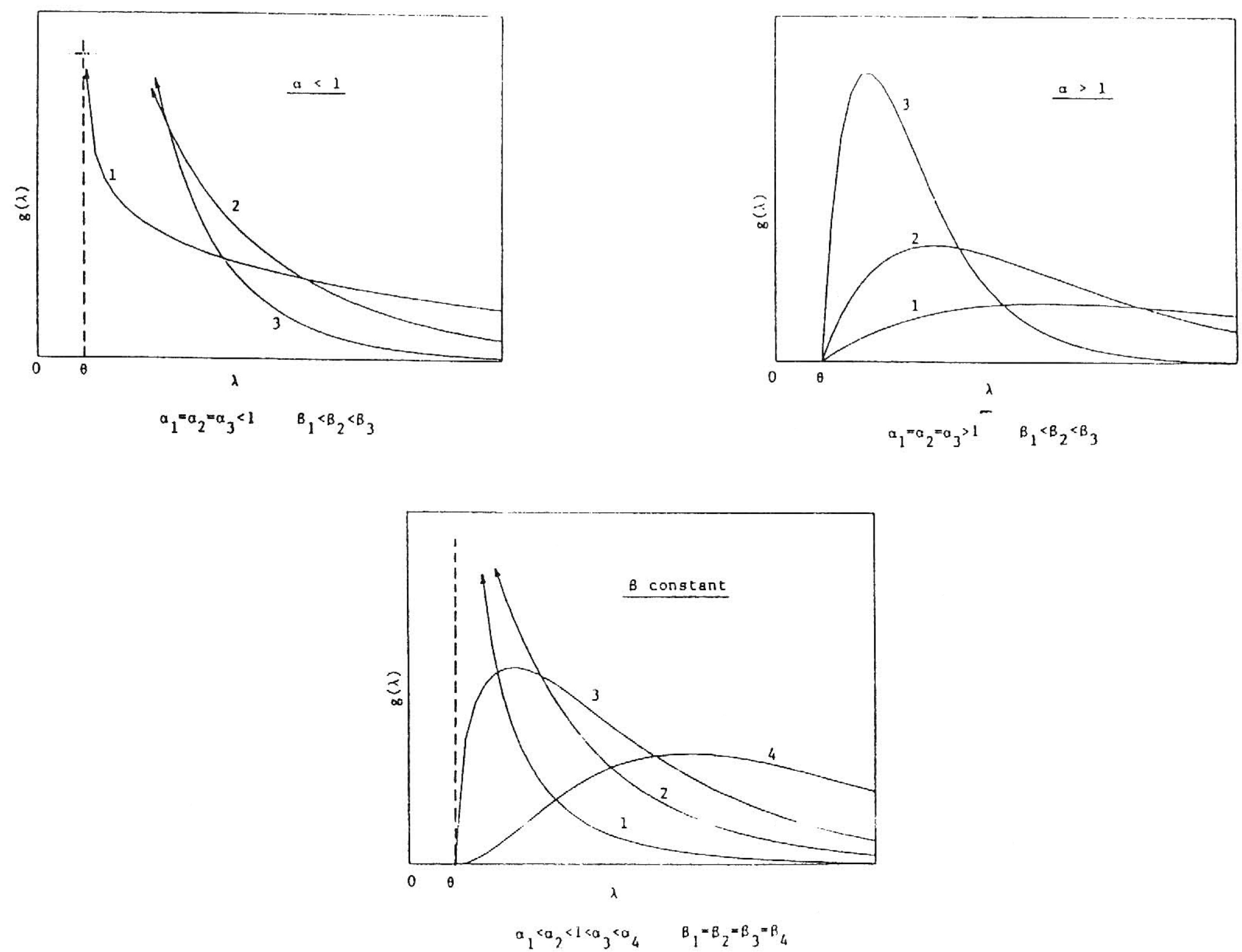

Fig. 3.1 The gamma density function and its variation with its parameters.

WEIBULL DISTRIBUTION [2]

(a) Density Function:

$$g(\lambda;\alpha,\beta,\theta) = \frac{\beta}{\alpha}\left(\frac{\lambda-\theta}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{\lambda-\theta}{\alpha}\right)^{\beta}\right) \tag{3.5}$$

(b) Parameters and Range:

(i) Parameters: shape $\beta>0$; scale $\alpha>0$; shift $\theta\geq 0$

(ii) Range: $0 \leq \lambda_{min} = \theta \leq \lambda < \infty$

(c) Shape of Density Function: (see Fig. 3.2)

(i) if $\beta>1$, $g(\lambda)$ has a mode at $\lambda_{min} = \theta + \alpha(1-\frac{1}{\beta})^{1/\beta}$

(ii) if $\beta<1$, $g(\lambda)$ decreases monotonically and becomes unbounded as λ approaches $\lambda_{min} = \theta$

(d) Cumulative Distribution:

$$G(\lambda;\alpha,\beta,\theta) = 1 - \exp\left(-\left(\frac{\lambda-\theta}{\alpha}\right)^{\beta}\right) \tag{3.6}$$

(e) Moments about Zero:

Use the recursion relation

$$\omega'_k \equiv E[\lambda^k] = \alpha^k\,\Gamma(1+\frac{k}{\beta}) - \sum_{i=1}^{k}(-1)^i\theta^i\binom{k}{i}\omega'_{k-i},\quad k=1,2\ldots \tag{3.7}$$

with $\omega'_0 = 1$.

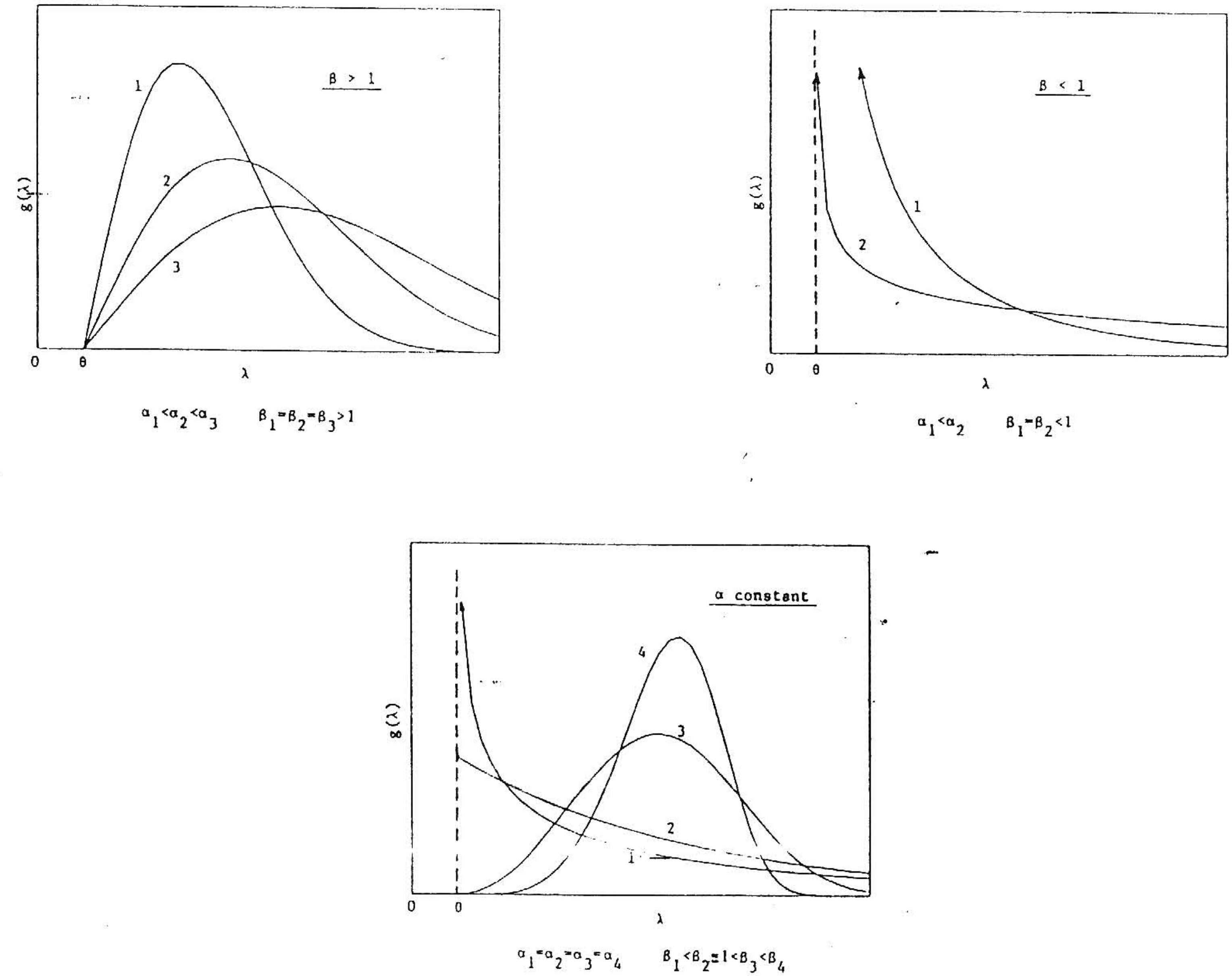

Fig. 3.2 The Weibull density function and its variation with its parameters.

LOGNORMAL DISTRIBUTION [2]

$[\ln(\lambda-\theta) \sim \text{Normal}(\alpha,\beta)]$

(a) Density Function:

$$g(\lambda;\alpha,\beta,\theta) = [(\lambda-\theta)\ \sqrt{2\pi}\beta]^{-1}\ \exp[-\ln(\lambda-\theta)-\alpha)^2/2\beta^2] \qquad (3.8)$$

(b) Parameters and Range:

(i) Parameters: $-\infty<\alpha,\beta<\infty$; shift $\theta\leq 0$

(ii) Range: $\lambda\geq\theta$

(c) Shape of Density Distribution: (see Fig. 3.3)

(i) $g(\lambda)$ has a mode at $\lambda_m = \theta + \exp[-\beta^2+\alpha]$

(ii) increasing β or decreasing α shifts $g(\lambda)$ to smaller λ-values

(d) Cumulative Distribution:

$$G(\lambda;\alpha,\beta,\theta) = \frac{1}{2}\ [1 - \text{Sign}(\lambda^*)\text{erf}(|\lambda^*|/\sqrt{2})] \qquad (3.9)$$

where $\lambda^* \equiv [-\ln(\lambda-\theta) + \alpha]/\beta$

(e) Moments about Zero:

Use the recursion relation

$$\omega'_k \equiv E[\lambda^k] = \exp[k\alpha + k^2\beta^2/2] - \sum_{i=1}^{k} (-1)^i\theta^i \binom{k}{i}\omega'_{i-1},\ k=1,2,\ldots \qquad (3.10)$$

with $\omega'_0 \equiv 1$.

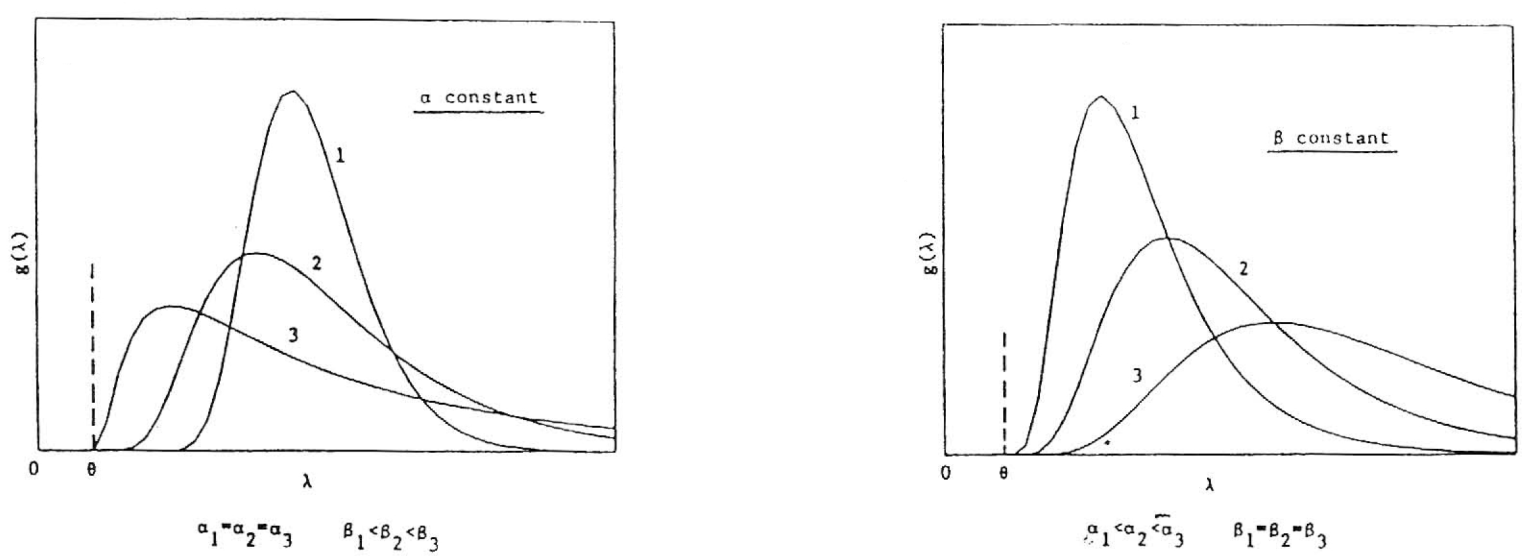

Fig. 3.3 The shifted lognormal density function and its variation with its parameters.

LOGBETA DISTRIBUTION [2]

$[\ln\lambda \sim \text{Beta}(\alpha,\beta,a,b)]$

(a) Density Function:

$$g(\lambda;\alpha,\beta,a,b) = \frac{\Gamma(\alpha+\beta)}{\lambda\Gamma(\alpha)\Gamma(\beta)} (b-a)^{1-\alpha-\beta} (\ln\lambda-a)^{\alpha-1} (b-\ln\lambda)^{\beta-1} \tag{3.11}$$

(b) Parameters and Range:

Parameters: shape $\alpha,\beta>0$; range $-\infty<a<b<\infty$

Range: $0 \leq \lambda_{min} \leq \lambda \leq \lambda_{max} < \infty$; where $\lambda_{min} = e^a$ and $\lambda_{max} = e^b$

(c) Shape of Density Distribution: (see Fig. 3.4)

(i) if $\alpha<1$, $g(\lambda)$ is unbounded at $\lambda_{min} = e^a$

(ii) if $\beta<1$, $g(\lambda)$ is unbounded at $\lambda_{max} = e^b$

(iii) if $\alpha>1$ and $\beta>1$, $g(\lambda)$ has a mode between λ_{min} and λ_{max}

(iv) if $\alpha>1$ and $\beta<1$, $g(\lambda)$ either has both a minimum and a maximum, or increases monotonically

(v) if $\alpha<1$ and $\beta>1$, $g(\lambda)$ decreases monotonically

(d) Cumulative Distribution:

$$G(\lambda;\alpha,\beta,a,b) = I_{\lambda^*}(\alpha,\beta) \tag{3.12}$$

where $\lambda^* \equiv [\ln\lambda-a]/(b-a)$ and the "incomplete beta function" is defined as

$$I_p(\alpha,\beta) \equiv \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^p x^{\alpha-1} (1-x)^{\beta-1} dx.$$

(e) Moments about Zero:

Use numerical techniques to evaluate

$$\omega_k' \equiv E[\lambda^k] = \int_{\lambda_{min}}^{\lambda_{max}} \lambda^k g(\lambda;\alpha,\beta,a,b)\, d\lambda$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\int_0^1 y^{\alpha-1}\{(1-y)^{\beta-1}[\bar{f}(y)-\bar{f}(1)y^{1-\alpha}]-\bar{f}(0)\}dy + \{\frac{\bar{f}(1)}{\beta} + \frac{\bar{f}(0)}{\alpha}\} \right) \tag{3.13}$$

where $\bar{f}(y) \equiv \exp[k\{a+(b-a)y\}]$.

