THE DISTINCTION OF SIMULATED FAILURE DATA BY THE LIKELIHOOD RATIO TEST

by

DARRYL D. DRAYER

B.S., Kansas State University, 1980

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

1981

Approved by:

N. Dean Eckhoff

Major Professor

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

In some applications failure rates are viewed as random variables, following a known type of distribution, e.g., a gamma distribution. Often the parameters of the known type of failure rate distribution are unknown. The integration of the failure rate distribution times the probability of failure in a test time (e.g., a Poisson distribution) over all failure rate space (0 to ∞) yields a marginal probability distribution, i.e., the probability of failure during a given test time for a process with a known type of failure rate distribution. This marginal distribution leads to Bayesian analysis, or Bayesian estimation [1]. In Bayesian estimation sample evidence obtained through direct observation is combined with prior information about the failure rate distribution, to make estimates of the parameters of their distribution.

If a group of items or components are required to perform a certain task, the components either operate or fail to operate. For example, when switched on a light bulb will light or remain dark. These kinds of components may be viewed in two ways, classical or Bayesian.

When a group of similar components is viewed in a classical sense all of the components are assumed to be completely identical. Because they are assumed identical all of the components are assumed to have the same probability of failure when energized or the same probability of failure per unit time. For example, all the light bulbs produced in a certain batch may be assumed to have a 0.01% probability of failing to light when turned on (i.e., once per one hundred tries the light bulb will fail) and may have a 25% probability of failing during a year of operation. If the same group of components is viewed in a Bayesian sense, the components are not assumed to be identical but are viewed as having differences. Thus, instead of there being only one probability of failure on demand or only one failure rate, each component has its own probability of failure or failure rate. The collection of probabilities, one for each component, taken together form a failure rate probability distribution.

A failure rate distribution may be generated experimentally by collecting data on the failures of components. For example a collection of light bulbs may be set up to be turned on and off. A count made of the number of times a given light bulb is turned on and off before it fails. The collection of the number of times of failure during a testing time describes the failure rate distribution.

Much work has been done with functions which can be used to model failure rate distributions, such as which functions are suitable and how can the parameters of a failure rate distribution be estimated [2]. One of the desirable characteristics of a function which is to be used as a failure rate distribution is that it may assume many different shapes by varying the distribution parameters. This variability of shape is required so that the function used will match the actual shape of the fail-rate distribution.

After a modeling function has been chosen a group of components may be observed and a failure rate found for each component. With this information the parameters of the modeling function for the failure rate distribution may be estimated for that group of components.

Another separate group of components may be observed and a failure rate found for each component. From the collection of failure rates, estimates may be made of parameters of the failure rate distribution for that group of components. If the two groups of components tested are similar, e.g., 100 watt light bulbs made by different companies, the question may be asked, are all the components from the same failure rate distribution or does each collection of components come from a different failure rate distribution?

In order to answer this question a test must be made of the equality of parameters between the two collections. This question is of major importance to this research work.

In this work an attempt to distinguish between simulated failure data from different failure rate distribution is made. In addition an attempt is made to answer the questions how similar can failure rate distributions be before data from each cannot be distinguished?

In order to answer this question sets of failure data were simulated from an assumed failure rate distribution function with known parameters. The simulated data were used to estimate parameters of the failure rate distributions for the data sets. The likelihood ratio test was used to compare data sets. From the results of the likelihood ratio tests, power curves were generated. The power curves were used to draw conclusions about how dissimiliar distributions must be in order to be confident that data are from different distributions.

2 THEORY

2.1 Failure Models

The safe operation of a nuclear power plant requires the use of hundreds of systems and components. These systems and components can be divided into two groups, those which are normally active and those which are normally inactive. Normally active components would include items such as cooling pumps, motors, monitors, plus hundreds of others. These components remain in use the entire time the plant is operating. The other class of components, those which are normally inactive, are only energized under very special circumstances. Probably the best example of a normally inactive component is a standby diesel generator. The generator sits idle not operating until the plant loses power, at which point the generator is required to start with little delay.

A failure rate or a mean time to failure may be associated with components which are normally active. A failure rate (λ) is a measure of the average number of failures per unit time. An alternate way of viewing the failure rate the probability of experiencing F failures in T component hours of operation. For a given pump this means failing F times in T total hours of operation. The mean time to failure is the inverse of the failure rate ($1/\lambda$) and gives the expected length of time for a component to operate before failing. One would like to have a model which estimates the failure rate and thus the mean time to failure for a given group of components.

For components which are normally inactive one wishes to find the probability of a failure-on-demand (p). This means one wants to find the probability that when a component or system is activated, energized, or turned on, it will fail to do what it is intended to do. For example, the probability of failure-on-demand for a diesel generator is the probability that the diesel will not start when the start button is depressed.

2.1.1 Homogeneous Model

Two different models may be used to describe component and system failures. The first of these models, the homogeneous model, is the simpler of the two. In this model the failure rate, λ, or the failure-on-demand probability, p, (depending on which type of component is being considered) is assumed to be some unknown constant. This unknown is assumed to be equal for all elements in that particular class of components. For example, the failure rate of pumps of type x is assumed to be λ for each and every pump of type x.

For the failure-on-demand case the probability of observing F failures in T demands is given by the binominal distribution

$$f(F|p,t) = \binom{T}{F} p^F (1-p)^{T-F}, \qquad (2.1)$$

where p is the probability of failure-on-demand and is a fixed but generally unknown constant.

For the failure rate case it is assumed that a component which fails is immediately brought back into operation with its λ remaining unchanged. With this information the probability of having F failures in T amount of time is given by the Poisson distribution,

$$f(F|\lambda,T) = \frac{(\lambda T)^F e^{-\lambda T}}{\Gamma(F+1)}, \qquad (2.2)$$

where (F+1) is the gamma function.

The maximum likelihood method (to be explained later) yields an estimate for λ (for failures with Poisson probability of occurring) of

$$\hat{\lambda} = \sum_{i=1}^{n} F_i / \sum_{i=1}^{n} T_i, \qquad (2.3)$$

i.e., the total number of failures divided by the total test time. For components which are designed to have a very low probability of failure the classical homogeneous model often fails, i.e., it predicts failure rates to be zero, an uncomfortable prediction.

2.1.2 Compound Model

The second more complex model known as the compound model is better equipped to handle low probabilities of failure. In this model the failure parameter, either λ or p is assumed to vary from component-to-component within the class but remains constant for the component. The parameters are assumed to be distributed as some function $g(\xi)$, where ξ is either λ or p. This distribution is called the failure rate 1 on-demand prior distribution, (or prior distribution, for simplicity) since it may be determined by previous knowledge of the components in a given class. With $g(\xi)$ known, the probability of experiencing F failure in T demands (or in T amount of time) is given by the marginal distribution

$$h(F|T) = \int_{\text{all}\,\xi} f(F|\xi,T)\; g(\xi)\; d\xi, \qquad (2.4)$$

where $f(F|\xi,T)$ is called the likelihood or conditional distribution. For the failure rate case of $f(F|\xi,T)$ is the Poisson distribution [given previously Eq. (2.3)] and for the failure-on-demand case $f(F|\xi,)$ is the binomial distribution [Eq. (2.1)].

The marginal distribution, for $f(F|\xi,T)$ a Poisson and $g(\xi)$ a gamma distribution [see Eq. (2.7)] is given by

$$h(F|T) = \frac{\Gamma(F+\alpha)}{\Gamma(\alpha)\Gamma(F+1)} \frac{(T^F t^\alpha)}{(T+t)^{F+\alpha}} . \qquad (2.5)$$

This compound model yields a fairly simple analytical expression; the gamma and Poisson distributions are said to be conjugates.

The prior distribution, which is the distribution of the failure parameter ξ among some class of components, must be described by some function. There are many requirements that this function should satisfy [3] one of which is that it should have several free parameters. The free parameters should be able to vary the shape of the function over a wide range. Thus once a function has been chosen the task is to vary these parameters to fit the data. Six functions have been suggested as possible candidates for the failure-on-demand case. These functions are: a beta, the lognormals, two log gammas, and a logbeta. For failure-rate case suggested candidates are: gamma, Weibull, lognormal, and logbeta. The function most commonly used for the failure rate case and the one used here is the gamma function. Only the failure rate with the gamma distribution will be studied in this work because it is by far the most common one and the mathematics are greatly simplified.

2.2 Notes on the Gamma Distribution

Since the gamma distribution is to be used as the prior distribution, a brief discussion about the distribution is in order. The gamma distribution is given by

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma^{(\alpha)}-\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, \text{ for } x > 0,\ \alpha > 0,\ \beta > 0, \quad (2.6)$$

$$= 0 \qquad , \text{ elsewhere},$$

where $\Gamma(\alpha)$ is the gamma function evaluated at α. For convenience sometimes a parameter, τ, is defined, as $\tau \equiv 1/\beta$. Hence Eq. (2.6) becomes

$$f(x) = \frac{\tau^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-x\tau}, \text{ for } x > 0,\ \infty > 0,\ \beta > 0$$

$$= 0 \qquad , \text{ elsewhere}. \quad (2.7)$$

The mean and variance of the gamma distribution are given by [4]:

$$\mu = \alpha\beta, \quad (2.8)$$

and

$$\sigma^2 = \alpha\beta^2. \quad (2.9)$$

The mode or maximum point of the distribution is found at $x = (\alpha-1)$, provided that $\alpha > 1$ (the only case for which a mode exists). For the gamma distribution with $\alpha \leq 1$, the shape is nonmodel, exponentially shaped.

One of the main reasons the gamma distribution was chosen as a prior distribution is because of the large number of shapes the distribution can assume. The parameter, α, is called the shape parameter while β is called the scale parameter. By varying the value of α the shape of the distribution is changed. A listing is given below of how the distribution behaves for different ranges of the shape parameter [5]. Some typical plots are shown in Figs. 2.1, 2.2, 2.3, 2.4, and 2.5.

For

$\alpha < 1$ $\qquad f(0) = \infty$

$\qquad f'(0) = \infty$

THIS BOOK CONTAINS NUMEROUS PAGES WITH DIAGRAMS THAT ARE CROOKED COMPARED TO THE REST OF THE INFORMATION ON THE PAGE.

THIS IS AS RECEIVED FROM CUSTOMER.

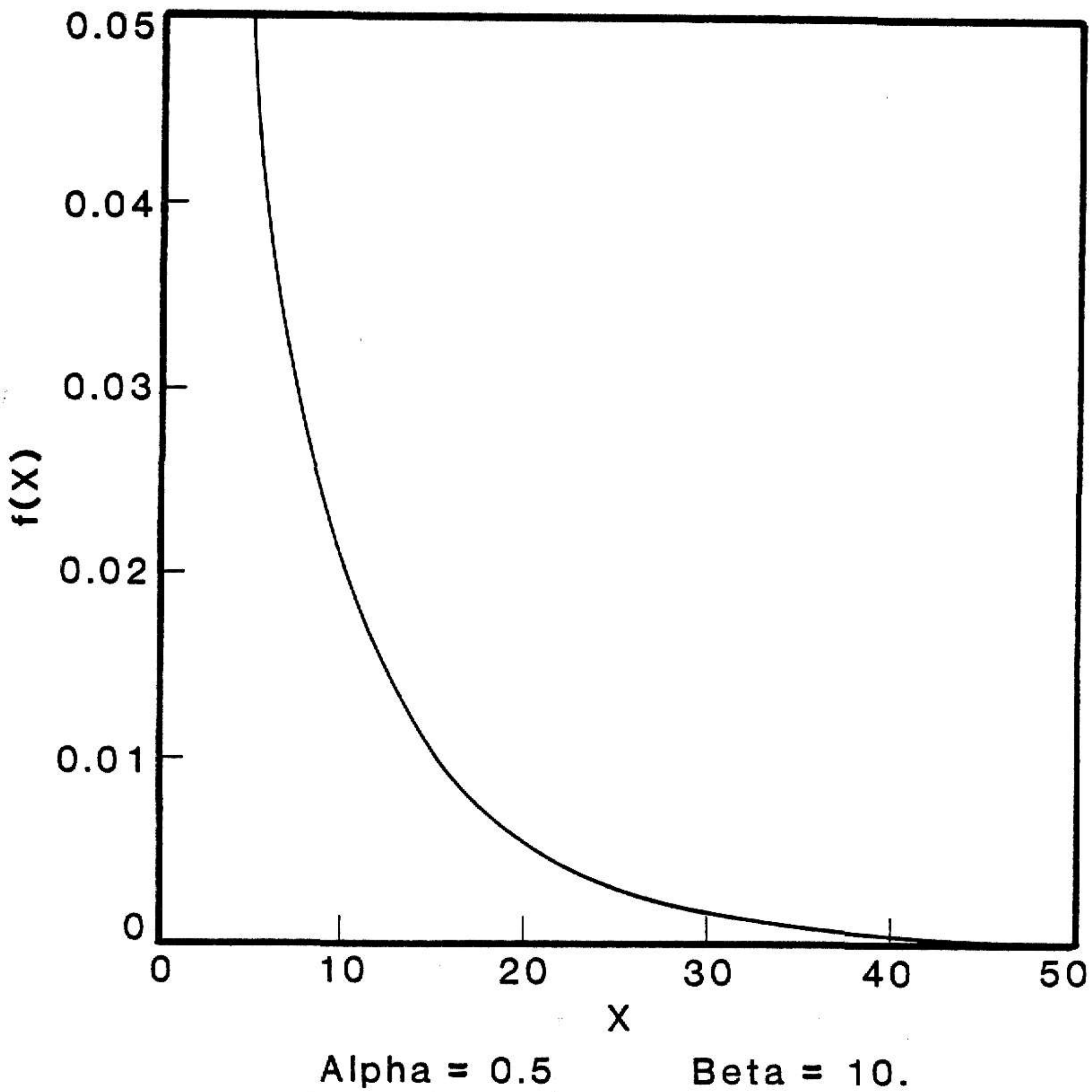

Figure 2.1 Gamma distribution with given parameters

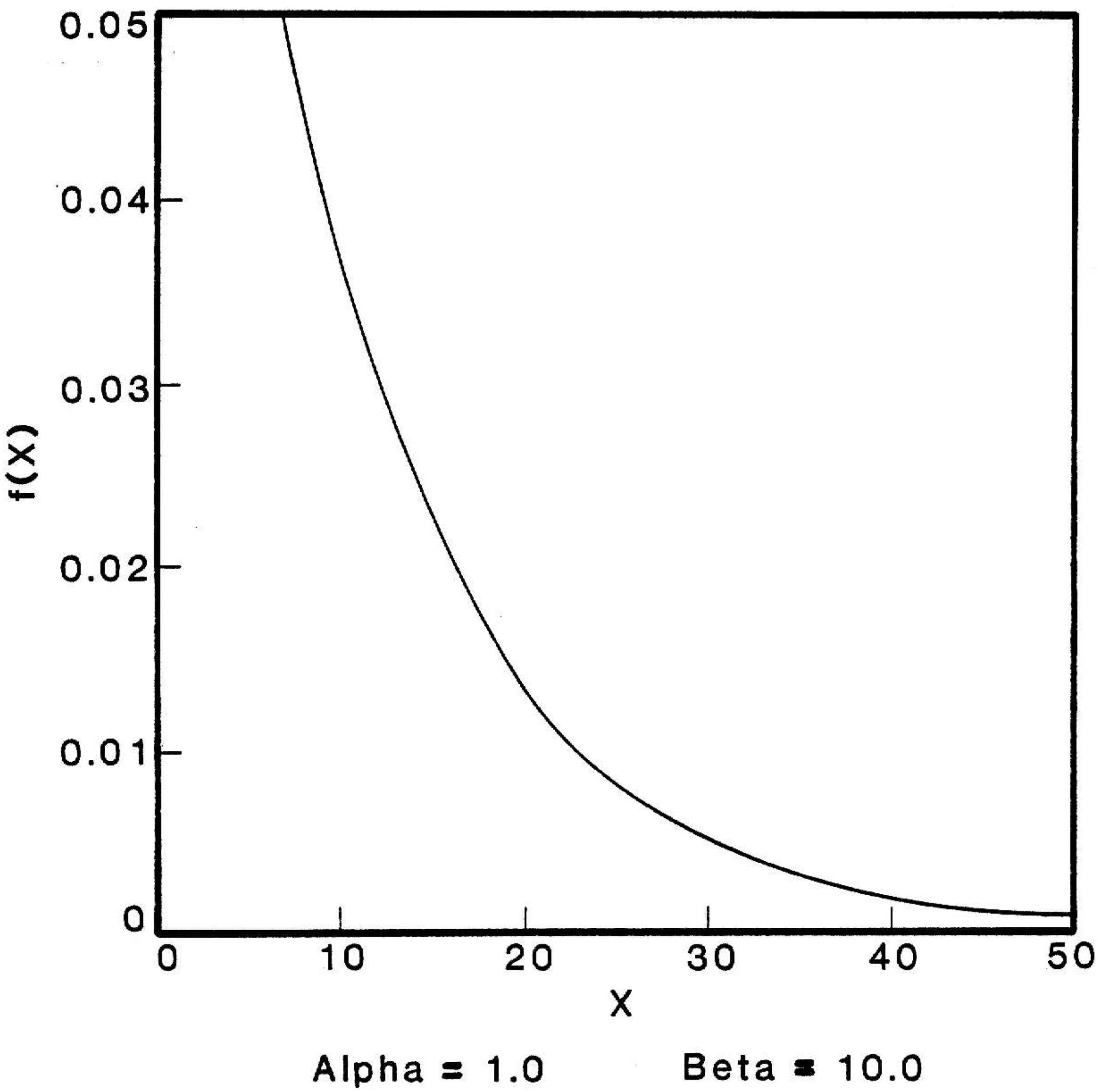

Figure 2.2 Gamma distribution with given parameters

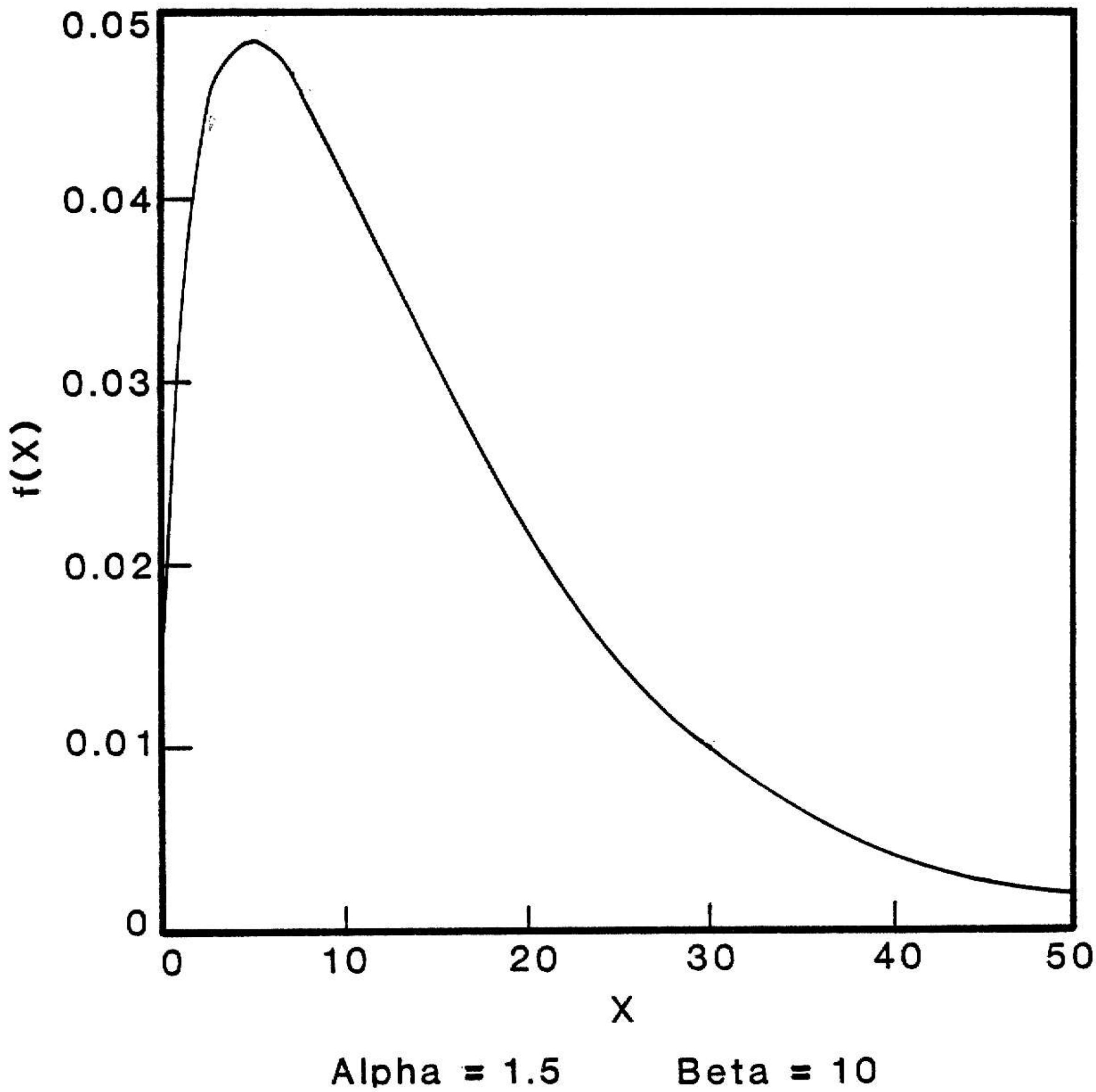

Figure 2.3 Gamma distribution with given parameters

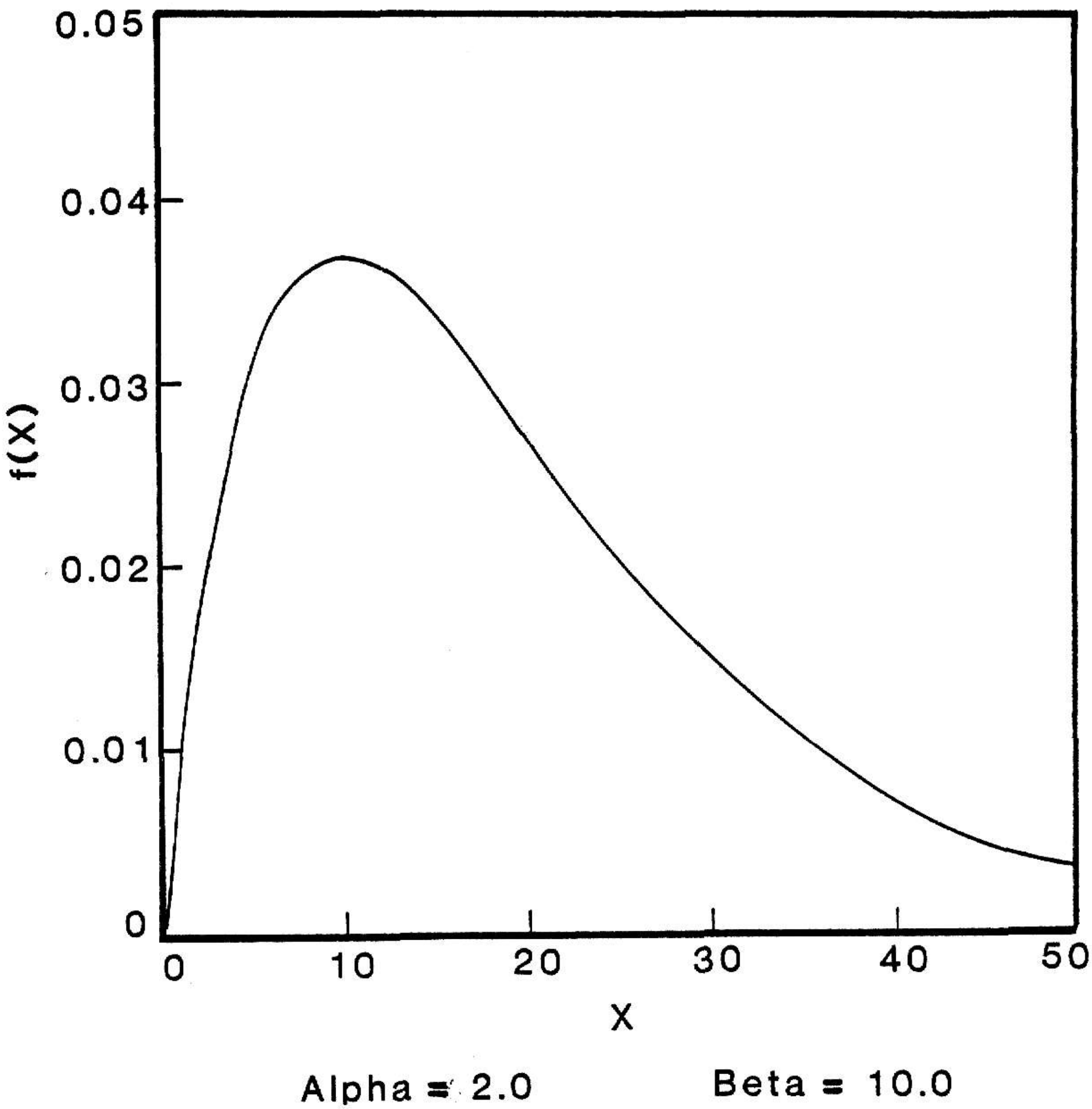

Figure 2.4 Gamma distribution with given parameters

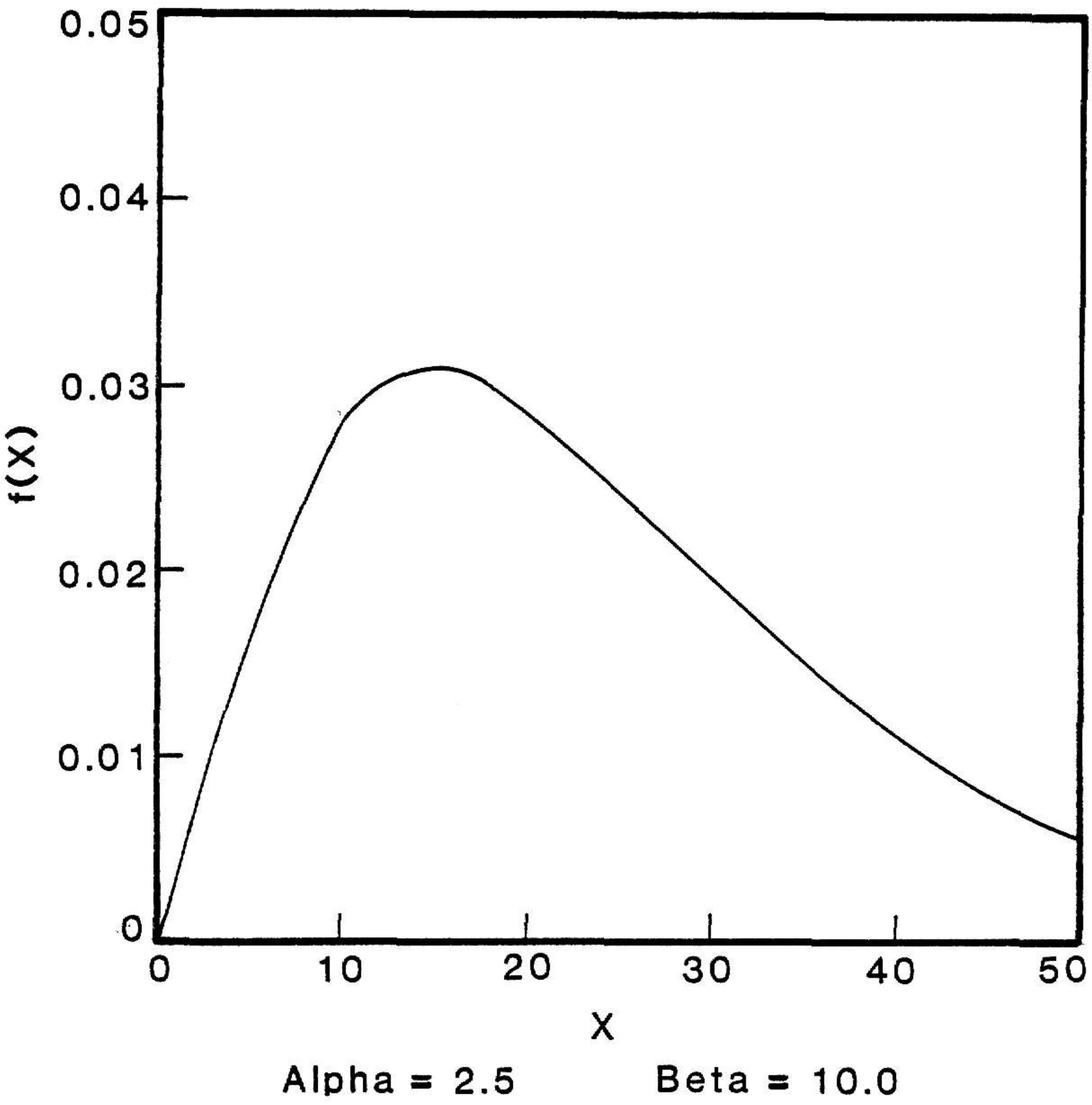

Figure 2.5 Gamma distribution with given parameters

<u>$\alpha = 1$</u> $\quad f(0) = \frac{1}{\beta}$

$f'(0) = \frac{-1}{\beta^2}$

<u>$1 < \alpha < 2$</u> $\quad f(0) = 0$

$f'(0) = \infty$

<u>$\alpha = 2$</u> $\quad f(0) = 0$

$f'(0) = \frac{1}{\beta^2}$

<u>$\alpha > 2$</u> $\quad f'(0) = 0$

$f''(0) = 0$

2.3 Estimation of Gamma Parameters

Several methods may be used for estimating the parameters of a gamma distribution from a set of failure rate data. In the program GAMMA MODIFIED (see Section 3.2.1) three empirical estimation techniques are used: i) matching the data moments to those of the prior distribution; ii) matching the data moments to those of the marginal distribution, and iii) maximizing the likelihood of the compound marginal distribution.

