INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Robust procedures in survival analysis and reliability

Zhou, Xiao Hua, Ph.D.
The Ohio State University, 1991
ROBUST PROCEDURES IN SURVIVAL ANALYSIS AND RELIABILITY

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By
Xiao Hua Zhou, B.Sc., M.Sc.

The Ohio State University
1991

Dissertation Committee:
Professor Sue Leurgans
Professor Saul Blumenthal
Professor John P. Klein

Approved By
Sue Leurgans
Advisor

Biostatistics
Graduate Program
To my parents
ACKNOWLEDGMENT

I would like to thank my advisors, Professor Leurgans and Professor Blumenthal, for their guidance, encouragement and support throughout the research for this dissertation.
VITA

March 15, 1963 .......... Born- Sichuan, People's Republic of China

1984 ......................... B.Sc., Sichuan University
                        Sichuan, People's Republic of China

1984-1985 ..................... Instructor, Sichuan University
                        Sichuan, People's Republic of China

1987 .......................... M.Sc., The University of Calgary
                        Calgary, Canada

1987-1988 ..................... University Fellow, Ohio State University

1988-present .................. Graduate Associate, Ohio State University

FIELDS OF STUDY

Major Field: Biostatistics
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>VITA</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>CHAPTER 1. INTRODUCTION</td>
<td></td>
</tr>
<tr>
<td>1. Overview of Robust Procedures</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Effect of Censoring on Statistical Procedures</td>
<td>9</td>
</tr>
<tr>
<td>1.3 Minimax Tests</td>
<td>11</td>
</tr>
<tr>
<td>1.4 Reliability Tests for Large Series Systems</td>
<td>13</td>
</tr>
<tr>
<td>1.5 Contents of Chapters</td>
<td>22</td>
</tr>
<tr>
<td>CHAPTER 2. A MINIMAX TEST PROBLEM FOR UNCENSORED DATA</td>
<td>25</td>
</tr>
<tr>
<td>CHAPTER 3. A MINIMAX TEST PROBLEM FOR CENSORED DATA</td>
<td>35</td>
</tr>
<tr>
<td>CHAPTER 4. MINIMAX TESTS</td>
<td>47</td>
</tr>
<tr>
<td>4.1 Minimax Tests for One Observation</td>
<td>47</td>
</tr>
<tr>
<td>4.2 Minimax Tests for More Than One Observation</td>
<td>50</td>
</tr>
<tr>
<td>CHAPTER 5. PROOFS OF THEOREM 3.1 AND THEOREMS IN CHAPTER IV</td>
<td>53</td>
</tr>
<tr>
<td>5.1 Proofs for One Observation</td>
<td>53</td>
</tr>
<tr>
<td>5.2 Proofs for More Than One Observation</td>
<td>60</td>
</tr>
</tbody>
</table>
7. Fortran Program for Calculating the Average of Estimated OC-Curves and the Nominal OC-Curve going through \( \beta \) at \( \theta = \theta_1 \) ............................................................... 165

8. Fortran Program for Calculating the Distance Between the Average of Estimated OC-Curves and the Nominal OC-Curve going through \( 1 - \alpha \) at \( \theta = \theta_0 \) ................................................................................. 168

9. Fortran Program for Calculating the Distance Between the Average of Estimated OC-Curves and the Nominal OC-Curve going through \( \beta \) at \( \theta = \theta_1 \) ................................................................................. 175

APPENDIX B....................................................................................... 182

Figure 4. Asymptotic Relative Efficiency of Mean Absolute Deviate with respect to Square Deviate ....................................................... 182

Figure 5. Levels and Powers of Sign Test and Wilcoxon Test ........................................................................................................ 183

Figure 6. Levels of the Minimax Test ........................................... 184

Figure 7. Maximal and Ideal of The Minimax Test ........ 185

Figure 8. OC-Curves of the One-Stage Reliability Demonstration Test for \( \delta^* = 1.5 \) and \( \delta = 1, \) or, \( 1.5, \) or \( 2 \) ................................................................. 186

Figure 9. OC-Curves of the One-Stage Reliability Demonstration Test for \( \delta^* = 5 \) and \( \delta = 4, \) or, \( 5, \) or \( 6 \) ................................................................. 187

Figure 10. The Average of 100 Estimated OC-Curves and the Nominal OC-Curve for \( \delta = 2 \) ................................................................. 188

Figure 11. The Average of 100 Estimated OC-Curves and the Nominal OC-Curve for \( \delta = 3 \) ................................................................. 189
Figure 12. The Average of 100 Estimated OC-Curves and the Nominal OC-Curve for $\delta=4$ .......... 190

LIST OF REFERENCES......................................................................................... 191
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Truncation Constants as Functions of $\varepsilon$</td>
<td>34</td>
</tr>
<tr>
<td>2. Type I and Type II Errors</td>
<td>98</td>
</tr>
<tr>
<td>3. Distance Between Average of Estimated OC-Curve and Nominal OC-Curve for $r=0$ in (10.2)</td>
<td>128</td>
</tr>
<tr>
<td>4. Distance Between Average of Estimated OC-Curve and Nominal OC-Curve for $r=1$ in (10.2)</td>
<td>129</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The Superposition of Two Processes</td>
<td>15</td>
</tr>
<tr>
<td>2. Relationship of $Q_j$, $F_j$ and $P_j^*$</td>
<td>29</td>
</tr>
<tr>
<td>3. Relationship of $Q_j$, $F_j$ and $P_j^*$</td>
<td>51</td>
</tr>
<tr>
<td>4. Asymptotic Relative Efficiency of Mean Absolute Deviate With Respect To Square Deviate</td>
<td>182</td>
</tr>
<tr>
<td>5. Levels and Powers of Sign Test and Wilcoxon Test</td>
<td>183</td>
</tr>
<tr>
<td>6. Levels of the Minimax Test</td>
<td>184</td>
</tr>
<tr>
<td>7. Maximal power and Ideal power of the Minimax Test</td>
<td>185</td>
</tr>
<tr>
<td>8. OC-Curves of the One-Stage Reliability Demonstration Test for $\delta^*=1.5$ and $\delta=1$, or, 1.5, or 2</td>
<td>186</td>
</tr>
<tr>
<td>9. OC-Curves of the One-Stage Reliability Demonstration Test for $\delta^*=5$ and $\delta=4$, or, 5, or 6</td>
<td>187</td>
</tr>
<tr>
<td>10. The Average of 100 Estimated OC-Curves and the Nominal OC-Curve for $\delta=2$</td>
<td>188</td>
</tr>
<tr>
<td>11. The Average of 100 Estimated OC-Curves and the Nominal OC-Curve for $\delta=3$</td>
<td>189</td>
</tr>
<tr>
<td>12. The Average of 100 Estimated OC-Curves and the Nominal OC-Curve for $\delta=4$</td>
<td>190</td>
</tr>
</tbody>
</table>
Introduction

In statistics, we often make some assumptions about the model that governs the distribution of the collected data. For example, we may assume that data is sampled from a normal density. However, the true distribution of the data is seldom the assumed distribution, and there are often some small discrepancies between the assumed distribution and the true distribution. A procedure is robust if the property one is interested in is reasonably good at the assumed model; and if this property does not change much when small deviations from the assumed model do occur. For example, the property one is interested in for a test might be the power of a test or the level of a test, and for an estimator, the property of interest might be the asymptotic variance. In this dissertation, we shall show how to apply the concept of robustness to solve problems in both survival analysis and reliability. More specifically, we find a minimax test for censored data and find a two-stage demonstration test for a large series system.
1.1. Overview of robust procedures.

In this section, we shall discuss why we need robust procedures, what robust procedures are and how we derive robust procedures. Also, we briefly compare robust procedures with non-parametric procedures. Finally, we discuss the effects of censoring on robust procedures.

Newcomb (1886) was the first person who realized in practice that large errors occur more frequently than the normal law indicated that they would and introduced a mixture of distributions to represent the true underlying distribution of the data. Pearson (1931) discovered that some standard tests for equality of variances are highly sensitive to deviations from normality. With the same test problems, Box (1953) later introduced the term "robustness."

There are two widely used estimators for the dispersion of a population. They are defined as follows:

Mean Absolute Deviation: \[ d_n = \frac{1}{n} \sum_{i=1}^{n} |X_i - \bar{X}|, \]

and

Mean Square Deviation: \[ s_n = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2}, \]

where \( \bar{X} \) is the sample mean.
Eddington (1914, p. 147) and Fisher (1920, footnote on p. 762) disagreed about the relative merits of $d_n$ and $s_n$. Eddington claimed that $d_n$ is a safer criterion of accuracy than $s_n$ for data arising from astronomy. Fisher showed that the ratio of asymptotic standard errors of $d_n$ and $s_n$ is $\sqrt{\pi/2}$ if the data is from a Gaussian distribution; that is, for exact normal observations, $s_n$ is about 14% more efficient than $d_n$. It looked like Fisher had settled this problem. However, Tukey (1960) raised this question again in the context of contamination. In his paper, Tukey introduced a mixture model to represent a small contamination. For small $\varepsilon > 0$, each observation $x$ is from $N(\mu, \sigma^2)$ with probability $1 - \varepsilon$, and from $N(\mu, 9\sigma^2)$ with probability $\varepsilon$, where $N(\mu, 9\sigma^2)$ is a distribution for contamination. We summarize Huber's (1981) result in Figure C.1 in Appendix C. From Figure C.1, we see that under Tukey's mixture model, the Asymptotic Relative Efficiency (ARE) of $d_n$ with respect to $s_n$ is always greater than 1 for $0.01 < \varepsilon < 0.1$, and this ARE is 2 when $\varepsilon = 0.05$.

According to Webster's *Third New International Dictionary* (1986) the word "robustness" means "the state of having strength or vigorous health". Box (1953) borrowed this word "robustness" and used it in statistics. Since then, the word "robustness" has been given many different interpretations of
being sturdy and strong. For example, Huber (1981) defined "robustness" as insensitivity against small deviations from the distribution assumptions. According to Huber, a robust procedure should possess the following properties:

1. It achieves a reasonably good (optimal or almost optimal) efficiency at the assumed model.
2. Its performance is impaired only slightly if the discrepancy between the assumed distribution and the true distribution is small.
3. A larger discrepancy between the true underlying model and the assumed model should not cause a disaster.

However, some people refer to robustness as the robustness of validity: that is, the actual confidence levels should be close to the nominal level. Others refer to robustness as the robustness of performance: that is, power should be reasonably good at the assumed model and the power changes little when the discrepancy between the assumed distribution and the true distribution model is small.

In the theory of robustness, there are four widely used approaches to get a robust procedure. Next, we discuss these four approaches. Most of the following materials are from Hampel (1986).
The first approach is used to search for a new robust procedure. With this approach, the assumed parametric model is replaced by a super-model that is obtained by augmenting this parametric model by adding more parameters, such as shape or location parameters. The next step is to find an optimal procedure for the assumed parametric model and to show either by mathematics or by simulation study that this procedure still behaves reasonably well even in the super-model.

This approach has two weaknesses. The first weakness is that finding a robust procedure heavily depends on finding the right statistical procedure for the assumed parametric model. The second weakness is that possible deviant distributions allowed from the assumed parametric model are too narrow because some possible deviant distributions may not share the same parametric form as the assumed parametric distribution.

The second approach called an adapted robust approach is used to search for new robust procedures when there is a nuisance parameter involved. With this approach, we use a sample to estimate the nuisance parameter; then, we proceed to find a robust procedure. For example, suppose we have a batch of new series systems, all of whose component failure distributions are identical gammas with unknown shape and scale parameters. We want to test a hypothesis about the unknown scale parame-
ter. Using the second approach to find a robust testing procedure, we first estimate the unknown nuisance shape parameter from the first sample; then we derive a testing procedure based on the estimated value of the nuisance parameter. In the second part of my dissertation, we shall illustrate how to use this approach to obtain robust procedures in reliability.

The third approach was proposed by Huber (1964). Huber proposed a gross error model, small but with narrow full neighborhoods, and introduced a class of estimators, M-estimators. Huber defined a minimax problem for location problems among all M-estimators. Then, Huber solved this minimax problem and found a minimax M-estimator, which has the smallest asymptotic variance over the set of all symmetric distributions among all M-estimators.

Huber (1967) proposed the fourth approach. This approach is used to find a robust procedure for the finite sample case. In this paper, Huber derived the minimax test for one contamination neighborhood versus another contamination neighborhood. This test is minimax in the sense that it maximizes the minimum exact power over all alternatives in the second neighborhood, given a bound on the level of the test. Huber also showed that such a minimax test is often the ordinary likelihood ratio test between a least favorable pair of hypotheses.
The fifth approach is called the infinitesimal approach, and was proposed by Hampel (1968). This approach can be used to evaluate the robustness of a procedure. Because many robustness criteria, such as asymptotic variance or efficacy of a test, can be expressed in terms of an influence function, the infinitesimal approach can also be used to derive a robust procedure. For example, Hampel (1968) found an optimal robust estimator by minimizing asymptotic variance given a bound on the influence function. The infinitesimal approach is based on three essential concepts: qualitative robustness, influence function (with many derived concepts), and breakdown points. Huber (1972) liked Hampel’s idea so much that he said that Hampel's influence function was the most important single heuristic tool for constructing robust estimators with specific properties. In the same paper, he also explained these three concepts by an analogy to the stability of a bridge: (i) qualitative robustness means a small perturbation should have small effects on the bridge; (ii) the influence function measures the effect of an infinitesimal perturbation on the bridge; (iii) the breakdown point measures how big the perturbation can be before the bridge breaks.

One of the important properties of nonparametric procedures is that they are distribution-free under the null hypothesis and
some strict assumptions about a model. For example, the linear rank test for the one-sample location problem is distribution-free under the null hypothesis if the underlying distribution is symmetric about zero. However, the symmetry of the distribution can be easily destroyed by small contaminations. Rieder (1981) studied robustness of one and two sample rank tests against gross error neighborhoods. Rieder found maximal asymptotic level and minimal asymptotic power over the neighborhoods. We apply Rieder's results to a one-sample Wilcoxon test and sign test. We then study the impact of gross error contamination on the levels and powers of Wilcoxon and sign tests. Our results are summarized in Figure C.2. in Appendix. From the graph, we can conclude that the levels of both the sign test and Wilcoxon test increase as the fraction of contamination $\varepsilon$ increases. The level of the sign test is more stable than that of the Wilcoxon test. The Wilcoxon has greater power than the sign test for small fractions of contamination $\varepsilon$ (including no contamination, $\varepsilon=0$). However, as $\varepsilon$ increases, the power of the Wilcoxon test drops more rapidly than that of the sign test.

Eplett (1980) showed that the limiting power of the twosample Mann-Whitney test is fairly insensitive to the gross-error contamination, provided that the underlying distribution $F$ does not have very heavy tails. Eplett also showed that
the limiting power of the normal score's test is sensitive to gross error contamination.

1.2. Effect of censoring on statistical procedures.

In this section, we shall discuss the effect of censoring on statistical procedures.

In survival analysis, the response of interest, the time until some specific event, can not always be fully observed. Instead, the existence of censoring can terminate the observations before the event occurs. For example, in a clinical trial studying cancer or heart disease, some patients may withdraw from the trial or some patients may still be alive when the data is analyzed. When this happens, we call these terminated observations censored observations. These censored observations do not provide direct information about the time the event occurs. We need to reexamine the robust techniques developed in the complete data setting.

The most common model for the censoring type is arbitrary right censoring, with the censoring effects assumed noninformative with respect to the survival time. Using the noninformative censoring model, Emerson (1981) studied the effects of different uniform censoring distributions on the robustness of two confidence intervals for the median lifetime. He applied
the interval derived for exponential life to data simulated from a Weibull distribution. He found out that as the departure from an exponential distribution increases, the coverage probability for the true median moves away from 1-α, the nominal level of coverage probability. Also, she discovered that as censoring increases, the coverage probabilities improve, and at 50% censoring the coverage probabilities are very close to the nominal coverage probabilities under the exponential distribution. However, Hardner (1985) showed that when the probability of censoring is greater than 50%, the coverage probabilities get increasingly worse and do not seem to improve until the probability of censoring exceeds 90%. From these results, we see that censoring complicates the problem of robustness. This point of view is strengthened by Green and Crowley (1986). They used the M-estimator for censored data defined by Reid (1981) to be solution $T_n$ to the equation

$$\int \Psi(x-T_n) dG_n(x) = 0,$$

where $G_n$ is the Kaplan-Meier estimator of the cumulative distribution function of true underlying life time. Green and Crowley showed that there is no efficient M-estimator when the data is censored.
1.3. Minimax Tests.

In this section, we shall discuss what minimax tests are, why we need them, and how we search for minimax tests. Since part of this dissertation studies the minimax approach, we would like to discuss this approach in more detail, and to explain why the minimax approach can be thought of as an optimally robust approach. Most of the following materials are adapted from Huber (1981) and Hampel (1986).

We know that robustness means insensitivity to small deviation from the assumptions. A quantitative measure of robustness can be associated with the maximum degradation of performance possible for small deviations from the assumptions. Therefore, one type of optimal robust procedure is to minimize this maximum degradation. Thus, a minimax procedure is an optimally robust procedure.

There are two major kinds of minimax results. One is an exact result for finite samples. Another kind is an asymptotic result. There are two exact, finite sample minimax results available; the first result is for testing, and the resulting test is called the minimax test, and the second result is for interval estimates of location. One of disadvantages of the minimax test is that unless the sample size is very small, the actual level and minimum power of the minimax test are hard to calculate. One
disadvantage of asymptotic minimax estimators is that these minimax estimates exist only if the ideal parametric distribution is symmetric. Anscombe (1960) introduced the idea of insuring against losses of power caused by small deviations from the ideal model. Of course, to protect against power losses, some efficiency at the ideal model must be sacrificed. So, the questions are how much efficiency we are willing to sacrifice and how big a deviation should be allowed before we turn down a request for insurance against power losses. One possible approach is to fix a neighborhood of the ideal model and minimize the worst power loss that can occur within that neighborhood.

A systematic way to search for a robust procedure is to try to find a procedure whose worst behavior is the best among all procedures one is considering. Next, we will describe a way to formalize this idea.

First, determine a class of statistical procedures one is interested in, such as the class of all measurable tests or a class of invariant estimators. Second, choose a set of distributions that represent possible small deviations from the assumed parametric distribution. Third, select a quantity that measures the performance of a procedure. This quantity might be the level of a test or the power of a test. Fourth, for each procedure in the class, calculate the worst performance over all
possible small deviations specified in the second step. Fifth, try to find a procedure in the class of statistical procedures selected in the first step that minimizes the worst performance that can occur over the set of possible small deviations.

If we can solve the minimax problem above, the resulting procedure is a robust procedure against the set of small deviations one chooses and the class of selected procedures. If both the set of small deviations and class of defined procedures are large, then one can simply call a minimax procedure an optimally robust procedure.

When the data are censored, there is often limited information about the tail of the distribution of the time to an event. Therefore, we should find a procedure that is insensitive to the tail of the distribution. We need a robust test for censored data. One natural approach to search for a robust test is to formulate and solve a minimax problem for censored data, as Huber (1965) did for uncensored data, for which she found a test procedure that minimizes the maximal risk associated with using this test procedure.

1.4. Reliability tests for large series systems.

In the study of equipment reliability, one is often interested in large complex systems that can be represented as a collec-
tion of several components or subsystems that are in series, in the sense that the failure of any subsystem will result in the failure of the system. It may be helpful to understand the nature of this large complex system by imagining a complex system as a large number of sockets into each of which there is inserted a component. The system is in operation at time zero and it remains in operation from then on until a component fails. Components in different sockets may or may not be alike. However, each is subject to malfunction. The time \( t \) from the moment at which a component is inserted into its socket, up to the moment of its failure, obeys a probability distribution called the "failure law" of that component. Whenever a component fails, the system will fail. When failure happens, the faulty component is replaced immediately by a new one of the same kind. The process of failure detection, trouble location and replacement is assumed to consume no appreciable time (or if it does, that time is absorbed into the time between failures). As time goes on, at each socket there develops an unending sequence of failures that constitute a renewal random process. Therefore, a large complex system can be modeled as a superimposed renewal process. For example, a series system of two components can be represented as follows:
Component 1

Component 2

Series system

Here X stands for a failure time.

Figure 1. The superposition of two processes

Drenick (1960) showed that the superposition of a large number of independent renewal processes leads at time equilibrium to a homogeneous Poisson process (HPP). In other words, Drenick's theorem says that whatever the form of the component failure distribution of a repairable series system is, the system failure rate will tend towards a constant as the cumulative time of operation of the system becomes large and the total number of components involved in the test becomes large.

It is common for consumers, such as the Military, to test samples from suppliers to make sure that the product they want to purchase conforms to contractual specifications. These specifications are usually stated in terms of some parameter, such
as Mean Time Between Failure. Reliability demonstration tests are procedures to carry out hypothesis tests about the specified values of parameters.

One widely used collection of reliability demonstration tests is given by MIL-STD 781D, published by The Department of Defense. Plans of MIL-STD 781D have the following form:
First, the number of repairable systems to test and the testing time are given. Second, a constant is given. If the counted number of failures during the testing time exceeds the constant, then the hypothesis that the system conforms to the specifications will be rejected.

To investigate the validity of the tests described in MIL-STD 781D, we look at the special case of exponentially distributed failure times.

Suppose we have a series system with m components. Each component has a life distribution \( F_i(s) = 1 - e^{-\lambda_i s} \), \( i = 1, \ldots, m \). Whenever a component fails, it is immediately replaced by a similar component having the same failure distribution. All components are assumed stochastically independent. The sequence of failures of each component follows a Poisson process with mean rate \( \lambda \). Since the system fails each time any one of the m components comprising it fails, the sequence of system failures is obtained as a superposition of the m individual component
Poisson processes. Since the Mean Time To Failure (MTTF) of each component is $1/\lambda$, the Mean Time Between Failures (MTBF) of the system, $\theta$, is $1/(m\lambda)$. Let $N_i(s)$ be the number of failures of component $i$ during the period $[0,s]$ and let $T_{i1}, \ldots, T_{iN_i(s)}$ be the Time Between Failure (TBF) of component $i$ during the time period $[0,s]$, $i=1,2,\ldots,m$. Suppose the data available for the system consists of TBFs for all components, $T_{11},\ldots,T_{1N_1(s)},\ldots,T_{m1},\ldots,T_{mN_m(s)}$ and the number of failures for each component, $N_1(s),\ldots,N_m(s)$. We are interested in the reliability of the systems on the test. One criteria for the reliability is MTBF of the systems $\theta$. We want to test a hypothesis about $\theta$ based on data from new systems. That is, we want to test

$$H_0: \theta=\theta_0 \quad \text{versus} \quad H_1: \theta=\theta_1 \quad (\theta_0 > \theta_1).$$

The above hypotheses are equivalent to the following hypotheses

$$H_0: \lambda=\lambda_0 \quad \text{versus} \quad H_1: \lambda=\lambda_1 \quad (\lambda_0 < \lambda_1).$$

From reliability theory (see Barlow and Proschan (1981)), we know that $N_1(s),\ldots,N_m(s)$ are independently and identically Poisson processes with mean rate $\lambda$. Let us calculate the joint density function of $T_{11},\ldots,T_{1N_1(s)},\ldots,T_{m1},\ldots,T_{mN_m(s)},N_1(s),\ldots,N_m(s)$.

Since $(T_{i1},\ldots,T_{iN_i(s)},N_i(s))$ are independently and identically distributed for all $i=1,\ldots,m$,
\[ P\{ T_{ij} \in [s_{ij}, s_{ij} + \Delta_{ij}], N_i(s) = n_i \ ; \ j = 1, \ldots, n_i, i = 1, \ldots, M \} = \]
\[ \prod_{i=1}^{m} P( T_{ij} \in [s_{ij}, s_{ij} + \Delta_{ij}], N_i(s) = n_i \ ; \ j = 1, \ldots, n_i) = \]
\[ \prod_{i=1}^{m} \prod_{j=1}^{n_i} P( T_{ij} \in [s_{ij}, s_{ij} + \Delta_{ij}], T_i(n_i+1) > s - \sum_{j=1}^{n_i} s_{ij} ; 1 \leq j \leq n_i) = \]
\[ \prod_{1 \leq i \leq n} \prod_{j=1}^{n_i} (\exp(-\lambda s_{ij}) - \exp(-\lambda(s_{ij} + \Delta_{ij}))) \exp(-\lambda(s - \sum_{j=1}^{n_i} s_{ij})) . \]

Since
\[ \exp(-\lambda s_{ij}) - \exp(-\lambda(s_{ij} + \Delta_{ij})) \]
\[ \Delta_{ij} \rightarrow \Delta_{ij} = 0 \]
\[ \lambda \exp(-\lambda s_{ij}) \]
and
\[ f(s_{11}, s_{1m_1}, \ldots, s_{nm_n}, m_1, \ldots, m_n) = \]
\[ \lim_{\Delta_{ij} \rightarrow 0} \prod_{(i,j)} \prod_{i=1}^{m} \prod_{j=1}^{n_i} P( T_{ij} \in [s_{ij}, s_{ij} + \Delta_{ij}], N_i(s) = n_i \ ; \ j = 1, \ldots, n_i, i = 1, \ldots, m ) \]

the joint density function of
\[ T_{11}, \ldots, T_{1N_1(s)}, \ldots, T_{n1}, \ldots, T_{nN_m(s)}, N_1(s), \ldots, N_m(s) \]
is given by
\[ f(s_{11}, \ldots, s_{1m_1}, \ldots, s_{nm_n}, m_1, \ldots, m_n) = \lambda^n \exp(-\lambda s), \]
where
It follows from the Neyman-Pearson lemma that the Most Powerful (MP) level \( \alpha \) test procedure for \( H_0 \) versus \( H_1 \) is to accept \( H_0 : \theta = \theta_0 \) if
\[
\frac{\lambda_1}{\lambda_0} N(s) \exp((\lambda_0 - \lambda_1)s) < d \quad \text{or} \quad N(s) < c_s,
\]
where \( c_s \) is determined by
\[
P_{\theta_0}(N(s) < c_s) = 1 - \alpha.
\]
We conclude that to derive the MP test for testing a hypothesis about the system MTBF, we do not need to know each component's TBFs \( T_{ij}, j=1, \ldots, N_i(s) \) and the number of failures for each individual component, \( N_i(s), i=1, \ldots, m \). The only data we need is the total number of system failures \( \sum_{i=1}^{m} N_i(s) \).

Based on (1.2)-(1.4), the test design, given in MIL-STD-781D is for a large complex series system that has exponential system TBF's. The test requires the counting of the number of system failures \( N_S(t) \) during the time period \((s, s+t)\). We accept the hypothesis that system MTBF is \( \theta_0 \) if \( N(t) < c \), for some \( c \).
The constant failure rate assumption for a mature system is justified by Drenick's theorem.

If someone utilizes the testing procedures defined by MIL-STD-781D for a large immature complex system that contains components exhibiting wear-out failure model, what would happen to the OC-Curve of the test? Störmer (1969) and Blumenthal, Greenwood, and Herbach (1971) showed independently that if all component processes have the same age $t$ and $t < \infty$ (that is, the system is not in equilibrium state) and if $m$ is large, then the approximate distribution of TBF is exponential with a scale parameter that is a function of the renewal density associated with the component failure density. Kasouf (1979) showed that most statistical procedures defined by MIL-STD-781D were very sensitive to departure from the initial assumption of constant failure rate of the system. Kasouf also noticed that applying these techniques to the reliability demonstration test, when the assumption of constant failure rate is not satisfied, may result in a substantial increase in the probability of accepting unreliable equipment. Blumenthal, Greenwood, and Herbach (1984) showed that for a complex system whose components tend to wear out as they age, the transient reliability can be much higher than the steady state reliability. In fact, a system with unacceptable equilibrium reliability can be very
likely to pass a reliability demonstration test when the system age is ignored in designing the test. Since testing time in the laboratory is much shorter than the time of use in the field, laboratory test reliability is that of a new system, but the observed reliability of field use is essentially that of an equilibrium system. Therefore time related wear-out failure modes of components will account for part of the discrepancy noted between the field and laboratory MTBF's.

Reliability demonstration tests in MIL-STD-781D are not suitable for new series systems. It is worthwhile to derive a correct reliability demonstration test for equilibrium MTBF when testing systems that are brand new ones, with each component having increasing hazard rate. We consider only brand new systems. For partially aged systems, it is very hard to describe how to redesign the reliability demonstration test to achieve the desired OC-Curve due to the difficulty of knowing the age of the systems and the complex relation of the renewal function to the underlying component failure density. Blumenthal and Zhou (1989) used the Poisson distribution with an appropriate mean to approximate the distribution of the number of system failures and derived a reliability demonstration test for brand new systems. The OC-Curve derived from this test can apply to any of the following three settings:
(1) The decision relates to a large batch of systems from which the tested ones are a random sample that gives information about the equilibrium behavior of all members of the batch.
(2) The decision relates to the equilibrium behavior of other systems manufactured under the same conditions as the test sample.
(3) Accept-reject decisions based on the test only relate to the tested systems. That is, whether they pass or fail the test determine only if equilibrium reliability of the tested systems will be satisfactorily high.