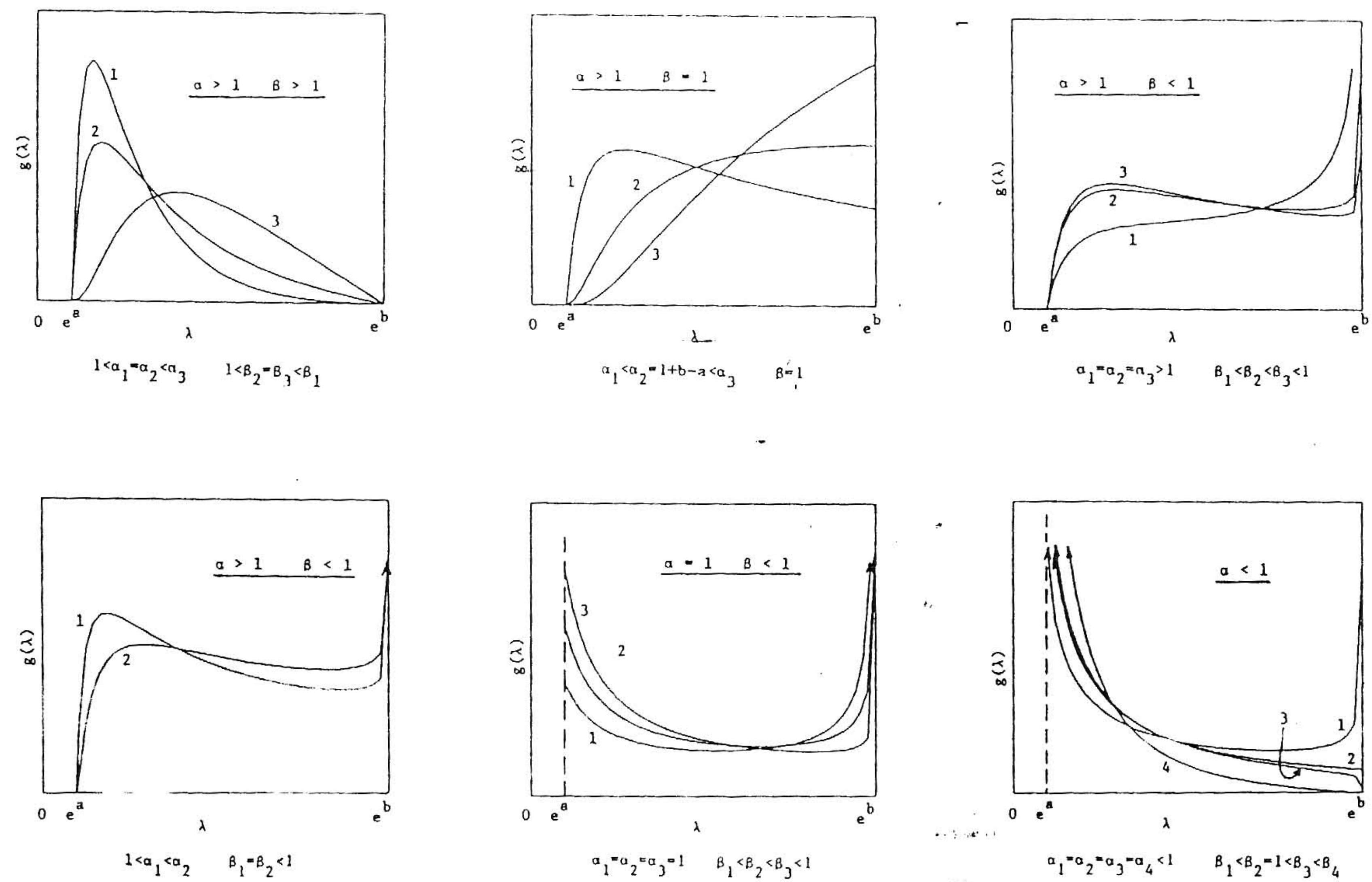

Fig. 3.4 The logbeta density function and its variation with its parameters.

LOGGAMMA DISTRIBUTION

$[-\ln\lambda \sim \text{Gamma}(\alpha,\beta,\theta)]$

(a) Density Function:

$$g(\lambda;\alpha,\beta,\theta) = [\Gamma(\alpha)\beta^{\alpha}\lambda]^{-1}\ (-\ln\lambda-\theta)^{\alpha-1}\ \exp[-(-\ln\alpha-\theta)/\beta] \qquad (3.14)$$

(b) Parameters and Range:

(i) Parameters: $\alpha,\beta,>0;\quad \theta\geq 0$

(ii) Range: $0 \leq \lambda \leq \lambda_{max}$ where $\lambda_{max} = e^{-\theta}$

(c) Shape of Density Distribution: (see Fig. 3.5)

(i) if $\alpha>1$ and $\beta<1$, have mode at $\exp[-\{\theta+(1-\alpha)/(1-\frac{1}{\beta})\}]$

(ii) if $\alpha<1$ and $\beta>1$, $g(\lambda)$ has minimum at $\exp[-\{\theta+(1-\alpha)/(1-\frac{1}{\beta})\}]$

(iii) if $\alpha<1$, $g(\lambda)$ is unbounded at $\lambda_{max} = e^{-\theta}$

(iv) if $\beta>1$, $g(\lambda)$ is unbounded at $\lambda=0$

(v) if $\alpha>1$ and $\beta>1$, $g(\lambda)$ decreases monotonically

(vi) if $\alpha<1$ and $\beta<1$, $g(\lambda)$ increases monotonically

(d) Cumulative Distribution:

$$G(\lambda;\alpha,\beta,\theta) = 1 - P(\alpha,[-\ln\lambda-\theta]/\beta) \qquad (3.15)$$

where the "normalized incomplete gamma function" is defined by Eq. (3.3).

(e) Moments about Zero:

$$\omega_k' \equiv E[\lambda^k] = \frac{e^{-\theta k}}{(1+\beta k)} \qquad (3.16)$$

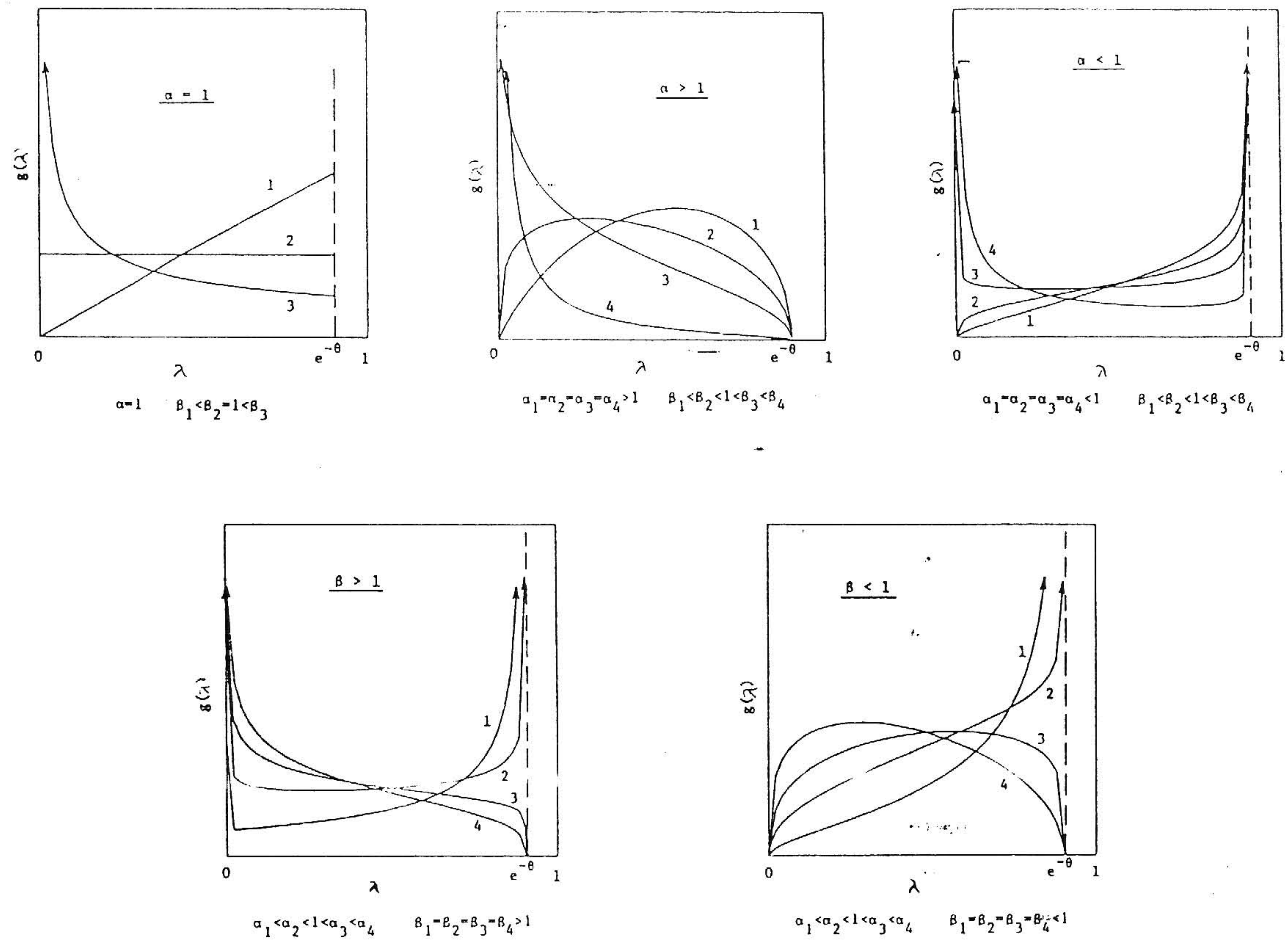

Fig. 3.5 The loggamma density function and its variation with its parameters.

LOGWEIBULL DISTRIBUTION

[$-\ln\lambda \sim \text{Weibull}(\alpha,\beta,\theta)$]

(a) Density Function:

$$g(\lambda;\alpha,\beta,\theta) = \alpha[\beta\lambda]^{-1} \left(\frac{-\ln\lambda-\theta}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{-\ln\lambda-\theta}{\alpha}\right)^{\beta}\right) \tag{3.17}$$

(b) Parameters and Range:

(i) Parameters: shape $\beta>0$; scale $\alpha>0$; shift $\theta\geq 0$

(ii) Range: $0 \leq \lambda_{min} = \theta \leq \lambda \leq \lambda_{max} = e^{-\theta}$

(c) Shape of Density Function: (see Fig. 3.6)

(i) if $\beta<1$, $g(\lambda)$ unbounded at $\lambda=0$ and $\lambda_{max} = e^{-\theta}$

(ii) if $\beta=1$ and $\alpha<1$, $g(\lambda)$ increases monotonically

(iii) if $\beta=1$ and $\alpha>1$, $g(\lambda)$ decreases monotonically

(iv) if $\beta>1$, $g(\lambda)$ bounded

(d) Cumulative Distribution:

$$G(\lambda;\alpha,\beta,\theta) = \exp\left(-\left(\frac{-\ln\lambda-\theta}{\alpha}\right)^{\beta}\right) \tag{3.18}$$

(e) Moments about Zero:

Use numerical techniques to evaluate

$$\omega_k' \equiv E[\lambda^k] = \int_{\lambda_{min}}^{\lambda_{max}} \lambda^k\, g(\lambda;\alpha,\beta,\theta)\, d\lambda$$

$$= \alpha\beta^{-1} \int_0^{e^{-\theta}} \lambda^{k-1} \left(\frac{-\ln\lambda-\theta}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{-\ln\lambda-\theta}{\alpha}\right)^{\beta}\right) d\lambda \tag{3.19}$$

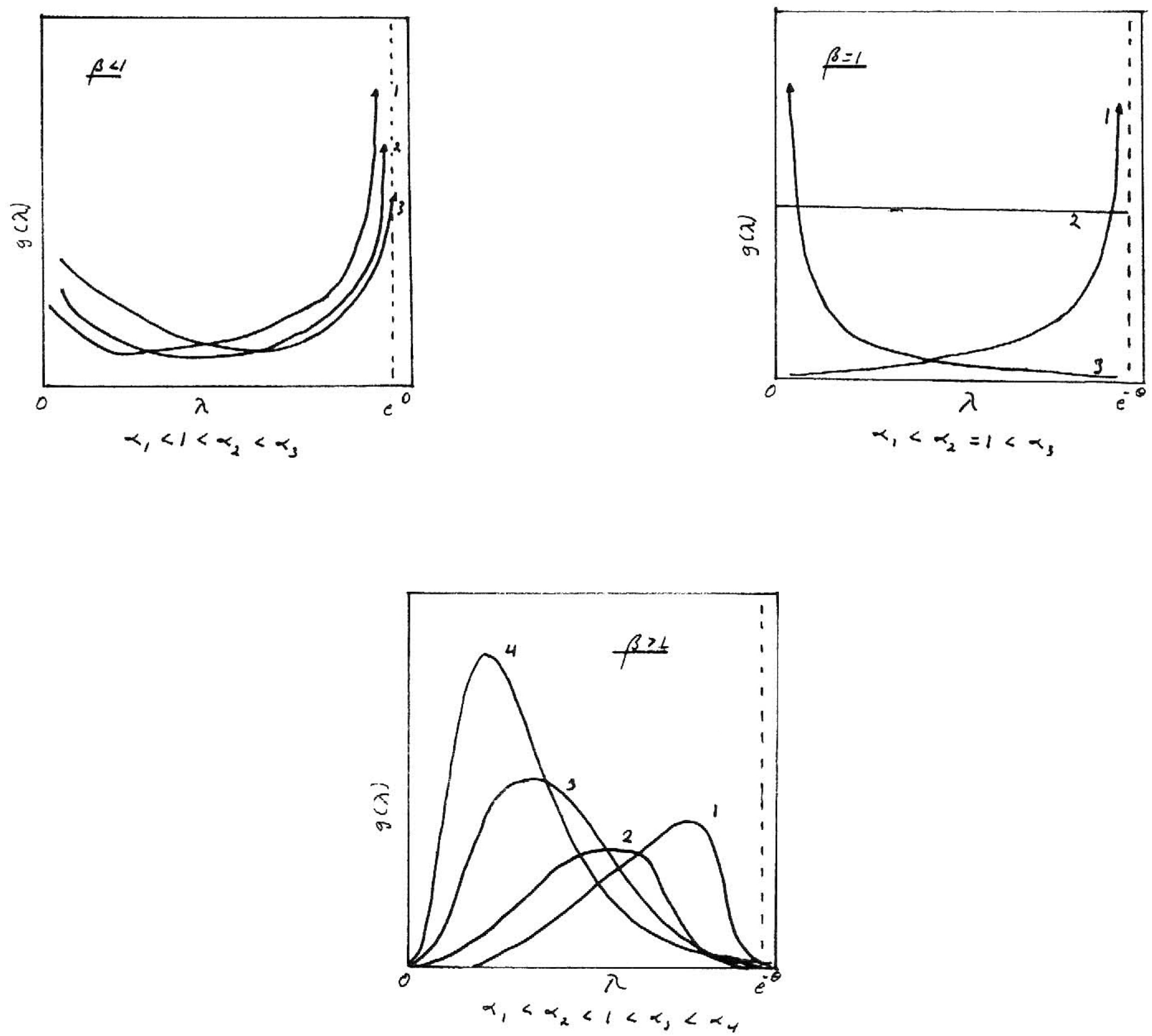

Fig. 3.6 The logWeibull density function and its variation with its parameters.

3.3 EVALUATION OF THE MARGINAL AND POSTERIOR DISTRIBUTIONS FOR THE COMPOUND MODEL

Once the prior distribution has been selected and values of the prior parameters $\underline{\theta}$ determined, the marginal and posterior distribution can be evaluated for given values of F and T. The marginal distribution is given as

$$h(F;T,\underline{\theta}) = \int_{\lambda_{min}}^{\lambda_{max}} f(F|\lambda;T)\ g(\lambda;\underline{\theta})\ d\lambda \tag{3.20}$$

and posterior distribution as

$$u(\lambda|F;T,\underline{\theta}) = f(F|\lambda;T)\ g(\lambda;\underline{\theta})/h(F;T,\underline{\theta})\ . \tag{3.21}$$

Generally, if a non-conjugate distribution is selected as the prior distribution, the integration required to evaluate the marginal distribution of Eq. (3.20) cannot be performed analytically, and numerical techniques must be used. One must be careful when numerically evaluating Eq. (3.20) since the prior distribution $g(\lambda)$ may become singular at one or both endpoints. Further complications arise if the range of integration for Eq. (3.20) is infinite (i.e., $\lambda_{max} = \infty$), as is generally the case for the present failure rate problem. The problem of numerically evaluating an integral at infinity is also encountered when the cumulative of the posterior distribution is to be evaluated, i.e.,

$$U(\lambda|F;T,\underline{\theta}) = I(\lambda)/I(\lambda_{max}) \tag{3.22}$$

where the "convolution integral" $I(\lambda)$ is defined as

$$I(\lambda) \equiv \int_{\lambda_{min}}^{\lambda_{max}} f(F|\lambda';T)\ g(\lambda';\underline{\theta})\ d\lambda'\ . \tag{3.23}$$

In order to evaluate numerically Eqs. (3.20)-(3.22), it is necessary to be able to compute $I(\lambda)$ for any λ between λ_{min} and λ_{max}.

For all of the prior families summarized in the previous section, it is necessary to use a numerical quadrature scheme to perform the integration required for $I(\lambda)$. To simplify the numerical integration of Eq. (3.23) one should first attempt to remove any potential singularities by performing analytical rearrangement of the integrand. Each prior distribution must be treated on a case-by-case basis, since there appears to be no specific technique for which each distribution may be systematically evaluated. In this section, explicit procedures are presented for evaluating $I(\lambda)$ for each of the six prior families considered in this study.

3.3.1 Explicit Expresions for $I(\lambda)$

As was mentioned previously, the prior families are generally defined for failure rates between λ_{min} and infinity. Consequently, difficulties are often encountered when numerically evaluating the convolution integral

$$I(\lambda) = \frac{T^F}{\Gamma(F+1)} \int_{\lambda_{min}}^{\lambda} \lambda'^F \, e^{-\lambda' T} \, g(\lambda';\underline{\theta}) \, d\lambda' \qquad (3.24)$$

over the possible large range of integration.

It is generally more convenient to simplify the numerical analysis of the convolution integrals by using a systematic method to obtain an effective finite upper-bound for λ, such that the integral of Eq. (3.24) can be set to zero for $\lambda > \lambda_{max}$, i.e., find an upper bound λ_{max} such that

$$\int_{\lambda_{min}}^{\infty} \lambda'^F \, e^{-\lambda' T} \, g(\lambda'|\underline{\theta}) \, d\lambda' = \int_{\lambda_{min}}^{\lambda_{max}} \lambda'^F \, e^{-\lambda' T} \, g(\lambda';\underline{\theta}) \, d\lambda' + \varepsilon \, , \qquad (3.25)$$

where ε is some small error term. With an effective finite λ_{max}, the convolution integral of Eq. (3.24) becomes

$$I(\lambda) \simeq \frac{T^F}{\Gamma(F+1)} \int_{\lambda_{min}}^{\min[\lambda,\lambda_{max}]} \lambda'^F \, e^{-\lambda' T} \, g(\lambda';\underline{\theta}) \, d\lambda' \qquad (3.26)$$

to within an error of ε. The convolution integral of Eq. (3.26) lends itself to numerical evaluation more easily than does Eq. (3.24).