2.3.1 Matching Moments to the Prior Method (MMPM)

In the matching moments to the prior distribution method, the estimated failure rate for each component is given by $\hat{\lambda} = F_i/T_i$, for $i = 1, 2, \ldots, n$. The mean and the variance of the failure rate estimators $\bar{\lambda}$ and S^2 are calculated. These estimators are equated to the mean and variance of the gamma distribution (shown previously as $\mu = \alpha\beta$ and $\sigma^2 = \alpha\beta^2$).

Explicitly $\bar{\lambda}$ and S^2 may be written as

$$\bar{\lambda} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{F_i}{T_i}\right) \qquad (2.10)$$

and

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\lambda_i - \overline{\lambda})^2 = \frac{1}{n-1} \sum_{i=1}^{n} ((\frac{F_i}{T_i}) - \overline{\lambda})^2. \quad (2.11)$$

Equating $\overline{\lambda}$ and S^2 to the mean and variance of the gamma distribution gives

$$\mu(\alpha,\beta) = \overline{\lambda} \quad (2.12)$$

$$\sigma^2(\alpha,\beta) = S^2 \quad (2.13)$$

or

$$\hat{\alpha}\ \hat{\beta} = \overline{\lambda} \quad (2.14)$$

and

$$\hat{\alpha}\ \hat{\beta}^2 = S^2. \quad (2.15)$$

Solving Eq. (2.13) and (2.14) simultaneously yields the expressions

$$\hat{\alpha} = \overline{\lambda}^2/S^2 \quad (2.16)$$

and

$$\hat{\beta} = S^2/\overline{\lambda}, \quad (2.17)$$

for the estimates of the prior distribution parameters. The estimates, $\hat{\alpha}$ and $\hat{\beta}$, will always to positive.

2.3.2 Marginal Matching Moments Method (MMMM)

The marginal distribution $h(F_i|T_i,\alpha,\beta)$, gives the probability of having F failures in time T when a distribution with parameters α and β is specified. Thus the marginal distribution can be used instead of the prior distribution when estimating α and β by matching the data moments to the moments of the marginal distribution.

The data mean and variance, $\bar{\lambda}$ and S^2, are equated to the expected value of mean and variance for the marginal distribution. If w_1 and w_2 are defined as the mean and variance of any prior distribution, it can be shown for any prior distribution that the expected values of $\bar{\lambda}$ and S^2 are [2]

$$E(\bar{\lambda}) = w_1 \tag{2.18}$$

and

$$E(S^2) = w_2 + \frac{w_1}{n} \sum_{i=1}^{n} T_i{-1}. \tag{2.19}$$

Thus when the prior distribution is the gamma distribution with $w_1 = \alpha\beta$ and $w_2 = \alpha\beta^2$ the marginal matching moments equations are

$$E(\bar{\lambda}) = \bar{\lambda} = \alpha\beta \tag{2.20}$$

and

$$E(S^2) = S^2 = \alpha\beta^2 + \frac{\alpha\beta}{n} \sum_{i=1}^{n} T_i^{-1} \tag{2.21}$$

Solving these equation simultaneously yields the following for $\hat{\alpha}$ and $\hat{\beta}$.

$$\hat{\alpha} = \bar{\lambda}^2/[S^2 - \frac{\bar{\lambda}}{n} \sum_{i=1}^{n} T_i^{-1}] \tag{2.22}$$

$$\hat{\beta} = [S^2 - \frac{\bar{\lambda}}{n} \sum_{i=1}^{n} T_i^{-1}]/\bar{\lambda}. \tag{2.23}$$

These estimators are not necessarily positive. In the situation where $S^2/\lambda < \frac{1}{n} \sum_{i=1}^{n} T_i^{-1}$ negative values for the estimators will result. Negative values for $\hat{\alpha}$ and $\hat{\beta}$ have no meaning (i.e., the gamma distribution requires α and β to be greater than zero); so if negative estimators are encountered they should be ignored (i.e., the analysis of a set of data yielding negative estimators by MMMM is ignored).

2.3.3 Marginal Maximum Likelihood Method (MMMLM).

The maximum likelihood estimators are the values of α and β which maximize the likelihood function for the given failure data (F_i,T_i). The likelihood function and maximum likelihood estimators will be discussed in detail later but for continuity they are discussed briefly here.

In the compound model each component is assumed to be independent. Thus the probability L of observing (F_1,T_1), (F_2,T_2), ..., (F_n,T_n) is the product of the probability of observing each pair of (F_i,T_i) separately or

$$L(\alpha,\beta) = \prod_{i=1}^{n} \frac{\Gamma(F_i+\alpha)(T_i\beta)^{F_i}}{\Gamma(\alpha)\Gamma(F_i+1)(1+T_i\beta)(F_i+\alpha)(T_i\alpha)^{F_i}} \qquad (2.24)$$

Thus the maximum likelihood parameter estimators are those values of α and β which maximuze L (α,β) for the given failure data. It can be shown [3] that maximizing these equations leads to

$$\alpha = [\frac{F}{\beta} - \sum_{i=1}^{n} \frac{T_iF_i}{1+\beta T_i}] \bigg/ \sum_{i=1}^{n} \frac{T_i}{1+\beta T_i} \qquad (2.26)$$

and

$$\sum_{i=1}^{n} [\sum_{i=1}^{F_{i-1}} \frac{1}{\alpha+j} - \ln\,(1+\beta T_i)] = 0, \qquad (2.27)$$

where $F \equiv \Sigma F_i$ the total number of failures.

These equations cannot be solved analytically for α and β but they may be solved by numerical means [3]. If Eq. (2.26) is substituted into

Eq. (2.27) the resulting equation is a function of β only. The resulting only. The resulting equation may be solved for $\hat{\beta}$, which may then be substituted into Eq. (2.27) in order to find $\hat{\alpha}$. The value for $\hat{\beta}$ is found in the program GAMMA MODIFIED (see Section 3.2.1) using the subroutine DRTMI. DRTMI is a modification of the routine taken from the IBM "Scientific Subroutine Package".

2.4 Hypothesis Testing Concepts

Much of the work done with distributions such as the gamma, logbeta, lognormal, and many others is devoted to trying to estimate the values of the parameters of the function. For example, one may want to find the mean life of an electrically operated valve. The ultimate purpose of this information is generally to help answer a question about the valve. The question could be: will the valve have a long enough life for the intended application; is one company's valve better than another company's; does one manufacturing process create a longer lived valve than others; or many other possible questions. Answering any of these questions or making a decision as to which valve or process is better, should not be made by a guess but should be backed by clear statistical reasoning.

An experiment involving random variables may be performed and decisions may be made based on the values assumed by the random variables. The experiment may be a survey, a sample of product output, pressing the start button on a diesel generator, measuring the lifetime of a valve or almost anything. If the experiments are repeated the sample results will usually be different, e.g., a diesel may start 18 out of 20 times in the first experiment but only 15 out of 20 in the second. This demonstrates that there is a certain amount of uncertainty in the results. Chance occurance in a random sample does not follow a regular pattern. Thus one attempts to formulate procedures and rules of action for making decisions. The purpose is not to make statements which are 100% correct, that would be impossible, but to say things, make inferences which have a certain probability of being correct.

A statistical hypothesis is an assertion or conjecture about the distribution of one or more random variables. There are two types of statistical hypothesis: 1) a hypothesis that the population has a certain distribution function, e.g., normal or chi squared, 2) a hypothesis that a population parameter has a specified value, e.g., $\mu = 20$ or $\sigma_x = 15$. In this work we are concerned with the second type of hypothesis.

Suppose a new process has been proposed for making valves. It is hoped that the new process will produce valves with an operating lifetime longer than valves produced with the old process or at least as long. A hypothesis may be formulated which says: the proposed process is no better than the standard process. We hope that the hypothesis will be rejected.

There are several steps required in testing whether a hypothesis can be accepted or rejected.

1. Make a hypothesis regarding an event A of interest.
2. Select or construct a mathematical model that describes that probability, P(A).
3. Formulate a decision rule based on the model that specifies when to accept and when to reject the hypothesis.

Many times a primary or null hypothesis, H_o, is tested against some alternative hypothesis. This says that if the null hypothesis is not accepted the alternate hypothesis is accepted. For example, a pipe may have a diameter specification of 0.5 ± 0.004 in. A primary hypothesis is that the pipe being produced is centered at 0.500 in. An alternate hypothesis may be that the pipe being produced is centered at 0.503 in. The null hypothesis may be written as

$$H_o: \mu = 0.500,$$

and the alternate hypothesis may be written as

$$H_1: \mu = 0.503.$$

In another notation $\mu_o = 0.500$ and $\mu_1 = 0.503$.

If the two hypotheses, the null and the alternate, completely define the two distributions and all the other parameters are known then the hypothesis is known as a simple hypothesis. Suppose that under the null hypothesis the density function is $f(x;\mu_o)$ and under the alternate hypothesis the density function is $f(x;\mu_1)$. It is assumed that if there are any other parameters besides the mean μ they are known. Thus the two hypotheses completely define the distribution. If instead any of the parameters are not specified, the hypothesis is known as a composite hypothesis.

When a hypothesis is tested there are two possible actions, accept or reject H_o. However, when H_o is accepted or rejected there are four possible outcomes:

1. rejecting H_o when H_o is true
2. accepting H_o when H_o is true
3. accepting H_o when H_o is false
4. rejecting H_o when H_o is false.

The probabilities of these outcomes may be written as:

1. α = P(rejecting H_o when H_o is true)
2. $1-\alpha$ = P(accepting H_o when H_o is true)
3. β = P(accepting H_o when H_o is false)
4. $1-\beta$ = P(rejecting H_o when H_o is false)

The two possible decisions errors 1. and 3., above are called type I and type II errors. A type I error (or an error of the first kind) is a rejection of the null hypothesis when in fact it is true (Case 1, above). A type II error is an acceptance of H_o when it is false (Case 3, above). Thus α = P (type I error) and β = P(type II error). These concepts may be illustrated by a plot of density function.

Suppose that we plot the density function of some null hypothesis say $p\{x;\theta\}$ (see Fig. 2.6).

Because the plot is of a density function we know that the probability that an observation x taken at random from a population with the density function $p\{x\}$ lies between x_a and x_b is given by the area under the curve from x_a to x_b. To find that area one evaluates the integral

$$\int_{x_a}^{x_b} p\{x\}\, dx, \tag{2.28}$$

which may be graphically shown as in Fig. 2.7. Any probability density must satisfy the condition $p\{x\} \geq 0$ and

$$\int_{-\infty}^{\infty} p\{x\}\, dx = 1. \tag{2.29}$$

Suppose that we make some null hypothesis, H_o, which has a density function $p\{x,\theta_o\}$. We also suppose that the alternative hypothesis H_1 has a density function $P\{x,\theta_1\}$. The two hypotheses completely describe the distributions, if there are other parameters they are assumed known. We are going to take some observation x and we want a criterion for accepting or rejecting the null hypothesis. To do this we choose some value of x known as x_c. Choosing this x_c divides the density function into two

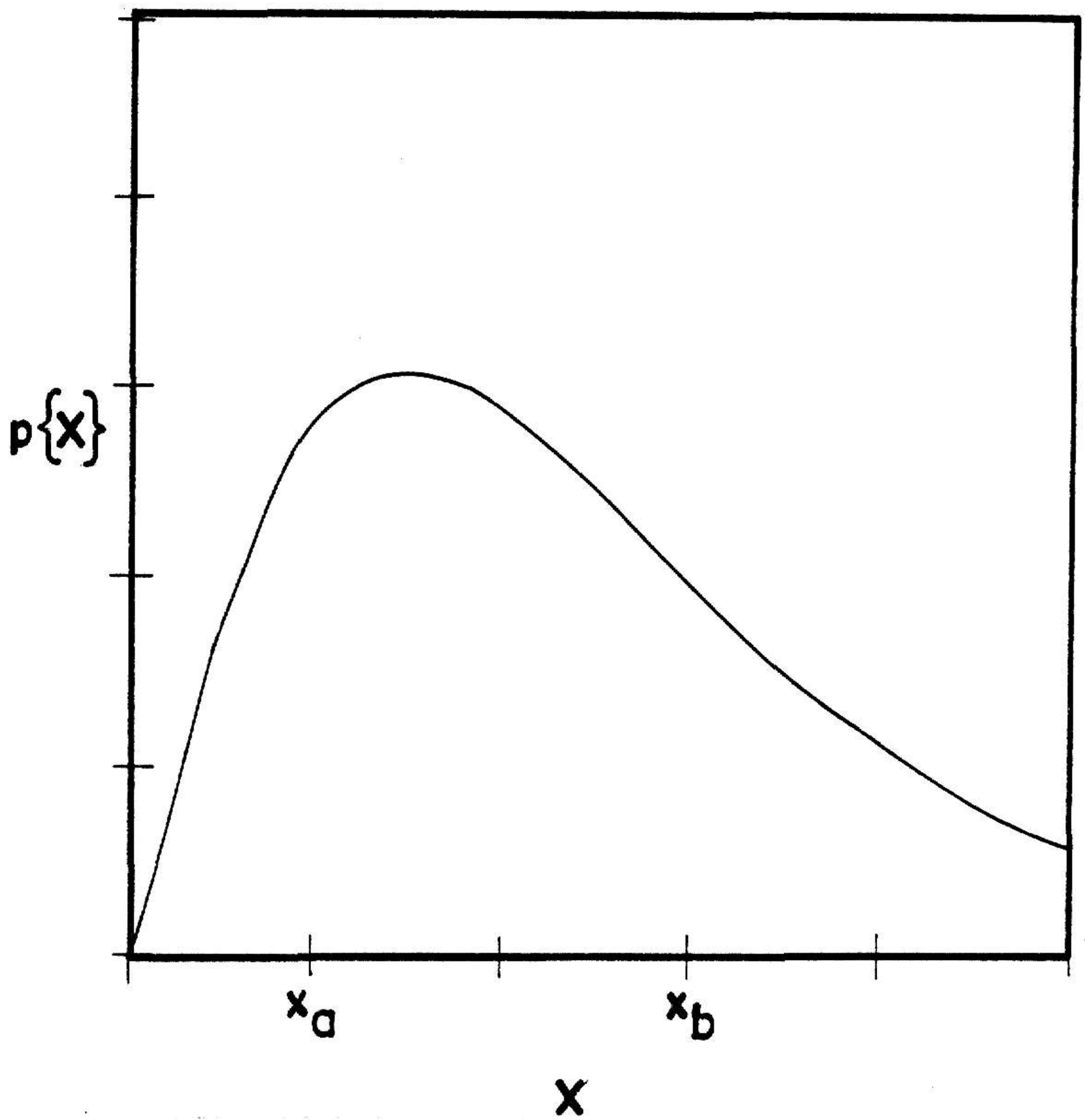

Figure 2.6 Density function

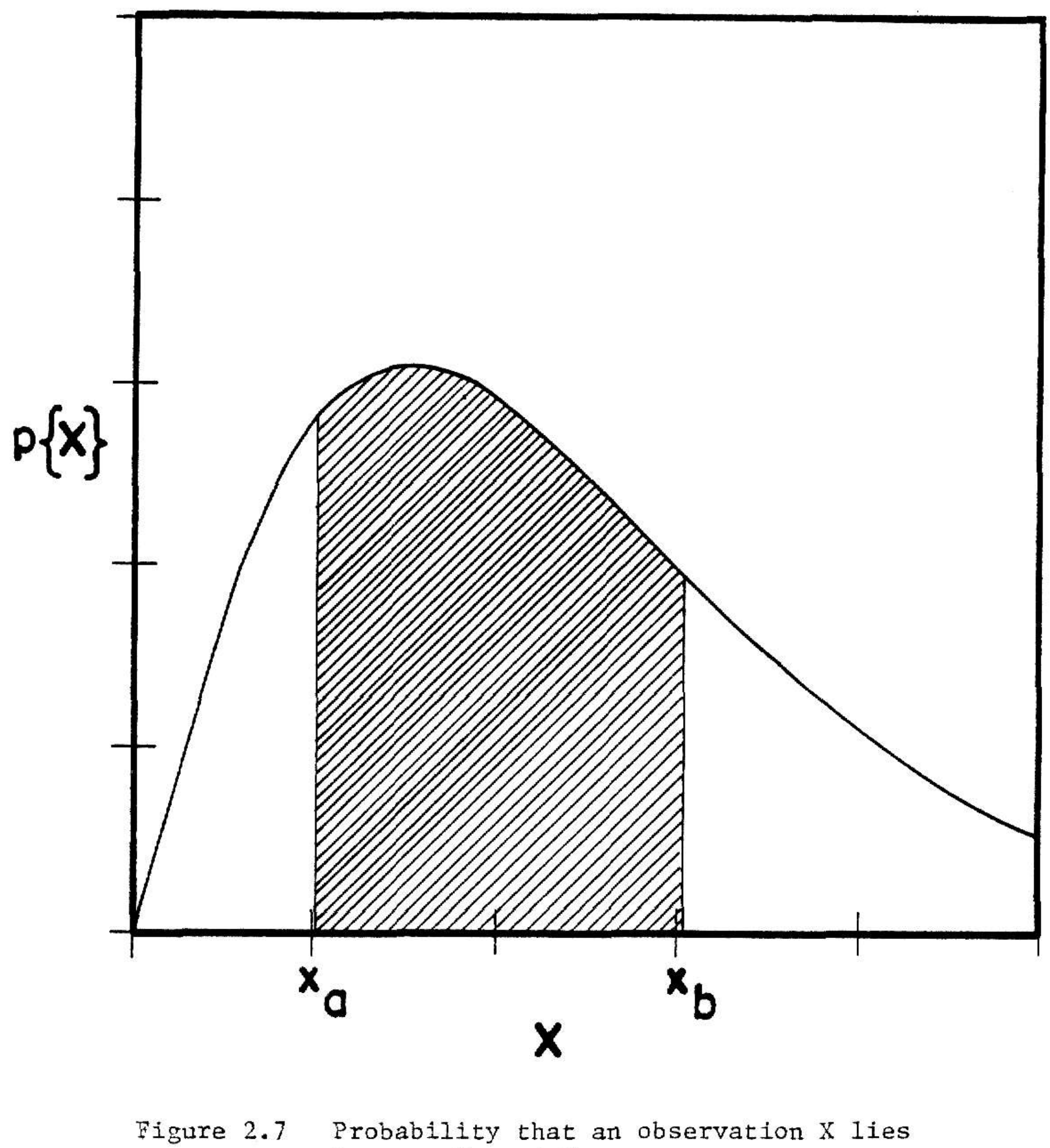

Figure 2.7 Probability that an observation X lies between X_a and X_b.

regions. If the observation falls in the region above x_c, the critical region, the null hypothesis is rejected.

The area of the critical region in the plot of the null hypothesis is α. In this region the null hypothesis has been rejected when in fact it is true, thus a type I error has been committed. The value for α can be controlled by choosing x_c. Commonly chosen values for α are 0.05 or 0.01 so that an error is only committed rarely. The numerical value for α is known as the significance level of the test. There are examples in everyday life where some hypothesis is made and tested with a small value of α. Probably the best example is the court room.

If a person is on trial for some crime they allegecly committed, the null hypothesis is made that the person is, in fact, innocent. The prosecution must prove beyond a reasonable doubt that the person is, in fact, guilty. The burden of proof is on the prosecution. The statement, "proof must be beyond a reasonable doubt," is equivalent to saying "α must be small". The court would rather let some criminals go free so that an innocent person is only very rarely convicted*.

2.5 Likelihood Ratio Test

The test that is used in this work to test the equality (or inequality) of prior parameters is the likelihood ratio test. This test can be used to determine whether two functions can be thought of as the same function. To do this a value called the likelihood is found for each function and these values are compared. From this comparison a determination of equality of inequality may be made. The details of this process follow.

The likelihood function of n random variables x_1, x_2 ..., x_n is defined as $L = p(x_1:\theta) \cdot p(x_2:\theta) \cdots p(x_n:\theta)$, where $p(x;\theta)$ is some distribution function which has one free parameter θ. If the function has more than one free parameter, then the likelihood function may be rewritten as

$$L(\theta_1,\theta_2, \ldots, \theta_k) = p(x_1,\theta_1,\theta_2, \ldots, \theta_k) \cdot p(x_2,\theta_1,\theta_2, \ldots \theta_k)$$

$$\cdots p(x_n,\theta_1,\theta_2, \ldots, \theta_k) \qquad (2.30)$$

or

$$L(\theta_1,\theta_2, \ldots \theta_k) = \prod_{i=1}^{n} p(x_i;\theta_1,\theta_2, \ldots, \theta_k), \qquad (2.31)$$

where k is the number of free parameters.

For an example of how the likelihood ratio test operates, assume we are given the normal distribution $f(x;\mu,\chi^2) = (2\pi\sigma^2)^{\frac{1}{2}} \exp(-(x-\mu)^2/2\sigma^2)$, where the distribution parameters are the mean, μ, and variance, σ^2. From the distribution we take a random sample composed of n elements, $x_1,x_2, \ldots, x_n$.

The likelihood is then defined as

$$L \equiv \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-(x_i-\mu)^2 / 2\sigma^2]. \qquad (2.32)$$

When actual values for x_i are substituted and the function is evaluated, L is equal to a real value.

---

*The Wichita Eagle of May 25, 1981 records a case of an individual who was convicted twice (after his first trial was overturned) and was due for yet a third trial (the prosecution was convinced he would be convicted a third time when the actual criminal confessed). Certainly in this case the α value was too large, allowing for a trial of another sort to recover lost wages and reputation!

Continuing the example; say one has some feel for what the parameter values are, the parameters in this case being μ and σ^2. Say $\mu = 0$ and $\sigma^2 = 1$. With this knowledge a null hypothesis may be made H_o: $\mu = 0$ and $\sigma^2 = 1$. Some alternate hypothesis, H_1, may also be made, for example H_1: $\mu = 1$ and $\sigma^2 = 2$. The likelihood for the null hypothesis would be

$$L_o = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi 1}} \exp[-\frac{1}{2\cdot 1}(x_i-0)^2] \tag{2.33}$$

which reduces to

$$L_o = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp[-\frac{1}{2}(x_i)^2]. \tag{2.34}$$

The likelihood of the alternate hypothesis is

$$L_1 = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi 2}} \exp[-\frac{1}{2\,2}(x_i-1)^2 \tag{2.35}$$

which reduces to

$$L_1 = \prod_{i=1}^{n} \frac{1}{2\sqrt{\pi}} \exp[-\frac{1}{4}(x_i-1)^2]. \tag{2.36}$$

The likelihood ratio is defined as $L \equiv L_o/L_1$. In this case the likelihood ratio is

$$L = \prod_{i=1}^{n} 2^{\frac{1}{2}} \exp[-\frac{1}{2}(x_i)^2 + \frac{1}{4}(x_i-1)^2]. \tag{2.37}$$

The value for the likelihood ratio will lie somewhere between zero and infinity. If $L = 1$ then the value of the likelihood of the null hypothesis is equal to the value of the alternate hypothesis, i.e., $L_o = L_1$.

Likewise, if the likelihood ratio is less than one, the likelihood of the alternate hypothesis is greater than the likelihood of the null hypothesis.

If we are given some function we will want to find the values of the parameters which yield the maximum value for the likelihood. The usual method for finding a maximum of a function is applied. For a function with only one parameter this means taking the first derivative with respect to the parameter, setting equal to zero, and solving for the parameter. For functions with multiple parameters the likelihood may be maximized by setting the partial derivative of L with respect to the parameters equal to zero, i.e.,

$$\frac{\partial L}{\partial \theta_1} = \frac{\partial L}{\partial \theta_2} = \ldots = \frac{\partial L}{\partial \theta_n} = 0. \tag{2.38}$$

Many times when maximizing the likelihood it is more convenient to maximize its logarithm, which will give the same result. For multiple parameters this would be

$$\frac{\partial \ln L}{\partial \theta_1} = \frac{\partial \ln L}{\partial \theta_2} = \ldots = \frac{\partial \ln L}{\partial \theta_n} = 0. \tag{2.39}$$

The values of the parameters obtained are called the maximum likelihood estimators (MLEs). Sometimes the value of one parameter will depend upon the value of another parameter. In this case a set of simultaneous equations must be solved. The normal distribution again provides a good example.

The likelihood function of the normal is:

$$L = \frac{1}{\sqrt{2\Pi\sigma^2}} \exp[-\frac{1}{2\sigma^2} (x_1-\mu)^2] \ldots \frac{1}{\sqrt{2\Pi\sigma^2}} \exp[-\frac{1}{2\sigma^2} (x_n-\mu)^2], \tag{2.40}$$

or

$$L = (\frac{1}{\sqrt{2\Pi\sigma^2}})^n \exp[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\mu)]^2. \tag{2.41}$$

The logarithm of the likelihood is

$$\ln(L) = n\ln(\frac{1}{\sqrt{2\pi\sigma^2}}) - \frac{1}{2\sigma^2} [\sum_{i=1}^{n} (x_i-\mu)^2. \tag{2.42}$$

Taking the derivative with respect to the mean, μ, setting the derivative to zero, and solving yields

$$\sum_{i=1}^{n} (x_i-\hat{\mu}) = 0, \tag{2.43}$$

which may be reduced to

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i. \tag{2.44}$$

To obtain the maximum likelihood estimation of σ^2 the logarithm of L is differentiated with respect to σ^2; the result is equated to zero; and since $\hat{\mu}$ is already known, [Eq. (2.46)] substitution for $\hat{\mu}$ gives

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i-\hat{\mu})^2. \tag{2.46}$$

In the preceding the value of σ^2 was maximized; we might just as easily have decided to maximize σ. It can be shown [6] that when σ is maximized the result is

$$\hat{\sigma} = [\frac{1}{n} \sum_{i=1}^{n} (x_i - \frac{1}{n} \sum_{i=1}^{n} x_i).^{\frac{1}{2}} \tag{2.47}$$

Squaring, the result becomes

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \frac{1}{n} \sum_{i=1}^{n} x_i). \tag{2.48}$$

Comparing this result to that of $\widehat{\sigma^2}$, Eq. (2.46) we see that

$$\hat{\sigma}^2 = \widehat{\sigma^2}. \tag{2.49}$$

Equation (2.49) demonstrates a general property of maximum likelihood estimators called invariance. Invariance says that if some function of θ, say $f(\theta)$, is a single-valued function of θ, and $\hat{\theta}$ is the maximum likelihood estimator of θ, then $\widehat{f(\theta)}$ is the maximum likelihood estimator $f(\theta)$, [7], i.e.,

$$f(\hat{\theta}) = \widehat{f(\theta)}. \tag{2.50}$$

Assume that we have information on the failure rates of a certain type of pump that two companies produce. For company one we know that the distribution of the pump failure-rate has the form $g(\alpha_1, \beta_1)$. For company two the form is $g(\alpha_2, \beta_2)$. A random sampling of failure rates may be obtained from each distribution. For company one the sampling would consist of the points F_{i1}/T_{i1} where $i = 1, 2, \ldots, n$. For company two the set of failures per unit time consists of the points F_{i2}/T_{i2}, where i = 1, 2, ..., m. The question arises can all these data be thought of as coming from one distribution or must each company's data be treated separately. In other words when dealing with these pumps can we treat all the pumps as having the same failure rate distribution or will we have to say pumps from, company one have a certain failure rate distribution and those from company two a different distribution. To make a determination of whether the failure rate distributions for pumps from the two companies are the same, the likelihood ratio may be employed.