1.5. Content of Chapters.

We first give the contents of chapters covering minimax tests for censored data. Then, we give the contents of chapters covering reliability demonstration tests for large new series systems. In Chapter II to Chapter VI of this dissertation, we investigate whether this minimax test can be extended to censored data. The answer turns out not to be as pleasant as we wished. We show that if the contamination distribution may be any distribution, then there is no minimax test. However, we do show that under some restrictions on the contamination distributions, a minimax test does exist.
In Chapter II, we review Huber's minimax tests for uncensored data. In Chapter III, we define minimax test for censored data and we define a least favorable pair for censored data. Then, we propose a candidate least favorable pair for censored data. In Chapter IV, we propose a candidate test and show there is no minimax test for full contamination neighborhoods in a single observation. We also derive the necessary and sufficient restrictions on contamination distributions so that our candidate test is a minimax test in a single observation. Also, we derive sufficient restrictions on the set of contamination distributions such that our candidate test is a minimax test when we have \( n \geq 2 \) observations. In Chapter V, we give proofs of results in Chapter III and Chapter IV. In Chapter VI, we derive a minimax test in a simpler censoring setting: Type-I censoring. Also, we work out in detail an example where both the true life time and the censoring time are exponentially distributed. This example shows that neither the power of the minimax test nor the level of the minimax test depends heavily on the censoring distribution, even though we assume the censoring distribution is known.

In Chapter VIII, IX and X, we shall develop a two-stage reliability test for the system MTBF in equilibrium based on count
data on brand new systems, each component having gamma failure distribution with shape parameter $\delta$, $\delta \geq 1$. In Chapter VII, we review the reliability demonstration test proposed by Blumethal and Zhou (1989) for the known shape parameter $\delta$. Then we show the OC-Curve of this reliability demonstration test is sensitive to the assumed value of the shape parameter $\delta$. Thus, we need to develop a robust demonstration test that is insensitive to the assumed value of the shape parameter $\delta$. In Chapter VIII, we derive a two-stage reliability demonstration test, and show that as the number of systems in the first stage test tends to infinity, the OC-Curve derived from this two-stage reliability demonstration test tends in probability to the nominal OC-Curve derived for a single-stage procedure assuming the shape parameter to be known. In Chapter IX, we report a simulation study of small first stage test size properties that confirms our limiting results and give some guidance on how to choose the number of systems for the first stage test.

In the last Chapter of the dissertation, we shall contrast the two approaches with robustness. We give overall conclusions, and propose some future research problems.
CHAPTER II.

A MINIMAX TEST PROBLEM FOR UNCENSORED DATA

Uncensored data is a special case of censored data. Huber (1965) has defined and solved a minimax test problem for uncensored data. In this Chapter, we shall review Huber's result in the general setting of censored data. Suppose we have n patients. Let $T_i$ be the survival time of the $i^{th}$ patient, and $C_i$ denote the censoring time associated with the $i^{th}$ patient. Assume that $T_1, T_2, \ldots, T_n$ are independently, identically distributed [abbreviated as i.i.d.] random variables, each with survival function $\bar{F}$, and that $C_1, C_2, \ldots, C_n$ are i.i.d., each with survival function $\bar{G}$. Also assume that $T_i$ and $C_i$ are independent.

We can only observe

$$(Y_1, \delta_1), \ldots, (Y_n, \delta_n),$$

where

$$Y_i = \min(T_i, C_i),$$

and

$$\delta_i = \begin{cases} 
1 & \text{if } T_i \leq C_i \\
0 & \text{if } T_i > C_i 
\end{cases}.$$

For notational convenience, from now on we shall use the same capital letter for both a survival function (subsurvival function)
and a probability measure (submeasure). For example, if we write \( \tilde{G}(x) \), then \( \tilde{G} \) denotes the survival function (subsurvival function). If we write \( \tilde{G}(B) \), where \( B \) is a Borel set, then \( \tilde{G} \) denotes a probability measure (submeasure). Let \( f \) be the density of \( \hat{F} \) with respect to a carrier measure \( \mu \), defined on \( (R^+,B) \), where \( R^+ = (0,\infty) \) and \( B \) is the Borel field on \( R^+ \). Let \( g \) be the density of \( \tilde{G} \) with respect to the same measure \( \mu \). Then, the likelihood of our full sample is

\[
L(y, \delta) = \prod_{i=1}^{n} f(y_i)^{\delta_i} \hat{F}(y_i)^{1-\delta_i} g(y_i)^{1-\delta_i} \tilde{G}(y_i)^{\delta_i},
\]

where

\[
(y, \delta) = (y_1, \ldots, y_n; \delta_1, \ldots, \delta_n).
\]

Let \( \tilde{F}_0 \) and \( \tilde{F}_1 \) be two distinct probability measures on \((R^+,B)\) in which we are interested. In order to formalize the possibility of unknown small deviations from the idealized model \( \tilde{F}_j, j=0,1 \), the idealized model \( \tilde{F}_j \) is replaced by a mixture model:

\[
P_j = \{ Q_j' | Q_j' = (1-\epsilon) \tilde{F}_j + \epsilon L_j : L_j \subseteq L_j \}, j=0,1,
\]

where \( L_j \) is a specific set of survival functions, and the contamination fraction \( \epsilon \) is fixed and known. Each set \( P_j \) defined by (2.7) is called a contamination neighborhood of \( \{F_j\} \) or a gross error model. If \( X_1, \ldots, X_n \) are i.i.d with a common survival func-
tion $Q'$, then our null hypothesis $H_0$ and alternative hypothesis $H_1$ are defined by

$$H_0: Q \in P_0 \text{ versus } H_1: Q \in P_1,$$

where the $P_j$ are defined by (2.7).

If we observe the survival time of all patients, we call this special case of censored data uncensored data. Huber (1965) defined the contamination neighborhoods of idealized models by (2.7), where $L_j$ is the set of all survival functions with densities, and showed there exists a least favorable pair $Q_j \in P_j$, $j=0,1$, such that the Neyman-Pearson likelihood ratio test of $Q_0$ versus $Q_1$ is also a minimax test of $P_0$ versus $P_1$. Huber calls the pair $(Q_0, Q_1)$ as a least favorable pair. Since the least favorable pair played an important role in finding a minimax test for uncensored data, we review Huber’s results for the least favorable pairs (LFP) for uncensored data. The densities of $Q_0$ and $Q_1$ are defined as follows:

$$(2.8) \quad q_0(t) = \begin{cases} 
(1-\epsilon)f_0(t) & \text{for } \frac{f_1(t)}{f_0(t)} < c'' \\
\frac{1-\epsilon}{c''}f_1(t) & \text{for } \frac{f_1(t)}{f_0(t)} \geq c''
\end{cases}$$

and
\[ q_1(t) = \begin{cases} (1-\varepsilon)f_1(t) & \text{for } \frac{f_1(t)}{f_0(t)} > c' \\ (1-\varepsilon)c'f_0(t) & \text{for } \frac{f_1(t)}{f_0(t)} \leq c' \end{cases} \]

where \( \varepsilon \) is the same one as in (2.7), and \( c' \) and \( c'' \) are determined by (2.9):

\[ \int q_0(t)dt = 1 \text{ and } \int q_1(t)dt = 1. \]

From (2.8) and (2.9), it follows that

\[ \frac{q_1(t)}{q_0(t)} = \begin{cases} c' & \text{for } \frac{f_1(t)}{f_0(t)} < c' \\ \frac{f_1(t)}{f_0(t)} & \text{for } c' \leq \frac{f_1(t)}{f_0(t)} \leq c'' \\ c'' & \text{for } \frac{f_1(t)}{f_0(t)} > c'' \end{cases} \]

Note that

\[ Q_j(t) = \int_{-\infty}^{\infty} q_j(x)dx, \quad j = 0, 1. \]

Thus, we conclude that if a least favorable pair \((Q_0, Q_1)\) exists, then

\[ 0 \leq c' \leq c'' \leq \infty. \]

From (2.10), we know that \( q_1(t)/q_0(t) \) is a truncated version of \( f_1(t)/f_0(t) \) by \( c' \) from below and \( c'' \) from above.

We could represent the relationships of \( Q_j, \bar{F}_j, \) and \( P_j \) by Figure 2. Note that \((Q_0, Q_1)\) is the hardest pair to test among all pairs in \((P_0, P_1)\).
Figure 2. Relationship of $Q_j$, $F_j$, and $P_j$.

The least favorable pair is denoted by $(Q_0, Q_1)$. Ideal survival functions are $\bar{F}_j$, the contamination neighborhoods are $P_j$.

Proposition 2.1 below from Huber (1965, p1755) gives a way to find c' and c''. Define a function of c, $B(c; f_0, f_1)$, as follows

$$B(c; f_0, f_1) = \begin{cases} \int f_0(t) d\mu + \frac{1}{c} \int f_1(t) d\mu & \text{if } f_1(t)/f_0(t) < c \\ \int f_1(t) d\mu & \text{if } f_1(t)/f_0(t) \geq c \end{cases}$$

The expression of $B(c; f_0, f_1)$ will be used to solve c' and c''.
Proposition 2.1:
(a) For any fixed densities \( f_0 \) and \( f_1 \), \( B(c;f_0,f_1) \) is a continuous function
of \( c \); \( B(c;f_0,f_1)=1 \) for \( c \geq c_1 \); \( B(c;f_0,f_1) \) is strictly decreasing for
\( 0 \leq c \leq c_1 \), where \( c_1 = \text{esssup}_{x \geq 0} \left( \frac{f_1(x)}{f_0(x)} \right) \).
(b) If \( P_0 \neq P_1 \), then \( c'' \) is the unique solution of the following equation:
\[
(2.12) \quad B(c;f_0,f_1) = \frac{1}{1-\epsilon},
\]
and \( c' \) is the unique solution of the following equation:
\[
(2.13) \quad B(1/c;f_1,f_0) = \frac{1}{1-\epsilon}.
\]
For \( \epsilon \) sufficiently small, \( c' < c'' \).

The next lemma gives the properties of \( c' \) and \( c'' \).

Lemma 2.1: Let \( c''=c''(\epsilon) \) and \( c'=c'(\epsilon) \) be solutions of (2.12) and
(2.13) respectively. Then, for any fixed \( \epsilon \in (0,1) \), \( c'' \) is a decreasing function of \( \epsilon \), and \( c' \) is an increasing function of \( \epsilon \).

Proof: Since \( f_0 \) and \( f_1 \) are fixed, we abbreviate \( B(c;f_0,f_1) \) by
\( B(c) \). Let \( 0 < \epsilon_1 < \epsilon_2 < 1 \); then the values \( c''(\epsilon_j) \) satisfy the following equations:
\[
(2.14) \quad B(c''(\epsilon_j)) = \frac{1}{1-\epsilon_j}, \quad j=1,2.
\]
Since \( 1/(1-\epsilon) \) is an increasing function of \( \epsilon \) for \( 0 < \epsilon < 1 \),
\[
(2.15) \quad B(c''(\epsilon_1)) < B(c''(\epsilon_2)).
\]
Since Proposition 3.1 implies that $B(c)$ is a decreasing function of $c$, it follows from (2.15) that
\[ c''(\varepsilon_1) > c''(\varepsilon_2). \]
This completes the proof of Lemma 3.1 for $c''$. The proof for $c'$ is similar.

Since $c'$ and $c''$ are determined by (2.12) and (2.13), the next example show how to calculate $c'$ and $c''$ and how $c'$ and $c''$ depend on $\varepsilon$. Since the exponential distributions are important survival distributions in survival analysis, in the next example, we assume that the true survival function is an exponential function. We use Newton's method to solve the non-linear equations (2.12) and (2.13) for $c''$ and $c'$ respectively. To employ Newton's method, we need the derivatives of $B(c; f_0, f_1)$, and $B(1/c; f_1, f_0)$ with respect to $c$.

**Example 2.1:** Assume that
\[ \tilde{F}_j(t) = \exp(-\lambda_j t), \quad j = 0, 1, \]
where $\lambda_1 < \lambda_0$. Then, if the carrier measure $\mu$ is the Lebesgue measure,
\[ (2.16) \quad f_1(t)/f_0(t) = (\lambda_1/\lambda_0) \exp[(\lambda_0 - \lambda_1)t]. \]

From (2.16), it follows that
\[ (2.17) \quad f_1(t)/f_0(t) > c \iff t > \frac{1}{\lambda_0 - \lambda_1} \ln(c\lambda_0/\lambda_1). \]

If $c \leq \lambda_1/\lambda_0$, then
\[
\frac{1}{\lambda_0 - \lambda_1} \ln(c \lambda_0/\lambda_1) < 0 \quad \text{for} \quad \lambda_1 < \lambda_0.
\]

But since \( t \) is nonnegative, (2.17) implies that
\[
\{ t : f_1(t)/f_0(t) > c \} = \mathbb{R}^+.
\]

Now we want to find the explicit expression to solve (2.12) and (2.13). Denote \( B(c; f_0, f_1) \) by \( B_1(c) \) when \( d\mu \) is replaced by \( dt \).

Then,
\[
B_1(c) = 1/c \quad \text{if} \quad c \leq \lambda_1/\lambda_0. \quad \text{If} \quad c > \lambda_1/\lambda_0, \quad \text{then, by (2.17) we have}
\]

\[
B_1(c) = \frac{1}{\lambda_0} \exp(-\lambda_0 t) dt + (1/c) \int_0^\infty \lambda_1 \exp(-\lambda_1 t) dt = u(c, \lambda_0, \lambda_1)
\]

\[
= 1 - \exp(-\lambda_0 u(c, \lambda_0, \lambda_1)) + \frac{1}{c} \exp(-\lambda_1 u(c, \lambda_0, \lambda_1)),
\]

where
\[
(2.18) \quad u(c, \lambda_0, \lambda_1) = \frac{\ln(c \lambda_0/\lambda_1)}{\lambda_0 - \lambda_1}.
\]

So,

\[
B_1(c) = 1 - (\lambda_1/(c \lambda_0))^{\lambda_0/(\lambda_0 - \lambda_1)} + (1/c)(\lambda_1/(c \lambda_0))^{\lambda_1/(\lambda_0 - \lambda_1)} \quad \text{for} \quad c > \lambda_1/\lambda_0.
\]

Define \( a_j \) by (2.19) below,
\[
(2.19) \quad a_j = \lambda_j/\lambda_0, \quad j = 0, 1, \quad r = \lambda_1/\lambda_0.
\]

Then, \( a_0 r = a_1 \), and
Because we will use Newton's method to solve non-linear equations (2.12) and (2.13), we need the derivative of $B_1(c)$. Differentiating (2.20) gives us

$$\frac{dB_1(c)}{dc} = \begin{cases} 
(a_1/c^2)(r/c)^{a_0-1} - ((1+a_1)/c^2)(r/c)^{a_1} & \text{for } c > r \\
-1/c^2 & \text{for } c \leq r
\end{cases}$$

From (2.20) and (2.21) and using Newton's method, we can solve $c''$.

Next, we solve for $c'$. Let $B_2(c) = B(1/c,f_1,f_0)$. Using the same method as we used to get (2.20) and (2.21), we obtain

$$B_2(c) = \begin{cases} 
c[1-(r/c)^{a_0}] + (r/c)^{a_1} & \text{for } c > r \\
1 & \text{for } c \leq r
\end{cases}$$

and

$$\frac{dB_2(c)}{dc} = \begin{cases} 
1 + (a_0-1)(r/c)^{a_0} - (a_1/c)(r/c)^{a_1} & \text{for } c > r \\
0 & \text{for } c \leq r
\end{cases}$$

Appendix A.1 contains a Fortran implementation of Newton's method for solving $c'$ and $c''$. Solutions for (2.12) and (2.13) for different $\epsilon$'s are illustrated in Table 1.
Table 1
Truncation constants $c'$ and $c''$ as functions of $\varepsilon$.

<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>$c''$</th>
<th>$c'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.040000</td>
<td>1.526285</td>
<td>0.822878</td>
</tr>
<tr>
<td>0.080000</td>
<td>1.194349</td>
<td>0.906784</td>
</tr>
<tr>
<td>0.100000</td>
<td>1.100642</td>
<td>0.945333</td>
</tr>
<tr>
<td>0.129000</td>
<td>1.000095</td>
<td>0.999940</td>
</tr>
<tr>
<td>0.130000</td>
<td>0.997143</td>
<td>1.001813</td>
</tr>
</tbody>
</table>

Survival functions $\bar{F}_j(t) = \exp(-\lambda_j t)$, with $\lambda_0 = 3.0$, $\lambda_1 = 2.0$. The truncation $c'$, and $c''$ are the solutions of (2.13) and (2.12) respectively.

Table 1 displays how $c''$ and $c'$ depend on $\varepsilon$. In the table, we see that $c''$ is a decreasing function of $\varepsilon$, and $c'$ is an increasing function of $\varepsilon$. Also, $c'' > c'$ for $0 < \varepsilon \leq 0.129$, and $c'' < c'$ for $0.13 \leq \varepsilon < 1$. For $\varepsilon > 0.13$, the sets $P_0$ and $P_1$ overlap.
In the previous Chapter, we discussed a minimax test problem for uncensored data. In this Chapter, we will discuss a minimax test problem for censored data. We define a minimax test for censored data. Since we know from Chapter II that a least favorable pair play an important role in finding a minimax test; therefore, we shall define a least favorable pair for censored data. Then, we propose a candidate least favorable pair for censored data.

If \(i\) th patient is censored, then we can not observe the survival time of \(i\) th patient and we can observe the censoring time. Let \(\phi = \phi(y, \delta)\) be any randomized test of \(H_0\) versus \(H_1\). Let \(R_j(\{(Q_j, \tilde{G})\}; \phi)\) be the risk associated with \(\{(Q_j, \tilde{G})\}; \phi)\). Let \(r_j\) be the constant loss associated with wrongly rejecting \(H_j\), and the \(f_j'\) be density of \(Q_j'\) with respect to \(\mu\). Let \(L'(y, \delta) = L(y, \delta)\) be defined in (2.3) with \(f\) replaced by \(f_j'\) and \(\tilde{F}\) replaced by \(Q_j'\). Then, we have

\[
R_0(\{(Q_0', \tilde{G})\}; \phi) = r_0 E_{\{(Q_0', \tilde{G})\}}(\phi),
\]

and
\[ R_1(\{(Q_j,\tilde{G})\};\phi) = r_1(1 - E[\{(Q_j,\tilde{G})\}\phi]) , \]

where
\[ E[\{(Q_j,\tilde{G})\}](\phi) = \sum_{(\delta_1, \ldots, \delta_n) = \{(0, 1)^n\}} \int \phi(y; \delta)L'(y, \delta)dy_1 \ldots dy_n. \]

Next, we shall give the definition for a minimax test for censored data.

**Definition 3.1:** A test procedure \( \phi \) is called a minimax test of level \( \alpha \) for testing \( H_0 \) it satisfies the following two conditions:

\[(3.3) \quad \sup_{\mathcal{P}_0} R_0((F, G); \phi) \leq \alpha, \]
\[(3.4) \quad \text{For any } \gamma \text{ such that } \sup_{\mathcal{P}_0} R_0((F, G); \gamma) \leq \alpha, \sup_{\mathcal{P}_0} R_1((F, G); \phi) \leq \sup_{\mathcal{P}_0} R_1((F, G); \gamma). \]

The requirement (3.3) says that the size of the test \( \phi \) is \( \alpha \).
The requirement (3.4) says that \( \phi \) is the test which minimizes the maximum risk that can occur over all alternatives for all tests whose maximal levels over all hypotheses is less than \( \alpha \).

Since we have
\[ \sup_{S \in \mathcal{P}_1} R_1((S, \tilde{G}); \phi) = r_1 \sup_{S \in \mathcal{P}_1} (1 - E(S, \tilde{G})\phi(T, \delta)) = r_1 - r_1 \inf E(S, \tilde{G})\phi(T, \delta)), \]
minimizing $\sup_{S \in P_1} R_1((S, \tilde{G}); \phi)$ is equivalent to maximizing

$\inf_{S \in P_1} E(S, \tilde{G}) \phi(T, \delta)$. For this reason, Hampel, et al. (1986), called

the $\phi$ that satisfies (3.3) and (3.4) a maximin test.

Huber (1965) showed that a least favorable pair $(Q_0, Q_1)$ exists with $Q_j \in P_j$, $j=0,1$, when there is no censoring. The minimax test is just the likelihood ratio test of $Q_0$ versus $Q_1$. If there is censoring, then we may not observe the lifetimes of all patients. To solve the minimax test problem above we would like to use some ideas Huber (1965) used to solve the minimax test problem for uncensored data. Huber (1965) solved the minimax problem by using idea of least favorable pairs that are hardest to distinguish in contamination neighborhoods for uncensored data. Since the least favorable pair plays an important role in finding a minimax test for uncensored data, we would like to use the least favorable pairs approach to solve the minimax problem for censored data. To do so, we need to extend the concept of least favorable pair for uncensored data to censored data.
Definition 3.2: A pair \((Q_0, Q_1)\) is called a least favorable pair for censored data with respect to \(\varphi\) if the requirements (3.5) to (3.7) are satisfied:

(3.5) \[ E(Q_0, G)\varphi(T, \delta) \leq E(Q_0, G)\varphi(T, \delta) = \alpha \text{ for all } Q_0 \in P_0. \]

(3.6) \[ E(Q_1, G)\varphi(T, \delta) \geq E(Q_1, G)\varphi(T, \delta) \text{ for all } Q_1 \in P_1. \]

(3.7) \[ \gamma(T, \delta), \]

where \(\gamma\) is any randomized test satisfying

(3.8) \[ \max_{Q_0 \in P_0} E(Q_0, G)\gamma(T, \delta) \leq \alpha. \]

If there is no censoring, then \(E(Q_0, G)\varphi(T, \delta) = E(Q_1, G)\varphi(T, \delta)\), if \(\varphi\) is the probability ratio test of \(Q_0\) versus \(Q_1\). Therefore, if \((Q_0, Q_1)\) is a least favorable pair, then

\[ E(Q_1, G)\varphi(T, \delta) \leq E(Q_1, G)\varphi(T, \delta) \leq E(Q_0, G)\varphi(T, \delta) \leq E(Q_1, G)\varphi(T, \delta). \]

So, \((Q_0, Q_1)\) is the hardest pair of simple hypotheses to test.

Therefore, our definition is a generalization of the definition for a least favorable pair for uncensored data.

The next lemma guarantees that a minimax test exists if a least favorable pair exists for censored data.

Lemma 3.1: If a pair \((Q_0, Q_1)\) is a least favorable pair for censored data with respect to \(\varphi\), then \(\varphi(T, \delta)\) is a minimax test and

\[ \sup_{\text{all } \varphi \text{ satisfying (3.8)}} \{ \inf_{Q_1 \in P_1} E(Q_1, G)\varphi(T, \delta) \} = E(Q_1, G)\varphi(T, \delta). \]
Proof of lemma 3.1: Taking the maximum over \( Q_0 \) in \( P_0 \) on both sides of (3.5) gives us

\[
\max_{Q_0 \in P_0} E(Q_0, \tilde{G})\varphi(T, \delta) = E(Q_0, \tilde{G})\varphi(T, \delta) = \alpha,
\]

which implies that (3.3) is true. Next we shall show that (3.4) is true also. Similarly, taking the minimum over \( P_1 \) of both sides (3.6) gives us

\[
\min_{Q_1 \in P_1} E(Q_1, \tilde{G})\varphi(T, \delta) = E(Q_1, \tilde{G})\varphi(T, \delta).
\]

For any \( \phi \) satisfying (3.8), we have \( E(Q_0, \tilde{G})\varphi(T, \delta) \leq \alpha \).

Combining the hypothesis (3.7) with the equation above gives

\[
\min_{Q_1 \in P_1} E(Q_1, \tilde{G})\varphi(T, \delta) \geq E(Q_1, \tilde{G})\varphi(T, \delta)
\]

But,

\[
E(Q_1, \tilde{G})\varphi(T, \delta) \geq \min_{Q_1 \in P_1} E(Q_1, \tilde{G})\varphi(T, \delta).
\]

So,

\[
(3.9) \quad \min_{Q_1 \in P_1} E(Q_1, \tilde{G})\varphi(T, \delta) \geq \min_{Q_1 \in P_1} E(Q_1, \tilde{G})\varphi(T, \delta).
\]

Thus, we have shown

\[
\max_{\text{all } \phi \text{ satisfying (3.8)}} \{ \min_{Q_1 \in P_1} E(Q_1, \tilde{G})\varphi(T, \delta) \} = E(Q_1, \tilde{G})\varphi(T, \delta).
\]

Since the loss functions is 0 or 1,

\[
R_0((S, \tilde{G}); \varphi) = r_0 E(S, \tilde{G})\varphi(T, \delta),
\]

and
\[ R_1((S, \tilde{G}); \varphi) = r_1(1 - E(S, \tilde{G}) \varphi(T, \delta)). \]

So, from (3.9) we conclude that (3.4) is true also. This completes the proof of Lemma 3.1.

We propose a candidate minimax test for censored data as follows:

\[(3.10) \quad \varphi(y, \delta) = \begin{cases} 
1 & \text{if } \sum_{i=1}^{n} \left( \delta_i \ln \left( \frac{q_1(y_i)}{q_0(y_i)} \right) + (1 - \delta_i) \ln \left( \frac{Q_1(y_i)}{Q_0(y_i)} \right) \right) > \ln k, \\
0 & \text{otherwise}
\end{cases} \]

where \( k \) is determined by

\[(3.11) \quad E(Q_0, G) \varphi(Y, \delta) = \alpha. \]

Recall that if \( \delta_i = 1 \), we observe the survival time of \( i \)th patient, and if \( \delta_i = 0 \), we observe the censoring time of \( i \)th patient. And, if there is no censoring, that is, if \( \delta_i = 1 \) for \( i = 1, \ldots, n \), then our candidate test reduces to the probability ratio test of \( Q_0 \) versus \( Q_1 \), which is the minimax test for uncensored data. Thus, the proposed candidate test is the generalization of the minimax test proposed by Huber for censored data.

Next, we begin to search for a least favorable pair for censored data with respect to \( \varphi(y, \delta) \), defined by (3.10). An obvious candidate is the pair \((Q_0, Q_1)\) defined by (3.1).

If there is censoring, we shall show that \((Q_0, Q_1)\) will not be a least favorable pair for the those contamination neighborhoods.
where a contamination distribution can be any distribution. Then, we shall give conditions on contamination distributions such that the pair \((Q_0, Q_1)\) is a least favorable pair.

Before we proceed, we need some definitions. Let us define \(\langle G, S \rangle\) to be the mapping from \(A \times A\) into \(C[0,\infty)\) such that

\[
\langle G, S \rangle(t) = \int G(x)[-dS(x)],
\]

where \(A\) is the set of absolutely continuous survival functions and \(C[0,\infty)\) is the set of all continuous functions. Let the symbol \(\langle G, S \rangle\) denote either the subsurvival function of (3.12) or the induced measure. Applying integration by parts, we can show that

\[
\langle \tilde{G}, \tilde{S} \rangle(t) + \langle \tilde{S}, \tilde{G} \rangle(t) = \tilde{G}(t)\tilde{S}(t) \text{ for all } t \geq 0.
\]

Next, we define two density functions of contamination distributions to be used to define our contamination neighborhoods. These two density functions are defined as follows:

\[
(3.14) \quad h_0(t) = \begin{cases} \frac{(1-\varepsilon)}{\varepsilon}(f_1(t)/c' - f_0(t)) & \text{for } f_1(t) > c''f_0(t) \\ 0 & \text{otherwise} \end{cases}
\]

\[
(3.15) \quad h_1(t) = \begin{cases} \frac{(1-\varepsilon)}{\varepsilon}(c'f_0(t) - f_1(t)) & \text{for } f_1(t) \leq c'f_0(t) \\ 0 & \text{otherwise} \end{cases}
\]

where \(c'\) and \(c''\) are determined so that \(h_0\) and \(h_1\) are densities on \([0,\infty)\). Therefore, \(c''\) and \(c'\) are determined uniquely by (3.16):
\begin{equation}
\int \frac{f_1(t)}{f_0(t)} d\mu = \frac{f_1(t)}{f_0(t)} > c'' \\
= \int (c'f_0(t)-f_1(t))d\mu = \frac{e}{1-e}.
\end{equation}

Let \( L_j \) be survival functions of \( h_j \). It is easily shown that

\begin{equation}
Q_j = (1-e)\tilde{F}_j + eL_j,
\end{equation}

where \( Q_0 \) and \( Q_1 \) are the least favorable pair defined by Huber.

Next, we shall show that a mixture model for true lifetimes leads to a mixture model for both subsurvival functions \( <\tilde{F}_j, \tilde{G}> \) and \( <\tilde{G}, \tilde{F}_j> \).