In this study, the following form of Chebyshev's inequality is utilized for determining λ_{max} for a non-negative density function $w(\lambda)$ defined on the interval $\lambda\varepsilon(0,\infty)$ with unit area,

$$\mathrm{Prob}\{|\lambda-\mu| \geq r\sigma\} \leq 1/r^2 \qquad (3.27)$$

where μ and σ are the mean and standard deviation of w and r is any positive number. By restricting r to $r \geq \mu/\sigma$, this inequality reduces to

$$\int_{\mu+r\sigma}^{\infty} w(\lambda) \, d\lambda \leq 1/r^2 \ , \qquad (3.28)$$

since $\mathrm{Prob}\{\lambda<\mu-r\sigma<0\} = 0$. If one then chooses $r = \max[\mu/\sigma, 1/\sqrt{\varepsilon}]$ then

$$\int_{\mu+r\sigma}^{\infty} w(\lambda) \, d\lambda \leq \varepsilon \ . \qquad (3.29)$$

Since $\int_0^{\infty} w(\lambda) d\lambda = 1$, this last result implies

$$1-\varepsilon \leq \int_0^{\mu+r\sigma} w(\lambda) \, d\lambda \leq 1. \qquad (3.30)$$

However, to find $\lambda_{max} = \mu+r\sigma$ by this method, the mean and standard deviation of $w(\lambda)$ must first be known. Another alternative estimate of λ_{max} can be obtained by using another form of Chebyshev's inequality,

$$\mathrm{Prob}\{\lambda \geq K\} \equiv \int_K^{\infty} w(\lambda) \, d\lambda \leq \mu/K \ , \qquad (3.31)$$

where K is any positive number. If K is chosen to be μ/ε, then

$$1-\varepsilon \leq \int_0^{\mu/\varepsilon} w(\lambda)\, d\lambda \leq 1. \tag{3.32}$$

Hence, if only the mean of w is known, then one could choose $\lambda_{max}=\mu/\varepsilon$.

Given a non-negative function $w(\lambda)$ with unit area on the interval $\lambda\varepsilon(o,\infty)$, the effective value for λ_{max} on the integral $\int_0^{\infty} w(\lambda)\, d\lambda$ can be selected as

$$\lambda_{max} = \min\left[\mu/\varepsilon,\ \lambda + \sigma \max(\mu/\sigma,\ 1/\sqrt{\varepsilon})\right] \tag{3.33}$$

so that

$$1-\varepsilon \leq \int_0^{\lambda_{max}} w(\lambda)\, d\lambda \leq 1. \tag{3.34}$$

One can use the above results for finding an effective upper limit for an integral over an infinite range to solve the problem of the convolution integral of Eq. (3.24) which can be written as

$$I(\lambda) = \int_0^{\lambda} f(F|\lambda;T)\ g(\lambda;\underline{\theta})\ d\lambda\ , \tag{3.35}$$

where f is the Poisson distribution, Eq. (2.1), and $g(\lambda)$ the prior distribution. To obtain an effective λ_{max} for the convolution integral, one needs to apply the above results to the Poisson and prior distributions separately and use the maximum of the two values obtained. Both of these distributions are non-negative and normalized such that

$$T \int_0^{\infty} f(F|\lambda;T)\ d\lambda = 1 \tag{3.36}$$

and

$$\int_0^{\infty} g(\lambda;\underline{\theta})\ d\lambda = 1. \tag{3.37}$$

The mean and variance for Eq. (3.36) are (F+1)/T and $(F+1)/T^2$, respectively, so that from Eqs. (3.33) and (3.34)

$$1-\varepsilon \le T \int_0^{\lambda_1} f(F|\lambda;T)\, d\lambda \le 1 , \tag{3.38}$$

where

$$\lambda_1 \equiv \min\left(\frac{F+1}{\varepsilon T},\ \left[\frac{F+1}{T} + \frac{\sqrt{F+1}}{T}\max(\sqrt{F+1},\ 1/\sqrt{\varepsilon})\right]\right) . \tag{3.39}$$

Similarly, if μ and σ^2 are the mean and variance for Eq. (3.37), then

$$1-\varepsilon \le \int_0^{\lambda_2} g(\lambda;\underline{\theta})\, d\lambda \le 1 . \tag{3.40}$$

where

$$\lambda_2 \equiv \min[\mu/\varepsilon,\ \max(2\mu,\ \mu+\sigma/\sqrt{\varepsilon})] . \tag{3.41}$$

Having determined the upper limits λ_1 and λ_2 for the Poisson and prior functions, respectively, the effective upper limit for the convolution integral of Eq. (3.35) can be taken as $\lambda_{max} = \max[\lambda_1, \lambda_2]$.

Although λ_{max} now requires some computational effort to determine λ_1 and λ_2 from Eqs. (3.39) and (3.41), the numerical evaluation of Eq. (3.24) is considerably more efficient than that of Eq. (3.26) particularly for distributions describing low failure probability events.

3.3.1.1 Gamma and Poisson Convolution Integral [2]

For the gamma prior distribution of Eq. (3.1), Eq. (3.26) becomes

$$I(\lambda) = \frac{T^F \beta^\alpha e^{-\theta T}}{\Gamma(F+1)\Gamma(\alpha)} \int_0^{\lambda^*-\theta} (\lambda+\theta)^F \lambda^{\alpha-1} e^{-\lambda(T+\beta)}\, d\lambda, \tag{3.42}$$

where $\lambda^* \equiv \min[\lambda, \lambda_{max}]$. For $\alpha>1$ the above integrand is bounded and numerical quadrature may be applied directly. However, if $0<\alpha<1$ the integrand becomes weakly singular at $\lambda=0$ and it is preferable to integrate the above results by parts before applying numerical quadrature. The result is

$$I(\lambda) = \frac{T^F \beta^\alpha e^{-\theta T}}{\Gamma(F+1)\Gamma(\alpha+1)} \left\{ \int_0^{\lambda^*-\theta} \lambda^\alpha (\lambda+\theta)^F \; e^{-\lambda(T+\beta)} \left\{ \frac{F}{\lambda+\theta} - (T+\beta) \right\} d\lambda \right.$$
$$\left. + (\lambda^*-\theta)^\alpha (\lambda^*)^F \; e^{-(\lambda^*-\theta)(T+\beta)} \right\} , \qquad (3.43)$$

Although the integrand in this expression is no longer positive, it is bounded over the entire integration range and hence suitable to be evaluated by numerical integration.

For the special case that the shift parameter $\theta=0$, the gamma distribution becomes the natural conjugate to the Poisson distribution and Eq. (3.42) reduces to (using the incomplete gamma function if Eq. (3.3))

$$I(\lambda) = \frac{\Gamma(F+\alpha)}{\Gamma(\alpha)\Gamma(F+1)} \; \frac{T^F \beta^\alpha}{(T+\beta)^{F+\alpha}} \; P(F+\alpha, \lambda^*(T+\beta)). \qquad (3.44)$$

As $\lambda \to 0$, this reduces to the marginal distribution

$$I(\infty) \equiv h(F;T,\alpha,\beta) = \frac{T(F+\alpha)}{\Gamma(\alpha)\Gamma(F+1)} \; \frac{T^F \beta^\alpha}{(T+\beta)^{F+\alpha}} \; . \qquad (3.45)$$

3.3.1.2 *Weibull and Poisson Convolution Integral* [2]

Upon substitution of the explicit expression for the Weibull prior distribution of Eq. (3.5) into Eq. (3.26), the convolution integral becomes

$$I(\lambda) = \frac{\beta T^F e^{-\theta T}}{\alpha^\beta \Gamma(F+1)} \int_0^{\lambda^*-\theta} \lambda^{\beta-1} (\lambda+\theta)^F \; e^{-\lambda T} \; e^{-(\lambda/\alpha)^\beta} \; d\lambda \qquad (3.46)$$

where again $\lambda^* \equiv \min[\lambda, \lambda_{max}]$. The integrand is positive and bounded if $\beta>1$, or if $F>1$ when $\theta=0$ and $\beta<1$. For these cases numerical quadrature can be used to evaluate this expression.

If $\beta<1$ and $\theta>0$, the above integral can be rewritten as

$$I(\lambda) = \frac{\beta T^F e^{-\theta T}}{\alpha^\beta \Gamma(F+1)} \left\{ \int_0^{\lambda^*-\theta} \lambda^{\beta-1}\, e^{-\lambda T}\, \{(\lambda+\theta)^F e^{-(\lambda/\alpha)^\beta} - \theta^F\}\, d\lambda \right.$$
$$\left. + \theta^F \Gamma(\beta)\, P(\beta,(\lambda^*-\theta)T)/T^\beta \right\} , \qquad (3.47)$$

where the normalized incomplete gamma function P is given by Eq. (3.3). For the special case that the shift parameter θ vanishes and the number of failures F=0, this result reduces to

$$I(\lambda) = \beta\alpha^{-\beta} \left\{ \int_0^{\lambda^*} \lambda^{\beta-1}\, e^{-\lambda T}\, \{e^{-(\lambda/\alpha)^\beta} -1\}\, d\lambda + \Gamma(\beta) P(\beta,\lambda^* T)/T^\beta \right\}. \qquad (3.48)$$

The integrands in Eqs. (3.47) and (3.48) have no singularities and hence are amenable to evaluation by numerical techniques.

3.3.1.3 Lognormal and Poisson Convolution Integral [2]

The lognormal prior distribution of Eq. (3.8) gives for the convolution integral of Eq. (3.26)

$$I(\lambda) = T^F\, e^{-\theta T}\, [\Gamma(F+1)\, \sqrt{2\pi}\beta]^{-1} \int_0^{\lambda^*-\theta} \frac{(\lambda+\theta)^F}{\lambda}\, e^{-\lambda T}\, \exp[-(\ln\lambda-\alpha)^2/2\beta^2]\, d\lambda \qquad (3.49)$$

where $\lambda^* \equiv \min[\lambda,\lambda_{max}]$. It can be shown that the integrand vanishes as $\lambda\to 0$ for all values of α,β,F and T, and thus Eq. (3.49) can be used directly with numerical integration techniques to evaluate $I(\lambda)$.

3.3.1.4 Logbeta and Poisson Convolution Integral [2]

For the logbeta prior distribution of Eq. (3.11), Eq. (3.26) becomes

$$I(\lambda) = \frac{T^F \Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(F+1)}\, (b-a)^{1-\alpha-\beta} \int_{e^a}^{\lambda^*} \lambda^{F-1} e^{-\lambda T} (\ln\lambda-a)^{\alpha-1} (b-\ln\lambda)^{\beta-1}\, d\lambda \qquad (3.50)$$

where $\lambda^* \equiv \min[\lambda, \lambda_{max}, e^b]$. For $\alpha,\beta>1$ this integrand is positive and bounded; and, hence, the integral can be directly evaluated with numerical quadrature.

However, if $\alpha<1$ or $\beta<1$, singularities appear in the integrand at $\lambda=e^a$ or $\lambda=e^b$. These integrable singularities can be removed by defining $x = (\ln\lambda-a)/(b-a)$ and adding and subtracting the singularities to give

$$I(\lambda) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(F+1)} \left[\int_0^{x^*} v(x)dx + \frac{W(0)}{\alpha}(x^*)^\alpha + \frac{w(1)}{\beta}[1-(1-x^*)^\beta]\right] , \tag{3.51}$$

where

$$x^* \equiv (\ln\lambda^*-a)/(b-a) , \tag{3.52}$$

$$v(x) \quad (1-x)^{\beta-1} \{x^{\alpha-1} [w(x)-w(0)(1-x)^{1-\beta}]-w(0)\} , \tag{3.53}$$

$$w(x) = \{T \exp[a+(b-a)x]\}^F \exp[-T e^{a+(b-a)x}] , \tag{3.54}$$

$$v(0) = -w(1) = -(Te^b)^F \exp[-Te^b] , \tag{3.55}$$

and

$$v(1) = -w(0) = -(Te^a)^F \exp[-Te^a] . \tag{3.56}$$

While this result is considerably more complicated than Eq. (3.50), the integrand in Eq. (3.51) is now bounded for all positive α and β and therefore suitable for numerical evaluation.

3.3.1.5 Loggamma and Poisson Convolution Integral

For the loggamma prior distribution of Eq. (3.14), the convolution integral becomes

$$I(\lambda) = \frac{T^F \beta^{-\alpha}}{\Gamma(F+1)\Gamma(\alpha)} \int_0^{\lambda^*} \lambda^{F-1} e^{-\lambda T}(-\ln\lambda-\theta)^{\alpha-1} \exp[-(-\ln\lambda-\theta)/\beta] \, d\lambda . \tag{3.57}$$

Where $\lambda^* \equiv \min[\lambda, \lambda_{max}, e^{-\theta}]$. This form of the convolution integral is bounded for most realistic values of the loggamma parameters. If $\beta>1$ and $F=0$, the integrand becomes singular. For this case, adding and subtracting the singularity at $\lambda=0$ gives

$$I(\lambda) = \frac{T^F \beta^{-\alpha}}{\Gamma(F+1)\Gamma(\alpha)} \int_0^{\lambda} (e^{-\lambda T}-1)\ (-\ln\lambda-\theta)^{\alpha-1}\ \exp\left[-(-\ln\lambda-\theta)/\beta\right] \frac{d\lambda}{\lambda}$$
$$+ \frac{T^F}{\Gamma(F+1)} \left[1-P(\alpha, \{-\ln\lambda-\theta\}/\beta)\right] \tag{3.58}$$

where P is the normalized incomplete gamma function defined by Eq. (3.3). The integral in this result is now bounded as $\lambda \to 0$.

3.3.1.6 LogWeibull and Poisson Convolution Integral

For the logWeibull prior distribution of Eq. (3.17), the convolution integral becomes

$$I(\lambda) = \frac{T^F \beta}{\Gamma(F+1)\alpha} \int_0^{\lambda_{max}} e^{-\lambda T}\ \lambda^{F-1} \left(\frac{-\ln\lambda-\theta}{\alpha}\right)^{\beta-1} \exp\left[-(\{-\ln\lambda-\theta\}/\alpha)^{\beta}\right] d\lambda \tag{3.59}$$

where $\lambda_{max} \equiv \min[\lambda_1, \lambda_2, e^{-\theta}]$ λ_1 and λ_2 having been determined from the results obtained by using the first form of Chebyshev's inequality, i.e., $\lambda_{max} = \mu + r\sigma$. Numerical difficulties often occurred in the evaluation of the convolution integral of Eq. (3.59) when the value for the error term is very small ($\sim 10^{-10}$), thereby producing an effective λ_{max} larger than required. Using a larger error term resulted in a lower value of λ_{max} that is more applicable for evaluating the convolution integral of Eq. (3.59) without any significant loss of accuracy in the final results.

If $\beta<1$ and F=0 the convolution integral of Eq. (3.59) becomes singular at $\lambda=0$. By adding and subtracting the singularity at $\lambda=0$ one obtains

$$I(\lambda) = \frac{T^F\beta}{\Gamma(F+1)\alpha}\left(\int_0^{\lambda_{max}} x^F\ (v(x) - v(\theta))\ dx + v(\theta)e^{-\theta}\right) \tag{3.60}$$

where

$$v(x) = x^{-1}\ e^{-xT}\left(\frac{-\ln x-\theta}{\alpha}\right)^{\beta-1}\ \exp\left[-(\{-\ln x-\theta\}/\alpha)^{\beta}\right]\ . \tag{3.61}$$

The integral in this result is now bounded as $\lambda\to0$.

Chapter 4

ESTIMATION OF PRIOR PARAMETERS FOR THE COMPOUND MODEL

An important factor to consider when using the compound model, is how does one best estimate values for the free parameters of the prior distribution. Appropriate values for the parameters must be selected so that the resulting compound model agrees as closely as possible with the observed failure data for the class of components under consideration.

Many methods are available for determining the prior parameters from prior failure data. In this study, three standard methods for estimating (from given failure data) the prior parameters for arbitrary prior families are used. They are

(i) matching the moments of the prior distribution to those of the observed failure data,

(ii) matching the moments of the marginal distribution to those of the observed failure data, and

(iii) maximizing the likelihood function of the marginal distribution.

In these particular applications of the estimation techniques, it is assumed that all range or shift parameters of the prior distribution are to be specified *a priori* so that only shape and scale parameters are to be estimated from the data. By so specifying the shift parameter, only two parameters need be estimated for each of the six prior families discussed in Chapter 3.

4.1 MATCHING MOMENTS METHODS

This method consists of matching the sample moments of the given data to the corresponding distribution moments. Since two parameters need be estimated, this method involves matching any two sample moments of the data to the corresponding moments of the distribution. In practice, the first two moments are usually used. The distribution moments, as will be shown later, are functions of the two equations with two unknowns being the parameters of the distribution. Often, the application of this method results in equations that are non-linear and thus requires the use of numerical techniques to obtain values for the parameters. Due to the different forms of the prior families considered in this study, this method must be applied on a case-by-case basis. In this section, the method of matching moments is applied to the prior and marginal distributions of the compound model for each of the six prior families presented in Chapter 3.

4.1.1 Prior Matching Moments Method (PMMM)

In this subsection, the prior parameters estimates are obtained by matching the mean $\bar{\lambda}$ and variance S^2 of the given data to the corresponding mean μ and variance σ^2 of the prior distribution. Explicitly, $\bar{\lambda}$ and S^2 are defined as

$$\bar{\hat{\lambda}} = \frac{1}{n} \sum_{i=1}^{n} (F_i/T_i) \qquad (4.1)$$

and

$$s_{\hat{\lambda}}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (F_i/T_i - \bar{\hat{\lambda}})^2 \tag{4.2}$$

where n is the number of components in the given data set and F_i is the number of failures experienced in test time T_i by individual components in the set.