Two hypotheses are made. The null hypothesis, H_o, is that all the data are from the same distribution. The alternate hypothesis, H_1, is that the data are from different distributions. Stated explicitly;

$$H_o: \ \alpha_1 = \alpha_2 \text{ and } \beta_1 = \beta_2 \tag{2.51}$$

$$H_1: \ \alpha_1 \neq \alpha_2 \text{ or } \beta_1 \neq \beta_2 \text{ or both.}$$

The likelihood of the null hypothesis is

$$L_o = \prod_{i=1}^{n} h(\frac{F_{i1}}{T_{i1}} ; \alpha_o, \beta_o) \prod_{i=1}^{m} h(\frac{F_{i2}}{T_{i2}} ; \alpha_o, \beta_o). \tag{2.52}$$

Since the null hypothesis says that all the data are from the same distribution, some $\hat{\alpha}_o$ may be found to maximize L_0. The likelihood of the alternate hypothesis is

$$L_1 = \prod_{i=1}^{n} h(\frac{F_{i1}}{T_{i1}}; \alpha_1, \beta_1) \prod_{i=1}^{m} h(\frac{F_{i2}}{T_{i2}}; \alpha_2, \beta_2) \tag{2.53}$$

The alternate hypothesis says that the parameters of the first distribution are not necessarily equal to the parameters of the second one; thus the maximum likelihood estimators are not necessarily the same for both groups. Equation (2.53) can be thought of as the likelihood of a function with four free parameters $\alpha_1, \alpha_2, \beta_1$, and $_2$. The maximum likelihood estimators are found in the standard way. The values which maximize L_1 are $\alpha_1, \alpha_2, \beta_1$, and β_2.

The likelihood ratio, L, is L_o/L_1 with the MLEs applied. The likelihood ratio will always be less than or equal to one. When the MLEs are found for Eq. (2.52) the entire function is maximized. In Eq. (2.53), when the estimators are found, both sections of the function are maximized separately. This means that if the distributions are the same maximizing each section separately makes no difference as compared to maximizing the entire function. If the distributions are different, maximizing separately will yield a larger value than maximizing as one. Thus if L is close to unity then we do not have to reject H_o, it can be said that the distributions are the same. On the other hand if L is close to zero, H_o must be rejected. We must say the data came from different distributions.

How close is close? To find how close is close enough to accept or reject, we define a test statistic T as $T \equiv -2\log L$. This test statistic can be shown to be distributed as Chi-squared, distribution with the number of degrees of freedom equal to the number of free parameters [4]. Thus, if we want to be 95% confident of our hypothesis H_o, we will say reject H_o, if $T > \chi^2$ (2,0.05), where 2 is the number of degrees of freedom, and 0.05 is the significance level.

2.6 Data Simulation

Often it is necessary to know how some system will perform or react to a stimulus or change, but direct experimentation is impossible. An example of this problem would be to try to predict how some epidemic is going to spread. Obviously one does not want to infect a population to find out what will happen. In order to predict what may happen without performing an actual experiment, simulation techniques may be used. In our case we need to make estimates of how many failures we expect when the probability of failure is described by a certain distribution.

When the prior distribution being used is the gamma distribution, we say previously (Eq. (2.5)) that the marginal distribution is:

$$h(F|T,\alpha,\beta) = \frac{(F+\alpha)}{\Gamma(\alpha)\Gamma(F+1)} \frac{T^F t\alpha}{(T+\tau)^{F+\alpha}}, \qquad (2.54)$$

where $\tau = 1/\beta$. This gives the probability of having F failures in a given time T with the parameters of the gamma function being α and β. Since F can only have integer values it is not difficult to form a discrete cumulative distribution. An example of a cumulative distribution generated from a distribution with the parameters; T = 10,000 hrs., α = 1.1, and β = 20,000 is shown in Fig. 2.8

The cumulative distribution is formed by starting with the probability of zero failures and adding to it the probability of one failure and then adding to that two failures etc. until the cumulative probability approaches one. Thus, the cumulative distribution is the sum (for the discrete case) or area (for the continuous case) under the probability distribution function up to a specified value of the independent variable. The value of the cumulative distribution will range between zero and unity.

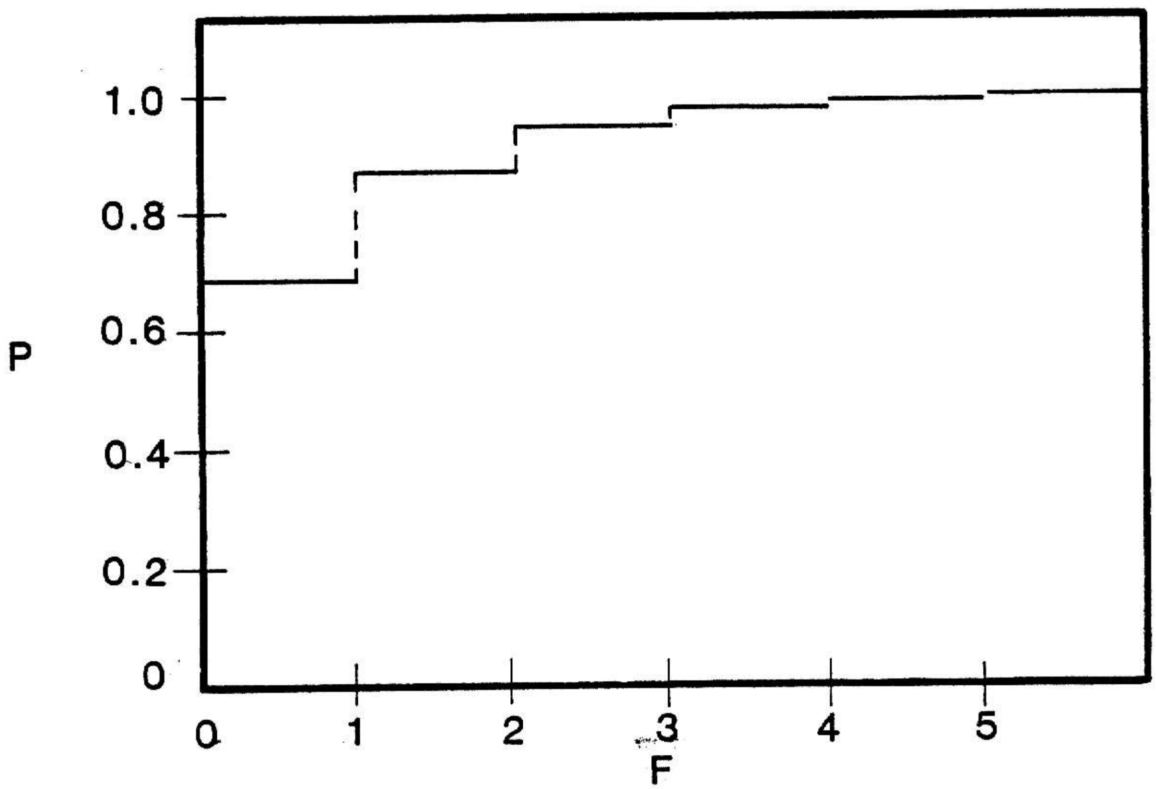

Figure 2.8 Cummulative distribution generated from distribution with T = 10,000, α = 1.1, β = 20,000.

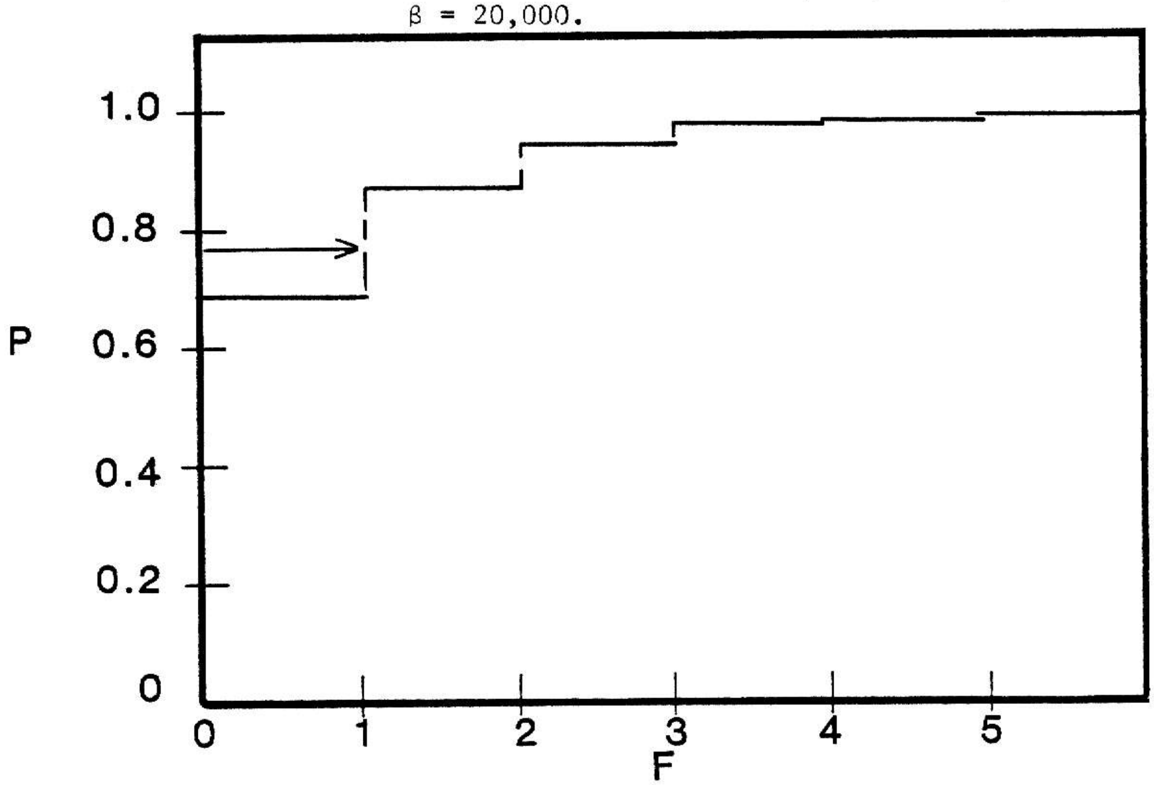

Figure 2.9 Number of failures corresponding to a generated random number of 0.781.

For simulation of data a convenient procedure is to generate a random number between zero and unity. Use this number as the value of the cumulative distribution for which a value of the independent variable (in our case number of failures) is sought. For example, if the generated random number is 0.781 the corresponding number of failures (using the above distribution) is 1, as shown in Fig. 2.9. This is the simulation of one datum from a distribution with T = 10,000 hrs., $\alpha = 1.1$, and $\beta = 20{,}000$ ($\tau = 5\text{x}10^{-5}$)

If this process is repeated a set of failures is produced. Thus data from different distributions may be generated allowing one to test hypotheses about the distributions knowing what the true results should be.

2.7 Power

Suppose that data have been simulated from a normal distribution with a mean $\mu_1 = 0$ and a variance of $\sigma^2 = 1$, i.e., N(0,1). Data are then simulated from another normal distribution with a mean $\mu_2 = 1$ and the same variance, i.e., N(1,1). A null hypothesis is made that all the data came from the same distribution.

$$H_0: \ \mu_1 = \mu_2,$$

and the alternative hypothesis is made that the data come from different distributions

$$H_1: \ \mu_1 \neq \mu_2.$$

In order to test these hypotheses, the likelihood ratio test is used. After all the estimators have been found and the likelihood ratio evaluated, the test statistic is evaluated and a determination is made either to accept or reject the null hypothesis at a certain significance level. If data are simulated and the process is repeated many times for the same distributions, a count can be made of the number of times the null hypothesis is rejected and the number of sets of data. The simulated data are used to find failure rate distribution parameters; often no estimators are found (for reasons to be discussed below). Thus the number of sets of data to be used is that number of data sets which yields estimators. Power can then be defined as the ratio of the number of times H_0 is rejected to the number of data sets which yield estimators.

If in the example, the mean of the second distribution was the same as the mean of the first, $\mu_1 = \mu_2 = 0$, the power would just be equal to the probability of a type I error, α, i.e., the probability of rejecting H_0 when it is in fact true. Ideally when the two parameters being tested are identical the power should be equal to zero. Similarly when the two parameters are not equal (e.g. μ_1, = 0, μ_2, = 0.1) the power should ideally equal unity.

Power may be generated for any combination of parameter values from the first and second distributions. If the parameter from the first distribution is held constant and the value of the parameter from the second function is allowed to vary, a power function is produced. The power function may be illustrated by plotting power versus the value of the second parameter or by plotting power versus the difference between the value of the parameter of the second function and the value of the parameter of the first as shown in Figs. 2.10 and 2.11 for $\alpha = 0.05$. Note when $\mu_1 = \mu_2$ the value of power equals 0.05. As μ_2 gets further away from μ_1

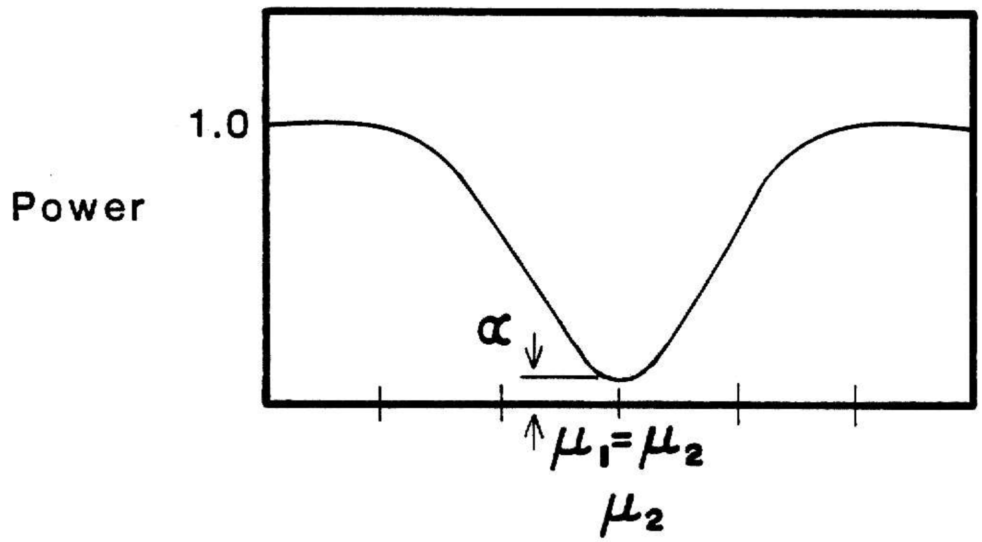

Figure 2.10 Power versus value of second parameter μ_2.

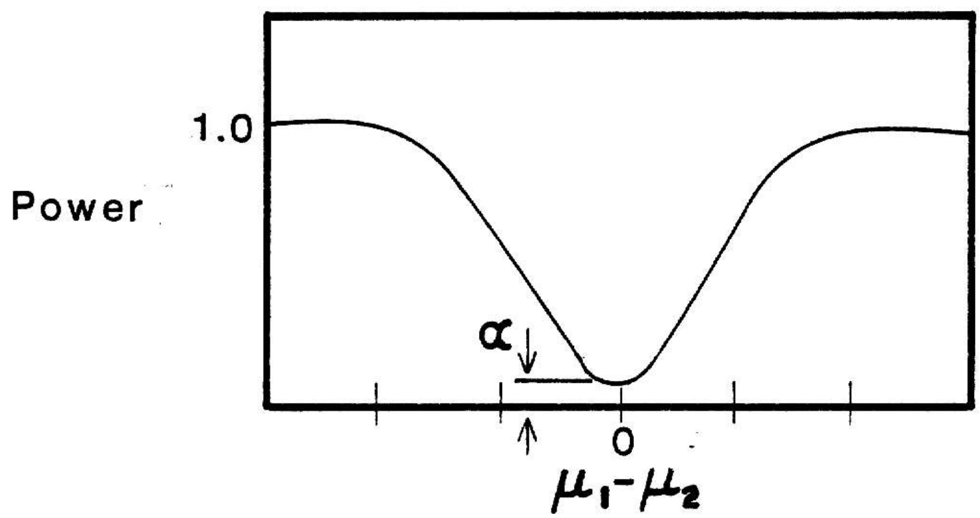

Figure 2.11 Power versus value of difference between standard parameter and second parameter.

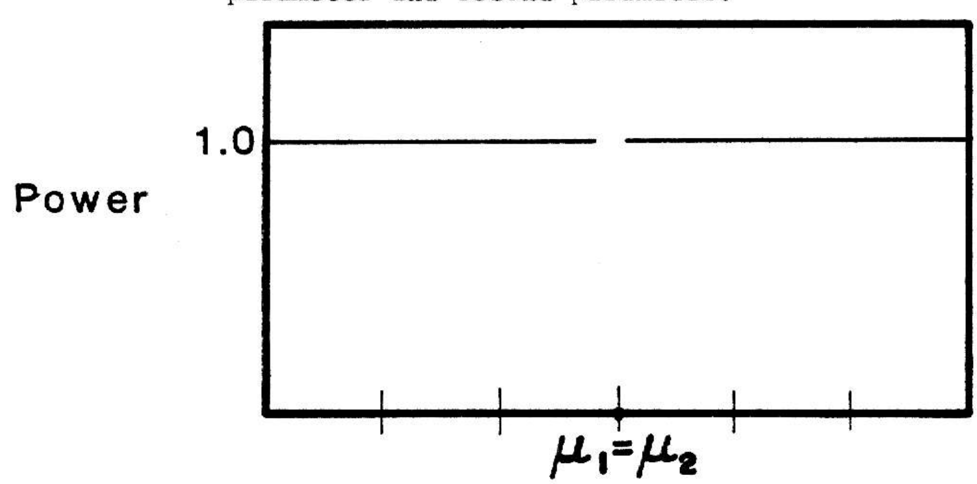

Figure 2.12 Ideal Power Curve.

the power goes to unity. An ideal power curve is shown in Fig. 2.12. In this case the hypothesis test always yields the correct conclusion.

3 CODE OPERATION

In this section a general overview is given of the operation of the computer codes used in this work. No attempt is made to detail code operation line by line. Instead this section is designed to outline the major steps in the programs and describe how to use the codes. A listing of the codes may be found in the Appendix.

3.1 Generate

3.1.1 Description

The purpose of the code GENERATE is to produce sets of simulated failure data. Given two sets of parameters for gamma prior distributions the program GENERATE will simulate failure data for each distribution and write the simulated data onto magnetic tape. The program may be thought of as being in two major sections: 1) determination of the cumulative distribution and 2) simulation of data.

The marginal distribution [Eq. (2.5)] shown earlier as

$$h(F|T,\alpha,\beta)) = \frac{\Gamma(F+\alpha)}{\Gamma(\alpha)\Gamma(F+1)} \frac{T^F \tau^\alpha}{(T+\tau)^{F+\alpha}}, \tag{3.1}$$

is used to form the cumulative distribution. Initially given values for time (T), alpha (α), and beta (β) are read in from cards. The number of failures (F) is set equal to zero and from Eq. (3.1), the probability of having zero failures in T hours of operation with parameters α and β is calculated. This probability of zero failures is the first point in the discrete cumulative distribution. F is incremented by one and a new probability is found (i.e., the probability of having 1 failure in T hours of operation with parameters α and β). This probability of one failure is added to the probability of zero failure to yield the second point in the cumulative distribution. This process of incrementing F, calculating a probability, and summing is continued until the cumulative probability is very close to unity (0.99999 for this work). The cumulative distribution is printed out, and the process is repeated for the second set of parameters.

The second section of the program simulates failure data. A random number between zero and unity is generated using the subroutine RANDU. A determination is made of the value of F in the cumulative distribution corresponding to the random number. Thus a simulated number of failures is produced which is paired with the corresponding operation time of the distribution being used. Two groups of failure data pairs are produced, one for each of the distributions being tested. After the two groups of pairs have been produced they are printed out and also written onto magnetic tape.

3.1.2 Input

In order to operate GENERATE four input cards are required.

<u>Card 1</u>

First distribution parameters; TIME (in hrs.), ALPHA, BETA.

FORMAT (E13.5,7X,E13.5,7X,E13.5)

<u>Card 2</u>

Second distribution parameters; TIME2 (in hrs.), ALPHA2, BETA2

FORMAT (E13.5,7X,E13.5,7X,E13.5)

<u>Card 3</u>

Number of failure data pairs per set

FORMAT (I6)

Card 4

Number of data sets required

FORMAT (I6)

3.1.3 Sample Output

Input parameters

First distribution

TIME = 10,000 (hrs.)

ALPHA = 1.5

BETA = 0.00005

Second distribution

TIME = 10,000 (hrs.)

ALPHA = 1.5

BETA = 0.00001

Number of failure data pairs per set = 10

Number of data sets = 250

3.2 GAMMA MODIFIED

3.2.1 Description

The program GAMMA MODIFIED is a modification of the program GAMMA8 written by J.K. Shultis in June, 1980. The original GAMMA8 program worked with failure data to give estimates of the gamma prior distribution parameters for both the homogeneous and compound model plus other properties of the data. The GAMMA MODIFIED program retains the parameter estimation for the compound model and has sections added to find likelihood ratios and power.

GAMMA MODIFIED reads the sets of failure rate data put on tape by GENERATE. The data are on tape in the following order: a group from the first distribution, a group from the second distribution, etc. In order

Table 3.1 Sample output from GENERATE

```
TIME=  0.10000D 05HRS.      ALPHA=  0.15000D 01    BETA=  0.50000D-04

CUMULATIVE DISTRIBUTION
F=        CUMULATIVE PROBABILITY
  0         0.54433D 00
  1         0.81650D 00
  2         0.92990D 00
  3         0.97400D 00
  4         0.99054D 00
  5         0.99660D 00
  6         0.99879D 00
  7         0.99957D 00
  8         0.99985D 00
  9         0.99995D 00
 10         0.99998D 00
 11         0.99999D 00

TIME2=  0.10000D 05HRS.      ALPHA2=  0.15000D 01    BETA2=  0.10000D-04

CUMULATIVE DISTRIBUTION
F=        CUMULATIVE PROBABILITY
  0         0.86678D 00
  1         0.98498D 00
  2         0.99841D 00
  3         0.99984D 00
  4         0.99998D 00
  5         0.10000D 01

SET NUMBER  1

FIRST DISTRIBUTION         NF     NT
                            0      10000
                            0      10000
                            0      10000
                            1      10000
                            1      10000
                            0      10000
                            0      10000
                            0      10000
                            0      10000
                            4      10000

SECOND DISTRIBUTION
                            0      10000
                            0      10000
                            0      10000
                            0      10000
                            1      10000
                            0      10000
                            0      10000
                            0      10000
                            0      10000
                            0      10000

SET NUMBER  2

FIRST DISTRIBUTION         NF     NT
                            3      10000
                            1      10000
```

0	10000
2	10000
0	10000
1	10000
1	10000
0	10000
1	10000
0	10000

SECOND DISTRIBUTION

0	10000
1	10000
2	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000

SET NUMBER 3

FIRST DISTRIBUTION

NF	NT
0	10000
0	10000
1	10000
0	10000
1	10000
0	10000
0	10000
0	10000
0	10000
3	10000

SECOND DISTRIBUTION

0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000

SET NUMBER 4

FIRST DISTRIBUTION

NF	NT
0	10000
0	10000
1	10000
0	10000
0	10000
1	10000
0	10000
3	10000
1	10000
0	10000

SECOND DISTRIBUTION

0	10000
0	10000
0	10000
0	10000
1	10000
0	10000
0	10000
0	10000
0	10000
1	10000

SET NUMBER 5

FIRST DISTRIBUTION

NF	NT
0	10000
0	10000
3	10000
0	10000
0	10000
1	10000
0	10000
5	10000
1	10000
1	10000

SECOND DISTRIBUTION

0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
1	10000
1	10000

SET NUMBER 6

FIRST DISTRIBUTION

NF	NT
0	10000
1	10000
0	10000
4	10000
0	10000
1	10000
2	10000
3	10000
1	10000
1	10000

SECOND DISTRIBUTION

0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000

NF	NT
0	10000
0	10000
0	10000

SET NUMBER 7

FIRST DISTRIBUTION

NF	NT
0	10000
0	10000
0	10000
0	10000
2	10000
0	10000
1	10000
0	10000
0	10000
3	10000

SECOND DISTRIBUTION

0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
1	10000
0	10000
0	10000
0	10000

SET NUMBER 8

FIRST DISTRIBUTION

NF	NT
0	10000
2	10000
1	10000
0	10000
1	10000
1	10000
2	10000
2	10000
1	10000
3	10000

SECOND DISTRIBUTION

1	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000
0	10000

SET NUMBER 9

FIRST DISTRIBUTION

NF	NT
1	10000

0	10000
0	10000
0	10000
0	10000
1	10000
0	10000
0	10000
4	10000
1	10000

SECOND DISTRIBUTION

0	10000
1	10000
0	10000
1	10000
0	10000
0	10000
1	10000
0	10000
0	10000
1	10000

SET NUMBER 10

FIRST DISTRIBUTION

NF	NT
0	10000
0	10000
0	10000
0	10000
0	10000
1	10000
0	10000
0	10000
0	10000
5	10000

SECOND DISTRIBUTION

0	10000
0	10000
0	10000
0	10000
0	10000
1	10000
0	10000
0	10000
0	10000
0	10000

to find maximum likelihood ratio between two groups of data it was shown in Section 2.5 that estimators must be found for each group taken separately and then for the combination of the two groups. To determine the likelihood ratio between data from the first and second distributions, GAMMA MODIFIED finds estimators for the first group of data, the second group, and for the combination of the first and second group. Next estimators are found for the third group, the fourth group, and finally the combination of the third and fourth groups. This process is repeated until estimators have been found for each pair of data groups (one from the first distribution and one group from the second distribution). Estimators are calculated in each case using all three techniques, matching data moments to the prior distribution, matching data moments to the marginal distribution, nad maximizing the likelihood function.

After the three sets of estimators are found for each of the two groups of data, the likelihood ratio is evaluated using each set of estimators. Since there is now a set of likelihood ratios for each estimation technique three values of power are calculated. The test statistic $t = -2 \cdot \ln(L)$ (where L is the likelihood ratio) is calculated for each likelihood ratio and tested against the vlaue of χ^2 (2,0.05). If $t > \chi^2$ (2,0.05) the null hypothesis that all the data came from the same distribution is rejected. Thus for each estimation technique the value of power is then found as the number of times t is found to be significant (i.e., H_o is rejected) divided by the number of likelihood ratios found.