\textbf{Lemma 3.2:} For any survival function \( L'_j \), the following conditions are equivalent:

\begin{equation}
Q'_j = (1-e)\tilde{F}_j + eL'_j
\end{equation}

\begin{equation}
\begin{cases}
<Q'_j \tilde{G}> = (1-e)<\tilde{F}_j, \tilde{G}> + e<L'_j \tilde{G}>
\\
<L'_j \tilde{G}> = (1-e)<\tilde{G}, \tilde{F}_j> + e<L'_j \tilde{G}>
\end{cases}
\end{equation}
\[(3.20) \begin{cases} 
<Q_j\tilde{G}> = (1-\epsilon)<\tilde{F}_j, \tilde{G}> + \epsilon<L_j\tilde{G}> \\
W_j' = (1-\epsilon)W_j + \epsilon\tilde{G}L_j'
\end{cases}\]

\[(3.21) \begin{cases} 
<Q'_j\tilde{G}'> = (1-\epsilon)<\tilde{G}, \tilde{F}_j'> + \epsilon<L_j\tilde{G}'> \\
W_j' = (1-\epsilon)W_j + \epsilon\tilde{G}L_j'
\end{cases}\]

where

\[(3.22) \quad W_j' = \tilde{G}Q'_j \text{ and } W_j = \tilde{G}Q_j.\]

**Proof:** We shall show that Lemma 3.2 holds by establishing the following sequence of implications:

\[(3.18) \Rightarrow (3.21) \Rightarrow (3.20) \Rightarrow (3.19) \Rightarrow (3.18).\]

If (3.18) is true, then from (3.22)

\[W_j' = (1-\epsilon)\tilde{G}\tilde{F}_j + \epsilon\tilde{G}L_j;\]

then from (3.12) we have

\[<\tilde{G}, Q'_j>(t) = \int_{\mathbb{R}} G(x)[-dQ'_j(x)] = (1-\epsilon)\int_{\mathbb{R}} G(x)[-d\tilde{F}_j(x)] + \epsilon\int_{\mathbb{R}} G(x)[-dL_j(x)] =
\]

\[= (1-\epsilon)<\tilde{G}, \tilde{F}_j>(t) + \epsilon<\tilde{G}, L_j>.\]

That is, we have shown that (3.18) implies (3.21).

If (3.21) is true, then, we have for any \(t \geq 0\)

\[\tilde{G}(t)Q'_j(t) = <Q'_j, \tilde{G}>(t) = (1-\epsilon)\tilde{G}(t)\tilde{F}_j(t) - (1-\epsilon)<\tilde{F}_j, \tilde{G}>(t) + \epsilon\tilde{G}(t)L_j(t) +
\]

\[+ \epsilon\tilde{G}(t)L_j(t) - \epsilon<L_j, \tilde{G}>(t).\]

Each side of this equation can be rewritten to give:
\(<Q_j^i, \tilde{G}> (t) - W_j^i (t) = (1-\varepsilon) <\tilde{F}_j, \tilde{G}> (t) + \varepsilon <L_j^i, \tilde{G}> (t) - W_j^i (t),
\)
or
\(<Q_j^i, \tilde{G}> = (1-\varepsilon) <\tilde{F}_j, \tilde{G}> + \varepsilon <L_j^i, \tilde{G}>.
\)

Therefore, we have shown that (3.21) implies (3.20).

Note that (3.20) and (3.21) have the same second equation, and the first equation in (3.20) can be obtained from the first equation in (3.21) by switching the ordering of the survival functions in every \(<,>\) expression. Thus, by the same method as used to show that (3.21) implies (3.20), we can show that (3.20) implies (3.19).

If (3.19) is true, then by adding two equations in (3.19) and by the decomposition of \(Q_j\), given by (3.17), we have
\[Q_j^i = (1-\varepsilon) \tilde{F}_j + \varepsilon L_j^i.\]

Thus, we have shown that (3.19) implies (3.18). Therefore, this completes the proof of lemma 3.2.

Next, we shall show that \((Q_0, Q_1)\) can not be a least favorable pair if the contamination distribution can be any distribution. We shall give sufficient and necessary conditions on the contamination distributions such that \((Q_0, Q_1)\) is a least favorable pair for \(n=1\). To simplify our notations, we define the two sets of real numbers as follows:
\[(3.23) \ A_t = \{s: Q_1 (s)/Q_0 (s) > t\}, \text{ and } B_t = \{s: q_1 (s)/q_0 (s) > t\}.
\]
Note that $A_t$ is the set of all real numbers such that the ratio of survival functions is greater than $t$, and $B_t$ is the set all real numbers such that the ratio of density functions is greater than $t$. Let us define the two sets of survival functions, $L_0^*$ and $L_1^*$, as follows:

(3.24) $L_0^* = \{ \text{survival functions } L_0^* : <\bar{G}, L_0^* > (B_k) + <L_0, \bar{G}>(A_k) \leq <\bar{G}, L_0>(B_k) + <L_0, \bar{G}>(A_k) \}$,

(3.25) $L_1^* = \{ \text{survival functions } L_1^* : <\bar{G}, L_1^* > (B_k) + <L_1, \bar{G}>(A_k) \geq <\bar{G}, L_1>(B_k) + <L_1, \bar{G}>(A_k) \}$,

where $k$ is determined by (3.26):

(3.26) $\alpha = \int_{B_k} \bar{G}(t) \mu_0(t) \, dt + \int_{A_k} g(t) \mu_0(t) \, dt = <\bar{G}, Q_0>(B_k) + <Q_0, \bar{G}>(A_k)$,

for $0 < \alpha < 1$.

From now on, $k$ is determined implicitly by (3.26) and we call $L_0^*$ and $L_1^*$ contamination distribution neighborhoods.

Theorem 3.1: For one observation, the pair $(Q_0, Q_1) \in (P_0, P_1)$ is a least favorable pair for censored data with respect to $\varphi$, defined by (3.10), if and only if the contamination neighborhoods $P_j$ are defined by (2.7), where $L_0^*$ and $L_1^*$ are defined by (3.24) and (3.25) respectively.
A proof of this theorem will be given in section V. From this theorem, we can see that although for uncensored data, the pair $(Q_0, Q_1)$ is a least favorable pair for the contamination neighborhoods that contain all possible contamination distributions, this same pair $(Q_0, Q_1)$ is not a least favorable pair for these same contamination neighborhoods for censored data. Therefore, if the contamination distribution can be any distribution, then the approach of using a least favorable pair to find a minimax test will not work in censored data. However, if the contamination distribution can be described by (3.24) and (3.25), then we can find a minimax test for censored data by using a least favorable pair.
CHAPTER IV
MIMIMAX TESTS.

In this Chapter, we shall propose a minimax test for the contamination neighborhoods defined by (3.24) and (3.25). Also, we shall simplify the contamination neighborhoods defined by (3.24) and (3.25) so we can see more clearly the structure of these contamination distributions. The structure of our contamination neighborhoods is very simple if the censoring is of type-I. The details of this result are given in section VI later. Our contamination neighborhoods are defined by

\[ P_j = \{ Q_j | Q_j = (1-\varepsilon)F_j + \varepsilon L_j : L_j \in L_j \} \quad j=0,1. \]

Proofs of results in this section will be given in Chapter V.

4.1. Minimax tests for one observation.

If \( n=1 \), then \( \varphi \), defined by (3.10), reduces to

\[ \varphi(t,\delta) = \begin{cases} 1 & \delta \ln(q_1(t)/q_0(t)) + (1-\delta)\ln(Q_1(t)/Q_0(t)) > \ln k \\ 0 & \text{Otherwise} \end{cases} \]

The next theorem says that \( \varphi(t,\delta) \) is a minimax test for the contamination distributions, defined by (3.24) and (3.25).
Let
\[ P_j^* = P_j \text{, when } L_j = L_j^*, j=0,1. \]

Then, we have the following theorem:

**Theorem 4.1**: The candidate test \( \varphi(t, \delta) \) is a minimax test for the contamination neighborhoods \( P_0^* \) versus \( P_1^* \) and

\[
\inf_{Q_1 \in P_1^*} E(Q_1; \tilde{G}) \varphi = E(Q_0; \tilde{G}) \varphi.
\]

**Proof**: By theorem 3.1 and lemma 4.2, it follows easily that \( \varphi(t, \delta) \) is a minimax test for the contamination neighborhoods, defined by (4.2). This completes the proof of theorem 4.1.

Because the restrictions on contamination distributions are complicated, we would like to simplify these conditions. We want to find such subsets of \( L_j^* \) that have a simpler structure than \( L_j^* \). Let us define these subsets as follows:

\[
\begin{align*}
L_2 &= \{ \text{survival functions } L_0' : & <\tilde{G}, L_0'(B_k) & \leq <\tilde{G}, L_0(0)>, \\
& & <L_0', \tilde{G}> (A_k) & \leq <L_0, \tilde{G}> (A_k) \}, \\
L_3 &= \{ \text{survival functions } L_1' : & <L_1', \tilde{G}> (A_k) & \geq <L_1, \tilde{G}> (A_k) \}, \\
L_4 &= \{ \text{survival functions } L_0' : & <L_0', \tilde{G}> (0) = <L_0, \tilde{G}> (0) \\
& & <L_0', \tilde{G}> (A_x) & \leq <L_0, \tilde{G}> (A_x) \text{ for } c' \leq x < c'' \},
\end{align*}
\]

and
(4.6) \[ L_5 = \{ \text{survival functions } L_1' : <L_1', \tilde{G}> (0) = <L_1, \tilde{G}> (0) \]
\[ <L_1', \tilde{G}> (A_x) \geq <L_1, \tilde{G}> (A_x) \text{ for } c' \leq x < c'' \}, \]

where

\[ (4.7) \quad c' = \min_{t \geq 0} \frac{Q_1(t)}{Q_0(t)}, \text{ and } c'' = \max_{t \geq 0} \frac{Q_1(t)}{Q_0(t)}. \]

\[ (4.8) \quad P_j = \{(1 - \epsilon) \tilde{F}_0 + \epsilon L_0' : L_0' \in L_j \}, \quad j = 2, 3, 4, 5. \]

The next proposition says that \( L_2 \) and \( L_4 \) are indeed subsets of \( L_0^* \), and \( L_3 \) and \( L_4 \) are subsets of \( L_1^* \).

**Proposition 4.1:** We have the following results:

1. \( Q_0 \in P_2 \cap P_4 \), and \( Q_1 \in P_3 \cap P_5 \).
2. \( L_0^* \supseteq L_2 \supseteq L_4 \), and \( L_1^* \supseteq L_3 \supseteq L_5 \),

where \( L_0^* \) and \( L_1^* \) are defined by (3.24) and (3.25) respectively.

Because \( L_2 \) and \( L_4 \) are the subsets of \( L_0^* \), and \( L_3 \) and \( L_4 \) are the subsets of \( L_1^* \), and \( Q_0 \in P_2 \cap P_4 \), and \( Q_1 \in P_3 \cap P_5 \), we have the following result.

**Proposition 4.2:** \( \phi \) is a minimax test for the contamination neighborhoods \( P_2 \) versus \( P_3 \), or \( P_4 \) versus \( P_5 \), defined by (4.8).
4.2. Minimax tests for more than one observation.

If there is more than one observation, then it can be shown that \((Q_0, Q_1)\), defined by (3.11), is not a least favorable pair in \((P_0^*, P_1^*)\). However, we shall show that \((Q_0, Q_1)\) is a least favorable pair in \(P_4^*\) and \(P_5^*\), defined by (4.8).

We summarize the above discuss in the following theorem:

**Theorem 4.2:** Let \(\varphi(t, \delta)\) be defined by (3.24). Then,

1. If \(n=1\), then \(\varphi(t, \delta)\) is a minimax test for the contamination neighborhoods \(P_0^*\) and \(P_1^*\).

2. If \(n>1\), then \(\varphi(t, \delta)\) is a minimax test for the contamination neighborhoods \(P_4^*\) and \(P_5^*\).

Next, we would like to illustrate the contamination neighborhoods \(P_0^*\) and \(P_1^*\) in Figure 5.1. Also, in Figure 5.1, we compare \(P_0^*\) and \(P_1^*\) with the contamination neighborhoods where a contamination distribution can be any distributions with densities. Finally, in Figure 3, we represent the relations of least favorable \((Q_0, Q_1)\) with idea survival functions \(\tilde{F}_j\).
Figure 3. Relationship of $Q_j$, $\tilde{F}_j$, $P_j^*$.  

The whole circle represents the set of all survival functions with densities; $P_0^*$ is the unshaded area of the whole circle on the left side; and $P_1^*$ is the unshaded area the circle on the right side.

The critical point in our minimax test is $k$, which is determined by (3.26). The next theorem gives a lower bound for this $k$.

Define $\bar{V}_1(G) = 1 - (Q_{jC}(R^+))^n$ and $\bar{V}_2(G) = 1 - (Q_{jU}(R^+))^n$,

and $b^* = \min_{t \geq 0} q_1(t)/q_0(t)$, and $b^{**} = \min_{t \geq 0} Q_1(t)/Q_0(t)$.  

Theorem 4.3: Let \( k \) be determined by (3.26).

If \( 0 < \alpha < V_1(G) \), then \( \text{In} k \geq b' \).

If \( 0 < \alpha < V_2(G) \), then \( \text{In} k \geq b'' \).

For one observation, the contamination neighborhoods \( L_0 \) and \( L_1 \) depend on the censoring distribution \( G \), and if there is no censoring, then \( L_j = \{ \text{all survival functions} \} \), \( j = 0, 1 \). However, the dependence of \( L_j \) on \( G \) is undesirable in applications. Next, we reduce \( L_0 \) and \( L_1 \) to some contamination neighborhoods that do not involve \( G \). The next corollary says that we can find a subset such that this subset does not involve \( G \). Define

\[
L_{20} = \{ \text{survival functions } L_0' : \quad -dL_0'(t)/d\mu \leq -dL_0(t)/d\mu \quad \text{for } t \in B_k \\
L_0'(t) \leq L_0(t) \quad \text{for } t \in A_k \}. 
\]

(4.39) \( L_{31} = \{ \text{survival functions } L_1' : L_1'(t) \geq L_1(t) \quad \text{for } t \in A_k \}. \)

Corollary 4.1:

\( L_2 \supseteq L_{20} \) and \( L_3 \supseteq L_{31} \).
CHAPTER V
PROOFS OF THEOREM 3.1 AND THEOREMS IN CHAPTER IV.

In Chapter III and IV, we gave our least favorable pair and minimax test results without proofs. In this Chapter we give the details of the proofs of those results. This Chapter is organized as follows:

First, we give a proof of theorem 3.1. Second, we show our minimax test result for only one observation. Third, we prove the minimax test results for censored data whose sample size is more than 2.

5.1. Proofs for one observation

The main task in proving the theorem 3.1 in Chapter III and the theorems in Chapter IV is to calculate $E(Q_0,\hat{G})\varphi(T,\delta)$, and $E(Q_0,\hat{G})\varphi(T,\delta)$. To calculate $E(Q_0,\hat{G})\varphi(T,\delta)$ we need to know the expressions for $<\hat{G},Q_1>(B_k)$ and $<\hat{G},Q_0>(B_k)$, defined by (3.23). The next lemma give the explicit expressions for $<\hat{G},Q_1>(B_k)$ and $<\hat{G},Q_0>(B_k)$.
Lemma 5.1: For any \( k \geq 0 \), we have
\[
\langle \tilde{G}, Q_1 \rangle(B_k) =
\begin{cases}
(1 - \varepsilon) \langle \tilde{F}_1, \tilde{G} \rangle(0) + \varepsilon \langle \tilde{G}, L_1 \rangle(0) & \text{if } k < c' \\
(1 - \varepsilon)(\langle \tilde{F}_1, \tilde{G} \rangle(B_k) + \varepsilon \langle \tilde{G}, L_0 \rangle(0)) & \text{if } c' \leq k < c'' \\
0 & \text{if } k \geq c''
\end{cases}
\]
\[
\langle \tilde{G}, Q_0 \rangle(B_k) =
\begin{cases}
(1 - \varepsilon) \langle \tilde{F}_0, \tilde{G} \rangle(0) + \varepsilon \langle \tilde{G}, L_0 \rangle(0) & \text{if } k < c' \\
0 & \text{if } k \geq c''
\end{cases}
\]

Proof: From (3.17) and the definition of \( \langle \rangle \), given by (3.12), it follows that
\[
\langle \tilde{G}, Q_0 \rangle = (1 - \varepsilon) \langle \tilde{F}_0, \tilde{G} \rangle + \varepsilon \langle \tilde{G}, L_0 \rangle, \quad j = 0, 1.
\]
If \( c' \leq k < c'' \), then from the expression for \( q_1(t)/q_0(t) \), (3.3), we have
\[
\{ t: q_1(t)/q_0(t) > k \} = \{ t: f_1(t)/f_0(t) \geq c'' \} \cup \{ t: k < f_1(t)/f_0(t) < c'' \} = \{ t: f_1(t)/f_0(t) > k \}.
\]

or
\[
B_k = \{ t: q_1(t)/q_0(t) > k \} = \{ t: f_1(t)/f_0(t) > k \} \text{ for } c' \leq k < c''.
\]
By the definition of \( h_0(t) \), given by (3.14), we know that \( h_0(t) > 0 \) if and if \( f_1(t)/f_0(t) > c'' \). From (2.10) we know that \( q_1(t)/q_0(t) = c'' \) for \( h_0(t) > 0 \). Thus, we have
(5.5) For \( k < c'' \), \( B_k = \{ t : q_1(t)/q_0(t) > k \} \supset \{ t : h_0(t) > 0 \} \).

So, for \( c' \leq k < c'' \), from (5.3) we have

\[
\langle \tilde{G}, L_0 \rangle (B_k) = \int_{B_k} \tilde{G}(t)h_0(t)\,d\mu = \int_{\{ t : h_0(t) > 0 \} \cap B_k} \tilde{G}(t)h_0(t)\,d\mu = 0 .
\]

So, from (5.3) we obtain

\[
\langle \tilde{G}, Q_0 \rangle (B_k) = (1 - \varepsilon) \langle \tilde{G}, \tilde{P}_0 \rangle (B_k) + \varepsilon \langle \tilde{G}, L_0 \rangle (0) .
\]

Thus, we have shown (5.2) for \( c' \leq k < c'' \).

Next, we shall show (5.1) for \( c' \leq k < c'' \). By the definition of \( h_1(t) \), defined by (3.14), and from (3.3), we know that

\[
\{ t : h_1(t) = 0 \} \supset \{ t : q_1(t)/q_0(t) > k \} \text{ for } c' \leq k < c'' ,
\]

which implies that

\[
\langle \tilde{G}, L_1 \rangle (B_k) = \int_{\{ t : q_1(t)/q_0(t) > k \}} \tilde{G}(t)h_1(t)\,d\mu = 0 .
\]

So, from (5.3) we have

\[
\langle \tilde{G}, Q_1 \rangle (B_k) = (1 - \varepsilon) \langle \tilde{G}, \tilde{P}_1 \rangle (B_k) ,
\]

which shows (5.1) that for \( c' \leq k < c'' \).

Finally, we shall that (5.1) and (5.2) are true for \( k < c' \) or \( k \geq c'' \). By (3.3) we know that

\[
c' \leq q_1(t)/q_0(t) \leq c'' \text{ for all } t \geq 0 .
\]

So,
Therefore, by (5.3) we know that (5.1) and (5.2) are true for \( k < c' \) or \( k \geq c'' \).

This completes the proof of lemma 5.1.

**Proof of theorem 3.1:** First, we express \( E(Q_0, \tilde{G})\varphi(T, \delta) \) in terms of \( <, > \) functions, defined by (3.12), then use the results from Lemma 5.1 to show theorem 4.1. To show theorem 4.1, we divide our proof into two parts. First, we show sufficient conditions, then we shall show necessary conditions.

For any \( Q_j^i \in P_j \) there exists a \( L_j^i \in L_j \) such that

\[
Q_j^i = (1 - \varepsilon) F_j + \varepsilon L_j^i,
\]

for \( j = 0, 1 \). By the definition of \( <, > \) functions, given by (3.19), we have that

\[
<\tilde{G}, Q_j^i> = (1 - \varepsilon) <\tilde{G}, F_j> + \varepsilon <\tilde{G}, L_j>^i,
\]

and

\[
<\tilde{G}, Q_j^i> = (1 - \varepsilon) <F_j, \tilde{G}> + \varepsilon <L_j, \tilde{G}>^i.
\]

Therefore,

\[
(5.7) \quad <\tilde{G}, Q_j^i>(B_k) = (1 - \varepsilon) <\tilde{F}_j, \tilde{G}> + \varepsilon <\tilde{G}, L_j>^i(B_k),
\]

and

\[
(5.8) \quad <Q_j^i, \tilde{G}>(A_k) = (1 - \varepsilon) <\tilde{F}_j, \tilde{G}>(A_k) + \varepsilon <L_j, \tilde{G}>(A_k).
\]

Therefore, by (5.7) and (5.8) we have that

\[
(5.9) \quad E(Q_0^i, \tilde{G})\varphi(T, \delta) = <\tilde{G}, Q_0^i>(B_k) + <Q_0^i, \tilde{G}>(A_k).
\]
Next, we shall show sufficient conditions in theorem 3.1. That is we shall show that if the contamination neighborhoods are given by $P_0^*$ and $P_1^*$, defined by (4.2), then $Q_0$ and $Q_1$ are a least favorable pair. To show that $Q_0$ and $Q_1$ are a least favorable pair, we need to show (3.5) to (3.7).

If $Q_0' \in L_0^*$, defined by (3.24), then by the definition of $L_0^*$, we have

$$\langle \tilde{G}, Q_0' \rangle (B_k) + \langle Q_0', \tilde{G} \rangle (A_k) \leq \langle \tilde{G}, Q_0 \rangle (B_k) + \langle Q_0, \tilde{G} \rangle (A_k).$$

Then, by (5.9) we have that

$$E(Q_0', \tilde{G}) \varphi(T, \delta) \leq E(Q_0, \tilde{G}) \varphi(T, \delta) = \alpha.$$ 

Thus, we have shown that (3.5) is true. Similarly, we show (3.6) is true.

Next, we shall show that (3.7) is true also. Let $\phi$ be any randomized test with maximal level $\alpha$, that is

$$\sup_{Q_0' \in P_0} E(Q_0', \tilde{G}) \varphi(T, \delta) \leq \alpha.$$ 

By the definition of $\varphi(t, 0)$ and $\varphi(t, 1)$, given by (3.10), and the fact $0 \leq \varphi(t, 1) \leq 1$ for any critical function, we have

$$\int (\varphi(t, 1) - \varphi(t, 1))(q_{1u}(t) - kq_0u(t))d\mu \geq 0 \quad (5.10)$$

and

$$\int (\varphi(t, 0) - \varphi(t, 0))(q_{1c}(t) - kq_0c(t))d\mu \geq 0 \quad (5.11)$$

where $q_{1u}(t) = -d<\tilde{G}, \tilde{F}_j>(t)/dt$, and $q_{1c}(t) = -d<\tilde{F}_j, \tilde{G}>(t)/dt$, $j = 0, 1$.

The sum of (5.10) and (5.11) gives
(5.12) \[ \int \phi(t,1)q_{1u}(t)d\mu + \int \phi(t,0)q_{1c}(t)d\mu - k[\int \phi(t,1)q_{0u}(t)d\mu + \int \phi(t,0)q_{0c}(t)d\mu] \geq 0. \]

\[ E(Q_j,\tilde{G})\phi(T,\delta) = \int \phi(t,1)q_{j_u}(t)d\mu + \int \phi(t,0)q_{j_c}(t)d\mu, \]

and

\[ E(Q_j,\tilde{G})\phi(T,\delta) = \int \phi(t,1)q_{j_u}(t)d\mu + \int \phi(t,0)q_{j_c}(t)d\mu, \]

substituting \( E(Q_j,\tilde{G})\phi(T,\delta) \) and \( E(Q_j,\tilde{G})\phi(T,\delta) \) in (5.12) gives us

(5.13) \[ E(Q_1,\tilde{G})\phi(T,\delta) - E(Q_1,\tilde{G})\phi(T,\delta) \geq k[E(Q_0,\tilde{G})\phi(T,\delta) - E(Q_0,\tilde{G})\phi(T,\delta)] \]

Because \( E(Q_0,\tilde{G})\phi(t,\delta) \leq \alpha = E(Q_0,\tilde{G})\phi(T,\delta) \),

\[ E(Q_1,\tilde{G})\phi(T,\delta) \geq E(Q_1,\tilde{G})\phi(T,\delta). \]

So, we have shown (3.7) is true also. Since we have now shown that all the hypotheses in the definition 3.1 hold, we have shown that \( (Q_0, Q_1) \) is a least favorable pair for censored data for contamination neighborhoods, \( P_0^* \) and \( P_1^* \), defined by (4.2) with respect to \( \phi \).

Finally, we show the necessary condition in theorem 4.1. That is we shall show if \( Q_0 \) and \( Q_1 \) are a least favorable pair for any contamination distribution neighborhoods \( L_0'' \) and \( L_1'' \), then \( L_0^* \supseteq L_0'' \), and \( L_1^* \supseteq L_1'' \). If \( Q_0 \) and \( Q_1 \) are a least favorable pair for any contamination distribution neighborhoods \( L_0'' \) and \( L_1'' \)
with respect to $\varphi$, then, from (3.5) in the definition 4.1, for any $Q_0' \in L_0''$ we have

$$E(\tilde{G},Q_0') \varphi(T,\delta) \leq E(\tilde{G},Q_0) \varphi(T,\delta).$$

Substituting (5.14) by (5.9) gives

$$<\tilde{G},Q_0'> (B_k) + <Q_0',\tilde{G}> (A_k) \leq <\tilde{G},Q_0> (B_k) + <Q_0,\tilde{G}> (A_k).$$

Therefore, by the definition of $L_0^*$, given by (3.24), we conclude that $Q_0' \in L_0^*$. Thus, we have shown that $Q_0' \in L_0^*$ whenever $Q_0' \in L_0''$, which implies that $L_0^* \supseteq L_0''$. Similarly, we show that $L_1 \supseteq L_1''$. This completes the proof of the theorem 3.1.

**Proof of proposition 4.1:** By the definitions of $L_j$, $j=2, 3, 4, 5$, it is easy to see that (1) of theorem 4.1 holds. For any $P_0' \in P_4$, there exists a $L_0' \in L_4$, defined by (4.5), such that

$$Q_0' = (1-\varepsilon) \tilde{F}_0 + \varepsilon L_0'.$$

Since $L_0' \in L_4$, $<L_0',\tilde{G}> (0) = <L_0,\tilde{G}> (0)$.

Note that $<L_0',\tilde{G}> (0) + <\tilde{G},L_0'> (0) = 1$. Therefore,

$$<\tilde{G},L_0'> (B_k) \leq <\tilde{G},L_0'> (0) = <L_0,\tilde{G}> (0).$$

Thus, $L_0' \in L_4$, which implies that $L_0' \in P_2$. So, $P_2 \supseteq P_4$.

Similarly, we show that $P_3 \supseteq P_5$, $P_1 \supseteq P_3$, and $P_0 \supseteq P_2$.

This completes the proof of proposition 4.1.
5.2. Proofs for more than one observation.

In this section, we shall provide proofs for the second part of Theorem 4.2 and Theorem 4.3. The strategy of the proofs is to write the power function of the candidate test in terms of random variables defined by individual survival functions. Then, we use the result for one observation to show the results for more than one observation.