The K-th "moment about zero", ω_k', for the prior distribution is defined as

$$\omega_k' \equiv E[\lambda^k] = \int_{\lambda_{min}}^{\lambda_{max}} \lambda^k \, g(\lambda;\underline{\theta}) \, d\lambda \ . \tag{4.3}$$

The expressions for ω_k', k = 1,2,3... were given for all six prior families in Section 3.1. The mean ω_1 and the variance ω_2 for the prior distributions are found to be

$$\text{(i)} \quad \omega_1 = \omega_1' \qquad \text{(mean)} \ , \tag{4.4}$$

$$\text{(ii)} \quad \omega_2 = \omega_2' - (\omega_1)^2 \quad \text{(variance)} \ . \tag{4.5}$$

If more than two parameters were to be estimated, one could obtain the third and fourth central moments by calculating first

$$\text{(iii)} \quad \omega_3 = \omega_3' - 3\omega_2'\omega_1 + 2\omega_1^3 \ , \tag{4.6}$$

$$\text{(iv)} \quad \omega_4 = \omega_4' - 4\omega_3'\omega_1 + 6\omega_2' - 3\omega_1^4 \ , \tag{4.7}$$

and then use the "shape factors", α_k, defined as [7]

$$\alpha_k \equiv \omega_k \, \{\omega_2\}^{-k/2}, \quad k>2 \ . \tag{4.8}$$

The first two shape factors are

$$\alpha_3 = \omega_3 \, \omega_2^{-3/2} \qquad \text{("index of skewness")} \tag{4.9}$$

and

$$\alpha_4 = \omega_4 \, \omega_2^{-2} \qquad \text{("index of kurtosis").} \tag{4.10}$$

Skewness is a measure of asymmetery. Kurtosis is a measure of the degree of flatness of a density near its center. However, in this study only two parameters need be estimated.

By equating the mean and variance of Eqs. (4.1) and (4.2) to the corresponding mean and variance of the prior distribution, the prior matching moments equations become

$$\omega_1(\alpha,\beta) = \bar{\hat{\lambda}} \tag{4.11}$$

and

$$\omega_2(\alpha,\beta) = S^2_{\hat{\lambda}} \,. \tag{4.12}$$

Solving these two equations for the parameters $\underline{\theta} \equiv \alpha$ and β will thus yield estimates for the prior parameters.

4.1.2 Marginal Matching Moments Method (MMMM)

This method consists of matching the data mean and variance, $\bar{\hat{\lambda}}$ and $S^2_{\hat{\lambda}}$, to the mean and variance, $\omega_1^*(\alpha,\beta)$ and $\omega_2^*(\alpha,\beta)$, corresponding to the marginal distribution. The task of obtaining analytical expressions for $\omega_1^*(\alpha,\beta)$ and $\omega_2^*(\alpha,\beta)$ appears at first glance to be difficult. However, it can be shown, for any prior distribution, the expected values of $\omega_1^*(\alpha,\beta)$ and $\omega_2^*(\alpha,\beta)$ can be expressed in terms of the prior distribution [2] simply by replacing $S^2_{\hat{\lambda}}$ by $S^2_{\hat{\lambda}} - \overline{T^*\lambda}$ where T^* is given by

$$T^{*} \equiv \frac{1}{n} \sum T_i^{-1} \quad . \tag{4.13}$$

The marginal matching moments equations become

$$\omega_1^{*}(\alpha,\beta) = \overline{\hat{\lambda}} \tag{4.14}$$

and

$$\omega_2^{*}(\alpha,\beta) = S_{\hat{\lambda}}^2 - T^{*}\overline{\hat{\lambda}} \quad . \tag{4.15}$$

Unfortunately, there are data sets for which negative values for α and β will be obtained. If the testing times of the components are short, T^{*} increases which increases the likelihood of obtaining a negative value for $[S_{\hat{\lambda}}^2 - T^{*}\overline{\hat{\lambda}}]$. Most prior families do not permit negative values for α and β. For these situations, the MMMM fails to estimate acceptable parameters.

4.1.3 Matching Moment Estimators for Particular Prior Distributions

1. Gamma Distribution: (shift parameter θ given) [2]

From Eq. (3. 4), the marginal matching moments equations for this distribution are

$$\frac{\alpha}{\beta} + \theta = \overline{\hat{\lambda}} \tag{4.16}$$

and

$$\frac{\alpha(\alpha+1)}{\beta^2} + \frac{2\theta\alpha}{\beta} + \theta^2 = S^2 + \overline{\hat{\lambda}}^2 \quad . \tag{4.17}$$

Solution of these equations for the parameter estimators yields

$$\hat{\alpha} = (\overline{\hat{\lambda}}-\theta)^2/S^2 \tag{4.18}$$

and

$$\hat{\beta} = (\overline{\hat{\lambda}}-\theta)/S^2 \, . \tag{4.19}$$

2. Weibull Distribution: (shift parameter θ given) [2]

From the moments given by Eq. (3.7) and from Eq. (4.4), the matching moments equations of Eqs. (4.11) and (4.12) become

$$\alpha\Gamma(1 + \frac{1}{\beta}) + \theta = \overline{\hat{\lambda}} \tag{4.20}$$

and

$$\alpha^2[\Gamma(1 + \frac{2}{\beta}) - \Gamma^2(1 + \frac{1}{\beta})] = S^2 \, , \tag{4.21}$$

From the first equation, one obtains

$$\alpha = (\overline{\hat{\lambda}}-\beta)/\Gamma(1 + \frac{1}{\beta}) \, , \tag{4.22}$$

which, upon substitution into Eq. (4.21), yields

$$S^2 \; \Gamma^2(1 + \frac{1}{\beta}) - (\overline{\hat{\lambda}}-\theta)^2 \; [\Gamma(1 + \frac{2}{\beta}) - \Gamma^2(1 + \frac{1}{\beta})] = 0. \tag{4.23}$$

This non-linear equation for β may be solved by standard numerical methods for finding the zeros of a function of one variable (e.g., Mueller's iteration method [6]). Once $\hat{\beta}$ is obtained, $\hat{\alpha}$ can be evaluated directly from Eq. (4.22).

3. Lognormal Distribution: (shift parameter θ given) [2]

From Eqs. (3.10) and (4.4) the moments ω_1 and ω_2 are obtained so that the matching moments equations become

$$\theta + \exp[\alpha + \tfrac{1}{2} \beta^2] = \overline{\hat{\lambda}} \tag{4.24}$$

and

$$\exp[2\alpha + 2\beta^2] + \exp[2\alpha + \beta^2] = S^2 \qquad (4.25)$$

Solution of these two equations yields

$$\hat{\alpha} = ln(\overline{\hat{\lambda}-\theta}) - \tfrac{1}{2}\, ln[1 + S^2/(\overline{\hat{\lambda}-\theta})^2] \qquad (4.26)$$

and

$$\hat{\beta} = \left\{ ln[1 + S^2/(\overline{\hat{\lambda}-\theta})] \right\}^{\frac{1}{2}} . \qquad (4.27)$$

4. Logbeta Distribution: (range parameters a and b given) [2]

Explicit expressions for the moments ω_1 and ω_2 are not available and the moments *must* be obtained from Eq. (3.13) by performing the integration numerically. Numerical techniques are used to solve the matching moments equations for the estimators $\hat{\alpha}$ and $\hat{\beta}$.

In a previous study [2], Eqs. (4.11) and (4.12) were solved by considering an equivalent minimization problem in which the estimators $\hat{\alpha}$ and $\hat{\beta}$ are selected as those values of α and β which minimize the non-negative function

$$F(\alpha,\beta) \equiv [\omega_1(\alpha,\beta) - \overline{\hat{\lambda}}]^2 + \left\{\omega_2'(\alpha,\beta) - S^2 - \overline{\hat{\lambda}}\right\}^2 . \qquad (4.28)$$

This minimization is performed by using a standard simplex search technique [8], and the minimizing values of α and β are then checked to verify that F vanishes at the minimum, thereby ensuring that a solution to the matching moments equations has been obtained.

5. Loggamma Distribution: (shift parameter θ given)

From Eq. (3.16) the matching moments equations, Eqs. (4.11) and (4.12), become

$$\bar{\hat{\lambda}} = [1+\beta]^{-\alpha} e^{-\theta} \tag{4.29}$$

and

$$\bar{\hat{\lambda}} + S^2 = [1+2\beta]^{-\alpha} e^{-2\theta} \tag{4.30}$$

The first equation can be solved for α, namely

$$\alpha = -(\theta + \ln\bar{\hat{\lambda}})/\ln(1+\beta) \tag{4.31}$$

Substitution of this result into Eq. (4.30) gives the following equation for β:

$$[2\theta + \ln(\bar{\hat{\lambda}} + S^2)]\ln(1+\beta) - [\theta + \ln\bar{\hat{\lambda}}]\ln(1+2\beta) = 0 \tag{4.32}$$

This equation with one unknown can be readily solved by standard numerical techniques which find zeros of a function of a single variable. Once $\hat{\beta}$ is found, $\hat{\alpha}$ is then obtained from Eq. (4.31).

6. Log<u>Weibull Distribution:</u> (shift parameter θ given)

For this prior distribution, explicit expressions for the moments ω_1 and ω_2 as functions of the parameters α and β could not be found analytically. The numerical solution of the matching moments equations can be performed in a similar manner as used for the logbeta distribution by minimizing the non-negative function

$$F(\alpha,\beta) \equiv [\omega_1(\alpha,\beta) - \bar{\hat{\lambda}}]^2 + \left(\omega_2'(\alpha,\beta) - S^2 - \bar{\hat{\lambda}}^2\right)^2 . \tag{4.33}$$

Due to the absolute difference convergence method used by the standard simplex search technique, Eq. (4.33) was normalized and redefined as

$$F(\alpha,\beta) \equiv [\omega_1(\alpha,\beta)/\bar{\hat{\lambda}} - 1]^2 + [\omega_2'(\alpha,\beta)/(S^2+\bar{\hat{\lambda}}^2) - 1]^2 \tag{4.34}$$

for this case. Further discussion concerning the convergence method used is covered in Chapter 5.

4.2 MARGINAL MAXIMUM LIKELIHOOD METHOD (MMLM)

Maximum likelihood estimators are the values of α and β which maximize the likelihood function for the given failure data. The marginal distribution of Eq. (2.3) gives the probability that a component, randomly selected from the class, will experience F failures in test time T. If each component is assumed to be independent, the probability L of observing $(F_1,T_1),(F_2,T_2),\ldots,(F_n,T_n)$ is the product of the probability of observing each pair of (F_i,T_i) separately or

$$L(\alpha,\beta) \equiv \prod_{i=1}^{n} h(F_i,T_i,\alpha,\beta) \;, \tag{4.35}$$

where h is the marginal distribution of Eq. (2.3). The maximum likelihood estimators of the prior parameters are those values of α and β which maximize $L(\alpha,\beta)$ for the given failure data. Equivalently, one may wish to maximize $\ln L(\alpha,\beta)$ - a more convenient form. Here, the prior parameters are the solutions of

$$\frac{\partial \ln L(\alpha,\beta)}{\partial\alpha} \equiv \sum_{i=1}^{n}\left[\frac{\frac{\partial}{\partial\alpha}h(F_i;T_i,\alpha,\beta)}{h(F_i;T_i,\alpha,\beta)}\right] = 0 \tag{4.36a}$$

and

$$\frac{\partial \ln L(\alpha,\beta)}{\partial\beta} \equiv \sum_{i=1}^{n}\left[\frac{\frac{\partial}{\partial\beta}h(F_i;T_i,\alpha,\beta)}{h(F_i;T_i,\alpha,\beta)}\right] = 0 \;. \tag{4.36b}$$

In this study and a previous study [1], the following second order central difference result was used to evaluate the derivatives of the marginal distribution:

$$\frac{df}{dx} \simeq \frac{f(x-2\Delta x) - 8f(x-\Delta x) + 8f(x+\Delta x) - f(x+2\Delta x)}{12\Delta x} + O(\Delta x)^4 \;, \tag{4.37}$$

where Δx is some suitably small increment of the variable.

Once the derivatives $\partial h/\partial\alpha$ and $\partial h/\partial\beta$ are computed numerically, values for the summations in Eqs. (4.36a) and (4.36b) can be found. Several methods are available for obtaining a solution to these maximum likelihood equations [10]. A convenient method used in this study was to find the minimum of the non-negative function

$$F(\alpha,\beta) \equiv \left(\frac{\partial\, ln L(\alpha,\beta)}{\partial\alpha}\right)^2 + \left(\frac{\partial\, ln L(\alpha,\beta)}{\partial\beta}\right)^2 . \tag{4.38}$$

by the simplex technique [8]. This particular method for obtaining the solution of the maximum likelihood equations was convenient in that the simplex technique had previously been incorporated into the study.

Chapter 5

SAFER COMPUTER CODE

In this chapter a general description is given of the SAFER computer code used in the analysis of conjugate and non-conjugate prior distributions in compound failure models. Due to the magnitude in size of the code, no attempt is made to detail code operation line by line.

In order to operate the code, one must obtain the users manual [11] for a complete description of the structure including input commands and options The manual also includes a flow diagram showing how the separate parts of the code are used and an example which illustrates input commands and output of the code.

In previous works [1-2] a good deal of time and effort was expended engineering this computer code to enable the user to perform safety analyses using a compound failure model. The safety analyst using SAFER has the option of choosing the prior family to represent the prior distribution used in the compound failure model. Hopefully, by creating such a computer code, the analyst will have available a code allowing greater ease in computations and greater latitude in the analysis.

5.1 Description

As discussed above, the prior distribution for the compound model analysis may be choosen as one of several distribution families (e.g. gamma, loggamma, Weibull, logbeta, etc.). The analyst may choose to estimate the unknown parameters of the selected prior distribution from the failure data by any or all of the procedures discussed in Chapter 4. i.e., (i) matching the data mean and variance to those of the selected prior distribution, (ii) matching the data mean and variance to those of the marginal distribution, and (iii) maximizing the likelihood function of the marginal distribution.

All needed procedures and numerical algorithms have been programmed into the code for proper evaluation of the convolution integrals in Chapter 3.

Several other functions are available with SAFER. A Chi-square and/or a Kolmogrov goodness-of-fit test can be performed in order to see how well the resulting statistical models describe the given failure data. An analysis of the posterior distribution estimated for each component as well as various types of confidence intervals, tolerance intervals, tables, plots, etc. can be requested.

5.2 Modifications

The description of the SAFER computer code discussed in the users manual [11] includes only four of the six currently existing possible choices of prior families. The original SAFER code was programmed to do the necessary analysis for the compound failure model with the following choices of prior family: gamma, Weibull, lognormal, and logbeta. In this study, the SAFER computer code was modified to include the prior families loggamma and logWeibull as choices for representing the prior distribution in the compound model. Procedures in each part of the program necessary for the complete analysis of the compound model were changed to accommodate loggamma and logWeibull distribution families.

There was no need to introduce any new numerical techniques into SAFER Only modifications of existing techniques available in SAFER were used.

5.2.1 Specific Changes Made in SAFER

Not only were steps added to the program to perform analyses on the loggamma-Poisson and logWeibull-Poisson compound failure models, but also other changes were made to the original SAFER code. The changes made should increase the efficiency of SAFER and decrease possible errors in future use of SAFER.

In a previous study [2], the simplex search technique was introduced into the SAFER code to solve a minimization problem. The estimators $\hat{\alpha}$ and $\hat{\beta}$ are selected as those values of α and β which minimize the non-negative function

$$F(\alpha,\beta) \equiv [\omega_1(\alpha,\beta) - \bar{\bar{\lambda}}]^2 + \left(\omega_2'(\alpha,\beta) - S^2 - \bar{\bar{\lambda}}^2\right)^2 . \qquad (5.1)$$

The method used to test for convergence was an absolute difference, as is seen in Eq. (5.1). In this study, the simplex search technique was used to determine estimates of α and β, for the logWeibull distribution, which minimize the non-negative function of Eq. (5.1). Realizing that components which exhibit inherently low failure rates also exhibit the characteristic $\omega_1(\alpha,\beta) >> \omega_2(\alpha,\beta)$, one can observe that the squared value of the second half of Eq. (5.1) may effectively be lost due to its magnitude relative to the squared value of the first half. This may result in an inequality when matching the variance of the data set to that of the distribution. In order to avoid possible error using the convergence method above, Eq. (5.1) was normalized to that of Eq. (4.34) i.e.,

$$F(\alpha,\beta) \equiv [\omega_1(\alpha,\beta)/\overline{\hat{\lambda}} - 1]^2 + [\omega_2'(\alpha,\beta)/(S^2+\overline{\hat{\lambda}}^2) - 1]^2 \tag{5.2}$$

Both squared values in Eq. (5.2) now carry the same weight for determining convergence.

The Mueller's iteration method, previously introduced into SAFER, was used to find the zeroeth solution of the non-linear equation for β in Eq. (4.23), for the case that the Weibull family is chosen as the prior distribution. This method was also utilized to solve for β in Eq. (4.32) for the case when the loggamma family is chosen to represent the prior distribution. From experience, it was found that a large change in β generally did not greatly effect the mean and variance of the loggamma distribution. Therefore, the users initial guess of β was often far from the true value. Mueller's iteration method requires that one chose two different parameter values such that the equation is positive for one value and negative for the other. This is done by multiplying the initial guess for the parameter by some multiplification factor to obtain one larger guess value and then divide the initial guess for the parameter by the same multiplication factor to obtain a second smaller guess. By using this multiplier <u>given</u> in SAFER one often did not obtain values for the parameter that would satisfy the above requirements for the method. The problem was corrected by increasing the multiplication and dividing factor of β to increase the upper-bound and decrease the lower-bound in an effort to enlarge the interval around the initial guessed value of β. This may increase the iterations required to obtain a solution in some cases. The advantage is that it reduces the possible error of the user when determining an initial value for the parameter.