Summarizing:

1) Failure rate data are read off tape.
2) Prior parameter estimators are found using the three estimation techniques.

3) Likelihood ratios are evaluated for each pair of data sets and for each estimation technique.

4) Three values of power are found, one for each estimation technique.

3.2.2 Input

On tape

Failure rate data prior; number of failures, time (in hours).

FORMAT (I8,I8)

On cards

Card 1

Number of points, number of sets.

FORMAT (I8,I8)

3.2.3 Sample Output

Input parameters

Number of points = 10

Number of sets = 10

3.3 GAMMA GENERATE

3.3.1 Description

When failure data are only needed for the determination of estimators and power, it is not necessary to record permanently the large sets of failure data required. It would be convenient to have one code which simulates failure data, produces estimators, and calculates power. This is the purpose of the code GAMMA GENERATE, to eleviate the need to write failure data onto tape and then read it off again. The code is simply a mating of the codes GENERATE and GAMMA MODIFIED, previously described.

The code operates as follows. Initial values are read for two distributions along with the number of failure data pairs per set and the number of data sets required, as in the program GENERATE. The parameters of the first distribution are held constant throughout the entire program

Table 3.2 Sample output from GAMMA MODIFIED

```
PCWER=  0.12500D+00      NUMBER OF PTS. USED=     8   ESTIMATICN METHCD NO.  1
PCWER=  0.C              NUMBER OF PTS. USED=     1   ESTIMATION METHCD NO.  2
PCWER=  0.0              NUMBER OF PTS. USED=     0   ESTIMATICN METHCD NC.  3
```

while the beta value of the second distribution (BETA2) may be allowed to vary. Cumulative distributions are formed for each set of parameters and from them sets of failure data are generated. The three parameter estimation techniques, matching moments to the prior, matching moments to the marginal, and maximum likelihood, are used to make parameter estimates. Power is then calculated for each estimation technique and the result is written out.

To this point operation has been exactly the same as in GENERATE and GAMMA MODIFIED. Instead of terminating the program at this point, the value of BETA2 is changed by a preset increment, a new cumulative distribution for the second set of parameters is calculated using the new value for BETA2, and the procedure for finding power is repeated. Incrementing BETA2 allows a set of powers to be generated. The size of the increment and the final value for BETA2 may be set within the program.

Output from the program consists of listing distribution parameters, cumulative distributions, the power for each combination of parameters using each estimation technique, and the number of points used in the calculation of power.

3.3.2 Input

Cards

Card 1

First distribution parameters: TIME (in hours), ALPHA, BETA.

FORMAT (E13.5,7X,E13.5,7X,E13.5)

Card 2

Second distribution parameters: TIME2 (in hours), ALPHA2, BETA2

FORMAT (E13.5,7X,E13,5,7X,E13.5)

Card 3

Number of failure data pairs per set.

FORMAT (I6)

Card 4

Number of data sets

FORMAT (13)

3.3.3 Sample Output

Input parameters

First distribution

TIME = 10,000 (hrs.)

ALPHA = 2.0

BETA = 0.00005

Number of failure data pairs per set = 10

Number of data sets = 250

Second distribution

TIME2 = 10,000 (hrs.)

ALPHA2 = 2.0

BETA2 = 0.00004

Number of failure data pairs per set = 10

Number of data sets = 250

3.4 GAMMAP GENERATE

3.4.1 Description

The code GAMMAP GENERATE is nearly the same as the code GAMMA GENERATE, the only difference being that matching moments to the prior is the only parameter estimation technique used, i.e., the section to estimate prior parameters using the techniques for matching moments to the marginal and maximum likelihood have been eliminated.

Table 3.3 Sample output from GAMMA GENERATE

```
TIME=  0.10000D+05HRS.      ALPHA=  C.20CC0D+C1   BETA=  C.5CCCCD-C4

CUMULATIVE DISTRIBUTION
F=        CUMULATIVE PROBABILITY
  0          0.44444D+00
  1          C.74C74D+C0
  2          0.88889D+00
  3          C.95473D+00
  4          C.98217D+00
  5          0.99314D+00
  6          0.99741D+00
  7          0.99903D+00
  8          C.99964D+00
  9          C.99987D+00
 10          0.99995D+00
 11          0.99998D+00
 12          C.99999D+00

TIME2=  0.10000D+05HRS.      ALPHA2=  0.20C0CC+C1   BETA2=  C.40C00D-C4

CUMULATIVE DISTRIBUTION
F=        CUMULATIVE PROBABILITY
  0          0.51020D+00
  1          0.80175D+00
  2          0.92670D+00
  3          C.97430D+00
  4          0.99130D+00
  5          0.99712D+00
  6          C.99907D+CC
  7          0.99970D+00
  8          C.99991D+00
  9          0.99997D+00
 10          C.99999D+00

POWER=  0.36CCCD-01      NUMBER OF PTS. USED=   250    ESTIMATION METHOD NO.  1

POWER=  0.86957D-02      NUMBER OF PTS. USED=   115    ESTIMATION METHOD NO.  2

POWER=  0.24390D-01      NUMBER OF PTS. USED=    82    ESTIMATION METHOD NO.  3
```

3.4.2 Input

Cards

Card 1

First distribution parameters: TIME (in hours), ALPHA, BETA.

FORMAT (E13.5,7X,E13,5,7X,E13.5)

Card 2

Second distribution parameters: TIME2 (in hours), ALPHA, BETA2.

Card 3

Number of failure data pairs per set.

FORMAT (I3)

3.4.3 Sample Output

Input parameters

First distribution

TIME = 10,000 (hrs.)

ALPHA = 2.0

BETA = 0.00005

Number of failure data pairs per set = 10

Number of data sets = 250

Second distribution

TIME2 = 10,000 (hrs.)

ALPHA2 = 2.0

BETA2 = 0.00004

Number of failure data pairs per set = 10

Number of data sets = 250

Table 3.4 Sample output from GAMMAP GENERATE.

TIME= 0.10000D+05HRS. ALPHA= 0.20000D+01 BETA= 0.50000D-04

CUMULATIVE DISTRIBUTION

F=	CUMULATIVE PROBABILITY
0	0.44444D+00
1	0.74074D+00
2	0.88889D+00
3	0.95473D+00
4	0.98217D+00
5	0.99314D+00
6	0.99741D+00
7	0.99903D+00
8	0.99964D+00
9	0.99987D+00
10	0.99995D+00
11	0.99998D+00
12	0.99999D+00

TIME2= 0.10000D+05HRS. ALPHA2= 0.20000D+01 BETA2= 0.32000D-03

CUMMULATIVE DISTRIBUTION

F=	CUMULATIVE PROBABILITY
0	0.56689D-01
1	0.14307D+00
2	0.24180D+00
3	0.34209D+00
4	0.43760D+00
5	0.52493D+00
6	0.60256D+00
7	0.67015D+00
8	0.72809D+00
9	0.77713D+00
10	0.81824D+00
11	0.85240D+00
12	0.88060D+00
13	0.90374D+00
14	0.92263D+00
15	0.93798D+00
16	0.95041D+00
17	0.96043D+00
18	0.96850D+00
19	0.97496D+00
20	0.98014D+00
21	0.98426D+00
22	0.98755D+00
23	0.99017D+00
24	0.99224D+00
25	0.99389D+00
26	0.99519D+00
27	0.99622D+00
28	0.99703D+00
29	0.99767D+00
30	0.99817D+00
31	0.99857D+00
32	0.99888D+00
33	0.99912D+00
34	0.99931D+00
35	0.99946D+00
36	0.99958D+00

37	0.99967D+00
38	0.99975D+00
39	0.99980D+00
40	0.99985D+00
41	0.99988D+00
42	0.99991D+00
43	0.99993D+00
44	0.99994D+00
45	0.99996D+00
46	0.99997D+00
47	0.99997D+00
48	0.99998D+00
49	0.99998D+00
50	0.99999D+00
51	0.99999D+00

POWER= 0.91600D+00 NUMBER OF PTS. USED= 250 ESTIMATION METHOD NO. 1

4 Data Simulation and Analysis

4.1 Generation of Power Curves

4.1.1 General Procedure

Below is a general procedure for obtaining a power curve for any general distribution with variable parameters. [Note that the value of the variable parameter chosen for testing must remain constant (in the standard distribution) throughout the production of the power curve.]

1) Choose which parameter will be the variable parameter.
2) Set all other parameters to constants.
3) Choose a value for which the variable parameter is to be tested against (a standard value).
4) Choose an initial starting value for the variable parameter.
5) Simulate data from the two separate distributions.
6) Find estimators of the parameters for each group of data separately and for the data from both distributions grouped together.
7) Find likelihood values for all data sets for which estimators were formed by all three estimation methods.
8) Calculate likelihood ratio calculate test statistic, compare to $\chi^2(2,0.05)$, and make decision to accept or reject H_o.
9) Repeat steps 5-8
10) Calculate power, i.e., number of times H_o was correctly rejected divided by the number of data sets for which estimators were calcuable.
11) Increment the value of the variable parameter in the second distribution and repeat steps 5-10.

The number of times steps 5-8 are repeated is set in the computer programs by the number of sets of data to be generated. The amount that the variable parameter is to be incremented (in this case BETA2) is also set within the programs.

4.1.2 Parameters set in Programs

In this work the gamma distribution was used. Variable parameters of the gamma distribution are ALPHA, BETA, and test time, TIME. BETA was chosen as the variable parameter to be tested.

The value of BETA in the standard distribution was set at 0.00005. Thus when data are simulated from the standard distribution and from some distribution where the value of beta (called BETA2 in the alternate distribution) has been varied, the null hypothesis is

H_o: All the data came from a distribution with a value of BETA of 0.00005.

The alternate hypothesis is

H_1: The data came from different distributions.

Power is then the ratio of the number of times H_o is correctly rejected to the total number of data sets for which estimators were found, i.e., the number of times data are found to come from different distributions divided by the number of times estimators are found.*

For example, say the MMMM (marginal matching moments method) was able to find estimators for only 151 sets of data out of 250 generated sets of data. With these 151 estimators it was found that the null hypothesis (all data came from a distribution with a beta of 0.00005) was rejected 31 times. Thus power equals 31/151 or 0.2053, not 31/250 or 0.124.

Throughout this work test time was held constant at 10,000 hours (which is approximately one year), although it is possible to vary time

---

*It is important to note that power has been defined as the number of times H_o is rejected over the number of times estimators were found. Power is not defined as the number of times H_o is correctly rejected over the number of sets of data. If estimators cannot be found for a particular set of data the set is not used in the calculation of power.

in the computer codes. Power curves were generated for value of ALPHA of 1.0, 1.5, 2.0, 2.5. ALPHA is the shape parameter of the gamma distribution so power curves were generated for each shape that the distribution may assume (see Sec. 2.2). The number of points per data set used throughout was ten. The test to accept or reject was made at the 0.05 significance level.

4.2 Results

The codes GAMMA GENERATE and GAMMAP MODIFIED were used to produce the following tables. Tables 4.1, 4.2, and 4.3 use MMPM (matching moments to the prior method), MMMM (matching moments to the marginal method), and MMLM (marginal maximum likelihood method), respectively, as the parameter estimation techniques. All have an ALPHA value of 1.5. Tables 4.4, 4.5, and 4.6 have an ALPHA of 2.0 and use the MMPM, MMMM, and MMLM, respectively. In Tables 4.7, 4.8, 4.9, and 4.10 the MMPM and the values for ALPHA of 1.0, 1.5, 2.0, and 2.5 were used. Each table is for one value of ALPHA and one estimation technique. Each table lists the value of BETA2, the power, and the number of points used in the calculation of power. The distribution with the given value of BETA2 is always compared to a distribution whose BETA value is 0.00005.

These data are used to plot power curves. Following each table is a corresponding plot of power versus BETA2.

Table 4.1 Power for alpha=1.5, using MMPM

BETA2 ($\times 10^{-5}$)	POWER	NO. OF POINTS USED
1.0	0.09290	183
2.0	0.03766	239
3.0	0.02449	245
4.0	0.01606	249
5.0	0.00803	249
6.0	0.01205	249
7.0	0.02000	250
8.0	0.06855	248
9.0	0.06800	250
10.0	0.09639	249
12.0	0.11200	250
14.0	0.27016	248
16.0	0.37200	250
18.0	0.44980	249
20.0	0.51200	250
22.0	0.61847	249
24.0	0.64516	248
26.0	0.73092	249
28.0	0.76800	250
30.0	0.82400	250
32.0	0.84800	250
34.0	0.88353	249

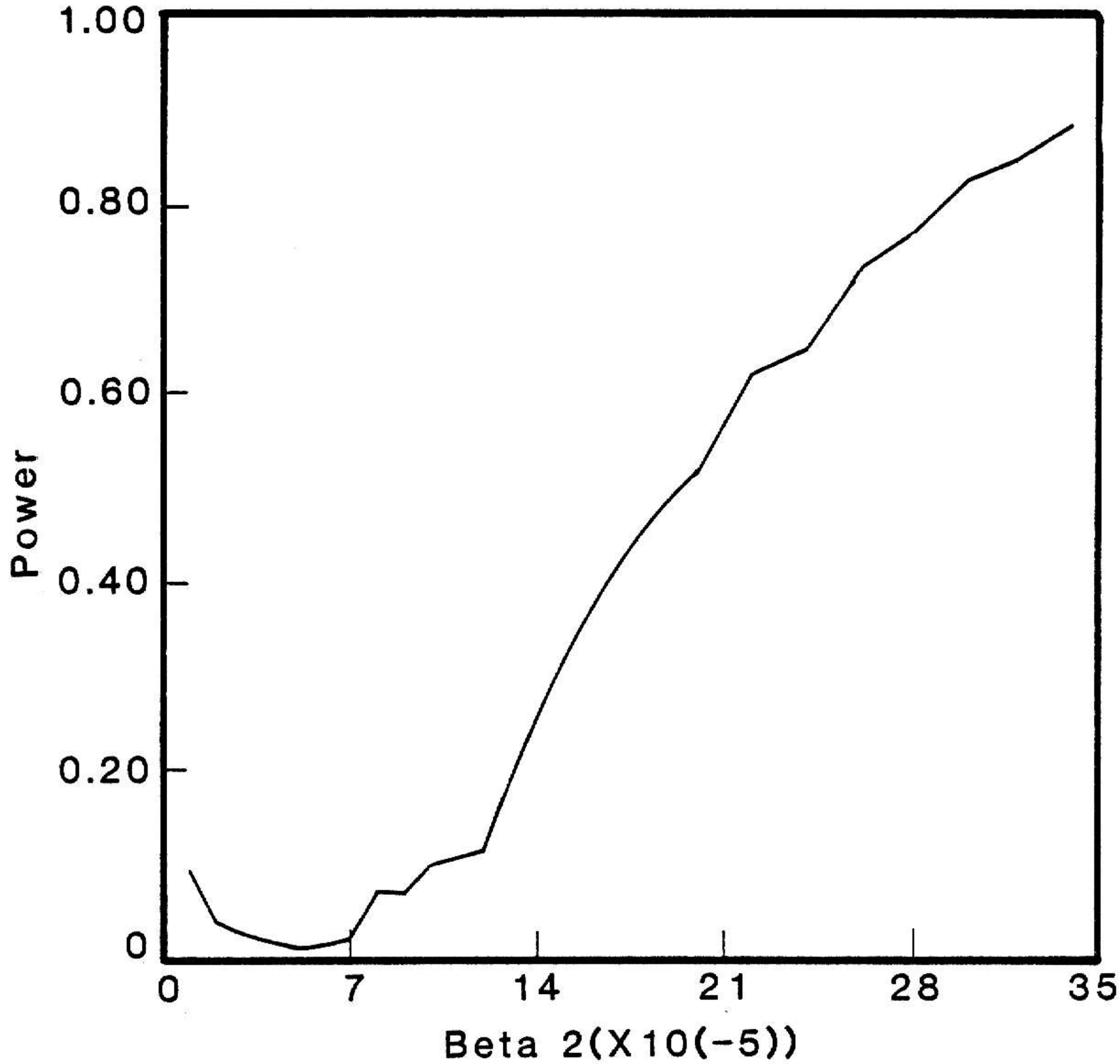

Figure 4.1 Power Curve for α = 1.5 using MMPM.

Table 4.2 Power for alpha=1.5, using MMMM

BETA2 ($x10^{-5}$)	POWER	NO. OF POINTS USED
1.0	0.03846	26
2.0	0.01613	62
3.0	0.00000	83
4.0	0.00000	111
5.0	0.00746	134
6.0	0.00000	104
7.0	0.01398	143
8.0	0.02290	131
9.0	0.06569	137
10.0	0.07857	140
12.0	0.10759	158
14.0	0.21192	151
16.0	0.31250	160
18.0	0.36306	157
20.0	0.51136	176
22.0	0.60403	149
24.0	0.55747	174
26.0	0.68072	166
28.0	0.71166	163
30.0	0.75301	166
32.0	0.73885	157
34.0	0.83133	166

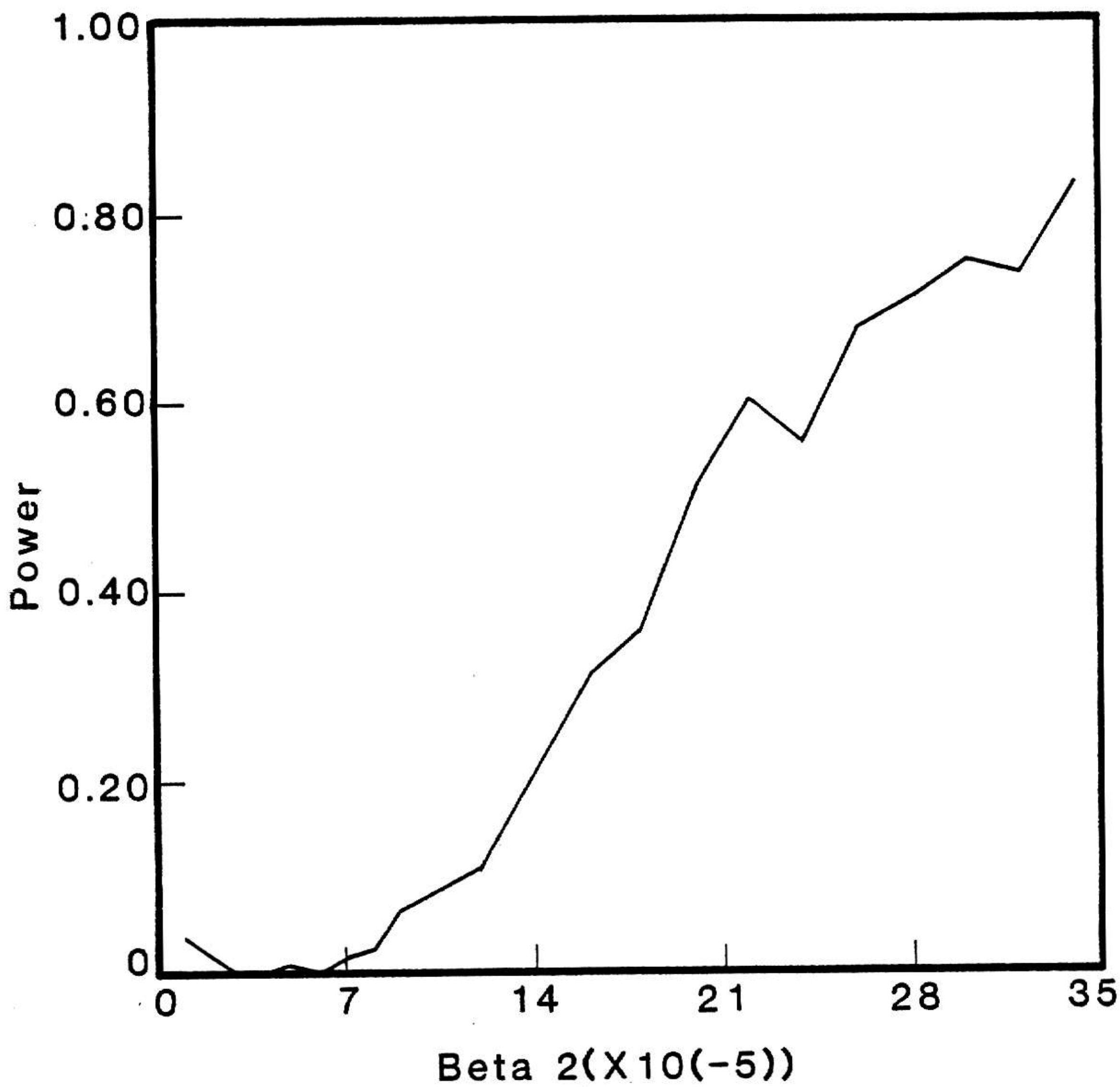

Figure 4.2 Power Curve for α = 1.5 using MMMM.

Table 4.3 Power for alpha=1.5, using MMLM

BETA2 ($x10^{-5}$)	POWER	NO. OF POINTS USED
1.0	0.11111	18
2.0	0.04878	41
3.0	0.0	60
4.0	0.06154	65
5.0	0.07292	96
6.0	0.05128	78
7.0	0.05882	102
8.0	0.07609	92
9.0	0.13333	90
10.0	0.25000	104
12.0	0.39655	116
14.0	0.54310	116
16.0	0.69167	120
18.0	0.73171	123
20.0	0.83007	153
22.0	0.93443	122
24.0	0.89583	144
26.0	0.92754	138
28.0	0.94853	136
30.0	0.97101	138
32.0	0.96825	126
34.0	0.99231	130

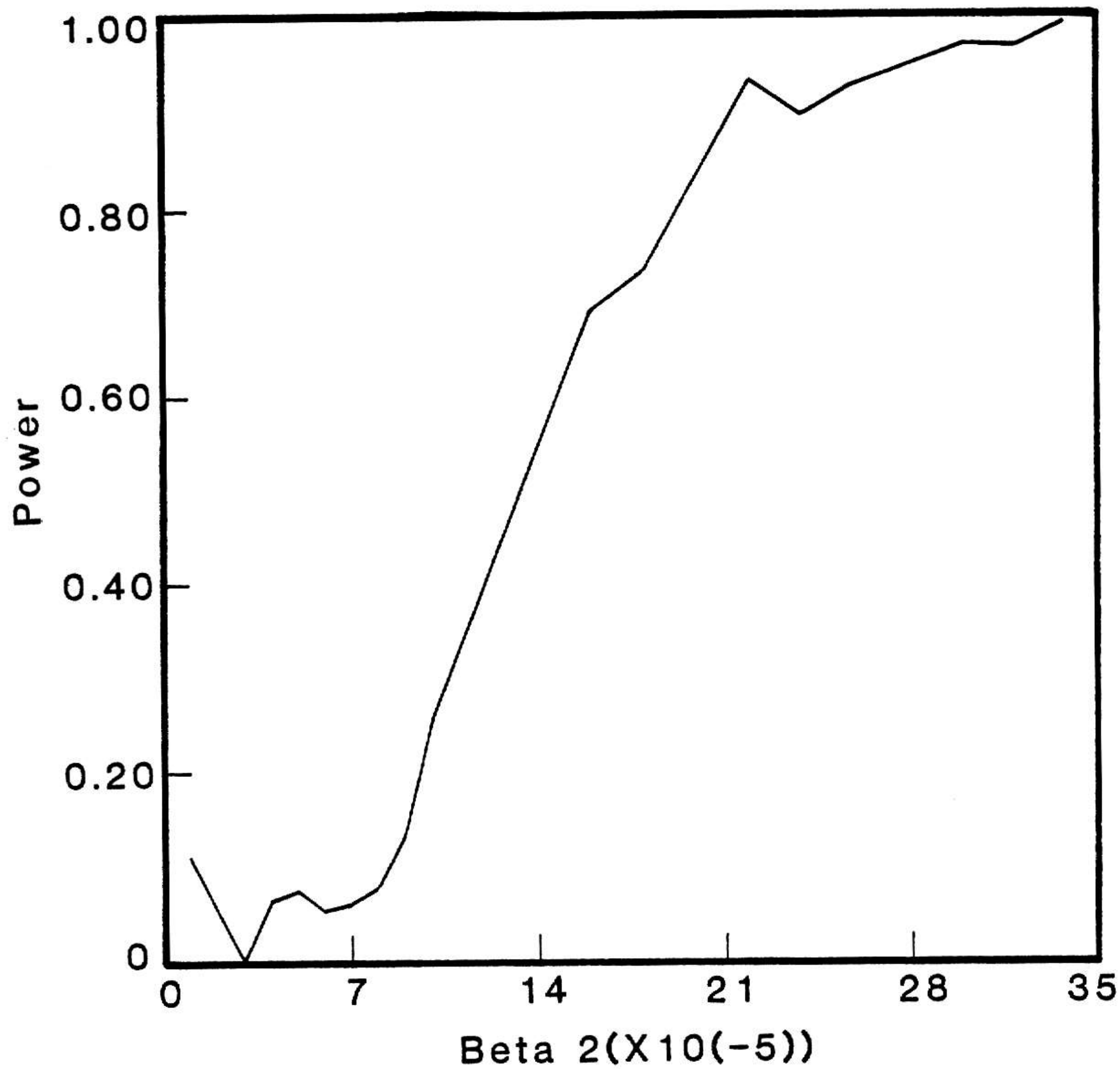

Figure 4.3 Power Curve for α = 1.5 using MMLM.

Table 4.4 Power for alpha=2.0, using MMPM

BETA2 $(x10^{-5})$	POWER	NO. OF POINTS USED
1.0	0.21154	208
2.0	0.07347	245
3.0	0.03239	247
4.0	0.01600	250
5.0	0.01200	250
6.0	0.02000	250
7.0	0.04418	249
8.0	0.06400	250
9.0	0.11200	250
10.0	0.09600	250
11.0	0.18870	249
12.0	0.25600	250
13.0	0.28800	250
14.0	0.33200	250
15.0	0.46400	250
16.0	0.48996	249
17.0	0.49200	250
18.0	0.58400	250
19.0	0.63600	250
20.0	0.63200	250
21.0	0.76000	250
22.0	0.74800	250
23.0	0.74400	250
24.0	0.82731	249
25.0	0.86400	250
26.0	0.87200	250
27.0	0.88000	250
28.0	0.94400	250
29.0	0.88755	249
30.0	0.92000	250
31.0	0.93200	250
32.0	0.94800	250
33.0	0.94800	250
34.0	0.97200	250
35.0	0.97200	250
36.0	0.99200	250
37.0	0.97600	250
38.0	0.99200	250
39.0	0.97600	250
40.0	0.96800	250

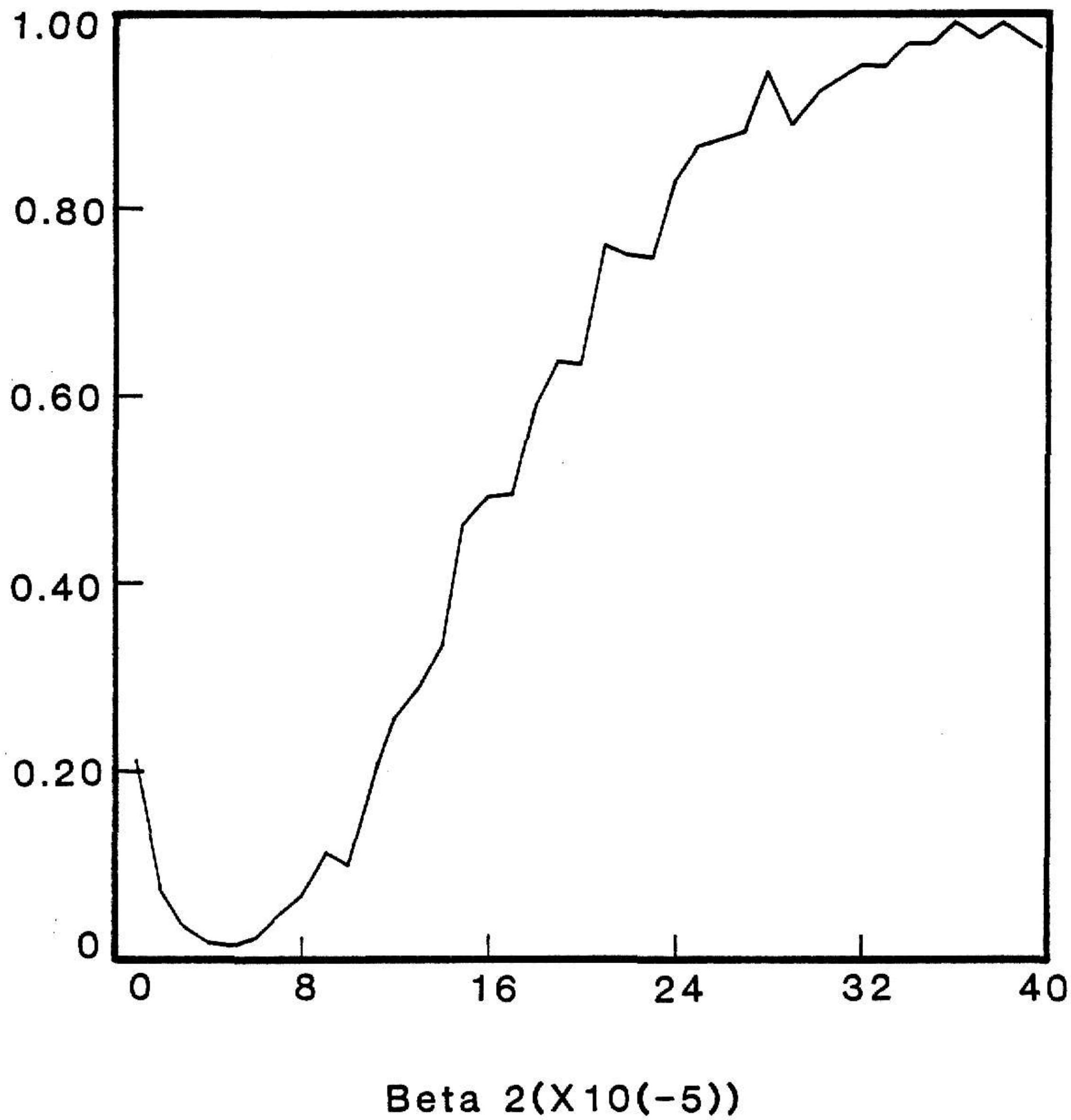

Figure 4.4 Power Curve for α = 2.0 for MMPM.