First, we need to define the random variables through individual subsurvival functions. For any $Q_j \in P_{j+1}$, $j=0,1$, we define two subsurvival functions of $Q_j$ by $Q_j^u$ and $Q_j^c$ such that

\begin{align}
Q_j^u(t) &= \int q_j^u(y)\tilde{G}(y)d\mu(y), \\
Q_j^c(t) &= \int q_j^c(y)g(y)d\mu(y).
\end{align}

Now, we define the random variables $r_j^u(y)$, $r_j^c(y)$ as follows:

\begin{align}
V_j^u(x) &= Q_j^u(\ln(q_1(t)/q_0(t))>x), \quad r_j^u(y) = \sup\{ x | V_j^u(x) > y \} \\
V_j^c(x) &= Q_j^c(\ln(q_1(t)/q_0(t))>x), \quad r_j^c(y) = \sup\{ x | V_j^c(x) > y \},
\end{align}

$j=0,1$. Also, we define two constants

\begin{align}
\beta_j^1 &= \min(Q_j^u((R^+)), Q_j(u((R^+))), \beta_j^2 = \min(Q_j^c((R^+)), Q_j(c((R^+))).
\end{align}

For $n=1$, from (3.10) we can obtain the following expression:

\[
E(Q_j^u(G)\varphi(T,\delta) = \int \varphi(t,1)q_j^u(t)d\mu + \int \varphi(t,0)q_j^c(t)d\mu
\]
\[ = Q'_{ju}(\ln(q_1(t)/q_0(t)>x) + Q'_{jc}(\ln(Q_1(t)/Q_0(t)>x). \]

Substituting (5.17) and (5.18) in the right-hand side of the above equation gives
\[ E(Q'_j G) \phi(T, \delta) = V_{ju}(x) + V_{jc}(x), \quad j=0,1. \]

Applying the probability integral transformation theory to the subsurvival functions above, we obtain two independent random variables \( Z_{j1} \) and \( Z_{j2} \) which have uniform distributions on 
\[ [0, \beta_{j1}] \] and \( [0, \beta_{j2}] \) respectively, such that
\[ (5.20) \quad P\{ r'_{ju}(Z_{j1})>x\}=V_{ju}(x), \text{ and } P\{ r'_{ju}(Z_{j1})>x\}=V_{ju}(x) \]
\[ (5.21) \quad P\{ r'_{jc}(Z_{j2})>x\}=V_{jc}(x), \text{ and } P\{ r'_{ju}(Z_{j2})>x\}=V_{jc}(x). \]

The next result translates our minimax result for one observation into a result in terms of random variables \( r_{ju}(z), r'_{ju}(z), r_{jc}(z), \) and \( r'_{jc}(z), j=0,1. \)

**Lemma 5.2:** For \( r_{ju}(z), r'_{ju}(z), r_{jc}(z), \) and \( r'_{jc}(z), j=0,1, \) defined by (5.17) and (5.18), we have
\[ (5.22) \quad r_{0u}(z) \leq r'_0 u(z) \quad \text{for } 0 \leq z \leq \beta_{11}, \]
\[ r_{1u}(z) \leq r'_1 u(z) \quad \text{for } 0 \leq z \leq \beta_{21}, \]
and
\[ (5.23) \quad r_{0c}(z) \geq r'_0 c(z) \quad \text{for } 0 \leq z \leq \beta_{12}, \quad r_{1c}(z) \geq r'_1 c(z) \quad \text{for } 0 \leq z \leq \beta_{22}. \]
Proof: We only show the first inequality (5.22). The other inequalities follow similarly. From Proposition 4.2, we know that for \( Q_0' \in P_4 \),
\[ V_{0u}(x) \leq V_{0u}(x) \text{ for any } x \geq 0. \]
Thus,
\[ r_{0u}(z) = \sup \{ x | V_{0u}(x) > z \} \leq \sup \{ \ln c' \leq x | V_{0u}(x) > z \} = r_{0u}(z). \]
We have shown the first inequality of (5.22). Using the same method used to show (5.22), we can show (5.23) easily. This completes the proof of Lemma 5.2.

Now, we prove the second part of Theorem 4.2.

Proof of the second part of Theorem 4.2:

To emphasize that there are more than one observation in the second part of Theorem 4.2, we write \( y \) as \( y_i \), and \( \delta \) as \( \tilde{\delta} \) in this proof. By Definition 3.2 and Lemma 3.1, we know that \( \phi(T, \tilde{\delta}) \) is a minimax test procedure for \( P_4 \) versus \( P_5 \) if we can show that (3.5) to (3.7) hold. We first show (3.5) and (3.6).

For any \( Q_j' \in P_{j+5} \),

Let
\[ L'_{i}(y, \tilde{\delta}) = \prod_{i=1}^{n} q_{j}'(y_i) \delta_i Q_{j}'(y_i)^{1-\delta_i} g(y_i)^{1-\delta_i} c(y_i) \delta_i, \]
and
\[
D_k = \left\{ \sum_{i=1}^{n} \ln(q_1(y_i)/q_0(y_i)) + \sum_{i=m+1}^{n} \ln(Q_1(y_i)/Q_0(y_i)) \right\}^{n}
\]

Then, we have

\[
E(Q_0 \mathcal{G} \varphi(x, \delta)) = \sum_{(\delta_1, \ldots, \delta_n) = ((0,1))^n} \int \varphi(x, \delta) L'(y, \delta) d\mu(y_1) \ldots d\mu(y_n)
\]

\[
= \sum_{i=1}^{n} \left( \left( \begin{array}{c} n \\ m \end{array} \right) Q_{j_1} \otimes^m \mathbb{R} \otimes^m \mathbb{R} \otimes (n-m) \right)
\]

\[
\sum_{m=0}^{n} \left( \left( \begin{array}{c} n \\ m \end{array} \right) P( \sum_{i=1}^{m} r_{j_1 u}(Z_{j1i}) + \sum_{i=m+1}^{n} r_{j_1 c}(Z_{j2i})) \right) \]

where \(Z_{j11}, \ldots, Z_{j1n}\) are i.i.d. uniform random variables on \([0, \beta_{j1}]\), \(Z_{j21}, \ldots, Z_{j2n}\) are i.i.d. uniform random variables on \([0, \beta_{j2}]\), and \(Q^\otimes m\) is the m-fold product measure of \(Q\). Therefore, we can express the expectation for \(n>1\) observations as the following sum:

(5.24) \[
E(Q_0 \mathcal{G} \varphi(x, \delta)) = \sum_{m=0}^{n} \left( \left( \begin{array}{c} n \\ m \end{array} \right) P( \sum_{i=1}^{m} r_{j_1 u}(Z_{j1i}) + \sum_{i=m+1}^{n} r_{j_1 c}(Z_{j2i})) \right)
\]

Therefore, from (5.22) and (5.23) of Lemma 5.2, we know that

\[
E(Q_0 \mathcal{G} \varphi(x, \delta)) \leq E(Q_0 \mathcal{G} \varphi(x, \delta))
\]

and

\[
E(Q_0 \mathcal{G} \varphi(x, \delta)) \geq E(Q_0 \mathcal{G} \varphi(x, \delta))
\]

which implies (3.5) and (3.6).
Next, we shall show (3.7). Let $\phi$ be any randomized test procedure that satisfies (3.8). Then,

$$E_{\{(Q_0, G)\}} \phi(y, \delta) \leq \alpha.$$  

(5.25)

To show (3.7), we need to show that

$$E_{(Q_1, G)} \varphi(T, \delta) \geq E_{(Q_1, G)} \phi(T, \delta).$$

Define

$$L_j(y, \delta) = \prod_{i=1}^{n} q_j(y_i) \delta_i Q_j(y_i)^{1-\delta_i}, \quad j=0,1.$$ 

By the definition of $\varphi$ (3.10), we know that $\varphi = 1$ if and only if

$$L_1(y, \delta) - kL_0(y, \delta) > 0. \quad \text{That is,}$$

$$\varphi(y, \delta) = \begin{cases} 1 & \text{if } L_1(y, \delta) - kL_0(y, \delta) > 0 \\ 0 & \text{otherwise} \end{cases}.$$ 

Since $\phi$ is a critical function, $0 \leq \phi \leq 1,$

$$\int (\varphi(y, \delta) - \phi(y, \delta)) (L_1(y, \delta) - kL_0(y, \delta)) \, d\mu(y_1) \ldots d\mu(y_n) \geq 0.$$ 

Or,

$$E_{(Q_1, G)} \varphi(y, \delta) - E_{(Q_1, G)} \phi(y, \delta) \geq \exp(k) (E_{(Q_0, G)} \varphi(y, \delta) - E_{(Q_0, G)} \phi(y, \delta)) = k(\alpha - E_{(Q_0, G)} \phi(y, \delta)).$$

Thus, from (5.25) we obtain

$$E_{\{(Q_1, G)\}} \phi(y, \delta) \geq E_{\{(Q_1, G)\}} \phi(y, \delta).$$

Thus, we have shown that (3.7) hold.
Applying Lemma 3.1, we have shown that \( \phi \) is a minimax test when we have more than one observation. Thus, we have shown Theorem 4.2 for any number of observations.

**Proof of theorem 4.3:**

To show this theorem, we shall use (5.24) to express the expectation for \( n>1 \) observations as the sum of individual probabilities that each only depends on one observation. Then, we compare these expectations with the expression that

\[
E_{\{(Q_0, \tilde{G})\}} \phi(T, \delta) = \alpha.
\]

Finally, we show our results.

Since \( Q_j \in P_{4+j} \), substituting \( Q_j \) in (5.24) by \( Q_j \) gives

\[
E_{\{(Q_j, \tilde{G})\}} \phi(T, \delta) = \sum_{m=0}^{n} \binom{n}{m} P \left\{ \sum_{i=1}^{m} r_{ju}(Z_{j1i}) + \sum_{i=m+1}^{n} r_{jc}(Z_{j2i}) > \ln k \right\}.
\]

Taking \( j=0 \) and dropping the first term in the right-hand side of the equation above gives us that

\[
E_{\{(Q_0, \tilde{G})\}} \phi(T, \delta) \geq \sum_{m=1}^{n} \binom{n}{m} P \left( \sum_{i=1}^{m} r_{0u}(Z_{01i}) + \sum_{i=m+1}^{n} r_{0c}(Z_{02i}) > \ln k \right) =
\]

\[
= \sum_{m=1}^{n} \binom{n}{m} Q_{ju} \otimes_{m} Q_{jc} \otimes (n-m)
\]

\[
\sum_{i=1}^{m} \ln(q_{1}(y_i)/q_{0}(y_i)) + \sum_{i=m+1}^{n} \ln(Q_{1}(y_i)/Q_{0}(y_i)) > \ln k).
\]

For \( \ln(k) < b' = \min q_{1}(t)/q_{0}(t) \), we have
\[
\sum_{i=1}^{m} \ln(q_1(y_j)/q_0(y_j)) > \ln k, \text{ for any } y.
\]

Thus, for \( \ln(k) < b' \), we have that
\[
\left\{ y : \sum_{i=1}^{m} \ln(q_1(y_j)/q_0(y_j)) + \sum_{i=m+1}^{n} \ln(q_1(y_j)/q_0(y_j)) > \ln k \right\} = (R^+) \odot \Lambda.
\]

Thus, for \( \ln k < b' \) we obtain

\[(5.26) \quad E_{((Q_0, G))}(T, \delta) \geq \sum_{m=1}^{n} \binom{n}{m} (Q_{ju}(R^+))^m x(Q_{jc}(R^+))^{n-m}.\]

Since
\[
1 = (Q_{ju}(R^+) + Q_{jc}(R^+))^n = \sum_{m=0}^{n} \binom{n}{m} (Q_{ju}(R^+))^m x(Q_{jc}(R^+))^{n-m} = (Q_{jc}(R^+))^n + \sum_{m=1}^{n} \binom{n}{m} (Q_{ju}(R^+))^m x(Q_{jc}(R^+))^{n-m},
\]
moving the first term in the right-hand side to the left-hand side gives
\[
\sum_{m=1}^{n} \binom{n}{m} (Q_{ju}(R^+))^m x(Q_{jc}(R^+))^{n-m} = 1 - (Q_{jc}(R^+))^n.
\]

Then, from (5.26), we see that
\[
E_{((Q_0, G))}(T, \delta) \geq 1 - (Q_{jc}(R^+))^n = V_{1n}(\mathcal{G}) \text{ for } k < b'.
\]

The critical value \( k \) is determined by (3.26). That is,
\[
E_{((Q_0, G))}(T, \delta) = \alpha.
\]

Therefore, for \( \ln k < b' \), we have \( V_{1n}(\mathcal{G}) \leq \alpha \). By contradiction, if
\( \alpha < V_{1n}(\tilde{G}) \), then \( \lnk \geq b' \).

If \( \lnk < b'' \), then we have
\[
E\{Q(0, \tilde{G})\} \varphi(T, \delta) \geq \sum_{m=0}^{n-1} \binom{n}{m} (Q_ju(R^+))^m x(Q_jc(R^+))^{n-m} = 1 - Q_ju(R^+)^n = V_{2n}(\tilde{G}).
\]

However, \( k \) is determined by the equation:
\[
\alpha = E\{Q(0, \tilde{G})\} \varphi(T, \delta).
\]

Thus, by contradiction, we obtain that if \( \alpha < V_{2n}(\tilde{G}) \), then \( \lnk \geq b'' \).

This completes the proof of Theorem 4.3.
CHAPTER VI
TYPE-I CENSORING AND AN EXAMPLE.

In this Chapter, we describe the contamination neighborhoods and minimax tests for two simple censoring patterns: Type I censoring and exponential censoring.

6.1. Type I censoring:

Some experiments are conducted in such way that an individual's life time will be known exactly only if it is less than some predetermined value. If this happens, the data are said to be type I censored.

Suppose that we have n individuals and that associated with the \(i\)th individual is a life time \(T_i\) and a fixed censoring time \(L\). The data are represented by the \(n\) pairs of random variables \((Y_i, \delta_i)\), where

\[ Y_i = \min(T_i, L) \quad \text{and} \quad \delta_i = \begin{cases} 1 & T_i \leq L \\ 0 & T_i > L \end{cases} \]

Assume that the \(T_i\) are i.i.d. with survival function \(\bar{F}\). Then, the joint density probability function of \((Y_i, \delta_i), i=0,...,n\) is
(6.1) \[ J=\prod_{i=1}^{n} f(t_i)^{\delta_i} \bar{F}(L)^{1-\delta_i}, \]

where \( f \) is the density function of \( \bar{F} \).

As before, we are interested in finding a minimax test for
\[
H_0 : \{Q_0\} \in P_0 \text{ versus } H_1 : \{Q_1\} \in P_1,
\]
where the contamination neighborhoods \( P_j \) are defined by (2.7), \( j=0,1 \).

Next we shall define a least favorable pair \( Q_j \). Let the \( Q_j \)'s be defined in terms of their densities \( q_0 \) and \( q_1 \):
\[
(6.3) \quad q_0(t) = \begin{cases} 
(1-\varepsilon)f_0(t) & \text{ if } f_1(t)/f_0(t) < c'' \\
(1/c'')(1-\varepsilon)f_1(t) & \text{ if } f_1(t)/f_0(t) \geq c''
\end{cases}
\]
and
\[
(6.4) \quad q_1(t) = \begin{cases} 
(1-\varepsilon)f_1(t) & \text{ if } f_1(t)/f_0(t) > c' \\
c'(1-\varepsilon)f_0(t) & \text{ if } f_1(t)/f_0(t) \leq c'
\end{cases}
\]

where \( c' \) and \( c'' \) are determined so that the densities integrate to 1:
\[
\int_0^\infty q_1(t)d\mu(t) = \int_0^\infty q_0(t)d\mu(t) = 1.
\]

Let \( h_0(t) \) be defined by (3.14), and \( h_1(t) \) be defined by (3.15). It is easy to check that
\[
Q_j = (1-\varepsilon)\bar{F}_j + \varepsilon H_j, j=0,1.
\]
From Lemma 2 of Huber (1965) we know that the following inequalities hold for any $k \geq 0$:

\[(6.5)\quad Q_0^1 (q_1/q_0 \geq k) \leq Q_0 (q_1/q_0 \geq k) \leq Q_1 (q_1/q_0 \geq k) \leq Q_1^1 (q_1/q_0 \geq k).\]

Now we shall give the minimax test for a finite sample of size $n$. Define

\[(6.6)\quad H_0 = \{ H_0^1 : H_0^1(L) \leq H_0(L) \},\]

where $H_0$ is the survival function of $h_0$.

\[(6.7)\quad H_1 = \{ H_1^1 : H_1^1(L) \geq H_1(L) \},\]

where $H_1$ is the survival function of $h_1$.

\[(6.8)\quad \varphi(t,\delta) = \begin{cases} 1 & \text{if} \sum_{i=1}^{n} \delta_i (\ln(q_1(t_i))/q_0(t_i)) > k \\ 0 & \text{otherwise} \end{cases},\]

where $k$ is determined by

\[(6.9)\quad E(\xi_0,\xi_1) \varphi(T,\delta) = \alpha .\]

Employing the same method used to show Theorem 4.1, we show the following result:

**Theorem 6.1**: The $\varphi(t,\delta)$ defined by (6.8) and (6.9) is the minimax test for

\[P_0 \quad \text{versus} \quad P_1,\]

where $H_j$ are defined by (6.6) and (6.7, $j=0,1$).
If no censoring exists, then we have that $L=\infty$ and $\delta_i = 1$ for all $i=1,\ldots,n$. Therefore, if there is no censoring, then the maxmin test defined by (6.8) reduces to the following form:

$$
\varphi(t, \delta) = \begin{cases} 
1 & \text{if } \sum_{i=1}^{n} \ln(q_1(t_i))/q_0(t_i) > k \\
0 & \text{otherwise}
\end{cases}
$$

And $H_j$ defined by (6.6) and (6.7) are reduced to following form:

$$
H_j = \{\text{all survival functions}\}.
$$

Thus, the minimax test in Theorem 6.1 is simply the minimax test given in Huber (1965) if there exist no censoring.

If censoring does exist, then minimax tests exist for $P_0$ versus $P_1$ whenever the censoring pattern in $P_0$ and $P_1$ is different in the way that censoring on $P_0$ is lighter than that on $P_1$.

6.2. Example.

In this section we shall illustrate the contamination neighborhoods for which minimax tests exist. We shall also derive the minimal power and level if both the underlying life distributions and the censoring distributions are exponential.

First, we calculate the critical value of our test $k$ for fixed $\alpha$ by solving equation (3.26). To solve the equation (3.26), we need expressions fro the sets $A_k$ and $B_k$, given by (3.23). To find $A_k$ and $B_k$, we need the ratios of the densities $q_1(t)/q_0(t)$, given
by (3.3), and the ratio of survival functions $Q_1(t)/Q_0(t)$. Second, after we have $k$, we find the contamination neighborhoods $P_0^*$ and $P_1^*$. Third, we calculate the power of the minimax tests in $P_1^*$, and the power of the test at the ideal alternative $F_1$.

Assume

\[(6.10) \quad f_j(t)=\lambda_j \exp(-\lambda_j t), \quad j=0,1 \quad \text{for } t>0,\]

where $\lambda_0 > \lambda_1$, and

\[(6.11) \quad g(t)=\eta \exp(-\eta t), \quad t>0.\]

Then,

\[f_1(t)/f_0(t) = (\lambda_0/\lambda_1) \exp((\lambda_0-\lambda_1)t).\]

Therefore,

\[f_1(t)/f_0(t) > c'' \quad \text{if and only if } t > d'',\]

where

\[(6.12) \quad d''=\frac{1}{\lambda_0-\lambda_1} \ln(c''\lambda_1/\lambda_0),\]

and

\[f_1(t)/f_0(t) \leq c' \quad \text{if and only if } t \leq d',\]

where

\[(6.13) \quad d'=\frac{1}{\lambda_0-\lambda_1} \ln(c'\lambda_1/\lambda_0).\]

By the definitions of $h_0$ and $h_1$, defined by equations (3.14) and (3.15), we have

\[h_0(t)=\begin{cases} (1-\varepsilon)/\varepsilon (f_1(t)/c''-f_0(t)) & \text{for } t > d'' \\ 0 & \text{for } t \leq d'' \end{cases},\]

and
Therefore, from (6.12), (6.13), and (3.1) we get

(6.14) \[ q_0(t) = \begin{cases} (1 - \varepsilon)f_1(t)/c'' & \text{for } t > d'' \\ (1 - \varepsilon)f_0(t) & \text{for } t \leq d'' \end{cases} \]

and

(6.15) \[ q_1(t) = \begin{cases} (1 - \varepsilon)c'f_0(t) & \text{for } t \leq d' \\ (1 - \varepsilon)f_1(t) & \text{for } t > d' \end{cases} \]

Since the \( Q_j \) is the survival function of \( q_j \), it follows that

(6.16) \[ Q_0(t) = \begin{cases} \int_0^t (1 - \varepsilon)f_1(x)/c'' dx & \text{for } t > d'' \\ \int_t^{d''} (1 - \varepsilon)f_0(x)dx + (1 - \varepsilon)/c'' \int_{d''}^\infty f_1(x)dx & \text{for } t \leq d'' \end{cases} \]

Since \( q_0 \) is the density function, integrating from 0 to \( \infty \) in (6.14) gives

(6.17) \[ (1 - \varepsilon)/c'' \tilde{F}_1(d'') - (1 - \varepsilon)\tilde{F}_0(d'') = \varepsilon. \]

Substituting (6.17) in (6.16) gives

(6.18) \[ Q_0(t) = \begin{cases} (1 - \varepsilon)\tilde{F}_1(t)/c'' & \text{for } t > d'' \\ (1 - \varepsilon)\tilde{F}_0(t) + \varepsilon & \text{for } t \leq d'' \end{cases} \]
Similarly, we get

\[
Q_1(t) = \begin{cases} 
(1-\epsilon)\bar{F}_1(t) & \text{for } t > d' \\
(1-\epsilon)\bar{F}_0(t) - c'(1-\epsilon) + 1 & \text{for } t \leq d'. 
\end{cases}
\] 

Therefore, dividing (6.19) by (6.18) gives

\[
\frac{Q_1(t)}{Q_0(t)} = \begin{cases} 
c'' & \text{for } t > d'' \\
\frac{\bar{F}_1(t)}{\bar{F}_0(t) + \frac{\epsilon}{1-\epsilon}} & \text{for } d' < t \leq d'' \\
\frac{(1-\epsilon)\bar{F}_0(t) + 1 - c'(1-\epsilon)}{(1-\epsilon)\bar{F}_0(t) + \epsilon} & \text{for } t \leq d'. 
\end{cases}
\]

Let us assume \(\lambda_0=3.0, \lambda_1=2.0\) and \(\epsilon=0.1\). Note that \(c''\) and \(c'\) are defined by (2.12) and (2.13), respectively. Thus, from Example 2.1, we obtain

\[
(6.21) \quad c' = 0.945333, \quad c'' = 1.100642.
\]

Substituting the values of (6.21) for \(c'\) and \(c''\) in (6.13) and (6.21) gives

\[
(6.22) \quad d' = 0.3492471, \quad d'' = 0.5013588.
\]
Next we establish a lemma that is to be used to calculate $A_k$. The lemma is about the monotonicity of $Q_1(t)/Q_0(t)$.

**Lemma 6.1:** The ratio of $Q_1(t)/Q_0(t)$, given by (6.20), is a continuous and increasing function of $t$ in $[0, \infty)$.

**Proof:** Since $\tilde{F}_j(t) = \exp(-\lambda_j t)$ is a continuous function of $t$ in $[0, \infty)$, to show $Q_1(t)/Q_0(t)$ is a continuous function of $t$, we only need to show that $Q_1(t)/Q_0(t)$ is continuous at $t=d^n$ and at $t=d^r$.

We need to show that

$$\frac{\tilde{F}_1(d^n)}{\tilde{F}_0(d^n)+\epsilon/(1-\epsilon)} = c^n,$$

and

$$\frac{(1-\epsilon)c'\tilde{F}_0(d') + 1 - c'(1-\epsilon)}{(1-\epsilon)\tilde{F}_0(d') + \epsilon} = \frac{\tilde{F}_1(d')}{\tilde{F}_0(d') + \epsilon/(1-\epsilon)}.$$

Since $\int q_0(t)dt = 1$, by (6.14) we get

$$((1-\epsilon)/c'')\tilde{F}_1(d'') + (1-\epsilon)(1-\tilde{F}_0(d'')) = 1.$$

Or, $((1-\epsilon)/c'')\tilde{F}_1(d'') = (1-\epsilon)\tilde{F}_0(d'') + \epsilon$, which implies (6.23). Thus, we have shown (6.23).

Now, we show (6.24). Since $\int q_1(t)dt = 1$,
(1 - ε) \( \bar{F}_1(d') \) + c'(1 - ε)(1 - \( \bar{F}_0(d') \)) = 1.

Or,

(1 - ε) \( \bar{F}_1(d') \) = c'(1 - ε)\( \bar{F}_0(d') \) + 1 - c'(1 - ε).

Thus,

\[
\frac{(1 - ε)c'\bar{F}_0(d')+1-c'(1-ε)}{(1-ε)\bar{F}_0(d') + ε} = \frac{(1 - ε)\bar{F}_1(d')}{(1-ε)\bar{F}_0(d') + ε},
\]

which is (6.24). Thus, we have established the continuity of \( Q_1(t)/Q_0(t) \) in \([0, \infty)\).

Next, we shall demonstrate that \( Q_1(t)/Q_0(t) \) is an increasing function of \( t \) in \([0, \infty)\). To show that \( Q_1(t)/Q_0(t) \) is an increasing function of \( t \) in \([0, \infty)\) we need to prove that

\[
\frac{(1 - ε)c'\bar{F}_0(t)+1-c'(1-ε)}{(1-ε)\bar{F}_0(t) + ε}
\]

is an increasing function of \( t \) of \( t \) in \([0, d']\), and that

\[
\frac{\bar{F}_1(t)}{\bar{F}_0(t) + ε/(1-ε)}
\]

is an increasing function of \( t \) in \((d', d'')\).

If we show that the derivative of

\[
\frac{\bar{F}_1(t)}{\bar{F}_0(t) + ε/(1-ε)}
\]

is positive in \((d', d']\), then
\[ \frac{\tilde{F}_1(t)}{\tilde{F}_0(t) + \varepsilon/(1-\varepsilon)} \]

is an increasing function of \( t \) of \( t \) in \([d', d'']\). Note that

\[
\frac{d}{dt} \left( \frac{\tilde{F}_1(t)}{\tilde{F}_0(t) + \varepsilon/(1-\varepsilon)} \right) = \frac{6\exp(-5t)-2\exp(-2t)(\exp(-3t)+\varepsilon/(1-\varepsilon))}{(\exp(-3t)+\varepsilon/(1-\varepsilon))^2} \]

Since the first ratio term is always positive for \( t \geq 0 \), to show the above quantity is positive, we need to show that the second term is positive.

Since \( \min_0\leq t \leq d'' \exp(-3t) = \exp(-3d'') = 0.2222224 > 0.2222222 = \exp(-3t) - 2\varepsilon/(1-\varepsilon) \),

\[ (6.25) \quad \frac{d}{dt} \left( \frac{\tilde{F}_1(t)}{\tilde{F}_0(t) + \varepsilon/(1-\varepsilon)} \right) > 0, \quad \text{for } 0 \leq t \leq d'', \]

or

\[ \frac{\tilde{F}_1(t)}{\tilde{F}_0(t) + \varepsilon/(1-\varepsilon)} \]

is an increasing function of \( t \) in \([d', d'']\).

Using the same method as used above, we show that

\[ \frac{(1-\varepsilon)c\tilde{F}_0(t) + 1-c'(1-\varepsilon)}{(1-\varepsilon)c\tilde{F}_0(t) + \varepsilon} \]

is an increasing function of \( t \) in \([0, d']\).
Therefore, $Q_1(t)/Q_0(t)$ is an increasing function of $t$ in $[0, \infty)$. This completes the proof of the lemma.

We know that $k$, which is the critical cut point of our mini-max test, is determined by (3.26). We now find this $k$ for different $\alpha$ by solving the equation (3.26). To solve (3.26), we need to find the sets $A_k$ and $B_k$. To find these two sets, we need some lemmas. Let us define

$$k_1 = \frac{(1-\varepsilon) \exp(-2d')}{(1-\varepsilon) \exp(-3d') + \varepsilon}. \quad (6.26)$$

$$t_1(k) = -(1/\lambda_0) \frac{(1-\varepsilon) \ln(1-c')}{(1-\varepsilon) (k-c') (1-\varepsilon) - \varepsilon}. \quad (6.27)$$

$$t_2(k) \text{ is the solution of the following equation:}$$

$$\frac{(1-\varepsilon) \exp(-2t)}{(1-\varepsilon) \exp(-3t) + \varepsilon} - k = 0. \quad (6.28)$$

$$t_3(k) = 1/(\lambda_0 - \lambda_1) \ln(k\lambda_1/\lambda_0). \quad (6.29)$$

Since $\varepsilon=0.1$, $c'=0.94533$ and $c''=1.100642$, from (6.25) we get $k_1 = 1.076852$.

Therefore, we have shown that

$$c' < 1.0 < k_1 < c''. \quad (6.30)$$

The next lemma gives properties of $t_2(k)$. 

Lemma 6.2: The solution of (6.25) \( t_2(k) \) has the following properties:

(6.31) \( t_2(k_1) = d' \) and \( t_2(c'') = d'' \).

(6.32) \( t_2(k) \) is an increasing function of \( k \) for \( k_1 < k < c'' \).

(6.33) \( t_2(k) \) is the unique solution of

\[
\exp(-2t)/(\exp(-3t) + \varepsilon/(1-\varepsilon)) - k = 0,
\]

and \( d' < t_2(k) < d'' \) for \( k_1 < k < c'' \).

Proof: From (6.26) we see that

(6.34) \( \exp(-2d')/(\exp(-3d') + \varepsilon/(1-\varepsilon)) - k = 0. \)

Therefore, \( d' \) is the solution of (6.27) for \( k = k_1 \). Thus, \( t_2(k_1) = d' \).

Since \( \tilde{F}_j(t) = \exp(-\lambda_j t) \), substituting \( \tilde{F}_j(t) \) in (6.23) by \( \exp(-\lambda_j t) \) gives

(6.35) \( \exp(-2d'')/(\exp(-3d'') + \varepsilon/(1-\varepsilon)) - c'' = 0. \)

Thus, \( t_2(c'') = d'' \).