Chapter 6

ANALYSIS OF DIFFERENT PRIOR DISTRIBUTIONS IN THE COMPOUND MODEL

When using the compound failure model to analyze the operation of a nuclear power plant, one would like the prior distribution describing the variation of the failure rates for a given set of components to be completely known, i.e., both the function family and its parameter values specified. Generally, one assumes the prior distribution belongs to a particular family of distributions and then proceeds to estimate appropriate values for the distribution parameters. By choosing the gamma family to represent the prior distribution, the resulting analysis of the compound model is relatively simple. Therefore, in most applications of the compound model, the gamma distribution is used.

The effect of choosing the conjugate gamma family to represent the prior distribution is generally unknown. Also, the properties of the estimators for the prior distribution are generally unknown. It is of interest to users of the compound failure model to determine what affect the choice of the prior family has on the model and to what degree the procedure used to estimate the prior parameters affects the results. The prupose of this study was to investigate further both of these questions for the failure rate problem.

Besides the gamma distribution, several other families of distributions are reasonable alternatives for representing the prior distributions. In this study, two families of distributions, the loggamma and logWeibull, were introduced as possible alternatives to the conjugate gamma distribution as the prior distribution. In order to obtain more conclusive results, three families of distributions, Weibull, lognormal, and logbeta, previously investigated in another study [1], are also included in this report.

Once a family of distributions has been selected, values for the prior parameters must be determined. In this study, the prior parameters were estimated from given component failure data by three methods, (PMMM, MMMM, and MMLM) for each of the prior families selected. From these results, it was then possible to determine the effect of the prior family selection and the parameter estimation technique on the compound failure model.

6.1 CONJUGATE ANALYSIS OF NON-CONJUGATE FAILURE DATA

Several different families of distributions could be used to describe the prior distribution in the compound failure model. It is of interest to the user of the compound failure model to have information of the economic feasibility of each family. Generally, the non-conjugate prior distributions require time consuming numerical methods to obtain results from the compound model. Therefore, it is of practical importance to determine how well the conjugate gamma distribution can be used to approzimate non-conjugate distributions in compound failure models.

In this section, failure data were generated from loggamma-Poisson and logWeibull-Poisson marginal distributions with known parameters. These failure data were then analyzed by theconjugate gamma-Poisson failure model and the results of the estimated gamma prior distributions then compared to results from the actual non-conjugate prior distributions.

6.1.1 Data Simulation Procedure

In this study, the loggamma-Poisson and logWeibull-Poisson marginal distributions were assumed with specific values of parameters α and β and component operation time T. With the marginal distribution of Eq. (2.3), simulated failure data were generated as follows. A random number between 0 and 1 was taken as the value of the cumulative marginal distribution. A value of F was then determined that yielded this value for the cumulative distribution of Eq. (2.3). By using a series of random numbers, a series of component failures F_i (all with same operation time T) were generated which were distributed according to the specific marginal distribution.

6.1.2 Results of Using a Conjugate Model to Analyze Data from Non-Conjugate Models

A total of 500 failure data were simulated from each of the two non-conjugate distributions. These data were grouped into 25 sets of sample size 20. Parameter values for the two non-conjugate prior distributions were chosen so that the prior means and standard deviations were 0.12(-4)[1] and 0.11(-4) per hour, respectively. The component operation time T was taken as 50,000 hours (almost 6 years). These values of the mean and standard deviation were chosen to represent components characterized by low failure rates.

[1] read 0.12(-4) as 0.12 times 10 to the -4 power, i.e., 0.000012

From the gamma parameters estimated by the PMMM and MMMM from the simulated non-conjugate failure data, various characteristics of the resulting estimated gamma distributions were compared to the same characterstics of the known non-conjugate prior distributions. These comparisons are summarized in Tables 6.1 and 6.2 for failure data obtained from the known loggamma-Poisson and logWeibull-Poisson marginal distributions, respectively. Only the two matching moments estimation methods (PMMM and MMMM) were used in the phase of the study. The marginal maximum likelihood estimation method applied to non-conjugate failure models requires a large amount of computational effort for the analysis of a large number of data samples. Therefore, this method was omitted in this phase of the study.

Tables 6.3, 6.4, and 6.5 present similar results from a previous study [1] where three families of distributions (the Weibull, lognormal, and logbeta) were selected as non-conjugate prior distributions. The procedure used to obtain the failure data and analyses by the conjugate gamma-Poisson failure model were the same as described above.

The results obtained from using the gamma-Poisson model to analyze failure data simulated from loggamma-Poisson and logWeibull-Poisson non-conjugate failure models often matched reasonably well with the actual values. These results agreed with those obtained in the previous study where the Weibull, lognormal, and logbeta non-conjugate distributions were selected as prior distributions for the compound model.

In the next section, the correct models are used to analyze these same sets of simulated failure data. It will be seen that perfect agreement is still not obtained. Furthermore, in some cases, the gamma-

Poisson model even yielded parameter estimators closer to the actual values than did the correct models. Given these results, one can conclude that the gamma-Poisson model has sufficient flexibility for analyzing failure data even though the data are believed to come from another failure distribution.

There was no significant difference in the results of the loggamma-Poisson model and the other non-conjugate results obtained in the previous works [1]. However, the logWeibull-Poisson model analysis consistently showed smaller variance (a more peaked distribution) whether the failure data were analyzed by using the gamma-Poisson model or with the logWeibull-Poisson model.

TABLE 6.1

Results of a gamma-Poisson analysis of data from a loggamma-Poisson distribution

parameter & True Value	Method	Mean	Standard Deviation	Sq. Root of MSE[a]	Bias
mean 1.20(-5)	PMMM	1.17(-5)	.360(-5)	.353(-5)	-.032(-5)
	MMMM	1.25(-5)	.350(-5)	.340(-5)	.050(-5)
standard deviation 1.10(-5)	PMMM	1.60(-5)	.391(-5)	.383(-5)	.502(-5)
	MMMM	.780(-5)	.392(-5)	.381(-5)	-.322(-5)
5-th percentile .215(-5)	PMMM	.015(-5)	.022(-5)	.021(-5)	-.200(-5)
	MMMM	.368(-5)	.267(-5)	.268(-5)	.153(-5)
25-th percentile .500(-5)	PMMM	.137(-5)	.120(-5)	.118(-5)	-.363(-5)
	MMMM	.694(-5)	.344(-5)	.335(-5)	.193(-5)
50-th percentile .881(-5)	PMMM	.557(-5)	.300(-5)	.294(-5)	-.324(-5)
	MMMM	1.07(-5)	.375(-5)	.364(-5)	.187(-5)
75-th percentile 1.52(-5)	PMMM	1.53(-5)	.536(-5)	.525(-5)	.007(-5)
	MMMM	1.61(-5)	.445(-5)	.433(-5)	.088(-5)
95-th percentile 3.25(-5)	PMMM	4.39(-5)	1.09(-5)	1.07(-5)	1.15(-5)
	MMMM	2.75(-5)	.967(-5)	.939(-5)	-.498(-5)

[a] Mean Squared Error

TABLE 6.2

Results of a gamma-Poisson analysis of data from a logWeibull-Poisson distribution

Parameter & True value	Method	Mean	Standard Deviation	Sq. Root of MSE[a]	Bias
mean	PMMM	1.14(-5)	.315(-5)	.309(-5)	-.056(-5)
1.20(-5)	MMMM	1.27(-5)	.270(-5)	.259(-5)	.075(-5)
standard	PMMM	1.52(-5)	.314(-5)	.334(-5)	.417(-5)
deviation	MMMM	.664(-5)	.415(-5)	.397(-5)	-.436(-5)
1.10(-5)					
5-th	PMMM	.019(-5)	.028(-5)	.027(-5)	-.444(-5)
percentile	MMMM	.516(-5)	.406(-5)	.389(-5)	.054(-5)
.463(-5)					
25-th	PMMM	.185(-5)	.133(-5)	.130(-5)	-.480(-5)
percentile	MMMM	.803(-5)	.407(-5)	.390(-5)	.138(-5)
.665(-5)					
50-th	PMMM	.581(-5)	.276(-5)	.270(-5)	-.332(-5)
percentile	MMMM	1.12(-5)	.346(-5)	.332(-5)	.210(-5)
.914(-5)					
75-th	PMMM	1.52(-5)	.467(-5)	.457(-5)	.169(-5)
percentile	MMMM	1.59(-5)	.308(-5)	.295(-5)	.236(-5)
1.35(-5)					
95-th	PMMM	4.18(-5)	.948(-5)	.929(-5)	1.38(-5)
percentile	MMMM	2.55(-5)	.865(-5)	.828(-5)	-.246(-5)
2.79(-5)					

[a] Mean Squared Error

TABLE 6.3

Results of a gamma-Poisson analysis of data from a Weibull-Poisson distribution

Parameter & True value	Method	Mean	Standard Deviation	Sq. Root of MSE[a]	Bias
mean 1.20(-5)	PMMM	1.20(-5)	.397(-5)	.389(-5)	.004(-5)
	MMMM	1.25(-5)	.395(-5)	.388(-5)	.053(-5)
	MMLM	1.24(-5)	.432(-5)	.421(-5)	.044(-5)
standard deviation 1.10(-5)	PMMM	1.81(-5)	.496(-5)	.864(-5)	.714(-5)
	MMMM	1.10(-5)	.568(-5)	.553(-5)	.004(-5)
	MMLM	1.39(-5)	.568(-5)	.623(-5)	.029(-5)
5-th percentile .083(-5)	PMMM	.009(-5)	.015(-5)	.075(-5)	-.074(-5)
	MMMM	.235(-5)	.352(-5)	.375(-5)	.152(-5)
	MMLM	.096(-5)	.145(-5)	.141(-5)	.013(-5)
25-th percentile .399(-5)	PMMM	.120(-5)	.106(-5)	.297(-5)	-.278(-5)
	MMMM	.490(-5)	.379(-5)	.380(-5)	.092(-5)
	MMLM	.329(-5)	.295(-5)	.294(-5)	-.070(-5)
50-th percentile .889(-5)	PMMM	.505(-5)	.287(-5)	.477(-5)	-.385(-5)
	MMMM	.910(-5)	.390(-5)	.380(-5)	.021(-5)
	MMLM	.770(-5)	.415(-5)	.419(-5)	-.119(-5)
75-th percentile 1.67(-5)	PMMM	1.53(-5)	.602(-5)	.608(-5)	-.147(-5)
	MMMM	1.65(-5)	.555(-5)	.541(-5)	-.021(-5)
	MMLM	1.65(-5)	.607(-5)	.588(-5)	-.020(-5)
95-th percentile 3.38(-5)	PMMM	4.78(-5)	1.38(-5)	1.95(-5)	1.41(-5)
	MMMM	3.44(-5)	1.39(-5)	1.35(-5)	.056(-5)
	MMLM	4.00(-5)	1.46(-5)	1.54(-5)	.621(-5)

[a] Mean Squared Error

TABLE 6.4

Results of a gamma-Poisson analysis of data from a lognormal-Poisson distribution

Parameter & True value	Method	Mean	Standard Deviation	Sq. Root of MSE[a]	Bias
mean 1.20(-5)	PMMM	1.20(-5)	.418(-5)	.409(-5)	-.004(-5)
	MMMM	1.25(-5)	.419(-5)	.411(-5)	.047(-5)
	MMLM	1.24(-5)	.459(-5)	.446(-5)	.038(-5)
standard deviation 1.10(-5)	PMMM	1.79(-5)	.524(-5)	.863(-5)	.694(-5)
	MMMM	1.07(-5)	.598(-5)	.582(-5)	-.023(-5)
	MMLM	1.32(-5)	.604(-5)	.625(-5)	.221(-5)
5-th percentile .230(-5)	PMMM	.010(-5)	.015(-5)	.221(-5)	-.221(-5)
	MMMM	.247(-5)	.351(-5)	.342(-5)	.017(-5)
	MMLM	1.21(-5)	.167(-5)	.195(-5)	-.109(-5)
25-th percentile .512(-5)	PMMM	.123(-5)	.110(-5)	.403(-5)	-.389(-5)
	MMMM	.505(-5)	.387(-5)	.377(-5)	-.006(-5)
	MMLM	.367(-5)	.324(-5)	.346(-5)	-.145(-5)
50-th percentile .884(-5)	PMMM	.506(-5)	.299(-5)	.478(-5)	-.377(-5)
	MMMM	.917(-5)	.406(-5)	.397(-5)	.034(-5)
	MMLM	.799(-5)	.442(-5)	.437(-5)	-.084(-5)
75-th percentile 1.51(-5)	PMMM	1.52(-5)	.628(-5)	.616(-5)	.006(-5)
	MMMM	1.64(-5)	.582(-5)	.580(-5)	.127(-5)
	MMLM	1.64(-5)	.637(-5)	.630(-5)	.128(-5)
95-th percentile 3.22(-5)	PMMM	4.74(-5)	1.46(-5)	2.08(-5)	1.51(-5)
	MMMM	3.37(-5)	1.47(-5)	1.44(-5)	.148(-5)
	MMLM	3.84(-5)	1.55(-5)	1.62(-5)	.622(-5)

[a] Mean Squared Error

TABLE 6.5

Results of a gamma-Poisson analysis of data from a logbeta-Poisson distribution

parameter & True value	Method	Mean	Standard Deviation	Sq. Root of MSE[a]	Bias
mean	PMMM	1.20(-5)	.418(-5)	.409(-5)	-.004(-5)
1.20(-5)	MMMM	1.25(-5)	.419(-5)	.411(-5)	.047(-5)
	MMLM	1.24(-5)	.459(-5)	.446(-5)	.038(-5)
standard	PMMM	1.79(-5)	.524(-5)	.863(-5)	.694(-5)
deviation	MMMM	1.07(-5)	.598(-5)	.582(-5)	-.023(-5)
1.10(-5)	MMLM	1.32(-5)	.604(-5)	.625(-5)	.221(-5)
5-th	PMMM	.009(-5)	.015(-5)	.237(-5)	-.237(-5)
percentile	MMMM	.247(-5)	.351(-5)	.342(-5)	.001(-5)
.246(-5)	MMLM	.121(-5)	.167(-5)	.204(-5)	-.125(-5)
25-th	PMMM	.123(-5)	.110(-5)	.415(-5)	-.401(-5)
percentile	MMMM	.505(-5)	.387(-5)	.377(-5)	-.019(-5)
.524(-5)	MMLM	.367(-5)	.324(-5)	.351(-5)	-.158(-5)
50-th	PMMM	.506(-5)	.299(-5)	.480(-5)	-.380(-5)
percentile	MMMM	.917(-5)	.406(-5)	.396(-5)	.031(-5)
.886(-5)	MMLM	.799(-5)	.442(-5)	.437(-5)	-.087(-5)
75-th	PMMM	1.52(-5)	.628(-5)	.616(-5)	.020(-5)
percentile	MMMM	1.64(-5)	.582(-5)	.583(-5)	.141(-5)
1.50(-5)	MMLM	1.64(-5)	.637(-5)	.633(-5)	.142(-5)
95-th	PMMM	4.74(-5)	1.46(-5)	2.10(-5)	1.54(-5)
percentile	MMMM	3.37(-5)	1.47(-5)	1.44(-5)	.180(-5)
3.19(-5)	MMLM	3.84(-5)	1.55(-5)	1.63(-5)	.654(-5)

[a] Mean Squared Error

6.2 NON-CONJUGATE ANALYSIS OF DATA FROM NON-CONJUGATE MODELS

The results of the previous section showed that the sample gamma-Poisson conjugate model could be successfully used to analyze failure data simulated from non-conjugate marginal distributions.

To confirm these results, failure data generated from the loggamma-Poisson and logWeibull-Poisson non-conjugate models were analyzed by the same compound failure models from which the data were generated.

Tables 6.6 and 6.7 present the results of using the correct non-conjugate models to analyze the simulated failure data from the loggamma-Poisson and logWeibull-Poisson marginal distributions, respectively. As in the previous section, only the two matching moments estimation methods (PMMM and MMMM) were used. Tables 6.8, 6.9, and 6.10 present the results of using the correct non-conjugate models to analyze the simulated failure data from the Weibull-Poisson, lognormal-Poisson, and logbeta-Poisson marginal distributions, respectively [1].

Of the five alternative prior models investigated in these two studies, the logbeta-Poisson and logWeibull-Poisson models presented the most difficulties in calculating parameter estimates even for the simple prior matching moments. Both models required a similar elaborate numerical algorithm to obtain estimates of α and β from both the PMMM and MMMM parameter estimation methods.