Table 4.5 Power for alpha=2.0, using MMMM

BETA2 $(x10^{-5})$	POWER	NO. OF POINTS USED
1.0	0.08108	37
2.0	0.0	80
3.0	0.0	100
4.0	0.0	119
5.0	0.00735	136
6.0	0.01818	110
7.0	0.02721	147
8.0	0.05436	129
9.0	0.06164	146
10.0	0.09589	146
11.0	0.20530	151
12.0	0.25806	155
13.0	0.29518	166
14.0	0.34783	184
15.0	0.43382	136
16.0	0.47674	172
17.0	0.48447	161
18.0	0.57558	172
19.0	0.60606	165
20.0	0.58491	159
21.0	0.73810	207
22.0	0.72000	150

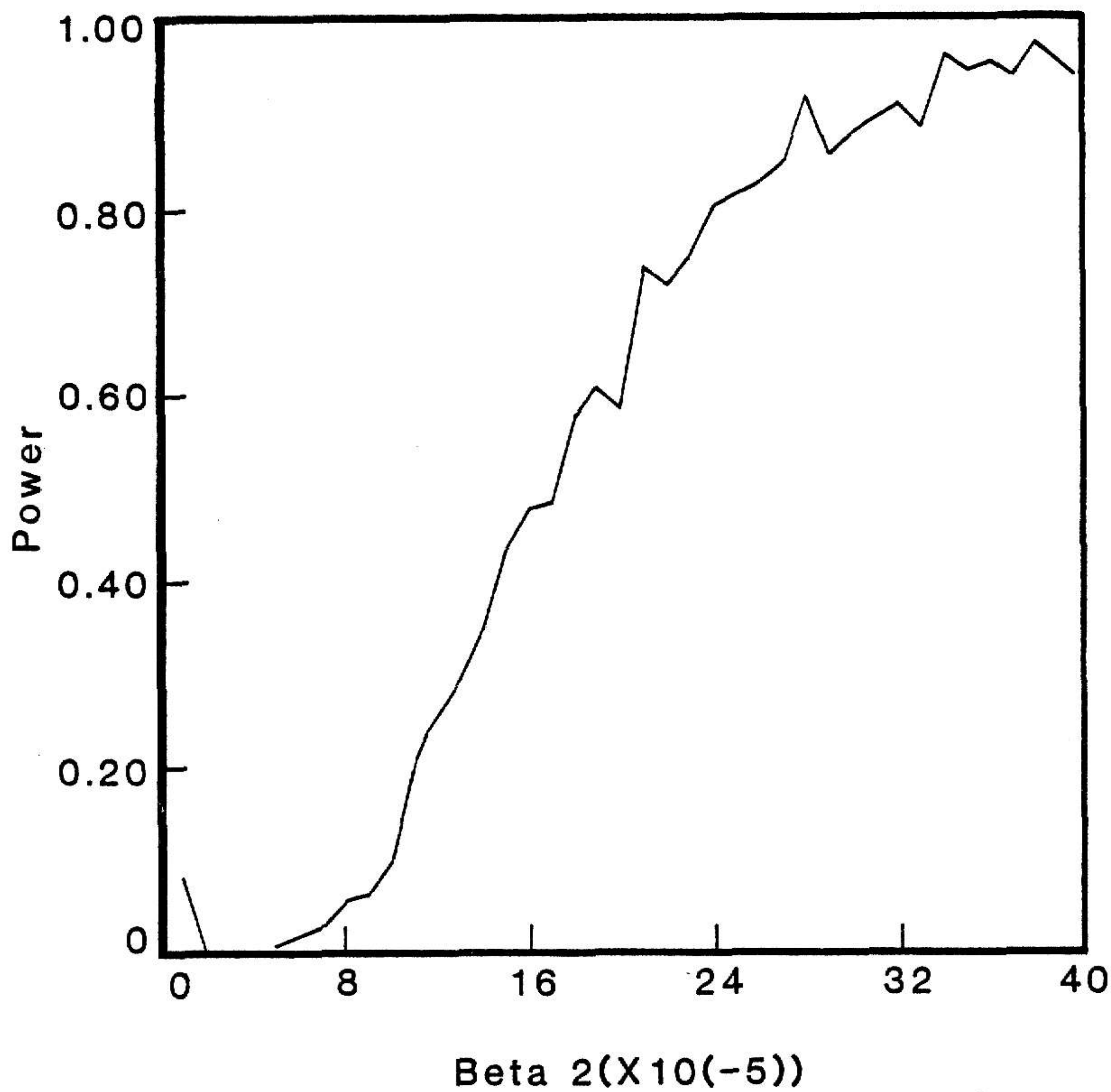

Figure 4.5 Power Curve for α = 2.0 using MMMM.

Table 4.6 Power for alpha=2.0, using MMLM

BETA2 (10^{-5})	POWER	NO. OF POINTS USED
1.0	0.17857	28
2.0	0.08929	56
3.0	0.01429	70
4.0	0.06060	66
5.0	0.01099	91
6.0	0.04168	72
7.0	0.08929	112
8.0	0.13592	103
9.0	0.22414	116
10.0	0.25439	114
11.0	0.42400	125
12.0	0.50388	129
13.0	0.54098	122
14.0	0.69863	146
15.0	0.74312	109
16.0	0.70803	137
17.0	0.80451	133
18.0	0.86330	139
19.0	0.86822	129
20.0	0.87200	125
21.0	0.91852	135
22.0	0.91870	123
23.0	0.93431	137
24.0	0.96552	145
25.0	0.95444	150
26.0	0.96296	135
27.0	0.96875	160
28.0	1.00000	123
29.0	0.99320	147
30.0	1.00000	141
31.0	0.99315	146
32.0	0.99291	141
33.0	0.98496	133
34.0	0.99291	141
35.0	0.98450	129
37.0	0.99248	133
38.0	0.98630	146
39.0	1.00000	138
40.0	1.00000	137

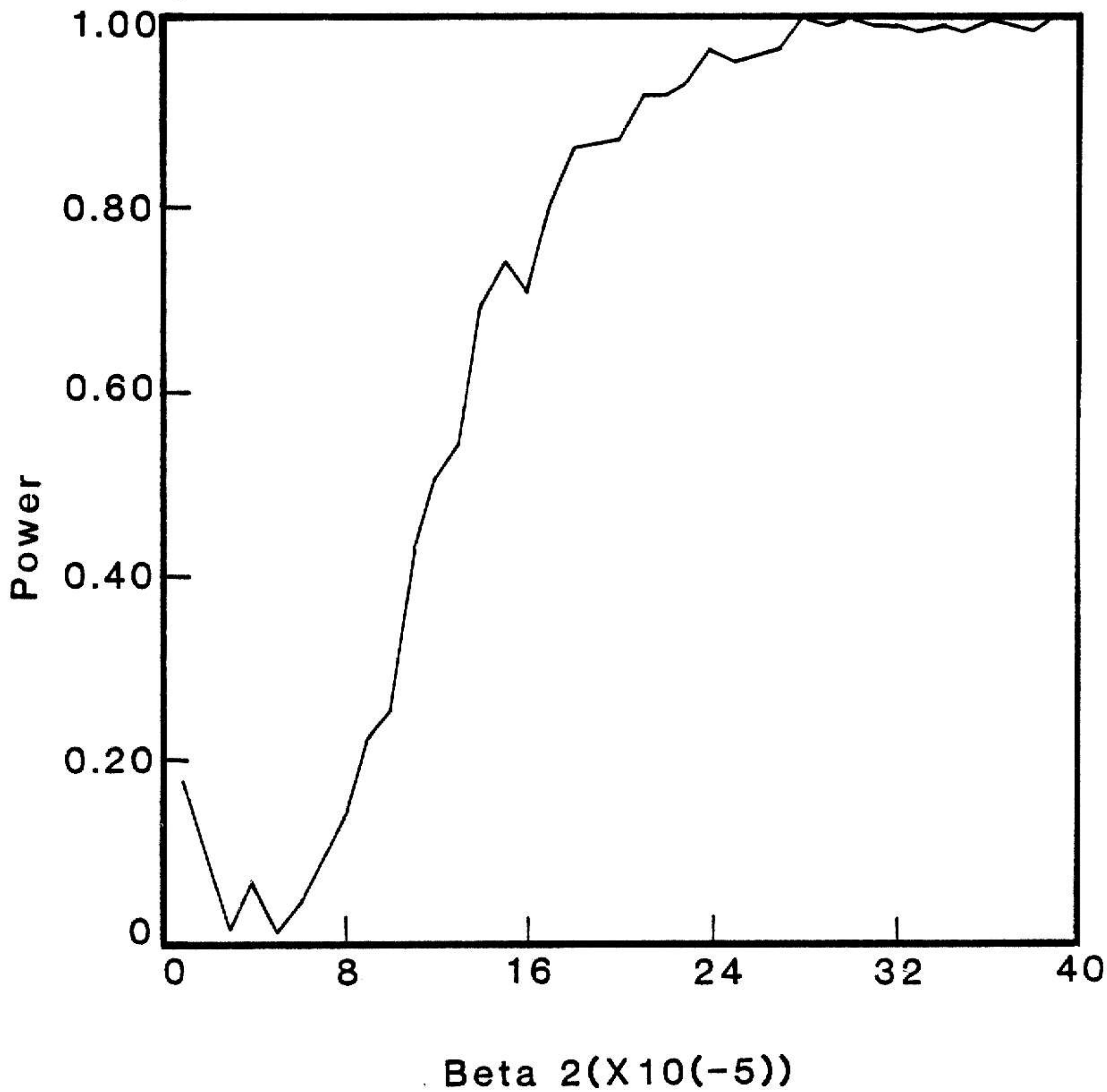

Figure 4.6 Power Curve for α = 2.0 using MMLM.

Table 4.7 Power for alpha=1.0, using MMPM

BETA2 ($x10^{-5}$)	POWER	NO. OF POINTS USED
0.1	0.03704	27
0.2	0.01923	52
0.3	0.03509	57
0.4	0.01266	79
0.5	0.02020	99
0.6	0.03704	108
0.7	0.03760	133
0.8	0.01653	121
0.9	0.02667	150
1.0	0.01936	155
2.0	0.00985	203
3.0	0.00000	226
4.0	0.00000	236
5.0	0.00823	243
6.0	0.01660	241
7.0	0.01397	242
8.0	0.03306	242
9.0	0.03674	245
10.0	0.05518	249
11.0	0.09465	243
12.0	0.06910	245
13.0	0.10976	248
14.0	0.11647	249
15.0	0.20988	243
16.0	0.21545	246
17.0	0.19919	246
18.0	0.25000	248
19.0	0.29918	244
20.0	0.28160	245
21.0	0.32530	249
22.0	0.41975	243
23.0	0.41463	246
24.0	0.41057	246
25.0	0.47581	248
26.0	0.53689	244
27.0	0.49388	245
28.0	0.54508	244
29.0	0.58367	245
30.0	0.63306	248
31.0	0.61538	247
32.0	0.61044	249

table 4.7 continued

BETA2 $(x10^{-5})$	POWER	NO. OF POINTS USED
33.0	0.67490	243
34.0	0.69512	246
35.0	0.69919	246
36.0	0.75806	248
37.0	0.75820	244
38.0	0.70612	245
39.0	0.75000	244
40.0	0.78776	245
41.0	0.80645	248
42.0	0.80162	247
43.0	0.82305	243
44.0	0.83065	248
45.0	0.85484	248
46.0	0.81967	244
47.0	0.85246	244
48.0	0.85950	242
49.0	0.86008	243
50.0	0.87347	245

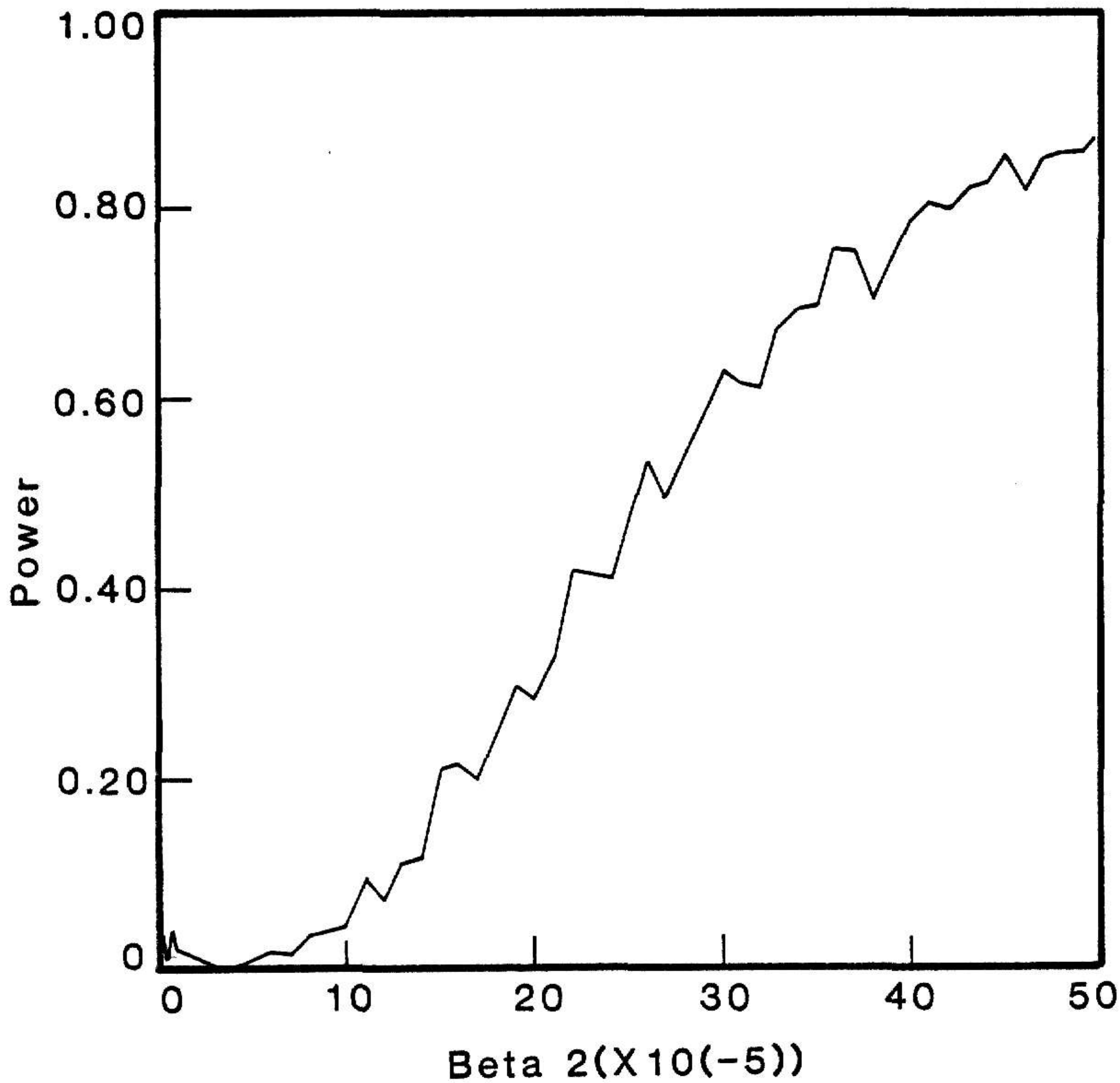

Figure 4.7 Power Curve for α = 1.0 using MMPM.

Table 4.8 Power for alpha=1.5, using MMPM

BETA2 $(x10^{-5})$	POWER	NO. OF POINTS USED
0.1	0.21875	32
0.2	0.10000	70
0.3	0.09302	86
0.4	0.14815	108
0.5	0.16438	146
0.6	0.08392	143
0.7	0.16352	159
0.8	0.08745	161
0.9	0.10053	189
1.0	0.09290	183
2.0	0.03766	239
3.0	0.0445	245
4.0	0.1606	249
5.0	0.00832	249
6.0	0.01205	249
7.0	0.02000	250
8.0	0.06855	248
9.0	0.06800	250
10.0	0.09639	249
11.0	0.07600	250
12.0	0.18952	248
13.0	0.22400	250
14.0	0.24900	249
15.0	0.27200	249
16.0	0.37750	249
17.0	0.37500	248
18.0	0.41767	249
19.0	0.49200	250
20.0	0.52800	250
21.0	0.53600	250
22.0	0.67871	249
23.0	0.64800	250
24.0	0.68800	250
25.0	0.68675	249
26.0	0.73200	250
27.0	0.73896	249
28.0	0.73577	246
29.0	0.78000	250
30.0	0.87200	250
31.0	0.81600	250
32.0	0.85600	250

table 4.8 continued

BETA2 $(\times 10^{-5})$	POWER	NO. OF POINTS USED
33.0	0.87500	249
34.0	0.87200	250
35.0	0.91968	249
36.0	0.86690	248
37.0	0.90760	249
38.0	0.92800	250
39.0	0.94400	250
40.0	0.93600	250
41.0	0.96386	249
42.0	0.95600	250
43.0	0.98000	250
44.0	0.95582	249
45.0	0.97200	250
46.0	0.94378	249
47.0	0.96341	246
48.0	0.97200	250
49.0	0.97200	250
50.0	0.98800	250

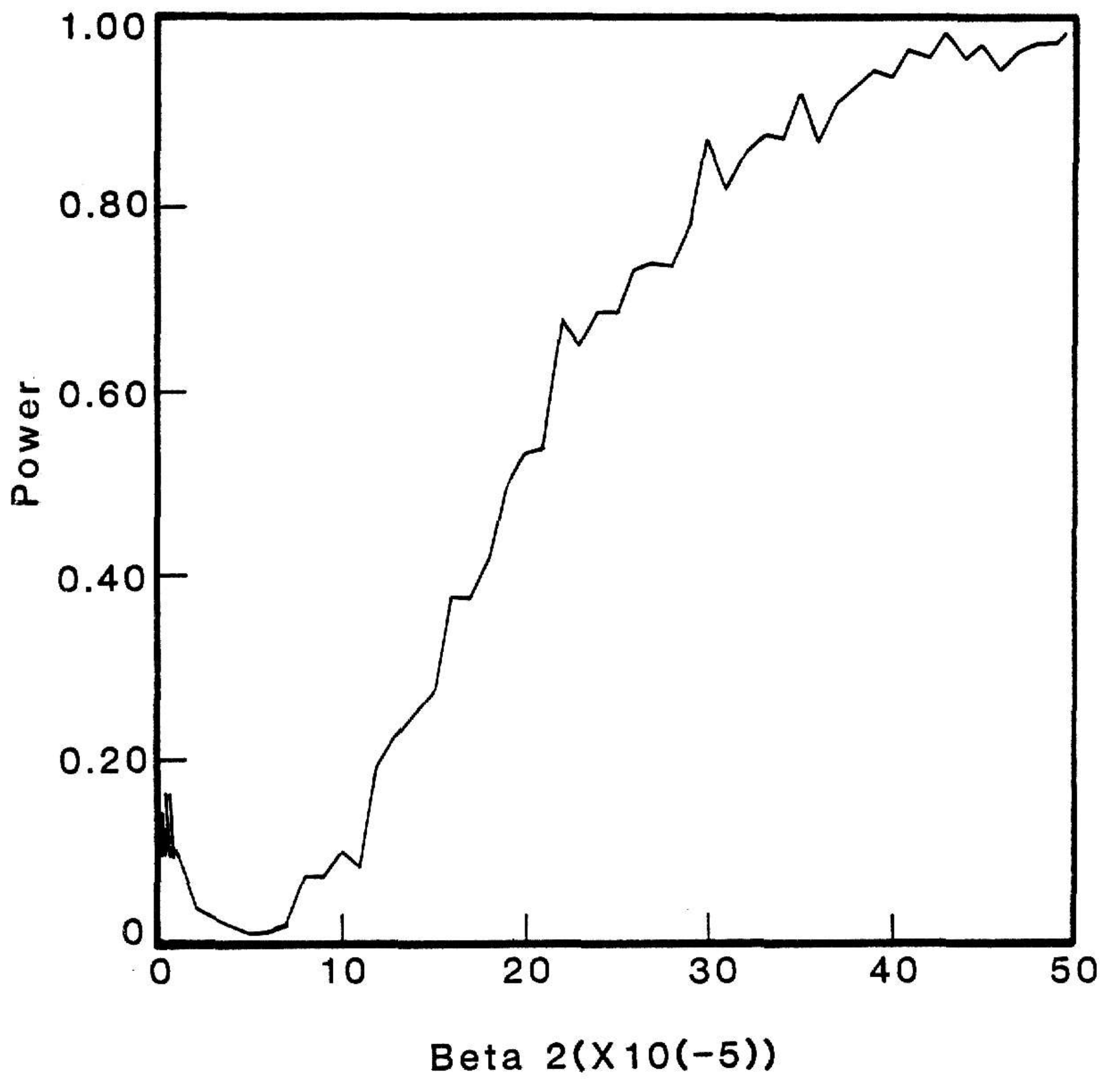

Figure 4.8 Power Curve for α = 1.5 using MMPM.

Table 4.9 Power for alpha=2.0, using MMPM

BETA2 $(x10^{-5})$	POWER	NO. OF POINTS USED
0.1	0.51282	39
0.2	0.30952	84
0.3	0.28696	115
0.4	0.33083	133
0.5	0.28902	173
0.6	0.24855	173
0.7	0.30270	185
0.8	0.25926	189
0.9	0.18779	213
1.0	0.16432	213
2.0	0.07347	245
3.0	0.03239	247
4.0	0.16000	250
5.0	0.12000	250
6.0	0.02000	250
7.0	0.04417	249
8.0	0.06400	250
9.0	0.11200	250
10.0	0.18000	250
11.0	0.18876	249
12.0	0.25600	250
13.0	0.28800	250
14.0	0.32300	250
15.0	0.46400	250
16.0	0.48996	249
17.0	0.49200	250
18.0	0.61044	249
19.0	0.62800	250
20.0	0.68000	250
21.0	0.74000	250
22.0	0.80800	250
23.0	0.74297	249
24.0	0.81200	250
25.0	0.85200	250
26.0	0.87600	250
27.0	0.86800	250
28.0	0.91200	250
29.0	0.91200	250
30.0	0.93600	250
31.0	0.93200	250
32.0	0.91600	250

table 4.9 continued

BETA2 $(x10^{-5})$	POWER	NO. OF POINTS USED
33.0	0.95984	249
34.0	0.94800	250
35.0	0.96400	250
36.0	0.96400	250
37.0	0.99600	250
38.0	0.97590	249
39.0	0.99600	250
40.0	0.99200	250
41.0	0.99600	250
42.0	0.98400	250
43.0	0.99600	250
44.0	0.98000	250
45.0	0.99200	250
46.0	0.99600	250
47.0	1.00000	250
48.0	0.99200	250
49.0	0.99600	250
50.0	1.00000	250

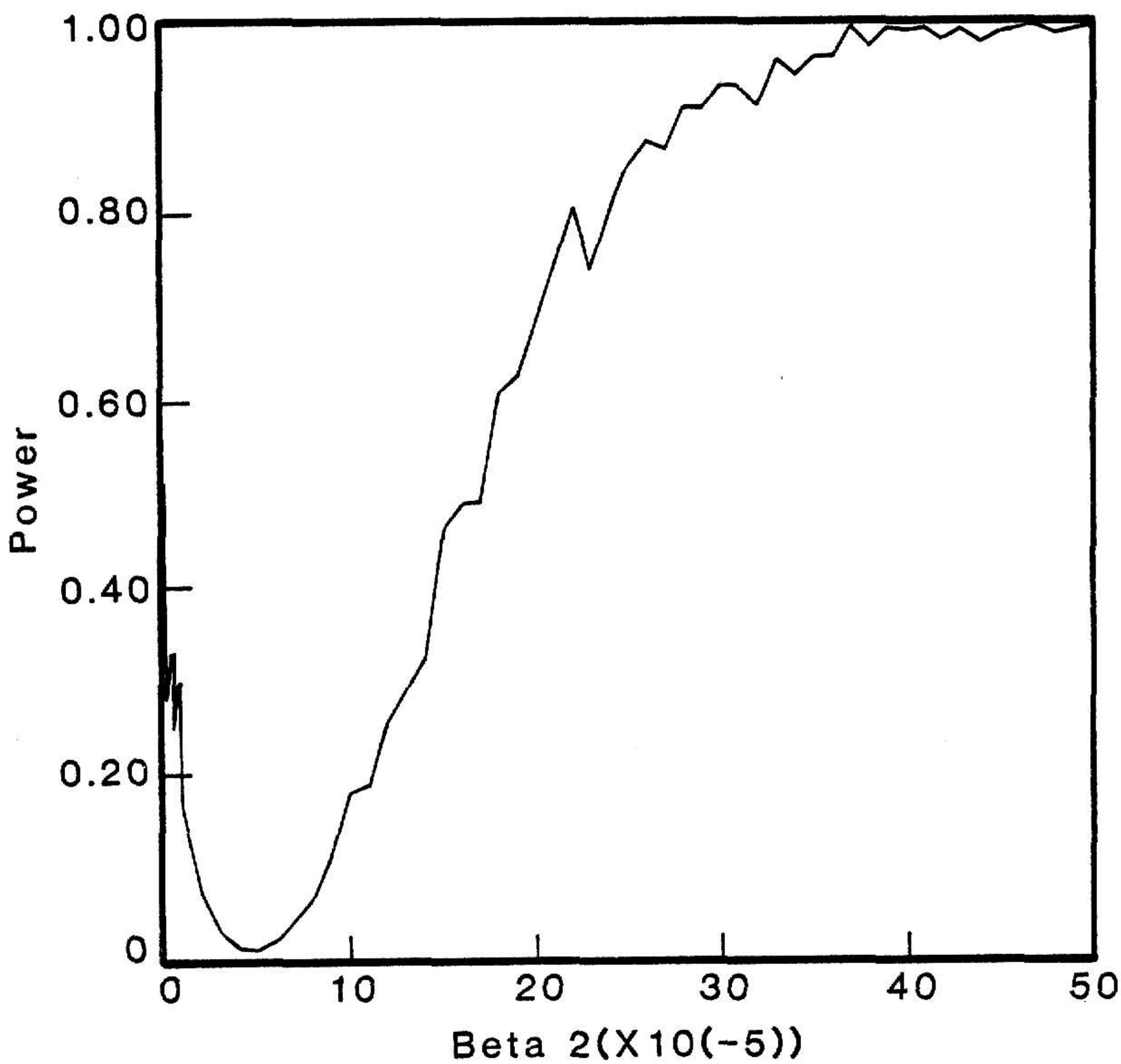

Figure 4.9 Power Curve for α = 2.0 using MMPM.