Substituting \( \tilde{F}_j(t) \) in (6.25) by \( \exp(-\lambda_j t) \) gives

\[
\frac{d}{dt} (\exp(-2t)/(\exp(-3t) + \varepsilon/(1-\varepsilon))) > 0 \quad \text{for } 0 \leq t \leq d''.
\]

Therefore, for any fixed \( k \),

\[
\exp(-2t)/(\exp(-3t) + \varepsilon/(1-\varepsilon)) - k
\]

is an increasing and continuous function of \( t \) in \([d', d'']\). From (6.34) we see that

\[
\exp(-2t)/(\exp(-3t) + \varepsilon/(1-\varepsilon)) \big|_{t = d'} = k_1.
\]

Thus, for \( k_1 < k < c'' \), we have

\[
\exp(-2t)/(\exp(-3t) + \varepsilon/(1-\varepsilon)) - k \big|_{t = d'} < 0.
\]
Since
\[ \frac{\exp(-2t)}{(\exp(-3t)+\varepsilon/(1-\varepsilon)) - k} \big|_{t=d''} = \frac{\exp(-2d'')}{(\exp(-3d'')+\varepsilon/(1-\varepsilon)) - k}, \]
from (6.35) we obtain
\[ \frac{\exp(-2t)}{(\exp(-3t)+\varepsilon/(1-\varepsilon)) - k} \big|_{t=d''} = c'' - k > 0 \text{ for } k < c''. \]
Thus, for \( k_1 < k < c'' \) the equation (6.29) has the unique solution \( t_2(k) \), and \( d' < t_2(k) < d'' \). This completes the proof of Lemma 6.2.

The next lemma characterizes the sets \( A_k \) and \( B_k \).

**Lemma 6.3:** We have the following expressions for \( A_k \) and \( B_k \):

\[ A_k = \begin{cases} \emptyset & \text{for } k \geq c'' \\ (t_2(k), \infty) & \text{for } k_1 \leq k < c'' \\ (t_1(k), \infty) & \text{for } 1 \leq k < k_1 \\ \mathbb{R}^+ & \text{for } k < 1 \end{cases} \]

\[ B_k = \begin{cases} \emptyset & \text{for } k \geq c'' \\ (t_3(k), \infty) & \text{for } c' \leq k < c'' \\ \mathbb{R}^+ & \text{for } k < c' \end{cases} \]

**Proof:** Lemma 6.1 tells us that \( Q_1(t)/Q_0(t) \) is an increasing function of \( t \). Therefore, from the expression for \( Q_1(t)/Q_0(t) \), given by (6.20) we see that if \( c' < 1 \leq k < k_1 \), then
\[ \{ t: Q_1(t)/Q_0(t) > k \} = (d', \infty) \cup \{ t \leq d': Q_1(t)/Q_0(t) > k \}. \]
If \( t \leq d' \), then
Thus, for \(1 \leq k < k_1\), we have shown that
\[
\{ t: \frac{Q_1(t)}{Q_0(t)} > k \} = (t_1(k), \infty).
\]
Using the same method used to prove the equation above, we show (6.36) for \(k \geq c''\) or \(k_1 < k < c''\) or \(k < 1\) and (6.37). This completes the proof of the lemma.

To calculate the critical value \(k\), we need to know the censoring distribution. Since an exponential distribution is commonly used for censoring distribution in statistical research, in our example, we assume the censoring distribution is an exponential distribution. Let us define
\[
G(t) = \exp(-\mu t), \quad \text{and} \quad w(\mu, k) = \int G(t) Q_0(t) dt + \int g(t) Q_0(t) dt.
\]

Recall the critical point \(k\) of our minimax test is determined by (3.26). Thus, from (6.39), we need to solve the following equation for \(k\):
\[
w(\mu, k) = \alpha \quad \text{for fixed } \mu \text{ and } \alpha.
\]

The next proposition gives an explicit expression for \(w(\mu, k)\).
Proposition 6.1: Let \( w(\mu,k) \) be defined by (6.39). Then,

\[
(6.41) \quad w(\mu,k) = \begin{cases} 
0 & \text{for } k \geq c'\prime \\
\frac{1-\varepsilon}{\lambda_0+\mu}(\lambda_0 E_3+\mu E_2)+\varepsilon \exp(-\mu t_2(k)) & \text{for } k_1 \leq k < c'' \\
\frac{1-\varepsilon}{\lambda_0+\mu}(\lambda_0 E_3+\mu E_1)+\varepsilon \exp(-\mu t_1(k)) & \text{for } 1 \leq k < k_1 \\
\frac{1-\varepsilon}{\lambda_0+\mu}(\lambda_0 E_3+\mu)+\varepsilon \exp(-\mu t_2(k)) & \text{for } c' \leq k < 1 \\
1 & \text{for } k < c' \end{cases}
\]

where

\( E_j = \exp(-(\lambda_0+\mu)t_j(k)) \), \( j = 1, 2, 3 \). For any fixed \( \mu \), \( w(\mu,k) \) is a decreasing function of \( k \).

Proof: Using (6.36) to substitute for the set over which the integral is integrated, we obtain

\[
(6.42) \quad \int_{B_k} G(t) q_0(t) dt = \begin{cases} 
0 & \text{for } k \geq c'\prime \\
\int_{t_3(k)}^{\infty} G(t) q_0(t) dt & \text{for } c' \leq k < c'' \\
\int_{0}^{\infty} G(t) q_0(t) dt & \text{for } k < c' \end{cases}
\]

Using (6.36) to substitute for \( A_k \) in the following integral gives
\[ g(t)Q_0(t)dt = \begin{cases} 
0 & \text{for } k \geq c' \\
\int_{t_2(k)}^\infty g(t)Q_0(t)dt & \text{for } k_1 \leq k < c'' \\
\int_{t_1(k)}^\infty g(t)Q_0(t)dt & \text{for } 1 \leq k < k_1 \\
\int_0^\infty g(t)Q_0(t)dt & \text{for } k < 1 
\end{cases} \]

Note that

\[ \int_a^\infty g(t)Q_0(t)dt + \int_a^\infty q_0(t)\hat{G}(t)dt = \hat{G}(a)Q_0(a). \]

Since \( d' \leq t_1(k) < d'' \) for \( c' \leq k \leq c'' \), adding (6.42) and (6.43) for \( w(.,.) \), given by (6.39), gives

\[ w(\mu,k) = \begin{cases} 
0 & \text{for } k \geq c' \\
\hat{G}(d'')Q_0(d'') + E_{5+} \int_{t_2(k)}^{d''} g(t)Q_0(t)dt & \text{for } k_1 \leq k < c'' \\
\hat{G}(d'')Q_0(d'') + E_{5+} \int_{t_1(k)}^{d''} g(t)Q_0(t)dt & \text{for } 1 \leq k < k_1 \\
\hat{G}(d'')Q_0(d'') + E_{5+} \int_0^{d''} g(t)Q_0(t)dt & \text{for } c' \leq k < 1 \\
1 & \text{for } k < c' 
\end{cases} \]

where
\[ E_5 = \int_{t_3(k)}^{d''} g(t)Q_0(t)\,dt. \]

Since \( t_3(k) < d'' \), substituting the right-hand side of (6.14) for \( q_0(t) \), we obtain
\[ E_5 = \int_{t_3(k)}^{d''} g(t)Q_0(t)\,dt = \int_{t_3(k)}^{d''} f_0(t)\ddot{G}(t)\,dt =
\]
\[ = (1 - \varepsilon) \int_{t_3(k)}^{d''} \lambda_0 \exp(-\lambda_0 t)\,dt = \frac{(1 - \varepsilon)\lambda_0}{\rho - \mu} \left( \exp(-\lambda_0 t_3) - \exp(-d'') \right). \]

Since \( t_2 < d'' \) and \( g(t) = \mu \exp(-\mu t) \) and \( \ddot{F}_0(t) = \exp(-\lambda_0 t) \), using the expression (6.16) for \( Q_0(t) \) and evaluating the following integral gives
\[ \int_{t_2(k)}^{d''} \lambda_0 \exp(-\lambda_0 t)\,dt = \int_{t_2(k)}^{d''} ((1 - \varepsilon)\ddot{F}_0(t) + \varepsilon g(t))\,dt =
\]
\[ = \varepsilon(\exp(-\mu t_2(k)) - \exp(-\mu d'')) + (1 - \varepsilon)(\mu/(\lambda_0 + \mu)\exp(-\lambda_0 t_2) - \exp(-d'')). \]

Adding the above two expressions gives us
\[ B(k) + \int_{t_2(k)}^{d''} g(t)Q_0(t)\,dt = (1 - \varepsilon)/(\lambda_0 + \mu)\lambda_0 \exp(-\lambda_0 t_3) +
\]
\[ + \mu \exp(-\lambda_0 t_2(k)) - (1 - \varepsilon)\exp(-d'') + \varepsilon(\exp(-\mu t_2(k)) - \exp(-\mu d'')). \]
Note that
\[ \tilde{G}(d'')Q_0(d'') = ((1-\varepsilon)\exp(-\lambda_0 d'') + \varepsilon)(\exp(-\mu d''). \]

Thus, using the above two expressions to substitute \( w(.) \) in (6.44) for \( k_1 \leq k < c'' \) gives
\[
w(\mu,k) = \frac{1-\varepsilon}{\lambda_0 + \mu}(\lambda_0 \exp(-\lambda_0 + \mu) t_3(k)) + \mu \exp(-\lambda_0 + \mu) t_2(k)) + \\
+ \varepsilon \exp(-\mu t_2(k)),
\]
for \( k_1 \leq k < c'' \). Thus, we have shown (6.41) for \( k_1 \leq k < c'' \). Using the same steps as used to show (6.41) for \( k_1 \leq k < c'' \), we can easily show other parts of the expression (6.41) for \( w(\mu,k) \).

This completes the proof of Proposition 6.1.

After we determine the critical point \( k \) of our test, we can look at the contamination neighborhoods \( P_0^* \) and \( P_1^* \), determined by (3.24) and (3.25) for fixed \( \mu \) and \( c' \). To determine \( P_0^* \) and \( P_1^* \), from (3.24), (3.25) and (4.2) we need to evaluate expressions \( \langle \ast, \rangle (B_k) \) and \( \langle \ast, \rangle (A_k) \), where the function \( \langle \ast, \rangle \) is defined by (3.12).

To find concrete solutions for \( P_0^* \) and \( P_1^* \), let us assume that \( \alpha = 0.04932 \) and \( \mu = 2.6 \), then, by using (6.41) and solving (6.40) for \( k \), we obtain \( k = 1.02 \). For \( k = 1.02 \), from (6.27) and (6.29) we obtain \( t_1(k) = 1.018 \) and \( t_3(k) = 0.425 \). Thus, from (6.36) and (6.37) in Lemma 6.3, we get
\[ A_k = (t_1(k), \infty) \text{ and } B_k = (t_3(k), \infty) \supseteq A_k. \]

Note that
\[ \langle \tilde{G}, L_0 \rangle (B_k) + \langle L_0, \tilde{G} \rangle (A_k) = \langle \tilde{G}, L_0 \rangle (t_3(k), t_1(k)) + 
\]
where \( l_0'(x) \) is the density of the survival function \( L_0'(x) \).

Thus, substituting \(<\tilde{G},L_0'>\)(B_k) + <\(L_0',\tilde{G}\>>(A_k)\) in (3.24) by (6.45) gives

\[
L_0^* = \{ L_0' \mid \int \tilde{G}(x)l_0'(x)dx + \tilde{G}(1.02)L_0'(1.02) \leq c_1 \},
\]

\[
c_1 = \int \tilde{G}(x)h_0(x)dx + \tilde{G}(1.02)L_0(1.02),
\]

Using the same method used to get (6.46), we can obtain an expression for \( L_1^* \):

\[
L_1^* = \{ L_1' \mid \int \tilde{G}(x)l_1'(x)dx + \tilde{G}(1.02)L_1'(1.02) \geq c_2 \},
\]

where

\[
c_2 = \int \tilde{G}(x)h_1(x)dx + \tilde{G}(1.02)L_1(1.02).
\]

Since from (6.22) we know that \( d'=0.3492471 \), by the expression (3.15) for \( h_1(t) \), we know that \( h_1(x)=0 \) for all \( x > d'=0.3492471 \).

Since \( 0.43 > d' \), \( h_1(x)=0 \) for all \( x > 0.43 \), which results in
\[ \int_0^{1.02} G(x)h_1(x)dx = 0, \]
\[ \int_{0.43}^{1.02} h(x)dx = 0. \]

and

\[ L_1(1.02) = \int_{1.02}^{\infty} h_1(x)dx = 0. \]

Thus, adding the equations above gives \( c_2 = 0 \). Since the left side of the inequality in (6.48) is always non-negative,

\[ L_1^* = \{ \text{all survival functions with densities} \}. \]

Thus,

\[ P_1^* = \{ Q_1' | Q_1'(t) = (1-\theta) \exp(-\lambda_1 t ) + L_1'(t) : L_1' \in L_1 \}. \]

Next, we want to see whether \( L_0^* \), defined by (6.35), contains some members of the exponential family,

\[ \{ \exp(-\eta x) , \eta > 0 \}. \]

Let \( L_0'(t) = \exp(-\eta t) \), then

\[ \int_{1.02}^{\infty} \tilde{G}(x)l_0'(x)dx = \frac{\eta}{\eta + \mu} \left[ \exp(-0.43(\eta + \mu)) - \exp(-1.02(\eta + \mu)) \right], \]

and

\[ \tilde{G}(1.02)L_0'(1.02) = \exp(-1.02(\eta + \mu)). \]

Thus, adding the two expressions above gives

\[ \int_{0.43}^{1.02} \tilde{G}(x)l_0'(x)dx + \tilde{G}(1.02)L_0'(1.02) = \frac{\eta}{\eta + \mu} \exp(-0.43(\eta + \mu)) + \]
\[ + \frac{\mu}{\eta + \mu} \exp(-1.02(\eta + \mu)). \]
Since \( \exp(-0.43(\eta+\mu)) \geq \exp(-1.02(\eta+\mu)) \), replacing
\( \exp(-1.02(\eta+\mu)) \) by \( \exp(-0.43(\eta+\mu)) \) in the above expression gives
\[
\int \tilde{G}(x)l_0'(x)dx + \tilde{G}(1.02)L_0'(1.02) \leq \exp(-0.43(\eta+\mu)).
\]
If \( \exp(-0.43(\eta+\mu)) \leq c_1 \), or
\[
\ln(c_1) \geq -0.43 - \mu,
\]
then it follows from (6.49) that
\[
\int \tilde{G}(x)l_0'(x)dx + \tilde{G}(1.02)L_0'(1.02) \leq c_1.
\]
Thus, by the expression (6.45) for \( L_0^* \), we see that
\[
L_0'(t) = \exp(-\eta t) \in L_0 \text{ if } \eta \text{ satisfies }
\[
\ln(c_1) \geq -0.43 - \mu.
\]
To get a lower bound for the set of values \( \eta \) for which \( \exp(-\eta t) \in L_0 \), we need to calculate \( c_1 \). Note that substituting \( Q_0 \) in (3.26) by the expression for \( Q_0 \) in (3.17) gives
\[
\alpha = (1-\varepsilon)(\langle \tilde{G}, F_0 \rangle(B_k) + \langle F_0, \tilde{G} \rangle(A_k)) + \varepsilon \langle \tilde{G}, L_0 \rangle(B_k) + \langle L_0, \tilde{G} \rangle(A_k).
\]
Substituting (7.47) for the right-hand side of (6.45) with \( L_0' \) replaced by \( L_0 \) gives
\[
c_1 = \langle \tilde{G}, L_0 \rangle(B_k) + \langle L_0, \tilde{G} \rangle(A_k)
\]
Thus, substituting (6.53) in (6.52) gives
\[ \alpha = (1 - \varepsilon) \left[ \int_{B_k} G(t) f_0(t) dt + \int_{A_k} g(t) F_0(t) dt \right] + \varepsilon c_1. \]

Therefore, using the equation above, we get \( c_1 = 0.04932 \) for \( \alpha = 0.04932 \) and \( \mu = 2.6 \). Thus,

\[ L_0(t) = \exp(-\eta t) \in L_0 \quad \text{for} \quad \eta \geq -\frac{\ln(c_1)}{0.43} - \mu = 4.56. \]

Then,

\[ P_1 \subseteq \{ Q_0 : Q_0(t) = (1 - \varepsilon) \exp(-\lambda_0 t) + \varepsilon \exp(-\eta t) : \eta \geq 4.56 \}. \]

Next, we calculate the minimal power of the minimax test over the contamination neighborhood \( P_1 \), and we call this power the minimal power. We shall compare this power with the power of the minimax test at the assumed alternative distribution, we shall call this power the ideal power.

Theorem 4.1 showed that the minimal power of the minimax test is given by

(6.54) \[ P(\mu, k) = \int_{B_k} G(t) q_1(t) dt + \int_{A_k} g(t) Q_1(t) dt. \]

The ideal power is given by

(6.55) \[ IP(\mu, k) = \int_{B_k} G(t) f_1(t) dt + \int_{A_k} g(t) F_1(t) dt. \]
Since we have expressions for $A_k$ and $B_k$, the same method used to evaluate the expression for $w(\mu, k)$ in the proposition 6.1 can be applied to obtain explicit expressions for $P(\mu, k)$ and $IP(\mu, k)$.

**Proposition 6.2:** Let $P(\mu, k)$ be defined by (6.54). Then, for any fixed $\mu$, $P(\mu, k)$ is a decreasing function of $k$ and

\[
(6.56) \quad P(\mu, k) = \begin{cases} 
0 & \text{for } k \geq c'' \vspace{0.5cm} \\
\frac{1-\epsilon}{\lambda_0 + \mu} (\lambda_1 E g + \mu E g) & \text{for } k_1 \leq k < c'' \vspace{0.5cm} \\
\frac{1-\epsilon}{\lambda_1 + \mu} (\lambda_1 E g + \mu E g) & \text{for } 1 \leq k < k_1 \vspace{0.5cm} \\
\frac{1-\epsilon}{\lambda_1 + \mu} (\lambda_1 E g) + E_{10} & \text{for } c' \leq k < 1 \vspace{0.5cm} \\
1 & \text{for } k < c' 
\end{cases}
\]

where $E_j = \exp(- (\lambda_1 + \mu)t_{j-6}(k))$, $j=7, 8, 9$, and

\[
E_{10} = (1-\epsilon)\mu/(\lambda_0 + \mu)(1-\exp(- (\lambda_0 + \mu)d')+(1-\epsilon)/(\lambda_1 + \mu)\exp(- (\lambda_1 + \mu)d')+
\quad + (1-c'(1-\epsilon))(1-\exp(-\mu d')).
\]
We plot \( w(\mu,k) \) versus \( k \) for different values of \( \mu \). The Fortran program to calculate the \( w(\mu,k) \) and critical \( k \) are in Appendix A.2. From Figure C.3 in Appendix C, we see that the critical point does not change too much as \( \mu \) varies. Particularly, for \( 0 < \alpha < 0.1 \), the critical point stays almost the same as \( \mu \) varies.

We plot the minimal power \( P(\mu,k) \) and the ideal power \( \text{IP}(\mu,k) \) versus \( k \) for 6 values of \( \mu \): 2.6, 3.0, .15, 1.0, 0.65 and 0.40. From Figure C.4 in Appendix C, we see that both the minimal power and the ideal power do not change too much as the mean of the censoring distribution varies. Also, the minimal power and the ideal power are close to each other.

From this example, we can conclude that our contamination neighborhoods contain rich enough distributions to include the possible contamination distributions. Also, we see that minimal power of our minimax test is robust against the censoring
distribution, even though we assume the censoring distribution is known.

CHAPTER VII

A ONE-STAGE RELIABILITY DEMONSTRATION TEST WITH KNOWN $\delta$.

In this Chapter, we shall describe the general testing problem for reliability of large series systems. Then, we propose a reliability demonstration test for known shape parameter $\delta$.

Suppose we have a system which has been in operation for a period of time $t$ and consists of $m$ components in series. We put this system on test for a time period $(t,t+s)$. Let $\mu =$ the component Mean Time To Failure (MTTF). The system has an equilibrium MTBF denoted by $\theta$, i.e. $\theta$ will represent MTBF for a "well aged" system. Component MTTF $\mu$ is regarded as fixed, but since components are replaced upon failure, system MTBF will in fact vary with the age $t$ of the system being tested. (The "age" of a series system would best be measured on a scale which reflects the distribution of component ages, and for identically distributed components, this can be done using $\mu r(t)$, where $r(t)$ is
the component renewal density). By the well-known renewal
theorem, we know that for a system in equilibrium, i.e. \( t=\infty \),
\( \theta=\mu/m \). We are interested in testing a hypothesis regarding the
equilibrium system MTBF \( \theta \), i.e. the goal is to test
(7.1) \( H_0: \theta=\theta_0 \quad \text{vs} \quad H_1: \theta=\theta_1 \) (\( \theta_1 < \theta_0 \))
with specified error probabilities \( \alpha \) and \( \beta \).

Now, let us put \( k \) new systems on test for a time period
\((0,s)\), where \( s \) does not depend on \( \theta \), and let \( N^k_0(s) \) be the number
of failures observed in the time period \((0,s)\). Let the component
failure distribution, \( F(x) \), be gamma, with density
\[ f(x)=dF(x)/dx = \frac{x^{\delta-1}}{\gamma^\delta \Gamma(\delta)} \exp(x/\gamma) , \delta \geq 1 , x \geq 0 \text{ and } \gamma > 0. \]
Then, the approximate distribution of \( N^k_0(s) \) is given by equation
(7.2) below (first published by Grigelionis(1964), also found in
equation (5.4) of Blumenthal, Greenwood, and Herbach (1973)).
(7.2) \[ P( N^k_0(s)=r ) = \frac{1}{r!}(E_\theta(N^k_0(s)))^r \exp(-E_\theta(N^k_0(s))) \text{ for large } n, \]
where \( n=km \).
(7.3) \[ E_\theta(N^k_0(s)) = nB(\delta)s^{\delta}/(m\theta)^\delta, \]
n=mk, and
(7.4) \[ B(\delta) = \frac{\delta^{\delta-1}}{\Gamma(\delta)}. \]

Next, we shall present the single stage test procedure for
known shape parameter \( \delta \), and demonstrate its sensitivity to the
assumed value of that parameter. By (7.2), (7.3) and the Neyman-Pearson Lemma, the most powerful test for \( H_0 \) versus \( H_1 \) is given by

\[
\text{accept } H_0 \text{ if } N_0^k(s) \leq c.
\]

For given \( \delta, m, \) and \( k, \) the values \( c \) and \( s \) are determined by the requirements

\[(7.5) \quad P_{\theta_0}(N_0^k(s) \leq c) \geq 1-\alpha \quad \text{and} \quad P_{\theta_1}(N_0^k(s) \leq c) \leq \beta,
\]

where \( \alpha<1-\beta. \) The OC-Curve of the test is given by

\[(7.6) \quad h_k(\theta, \delta, s) = P_{\theta}(N_0^k(s) \leq c) = \sum_{r=0}^{c} \frac{(1/r!)[E_{\theta}(N_0^k(s))]^r e^{-[E_{\theta}(N_0^k(s))]},}
\]

where \( E_{\theta}(N_0^k(s)) \) is given by (7.3). Note that for the Poisson

\[(7.7) \quad h_k(\theta, \delta, s) = P_{\theta}(N_0^k(s) \leq c) = P(\chi^2_{(c+1)} \leq 2E_{\theta}N_0^k(s)).
\]

Using (7.3) and (7.7), we can see that (7.5) is equivalent to the following requirement:

\[(7.8) \quad (\theta_1 m) \delta \chi^2_{(c+1)}(1-\beta) \leq 2n_1 B(\delta)(s) \delta \leq (\theta_0 m) \delta \chi^2_{(c+1)}(\alpha).
\]

From (7.8) we see that \( s \) can be chosen independently of \( \theta \) and that \( c \) must satisfy

\[(7.9) \quad (\theta_0/\theta_1) \delta \leq \chi^2_{(c+1)}(1-\beta)/\chi^2_{(c+1)}(\alpha).
\]

Next, we will give a way to choose \( c. \) From (7.9), we see that \( c \)
is a function of \( (\theta_0/\theta_1)^{\delta} \). Denote \( c = c((\theta_0/\theta_1)^{\delta}) \). For the increasing hazard rate distributions of interest to us, \( \delta > 1 \) and since the MTBF ratio \( (\theta_0/\theta_1) > 1 \), we see that \( (\theta_0/\theta_1)^{\delta} > (\theta_0/\theta_1) \).

Taking logs in (7.9), we find that \( c = i \) whenever \( a_i \leq \delta < a_{i-1} \), \( 0 \leq i \leq i^* \), where the \( a \)'s form a decreasing sequence, and \( a_{-1} = \infty \).

The \( a \)'s and \( i^* \) are defined by;

\[ i^* \text{ is the unique solution of } \]

\[ (7.10) \quad \chi_2^2(i^*+1)(1-\beta)/\chi_2^2(i^*+1)(\alpha) \leq (\theta_0/\theta_1) < \]

\[ < \chi_2^2(i^*)(1-\beta)/\chi_2^2(i^*)(\alpha), \]

and \( a_{i^*+1} \leq a_{i^*} < \ldots < a_0 < a_{-1} = \infty \), by

\[ (7.11) \quad a_i = (\ln[\chi_2^2(i+1)(1-\beta)/\chi_2^2(i+1)(\alpha)]/\ln[\theta_0/\theta_1])^{1/\delta}, \quad i=0,\ldots,i^*. \]

Thus, we have shown that \( c \leq i^* \), and

\[ (7.12) \quad c = c(\delta) = i \quad \text{if } \delta \in [a_i,a_{i-1}), \text{ for } i=0,\ldots,i^*. \]

Next we investigate what would happen to the OC-Curve, defined by (7.6), if we take either of the two extreme values of \( s \) given by (7.8).

**Lemma 7.1**: Let \( c \) be defined by (7.12). If \( s \) is defined by

\[ (7.13) \quad s = \theta_0 m(\chi_2^2(c+1)(\alpha)/2nB(\delta))^{1/\delta}, \]

then
(7.14) \[ h_k(\theta_0, \delta, s) = 1 - \alpha \]

and

\[ P(\chi^2_{2(c+1)} \geq V_c(1-\beta, \alpha) ) = \beta_1 \leq h_k(\theta_1, \delta, s) \leq \beta, \]

where

(7.15) \[ V_c(a, b) = \chi^2_{2(c)\{a\}}[\chi^2_{2(1+c)\{b\}}/\chi^2_{2(c)\{b\}}], \text{ for } c \geq 1, \]

\[ V_0(a, b) = \infty \text{ for } a > b, \]

\[ V_0(a, b) = 0 \text{ for } a < b. \]

If \( s \) is defined by

(7.16) \[ s = \theta_1 m(\chi^2_{2(c+1)(1-\beta)}/2nB(\delta))^1/\delta, \]

then,

(7.17) \[ h_k(\theta_1, \delta, s) = \beta, \]

and

\[ 1 - \alpha \leq h_k(\theta_0, \delta, s) \leq 1 - \alpha_1 = P(\chi^2_{2(c+1)} \geq V_c(\alpha, 1-\beta)), \]

where \( V_c(a, b) \) is defined by (7.15).

Proof of lemma 7.1:

From the definition of \( c \), we have

\[ \chi^2_{2(c+1)(1-\beta)}/\chi^2_{2(c+1)}(\alpha) \leq (\theta_0/\theta_1)^\delta < \chi^2_{2(c)(1-\beta)}/\chi^2_{2(c)}(\alpha), \]

Then by (7.3) we get that

\[ E_\theta N_0^k(s) = (1/2) (\theta_0/\theta)^\delta \chi^2_{2(1+c)}(\alpha). \]

Therefore,
Now, use relation (7.11), along with (7.18) to obtain (7.14). A similar argument gives (7.17), completing the proof.

**Remarks**: (A) The plan defined by the testing time \( s \), given by (7.13), protects consumers more than producers in the sense that satisfactory batches of MTBF \( \theta_0 \) are accepted with exactly the nominal probability \( 1-\alpha \) and that unsatisfactory batches of MTBF \( \theta_1 \) are only accepted with a probability that is less than the given number \( \beta \).

(B) The plan defined by the testing time \( s \), given by (7.16), protects producers more than consumers in the sense that unsatisfactory batches of MTBF \( \theta_1 \) are accepted with exactly the given probability \( \beta \) and that satisfactory batches of MTBF \( \theta_0 \) are accepted with a probability that is greater than the nominal \( 1-\alpha \).

To get an idea of the values of \( \beta_1 \) and \( \alpha_1 \), defined by (7.14) and (7.17) respectively, how far the values of the OC-Curve at \( \theta_0 \) differ from \( \alpha \) and \( \alpha_1 \), and how far the values of the OC-Curve at \( \theta_1 \) differ from \( \beta \), and \( \beta_1 \) we shall give an example.

**Example 7.1** \( \alpha=\beta=0.1 \) and \( \theta_0/\theta_1=2.0 \)

For different values of \( d \), we calculate values of \( c \) as a function of \( d \) from (7.12), and values of \( \beta_1 \) and \( \alpha_1 \) as functions of \( d \) from (7.14) and (7.17). Let \( s(0) \) be the value of \( s \) determined by

\[
(7.18) \quad (1/2) \chi^2_{2(1+c)}(1-\beta) \leq E_{\theta_1} N_0^k(s) \leq (1/2) V_0(1-\beta,\alpha).
\]
(7.13), $s(1)$ be the value of $s$ determined by (7.16), $OC(0)=h_k(\theta_0, \delta, s(0))$, and $OC(1)=h_k(\theta_1, \delta, s(1))$. Then the calculations are summarized in Table 2.