TABLE 6.6

Results of a loggamma-Poisson analysis of data from a loggamma-Poisson distribution

Parameter & True value	Method	Mean	Standard Deviation	Sq. Root of MSE[a]	Bias
alpha	PMMM	110.5	24.31	23.82	-88.69
199.2	MMMM	442.0	208.9	202.6	222.8
beta	PMMM	0.114	0.266	0.260	.506
.0585	MMMM	0.032	0.019	0.018	-.027
mean	PMMM	1.17(-5)	.360(-5)	.353(-5)	-.032(-5)
1.20(-5)	MMMM	1.26(-5)	.370(-5)	.359(-5)	.056(-5)
stan. dev.	PMMM	1.60(-5)	.391(-5)	.383(-5)	.502(-5)
1.10(-5)	MMMM	.843(-5)	.366(-5)	.354(-5)	-.257(-5)
5-th pctl.	PMMM	.099(-5)	.063(-5)	.062(-5)	-.116(-5)
.215(-5)	MMMM	.402(-5)	.229(-5)	.222(-5)	.187(-5)
25-th pctl.	PMMM	.309(-5)	.153(-5)	.150(-5)	-.191(-5)
.500(-5)	MMMM	.703(-5)	.312(-5)	.302(-5)	.203(-5)
50-th pctl.	PMMM	.664(-5)	.272(-5)	.267(-5)	-.217(-5)
.381(-5)	MMMM	1.04(-5)	.373(-5)	.361(-5)	.164(-5)
75-th pctl.	PMMM	1.39(-5)	.473(-5)	.463(-5)	-.131(-5)
1.52(-5)	MMMM	1.56(-5)	.455(-5)	.441(-5)	.038(-5)
95-th pctl.	PMMM	3.87(-5)	1.03(-5)	1.01(-5)	.626(-5)
3.25(-5)	MMMM	2.82(-5)	.847(-5)	.820(-5)	-.432(-5)

[a] Mean Squared Error

TABLE 6.7

Results of a logWeibull-Poisson analysis of data from a logWeibull-Poisson distribution

Parameter & True value	Method	Mean	Standard Deviation	Sq. Root of MSE[a]	Bias
alpha	PMMM	12.01	.360	.352	.243
11.77	MMMM	11.50	.288	.276	-.269
beta	PMMM	21.55	1.66	1.63	-4.15
25.70	MMMM	57.59	27.8	26.6	31.9
mean	PMMM	1.14(-5)	.315(-5)	.308(-5)	-.056(-5)
1.20(-5)	MMMM	1.29(-5)	.267(-5)	.256(-5)	.094(-5)
stan. dev.	PMMM	1.52(-5)	.341(-5)	.334(-5)	.417(-5)
1.10(-5)	MMMM	.599(-5)	.442(-5)	.423(-5)	-.501(-5)
5-th pctl.	PMMM	.346(-5)	.126(-5)	.124(-5)	-.116(-5)
.463(-5)	MMMM	.818(-5)	.330(-5)	.317(-5)	.355(-5)
25-th pctl.	PMMM	.536(-5)	.180(-5)	.176(-5)	-.129(-5)
.665(-5)	MMMM	.974(-5)	.318(-5)	.305(-5)	.309(-5)
50-th pctl.	PMMM	.782(-5)	.243(-5)	.238(-5)	-.131(-5)
.914(-5)	MMMM	1.15(-5)	.300(-5)	.288(-5)	.234(-5)
75-th pctl.	PMMM	1.25(-5)	.362(-5)	.354(-5)	-.101(-5)
1.35(-5)	MMMM	1.42(-5)	.290(-5)	.278(-5)	.073(-5)
95-th pctl.	PMMM	2.97(-5)	.750(-5)	.736(-5)	.174(-5)
2.79(-5)	MMMM	2.22(-5)	.600(-5)	.575(-5)	-.573(-5)

[a] Mean Squared Error

TABLE 6.8

Results of a Weibull-Poisson analysis of data from a Weibull-Poisson distribution

Parameter & True value	Method	Mean	Standard Deviation	Sq. Root of MSE[a]	Bias
alpha	PMMM	.938(-5)	.399(-5)	.495(-5)	-.304(-5)
1.24(-5)	MMMM	1.25(-5)	.445(-5)	.433(-5)	.004(-5)
tau	PMMM	.683	.126	.432	-.413
1.10	MMMM	2.04	2.44	2.56	.944
mean	PMMM	1.20(-5)	.397(-5)	.389(-5)	.004(-5)
1.20(-5)	MMMM	1.25(-5)	.395(-5)	.368(-5)	.053(-5)
stan. dev.	PMMM	1.81(-5)	.496(-5)	.864(-5)	.714(-5)
.10(-5)	MMMM	1.10(-5)	.568(-5)	.553(-5)	.005(-5)
5-th pctl.	PMMM	.017(-5)	.016(-5)	.068(-5)	-.066(-5)
.083(-5)	MMMM	.219(-5)	.329(-5)	.348(-5)	.136(-5)
25-th pctl.	PMMM	.168(-5)	.105(-5)	.253(-5)	-.231(-5)
.399(-5)	MMMM	.500(-5)	.371(-5)	.375(-5)	.101(-5)
50-th pctl.	PMMM	.561(-5)	.269(-5)	.421(-5)	-.328(-5)
.890(-5)	MMMM	.926(-5)	.393(-5)	.384(-5)	.037(-5)
75-th pctl.	PMMM	1.49(-5)	.569(-5)	.586(-5)	-.182(-5)
1.67(-5)	MMMM	1.65(-5)	.563(-5)	.548(-5)	-.022(-5)
95-th pctl.	PMMM	4.55(-5)	1.38(-5)	1.79(-5)	1.18(-5)
3.38(-5)	MMMM	3.40(-5)	1.39(-5)	1.36(-5)	.018(-5)

[a] Mean Squared Error

TABLE 6.9

Results of a lognormal-Poisson analysis of data from a lognormal-Poisson distribution

Parameter & True value	Method	Mean	Standard Deviation	Sq. Root of MSE[a]	Bias
alpha	PMMM	-12.03	.568	.683	-0.40
-11.63	MMMM	-11.67	.526	.514	-0.04
tau	PMMM	1.105	.148	.357	.326
.778	MMMM	0.734	.303	.299	-.044
mean	PMMM	1.20(-5)	.418(-5)	.409(-5)	-.004(-5)
1.20(-5)	MMMM	1.25(-5)	.429(-5)	.411(-5)	.047(-5)
stan. dev.	PMMM	1.79(-5)	.524(-5)	.863(-5)	.694(-5)
1.10(-5)	MMMM	1.07(-5)	.598(-5)	.582(-5)	-.023(-5)
5-th pctl.	PMMM	.085(-5)	.053(-5)	.053(-5)	-.162(-5)
.246(-5)	MMMM	.222(-5)	.212(-5)	.252(-5)	-.024(-5)
25-th pctl.	PMMM	.260(-5)	.138(-5)	.135(-5)	-.264(-5)
.524(-5)	MMMM	.449(-5)	.268(-5)	.326(-5)	-.075(-5)
50-th pctl.	PMMM	.573(-5)	.272(-5)	.266(-5)	-.314(-5)
.886(-5)	MMMM	.780(-5)	.334(-5)	.384(-5)	-.107(-5)
75-th pctl.	PMMM	1.27(-5)	.555(-5)	.545(-5)	-.226(-5)
1.50(-5)	MMMM	1.42(-5)	.541(-5)	.541(-5)	-.079(-5)
95-th pctl.	PMMM	4.06(-5)	1.68(-5)	1.66(-5)	.870(-5)
3.19(-5)	MMMM	3.59(-5)	1.73(-6)	1.81(-5)	.400(-5)

[a] Mean Squared Error

TABLE 6.10

Results of a logbeta-Poisson analysis of data from a logbeta-Poisson distribution

Parameter & True value	Method	Mean	Standard Deviation	Sq. Root of MSE[a]	Bias
alpha	PMMM	136.	39.3	150.	-145.
282	MMMM	136.(+1)	259.(+1)	274.(+1)	108.(+1)
tau	PMMM	70.9	17.0	71.5	-69.5
140	MMMM	653.	122.(+1)	130.(+1)	512.
mean	PMMM	1.20(-5)	.418(-5)	.409(-5)	-.004(-5)
1.20(-5)	MMMM	1.25(-5)	.420(-5)	.411(-5)	.047(-5)
stan. dev.	PMMM	1.79(-5)	.524(-5)	.863(-5)	.694(-6)
1.10(-5)	MMMM	1.07(-5)	.599(-5)	.583(-5)	-.023(-5)
5-th pctl.	PMMM	.100(-5)	.061(-5)	.143(-5)	-.130(-5)
.230(-5)	MMMM	.351(-5)	.319(-5)	.333(-5)	.120(-5)
25-th pctl.	PMMM	.303(-5)	.152(-5)	.257(-5)	-.209(-5)
.512(-5)	MMMM	.606(-5)	.342(-5)	.346(-5)	.095(-5)
50-th pctl.	PMMM	.649(-5)	.284(-5)	.364(-5)	-.234(-5)
.884(-5)	MMMM	.940(-5)	.379(-4)	.373(-5)	.056(-5)
75-th pctl.	PMMM	1.38(-5)	.531(-5)	.536(-5)	-.129(-5)
1.51(-5)	MMMM	1.51(-5)	.531(-5)	.517(-5)	.002(-5)
95-th pctl.	PMMM	4.04(-5)	1.31(-5)	1.52(-5)	.814(-5)
3.22(-5)	MMMM	3.16(-5)	1.33(-5)	1.30(-5)	-.065(-5)

[a] Mean Squared Error

The use of a loggamma-Poisson model to analyze failure data from a loggamma-Poisson marginal distribution (as seen from Table 6.6) gave better estimators of alpha from the PMMM, but better estimators of beta from the MMMM. The means of the estimated prior distribution from the two matching moments methods showed only slight disagreement due to the fewer successful estimates obtained from the MMMM. The mean of the standard deviations of the estimated prior distributions obtained by the MMMM have less bias than those from the PMMM. The percentile estimators of the PMMM are consistently lower than those for the MMMM. Only the 95-th percentile estimators of the PMMM are larger than those of the MMMM. Since the mean of the standard deviations was significantly lower for the MMMM than for the PMMM, the percentile estimators of the MMMM showed a more peaked distribution.

In comparison with results in Table 6.1, which were obtained by using a gamma-Poisson model to analyze failure data from a loggamma-Poisson marginal distribution, the means of the estimated prior distribution from the two matching moments methods are the same as those obtained when the correct model was used to analyze the failure data, (shown in Table 6.6). The means of the standard deviations of the estimated prior distributions obtained by the PMMM was the same as in Table 6.6, but less for the MMMM. The characteristics of the percentiles estimated were the same as those obtained when the correct model was used to analyze the failure data, shown in Table 6.6

The logWeibull-Poisson model yielded results that favored the prior matching moments estimation method. The mean of the PMMM estimated prior distribution has an advantage over that obtained from the MMMM, mainly because the MMMM failed to give successful estimates for 12 of the 25 data sets. The estimators for the PMMM of alpha and beta are less biased than those obtained from the MMMM. The PMMM gave a larger mean square error for the estimators of alpha than did the MMMM. The PMMM showed better results for the means and standard deviations of the estimated prior distribution. Only for the 75-th percentile did the MMMM showed smaller mean square errors for estimates of the 75-th and 95-th percentiles. In all, the PMMM did a better job than the MMMM for the given simulated failure data.

In comparison with results in Table 6.2, which were obtained by using a gamma-Poisson model to analyze failure data from a logWeibull-Poisson marginal distribution, the estimators of the means of the estimated prior distribution from the two matching moments methods are the same for both failure models. The estimators of the standard deviations were the same for the PMMM, but better for the MMMM when the gamma-Poisson model was used. With the PMMM, the correctly assumed model gave better estimators of the 5-th, 25-th, 50-th, and 75-th percentiles, but poorer estimators of the 95-th percentiles than did the gamma-Poisson model. For the estimates of the 5-th, 25-th, 50-th, and 95-th percentiles, the gamma-Poisson model yielded less biased estimators from the MMMM, but yielded greater MSE estimators for the 5-th, 25-th, and 95-th percentiles from the MMMM. In all, the gamma-Poisson model yielded poorer estimators for the PMMM, but better estimators for the MMMM.

The use of a Weibull-Poisson model to analyze failure data from a Weibull-Poisson marginal distribution (as is seen from Table 6.8) showed better results for MMMM than the PMMM. The gamma-Poisson model showed comparable results from each estimation method considered.

The lognormal-Poisson model also yielded results that favored the marginal matching moments estimation method. A comparison of the lognormal-Poisson results to those obtained with gamma-Poisson analysis (Table 6.4) shows the mean and standard deviation of the estimated prior distribution for both models are the same. Using a gamma-Poisson model yielded better estimates of the 5-th, 25-th, 50-th, and 95-th percentiles in terms of bias and MSE for the MMMM than those from a lognormal-Poisson model. However, the PMMM did a better job with the lognormal-Poisson model than with the gamma-Poisson model.

As shown in Table 6.10, the PMMM gave better estimates for the parameters alpha and beta for the logbeta prior distribution than did the MMMM. The mean of the PMMM estimated prior distribution was better than that of the MMMM, but the MMMM showed better results for estimates of the prior standard deviation. The 5-th, 25-th, and 5-th percentiles favor the PMMM. Using the gamma-Poisson model in Table 6.5 gave similar values for the mean and standard deviation of the estimated prior distribution. The results from the Table 6.5 showed poorer estimators of the 5-th, 25-th, and 50-th percentiles, but estimates from the MMMM gave comparable values of MSE for both parameter estimation methods.

In each of the three alternative models previously investigated [1,2] the MMMM failed to give successful estimates for 6 of the 25 simulated data sets.

6.3 COMPARISON OF USING SIX DIFFERENT PRIOR DISTRIBUTIONS WITH THE SAME FAILURE DATA

Another method to show the effect of using different prior distributions in the failure rate analysis is to plot the estimated prior distributions for a particular failure data set. In a previous study [1], failure data sets of size 5 were simulated from a gamma-Poisson marginal distribution by the method described in Section 6.1, but with the observed time T increased to 1,000,000 hours so as to raise the mean number of observed failures from 0.6 to 12. Therefore, the majority of simulated failure data no longer consisted of zero or one failure and the distributions of prior parameters estimators obtained from these data were smooth. Two of these data sets were chosen and analyzed with the conjugate gamma-Poisson failure model and three non-conjugate failure models, i.e., Weibull-Poisson, lognormal-Poisson, and logbeta-Poisson. For each failure model, estimates of the prior parameters were calculated by all three estimation methods and the estimated prior distributions were plotted. The 5-th, 25-th, 50-th, 75-th, and 95-th percentiles of each estimated prior distribution were also computed. In this study, the same two data sets were analyzed by the loggamma-Poisson and logWeibull-Poisson failure models using all three estimation methods and the estimated prior distributions were plotted along with those prior distributions previously estimated, i.e., gamma, Weibull, lognormal, and logbeta. The 5-th, 25-th, 50-th, 75-th, and 95-th percentiles of the estimated loggamma and logWeibull prior distributions were also computed.

The two failure data sets chosen for this phase of the study were (2,3,10,11,28) and (4,8,10,12,13). The first data sample was called good data since it yielded prior gamma parameter estimates with small bias for all three parameter estimation methods. The second data sample used was called bad data since it gave very poor prior parameter estimates. The prior parameter estimates for the prior distributions obtained from the good and bad data sets using each estimation method are given in Tables 6.11 and 6.12, respectively.

Figures 6.1, 6.2, and 6.3 shows the estimated prior distributions belonging to the six prior families combined from both studies, with parameters estimated by the PMMM, MMMM, and MMLM, respectively, for the good data sample. In order to better visualize the resulting prior distributions, the figures are presented both in linear and semi-log form. All three figures show the same characteristics, i.e., close similarity between the estimated gamma and Weibull distributions and between the loggamma, lognormal, and logbeta distributions. Since the 95-th percentile is an important characteristic in safety studies, the observation of considerable interest is that the high failure rate tail of the distributions, are very close to one another. This charateristic is confirmed by the results in Table 6.13 which gives the 5-th, 25-th, 50-th, 75-th, and 95-th percentiles of the estimated prior distributions of Figures 6.1, 6.2, and 6.3.

Figures 6.4, 6.5, and 6.6 and Table 6.14 present the results obtained from the bad data sample. These figures show characteristics comparable with those of the good data. Close similarity are particularly evident between the loggamma, lognormal, and logbeta distributions when the MMMM was used for this bad data set. The estimated prior distributions from the gamma and Weibull distributions show some minor differences. However, the results in Table 6.14 show close estimates of all five percentiles calculated for the different estimated distributions, except those obtained for the logWeibull prior distribution from the MMLM estimation method. From these results, one can conclude that any difficulty with estimating the prior distribution from the failure data is inherent in the data sample and not in the parameter estimation technique used.

TABLE 6.11

Parameter estimates for six prior models obtained by three estimation techniques for GOOD data

Prior[a] Dist.	Method	Alpha	Beta[b]
Gamma	PMMM	1.073	9.936(4)
	MMMM	1.191	1.103(5)
	MMLM	1.543	1.429(5)
Weibull	PMMM	1.096(-5)	1.036
	MMMM	1.096(-5)	1.036
	MMLM	1.167(-5)	1.267
Lognormal	PMMM	-11.76	0.811
	MMMM	-11.74	0.781
	MMLM	-11.78	0.860
Logbeta	PMMM	257.7	130.7
	MMMM	262.0	133.5
	MMLM	242.7	123.1
Loggamma	PMMM	187.2	6.298(-2)
	MMMM	203.2	5.788(-2)
	MMLM	184.2	6.369(-2)
LogWeibull	PMMM	11.89	25.49
	MMMM	11.87	26.26
	MMLM	12.15	16.07

[a] The range or shift parameters are selected so that the support of the prior distribution is $(0,\infty)$.

[b] For the gamma distribution, these represent $1/\beta$.

TABLE 6.12

Parameter estimates for six prior models obtained by three estimation techniques for BAD data

Prior[a] Dist.	Method	Alpha	Beta[b]
Gamma	PMMM	6.093	7.348(5)
	MMMM	25.99	2.765(6)
	MMLM	82.94	8.823(6)
Weibull	PMMM	1.055(-5)	2.849
	MMMM	1.014(-5)	5.921
	MMLM	9.996(-6)	7.786
Lognormal	PMMM	-11.64	3.687(-1)
	MMMM	-11.59	1.943(-1)
	MMLM	-11.58	0.988(-1)
Logbeta	PMMM	1.323(3)	6.595(2)
	MMMM	4.788(3)	2.371(3)
	MMLM	1.626(4)	8.038(3)
Loggamma	PMMM	9.788(2)	1.189(-2)
	MMMM	3.537(3)	3.278(-3)
	MMLM	9.126(3)	1.269(-3)
Logweibull	PMMM	11.76	48.17
	MMMM	11.67	83.81
	MMLM	11.59	357.09

[a] The range or shift parameters are selected so that the support of the prior distribution is $(0,\infty)$.