Table 4.10 Power for alpha=2.5, using MMPM

BETA2 ($x10^{-5}$)	POWER	NO. OF POINTS USED
0.1	0.68519	54
0.2	0.54082	98
0.3	0.50781	128
0.4	0.53459	159
0.5	0.47312	186
0.6	0.38421	190
0.7	0.47619	210
0.8	0.37864	206
0.9	0.37391	230
1.0	0.33190	232
2.0	0.11245	249
3.0	0.05622	249
4.0	0.04000	250
5.0	0.02000	250
6.0	0.02000	250
7.0	0.02800	250
8.0	0.04000	250
9.0	0.13200	250
10.0	0.11200	250
11.0	0.28000	250
12.0	0.31600	250
13.0	0.37200	250
14.0	0.43600	250
15.0	0.56000	250
16.0	0.56400	250
17.0	0.69200	250
18.0	0.72400	250
19.0	0.74800	250
20.0	0.80000	250
21.0	0.85200	250
22.0	0.81600	250
23.0	0.86400	250
24.0	0.90800	250
25.0	0.93600	250
26.0	0.92400	250
27.0	0.96400	250
28.0	0.95600	250
29.0	0.96400	250
30.0	0.97600	250
31.0	0.97200	250
32.0	0.98000	250

table 4.10 continued

BETA2 ($x10^{-5}$)	POWER	NO. OF POINTS USED
33.0	0.96800	250
34.0	1.00000	250
35.0	0.98800	250
36.0	0.99600	250
37.0	0.99600	250
38.0	1.00000	250
39.0	0.99200	250
40.0	0.99600	250
41.0	0.99200	250
42.0	0.99600	250
43.0	0.99200	250
44.0	1.00000	250
45.0	1.00000	250
46.0	1.00000	250
47.0	1.00000	250
48.0	1.00000	250
49.0	1.00000	250
50.0	1.00000	250

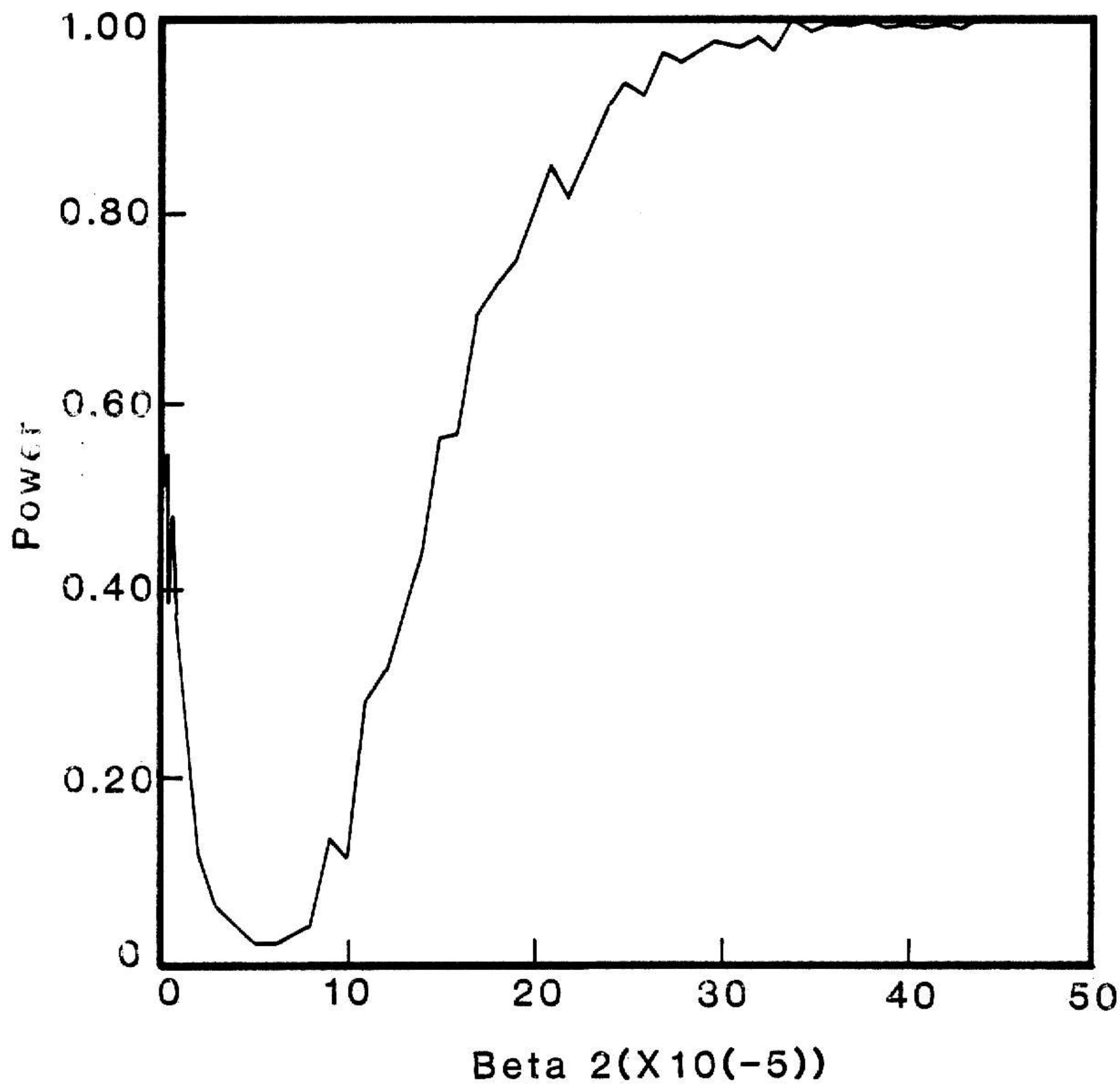

Figure 4.10 Power Curve for α = 2.5 using MMPM.

4.3 Analysis

For the MMMM and MMLM many sets of data were excluded from the calculation of power (see Tables 4.2, 4.3, 4.5, and 4.6). For example, from Table 4.3 (with $\alpha = 1.5$) the MMLM with BETA2 = 0.00002 yields only 41 out of 250 simulated data sets useful in the calculation of power. With the same ALPHA and BETA but using the MMMM (Table 4.2) 62 out of 250 simulated data sets were useful in the calculation of power. However for the same parameters, the MMPM (Table 4.1) yielded 239 (out of 250) useful data sets. This illustrates the tendency that over the entire range of BETA2 the MMMM and MMLM consistently yield nonuseful parameter estimates for a large number of sets of failure data. This is in contrast to the MMPM, which consistently uses a very large number of the failure data sets.

Checks were placed in the programs to find where and why the analysis of a failure data set failed. For MMMM many of the data sets yielded negative parameter estimates. As mentioned previously, a negative parameter estimate is not allowed (see definition of the gamma distribution) and therefore cannot be used for the calculation of any likelihood or a likelihood ratio. Thus, whenever a set of failure data yielded a negative estimator the data set was not used in the calculation of power.

For the MMLM many sets of data estimators could just not be found (a convergence problem). The subroutine DRTMI in the codes GAMMA GENERATE and GAMMA MODIFIED is responsible for finding the parameter estimates. The subroutine sets a minimum and maximum value between which the estimators are to be found and then proceeds in an iterative process to find them. If the parameter estimates cannot be found in this manner for a set of failure data a likelihood cannot be found for the set, and the set is dropped from the calculation of power. Thus any set of data which failed to yield parameter estimates was not used in the calculation of power.

For the MMPM the only time parameters could not be found was for failure data sets in which there were no failures. If in a given simulated failure data set, there are zero failures the MMPM cannot find parameter estimates, and, thus, a likelihood cannot be found. Hence, the set of failure data was not used to calculate the power. As the product of ALPHA times BETA decreases the probability of failures decreases so at lower values of ALPHA and BETA the number of data sets rejected by the MMPM increases. Even so the MMPM yields by far the largest number of parameter estimates which may be used in the calculation of power.

Whenever parameter estimates can be found it is possible to find likelihoods. By taking the natural logarithm of the likelihood function in the programs, problems with underflow have been eliminated in the calculation of likelihoods. Thus the major weakness of the programs does not lie in the calculation of likelihoods or in the calculation of likelihood ratios but in the parameter estimation techniques!

This work has shown that power curves can be generated from simulated failure data. The following observations about these power curves (Figs. 4.1 to 4.10) are of interest: 1). They are not symmetrical about BETA2 = 0.00005 (note: BETA1 = 0.00005), because the allowed beta parameter space is zero to infinity. Perhaps an equal logarithmic mesh for the Beta2 parameter between 0.00005 and zero would have revealed a power curve "pulse" which approaches unity as BETA2 approached zero. 2). All power curves exhibit a stochastic rather than a smooth behavior. For example, note that at BETA2 = 0.00005, where power should equal 0.05 (since BETA1 = 0.00005 and the level of significance equals 0.05), the power actually varies from 0.07 to 0.007. For ALPHA = 1.5 and the MMMM power curve, the power was zero for BETA2 equal to 0.00003, 0.00004, and 0.00006. Since the procedure for generating the power curves is stochastic, a repetition of the procedure would yield a distribution of power values for each value of BETA2; at BETA2 = 0.00005 the power distribution would have an expected value of 0.05. 3). All power curves appear to approach unity asymptotically for large values of BETA2. 4). Each power curve has a shape which approximates a negative pulse.

The power curve negative pulse behavior is of interest for further analysis. One of the ways a pulse width is commonly characterized (in spectroscopy) is by its full width at half maximum (FWHM) height. To help characterize a power curve consider its <u>full width at half minimum</u>, or since all the power curves have a maximum value of unity and a minimum value near zero (theoretically 0.05), full width at 0.5 (FW0.5).

For BETA2 values at which power equals 0.5 (end points of the FW0.5), the null hypothesis (i.e., the data sets are from the same gamma distribution) is correctly rejected only 50% of the time. Thus, if each half of the power pulse (divide the pulse at BETA2 = 0.00005) is thought of as a cumulative distribution, the derivative of each half would yield a curve with its most probable value at the BETA2 value which defines the FW0.5 end point.

Ideally (as shown in Fig. 2.12) a power curve for a two sided test should dip to zero at a single point and should be unity at all other points. Thus for the ideal case the FW0.5 equals zero. In actuality the smaller the value of FW0.5, or the steeper the sides of the power curve, the better the test is at differentiating between distributions. Thus a power curve with a FW0.5 of 0.00001 uses a technique which is better at differentiating between distributions than a technique which produces a power curve with a FW0.5 of 0.00025.

Table 4.11 gives a listing of FWO.5 for the generated power curves. For a given ALPHA value, the MMLM yields the smallest value of FWO.5. The values of FWO.5 generated by MMPM and MMMM are equal, and are 30% to 40% larger than the FWO.5 from the MMLM.

As the value of ALPHA increases the value of FWO.5 decreases. Thus, as the value of failure rate for which the gamma distribution is a maximum (the mode) increases, which occurs for gamma distributions as the ALPHA parameter increases, a data set from a gamma distribution with a BETA parameter nearer to that value of the parent distribution can be more correctly identified (i.e., H_o correctly rejected more often).

The power curve allows the analyst a choice of signifcance level or percentage of time which H_o is correctly rejected. For example, consider the power curve for ALPHA = 2.0 and the MMPM. The 0.05 significance level (95% confidence of correctly rejecting H_o) occurs at a BETA2 equal to about 0.00033, while the 0.01 significance level occurs at about 0.000365. Hence to answer the principal objective question for this work-how close together (or in shape) can failure rate distributions be to one another before they cannot be distinguished from one another - one must select the level of significance which is acceptable before a definitive answer is possible.

The MMPM may be the better parameter estimation procedure even though the FWO.5 for this method is larger than the other methods. Actual data are sometimes hard to obtain and or very expensive to obtain. If it is difficult to obtain data or the data are scarce, one would like to use as much of the data as possible. One does not want to waste or throw out data. As mentioned earlier the MMPMis able to find estimators for every case where there are some failures observed. For the MMMM and MMLM, however, allowable estimators are not found for many sets of data, hence,

Table 4.11 FW0.5 for generated power curves

Alpha	Estimation Technique	FW0.5* ($\times 10^{-5}$)
1.5	MMPM	17.6
1.5	MMMM	17.6
1.5	MMLM	13.3
2.0	MMPM	16.8
2.0	MMMM	16.8
2.0	MMLM	12.0
1.0	MMPM	25.0
1.5	MMPM	19.0
2.0	MMPM	17.5
2.5	MMPM	14.0

*FW0.5 is full width at 0.5. Note that power can vary between 1.0 and 0.0; thus, FW0.5 is nearly equivalent to a full width at half maximizying (a term often used in spectroscopy to characterize the width of spectral peaks).

both methods are wasteful of precious data. Thus even though the MMLM

both methods are wasteful of precious data. Thus even though the MMLM

appears to have the smallest value of FW0.5, if it cannot find a power because it cannot use the data, the estimation method is of questionable value. Hence, a compromise may be called for; perhaps the MMPM should be used.

Another advantage of using the MMPM is cost. The code GAMMA GENERATE is fairly expensive to run. All operations must be repeated three times, once for each estimation technique. The calculation of prior parameters using the MMLM increases the cost even more since an iterative process is required. The code GAMMAP MODIFIED, which is just the code GAMMA GENERATE with the MMMM and MMLM removed, costs less than one fourth the amount that the code GAMMA GENERATE costs to run for the analysis of equivalent failure data sets. Therefore the MMPM (as the parameter estimation technique) in the production of power curves is the best technique for two reasons: 1) MMPM uses more of the available data than either MMMM or MMLM and 2) MMPM costs less to produce results.

From all of the above, there conclusions may be drawn:

1) Power curves can be generated using simulated failure data.
2) As alpha increases, power curves have a smaller value of FW0.5 and thus are better at distinguishing between different distributions.
3) MMPM is the best estimation technique to use in the generation of power curves.

5 ACKNOWLEDGMENTS

The author wishes to express his thanks to Professors N. Dean Eckhoff, G.A. Milliken, and D. Grosh for their guidance, advice, and assistance during the course of this work. Thanks are given to Ms. Sherry Berner for typing this manuscript and to Ms. Barbara Wieliczka for producing the figures. Special thanks to my wife, Donna Drayer, for her encouragement and advice. This research was supported by a contract with EG & G, Inc. and the U.S. Nuclear Regulatory Commission.

6 REFERENCES

1. I. Miller and J.E. Freund, Probability and Statistics for Engineers, Pretice-Hall, 1977.

2. J.K. Shultis, D.E. Johnson, G.A. Milliken, and N.D. Eckhoff, "Non-Conjugate Prior Distributions in Compound Failure Models," KSU Report 2705c, Kansas State University, Manhattan, Kansas, 1980.

3. J.K. Shultis, D.E. Johnson, G.A. Milliken, and N.D. Eckhoff, "Gamma-A Code for the Analysis of Component Failure Rates with a Compound Posiion-Gamma Model," KSU Report 2075a, 1980.

4. A.M. Mood, F.A. Graybill, and D.C. Boes, Introduction to the Theory of Statistics, McGraw-Hill, New York, 1974.

5. D. Grosh, "The Gamma and Beta Functions," unpublished class notes from Introductory Probability and Statistics, fall 1980.

6. E.G. Kirkpatrick, Introductory Statistics and Probability for Engineering, Sciences, and Technology, Pretice-Hall, Englewood Clifs, New Jersey, 1974.

7. K.A. Brownlee, Statistical Theory and Methodology in Science and Engineering, John Wiley and Sons, New York, 1965.

APPENDIX

Computer Code Listings

Computer Code Listing for

GENERATE

```
C       ****************************************************************
C       *                                                              *
C       *  GENERATE                                                    *
C       *                                                              *
C       *                                                              *
C       *  GIVEN TWO SETS OF PARAMETERS (TIME, ALPHA, AND BETA) THIS   *
C       *  PROGRAM GENERATES THE CUMULATIVE MARGINAL DISTRIBUTION      *
C       *  FOR EACH SET.  FROM THE CUMULATIVE DISTRIBUTION THE PROGRAM *
C       *  SIMULATES FAILURE DATA (NUMBER OF FAILURES PER A GIVEN TIME). *
C       *  THE NUMBER OF DESIRED SETS OF FAILURE DATA MAY BE SPECIFIED *
C       *  ALONG WITH THE NUMBER OF DATA PAIRS PER SET.                *
C       *  AFTER THE FAILURE RATE INFORMATION IS GENERATED IT IS WRITTEN *
C       *  ONTO TAPE.                                                  *
C       *                                                              *
C       *                                                              *
C       *     INPUT CARDS                                              *
C       *                                                              *
C       *       CARD 1 FIRST DISTRIBUTION PARAMETERS TIME, ALPHA,BETA.  *
C       *       FORMAT(E13.5,7X,E13.5,7X,E13.5)                        *
C       *                                                              *
C       *       CARD 2  SECOND DISTRIBUTION PARAMETERS  TIME, ALPHA,   *
C       *       BETA.                                                  *
C       *       FORMAT(E13.5,7X,E13.5,7X,E13.5)                        *
C       *                                                              *
C       *       CARD 3 NUMBER OF DATA PAIRS PER SET.                   *
C       *       FORMAT(I6)                                             *
C       *                                                              *
C       *       CARD 4  NUMBER OF SETS.                                *
C       *       FORMAT(I6)                                             *
C       *                                                              *
C       ****************************************************************
C       PROGRAM IS IN DOUBLE PRECISION
        IMPLICIT REAL*8(A-H,O-Z)
        DIMENSION HARRAY(30),YARRAY(30),NF(10000),NT(10000),NF2(10000),NT2
       *(10000)
        INTEGER F,COUNT,FARRAY,COUNT2
C       LET THE STARTING VALUE FOR RANDU ARBITRARILY EQUAL 12345
        IX=12345
C
C       READ IN AND ESTABLISH CUMULATIVE DISTRIBUTION FOR ALPHA1 & BETA1
C       READ IN VALUES FOR TIME(=T) AND THE GAMMA FUNCTION
C       PARAMETERS (ALPHA AND BETA).
C
        READ(5,1)T,ALPHA,BETA
    1   FORMAT(E13.5,7X,E13.5,7X,E13.5)
        WRITE(6,2)T,ALPHA,BETA
    2   FORMAT('1','TIME=',E13.5,'HRS.',5X,'ALPHA=',E13.5,3X,'BETA=',E13.5
       /)
        TAU=1.0/BETA
        COUNT=0
        OLD=0
        DO 3 I=1,200
        COUNT=COUNT+1
        F=I-1
        TOP=(DGAMMA(F+ALPHA))*(T**F)*(TAU**ALPHA)
        BOT=(DGAMMA(ALPHA))*(DGAMMA(F+1.0D+00))*((T+TAU)**(F+ALPHA))
        OLD=OLD+(TOP/BOT)
        HARRAY(I)=OLD
```

```
      IF(HARRAY(I).GT.0.99999)GCTC 5
    3 CCNTINUE
      WRITE(6,4)
    4 FCRMAT(' ','NOT ENCUGH TIMES THROUGH LCCP TO GET CUMULATIVE DISTRI
     /BUTICN.GT.0.99999')
      GCTC 20CC
C
C     WRITE CUMULATIVE DISTRIBUTICN 1
C
    5 WRITE(6,6)
    6 FCRMAT('0','CUMULATIVE DISTRIBUTION')
      WRITE(6,7)
    7 FCRMAT(' ','F=',6X,'CUMULATIVE PROBABILITY')
      DC 9 I=1,COUNT
      F=I-1
      WRITE(6,8)F,HARRAY(I)
    8 FCRMAT(' ',I3,5X,E13.5)
    9 CCNTINUE
C
C     READ IN AND ESTABLISH CUMULATIVE DISTRIBUTICN FCR ALPHA2 & BETA2
C     READ IN VALUES FOR TIME(=T) AND THE GAMMA FUNCTICN
C     PARAMETERS (ALPHA AND BETA
C
      READ 10,T2,ALPHA2,BETA2
   10 FCRMAT(E13.5,7X,E13.5,7X,E13.5)
      WRITE(6,11)T2,ALPHA2,BETA2
   11 FCRMAT('-','TIME2=',E13.5,'HRS.',5X,'ALPHA2=',E13.5,3X,'BETA2=',E1
     /3.5)
      TAU2=1.0/BETA2
      CCUNT2=0
      OLD=0
      DC 12 I=1,200
      CCUNT2=CCUNT2+1
      F=I-1
      TOP=(DGAMMA(F+ALPHA2))*(T2**F)*(TAU2**ALPHA2)
      BCT=(DGAMMA(ALPHA2))*(DGAMMA(F+1.0D+00))*((T2+TAU2)**(F+ALPHA2))
      OLD=OLD+(TOP/BOT)
      YARRAY(I)=OLD
      IF(YARRAY(I).GT.0.99999)GCTC 14
   12 CCNTINUE
      WRITE(6,13)
   13 FCRMAT(' ','NOT ENCUGH TIMES THROUGH LOOP TO GET CUMULATIVE DISTRI
     /BUTICN.GT.0.99999')
      GCTC 200C
C
C     WRITE CUMULATIVE DISTRIBUTICN 2
C
   14 WRITE(6,15)
   15 FCRMAT('C','CUMULATIVE DISTRIBUTICN')
      WRITE(6,16)
   16 FCRMAT(' ','F=',6X,'CUMULATIVE PROBABILITY')
      DC 18 I=1,COUNT2
      F=I-1
      WRITE(6,17)F,YARRAY(I)
   17 FCRMAT(' ',I3,5X,E13.5)
   18 CCNTINUE
C
C
```

```
C     SECTICN TC GENERATE FRECUENCIES
C
C
C
C
C     READ NUMBER OF DATA PAIRS PER SET
C
      READ 19,N
   19 FCRMAT(I6)
C
C     READ N2, THE NUMBER OF SETS DESIRED
C
      READ 20,N2
   20 FCRMAT(I3)
      J1=0
      DC 1000 M=1,N2
      WRITE(6,21)M
   21 FCRMAT('C','SET NUMBER',I3)
C
C     DISTRIBUTICN NUMBER 1
C
      WRITE(6,121)
  121 FCRMAT('0','FIRST DISTRIBUTICN        NF      NT')
C     N IS THE NUMBER OF DATA PCINTS DESIRED
      DC 26 J=1,N
      J1=J1+1
C     GENERATE RANDCM NUMBERS
   22 CALL RANDU(IX,IY,U)
      IX=IY
C     SINCE CUMULATIVE CNLY GCES UP TO 0.99999, MUST CHECK TO SEE THAT
C     THE R.N. IS.LE.C.99999.
      IF(U.GT.C.99999)GOTO 22
C     FIND WHAT F CORRESPCNDS TC R.N.
      DC 23 K=1,COUNT
      UPPER=HARRAY(K)
      IF(U.LT.UPPER)GCTO 24
   23 CCNTINUE
   24 NF(J1)=K-1
      NT(J1)=T
      WRITE(6,125)NF(J1),NT(J1)
  125 FCRMAT(' ',19X,I8,2X,I8)
      WRITE(9,25)NF(J1),NT(J1)
   25 FCRMAT(I8,I8)
   26 CCNTINUE
C
C     DISTRIBUTICN NUMBER 2
C
      WRITE(6,126)
  126 FCRMAT('C','SECCND DISTRIBUTICN')
      DC 31 J=1,N
      J1=J1+1
C     GENERATE RANDCM NUMBERS
   27 CALL RANDU(IX,IY,U)
      IX=IY
C     SINCE CUMULATIVE CNLY GCES UP TO 0.99999, MUST CHECK TO SEE THAT
C     THE R.N. IS.LE.C.99999
      IF(U.GT.C.99999)GOTO 27
C     FIND WHAT F CORRESPCNDS TC R.N.
```

```
      DC 28 K=1,CCUNT2
      UPPER=YARRAY(K)
      IF(U.LT.UPPER)GCTO 29
28    CCNTINUE
29    NF2(J1)=K-1
      NT2(J1)=T2
      WRITE(6,130)NF2(J1),NT2(J1)
130   FCRMAT(' ',19X,I8,2X,I8)
      WRITE(9,30)NF2(J1),NT2(J1)
30    FCRMAT(I8,I8)
31    CCNTINUE
1000  CCNTINUE
      WRITE(6,2001)
2001  FCRMAT('1','END')
2000  STOP
      END


      SUBRCUTINE RANDU(IX,IY,YFL)
      REAL*8 YFL
      IY=IX*65539
      IF(IY)5,6,6
5     IY=IY+2147483647+1
6     YFL=IY
      YFL=YFL*.4656613E-9
      RETURN
      END
```