**Table 2. The values of Type-I and Type-II errors**

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>1.1</th>
<th>1.3</th>
<th>1.5</th>
<th>1.7</th>
<th>2.0</th>
<th>2.2</th>
<th>2.5</th>
<th>3.0</th>
<th>4.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c$</td>
<td>11</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$1-\alpha$</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>$OC(0)$</td>
<td>0.91</td>
<td>0.91</td>
<td>0.92</td>
<td>0.93</td>
<td>0.91</td>
<td>0.94</td>
<td>0.93</td>
<td>0.91</td>
<td>0.98</td>
</tr>
<tr>
<td>$1-\alpha_1$</td>
<td>0.91</td>
<td>0.92</td>
<td>0.93</td>
<td>0.93</td>
<td>0.95</td>
<td>0.95</td>
<td>0.96</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$OC(1)$</td>
<td>0.09</td>
<td>0.08</td>
<td>0.08</td>
<td>0.06</td>
<td>0.08</td>
<td>0.04</td>
<td>0.05</td>
<td>0.08</td>
<td>0.00</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.08</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>0.03</td>
<td>0.03</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Next, we shall study the OC-Curve, $h_k(\theta_0, \delta, s)$, when $s$ is chosen in between the two extreme values given by (7.13) and (7.16), and $c$ is given by (7.12) for fixed $\delta, \alpha$ and $\beta$.

Let us define a function of $r$ on $[0,1]$ as follows:
Let us choose a testing time $s(r; \delta)$ such that
\begin{equation}
R(r; \delta, c) = nB(\delta)(s(r; \delta)/m)^\delta \quad \text{for } 0 \leq r \leq 1.
\end{equation}
Since $0 \leq r \leq 1$, $s(r; \delta)$ satisfies (7.8). If $\delta \in [a_i, a_{i-1})$ for some $i$: $0 \leq i \leq i^*$, then we have
\begin{equation}
h_k(\theta, \delta, s(r; \delta)) = P\left( \chi^2_{2(1+i)} \geq (1/\theta)^{\delta} R(r; \delta, i) \right).
\end{equation}

The next lemma summarizes some properties of $h_k(\theta, \delta, s(r; \delta))$.

\textbf{Lemma 7.2:} For any fixed $r$: $0 \leq r \leq 1$, we have
(A) The OC-Curve $h_k(\theta, \delta, s(r; \delta))$ is a continuous function of $\delta$ in $(a_i, a_{i-1})$ for fixed $i$: $0 \leq i \leq i^*$.
(B) $h_k(\theta, a_i + \Delta, s(r; \delta)) \rightarrow h_k(\theta, a_i, s(r; \delta)) = P\left( \chi^2_{2(1+i)} \geq (1/\theta)^{\delta} R(r; \delta, i) \right)$ as $\Delta \rightarrow 0^+$,

and
\begin{equation}
h_k(\theta, a_i + \Delta, s(r; \delta)) \rightarrow P\left( \chi^2_{2(2+i)} \geq (1/\theta)^{\delta} R(r; \delta, i+1) \right) \quad \text{as } \Delta \rightarrow 0^-.
\end{equation}
(C) For fixed $\theta$, $\delta$ and $k$, $h_k(\theta, \delta, s(r; \delta))$ is an increasing function of $r$ on $[0,1]$.

\textbf{Proof:} For any $\delta \in (a_i, a_{i-1})$ for some $i$: $0 \leq i \leq i^*$, there exists $\Delta_0 > 0$ such that for all $\Delta$: $\Delta < \Delta_0$, we have $\delta + \Delta \in (a_i, a_{i-1})$. So, by (7.21) we have
\begin{equation}
h_k(\theta, \delta + \Delta, s(r)) = P\left( \chi^2_{2(i+1)} \geq (1/\theta)^{\delta + \Delta} R(r; \delta + \Delta, i) \right).
\end{equation}

Since
\[ R(r;\delta+\Delta,i) \to R(r;\delta,i) \text{ as } \Delta \to 0, \]

by continuity of the \( \chi^2 \) distribution, (A) follows.

If \( \delta = a_i \), then

\[ h_k(\theta,\delta+\Delta,s(r)) = P(\chi^2_{2(1+c)} \geq (1/\theta)\Delta+\delta R(r;\delta+\Delta,i)) \]

for small \( \Delta > 0 \), and

\[ h_k(\theta,\delta+\Delta,s(r)) = P(\chi^2_{2(1+c)} \geq (1/\theta)\Delta+\delta R(r;\delta+\Delta,i+1)) \]

for small \( \Delta < 0 \). Taking limits on the right sides of the above equalities gives (B).

Since \( \theta_0 > \theta_1 \), \( R(r;\delta,i) \) is a decreasing function of \( r \) in \([0,1]\), which with (7.21) gives (C). This completes the proof of Lemma 2.2.

Next we shall investigate the robustness of the proposed tests to departure from the assumed value of the shape parameter. Let \( c = c((\theta_0/\theta_1)^\delta \) be given by (7.12) and \( s \) be given by (7.13).

If \( \delta \) is known, using these values of \( c \) and \( s \) assures the desired OC-Curve. Now, let \( c \) and \( s \) be determined using \( \delta = \delta^* \), some assumed value, so that when \( \delta = \delta^* \), the desired OC-Curve is achieved. Denote the values of \( c \), and \( s \) which correspond to \( \delta^* \) as \( c^* \), and \( s^* \) respectively.

**Theorem 7.2:** The OC-Curve \( h_k(\theta,\delta,s(r)) \) has the following properties:

(i) The OC-Curve is an increasing function of \( \delta \) for \( \delta \geq 1 \), if
\( \theta \geq \delta s^*/m. \)  

(ii) The OC-Curve is a decreasing function of \( \delta \geq 1, \) if

\( \theta \leq \sqrt{\delta} s^*/m. \)

**Proof.**

Taking logs in (7.7) gives us that

\[
\ln E_0^k(N_0^k(s^*)) = \delta \ln (s^*/(m\theta)) + \ln(n) + \ln(\delta) - \ln \Gamma(\delta),
\]

and differentiating,

\[
\frac{d\ln E_0^k(N_0^k(s^*))}{d\delta} = \ln(s^*/(m\theta)) + \ln(\delta) + \frac{(\delta-1)/\delta}{\psi(\delta)},
\]

where

\[
\psi(\delta) = \frac{d \ln \Gamma(\delta)}{d \delta}.
\]

From the series expansions given in Abramowitz and Stegun (1965, Sections 6.1.41, 6.1.42, and 6.3.18), we have

\[
-\frac{1}{360\delta^3} < \ln \Gamma(\delta) - [\frac{1}{2}\ln \delta - \frac{1}{2}\ln 2\pi + \frac{1}{12\delta}] < 0,
\]

\[
0 < \delta \{\psi(\delta) - [\ln \delta - \frac{1}{2\delta} - \frac{1}{12\delta^2}]\} < \frac{1}{120\delta^3}.
\]

Therefore, we have

\[
0 < \frac{d\ln E_0^k(N_0^k(s^*))}{d\delta} - [\ln(s^*/(m\theta)) + 1 - 1/(2\delta)] < 1/(12\delta^2).
\]

Since \( \delta \geq 1, \) \((1/2\delta) \leq (1/2),\) and

\[
(7.24) \quad \ln(s^*/(\theta m)) + 1/2 < \frac{d\ln E_0^k(N_0^k(s^*))}{d\delta} < \ln(s^*/(m\theta)) + 1.
\]
The upper bound in (7.24) will be negative whenever \((s^*/(m\theta))\) < \((1/\theta)\). Thus, \(E_0(N_0^k(s^*))\) is decreasing in \(\delta\) when (7.22) holds.

From (7.7), we see that when \(E_0(N_0^k(s^*))\) decreases, the OC-Curve increases, showing (i).

Requiring that the lower bound in (7.24) be positive and arguing as above gives (ii), and completes the proof.

Remark: The implication of Theorem 2.1 is that for large \(\theta\), the OC-Curve is below the nominal one if \(\delta<\delta^*\) and above the nominal OC-Curve if \(\delta>\delta^*\), and vice versa for small \(\theta\). Thus, for any given \(\delta\), the actual and nominal OC-Curves cross somewhere. However, for large \(n\), the crossover point may be at a very small \(\theta\) value, hence for all practical purposes, the OC-Curve will shift either up or down with \(\delta\) for all \(\theta\).

Example 7.2. Consider systems consisting of 25 components, each with a gamma failure distribution having unknown shape parameter. Suppose \(k=13\), and \(n=325\). Assume that the desired ratio \(\theta_0/\theta_1\) is 1.25, and let \(\alpha=\beta=0.1\). Then, from Theorem 7.1 and Table 1 of Zhou and Blumenthal (1990), we obtain the following.

(i) For \(\delta^*=1.5\), \((s^*/m)=0.2323\ \theta_0\), and the bound in (7.22) can be rewritten as \((\theta_0/\theta)\leq1.58\). So, for all \(\theta\) such that \(0<\theta_0/\theta\leq1.58\),
the OC-Curve shifts up if $\delta > 1.5$ and down if $\delta < 1.5$.

(ii) For $\delta^* = 5$, as above, we see that for $\theta$ such that $0 < \theta_0 / \theta \leq 1.79$, the OC-Curve shifts up if $\delta > 5$ and down if $\delta < 5$.

When we are comparing the OC-Curve for a specified $\delta$ with that for $\delta^*$, a stronger statement can be made about the range of $\theta$ values for which one curve lies above or below the other than given by Theorem 8.1. To emphasize the role of $\delta$ in computing $E_0(N_k^0(s))$ from (7.6), we use the notation $E_{(\theta, \delta)}(N_k^0(s))$, so that $E_{(\theta, \delta^*)}(N_k^0(s^*))$ represents the expectation when the assumed value $\delta^*$ is correct. From (7.7), we can see that

\[ h_k(\theta, \delta, s) > (<) h_k(\theta, \delta^*, s^*) \]

if

\[
\frac{E_{(\theta, \delta)}(N_k^0(s^*))}{E_{(\theta, \delta^*)}(N_k^0(s^*))} < (> 1.
\]

Using (7.3), a direct evaluation the ratio of the expected value yields:

(a) If $\delta < \delta^*$, then $h_k(\theta, \delta, s^*) < (>) h_k(\theta, \delta^*, s^*)$, whenever $\theta < (>) (s^*/m) [(B(\delta^*)/B(\delta))^{1/(\delta^*-\delta)}]$,

(b) If $\delta > \delta^*$, then $h_k(\theta, \delta, s^*) < (>) h_k(\theta, \delta^*, s^*)$, whenever $\theta < (>) (s^*/m)[(B(\delta)/B(\delta^*))^{1/(\delta-\delta^*)}]$. 
Returning to example 7.1 above, for (i) we find that for \( \delta = 1 \),
\[ h_k(\theta, 1, s^*) < h_k(\theta, 1.5, s^*) \]
for all \( \theta \) such that \( 0 < \theta_0/\theta < 2.25 \), and for
\( \delta = 2 \),
\[ h_k(\theta, 1, s^*) < h_k(\theta, 1.5, s^*) \]
for all \( \theta \) such that \( 0 < \theta_0/\theta < 2.05 \).

For (ii) we find that if \( \delta = 4 \),
\[ h_k(\theta, 4, s^*) > h_k(\theta, 5, s^*) \]
for all \( \theta \) such that \( 0 < \theta_0/\theta \leq 1.99 \), and for \( \delta = 6 \),
\[ h_k(\theta, 6, s^*) > h_k(\theta, 5, s^*) \]
for all \( \theta \) such that \( 0 < \theta_0/\theta \leq 1.95 \).

An alternative way to look at the OC-Curve is to ask what value of \( (\theta_0/\theta) \) will give a particular value for the OC-Curve. Denote the value of the ratio \( (\theta_0/\theta) \) for which \( h_k(\theta, \delta^*, s^*) = \pi \) by \( (\theta_0/\theta)_{\delta^*} \) and the value such that \( h_k(\theta; \delta, s) = \pi \), by \( (\theta_0/\theta)_{\pi, \delta} \). Using
(7.3), (7.7) and (7.13) gives
(7.25) \[ (\theta_0/\theta)_{\delta^*} = \left( \chi_2^*(c^*+1)(1-\pi) / \chi_2^*(c^*+1)(\alpha) \right)^{1/\delta^*}, \]
(7.26) \[ \left( (\theta_0/\theta)_{\pi, \delta} / (\theta_0/\theta)_{\delta^*} \right) = \frac{B(\delta^*)^{1/\delta^*} 2}{B(\delta)^{1/\delta} (\chi_2^*(c^*+1)(1-\pi)/2\pi)^{1/\delta - 1/\delta^*}}. \]

For each of the \( \delta^* \) values of example 7.1, Figure 7 in Appendix C gives the OC-Curves for \( \delta = 1, 1.5 \) and 2 when \( \delta^* = 1.5 \) and \( \theta_0/\theta_1 = 1.25 \). Figure 7 show that the shift in the OC-Curve can be rather large for small errors in \( \delta \). Figure 8 in Appendix gives OC-Curves for \( \delta = 4, 5 \) and 6 when \( \delta^* = 5 \) and \( \theta_0/\theta_1 = 1.25 \). Again, Figure 8 show that the shift in the OC-Curve can be rather large for small errors in \( \delta \). Therefore, if we do not know
the value of $\delta$, we cannot use the testing procedure described in this section.

CHAPTER VIII

A TWO-STAGE RELIABILITY DEMONSTRATION TEST WITH UNKNOWN $\delta$.

Next, we shall develop a new testing procedure to use when the shape parameter $\delta$ is unknown. If the shape parameter is unknown, then we use a two-stage test procedure to test the hypothesis (7.1). We shall estimate $\delta$ from the first stage of sampling and use this estimate in place of $\delta$ in the second stage of sampling. The systems tested in the second stage are new and different from the systems tested in the first stage. One difference between our two-stage test procedure and the one-stage test procedure of Chapter VII is that the two-stage test procedure can be applied only to the first two of the three cases, given in section 1.4 of Chapter I. Next, we shall define an estimate of $\delta$, $\hat{\delta}_k$, and study its asymptotic properties.
8.1. An estimator of \( \delta \).

Put \( k \) newly minted systems, each system consisting of \( m \) components, on test and let \( s^* \) be the testing time, defined by

\[
s^* = m \theta_0 / \ln(k)
\]

Note that this testing time is longer than the testing time of Chapter VII. Let \( N_0^k(s^*) \) be the number of failures observed during time period \((0,s^*)\), and let \( n = km \). For large \( k \), \( N_0^k(s^*) \) has approximately a Poisson distribution with the mean

\[
E_\theta N_0^k(s^*) = [B(\delta)/(m \theta)]^\delta (s^n \theta^{-1/\delta})^\delta \theta = nB(\delta)(s^*)^\delta / (m \theta)^\delta = g_k(\delta, \theta).
\]

That is, approximately

\[
P[ N_0^k(s^*) = r ] = (1/r!)(E_\theta N_0^k(s^*))^r \exp(-E_\theta N_0^k(s^*)).
\]

Define

\[
g_k(\delta) = (nB(\delta)/\ln(k))^\delta = (n\theta^{-1}/\Gamma(\delta)(\ln(k))^\delta).
\]

Then, we have

\[
g_k(\delta, \theta) = (\theta_0/\theta)^\delta g_k(\delta)
\]

Now, let us define an estimate of \( \delta \) as follows:

\[
\hat{\delta}_k = \begin{cases} 
  1 & \text{if } N_0^k(s^*) > \frac{km}{\ln k} \\
  g_k^{-1}(N_0^k(s^*)) & \text{if } N_0^k(s^*) \leq \frac{km}{\ln k}
\end{cases}
\]

**Theorem 8.1:** If \( \theta = \theta_0 \), then \( \hat{\delta}_k \) is the M.L.E of \( \delta \).

See Chapter IX for a proof.
Note that if \( N_0^k(s^*) = 0 \), then \( \delta_k = \infty \), which would create some difficulty in application and in our simulation study later. Thus, we would like to modify \( \delta_k \) to make it bounded. The modified estimate is given as follows

\[
\delta_k = \begin{cases} 
1 & \text{if } N_0^k(s^*) > g_k(1) \\
g_k^{-1}(N_0^k(s^*)) & \text{if } 0 < N_0^k(s^*) \leq g_k(1) \\
m_k^* & \text{if } N_0^k(s^*) = 0 
\end{cases}
\]

where \( m_k^* \) is the smallest integer \( \geq g_k(1) \).

**Remark:** The reason for truncating \( \delta_k \) from above by \( m_k^* \) is that we want our estimate, \( \delta_k \), to be a nonincreasing function of \( N_0^k(s^*) \).

Hereafter, we use \( \delta_k \) defined by (8.6).

**Theorem 8.2:** For any fixed \( \theta \), we have

\[
\hat{\delta}_k \overset{P}{\longrightarrow} \delta \quad \text{as } k \to \infty.
\]

See the Chapter IX for a proof.

The next theorem says that scaled \( \hat{\delta}_k \) is asymptotically distributed as a normal random variable for \( \delta > 1 \).

**Theorem 8.3:** Define

\[
\mu_k(\delta; \theta) = \delta + \left[ \delta \ln(\theta / \theta_0) / (\ln(\ln k) - 1 + (1/2 \delta)) \right],
\]

and

\[
\sigma_k(\delta; \theta) = \left( 1 / \sqrt{g_k(\delta; \theta) \left( \ln(\ln k) - 1 + (1/2 \delta) \right)} \right).
\]

Then, as \( k \to \infty \),
(a) for \( \delta > 1 \), we have
\[
\frac{[\hat{\delta}_k - \mu_k(\delta; \theta)]/\sigma_k(\delta; \theta)]}{L} \rightarrow N(0,1)
\]
(b) for \( \delta = 1 \), we have
\[
\frac{\hat{\delta}_k - \mu_k(1; \theta)}{\sigma_k(1; \theta)} \rightarrow \begin{cases} N(0,1) & \text{if } \theta < \theta_0, \\ \min(X,0) & \text{if } \theta = \theta_0, \\ 0 & \text{if } \theta > \theta_0 \end{cases}
\]
where \( X \) is a \( N(0,1) \) random variable.

**Remarks**: As \( k \to \infty \), for \( \delta > 1 \),
1. If \( \theta = \theta_0 \), \( \frac{[\hat{\delta}_k - \delta]/\sigma_k(\delta)]}{L} \rightarrow N(0,1) \), where \( \sigma_k(\delta) = \sigma_k(\delta; \theta_0) \).
2. If \( \theta \neq \theta_0 \),
\[
\frac{[\hat{\delta}_k - \delta]/\sigma_k(\delta; \theta)] + \delta \ln(\theta/\theta_0)\sqrt{g_k(\delta; \theta)}}{L} \rightarrow N(0,1),
\]
and
\[
\ln(\theta/\theta_0)\sqrt{g_k(\delta; \theta)} \rightarrow \begin{cases} \infty & \text{if } \theta > \theta_0, \\ -\infty & \text{if } \theta < \theta_0 \end{cases}
\]
See Chapter IX for a proof.

**8.2. A two-stage test procedure**

Having derived an estimate of \( \delta, \hat{\delta}_k \), we shall give our two-stage test procedure when the shape parameter \( \delta \) is unknown. Let \( \hat{\delta}_k \) be the estimate defined by (8.6). As in (7.12), define
\[
\hat{\delta}_k = i \text{ if } \hat{\delta}_k \in [a_{i-1}, a_i]
\]
and let
\[
\hat{s}(r) = s(r, \hat{\delta}_k)
\]
where \( s(r, \delta) \) is defined by (7.20).
Now, we run the second stage test with \( k_1 \) new systems that are different from the ones used to estimate \( \delta \), each having \( m \) components and let \( N_0^{k_1}(\hat{s}_k) \) be the number of system failures in the interval \((0, \hat{s}(r))\). Our hypotheses are defined by (8.1), and the form of the test is:

\[
\text{Accept } H_0 \text{ if } N_0^{k_1}(\hat{s}_k) \leq \hat{c}_k.
\]

By (8.6) we know that the OC-Curve is given by

\[
(8.9) \quad h_{k_1}(\theta, \delta_k, \hat{s}(r)) = P_0\left(N_0^{k_1}(\hat{s}_k) \leq \hat{c}_k | \delta_k \right),
\]

where \( P_0\left(N_0^{k_1}(\hat{s}_k) \leq \hat{c}_k | \delta_k \right) \) is given by (8.7) and (8.3), with the appropriate substitutions.

The OC-Curve given by (8.9) has the following properties:

**Theorem 8.4:** If \( \delta \geq 1 \) and \( d \in (a_i, a_{i-1}) \) for some \( i : 0 \leq i \leq i^* \), then for the fixed \( \theta \), as \( k \to \infty \),

\[
(8.10) \quad h_{k_1}(\theta, \delta_k, \hat{s}(r)) \xrightarrow{P} h_{k_1}(\theta, \delta, s(r,d)) \tag{8.10}
\]

and

\[
(8.11) \quad E h_{k_1}(\theta, \delta_k, \hat{s}(r)) \xrightarrow{P} h_{k_1}(\theta, \delta, s(r,d)) \tag{8.11}
\]

**Proof:** By Lemma 8.1, we know that \( h_{k_1}(\theta, \delta, s(r,d)) \) is a continuous function of \( \delta \) in \((a_i, a_{i-1})\) for any fixed \( \theta \). Therefore, for any sufficiently small \( \epsilon > 0 \), there exists a \( \xi > 0 \) such that

\[
| h_{k_1}(\theta, x, s(r,x)) - h_{k_1}(\theta, \delta, s(r,d)) | < \epsilon, \quad \text{whenever } | x - \delta | < \xi.
\]

Because \( \delta_k \) is discrete,
\[ P(\mid h_{k_1}(\theta, \hat{\delta}_k, \hat{s}(r)) - h_{k_1}(\theta, \delta, s(r, \delta)) \mid < \varepsilon) = \sum_{|h_{k_1}(\theta, x, s(r, x)) - h_{k_1}(\theta, \delta, s(r, \delta))| < \varepsilon} P(\hat{\delta}_k = x) \leq \sum_{|x - \delta| < \varepsilon} P(\hat{\delta}_k = x) = P(|\hat{\delta}_k - \delta| < \xi) \rightarrow 1 \text{ as } k \rightarrow \infty, \text{ by Theorem 8.2.} \]

This completes the proof of (8.10).

If \( \delta \in (a_i, a_{i-1}) \) for some \( i: 0 \leq i \leq i^* \), then

\[ E h_{k_1}(\theta, \hat{\delta}_k, \hat{s}(r)) = \sum_{x \in (a_i, a_{i-1})} h_{k_1}(\theta, x, s(r, x)) P(\hat{\delta}_k = x) + \sum_{x \in (a_i, a_{i-1})} h_{k_1}(\theta, x, s(r, x)) P(\hat{\delta}_k = x) \]

\[ = h_{k_1}(\theta, s(r, \delta)) P(\hat{\delta}_k \in (a_i, a_{i-1})) + \]

\[ + \sum_{x \in (a_i, a_{i-1})} [h_{k_1}(\theta, x, s(r, x)) - h_{k_1}(\theta, s(r, \delta))] P(\hat{\delta}_k = x) + \]

\[ + \sum_{x \in (a_i, a_{i-1})} h_{k_1}(\theta, x, s(r, x)) P(\hat{\delta}_k = x) = \]

\[ = (I)_k + (II)_k + (III)_k, \]

where

\[(I)_k = h_{k_1}(\theta, s(r, \delta)) P(\hat{\delta}_k \in (a_i, a_{i-1})), \]

\[(II)_k = \sum_{x \in (a_i, a_{i-1})} [h_{k_1}(\theta, x, s(r, x)) - h_{k_1}(\theta, s(r, \delta))] P(\hat{\delta}_k = x), \]

\[(III)_k = \sum_{x \in (a_i, a_{i-1})} h_{k_1}(\theta, x, s(r, x)) P(\hat{\delta}_k = x). \]

Next, we shall show that as \( k \rightarrow \infty \),

(a) \( (I)_k \rightarrow h_{k_1}(\theta, s(r, \delta)) \)

(b) \( (II)_k \rightarrow 0 \), and (c) \( (III)_k \rightarrow 0. \)
By Theorem 8.2, and \( \delta \in (a_i, a_{i-1}) \), \( P(\delta_k \in (a_i, a_{i-1})) \to 1 \) as \( k \to \infty \), showing (a).

Since for small \( \xi > 0 \), \( (a_i, a_{i-1}) \supseteq (\delta - \xi, \delta + \xi) \), we can write (II)_k as the sum of two parts:

\[
(\text{II})_k = \sum_{x \in (\delta - \xi, \delta + \xi)} [h_{k1}(\theta, x, s(r, x)) - h_{k1}(\theta, \delta, s(r, \delta))] P(\delta_k = x) + \sum_{x \not\in (\delta - \xi, \delta + \xi)} [h_{k1}(\theta, x, s(r, x)) - h_{k1}(\theta, \delta, s(r, \delta))] P(\delta_k = x).
\]

Therefore, by (8.12) it follows that

\[
|\text{II}_k| \leq \varepsilon P(0 < |\delta_k - \delta| < \xi) + P(|\delta_k - \delta| \geq \xi) \to \varepsilon \text{ as } k \to \infty,
\]

by Theorem 8.2. Letting \( \varepsilon \downarrow 0 \) gives (b). Since \( |h_{k1}(\theta, \delta, s(r, \delta))| \leq 1 \), (III)_k \( \leq P(\delta_k \in (a_i, a_{i-1})) \to 0 \) as \( k \to \infty \), by Theorem 8.2 and \( a_i < \delta < a_{i-1} \), which implies (c). Thus, we have shown (8.11), completing the proof of Theorem 8.4.

When \( \delta = a_i \), \( h_{k1}(\theta, \delta_k, s(r)) \) is discontinuous and hence will not converge in probability to \( h_{k1}(\theta, \delta, s(r, \delta)) \) as \( k \to \infty \). However, we still have the following asymptotic result about \( h_{k1}(\theta, \delta_k, s(r)) \), when \( \delta = a_i \).

**Theorem 8.5**: If \( \delta = a_i \) for some \( i: 0 \leq i \leq i^* - 1 \), then as \( k \to \infty \),

\[
(8.13) \quad h_{k1}(\theta, \delta_k, s(r)) - q_{k1}(i, \theta, \delta_k) \to 0,
\]

and
\[(8.14)\] \(E h_{K1}(\theta, \delta_k, \hat{s}(r)) \rightarrow \left\{ \begin{array}{ll}
h_{K1}(\theta, a_i, s(r; a_i)) & \text{if } \theta > \theta_0 \\
(1/2)(h_{K1}(\theta, a_i, s(r; a_i)) + g_{K1}(r; \theta, a_i, i)) & \text{if } \theta = \theta_0 \\
g_{K1}(r; \theta, a_i, i) & \text{if } \theta < \theta_0 \\
\end{array} \right.\)

where
\[g_{K1}(r; \theta, a_i, i) = P(\chi^{2}_{2(i+2)} \geq (1/\theta)^a \cdot R(r; a_i, i+1)),\]

and
\[q_{K1}(i, \theta, x) = I_{[x \geq a_i]}(x) \ h_{K1}(\theta, a_i, s(r; a_i)) + I_{[x < a_i]}(x) \ g_{K1}(r; \theta, a_i, i).\]

\(R(r; a_i, i+1)\) is defined by \((8.19)\), and
\[I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases} \]

Proof: If \(\delta = a_i\) for some \(i: 0 \leq i \leq i^* - 1\), we have
\[P_{\theta}(|h_{K1}(\theta, \delta_k, \hat{s}(r)) - q_{K1}(i, \theta, \delta_k)| < \epsilon) = \sum_r P(\delta_k = x) \cdot |h_{K1}(\theta, a_i, s(r; a_i)) - q_{K1}(i, \theta, x)| < \epsilon\]

Note that,
\[h_{K1}(\theta, x, s(r; x)) - q_{K1}(\theta, x) = h_{K1}(\theta, x, s(r; x)) - h_{K1}(\theta, a_i, s(r; a_i)) \quad \text{if } x \geq a_i\]

\[= h_{K1}(\theta, x, s(r; x)) - g_{K1}(r; \theta, a_i, i) \quad \text{if } x < a_i.\]

By Lemma 8.1 we know that
\[h_{K1}(\theta, x, s(r; x)) \rightarrow h_{K1}(\theta, a_i, s(r; a_i)) \text{ as } x \rightarrow a_i^+,\]

and
\[h_{K1}(\theta, x, s(r; x)) \rightarrow g_{K1}(r; \theta, a_i, i) \text{ as } x \rightarrow a_i^-..\]
Therefore, there exists a small $\eta > 0$ such that
\[ |h_{k1}(\theta, x, s(r;x)) - h_{k1}(\theta, a_i, s(r;a_i))| < \epsilon/2 \quad \text{for} \quad x \in [a_i, a_i + \eta], \]
and
\[ |h_{k1}(\theta, x, s(r;x)) - g_{k1}(r, \theta, a_i, i)| < \epsilon/2 \quad \text{for} \quad x \in [a_i - \eta, a_i). \]
So, we have that
\[ P_\theta (|h_{k1}(\theta, s_k, \hat{s}(r)) - q_{k1}(i, \theta, \hat{s}_k)| < \epsilon) \rightarrow 1 \]
as $k \rightarrow \infty$, by Theorem 8.2 with $\delta = a_i$. Thus, for any $\epsilon > 0$,
\[ P_\theta (|h_{k1}(\theta, s_k, \hat{s}(r)) - q_{k1}(i, \theta, \hat{s}_k)| < \epsilon) \rightarrow 1 \quad \text{as} \quad k \rightarrow \infty, \]
completing the proof of (8.13).