[b] For the gamma distribution, these represent $1/\beta$.

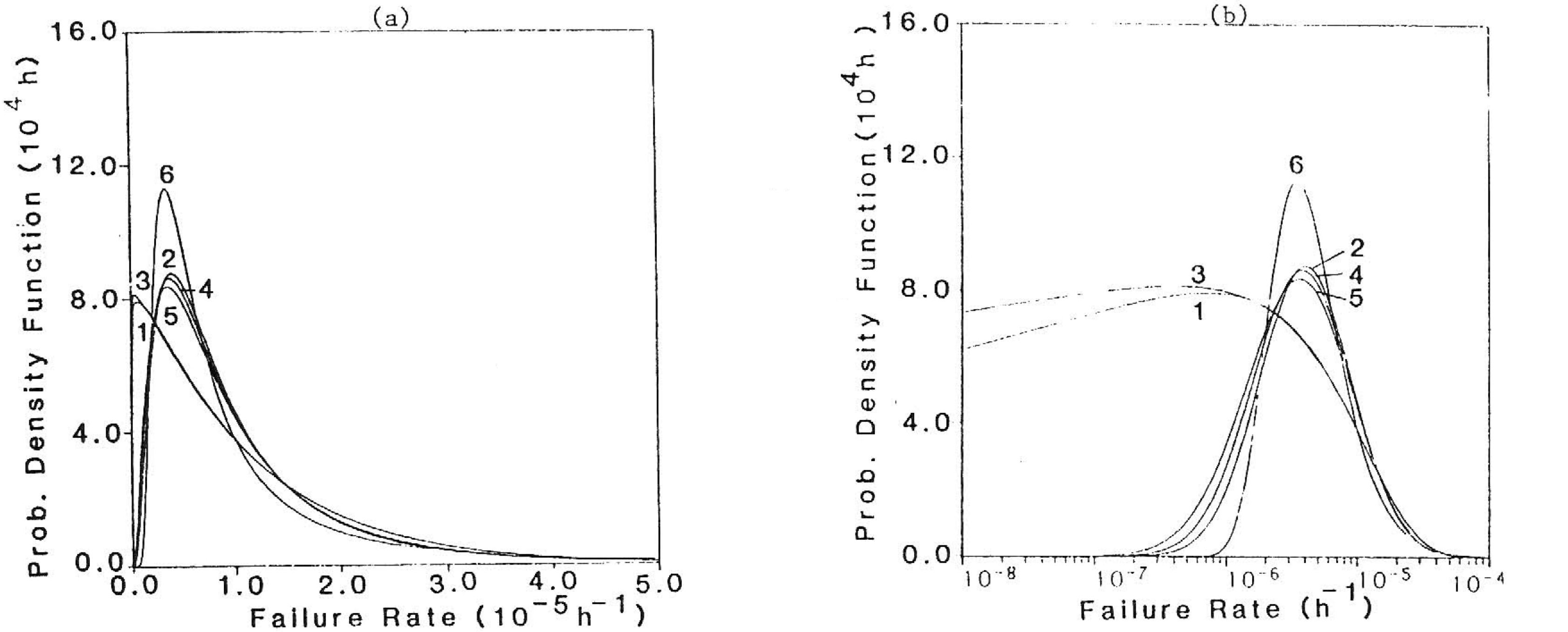

Figure 6.1: Estimated prior distributions obtained by the PMMM with GOOD data, (a) linear (b) semi-log. The prior distributions were assumed to belong to the families: (1) Gamma, (2) Lognormal, (3) Weibull, (4) Logbeta, (5) Loggamma, and (6) Log-Weibull.

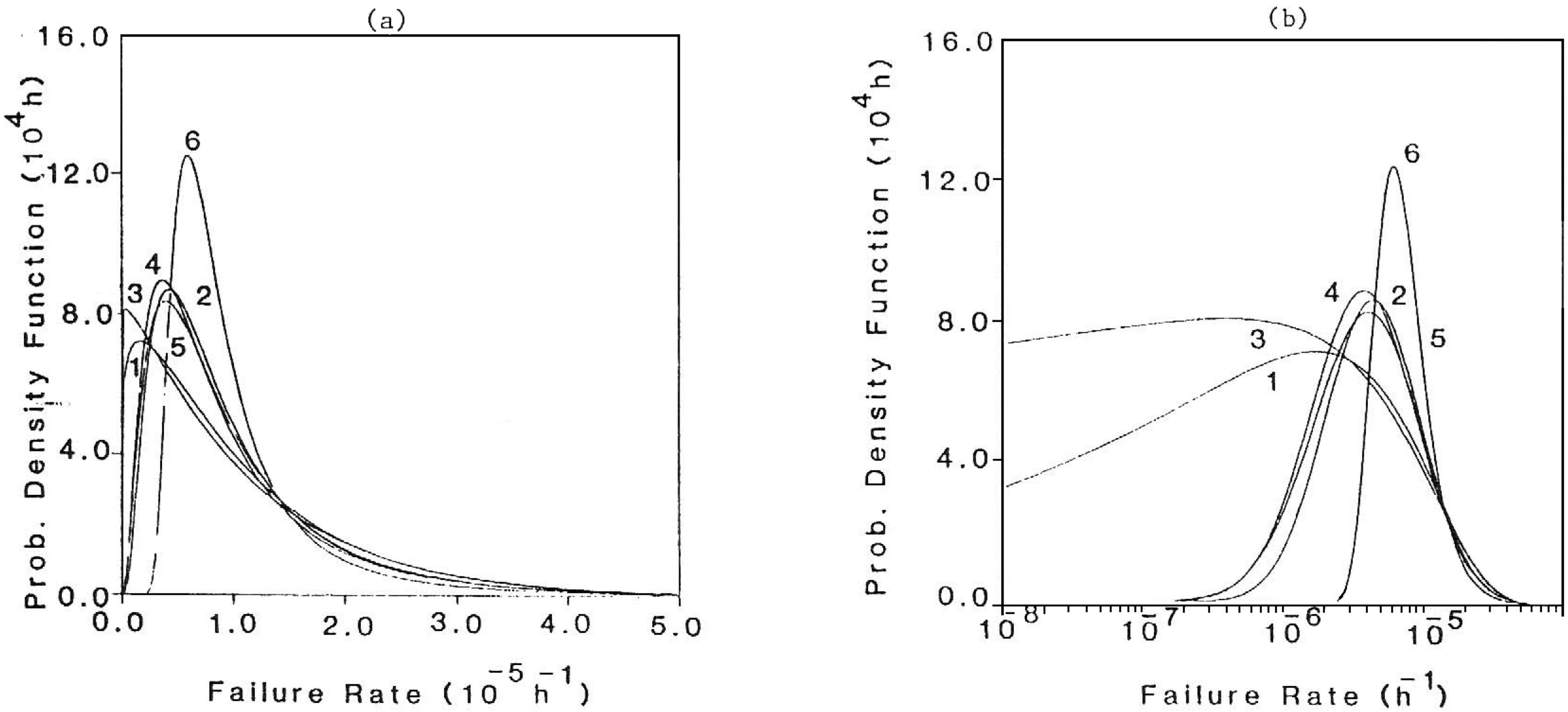

Figure 6.2: Estimated prior distributions obtained by the MMMM with good data, (a) linear (b) semi-log. The prior distributions were assumed to belong to the families: (1) Gamma, (2) Lognormal, (3) Weibull, (4) Logbeta, (5) Loggamma, and (6) Log-Weibull.

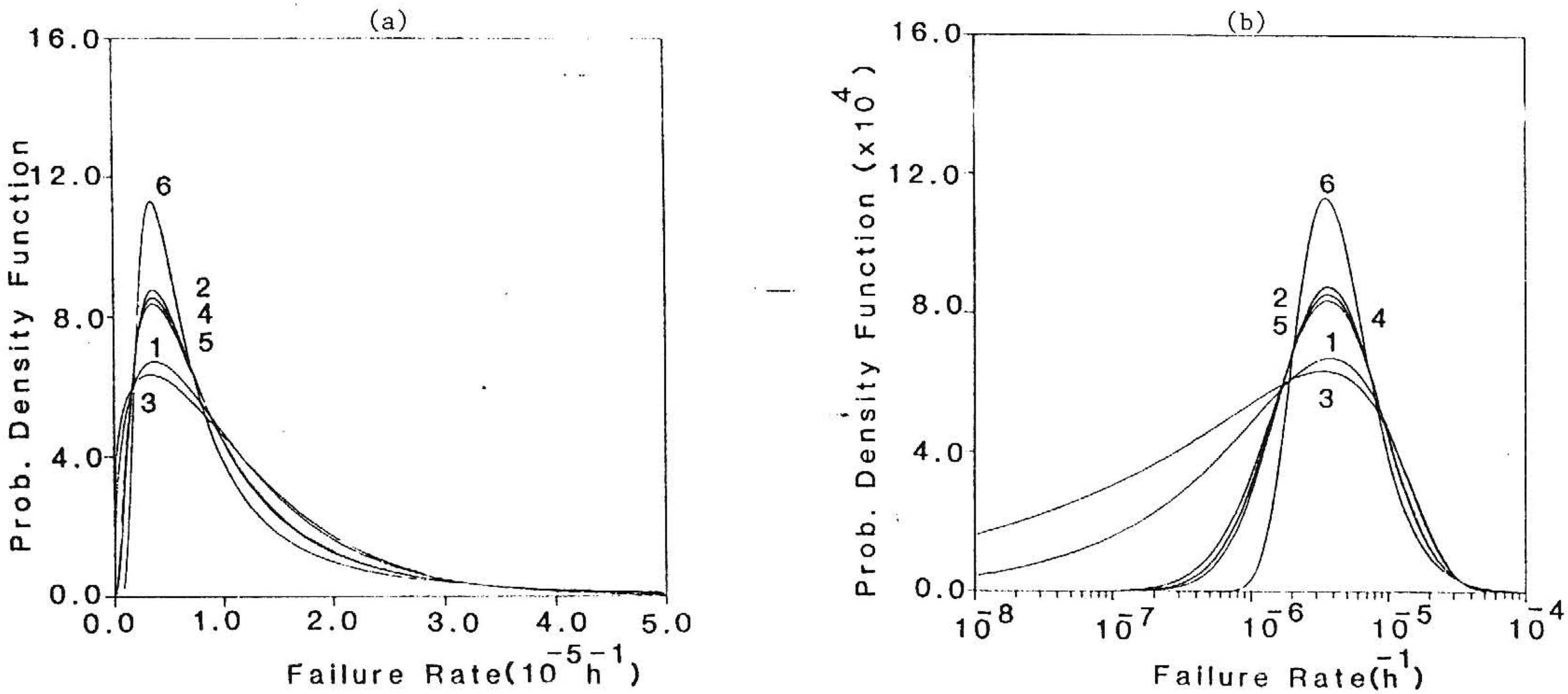

Figure 6.3: Estimated prior distributions obtained by the MMLM with GOOD data, (a) linear (b) semi-log. The prior distributions were assumed to belong to the families: (1) Gamma, (2) Lognormal, (3) Weibull, (4) Logbeta, (5) Loggamma, and (6) Log-Weibull.

TABLE 6.13

Percentiles of six prior distributions obtained from three estimation methods with GOOD data

Method	Prior Dist.	Percentiles 5-th	25-th	50-th	75-th	95-th
PMMM	gamma	.066(-5)	.333(-5)	.770(-5)	1.50(-5)	3.16(-5)
PMMM	Weibull	.062(-5)	.329(-5)	.769(-5)	1.50(-5)	3.16(-5)
PMMM	lognormal	.205(-5)	.449(-5)	.777(-5)	1.34(-5)	2.95(-5)
PMMM	logbeta	.190(-5)	.437(-5)	.773(-5)	1.36(-5)	2.99(-5)
PMMM	loggamma	.177(-5)	.428(-5)	.772(-5)	1.37(-5)	3.01(-5)
PMMM	logWeibull	.408(-5)	.591(-5)	.816(-5)	1.21(-5)	2.55(-5)
MMMM	gamma	.082(-5)	.366(-5)	.797(-5)	1.49(-5)	3.04(-5)
MMMM	Weibull	.074(-5)	.357(-5)	.798(-5)	1.50(-5)	3.05(-5)
MMMM	lognormal	.220(-5)	.470(-5)	.796(-5)	1.35(-5)	2.87(-5)
MMMM	logbeta	.185(-5)	.423(-5)	.745(-5)	1.30(-5)	2.85(-5)
MMMM	loggamma	.194(-5)	.451(-5)	.793(-5)	1.37(-5)	2.92(-5)
MMMM	logWeibull	.423(-5)	.605(-5)	.828(-5)	1.22(-5)	2.50(-5)
MMLM	gamma	.133(-5)	.445(-5)	.858(-5)	1.48(-5)	2.79(-5)
MMLM	Weibull	.112(-5)	.436(-5)	.874(-5)	1.51(-5)	2.78(-5)
MMLM	lognormal	.186(-5)	.429(-5)	.766(-5)	1.37(-5)	3.15(-5)
MMLM	logbeta	.181(-5)	.429(-5)	.773(-5)	1.38(-5)	3.11(-5)
MMLM	loggamma	.177(-5)	.430(-5)	.780(-5)	1.39(-5)	3.07(-5)
MMLM	logWeibull	.224(-5)	.412(-5)	.696(-5)	1.31(-5)	4.11(-5)

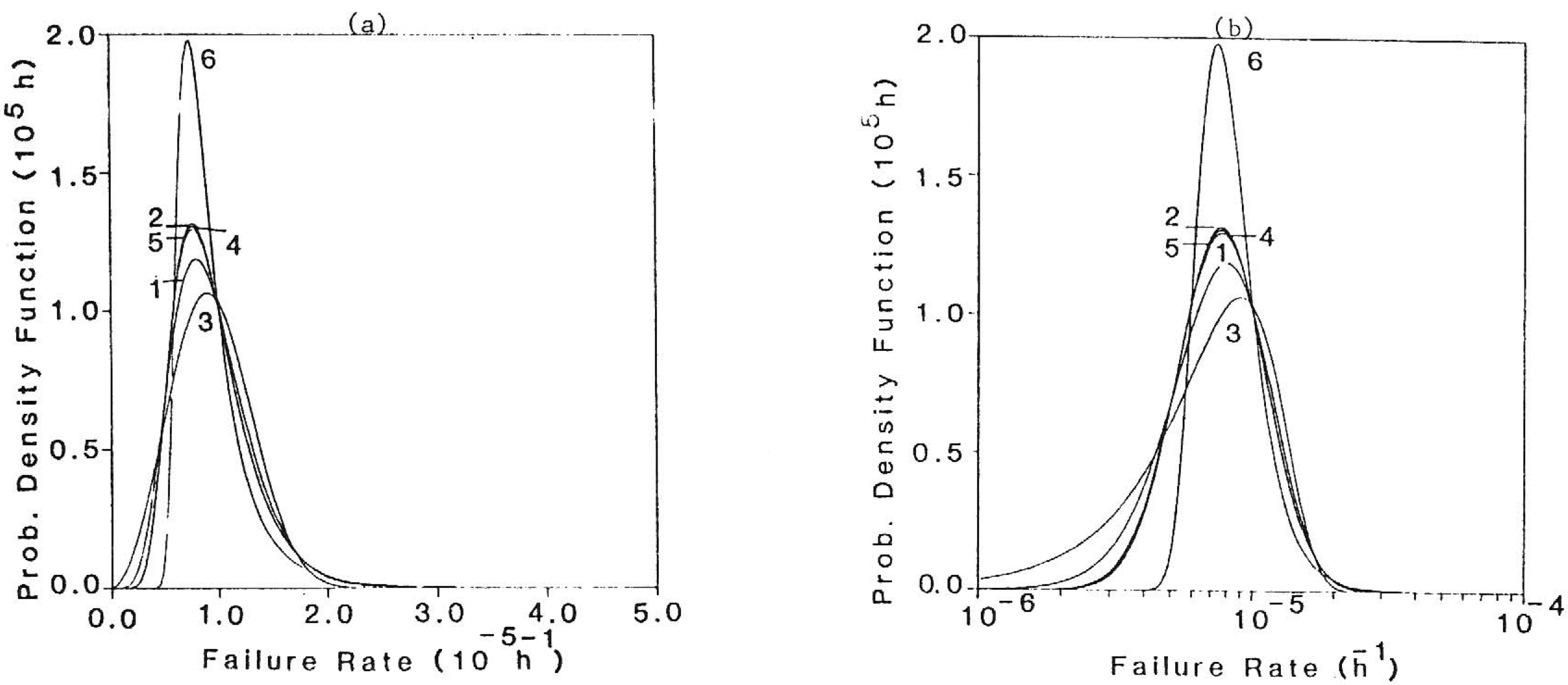

Figure 6.4: Estimated prior distributions obtained by the PMMM with BAD data, (a) linear (b) semi-log. The prior distributions were assumed to belong to the families: (1) Gamma, (2) Lognormal, (3) Weibull, (4) Logbeta, (5) Loggamma, and (6) Log-Weibull.