Computer Code Listing for

GAMMA MODIFIED

```
C     ******************************************************************
C     *                                                                *
C     *  GAMMA MODIFIED                                                *
C     *                                                                *
C     *     THIS PROGRAM IS A MCDIFICATICN CF THE PRCGRAM GAMMA8       *
C     *  WRITTEN BY J.K. SHULTIS IN JUNE OF 1980.  THIS PROGRAM READS  *
C     *  SETS OF FAILURE RATE DATA FROM TAPE, ESTIMATES PARAMETERS,    *
C     *  FINDS THE LIKELIHOOD RATIC BETWEEN SETS OF DATA, AND          *
C     *  CALCULATES POWER.                                             *
C     *  THE STATISTICAL MODEL ASSUMED IS THE CCMPOUND POISSON-GAMMA   *
C     *  MODEL IN WHICH THE FAILURE RATES FOR EACH COMPONENT MAY VARY  *
C     *  ACCORDING TC A GAMMA PRIOR DISTRIBUTICN.  THE PARAMETERS OF   *
C     *  THE GAMMA PRIOR ARE ESTIMATED FROM THE GIVEN ATTRIBUTE DATA BY*
C     *  THE FCLLOWING METHODS;                                        *
C     *   (1) MATCHING UNWEIGHTED DATA MOMENTS TO THE PRIOR            *
C     *   DISTRIBUTION                                                 *
C     *   (2) MATCHING UNWEIGHTED DATA MOMENTS TO THE MARGINAL         *
C     *   DISTRIBUTION                                                 *
C     *   DISTRIBUTION                                                 *
C     *                                                                *
C     *                                                                *
C     *  INPUT DATA                                                    *
C     *                                                                *
C     *     CN TAPE;                                                   *
C     *                                                                *
C     *       FAILURE RATE DATA, NUMBER OF FAILURES, NF, TIME (IN      *
C     *      HOURS), NT / FORMAT(I8,I8)                                *
C     *                                                                *
C     *                                                                *
C     *     CN CARDS;                                                  *
C     *                                                                *
C     *      CARD 1 NUMBER CF PCINTS NNN NUMBER CF SETS N2 / FORMAT    *
C     *      (I3,I3)                                                   *
C     *                                                                *
C     *      CARD 2 TITLE / FORMAT(10A8)                               *
C     *                                                                *
C     *                                                                *
C     *                 DARRYL DRAYER, KANSAS STATE UNIVERSITY 4/81    *
C     *                                                                *
C     ******************************************************************
C
      IMPLICIT REAL*8(A-H,O-Z)
      REAL*8 LAM(5000),LAML(5000),LAMU(5000)
      REAL*8 TITLE(10),ID(5000),LAMBAR,LAMLOW,LAMUP,LNPROB,VAR(4)
      REAL*8 ALP(4),BET(4),ALIKE(5,5000)
      REAL*8 ARATIO(3,5000)
      COMMON /DTA/ID,NF(5000),NT(5000),NN(5000),N,NDATA,IPRT,NITER,J
      EXTERNAL FN
C
C
C  SET PRINTER UNIT=6; SET READER UNIT=5; REL. ACC OF PERCENTILES
C  IS SET BY EPPS; MAX NUMBER OF ITERATIONS FCR EVALUATING PERCENTILE
C  SET BY VARIABLE NITER
      IPRT = 6
      IRDR = 5
```

```
      NITER = 20
C     READ IN NUMBER OF POINTS NNN, & NUMBER OF SETS, N2.
      READ 1,NNN,N2
    1 FORMAT(I3,I3)
C     READ DATA OFF TAPE & SET ARRAYS
      L=2*NNN*N2
      DO 4 I=1,L
      READ(9,2)NF(I),NT(I)
    2 FORMAT(I8,I8)
      NN(I)=1
      ID(I)=I
    4 CONTINUE
      J=1
      N2CONT=0
    5 NCOUNT=0
      NOLDJ=J
    6 NDATA=J+NNN-1
    7 CONTINUE
C
C
      PCT=0.1
      BMIN=0.0
      BMAX=0.0
      EPS=0.0001
C  CALCULATE TOTAL NUMBER OF COMPONENTS N
      N = 0
      DO 8 I=J,NDATA
      IF (NN(I).LE.0) NN(I) = 1
    8 N=N+NN(I)
C
C  CALCULATE MEAN AND VARIANCE OF THE DATA
      SUMT = 0.0
      SUMF = 0.0
      SUMFT = 0.0
      SUMFT2 = 0.0
      SUMTI = 0.0
      DO 9 I=J,NDATA
      TT = NT(I)
      SUMT = SUMT+TT*NN(I)
      SUMTI = SUMTI+NN(I)/TT
      SUMF = SUMF+NF(I)*NN(I)
      AA = NF(I)/TT
      SUMFT = SUMFT+AA*NN(I)
    9 SUMFT2 = SUMFT2+AA*NN(I)*AA
      IF(SUMFT.EQ.0.00)GOTO 100
      GOTO 101
  100 ALP(1)=0.0
      BET(1)=0.0
      ALP(2)=0.0
      BET(2)=0.0
      ALP(3)=0.0
      BET(3)=0.0
      GOTO 13
  101 UMEAN = SUMFT/N
      UVAR = (SUMFT2-UMEAN*UMEAN*N)/(N-1)
C
C
C
C***    BEGIN ANALYSIS FOR COMPOUND MODEL;
```

```
C
C  CALCULATE MATCHING MOMENTS ESTIMATORS
      BETAUP = UVAR/UMEAN
      ALPHUP = UMEAN/BETAUP
      BETAUM = (UVAR-UMEAN*SUMTI/N)/UMEAN
      IF(BETAUM.EQ.0.D0)GOTO 102
      ALPHUM = UMEAN/BETAUM
      GOTO 103
  102 ALPHUM=0.0
C
C***  UNWEIGHTED MATCHING MOMENTS ESTIMATORS TO THE PRIOR - METHOD 1
  103 AAA=ALPHUP
      BBB = BETAUP
      ALP(1)=AAA
      BET(1)=BBB
C
C***  UNWEIGHTED MATCHING MOMENTS TO THE MARGINAL - METHOD 2
   10 AAA = ALPHUM
      BBB = BETAUM
      ALP(2)=AAA
      BET(2)=BBB
      IF(BET(2).LE.0.D0)GOTO 12
C
C***  MAXIMUM LIKELIHOOD ESTIMATORS - METHOD 3
      IF (BMIN.EQ.0.0) BMIN = BETAUP*0.01
      IF (BMAX.EQ.0.0) BMAX = 100.0*BETAUP
      CALL CRTMI (BETA,FF,FN,BMIN,BMAX,EPS,NITER,IER)
      IF (IER.NE.0)GOTO 11
      CALL MLGAMP (FFF,ALPHA,BETA)
      AAA = ALPHA
      BBB = BETA
      ALP(3)=AAA
      BET(3)=BBB
      GOTO 13
   11 CONTINUE
      ALP(3)=0.0
      BET(3)=0.0
      GOTO 13
C
C
   12 ALP(3)=0.0
      BET(3)=0.0
   13 NCOUNT=NCOUNT+1
C
C     FIND LIKELIHOOD FOR EACH GROUP
C
      DO 25 I=1,3
      IF(BET(I).LE.0.D0)GOTO 19
      BET(I)=1.D0/BET(I)
      FIRST=1.D0
      DO 16 I2=J,NDATA
      AT=NT(I2)
      OOPS=NF(I2)+ALP(I)
      IF(OOPS.GT.35.D0)GOTO 17
      IF(ALP(I).LE.0.D0)GOTO 19
      IF(BET(I).LE.0.D0)GOTO 19
      IF(OOPS.LT.0.D0)GOTO 19
      X1=NF(I2)+ALP(I)
      X2=NF(I2)
```

```
      X3=NT(I2)
      X4=ALP(I)
      X5=BET(I)
      X6=NF(I2)+1.D0
      X7=NT(I2)+BET(I)
      TOP=DLGAMA(X1)+X2*DLOG(X3)+X4*DLOG(X5)
      BOT=DLGAMA(X4)+DLGAMA(X6)+X1*DLOG(X7)
      DIF=TOP-BOT
      IF(DIF.LT.-100.0D0)GOTO 23
      IF(DIF.GT.170.0D0)GOTO 21
      ALIKE(NCOUNT,I)=DEXP(DIF)*FIRST
C
C
      IF(DABS(ALIKE(NCOUNT,I)).LT.1.0D-50)ALIKE(NCOUNT,I)=0.0
      FIRST=ALIKE(NCOUNT,I)
   16 CONTINUE
      GOTO 25
   17 ALIKE(NCOUNT,I)=0
      WRITE(6,18)NCOUNT,I
   18 FORMAT('0','LIKELIHOOD RATIO FOR ALIKE(',I3,I3,') SET EQUAL TO ZER
     /O BECAUSE DGAMMA IS TOO LARGE TO EVALUATE')
      GOTO 25
   19 ALIKE(NCOUNT,I)=0
      GOTO 25
   21 WRITE(6,22)NCOUNT,I
      ALIKE(NCOUNT,I)=0
   22 FORMAT('0','ALIKE(',I3,I3,') SET EQUAL TO 0 BECAUSE OF POWER')
      GOTO 25
   23 WRITE(6,24)NCOUNT,I
   24 FORMAT(' ','ALIKE(',I3,I3,') SET EQUAL TO ZERO BECAUSE OF UNDERFLO
     /W')
      ALIKE(NCOUNT,I)=0
   25 CONTINUE
      IF(NCOUNT.EQ.2)GOTO 26
      IF(NCOUNT.EQ.3)GOTO 27
      J=J+NNN
      GOTO 6
   26 J=NOLDJ
      NDATA=(2*NNN)+J-1
      GOTO 7
   27 CONTINUE
      N2CONT=N2CONT+1
      DO 32 I=1,3
      DO 28 I30=1,3
      IF(ALIKE(I30,I).EQ.0.0)GOTO 30
   28 CONTINUE
      ARATIO(I,N2CONT)=ALIKE(3,I)/(ALIKE(1,I)*ALIKE(2,I))
      GOTO 32
   30 ARATIO(I,N2CONT)=0.0
   32 CONTINUE
      J=J+2*NNN
      IF(N2CONT.EQ.N2)GOTO 33
      GOTO 5
   33 CONTINUE
C
C
C
      DO 38 I=1,3
      NREJET=0
```

```
      SUBROUTINE MLGAMP (FN,ALPHA,BETA)
C***********************************************************************
C*                                                                     *
C* THIS SUBROUTINE EVALUATES ALPHA AND THE FUNCTION FN WHOSE ZERO      *
C* GIVES THE VALUE OF BETA FOR THE MAXIMUM LIKELIHOOD ESTIMATORS.      *
C*                                                                     *
C***********************************************************************
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON /DTA/IDUMMY(10000),NF(5000),NT(5000),NN(5000),N,NDATA,IPRT,
     /NITER,J2
C
C     CALCULATE ALPHA AS A FUNCTION OF BETA
      SUM1 = 0.
      SUM2 = 0.
      FF = 0.
      DO 1 J=J2,NDATA
      FF = FF+NF(J)*NN(J)
      A = NN(J)*NT(J)/(1.+BETA*NT(J))
      SUM1 = SUM1+A
    1 SUM2 = SUM2+A*NF(J)
      ALPHA = (FF/BETA-SUM2)/SUM1
C
C     CALCULATE F(ALPHA,BETA)
      FN = 0.
      DO 4 J=J2,NDATA
      A = DLOG(1.+BETA*NT(J))
      SUM = 0.
      NNN = NF(J)
      IF (NNN.EQ.0) GO TO 3
      DO 2 K=1,NNN
    2 SUM = SUM+1./(ALPHA+K-1.)
    3 FN = FN+(A-SUM)*NN(J)
    4 CONTINUE
      RETURN
      END
```

```
      REAL FUNCTION FN*8(X)
C***********************************************************************
C*                                                                     *
C* THIS FUNCTION EVALUATES THE AUXILARY LIKELIHOOD FUNCTION FN WHOSE   *
C* ZERO GIVES THE VALUE OF THE BETA PARAMETER.  THIS SUBROUTINE SIMPLY *
C* CALLS MLGAMP WHERE THE ACTUAL EVALUATION IS PERFORMED.              *
C*                                                                     *
C***********************************************************************
      REAL*8 ALPHA,X
      CALL MLGAMP (FN,ALPHA,X)
      RETURN
      END
```

```
      SUBROUTINE DRTMI (X,F,FCT,XLI,XRI,EPS,IEND,IER,NOOOPS)
C******************************* DRTMI ********************************
C*                                                                    *
C* THIS SUBROUTINE SOLVES A NONLINEAR EQUATION OF THE FORM FCT(X)=0   *
C* BY THE MUELLER'S ITERATION METHOD.                                 *
C* THIS PROGRAM IS TAKEN FROM THE IBM SSP (RTMI) AND MODIFIED TO WORK *
C* IN DOUBLE PRECISION.                                               *
C*                                                                    *
C*    INPUT VARIABLES:                                                *
C*         X    = RETURNED VALUE OF THE ZERO                          *
C*         F    = VALUE OF THE FUNCTION FCT AT THE ZERO               *
C*         FCT  = EXTERNAL FUNCTION WHOSE ZERO IS TO BE FOUND         *
C*         XLI  = LEFT BOUNDARY OF THE X-AXIS TO BE SEARCHED FOR ROOT *
C*         XRI  = RIGHT BOUNDARY OF X-AXIS TO BE SEARCHED FOR ROOT    *
C*         EPS  = ACCURACY OF DESIRED RESULT                          *
C*         IEND = MAXIMUM NUMBER OF ITERATIONS TO BE USED             *
C*         IER  = ERROR RETURN CODE; =0 IF DESIRED ACCURACY ACHIEVED  *
C*                                                                    *
C**********************************************************************
      IMPLICIT REAL*8(A-H,O-Z)
C
C
C     PREPARE ITERATION
      IER = 0
      XL = XLI
      XR = XRI
      X = XL
      TOL = X
      F = FCT(TOL)
      IF (F) 1,16,1
    1 FL = F
      X = XR
      TOL = X
      F = FCT(TOL)
      IF (F) 2,16,2
    2 FR = F
      IF (DSIGN(1.D0,FL)+DSIGN(1.D0,FR)) 25,3,25
C
C     BASIC ASSUMPTION FL*FR LESS THAN 0 IS SATISFIED.
C     GENERATE TOLERANCE FOR FUNCTION VALUES.
    3 I = 0
      TOLF = 100.*EPS
C
C     START ITERATION LOOP
    4 I = I+1
C
C     START BISECTION LOOP
      DO 13 K=1,IEND
      X = .5*(XL+XR)
      TOL = X
      F = FCT(TOL)
      IF (F) 5,16,5
    5 IF (DSIGN(1.D0,F)+DSIGN(1.D0,FR)) 7,6,7
C
C     INTERCHANGE XL AND XR IN ORDER TO GET THE SAME SIGN IN F AND FR
```

```
    6 TOL = XL
      XL = XR
      XR = TOL
      TOL = FL
      FL = FR
      FR = TOL
    7 TOL = F-FL
      A = F*TOL
      A = A+A
      IF (A-FR*(FR-FL)) 8,9,9
    8 IF (I-IEND) 17,17,9
    9 XR = X
      FR = F
C
C     TEST ON SATISFACTORY ACCURACY IN BISECTION LOOP
      TOL = EPS
      A = DABS(XR)
      IF (A-1.) 11,11,10
   10 TOL = TOL*A
   11 IF (DABS(XR-XL)-TOL) 12,12,13
   12 IF (DABS(FR-FL)-TOLF) 14,14,13
   13 CONTINUE
C     END OF BISECTION LOOP
C
C     NO CONVERGENCE AFTER IEND ITERATION STEPS FOLLOWED BY IEND
C     SUCCESSIVE STEPS OF BISECTION OR STEADILY INCREASING FUNCTION
C     VALUES AT RIGHT BOUNDS. ERROR RETURN.
      IER = 1
   14 IF (DABS(FR)-DABS(FL)) 16,16,15
   15 X = XL
      F = FL
   16 RETURN
C
C     COMPUTATION OF ITERATED X-VALUE BY INVERSE PARABOLIC INTERPOLATION
   17 A = FR-F
      DX = (X-XL)*FL*(1.+F*(A-TOL)/(A*(FR-FL)))/TOL
      XM = X
      FM = F
      X = XL-DX
      TOL = X
      F = FCT(TOL)
      IF (F) 18,16,18
C
C     TEST ON SATISFACTORY ACCURACY IN ITERATION LOOP
   18 TOL = EPS
      A=DABS(X)
      IF (A-1.) 20,20,19
   19 TOL = TOL*A
   20 IF (DABS(DX)-TOL) 21,21,22
   21 IF (DABS(F)-TOLF) 16,16,22
C
C     PREPARATION OF NEXT BISECTION LOOP
   22 IF (DSIGN(1.D0,F)+DSIGN(1.D0,FL)) 24,23,24
   23 XR = X
      FR = F
      GO TO 4
   24 XL = X
      FL = F
      XR = XM
```

```
      FR = FM
      GOTO 4
C     END OF ITERATION LOOP
C
C
C     ERROR RETURN IN CASE OF WRONG INPUT DATA
   25 IER = 2
      RETURN
      END
```

Computer Code Listing for

GAMMA GENERATE

```
C     ******************************************************************
C     *                                                                *
C     *  GAMMA GENERATE                                                *
C     *                                                                *
C     *                                                                *
C     *  GENERATE                                                      *
C     *  GIVEN TWO SETS OF PARAMETERS (TIME, ALPHA, AND BETA) THIS     *
C     *  SECTION GENERATES THE CUMULATIVE MARGINAL DISTRIBUTION        *
C     *  FOR EACH SET.  FROM THE CUMULATIVE DISTRIBUTION THE PROGRAM   *
C     *  SIMULATES FAILURE DATA (NUMBER OF FAILURES PER A GIVEN TIME). *
C     *  THE NUMBER OF DESIRED SETS OF FAILURE DATA MAY BE SPECIFIED   *
C     *  ALONG WITH THE NUMBER OF DATA PAIRS PER SET.                  *
C     *                                                                *
C     *  GAMMA                                                         *
C     *  THE SECOND SECTION OF THE PROGRAM IS A MODIFICATION OF THE    *
C     *  PROGRAM GAMMA8 WRITTEN BY J.K. SHULTIS IN JUNE OF 1980.  THIS *
C     *  SECTION TAKES SET OF FAILURE RATE DATA FROM GENERATE,         *
C     *  ESTIMATES PRIOR PARAMETERS, FINDS THE LIKELIHOOD RATIO BETWEEN*
C     *  SETS OF DATA, AND CALCULATES POWER.  THE STATISTICAL MODEL    *
C     *  ASSUMED IS THE COMPOUND POISSON-GAMMA MODEL IN WHICH THE      *
C     *  FAILURE RATES FOR EACH COMPONENT MAY VARY ACCORDING TO A      *
C     *  GAMMA PRIOR DISTRIBUTION.  THE PARAMETERS OF THE GAMMA PRIOR  *
C     *  ARE ESTIMATED BY THREE METHODS: MATCHING UNWEIGHTED DATA      *
C     *  MOMENTS TO THE PRIOR, MATCHING UNWEIGHTED DATA MOMENTS TO THE *
C     *  MARGINAL, AND MAXIMUM LIKELIHOOD.                             *
C     *                                                                *
C     *                                                                *
C     *     INPUT CARDS                                                *
C     *                                                                *
C     *       CARD 1  FIRST DISTRIBUTION PARAMETERS/ TIME, ALPHA, BETA.*
C     *       FORMAT(E13.5,7X,E13.5,7X,E13.5)                          *
C     *                                                                *
C     *       CARD 2  SECOND DISTRIBUTION PARAMETERS  TIME, ALPHA,     *
C     *       BETA.                                                    *
C     *       FORMAT(E13.5,7X,E13.5,7X,E13.5)                          *
C     *                                                                *
C     *       CARD 3  NUMBER OF FAILURE DATA PAIRS PER SET.            *
C     *       FORMAT(I6)                                               *
C     *                                                                *
C     *       CARD 4  NUMBER OF SETS.                                  *
C     *       FORMAT(I6)                                               *
C     *                                                                *
C     *       CARD 5  NUMBER OF DATA PAIRS PER SET & NUMBER OF SETS.   *
C     *       FORMAT(I3,I3)                                            *
C     *                                                                *
C     *                                                                *
C     *                DARRYL CRAYER, KANSAS STATE UNIVERSITY 4/81     *
C     *                                                                *
C     ******************************************************************
C     PROGRAM IS IN DOUBLE PRECISION
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION HARRAY(99),YARRAY(99),NF2(10000),NT2(10000)
      INTEGER F,COUNT,FARRAY,COUNT2
      REAL*8 ALP(3),BET(3),ALIKE(3,10000)
      REAL*8 ARATIO(3,10000)
```

```
      COMMON /DTA/NF(10000),NT(1000C),NN(10000),N,NDATA,IPRT,NITER,J
      EXTERNAL FN
C *****
C *****
C     GENERATE SECTICN
C *****
C *****
C     LET THE STARTING VALUE FCR RANDU ARBITRARILY EQUAL 12345
      IX=12345
C
C     READ IN AND ESTABLISH CUMULATIVE CISTRIBUTICN FOR ALPHA1 & BETA1
C     READ IN VALUES FOR TIME(=T) AND THE GAMMA FUNCTICN
C     PARAMETERS (ALPHA AND BETA).
C
      READ 100,T1,ALPHA,BETA
  100 FORMAT(E13.5,7X,E13.5,7X,E13.5)
      WRITE(6,101)T1,ALPHA,BETA
  101 FORMAT('1','TIME=',E13.5,'HRS.',5X,'ALPHA=',E13.5,3X,'BETA=',E13.5
     /)
      READ 102,T2,ALPHA2,BETA2
  102 FORMAT(E13.5,7X,E13.5,7X,E13.5)
C
C     READ NUMBER OF DATA PAIRS PER SET
C
      READ 103,N
  103 FORMAT(I6)
C
C     READ N2, THE NUMBER OF SETS DESIRED
C
      READ 104,N2
  104 FORMAT(I3)
C     READ IN NUMBER OF POINTS NNN, & NUMBER OF SETS, N2.
      READ 105,NNN,N2
  105   FORMAT(I3,I3)
      TAU=1.0/BETA
      COUNT=0
      OLD=0
      T=T1
      DO 106 I=1,200
      COUNT=COUNT+1
      F=I-1
      TOP=(DGAMMA(F+ALPHA))*(T**F)*(TAU**ALPHA)
      BOT=(DGAMMA(ALPHA))*(DGAMMA(F+1.0D+00))*((T+TAU)**(F+ALPHA))
      CLD=CLD+(TOP/BOT)
      HARRAY(I)=CLD
      IF(HARRAY(I).GT.0.99999)GOTO 109
  106 CONTINUE
      WRITE(6,107)
  107 FORMAT(' ','NOT ENOUGH TIMES THROUGH LOOP TO GET CUM.GT.0.99999')
      GOTO 2000
C
C     WRITE CUMULATIVE DISTRIBUTION 1
C
  108 CONTINUE
  109 WRITE(6,110)
  110 FORMAT('0','CUMULATIVE DISTRIBUTION')
      WRITE(6,111)
  111 FORMAT(' ','F=',6X,'CUMULATIVE PROBABILITY')
      DO 113 I=1,COUNT
```

```
      F=I-1
      WRITE(6,112)F,HARRAY(I)
  112 FORMAT(' ',I3,5X,E13.5)
  113 CONTINUE
C
C     READ IN AND ESTABLISH CUMULATIVE DISTRIBUTION FOR ALPHA2 & BETA2
C     READ IN VALUES FOR TIME(=T) AND THE GAMMA FUNCTION
C     PARAMETERS (ALPHA AND BETA
C
      WRITE(6,114)T2,ALPHA2,BETA2
  114 FORMAT('-','TIME2=',E13.5,'HRS.',5X,'ALPHA2=',E13.5,3X,'BETA2=',E1
     /3.5)
      TAU2=1.0/BETA2
      COUNT2=0
      CLD=0
      DO 115 I=1,200
      COUNT2=COUNT2+1
      F=I-1
      TOP=DLGAMA(F+ALPHA2)+F*DLOG(T2)+ALPHA2*DLOG(TAU2)
      BOT=DLGAMA(ALPHA2)+DLGAMA(F+1.0D+00)+(F+ALPHA2)*DLOG(T2+TAU2)
      CLD=CLD+DEXP(TOP-BOT)
      YARRAY(I)=CLD
      IF(YARRAY(I).GT.0.99999)GOTO 117
  115 CONTINUE
      WRITE(6,116)
  116 FORMAT(' ','NOT ENOUGH TIME THROUGH LOOP TO GET CUM.GT.0.99999')
      GOTO 2000
C
C     WRITE CUMULATIVE DISTRIBUTION 2
C
  117 WRITE(6,118)
  118 FORMAT('0','CUMULATIVE DISTRIBUTION')
      WRITE(6,119)
  119 FORMAT(' ','F=',6X,'CUMULATIVE PROBABILITY')
      DO 121 I=1,COUNT2
      F=I-1
      WRITE(6,120)F,YARRAY(I)
  120 FORMAT(' ',I3,5X,E13.5)
  121 CONTINUE
C
C
C     SECTION TO GENERATE FREQUENCIES
C
C
C
      J1=0
      DO 131 M=1,N2
C
C     DISTRIBUTION NUMBER 1
C
C     N IS THE NUMBER OF DATA POINTS DESIRED
      DO 125 J=1,NNN
      J1=J1+1
C     GENERATE RANDOM NUMBERS
  122 CALL RANDU(IX,IY,U)
      IX=IY
C     SINCE CUMULATIVE ONLY GOES UP TO 0.99999, MUST CHECK TO SEE THAT
C     THE R.N. IS.LE.0.99999.
      IF(U.GT.0.99999)GOTO 122
```

```
C     FIND WHAT F CORRESPONDS TO R.N.
      DO 123 K=1,COUNT
      UPPER=HARRAY(K)
      IF(U.LT.UPPER)GOTO 124
  123 CONTINUE
  124 NF(J1)=K-1
      NT(J1)=T1
      NN(J1)=1
  125 CONTINUE
C
C     DISTRIBUTION NUMBER 2
C
      DO 130 J=1,NNN
      J1=J1+1
C     GENERATE RANDOM NUMBERS
  126 CALL RANDU(IX,IY,U)
      IX=IY
C     SINCE CUMULATIVE ONLY GOES UP TO 0.99999, MUST CHECK TO SEE THAT
C     THE R.N. IS.LE.0.99999
      IF(U.GT.0.99999)GOTO 126
C     FIND WHAT F CORRESPONDS TO R.N.
      DO 127 K=1,COUNT2
      UPPER=YARRAY(K)
      IF(U.LT.UPPER)GOTO 128
  127 CONTINUE
  128 NF2(J1)=K-1
      NT2(J1)=T2
      NF(J1)=K-1
      NT(J1)=T2
      NN(J1)=1
  129 CONTINUE
  130 CONTINUE
  131 CONTINUE
C *****
C *****
C     GAMMA SECTION
C *****
C *****
C
C
C
C  SET PRINTER UNIT=6; SET READER UNIT=5; REL. ACC OF PERCENTILES
C  IS SET BY EPPS; MAX NUMBER OF ITERATIONS FOR EVALUATING PERCENTILE
C  SET BY VARIABLE NITER
      IPRT = 6
      IRDR = 5
      NITER=20
      L=2*NNN*N2
      J=1
      N2CONT=0
  200 NCOUNT=0
      NOLDJ=J
  201 NDATA=J+NNN-1
  202 CONTINUE
C
C
      PCT=0.1
      BMIN=0.0
      BMAX=0.0
```

```
      EPS=0.01
C  CALCULATE TCTAL NUMBER CF CCMPCNENTS N
      N = 0
      CO 203 I=J,NCATA
      IF (NN(I).LE.0) NN(I) = 1
  203 N=N+NN(I)
C
C  CALCULATE MEAN AND VARIANCE OF THE DATA
      SUMT = 0.0
      SUMF = C.0
      SUMFT = 0.0
      SUMFT2 = 0.0
      SUMTI = 0.0
      DO 204 I=J,NCATA
      TT = NT(I)
      SUMT = SUMT+TT*NN(I)
      SUMTI = SUMTI+NN(I)/TT
      SUMF = SUMF+NF(I)*NN(I)
      AA = NF(I)/TT
      SUMFT = SUMFT+AA*NN(I)
  204 SUMFT2 = SUMFT2+AA*NN(I)*AA
      IF(SUMFT.EQ.0.D0)GCTC 205
      GOTO 206
  205 ALP(1)=0.0
      BET(1)=C.0
      ALP(2)=0.0
      BET(2)=C.0
      ALP(3)=C.0
      BET(3)=C.0
      GOTO 211
  206 UMEAN = SUMFT/N
      UVAR = (SUMFT2-UMEAN*UMEAN*N)/(N-1)
C
C
C
C***    BEGIN ANALYSIS FOR CCMPCUNC MODEL;
C
C  CALCULATE MATCHING MOMENTS ESTIMATORS
      BETAUP = UVAR/UMEAN
      ALPHUP = UMEAN/BETAUP
      BETAUM = (UVAR-UMEAN*SUMTI/N)/UMEAN
      IF(BETAUM.EQ.0.D0)GCTC 207
      ALPHUM = UMEAN/BETAUM
      GOTC 208
  207 ALPHUM=C.0
C
C***   UNWEIGHTED MATCHING MCMENTS ESTIMATORS TC THE PRIOR - METHOD 1
  208 AAA=ALPHUP
      BBB = BETAUP
      ALP(1)=AAA
      BET(1)=BBB
C
C***   UNWEIGHTED MATCHING MCMENTS TC THE MARGINAL - METHCD 2
  209 AAA = ALPHUM
      BBB = BETAUM
      ALP(2)=AAA
      BET(2)=BBB
C
C***   MAXIMUM LIKELIHOCD ESTIMATCRS - METHCC 4
```

```
      IF (BMIN.EQ.0.0) BMIN = BETALP*0.01
      IF (BMAX.EQ.0.0) BMAX = 100.0*BETALP
      CALL DRTMI (BETA,FF,FN,BMIN,BMAX,EPS,NITER,IER)
      IF (IER.NE.0)GOTO 210
      CALL MLGAMP (FFF,ALPHA,BETA)
      AAA = ALPHA
      BBB = BETA
      ALP(3)=AAA
      BET(3)=BBB
      GOTO 211
  210  CONTINUE
      ALP(3)=0.0
      BET(3)=0.0
C
      GOTO 211
C
C
  211  NCOUNT=NCOUNT+1
C     FIND LIKELIHOOD FOR EACH GROUP
      DO 218 I=1,3
      IF(BET(I).LE.0.D0)GOTO 213
      BET(I)=1.D0/BET(I)
      FIRST=0.D0
      DO 212 I2=J,NDATA
      AT=NT(I2)
      OOPS=NF(I2)+ALP(I)
      IF(ALP(I).LE.0.D0)GOTO 213
      IF(BET(I).LE.0.D0)GOTO 213
      IF(OOPS.LT.0.D0)GOTO 213
      X1=NF(I2)+ALP(I)
      X2=NF(I2)
      X3=NT(I2)
      X4=ALP(I)
      X5=BET(I)
      X6=NF(I2)+1.D0
      X7=NT(I2)+BET(I)
      TOP=DLGAMA(X1)+X2*DLOG(X3)+X4*DLOG(X5)
      BOT=DLGAMA(X4)+DLGAMA(X6)+X1*DLOG(X7)
      DIF=TOP-BOT
      ALIKE(NCOUNT,I)=DIF+FIRST
C
C
      FIRST=ALIKE(NCOUNT,I)
  212 CONTINUE
      GOTO 218
  213 ALIKE(NCOUNT,I)=0
  218 CONTINUE
      IF(NCOUNT.EQ.2)GOTO 219
      IF(NCOUNT.EQ.3)GOTO 220
      J=J+NNN
      GOTO 201
  219 J=NOLDJ
      NDATA=(2*NNN)+J-1
      GOTO 202
  220 CONTINUE
      N2CONT=N2CONT+1
      DO 223 I=1,3
      DO 221 I30=1,3
      IF(ALIKE(I30,I).EQ.0.0)GOTO 222
```

```
  221 CONTINUE
      ARATIO(I,N2CONT)=ALIKE(3,I)-ALIKE(1,I)-ALIKE(2,I)
      GOTO 223
  222 CONTINUE
      ARATIO(I,N2CONT)=0.0
  223 CONTINUE
      J=J+2*NNN
      IF(N2CONT.EQ.N2)GOTO 224
      GOTO 200
  224 CONTINUE
C
C
C
      DO 229 I=1,3
      NREJET=0
      N3CONT=0
      DO 225 J=1,N2
      IF(ARATIO(I,J).EQ.0.0)GOTO 225
      N3CONT=N3CONT+1
      T=-2.0D0*ARATIO(I,J)
C     THIS IS WHERE THE SIGNIFICANCE LEVEL COMES IN
      CHI=5.991D0
      IF(CHI.GT.T)GOTO 225
      NREJET=NREJET+1
  225 CONTINUE
      IF(NREJET.EQ.0)GOTO 226
      IF(N3CONT.EQ.0)GOTO 226
      X10=NREJET
      X11=N3CONT
      POWER=X10/X11
      GOTO 227
  226 POWER=0.0
  227 CONTINUE
      WRITE(6,228)POWER,N3CONT,I,NREJET
  228  FORMAT('0','POWER=',D13.5,5X,'NUMBER OF PTS. USED=',I6,3X,'ESTIMA
     /ION METHOD NO.',I3,I3)
  229 CONTINUE
      BETA2=BETA2+0.00001
      IF(BETA2.LE.0.00005)GOTO 108
 2000 STOP
      END


       REAL FUNCTION FN*8(X)
C*********************************************************************
C*                                                                   *
C* THIS FUNCTION EVALUATES THE AUXILARY LIKELIHOOD FUNCTION FN WHOSE *
C* ZERO GIVES THE VALUE OF THE BETA PARAMETER.  THIS SUBROUTINE SIMPLY *
C* CALLS MLGAMP WHERE THE ACTUAL EVALUATION IS PERFORMED.           *
C*                                                                   *
C*********************************************************************
       REAL*8 ALPHA,X
       CALL MLGAMP (FN,ALPHA,X)
       RETURN
       END
```

```
      SUBROUTINE MLGAMP (FN,ALPHA,BETA)
C***********************************************************************
C*                                                                     *
C* THIS SUBROUTINE EVALUATES ALPHA AND THE FUNCTION FN WHOSE ZERO      *
C* GIVES THE VALUE OF BETA FOR THE MAXIMUM LIKELIHOOD ESTIMATORS.      *
C*                                                                     *
C***********************************************************************
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON /DTA/NF(10000),NT(10000),NN(10000),N,NDATA,IPRT,NITER,J2
C
C     CALCULATE ALPHA AS A FUNCTION OF BETA
      SUM1 = 0.
      SUM2 = 0.
      FF = 0.
      DO 1 J=J2,NDATA
      FF = FF+NF(J)*NN(J)
      A = NN(J)*NT(J)/(1.+BETA*NT(J))
      SUM1 = SUM1+A
    1 SUM2 = SUM2+A*NF(J)
      ALPHA = (FF/BETA-SUM2)/SUM1
C
C     CALCULATE F(ALPHA,BETA)
      FN = 0.
      DO 4 J=J2,NDATA
      A = DLOG(1.+BETA*NT(J))
      SUM = 0.
      NNN = NF(J)
      IF (NNN.EQ.0) GO TO 3
      DO 2 K=1,NNN
    2 SUM = SUM+1./(ALPHA+K-1.)
    3 FN = FN+(A-SUM)*NN(J)
    4 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE DRTMI (X,F,FCT,XLI,XRI,EPS,IEND,IER,NCCCPS)
C******************************** DRTMI *********************************
C*                                                                      *
C* THIS SUBROUTINE SOLVES A NONLINEAR EQUATION OF THE FORM FCT(X)=0     *
C* BY THE MUELLER'S ITERATION METHOD.                                   *
C* THIS PROGRAM IS TAKEN FROM THE IBM SSP (RTMI) AND MODIFIED TO WORK   *
C* IN DOUBLE PRECISION.                                                 *
C*                                                                      *
C*    INPUT VARIABLES:                                                  *
C*        X     = RETURNED VALUE OF THE ZERO                            *
C*        F     = VALUE OF THE FUNCTION FCT AT THE ZERO                 *
C*        FCT   = EXTERNAL FUNCTION WHOSE ZERO IS TO BE FOUND           *
C*        XLI   = LEFT BOUNDARY OF THE X-AXIS TO BE SEARCHED FOR ROOT   *
C*        XRI   = RIGHT BOUNDARY OF X-AXIS TO BE SEARCHED FOR ROOT      *
C*        EPS   = ACCURACY OF DESIRED RESULT                            *
C*        IEND = MAXIMUM NUMBER OF ITERATIONS TO BE USED                *
C*        IER   = ERROR RETURN CODE; =0 IF DESIRED ACCURACY ACHIEVED    *
C*                                                                      *
C************************************************************************
      IMPLICIT REAL*8(A-H,O-Z)
C
C
C     PREPARE ITERATION
      IER = 0
      XL = XLI
      XR = XRI
      X = XL
      TOL = X
      F = FCT(TOL)
      IF (F) 1,16,1
    1 FL = F
      X = XR
      TOL = X
      F = FCT(TOL)
      IF (F) 2,16,2
    2 FR = F
      IF (DSIGN(1.D0,FL)+DSIGN(1.D0,FR)) 25,3,25
C
C     BASIC ASSUMPTION FL*FR LESS THAN 0 IS SATISFIED.
C     GENERATE TOLERANCE FOR FUNCTION VALUES.
    3 I = 0
      TOLF = 100.*EPS
C
C
C     START ITERATION LOOP
    4 I = I+1
C
C     START BISECTION LOOP
      DO 13 K=1,IEND
      X = .5*(XL+XR)
      TOL = X
      F = FCT(TOL)
      IF (F) 5,16,5
    5 IF (DSIGN(1.D0,F)+DSIGN(1.D0,FR)) 7,6,7
C
```

```
C      INTERCHANGE XL AND XR IN ORDER TO GET THE SAME SIGN IN F AND FR
     6 TOL = XL
       XL = XR
       XR = TOL
       TOL = FL
       FL = FR
       FR = TOL
     7 TOL = F-FL
       A = F*TOL
       A = A+A
       IF (A-FR*(FR-FL)) 8,9,9
     8 IF (I-IEND) 17,17,9
     9 XR = X
       FR = F
C
C      TEST ON SATISFACTORY ACCURACY IN BISECTION LOOP
       TOL = EPS
       A = DABS(XR)
       IF (A-1.) 11,11,10
    10 TOL = TOL*A
    11 IF (DABS(XR-XL)-TOL) 12,12,13
    12 IF (DABS(FR-FL)-TOLF) 14,14,13
    13 CONTINUE
C      END OF BISECTION LOOP
C
C      NO CONVERGENCE AFTER IEND ITERATION STEPS FOLLOWED BY IEND
C      SUCCESSIVE STEPS OF BISECTION OR STEADILY INCREASING FUNCTION
C      VALUES AT RIGHT BOUNDS. ERROR RETURN.
       IER = 1
    14 IF (DABS(FR)-DABS(FL)) 16,16,15
    15 X = XL
       F = FL
    16 RETURN
C
C      COMPUTATION OF ITERATED X-VALUE BY INVERSE PARABOLIC INTERPOLATION
    17 A = FR-F
       DX = (X-XL)*FL*(1.+F*(A-TOL)/(A*(FR-FL)))/TOL
       XM = X
       FM = F
       X = XL-DX
       TOL = X
       F = FCT(TOL)
       IF (F) 18,16,18
C
C      TEST ON SATISFACTORY ACCURACY IN ITERATION LOOP
    18 TOL = EPS
       A = DABS(X)
       IF (A-1.) 20,20,19
    19 TOL = TOL*A
    20 IF (DABS(DX)-TOL) 21,21,22
    21 IF (DABS(F)-TOLF) 16,16,22
C
C      PREPARATION OF NEXT BISECTION LOOP
    22 IF (DSIGN(1.D0,F)+DSIGN(1.D0,FL)) 24,23,24
    23 XR = X
       FR = F
       GO TO 4
    24 XL = X
       FL = F
```

```
      XR = XM
      FR = FM
      GO TO 4
C     END CF ITERATICN LOOP
C
C
C     ERRCR RETURN IN CASE CF WRCNG INPUT CATA
   25 IER = 2
      RETURN
      END
```