Next, we shall prove (8.14). Observe that
\[ E_{k1}(i, \theta, \hat{s}_k) = h_{k1}(\theta, a_i, s(r;a_i))(1 - G_k(i;\theta)) + \]
\[ + g_{k1}(r; \theta, a_i, i)G_k^r(i;\theta), \]
where
\[ G_k^r(i;\theta) = P_\theta (\Delta_k = a_i). \]
Defining $AN_k$ as in (10.8), we have
\[ G_k^r(i;\theta) = P_\theta (AN_k > g_k(\delta;\theta)((\theta/\theta_0)\delta - 1)). \]
As $k \rightarrow \infty$, $g_k(\delta;\theta) \rightarrow \infty$ and $AN_k$ is asymptotically standard normal by (10.13), so that
\[ G_k^r(i; \theta) \rightarrow \begin{cases} 
0 & \text{if } \theta > \theta_0 \\
1/2 & \text{if } \theta = \theta_0 \\
1 & \text{if } \theta < \theta_0 
\end{cases} \]

Therefore,

\[
\begin{align*}
E q_{r}^{K_1}(i, \theta, \delta_k) & \\
& \begin{cases} 
  h_{K_1}(\theta, a_i, s(r; a_i)) & \text{if } \theta > \theta_0 \\
  (1/2)(h_{K_1}(\theta, a_i, s(r; a_i)) + g_{K_1}(r; \theta, a_i, i)) & \text{if } \theta = \theta_0 \\
  g_{K_1}(r; \theta, a_i, i) & \text{if } \theta < \theta_0 
\end{cases} 
\end{align*}
\]

Using (8.13) and the fact that \( h_{K_1}(\theta, a_i, s(r; a_i)) \) and \( q_{K_1}^{r}(i, \theta, \delta_k) \) are bounded completes the proof of (8.14) and Theorem 8.5.
CHAPTER IX
PROOFS OF RESULTS IN CHAPTER VIII.

To show the asymptotic results of the estimator \( \hat{\delta}_k \) of \( \delta \), we first need to show the monotonicity of \( g_k(\delta) \). Then, we use the monotonicity of \( g_k(\delta) \) to show the consistency of the estimator \( \hat{\delta}_k \) and the normality of the scaled estimator \( \hat{\delta}_k \).

The next lemma says that \( g_k(\delta) \) is a decreasing function of \( \delta \).

**Lemma 9.1:** For \( k > \exp(\exp(131/120)) \), then \( g_k(\delta) \) is strictly decreasing in \( \delta \geq 1 \). Also

\[
(9.1) \quad g_k(1) = \frac{n}{\ln(k)} \text{ and } g_k(\infty) = 0.
\]

**Proof:** From (8.4) we have that

\[
(9.2) \quad \ln(g_k(\delta)) = (\delta-1)\ln\delta - \ln\Gamma(\delta) + \ln(m_k) - \delta\ln(k),
\]

and

\[
(9.3) \quad \frac{d\ln(g_k(\delta))}{d\delta} = \ln(\delta) + \frac{\delta-1}{\delta} \psi(\delta) - \ln(k),
\]

where

\[
(9.4) \quad \psi(\delta) = \frac{d\ln\Gamma(\delta)}{d\delta}.
\]

From the series expansions given in Abramowitz and Stegun (1965, Sections 6.1.41, 6.1.42, and 6.3.18), we have
\[ 0 < \delta \{ \psi(\delta) - \ln\delta - \frac{1}{2\delta} - \frac{1}{128\delta^2} \} < \frac{1}{120\delta^3}, \]

and
\[ 0 < \delta \{ \psi(\delta) - \ln\delta - \frac{1}{2\delta} - \frac{1}{128\delta^2} \} < \frac{1}{120\delta^3}. \]

Thus, we have
\[ (9.5) \quad \psi(\delta) = \ln(\delta) - \frac{1}{2\delta} - a(\delta), \]

where
\[ \frac{1}{128^2} < a(\delta) < \frac{1}{128^2} + \frac{1}{120\delta^4} \quad \text{for all } \delta > 0; \]

and
\[ (9.6) \quad \ln\Gamma(\delta) = (\delta - 1/2)\ln(\delta) - \delta + (1/2)\ln(2\pi) + b(\delta), \]

where
\[ \frac{1}{128} < b(\delta) < \frac{1}{128} + \frac{1}{360\delta^3} \quad \text{for } \delta > 0. \]

Therefore,
\[ (9.7) \quad \frac{d\ln g_k(\delta)}{d\delta} < 1 - \ln(\ln k) + \frac{11}{120} = \frac{131}{120} - \ln(\ln k) \quad \text{for } \delta \geq 1. \]

Thus, for \( \ln(\ln k) > 131/120 \) we have \( \frac{d\ln g_k(\delta)}{d\delta} < 0 \) for \( \delta \geq 1. \)

So, \( g_k(\delta) \) is strictly decreasing for \( \delta \geq 1. \) Combining (9.5) and
\[ (9.7) \] gives us that
\[ (9.8) \quad \ln g_k(\delta) < -\frac{1}{2}\ln(\delta) + \delta - \delta\ln(\ln k) + \ln(k) \to \infty \quad \text{as } \delta \to \infty \]

since \( k > \exp(\exp(131/120)) \), \( \ln g_k(\delta) \to -\infty \) as \( \delta \to \infty. \)

Therefore, \( g_k(\infty) = 0. \) This completes the proof.
Proof of theorem 8.1: For \( \theta = \theta_0 \), the likelihood is given as

\[
L(\delta) = P(N_0^k(s^*) = r) = \frac{1}{r!}(g_k(\delta))^r \exp(-g_k(\delta)).
\]

Then,

\[
\frac{d \ln L(\delta)}{d \delta} = \frac{d g_k(\delta)}{d \delta} \left( r - g_k(\delta) \right).
\]

If \( r > g_k(1) \), then by the lemma 9.1, it follows that

\[
\frac{d g_k(\delta)}{d \delta} < 0
\]

and \( r - g_k(\delta) > 0 \) for all \( \delta \geq 1 \). Therefore, for \( r > g_k(1) \),

\[
\frac{d \ln L(\delta)}{d \delta} < 0 \quad \text{for all } \delta \geq 1.
\]

So, for \( r > g_k(1) \), \( \max_{\delta \geq 1} L(\delta) = L(1) \).

If \( r \leq g_k(1) \), by the lemma 9.1 it follows that there exists a \( \delta_k \) such that \( r - g_k(\delta_k) = 0 \). And

\[
\frac{d (\ln L(\delta))}{d \delta} = \begin{cases} 
> 0 & \text{if } 1 \leq \delta < \delta_k^* \\
< 0 & \text{if } \delta > \delta_k^*
\end{cases}
\]

Therefore, for \( r \geq g_k(1) \), we have shown \( \max_{\delta \geq 1} L(\delta) = L(\delta_k) \). This completes the proof of theorem 8.1.

Proof of theorem 8.2: We need to show that for any \( 0 < \varepsilon < 1 \)

\[
P_\theta( | \delta_k - \delta | > \varepsilon ) \to 0 \text{ as } k \to \infty.
\]

Note that \( P_\theta( | \delta_k - \delta | > \varepsilon ) = P_\theta(\delta_k < \delta - \varepsilon) + P_\theta(\delta_k > \delta + \varepsilon) \).

Since \( g_k(1) \to \infty \text{ as } k \to \infty \), for large \( k \), \( m^*_k > \delta + \varepsilon \). Therefore,

\[
P_\theta(\delta_k > \delta + \varepsilon) = P_\theta(N_0^k(s^*) = 0) + P_\theta(0 < N_0^k(s^*) \leq g_k(1),
\]

and
\[ N_0^k(s^*) < g_k(\delta + \epsilon) = \]
\[ = P_\theta(N_0^k(s^*) < g_k(\delta + \epsilon)) = P_\theta(AN_k < \sqrt{g_K(\delta; \theta)}(g_k(\delta + \epsilon)/g_k(\delta; \theta) - 1)) = \]
\[ = P_\theta(AN_k < \sqrt{g_K(\delta; \theta)}(\theta/\theta_0)\delta g_k(\delta + \epsilon)/g_k(\delta) - 1)), \]

where
\[ (9.8) \quad AN_k = (N_0^k(s^*) - g_k(\delta; \theta))/\sqrt{g_k(\delta; \theta)}, \]

and \( g_k(\delta; \theta) \) is given by (3.4) and (3.5). Thus,
\[ (9.9) \quad P_\theta(\delta_k > \delta + \epsilon) = \]
\[ = P_\theta(AN_k < \sqrt{g_K(\delta; \theta)}(\theta/\theta_0)\delta g_k(\delta + \epsilon)/g_k(\delta) - 1)). \]

For \( 1 < \delta - \epsilon \), then
\[ P_\theta(\delta_k < \delta - \epsilon) = P(N_0^k(s^*) > g_k(1)) + P_\theta(0 < N_0^k(s^*) \leq g_k(1) \]

and
\[ N_0^k(s^*) > g_k(\delta - \epsilon) = \]
\[ = P_\theta(N_0^k(s^*) > g_k(\delta - \epsilon)) = P_\theta(AN_k > \sqrt{g_K(\delta; \theta)}(g_k(\delta - \epsilon)/g_k(\delta; \theta) - 1)) = \]
\[ = P_\theta(AN_k > \sqrt{g_K(\delta; \theta)}(\theta/\theta_0)\delta g_k(\delta - \epsilon)/g_k(\delta) - 1)). \]

For \( 1 \geq \delta - \epsilon \), since \( \delta_k \geq 1 \),
\[ P_\theta(\delta_k < \delta - \epsilon) = 0 . \]

Therefore, we have shown
\[ (9.10) \quad P_\theta(\delta_k < \delta - \epsilon) \]
\[ = P_\theta(AN_k > \sqrt{g_K(\delta; \theta)}(\theta/\theta_0)\delta g_k(\delta - \epsilon)/g_k(\delta) - 1)) \quad \text{if } 1 < \delta - \epsilon \]
\[ = 0 \quad \text{if } 1 \geq \delta - \epsilon . \]
Note that as $k \to \infty$

\[(9.11) \quad g_k(\delta + x)/g_k(\delta) = (B(\delta + x)/B(\delta)) (1/\ln k)^{\delta} \to \begin{cases} 0 & \text{if } x > 0 \\ 1 & \text{if } x = 0 \\ \infty & \text{if } x < 0 \end{cases} \]

From (8.4) and (8.5) we know that

\[(9.12) \quad g_k(\delta; 0) \to \infty \text{ as } k \to \infty \]

Since $N_k^0(s^*)$ is an approximately Possion distribution with a mean $g_k(\delta; \theta)$, by (9.8) and (9.9) we have

\[(9.13) \quad \mathbb{A}N_k \overset{L}{\longrightarrow} N(0, 1) \text{ as } k \to \infty. \]

From (9.11) and (9.12) we get

\[
\sqrt{g_k(\delta; \theta)} ((\theta/\theta_0)^\delta g_k(\delta + \epsilon)/g_k(\delta - 1)) \to -\infty \text{ as } k \to \infty,
\]

and

\[
\sqrt{g_k(\delta; \theta)} ((\theta/\theta_0)^\delta g_k(\delta - \epsilon)/g_k(\delta - 1)) \to \infty \text{ as } k \to \infty.
\]

Therefore, by (9.13) we have that for $0 < \epsilon < 1$

\[
P_\theta( |\hat{\delta} - \delta| > \epsilon ) \to 0 \text{ as } k \to \infty.
\]

This completes the proof of theorem 8.2.*

**Proof of theorem 8.3:**

(a) Let \( \{x_k\} \) be a sequence to be determined later such that \( x_k \to 0 \) as \( k \to \infty \).

Note that

\[
g_k(\delta + x_k)/g_k(\delta; \theta) = (\theta/\theta_0)^\delta (g_k(\delta + x_k)/g_k(\delta)) =
\]

\[
= (\theta/\theta_0)^\delta (kB(\delta + x_k)/(\ln k)^{\delta + x_k})(1/(\ln k)^{x_k}) =
\]

So,
\[
\ln(g_k(\delta+x_k)g_k(\delta;\theta)) = \delta \ln(\theta/\theta_0) - x_k \ln(\ln k) + \ln(B(\delta+x_k)/B(\delta)).
\]

Let
\[
u(x) = \ln B(x),
\]
then
\[
u(\delta) = (\delta-1)\ln \delta - \ln \Gamma(\delta) \quad \text{and} \quad u'(\delta) = \ln \delta + (\delta-1)/\delta - \psi(\delta).
\]
So, by (3.10) and (3.11) it follows that
\[
u(\delta) = \frac{1}{2} \ln \delta + \delta \frac{1}{2} \ln(2\pi) - b(\delta),
\]
where \(0 < b(\delta) < 1/(\pi^2 \delta),\) and
\[
u'(\delta) = 1 - \frac{1}{2\delta} + a(\delta),
\]
where \(0 < a(\delta) < \frac{1}{\pi^2 \delta}.\)

Since \(u(\delta)\) and \(u'(\delta)\) are continuous functions of \(\delta,\) by Taylor's theorem it follows that
\[
u(\delta+x_k) = u(\delta) + x_k u'(\delta),
\]
where \(z_k\) is interior to the interval joining \(\delta\) and \(\delta+x_k.\)

Therefore,
\[
u(\delta+x_k) = u(\delta) + x_k u'(\delta) + x_k (u'(z_k) - u'(\delta)).
\]

Define
\[
w(z_k; \delta) = u'(z_k) - u'(\delta) + a(\delta),
\]
then
\[
\ln(g_k(\delta+x_k)g_k(\delta;\theta)) = \delta \ln(\theta/\theta_0) - x_k \ln(\ln k) + x_k (1 - 1/2\delta) + x_k w(z_k; \delta).
\]

Now, let
Since $g_k(\delta; \theta) \to \infty$ as $k \to \infty$, then $x_k \to 0$ as $k \to \infty$.

Combining (9.17) and (9.18) gives us that

$$\ln(g_k(\delta+x_k)/g_k(\delta; \theta)-x_k w(z_k; \theta)) = \ln(1 - b/\sqrt{g_k(\delta; \theta)}) \cdot \frac{\delta \ln(\theta/\theta_0)}{\ln(\ln k) - 1 + 1/2 \delta}.$$  

Thus,

$$g_k(\delta+x_k) = g_k(\delta; \theta) - 1 = -(1/\sqrt{g_k(\delta; \theta)}) \cdot \exp(x_k w(z_k; \delta)).$$

Since $x_k \to 0$ and $m_k^* \to \infty$ as $k \to \infty$, for $\delta > 1$ we have $\delta + x_k > 1$ and $m_k^* > \delta + x_k$ for large $k$. Thus,

$$P_\theta(\delta_k - \delta < x_k) = P(\delta_k < \delta + x_k) = P(\mathcal{N}_0(s^*) > g_k(1)) +$$

$$+ P_\theta(g_k(\delta+x_k) < \mathcal{N}_0(s^*) \leq g_k(1)) =$$

$$= P_\theta(\mathcal{N}_0(s^*) > g_k(\delta+x_k)).$$

By (9.17) and (9.18) we have that

$$P_\theta(\mathcal{N}_0(s^*) > g_k(\delta+x_k)) = P_\theta(\mathcal{A}_k > \sqrt{g_k(\delta; \theta)} \cdot (g_k(\delta+x_k)/g_k(\delta; \theta) - 1)) =$$

$$= P_\theta(\mathcal{A}_k > b \cdot \exp(x_k w(z_k; \delta))),$$

where $\mathcal{A}_k$ is defined by (9.8). Thus,

$$P_\theta(\delta_k - \delta < x_k) = P_\theta(\mathcal{A}_k > b \cdot \exp(x_k w(z_k; \delta))).$$

From (9.14) it follows that for large $k$, $w(z_k; \delta)$ is bounded for any fixed $\delta$ and $\theta$. Therefore,

$$\exp(x_k w(z_k; \delta)) \to 1 \text{ as } k \to \infty.$$

Since

$$\frac{\mathcal{N}_0^k(s^*) - g_k(\delta; \theta)}{\sqrt{g_k(\delta; \theta)}} \overset{L}{\to} \mathcal{N}(0, 1) \text{ as } k \to \infty,$$
by Slutsky's Theorem it follows that
\[ P_\theta(\delta_k < x_k) \to \Phi(b) \text{ as } k \to \infty. \]

Note
\[ \sqrt{g_k(\delta; \theta)} \left( \ln k - 1 + 1/2 \delta \right) \left( x_k - \frac{\delta \ln(\theta/\theta_0)}{\ln(\ln k) - 1 + 1/2 \delta} \right) = -\sqrt{g_k(\delta; \theta)} \ln(1-b/\sqrt{g_k(\delta; \theta)}) \to b \text{ as } k \to \infty. \]

So, by Slutsky's Theorem it follows that
\[ P_\theta(\sqrt{g_k(\delta; \theta)} (\ln(\ln k) - 1 + 1/2 \delta)(\delta_k \mu_k(\delta; \theta)) < b) \to \Phi(b) \text{ as } k \to \infty. \]

This completes the proof of the part (a).

(b) For \( \delta=1 \), \( x_k \), defined by (9.18), is reduced to

\[ x_k = \frac{-\ln(1-b/\sqrt{g_k(1; \theta)}) + \ln(\theta/\theta_0)}{\ln(\ln k) - 1/2} \]

where

\[ g_k(1; \theta) = (\theta_0 \theta)(k/\ln k). \]

Since \( g_k(1; \theta) \to \infty \text{ as } k \to \infty \), for large \( k \)

\[ -\ln(1-b/\sqrt{g_k(1; \theta)}) + \ln(\theta/\theta_0) < 0 \quad \text{if } \theta < \theta_0 \]

\[ > 0 \quad \text{if } \theta > \theta_0. \]

Combining (9.21) and (9.22) gives us that

\[ x_k < 0 \quad \text{if } \theta < \theta_0 \]

\[ > 0 \quad \text{if } \theta > \theta_0. \]

For \( \theta = \theta_0 \), we have

\[ x_k = \frac{-\ln(1-b/\sqrt{g_k(1; \theta_0)})}{\ln(\ln k) - 1/2}. \]

So,
By (9.23) and (9.24) we know that for \( \theta > \theta_0 \) or \( \theta = \theta_0 \) and \( b > 0 \)
\[ x_{k+1} > 1. \]

Therefore, using the same method as used to show part (a), we can show that
\[ P_\theta(\hat{\delta}_{k-1} < x_k) \to \Phi(b) \text{ if } \theta > \theta_0 \text{ and any } b, \]
and
\[ P_\theta(\hat{\delta}_{k-1} < x_k) \to \Phi(b) \text{ if } \theta = \theta_0 \text{ and } b > 0, \]

Since \( \hat{\delta}_k \geq 1 \), from (9.23) and (9.24) we obtain
\[ P_\theta(\hat{\delta}_{k-1} < x_k) = 0 \text{ if } \theta < \theta_0 \text{ or } \theta = \theta_0 \text{ and } b \leq 0. \]

From (9.25) and (9.26) we conclude that (i) is true. From (9.26) we conclude
\[ P_\theta(\hat{\delta}_{k-1} < x_k) \to \Phi(b) \text{ as } k \to \infty \text{ if } \theta = \theta_0 \text{ and } b > 0 \]
\[ \to 0 \text{ as } k \to \infty \text{ if } \theta = \theta_0 \text{ and } b \leq 0. \]

By (9.28) we conclude that (ii) is true. And From (9.27) we conclude (iii) is true also. This completes the proof part (b); this completes the proof of theorem 8.3.
CHAPTER X
SIMULATION STUDY.

In Chapter VIII we obtained asymptotic properties of the two stage test procedures as the sample size of the first stage test tends to infinity. Next, we shall study the behavior of the OC-Curve of the two stage test procedure when the sample size of the first stage test is finite. More, specifically, we want to see how close the OC-Curve of the two stage test is to the OC-Curve of the single stage test that we would use if we knew the value of $\delta$.

Let $k$ be the number of systems and $m$ be the number of components in the first stage. We take $m=15$ and $k$ to be one of 13 values: 70 (20) 310. For each $k$, we calculate the testing time $s^*$ for the first stage test, using (9.1). We assume that $N_k^0(s^*)$ has a Poisson distribution with a mean of $g_k(\delta,\theta)$, defined by (9.5). Let $\theta_0=100.0$, $\theta_1=50.0$, $\alpha=0.1$ and $\beta=0.1$. For each fixed value of three representative $\delta$ values: 2, 3, 4, and each of 17 representative $\theta$ values: 35 (5) 115, we generate a sample of size 100 Poisson random variables, each having mean $g_k(\delta,\theta)$. For each of the 100 generated numbers, $N_k^0(s^*)$, we calculate the estimate of $\delta$, $\delta_k$, defined by (9.16).
We calculate the OC-Curve for the second stage test as follows. Suppose the second stage test is run with \( k_1 \) new systems, each having \( m \) components, and that the testing time \( s(r; \delta_k) \) of the second stage test is defined by:

\[
(10.1) \quad s(r; \delta_k) = m \left( R_1(r; \delta_k, c)/n_1 B(\delta_k) \right)^{1/\delta_k},
\]

where \( R_1(r; \delta_k, c) \) is defined by (8.19), \( 0 \leq r \leq 1 \), and \( n_1 = mk_1 \).

If \( N_{00}^{k_1}(\delta_k) \) is the number of component failures during the second stage test, we would reject \( H_0: \theta = \theta_0 \) if \( N_{00}^{k_1}(\delta_k) > c \), with \( c = c(\delta_k) \), given by (8.17). Then, by (2.28) we obtain the OC-Curve of the second stage test, given \( \delta_k \), as

\[
(10.2) \quad h_1(\theta, \delta_k, s(r; \delta_k)) = P[\chi^2_{2(c+1)} \geq (1/\theta)^{\delta_k} R_1(r; \delta_k, c)].
\]

Note that with the testing time \( s(r; \delta_k) \) given by (10.1), the second stage OC-Curve is independent of the choice of \( k_1 \). Next, we calculate the average of \( h_1(\theta, \delta_k, s(r; \delta_k)) \) for the 100 \( \delta_k \)'s, and let \( \text{AEOC}(\delta_k, \theta, r) \) denote this average. Note \( \text{AEOC}(\delta_k, \theta, r) \) is a function of \( \theta \). Let \( \text{NOC}(\delta, \theta, r) \) denote the nominal OC-Curve obtained from the single stage test by assuming \( \delta \) is known, and using the same value of \( r \) as in the two-stage test. This \( \text{NOC}(\delta, \theta, r) \) is defined by (7.21). In this simulation study, we only take \( r = 0 \), or 1. By choosing \( r = 0 \) we force both the
AEOC($\delta_k,\theta,0$) and the NOC($\theta,\delta,0$) to go through $1-\alpha$ at $\theta=\theta_0$, and by choosing $r=1$ we force both the AEOC($\delta_k,\theta,0$) and the NOC($\theta,\delta,1$) to go through $\beta$ at $\theta=\theta_1$.

To measure the closeness of the average estimated OC-Curve, AEOC($\delta_k,\delta$), to the nominal OC-Curve NOC($\delta,\theta,r$), we use $L_2$-distance:

\begin{equation}
D_r(k,\delta)=\int (AEOC(\delta_k,\theta,r)-NOC(\delta,\theta,r))^2 d\theta.
\end{equation}

Since we only take $\theta=35$ (5) 115,

\begin{equation}
D_r(k,\delta)=5.0\left[\sum_{i=1}^{17} (AEOC(\delta_k,\theta_i,r)-NOC(\delta,\theta_i,r))^2\right],
\end{equation}

where $\theta_i=120-5i$, $i=1, \ldots, 17$. The $D_r(k,\delta)$ are summarized in Table 3 and Table 4 below.

From Table 3 and Table 4, we conclude that for a fixed $\delta$, $D_r(k,r)$ is a decreasing function of $k$. Note that $D_1(k,\delta)<D_0(k,\delta)$ for $k=70$ (20) 310 and $\delta=2.0, 3.0, 4.0$. Therefore, we conclude that the OC-Curve of the two stage test obtained by choosing $r=1$ has a faster convergence rate than the one obtained by choosing $r=0$. The smaller $\delta$ is, the faster is the convergence rate of the OC-Curve of the two stage test to the nominal OC-Curve.

Next, we want see the overall comparison of AEOC($\delta_k,\theta,r$) with NOC($\delta,\theta,r$). To achieve this goal we plot AEOC($\delta_k,\theta,r$) and NOC($\delta,\theta,r$) against $\theta$ for three values of $\delta$: 2, 3, and 4 and $k=70$. 
and 90 for $r=0$ and 1. From the plots given in Figures 9, 10, and 11 in Appendix B, we see that the variation of the estimated OC-Curve increases as $\delta$ increases. The best way to choose a testing time for the two stage test to protect both the consumers and the producers is setting $r=1$. However, the best way to choose a testing time for the one stage test to protect both consumers and producers is setting $r=0$. 
Table 3
Distance between Ave E.OC-Curve and N.OC-Curve
for r=0 in (10.2)

<table>
<thead>
<tr>
<th>k</th>
<th>δ=2.0</th>
<th>δ=3.0</th>
<th>δ=4.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>70.00</td>
<td>2.429172</td>
<td>1.813981</td>
<td>1.032769</td>
</tr>
<tr>
<td>90.00</td>
<td>2.215829</td>
<td>1.404263</td>
<td>0.640173</td>
</tr>
<tr>
<td>110.00</td>
<td>1.983507</td>
<td>1.09909</td>
<td>0.4484</td>
</tr>
<tr>
<td>130.00</td>
<td>1.899739</td>
<td>0.86361</td>
<td>0.366576</td>
</tr>
<tr>
<td>150.00</td>
<td>1.735508</td>
<td>0.743318</td>
<td>0.300785</td>
</tr>
<tr>
<td>170.00</td>
<td>1.66389</td>
<td>0.636755</td>
<td>0.233428</td>
</tr>
<tr>
<td>190.00</td>
<td>1.557049</td>
<td>0.586274</td>
<td>0.223692</td>
</tr>
<tr>
<td>210.00</td>
<td>1.506286</td>
<td>0.530339</td>
<td>0.202927</td>
</tr>
<tr>
<td>230.00</td>
<td>1.448364</td>
<td>0.484443</td>
<td>0.179257</td>
</tr>
<tr>
<td>250.00</td>
<td>1.369279</td>
<td>0.454816</td>
<td>0.175989</td>
</tr>
<tr>
<td>270.00</td>
<td>1.358541</td>
<td>0.392952</td>
<td>0.163294</td>
</tr>
<tr>
<td>290.00</td>
<td>1.269199</td>
<td>0.392952</td>
<td>0.150184</td>
</tr>
<tr>
<td>310.00</td>
<td>1.256077</td>
<td>0.375684</td>
<td>0.149925</td>
</tr>
</tbody>
</table>
Table 4
Distance between Ave E.OC-Curve and N.OC-Curve
for r=1 in (10.2)

<table>
<thead>
<tr>
<th>k</th>
<th>δ=2.0</th>
<th>δ=3.0</th>
<th>δ=4.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>70.00</td>
<td>1.998510</td>
<td>3.232622</td>
<td>3.999434</td>
</tr>
<tr>
<td>90.00</td>
<td>1.852584</td>
<td>2.886106</td>
<td>3.462489</td>
</tr>
<tr>
<td>110.00</td>
<td>1.743852</td>
<td>2.606819</td>
<td>3.064068</td>
</tr>
<tr>
<td>130.00</td>
<td>1.649409</td>
<td>2.380550</td>
<td>2.775407</td>
</tr>
<tr>
<td>150.00</td>
<td>1.565836</td>
<td>2.20818</td>
<td>2.545728</td>
</tr>
<tr>
<td>170.00</td>
<td>1.135264</td>
<td>1.858667</td>
<td>2.352411</td>
</tr>
<tr>
<td>190.00</td>
<td>0.755501</td>
<td>1.570134</td>
<td>2.182159</td>
</tr>
<tr>
<td>210.00</td>
<td>0.639889</td>
<td>1.331935</td>
<td>2.053303</td>
</tr>
<tr>
<td>230.00</td>
<td>0.386755</td>
<td>1.072474</td>
<td>1.933197</td>
</tr>
<tr>
<td>250.00</td>
<td>0.254567</td>
<td>0.641400</td>
<td>1.877126</td>
</tr>
<tr>
<td>270.00</td>
<td>0.172500</td>
<td>0.459895</td>
<td>1.790456</td>
</tr>
<tr>
<td>290.00</td>
<td>0.124895</td>
<td>0.445684</td>
<td>1.749434</td>
</tr>
<tr>
<td>310.00</td>
<td>0.119626</td>
<td>0.435072</td>
<td>1.526004</td>
</tr>
</tbody>
</table>
CHAPTER XII
CONCLUDING REMARKS.