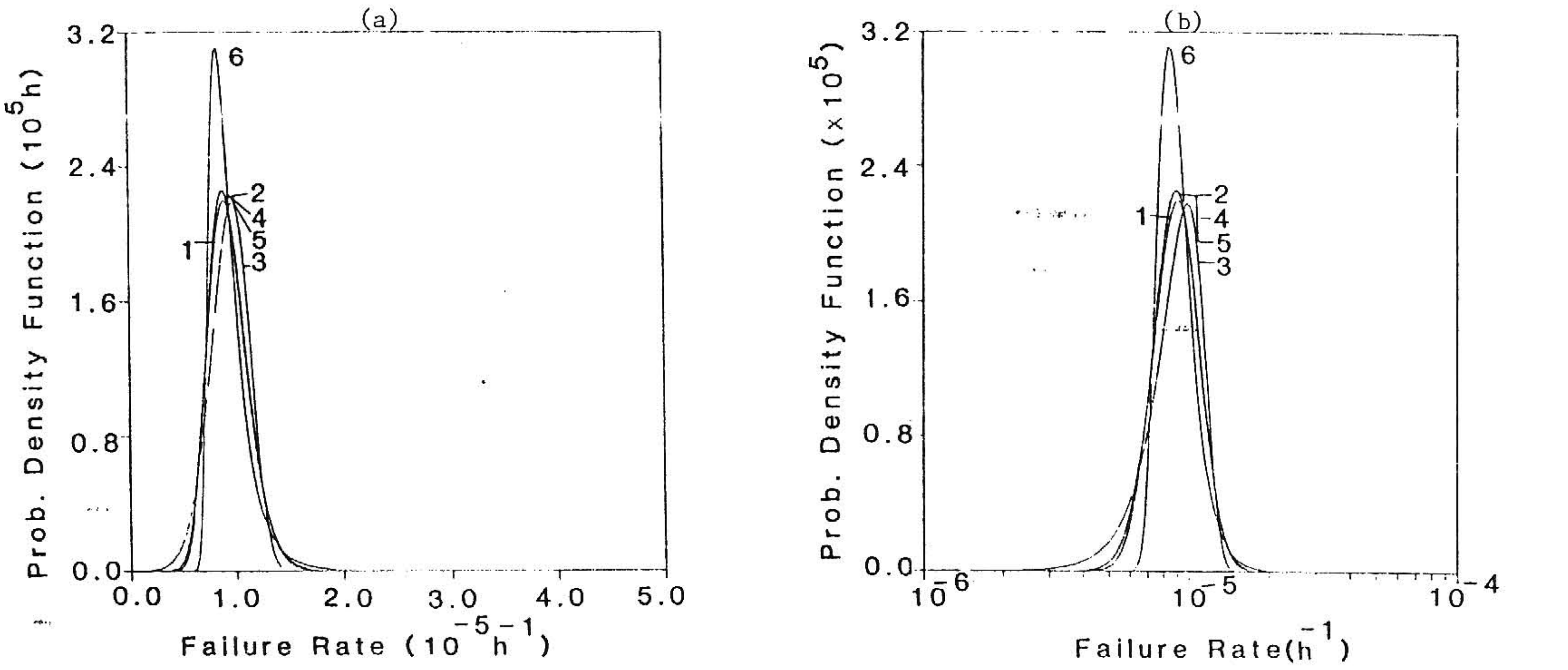

Figure 6.5: Estimated prior distributions obtained by the MMMM with BAD data, (a) linear (b) semi-log. The prior distributions were assumed to belong to the families: (1) Gamma, (2) Lognormal, (3) Weibull, (4) Logbeta, (5) Loggamma, and (6) Log-Weibull.

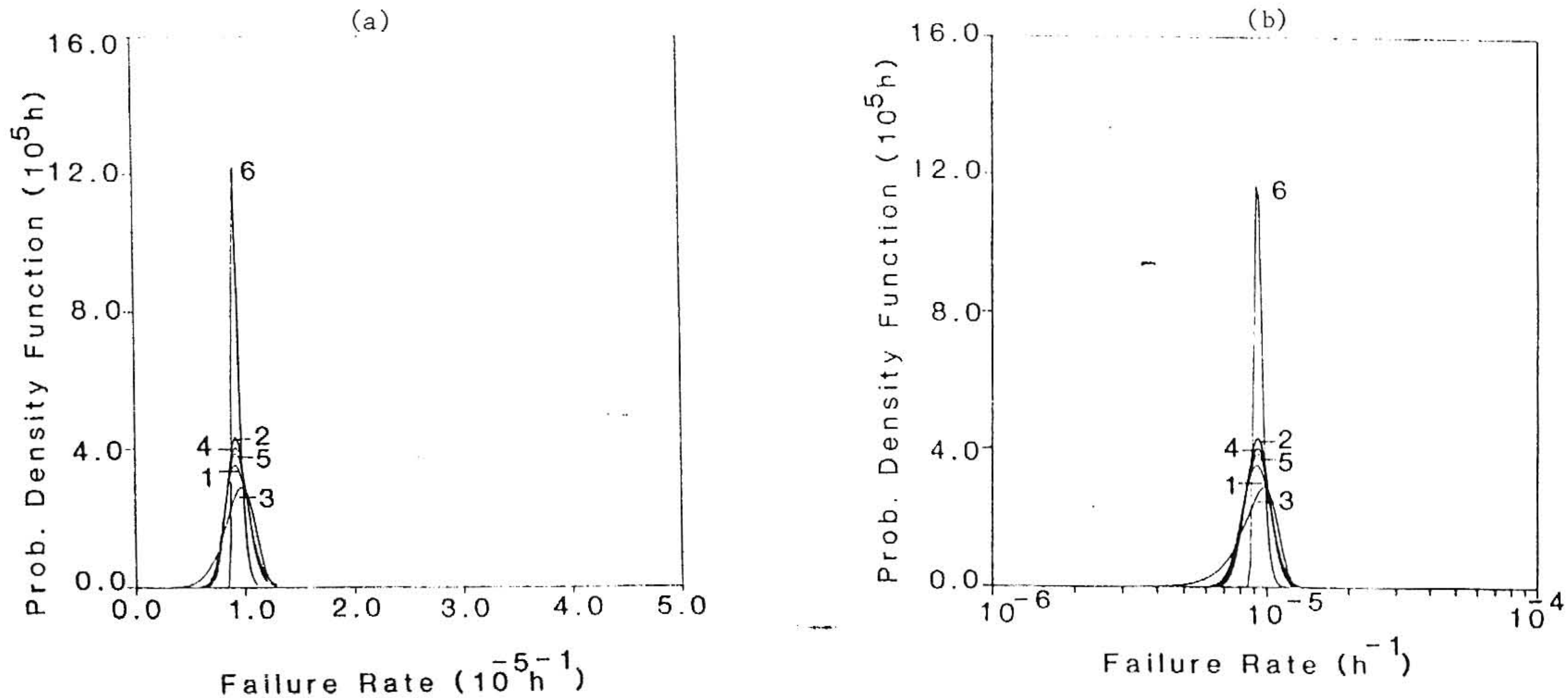

Figure 6.6: Estimated prior distributions obtained by the MMLM with BAD data, (a) linear (b) semi-log. The prior distributions were assumed to belong to the families: (1) Gamma, (2) Lognormal, (3) Weibull, (4) Logbeta, (5) Loggamma, and (6) Log-Weibull.

TABLE 6.14

Percentiles of six prior distributions obtained from three estimation methods with BAD data

Method	Prior Dist.	Percentiles				
		5-th	25-th	50-th	75-th	95-th
PMMM	gamma	.438(-5)	.681(-5)	.895(-5)	1.15(-5)	1.59(-5)
PMMM	Weibull	.372(-5)	.681(-5)	.928(-5)	1.18(-5)	1.55(-5)
PMMM	lognormal	.480(-5)	.685(-5)	.878(-5)	1.13(-5)	1.61(-5)
PMMM	logbeta	.476(-5)	.684(-5)	.879(-5)	1.13(-5)	1.61(-5)
PMMM	loggamma	.473(-5)	.684(-5)	.881(-5)	1.13(-5)	1.61(-5)
PMMM	logWeibull	.594(-5)	.719(-5)	.851(-5)	1.05(-5)	1.57(-5)
MMMM	gamma	.659(-5)	.810(-5)	.928(-5)	1.06(-5)	1.26(-5)
MMMM	Weibull	.614(-5)	.822(-5)	.953(-5)	1.07(-5)	1.22(-5)
MMMM	lognormal	.670(-5)	.809(-5)	.922(-5)	1.05(-5)	1.27(-5)
MMMM	logbeta	.669(-5)	.809(-5)	.923(-5)	1.05(-5)	1.27(-5)
MMMM	loggamma	.668(-5)	.809(-5)	.923(-5)	1.05(-5)	1.27(-5)
MMMM	logWeibull	.733(-5)	.816(-5)	.899(-5)	1.01(-5)	1.28(-5)
MMLM	gamma	.777(-5)	.868(-5)	.936(-5)	1.01(-5)	1.12(-5)
MMLM	Weibull	.683(-5)	.852(-5)	.954(-5)	1.04(-5)	1.15(-5)
MMLM	lognormal	.795(-5)	.875(-5)	.935(-5)	1.00(-5)	1.10(-5)
MMLM	logbeta	.785(-5)	.870(-5)	.934(-5)	1.00(-5)	1.11(-5)
MMLM	loggamma	.764(-5)	.860(-5)	.934(-5)	1.01(-5)	1.14(-5)
MMLM	logWeibull	.889(-5)	.912(-5)	.933(-5)	.960(-5)	1.02(-5)

Chapter 7

SUMMARY AND CONCLUSION

This chapter briefly summarizes this study and suggests areas for future research.

7.1 FAMILIES OF DISTRIBUTIONS

One of the objectives of this study was to examine the ability of several different prior distributions to model failure-rate data. This study incorporated the non-conjugate loggamma and logWeibull distributions into a compound failure-rate model, and thereby extended an earlier study[11] that used the conjugate gamma, non-conjugate Weibull, lognormal, and logbeta prior distributions.

In this study the non-conjugate loggamma and logWeibull distributions were chosen as possible alternatives to the conjugate gamma distribution in order to determine the effect the choice of the prior family has on the failure model. Data were simulated from these alternative distributions with their parameters at preset values, i.e.; the data were therefore generated from known distributions. The data simulation, as explained in section 6.1.1, was achieved by utilizing a random number generator program to obtain numbers between 0 and 1. These numbers were then taken as the values of the cummulative marginal distribution. Sets of failure data were then determined from the cummulative distribution thereby producing sets of failure data representative of known alternative distributions.

These sets of failure data were analyzed by using the conjugate gamma model. The same failure data sets were also analyzed by using the non-conjugate models from which they were simulated. The results of the conjugate gamma and non-conjugate analysis of the failure data were then compared. It was determined that the conjugate gamma model is as good, if not better, in analyzing failure data than the non-conjugate models, even when the failure data are known to come from the non-conjugate distribution.

7.2 ESTIMATION METHODS

In this study the results of an earlier study [11] were verified, namely that the method chosen to estimate values of the prior parameters is of greater importance than the choice of the distribution family used.

Three methods were used to estimate values of the prior parameters: (1) matching the data moments to those of the prior distribution (PMMM), (2) matching the moments of the marginal distribution to those of the data (MMMM), and (3) the maximum likelihood method based on the marginal distribution (MMLM).

The matching moments methods consisted of matching the sample moments of the given failure data to the corresponding distribution moments. Since two parameters needed to be estimated in order to determine the distribution shape and range, this produced two equations with two unknowns. These equations were solved either directly or by use of standard numerical techniques.

The Marginal Maximum Likelihood Method consists of determining the values of the two unknowns α and β which maximized the likelihood function for the given failure data. This is explained in section 4.2.

Each of the estimation techniques were used in estimating the parameters of the chosen prior families from the simulated failure data. Percentiles of the results and estimated prior distributions were compared to each other and to the known exact values. From these comparisons, it was determined that the method used to estimate the parameters of the chosen prior family has a greater effect than the selection of the prior family.

7.3 SUGGESTIONS FOR FUTURE STUDIES

7.3.1 Estimation Techniques

To augment this study and to obtain a better overall understanding of the different estimation techniques, one should in future work investigate several other methods of parameter estimation. For example, in the matching moments equations one could use the third and fourth moments "shape factors" (described in chapter 4). Kurtosis, a measure of the degree of flatness of a density near its center, could be utilized in estimating parameters of distributions.

7.3.2 Degradation of Components and Systems.

In this study, it was assumed the failure rate λ for a particular component was a constant and that the component if it should fail is immediately repaired and brought back into operation with its failure rate unchanged. It was also assumed that failures were independent of one another. These assumptions considerably simplified the development of the compound failure model, and experience with mechanical components show that they are doomed to fail with continual use. Components are degraded with use and it can be expected that their failure rates λ will generally change in time.

All plants are built with some form of redundant emergency back-up systems ("failure-on-demand problem"). That is, duplicate systems designed such that if one fails when signaled to operate, the other is required to start, i.e. back-up to a back-up system. Until lately is has been long standing with the Nuclear Regulatory Commission that if one of the safety back-up systems fail, then the remaining (2nd) redundant system must be tested every day to insure its operability until the first back-up system is repaired or replaced. The obvious conclusion to this practice is that the secondary system being tested every day will wear and its probability of failure will increase, i.e. pump failure, failed valve closure, etc., thereby , possibly leaving no back-up system and requiring the shutdown of the plant. One of the supporting truths to this is that a major portion of the wear and strain on a pump is during its start up and shut down - precisely the strains imposed by testing.

In order to include degradation of components and systems in the failure rate case, one would have a time dependent problem which would further complicate an already difficult task. The computational diffulties would mount with each step of mathematics required to compute results from a time-dependent compound failure model. However, with numerical procedures developed here and in earlier studies, it should be possible to reduce these more complex (and more realistic) failure models to numerics. One could change (increase) the value of λ after each test or use, based on the original constant failure rate assumed. This would account for the degradation of components during its operating mode. Since degradation is a reasonable assumption, this should produce more realistic results.

REFERENCES

1. J.K. Shultis, N.D. Eckhoff, D.E. Johnson, G.A. Milliken, D.D. Drayer, D.R. Gallup, and W. Buranapan, "Analysis of Failure Rate Data with Compound Models," NUREG/CR-2714, U.S. Nuclear Regulatory Commision, Washington (1982).

2. J.K. Shultis, D.E. Johnson, G.A. Milliken, and N.D. Eckhoff, "Use of Non-Conjugate Prior Distributions in Compound Failure Models," NUREG/CR-2374, U.S. Nuclear Regulatory Commission, Washington (1981).

3. A.M. Mood, F.A. Graybill, and D.C. Boes, <u>Introduction to the Theory of Statistics</u>, McGraw-Hill, New York, 1974.

4. M. Greenwood and A.U. Yule, J. Roy. Stat. Soc., <u>83</u>, 255 (1920).

5. J.C. Kleinman, "Proportions with Extraneous Variance: Single and Independent Samples," J. Am. Stat. Assoc., <u>68</u>, 46 (1973).

6. A. Hald, "Statistical Theory with Engineering Applications," Wiley, New York (1952).

7. N.L. Johnson and S. Kotz, "Discrete Distributions," Houghton Mifflin, New York (1969).

8. J. Kowalik and M.R. Osborne, "Methods for Unconstained Optimization Problems," Esevier, New York (1968).

9. R.W. Hornbeck, "Numerical Methods," Quantum, New York (1975).

10. M. Kendall and A. Stuart, "The Advanced Theory of Statistics," Vol. 2, 3rd ed., Griffin (Macmillan), New York (1973).

11. J.K. Shultis, N.D. Eckhoff, D.E. Johnson and G.A. Milliken, "SAFE-R and SAFE-D: Computer Codes for the Analysis of Failure Data," NUREG/CR-2375, U.S. Nuclear Regulatory Commision, Washington (1981).

ACKNOWLEDGMENTS

At this time I would like to express my appreciation to Mr. W. Buranapan and professors J. K. Shultis and M. Rubinson for their advise and guidance. Graditudes to Barbara Wieliczka for helping with the figures. My greatest thanks go to my wife Nancy and son Dayton, the true backbone of my achievements.

ANALYSIS OF COMPONENT FAILURE DATA BY NON-CONJUGATE COMPOUND FAILURE MODELS

by

JEFFREY H. SIMMONS

B.S., Kansas State University, 1981

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Nuclear Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1983

ABSTRACT

Compound or Bayesian failure models are often used to describe nuclear plant components with inherently low failure rates. In these models the failure rates of similar components are assumed to vary according to some unknown prior distribution, and yet to remain constant for each individual component. Usually, the conjugate gamma distribution is chosen to represent the prior distribution in the compound model because of the subsequent analytical simplification afforded by its use in the compound failure model.

Recently, the use of non-conjugate prior distributions in a compound failure-rate model was studied. The purpose of this study is to extend this earlier work by incorporating additional non-conjugate prior distributions in the compound failure-rate model. In particular, the use of non-conjugate loggamma and logWeibull distributions are studied as alternative distributions to represent the prior distribution.

Failure data were simulated from known loggamma-Poisson and log-Weibull-Poisson marginal distributions. Prior distributions were then estimated from these data by using the conjugate gamma-Poisson failure model as well as by the non-conjugate failure models from which the data was generated. Parameters for the various prior distributions were estimated from the simulated failure data by the following three methods: (i) matching the data moments to those of the prior distribution, (ii) matching the moments of the marginal distribution to those of the data, and (iii) the maximum likelihood method that estimates the prior parameters so as to maximize the probability of observing the actual data.

The main conclusion from this study is that the computationally simpler conjugate model can be used to describe failure data from components with inherently low failure probabilities even when the data is known to come from a non-conjugate distribution. Of far greater importance than the selection of the prior family is the method used to estimate the prior parameters from the failure data. For some data sets the three parameter estimation techniques are found to give quite different results.