Computer Code Listing for

GAMMAP GENERATE

```
C      ****************************************************************
C      *                                                              *
C      *  GAMMAP GENERATE                                             *
C      *                                                              *
C      *                                                              *
C      *  GENERATE                                                    *
C      *  GIVEN TWO SETS OF PARAMETERS (TIME, ALPHA, AND BETA) THIS   *
C      *  SECTION GENERATES THE CUMULATIVE MARGINAL DISTRIBUTION      *
C      *  FOR EACH SET.  FROM THE CUMULATIVE DISTRIBUTION THE PROGRAM *
C      *  SIMULATES FAILURE DATA (NUMBER OF FAILURES PER A GIVEN TIME). *
C      *  THE NUMBER OF DESIRED SETS OF FAILURE DATA MAY BE SPECIFIED *
C      *  ALONG WITH THE NUMBER OF DATA PAIRS PER SET.                *
C      *                                                              *
C      *  GAMMA                                                       *
C      *  THE SECOND SECTION OF THE PROGRAM IS A MODIFICATION OF THE  *
C      *  PROGRAM GAMMA8 WRITTEN BY J.K. SHULTIS IN JUNE OF 1980.  THIS *
C      *  SECTION TAKES SET OF FAILURE RATE DATA FROM GENERATE,       *
C      *  ESTIMATES PRIOR PARAMETERS, FINDS THE LIKELIHOOD RATIO BETWEEN*
C      *  SETS OF DATA, AND CALCULATES POWER.  THE STATISTICAL MODEL  *
C      *  ASSUMED IS THE COMPOUND POISSON-GAMMA MODEL IN WHICH THE    *
C      *  FAILURE RATES FOR EACH COMPONENT MAY VARY ACCORDING TO A    *
C      *  GAMMA PRIOR DISTRIBUTION.  THE PARAMETERS OF THE GAMMA PRIOR *
C      *  ARE ESTIMATED BY MATCHING UNWEIGHTED DATA MOMENTS TO THE    *
C      *  PRIOR.                                                      *
C      *                                                              *
C      *                                                              *
C      *     INPUT CARDS                                              *
C      *                                                              *
C      *       CARD 1  FIRST DISTRIBUTION PARAMETERS/ TIME, ALPHA, BETA.*
C      *       FORMAT(E13.5,7X,E13.5,7X,E13.5)                        *
C      *                                                              *
C      *       CARD 2  SECOND DISTRIBUTION PARAMETERS  TIME, ALPHA,   *
C      *       BETA.                                                  *
C      *       FORMAT(E13.5,7X,E13.5,7X,E13.5)                        *
C      *                                                              *
C      *       CARD 3  NUMBER OF FAILURE DATA PAIRS PER SET.          *
C      *       FORMAT(I6)                                             *
C      *                                                              *
C      *       CARD 4  NUMBER OF SETS.                                *
C      *       FORMAT(I6)                                             *
C      *                                                              *
C      *       CARD 5  NUMBER OF DATA PAIRS PER SET & NUMBER OF SETS. *
C      *       FORMAT(I3,I3)                                          *
C      *                                                              *
C      *                                                              *
C      *               DARRYL CRAYER, KANSAS STATE UNIVERSITY 4/81    *
C      *                                                              *
C      ****************************************************************
C      ****************************************************************
C      PROGRAM IS IN DOUBLE PRECISION
       IMPLICIT REAL*8(A-H,O-Z)
       DIMENSION HARRAY(99),YARRAY(99),NF2(10000),NT2(10000),NF(10000),
      /NT(10000),NN(10000)
       INTEGER F,COUNT,FARRAY,COUNT2
       REAL*8 ALP(1),BET(1),ALIKE(1,10000)
```

```
      REAL*8 ARATIC(1,10000)
C
C
C
C  SET PRINTER UNIT=6; SET READER UNIT=5; REL. ACC CF PERCENTILES
C  IS SET BY EPPS; MAX NUMBER CF ITERATICNS FCR EVALUATING PERCENTILE
C  SET BY VARIABLE NITER
      IPRT = 6
      IRCR = 5
      NITER = 20
C********
C********
C     GENERATE SECTION
C********
C********
C     LET THE STARTING VALUE FCR RANDU ARBITRARILY EQUAL 12345
      IX=12345
C
C     READ IN AND ESTABLISH CUMMULATIVE DISTRIBUTION FCR ALPHA1 & BETA1
C     READ IN VALUES FOR TIME(=T) AND THE GAMMA FUNCTION
C     PARAMETERS (ALPHA AND BETA).
C
      READ 100,T1,ALPHA,BETA
  100 FORMAT(E13.5,7X,E13.5,7X,E13.5)
      WRITE(6,101)T1,ALPHA,BETA
  101 FORMAT('1','TIME=',E13.5,'HRS.',5X,'ALPHA=',E13.5,3X,'BETA=',E13.5
     /)
      READ 102,T2,ALPHA2,BETA2
  102 FORMAT(E13.5,7X,E13.5,7X,E13.5)
C
C     READ NUMBER CF DATA PAIRS PER SET
C
      READ 103,N
  103 FORMAT(I6)
C
C     READ N2, THE NUMBER CF SETS DESIRED
C
      READ 104,N2
  104 FORMAT(I3)
C     READ IN NUMBER OF POINTS NNN, & NUMBER CF SETS, N2.
      READ 105,NNN,N2
  105 FORMAT(I3,I3)
      TAU=1.0/BETA
      COUNT=0
      CLD=0
      T=T1
      DO 106 I=1,200
      COUNT=COUNT+1
      F=I-1
      TOP=(DGAMMA(F+ALPHA))*(T**F)*(TAU**ALPHA)
      BOT=(DGAMMA(ALPHA))*(DGAMMA(F+1.0D+00))*((T+TAU)**(F+ALPHA))
      CLD=CLD+(TOP/BOT)
      HARRAY(I)=CLD
      IF(HARRAY(I).GT.0.99999)GOTO 109
  106 CONTINUE
      WRITE(6,107)
  107 FORMAT(' ','NOT ENOUGH TIMES THROUGH LOOP TO GET CUM.GT.0.99999')
      GOTO 2000
C
```

```
C      WRITE CUMMULATIVE DISTRIBUTICN 1
C
  108 CONTINUE
  109 WRITE(6,110)
  110 FORMAT('0','CUMULATIVE DISTRIBUTION')
      WRITE(6,111)
  111 FORMAT(' ','F=',6X,'CUMULATIVE PROBABILITY')
      DO 113 I=1,COUNT
      F=I-1
      WRITE(6,112)F,HARRAY(I)
  112 FORMAT(' ',I3,5X,E13.5)
  113 CONTINUE
C
C      ESTABLISH CUMULATIVE CISTRIBUTION FOR ALPHA2 & BETA2
C
      WRITE(6,114)T2,ALPHA2,BETA2
  114 FORMAT('-','TIME2=',E13.5,'HRS.',5X,'ALPHA2=',E13.5,3X,'BETA2=',E1
     /3.5)
      TAU2=1.C/BETA2
      COUNT2=0
      CLD=0
      CO 115 I=1,200
      COUNT2=CCUNT2+1
      F=I-1
      TOP=DLGAMA(F+ALPHA2)+F*DLCG(T2)+ALPHA2*CLOG(TAU2)
      BOT=DLGAMA(ALPHA2)+DLGAMA(F+1.00+00)+(F+ALPHA2)*CLCG(T2+TAU2)
      CLC=CLC+CEXP(TCP-BOT)
      YARRAY(I)=OLD
      IF(YARRAY(I).GT.0.99999)GCTC 117
  115 CONTINUE
      WRITE(6,116)
  116 FORMAT(' ','NOT ENOUGH TIME THROUGH LCCP TC GET CUM.GT.0.99999')
      GOTO 2CCC
C
C      WRITE CUMMULATIVE DISTRIBUTICN 2
C
  117 WRITE(6,118)
  118 FORMAT('0','CUMMULATIVE DISTRIBUTION')
      WRITE(6,119)
  119 FORMAT(' ','F=',6X,'CUMULATIVE PROBABILITY')
      DO 121 I=1,CCUNT2
      F=I-1
      WRITE(6,120)F,YARRAY(I)
  120 FORMAT(' ',I3,5X,E13.5)
  121 CONTINUE
C
C
C      SECTICN TO GENERATE FREQUENCIES
C
C
C
      J1=0
      CO 1000 M=1,N2
C
C      DISTRIBUTICN NUMBER 1
C
C      N IS THE NUMBER OF DATA PCINTS DESIREC
      CO 125 J=1,NNN
      J1=J1+1
```

```
C      GENERATE RANCCM NUMBERS
  122 CALL RANCU(IX,IY,U)
      IX=IY
C      SINCE CUMULATIVE ONLY GCES UP TO 0.99999, MUST CHECK TO SEE THAT
C      THE R.N. IS.LE.0.99999.
      IF(U.GT.C.99999)GCTO 122
C      FIND WHAT F CORRESPCNDS TC R.N.
      DO 123 K=1,COUNT
      UPPER=HARRAY(K)
      IF(U.LT.UPPER)GOTO 124
  123 CONTINUE
  124 NF(J1)=K-1
      NT(J1)=T1
      NN(J1)=1
  125 CCNTINUE
C
C      CISTRIBUTICN NUMBER 2
C
      DO 130 J=1,NNN
      J1=J1+1
C      GENERATE RANDCM NUMBERS
  126 CALL RANCU(IX,IY,U)
      IX=IY
C      SINCE CUMULATIVE ONLY GOES UP TO C.99999, MUST CHECK TC SEE THAT
C      THE R.N. IS.LE.0.99999
      IF(U.GT.C.99999)GCTO 126
C      FIND WHAT F CCRRESPCNCS TO R.N.
      DO 127 K=1,COUNT2
      UPPER=YARRAY(K)
      IF(U.LT.UPPER)GCTC 128
  127 CONTINUE
  128 NF2(J1)=K-1
      NT2(J1)=T2
      NF(J1)=K-1
      NT(J1)=T2
      NN(J1)=1
  129 CCNTINUE
  130 CONTINUE
 1C00 CONTINUE
C********
C********
C      GAMMAP SECTICN
C********
C********
      L=2*NNN*N2
      J=1
      N2CCNT=0
  200 NCCUNT=0
      NCLDJ=J
  201 NDATA=J+NNN-1
  202 CONTINUE
C
C
      PCT=0.1
      BMIN=0.0
      BMAX=0.0
      EPS=0.0001
C  CALCULATE TCTAL NUMBER CF COMPONENTS N
      N = 0
```

```
      DO 203 I=J,NDATA
      IF (NN(I).LE.0) NN(I) = 1
  203 N=N+NN(I)
C
C  CALCULATE MEAN AND VARIANCE OF THE DATA
      SUMT = 0.0
      SUMF = 0.0
      SUMFT = 0.0
      SUMFT2 = 0.0
      SUMTI = 0.0
      DO 204 I=J,NDATA
      TT = NT(I)
      SUMT = SUMT+TT*NN(I)
      SUMTI = SUMTI+NN(I)/TT
      SUMF = SUMF+NF(I)*NN(I)
      AA = NF(I)/TT
      SUMFT = SUMFT+AA*NN(I)
  204 SUMFT2 = SUMFT2+AA*NN(I)*AA
      IF(SUMFT.EQ.0.D0)GOTO 205
      GOTO 206
  205 ALP(1)=0.0
      BET(1)=0.0
      GOTO 208
  206 UMEAN = SUMFT/N
      UVAR = (SUMFT2-UMEAN*UMEAN*N)/(N-1)
C
C
C
C***     BEGIN ANALYSIS FOR COMPOUND MODEL;
C
C  CALCULATE MATCHING MOMENTS ESTIMATORS
      BETAUP = UVAR/UMEAN
      ALPHUP = UMEAN/BETAUP
C
C***  UNWEIGHTED MATCHING MOMENTS ESTIMATORS TO THE PRIOR - METHOD 1
  207 AAA=ALPHUP
      BBB = BETAUP
      ALP(1)=AAA
      BET(1)=BBB
C
C***  UNWEIGHTED MATCHING MOMENTS TO THE MARGINAL - METHOD 2
  208 NCOUNT=NCOUNT+1
C     FIND LIKELIHOOD FOR EACH GROUP
      I=1
      IF(BET(I).LE.0.D0)GOTO 210
      BET(I)=1.D0/BET(I)
      FIRST=0.D0
      DO 209 I2=J,NDATA
      AT=NT(I2)
      OOPS=NF(I2)+ALP(I)
      IF(ALP(I).LE.0.D0)GOTO 210
      IF(BET(I).LE.0.D0)GOTO 210
      IF(OOPS.LT.0.D0)GOTO 210
      X1=NF(I2)+ALP(I)
      X2=NF(I2)
      X3=NT(I2)
      X4=ALP(I)
      X5=BET(I)
      X6=NF(I2)+1.D0
```

```
      X7=NT(I2)+BET(I)
      TOP=DLGAMA(X1)+X2*DLOG(X3)+X4*DLOG(X5)
      BOT=DLGAMA(X4)+DLGAMA(X6)+X1*DLOG(X7)
      DIF=TOP-BOT
      ALIKE(NCOUNT,I)=DIF+FIRST
C
C
      FIRST=ALIKE(NCOUNT,I)
  209 CONTINUE
      GOTO 211
  210 ALIKE(NCOUNT,I)=0
  211 CONTINUE
      IF(NCOUNT.EQ.2)GOTO 212
      IF(NCOUNT.EQ.3)GOTO 213
      J=J+NNN
      GOTO 201
  212 J=NOLDJ
      NDATA=(2*NNN)+J-1
      GOTO 202
  213 CONTINUE
      N2CONT=N2CONT+1
      DO 214 I30=1,3
      IF(ALIKE(I30,I).EQ.0.0)GOTO 215
  214 CONTINUE
      ARATIO(I,N2CONT)=ALIKE(3,I)-ALIKE(1,I)-ALIKE(2,I)
      GOTO 216
  215 CONTINUE
      ARATIO(I,N2CONT)=0.0
  216 CONTINUE
      J=J+2*NNN
      IF(N2CONT.EQ.N2)GOTO 217
      GOTO 200
  217 CONTINUE
C
C
C
      NREJET=0
      N3CONT=0
      DO 218 J=1,N2
      IF(ARATIO(I,J).EQ.0.0)GOTO 218
      N3CONT=N3CONT+1
      T=-2.0D0*ARATIO(I,J)
C     THIS IS WHERE THE SIGNIFICANCE LEVEL COMES IN
      CHI=5.991D0
      IF(CHI.GT.T)GOTO 218
      NREJET=NREJET+1
  218 CONTINUE
      IF(NREJET.EQ.0)GOTO 219
      IF(N3CONT.EQ.0)GOTO 219
      X10=NREJET
      X11=N3CONT
      POWER=X10/X11
      GOTO 220
  219 POWER=0.0
  220 CONTINUE
      WRITE(6,221)POWER,N3CONT,I
  221 FORMAT('0','POWER=',D13.5,5X,'NUMBER OF PTS. USED=',I5,3X,'ESTIMAT
     /ION METHOD NO.',I3)
  222 CONTINUE
```

```
      BETA2=BETA2+0.00001
      IF(BETA2.LE.0.00033)GOTO 108
 2000 STOP
      END
```

THE DISTINCTION OF SIMULATED

FAILURE DATA BY THE

LIKELIHOOD RATIO TEST

by

DARRYL D. DRAYER

B.S., Kansas State University, 1980

AN ABSTRACT OF

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

1981

ABSTRACT

The objective of this work was to use the likelihood ratio test to answer the question: how close together (or in shape) can failure rate distributions be to one another before they cannot be distinguished from one another.

To answer this question a Bayesian (or compound model) analysis approach was used, i.e., a marginal probability distribution of number of failures given test time was formed integrating the product of a gamma failure rate distribution and a Poisson likelihood distribution over all figure rate space (0 to ∞). Faiure data sets of sample size ten were generated by the inverse transformation method, i.e., a random number from a uniform distribution was used as the marginal distribution cumulate with known parameters from which number of failures was extracted. The various data sets were used to form likelihood ratio tests from which power curves, i.e., percentage of time that data sets from different failure rate distributions were correctly identified as coming from different distributions versus a measure (one of the parameters of the gamma distribution) of the difference in the distributions, were formed. The power curves were characterized by the full width at one half, since the power curve should vary from unity (100% correct detection) to zero (100% incorrect detection), to form conclusions about the simulated data, from which an answer to the objective question can be formulated.

For the matching moments to the prior distribution method (MMPM), approximately 96% of the 250 data sets were useful, i.e., only about 4% of the data sets had no fialure, hence, failed to yield prior distribution estimators, the only time for which the MMPM compound model analysis fails. For the marginal maximum likelihood method (MMLM), only about 16% of the

data set yielded useful data, i.e., for about 84% of the data sets MMLM failed to yield estimators because of iteration problems. Of the 250 data sets (sample size 10) only about 25% yielded acceptable (greater than zero) estimators by the marginal matching moments method (MMMM).

The power curves were all shaped like a negative pulse with the minimum power value generally occurring at BETA2 equal to 0.00005 and ranging in value from 0.007 to 0.012. The MMMM and MMLM had lower FW0.5 values than the MMPM. Hence, these two methods should be capable of correctly rejecting the null hypothesis (discerning between data sets from different gamma distributions) for gamma distributions closer together (or more similar in shape) more often than the MMPM. This ability notwithstanding, the MMPM for estimating parameters of the gamma distribution for modeling failure rate data was judged best because of the percentage of data sets for which allowable estimators were found and the cost of analyzing data by the MMPM is considerably less than the other two methods.