We have applied the concept of robustness to both survival analysis and reliability. We summarize what we did as follows.

In survival analysis, we demonstrated that if a contamination distribution can be any distribution, then no minimax tests exist for censored data. Also, we showed that minimax tests for censored data do exist for some contamination neighborhoods. The difficulty of finding a minimax test for censored data is that we may not be able to observe the survival time of a patient. Thus, even though we have only two contamination neighborhoods of the survival function, we actually have three contamination neighborhoods for three subsurvival functions generated by the survival function and censoring function. The least favorable pairs from three contamination neighborhoods do not agree.

In reliability, first, we have proven that the OC-Curve of the two-stage reliability demonstration test tends to the nominal OC-Curve in probability as the sample size of the first stage
test tends to infinity. Here, the nominal OC-Curve is obtained by assuming that we know the shape parameter. Second, we conducted a simulation study for finite sample size for the first stage test. We found that the average of the estimated OC-Curve of the two-stage reliability demonstration test is close to the nominal OC-Curve for the small sample size used in the first stage test. Also, we found that the variation of the average estimated OC-Curve of the two-stage reliability demonstration test increases as the shape parameter \( \delta \) increases.

In conclusion, we have demonstrated that there are many different approaches to derive robust procedures. Which robust procedure you should choose depends on what kinds of contamination you hypothesize. In the part of my dissertation dealing with survival analysis, we decided to use minimax tests to find a robust procedure because we knew that the contaminations lie in some neighborhoods of true survival functions. We found that it is difficult to find a minimax test for censored data. In the part of my dissertation dealing with reliability, we used the concept of adapted robust procedures to find a robust procedure because we knew that the contaminations are due to the unknown shape parameter. It turned out that we could find a two-stage reliability demonstration test that has some nice asymptotical properties.
We derived minimax tests for censored data only for one sample. In some cases of clinical trials, people like to compare survival functions for two groups of patients using two different drugs respectively. The first research problem needed to be solved is how to extend our minimax tests to the two-sample case. The difficulty of deriving minimax tests for censored data for two-sample case will be that there will be too many contamination neighborhoods. We note that the computation for finding minimal power and ideal power of our minimax test is not very straightforward. Therefore, the second research we would like to work on is the development of an easy algorithm for computing the minimal power and ideal power of our minimax test by approximation.

In reliability, our two-stage reliability demonstration test is based on the assumption that the number of failures during a fixed testing period has a Poisson process for large numbers of systems on the test. However, Blumenthal (1990) has shown that if the number of systems on the test is not large, the Binomial approximation is better than the Poisson approximation. The third problem we want to solve is to develop a two-stage reliability demonstration test for a small number of systems on the test based on a Binomial approximation. If we can only have one sample available, then we can not use a
two-stage test. Thus, the fourth research problem is to use a Bayesian approach to develop a reliability demonstration test for new systems. That is, we put a prior, such as Jeffrey’s prior on the shape parameter $\delta$. 
Appendix A.
Fortran Programs

1. Fortran program for calculating the level and power of Sign and Wilcoxon tests.

real e, level1(26), level2(27), power1(27), power2(27)
real cut, cdelta, temp(4), ee(9)

* "1" is for Sign test and "2" is for Wilcoxon test

parameter (pi=3.141256)
alpha=0.05
cdelta=1.0
do 1 i=1,16
e=0.01*(i-1)
cut=anorin(1-alpha)
level1(i)=1.0-anordf(cut-e)
level2(i)=1.0-anordf(cut-sqrt(3.0)*e)
power1(i)=1.0-anordf(cut-cdelta*sqrt(2.0/pi)+e)
power2(i)=1.0-anordf(cut-cdelta*sqrt(3.0/pi)+sqrt(3.0)*e)
write(6,5)e, level1(i), level2(i), power1(i), power2(i)
5 format(f7.4,f7.4,4x,f7.4,4x,f7.4,4x,f7.4)
1 continue
   write(6,*)
end
2. This program is used to calculate the truncation constants $c'$ and $c''$.

dimension ak(33), aw(33), tt(33)
real lamda0, lamda1, mu, inx0, inx1, k, k1, b1, b2, d1, d2, root
external t3, t1, t2, bisect, wt2
intrinsic exp
parameter (e=0.1, lamda0=3.0, lamda1=2.0)

mu=2.6
b1=0.945333
b2=1.100642
d1=t3(b1)
d2=t3(b2)
k1=exp(-lamda1*d1)/(exp(-lamda0*d1)+e/(1-e))
do 70 i=1,31
ai=i-1

ak(i)=(ai/30.0)*b1+b2*((30.0-ai)/30.0)
k=ak(i)
inx0=d1
inx1=d2
tol=1e-6
maxtis=30
if (k .ge. k1 .and. k .lt. b2) then
\begin{verbatim}
    tt(i) = t2(bisect, wt2, k, inx0, inx1, tol, maxtis, status, root, resid, 
    c k1, b2, d1, d2)
    ttt2 = tt(i)
    www2 = \exp(- (\lambda_0 + \mu) * ttt2)
    aww2 = \exp(- \mu * ttt2)
    endif
    if (k \geq 1.0 \land k \lt k1) then
        ttt1 = t1(k, b1, k1, d1)
        www1 = \exp(- (\lambda_0 + \mu) * ttt1)
        sww1 = \exp(- \mu * ttt1)
    endif
    if (k \geq b1 \land k \lt b2) then
        ttt3 = t3(k)
        www3 = \exp(- (\lambda_0 + \mu) * ttt3)
    endif
    if (k \gt b2) then
        w = 0.0
    else if (k \geq k1 \land k \leq b2) then
        w = (1 - e) / (\lambda_0 + \mu) * (\lambda_0 * www3 + \mu * www2) + sww2
    else if (k \geq 1.0 \land k \lt k1) then
        w = (1 - e) / (\lambda_0 + \mu) * (\lambda_0 * www3 + \mu * www1) + sww1
    else if (k \geq b1 \land k \lt 1.0) then
        w = (1 - e) / (\lambda_0 + \mu) * (\lambda_0 * www3 + \mu) + e
\end{verbatim}
else
w=1.0
end if
aw(i)=w
  write(6,80)aw(i),ak(i)
80  format(f8.6,4x,f8.6)
70  continue
888 continue
     stop
end
real function wt2 (k,t)
real k,lambda1,lambda0
parameter (lambda0=3.0,lambda1=2.0,e=0.1)
intrinsic exp
   wt2=exp(-lambda1*t)-k*exp(-lambda0*t)-k*e/(1-e)
return
end
real function t3 (k)
real lambda0,lambda1,k
intrinsic alog
parameter (lambda0=3.0,lambda1=2.0)
   t3=1.0/(lambda0-lambda1)*alog(k*lambda0/lambda1)
return
end
real function t1 (k,b1,k1,d1)
real k,k1
intrinsic alog
parameter (lamda0=3.0,e=0.1)
if (k .lt. k1 .and. k .ge. 1.0) then
  t1=-(1/lamda0)*alog((1-b*(1-e)-k*e)/((k-b1)*(1-e)))
else
  t1=0.0
end if
return
end

real function t2 (bisect,wt2,k,inx0,inx1,tol,maxtis,c status,root,resid, k1,b2,d1,d2)
real d1,d2,inx0,inx1,k,k1,b2
integer maxtis
external bisect,wt2
if (k .eq. k1) then
  t2=d1
end if
if (k .eq. b2) then
  t2=d2
end if
if (k .gt. k1 .and. k .lt. b2) then
   call bisect (wt2, k, inx0, inx1, tol, maxtis, status, root, resid)
   t2 = root
end if
return
end

subroutine bisect(f, k, inx0, inx1, tol, maxtis, status, root, resid)
   integer count, limit, maxtis, solved, nobrac, status
   real f, f0, f1, inx0, inx1, lastx0, lastx1, x1, x0, xmid, k
   parameter (solved=0, limit=1, nobrac=2)
   intrinsic abs
   external f
   x0 = inx0
   x1 = inx1
   f0 = f(k, x0)
   f1 = f(k, x1)
   if (f0*f1 .gt. 0) then
      status = nobrac
   else
      do 10 count = 1, maxtis
         xmid = (x0 + x1)/2
         fmid = f(k, xmid)
         if (fmid .eq. 0) then
            break
         end if
         do 9 count = 1, limit
            xmid = (x0 + x1)/2
            fmid = f(k, xmid)
            if (fmid .eq. 0) then
               break
            end if
         end do
      end do
   end if
10 continue
end subroutine bisect
status=solved
go to 11
endif
if (f0*fmid.lt. 0) then
x1=xmid
endif
if (f0*fmid.gt. 0) then
x0=xmid
f0=f(k,x0)
endif
if (abs(x0-x1).lt.tol) then
status=solved
endif
continue
status=limit
11 if (status.eq.solved) then
root=(x0+x1)/2
resid=f(k,root)
endif
if (status.eq.limit) then
lastx0=x0
lastx1=x1
3. The following program is used to calculating the minimal power of the minimax test.

dimension ak(33),aw(33),tt(33),tem(34)
real lamda0, lamda1,mu,inx0,inx1,k,k1,b1,b2,d1,d2,root
integer status,solved,count
external t3,t1,t2,bisect,wt2,trunc
intrinsic exp
parameter (e=0.1,lamda0=3.0,lamda1=2.0)
mu=1.5
   call trunc (e,lamda0,lamda1, b1,b2)
   write(6,234)b1,b2
234  format('b1=',f7.4,4x,'b2=',f7.4,)
   d1=t3(b1)
   d2=t3(b2)
      k1=exp(-lamda1*d1)/(exp(-lamda0*d1)+e/(1-e))
do 70 i=1,31
   ai=i-1
   ak(i)=(ai/30.0)*b1+b2*((30.0-ai)/30.0)
k=ak(i)
inx0=d1
inx1=d2
tol=1e-6
maxtis=30

\[ tt(i) = t2(\text{bisect}, wt2, k, inx0, inx1, tol, maxtis, \text{status, root, resid, }\]
c k1,b2,d1,d2)\]

\[ ttt2 = tt(i) \]
\[ ttt1 = t1(k, b1, k1, d1) \]
\[ ttt3 = t3(k) \]
\[ www1 = \exp(-(\lambda_0 + \mu) * ttt1) \]
\[ www2 = \exp(-(\lambda_0 + \mu) * ttt2) \]
\[ www3 = \exp(-(\lambda_0 + \mu) * ttt3) \]
\[ sww1 = e * \exp(-\mu * ttt1) \]
\[ sww2 = e * \exp(-\mu * ttt2) \]

if (k .gt. b2) then
  w=0.0
else if (k .ge. k1 .and. k .le. b2) then
  w=(1-e)/(\lambda_0 + \mu)*(\lambda_0 * www3 + \mu * www2) + sww2
else if (k .ge. 1.0 .and. k .lt. k1) then
  w=(1-e)/(\lambda_0 + \mu)*(\lambda_0 * www3 + \mu * www1) + sww1
else if (k .ge. b1 .and. k .lt. 1.0) then
  w=(1-e)/(\lambda_0 + \mu)*(\lambda_0 * www3 + \mu) + e
else
w=1.0
end if
aw(i)=w
write(6,80)aw(i),ak(i)
80 format(f8.6,4x,f8.6)
70 continue
888 continue
stop
end
end subroutine trunc (e, lamda0, lamda1, b1, b2)
real e, lamda0, lamda1, b1, b2, r
integer n, m
a0=lamda0
a1=lamda1
c0=a0/(a0-a1)
c1=a1/(a0-a1)
r=a1/a0
n=1
if(1-e .le. a1/a0) then
xnew=1-e
b2=xnew
else
x=r+0.5
2 \[ ff = 1.0-(r/x)\cdot c0+(1.0/x)*(r/x)\cdot c1 \]
   \[ dff = (c1/x**2)*(r/x)**(c0-1)-((1+c1)/x**2)*(r/x)**c1 \]
   \[ xf = ff-1.0/(1-e) \]
   \[ xnew = x-xf/dff \]
   if(abs(x-xnew) .lt. 1e-6 .or. n .gt. 20) go to 3
   n=n+1
   x=xnew
   go to 2
3   b2=xnew
   endif
   y=a1/a0 +0.2
   m=1
5   \[ cf = y*(1-(r/y)**c0)+(r/y)**c1 \]
   \[ dcf = 1+(c0-1.0)*(r/y)**c0-(c1/y)*(r/y)**c1 \]
   \[ ycf = cf-1.0/(1-e) \]
   \[ ynew = y-ycf/dcf \]
   if(abs(y-ynew) .lt. 1e-6 .or. m .gt. 20) go to 6
   m=m+1
   y=ynew
   go to 5
6   b1=ynew
   return
   end
real function wt2 (k,t)
real k, lamda1, lamda0
parameter (lamda0=3.0, lamda1=2.0, e=0.1)
intrinsic exp
wt2=exp(-lamda1 * t)/(exp(-lamda0 * t)+e/(1-e))-k
return
end

real function t3 (k)
real lamda0, lamda1, k
intrinsic alog
parameter (lamda0=3.0, lamda1=2.0)
t3=1.0/(lamda0-lamda1)*alog(k*lamda0/lamda1)
return
end

real function t1 (k,b1,k1,d1)
real k, k1
intrinsic alog
parameter (lamda0=3.0, e=0.1)
if (k .eq. k1) then
t1=d1
end if
if (k .eq. 1.0) then
t1=0.0
end if
if (k.ge.1.0.and.k.lt.k1) then
   t1=-(1/lamda0)*alog((1-b1)/(k-b1)*(1-e)-e/(1-e))
end if
return
end
real function t2 (bisect,wt2,k,inx0,inx1,tol,maxtis,
c status,root,resid, k1,b2,d1,d2)
real d1,d2,inx0,inx1,k,k1,b2
integer maxtis
external bisect,wt2
if (k.eq.k1) then
   t2=d1
end if
if (k.eq.b2) then
   t2=d2
end if
if (k.gt.k1.and.k.lt.b2) then
   call bisect (wt2,k,inx0,inx1,tol,maxtis, status,root,resid)
t2=root
end if
return
end

subroutine bisect(f,k,inx0,inx1,tol,maxtis, status,root,resid)
integer count,limit,maxtis,solved,nobrac,status
real f,f0,f1,inx1,inx0,lastx0,lastx1,x1,x0,xmid,k
parameter (solved=0,limit=1,nobrac=2)
intrinsic abs
external f
x0=inx0
x1=inx1
f0=f(k,x0)
f1=f(k,x1)
if (f0*f1 .gt. 0) then
status=nobrac
else
   do 10 count=1,maxtis
      xmid=(x0+x1)/2
      fmid=f(k,xmid)
      if (fmid .eq. 0) then
         status=solved
         go to 11
      endif
      if (f0*fmid .lt. 0 ) then
         x1=xmid
      endif
   enddo
10 continue
endif
end
endif

if (f0*fmid > 0) then
  x0=xmid
  f0=f(k,x0)
endif

if (abs(x0-xl) < tol) then
  status=solved
  go to 11
endif

continue

status=limit

if (status eq solved) then
  root=(x0+x1)/2
  resid=f(k,root)
endif

if (status eq limit) then
  lastx0=x0
  lastx1=x1
endif

endif

return

end
4. The following Fortran program is used to calculate the power of the minimax test at $\bar{F}_1$.

dimension ak(33),aw(33),tt(33),tem(34)
    real lamda0, lamda1,mu,inx0,inx1,k,k1,b1,b2,d1,d2,root
    integer status,solved,count
    external t3,t1,t2,bisect,wt2
    intrinsic exp
    parameter (e=0.1,lamda0=3.0,lamda1=2.0)
    mu=0.4
    b1=0.945333
    b2=1.100642
    d1=t3(b1)
    d2=t3(b2)
    k1=exp(-lamda1*d1)/(exp(-lamda0*d1)+e/(1-e))
    do 70 i=1,31
      ai=i-1
      ak(i)=(ai/30.0)*b1+b2*((30.0-ai)/30.0)
      k=ak(i)
      inx0=d1
      inx1=d2
      tol=1e-6
      maxtis=30
tt(i) = t2(bisect, wt2, k, inx0, inx1, tol, maxtis, status, root, c resid, k1, b2, d1, d2)

ttt2 = tt(i)

www1 = exp(-(lamda1 + mu) * ttt1)

www2 = exp(-(lamda1 + mu) * ttt2)

www3 = exp(-(lamda1 + mu) * ttt3)

if (k .gt. b2) then
  w = 0.0
else if (k .ge. k1 .and. k .le. b2) then
  w = 1.0 / (lamda1 + mu) * (lamda1 * www3 + mu * www2)
else if (k .ge. 1.0 .and. k .lt. k1) then
  w = 1.0 / (lamda1 + mu) * (lamda1 * www3 + mu * www1)
else if (k .ge. b1 .and. k .lt. 1.0) then
  w = 1.0 / (lamda1 + mu) * (lamda1 * www3) + mu / (lamda1 + mu)
else
  w = 1.0
endif

aw(i) = w

write(6, 80) aw(i), ak(i)

80 format(f8.6, 4x, f8.6)

70 continue
888     continue
     stop
     end

     real function wt2 (k,t)
     real k,lamda1,lamda0
     parameter (lamda0=3.0,lamda1=2.0,e=0.1)
     intrinsic exp
     wt2=exp(-lamda1*t)/(exp(-lamda0*t)+e/(1-e))-k
     return
     end

     real function t3 (k)
     real lamda0,lamda1,k
     intrinsic alog
     parameter (lamda0=3.0,lamda1=2.0)
     t3=1.0/(lamda0-lamda1)*alog(k*lamda0/lamda1)
     return
     end

     real function t1 (k,b1,k1,d1)
     real k,k1
     intrinsic alog
     parameter (lamda0=3.0,e=0.1)
     if (k .eq. k1) then
     t1=d1
end if
if (k .eq. 1.0) then
    t1=0.0
end if
if (k .ge. 1.0 .and. k .lt. k1) then
    t1=-(1/lambda0)*alog((1-b1)/(k-b1)*(1-e)-e/(1-e))
end if
return
end

real function t2 (bisect, wt2, k, inx0, inx1, tol, maxtis, c status, root, resid, k1, b2, d1, d2)
real d1, d2, inx0, inx1, k, k1, b2
integer maxtis
external bisect, wt2
if (k .eq. k1) then
    t2=d1
end if
if (k .eq. b2) then
    t2=d2
end if
if (k .gt. k1 .and. k .lt. b2) then
    call bisect (wt2, k, inx0, inx1, tol, maxtis, status, root, resid)
    t2=root
end if
return
end

subroutine bisect(f,k,inx0,inx1, tol, maxtis, status, root, resid)
integer count, limit, maxtis, solved, nobrac, status
real f, f0, f1, inx0, inx1, lastx0, lastx1, x1, x0, xmid, k
parameter (solved=0, limit=1, nobrac=2)
intrinsic abs
external f

x0 = inx0
x1 = inx1
f0 = f(k, x0)
f1 = f(k, x1)
if (f0 * f1 .gt. 0) then
    status = nobrac
else
    do 10 count = 1, maxtis
        xmid = (x0 + x1)/2
        fmid = f(k, xmid)
        if (fmid .eq. 0) then
            status = solved
            go to 11
        endif
    enddo
endif
if (f0*fmid .lt. 0) then
  x1=xmid
endif

if (f0*fmid .gt. 0) then
  x0=xmid
  f0=f(k,x0)
endif

if ( abs(x0-x1) .lt. tol ) then
  status=solved
  go to 11
endif

10 continue
  status=limit
11 if ( status .eq. solved) then
  root=(x0+x1)/2
  resid=f(k,root)
endif

if (status .eq. limit) then
  lastx0=x0
  lastx1=x1
endif
endif
5. The following program is used to calculate the minimal power of the minimax test over the alternative contamination neighborhood.

```fortran
dimension ak(33),aw(33),tt(33),tem(34)
real lamda0, lamda1, mu,inx0,inx1,k,k1,b1,b2,d1,d2,root
integer status,solved,count
external t3,t1,t2,bisect,wt2
intrinsic exp
parameter (e=0.1,lamda0=3.0,lamda1=2.0)
mu=3.0
b1=0.945333
b2=1.100642
d1=t3(b1)
d2=t3(b2)
k1=exp(-lamda1*d1)/(exp(-lamda0*d1)+e/(1-e))
do 70 i=1,31
ai=i-1
ak(i)=(ai/30.0)*b1+b2*((30.0-ai)/30.0)
k=ak(i)
inx0=d1
```

return
end

inx1=d2
tol=1e-6
maxtis=30
    tt(i)=t2(bisect,wt2,k,inx0,inx1,tol,maxtis, status,root,resid,
c k1,b2,d1,d2)
    ttt2=tt(i)
    ttt1=t1(k,b1,k1,d1)
    ttt3=t3(k)
    www1=exp(-(lamda1+mu)*ttt1)
    www2=exp(-(lamda1+mu)*ttt2)
    www3=exp(-(lamda1+mu)*ttt3)
    ter1=mu*(1-e)/(lamda0+mu)*(1-exp((-d1)*(lamda0+mu)))
    ter2=(1-e)*mu/(lamda1+mu)*exp(-1-(lamda1+mu)*d1)
    term1=ter1+ter2
    term2=(1-b1*(1-e))*(1-exp(-mu*d1))
if (k .gt. b2) then
        w=0.0
else if (k .ge. k1 .and. k .le. b2) then
        w=(1-e)/(lamda1+mu)*(lamda1*www3+mu*www2)
else if (k .ge. 1.0 .and. k .lt. k1) then
        w=(1-e)/(lamda1+mu)*(lamda1*www3+mu*www1)
else if (k .ge. b1 .and. k .lt. 1.0) then
        w=(1-e)/(lamda1+mu)*(lamda1*www3)+term1+term2
else
w=1.0
endif
aw(i)=w
write(6,80)aw(i),ak(i)
80 format(f8.6,4x,f8.6)
70 continue
888 continue
stop
end

real function wt2 (k,t)
real k,lamda1,lamda0
parameter (lamda0=3.0,lamda1=2.0,e=0.1)
intrinsic exp
wt2=exp(-lamda1*t)/(exp(-lamda0*t)+e/(1-e))-k
return
end

real function t3 (k)
real lamda0,lamda1,k
intrinsic alog
parameter (lamda0=3.0,lamda1=2.0)
t3=1.0/(lamda0-lamda1)*alog(k*lamda0/lamda1)
return
end

real function t1 (k,b1,k1,d1)
real k,k1
intrinsic alog
parameter (lamda0=3.0,e=0.1)
if (k .eq. k1) then
  t1=d1
end if
if (k .eq. 1.0) then
  t1=0.0
end if
if (k .ge. 1.0 .and. k .lt. k1) then
  t1 =-(1/lamda0)*alog((1 -b1 )/(k-b1 )*(1 -e)-e/(1 -e))
end if
return
end

real function t2 (bisect,wt2,k,inx0,inx1,tol,maxtis,
c status,root,resid, k1,b2,d1,d2)
real d1,d2,inx0,inx1,k,k1,b2
integer maxtis
external bisect,wt2
if (k .eq. k1) then
  t2=d1
end if
if (k .eq. b2) then
  t2=d2
end if
if (k .gt. k1 .and. k .lt. b2) then
  call bisect (wt2,k,inx0,inx1,tol,maxtis, status,root,resid)
  t2=root
end if
return
end

subroutine bisect(f,k,inx0,inx1,tol,maxtis, status,root,resid)
integer count,limit,maxtis,solved,nobrac,status
real f,f0,f1,inx0,inx1,lastx0, lastx1,x1,x0,xmid,k
parameter (solved=0,limit=1,nobrac=2)
intrinsic abs
external f
x0=inx0
x1=inx1
f0=f(k,x0)
f1=f(k,x1)
if (f0*f1 .gt. 0) then
  status=nobrac
else
do 10 count=1,maxtis
  xmid=(x0+x1)/2
  fmid=f(k,xmid)
  if (fmid .eq. 0) then
    status=solved
    go to 11
  endif
  if (f0*fmid .lt. 0 ) then
    x1=xmid
  endif
  if (f0*fmid .gt. 0) then
    x0=xmid
    f0=f(k,x0)
  endif
  if ( abs(x0-x1) .lt. tol ) then
    status=solved
    go to 11
  endif
  10 continue
  status=limit
  11 if ( status .eq. solved) then
    root=(x0+x1)/2
6. The following program is used to calculate the average of 100 estimated OC-Curves and the nominal OC-Curve going through $1-\alpha$ at $\theta_0$.

```fortran
resid=f(k,root)
endif
if (status .eq. limit) then
  lastx0=x0
  lastx1=x1
endif
endif
return
end

dimension z(500),ir(1000),ngn(1000,500),xnew(4567)
dimension est(100,500),theta(500),xp(1000,500)
dimension stat(1000,500),a(15),sum(300)
dimension dis(324)
integer nr,iseed, nok
real xtheta,z,x
nth=17
do 2000 i=1,nth
  theta(i)=120.00-5.0*i
2000 continue
```
2000 continue
    m=15
    do 9000 nok=2,14
       ak=30.0+20.0*nok
       dis(nok)=0.0
    nr=100
    xnr=nr
    am=m
    algk=alog(ak)
    delta=4.0
    del=delta
    dg=gamma(delta)
    xy=am*ak/algk
      write(6,*) ' '
      write(6,5000) delta,m,ak
5000  format(2x,'delta=',f4.2,2x,'m=',i3,3x,'k=',f6.1)
    mstar=int(xy)+1
    do 850 j=1,nth
       v=theta(j)
       vV=(del**(del-1))/dg
          z(j)=(100.0/v)**del*(am*ak/alog(ak)**del)*vV
    850   continue
    do 2 j=1,nth
iseed=0

xtheta=z(j)
call rnset(iseed)
call rnpoi(nr,xtheta,ir)
do 1 i=1,nr
   ngn(i,j)=ir(i)
1  continue
2  continue
do 4 i=1,mstar
   ai=i
   y=ai
   n=1
   x=1.20
3  aag=gamma(x)
   ag=((ak*am)/(aag*x))*(x/algk)**x
   af=ag-y
   adf=(alog(x)-alog(algk)-1.0/x+1.0-psi(x))*ag
   xnew(i)=x-af/adf
   if(abs(x-xnew(i)) .lt. 1e-6 .or. n .gt. 20) go to 4
   n=n+1
   x=xnew(i)
go to 3
4  continue
do 9 j=1,nth
  do 8 i=1,nr
    if(ngn(i,j) .eq. 0)then
      est(i,j)=mstar
    else
      if(ngn(i,j) .gt. mstar)then
        est(i,j)=1
      else
        do 5 k=1,mstar
           if(ngn(i,j) .eq. k) go to 88
        5 continue
    88 est(i,j)=xnew(k)
    endif
  endif
8 continue
9 continue
  write(6,5001)
5001 format(1x,'theta',4x,'Ave E.OC-Curve',4x,'N.OC-Curve')
  write(6,*)
end
7. The following program is used to calculate the average of 100 estimated OC-Curves and the nominal OC-Curve going through $\beta$ at $\theta_1$.

```
dimension z(500),ir(1000),ngn(1000,500),xnew(4567),
dimension est(100,500),theta(500),xp(1000,500)
dimension stat(1000,500),a(15),sum(300)
dimension dis(324)
integer nr,iseed, nok
real xtheta,z,x
nth=17
do 2000 i=1,nth
   theta(i)=120.00-5.0*i
2000 continue
m=15
do 9000 nok=2,14
   ak=30.0+20.0*nok
   dis(nok)=0.0
9000 nr=100
xnr=nr
am=m
algk=alog(ak)
delta=4.0
del=delta
```


dg=gamma(delta)

xy=am*ak/alogk

write(6,*) ' '

write(6,5000) delta,m,ak

5000 format(2x,'delta=',f4.2,2x,'m=',i3,3x,'k=',f6.1)

mstar=int(xy)+1

do 850 j=1,nth

v=theta(j)

vv=(del**(del-1)/dg)

z(j)=(100.0/v)**(am*ak/alog(ak)**del)*vv

850 continue

do 2 j=1,nth

iseed=0

xtheta=z(j)

call rnset(iseed)

call rnpoi(nr,xtheta,ir)

do 1 i=1,nr

ngn(i,j)=ir(i)

1 continue

2 continue

do 4 i=1,mstar

ai=i

y=ai
n=1
x=1.20

3 aag=gamma(x)
   ag=((ak*am)/(aag*x))*(x/algk)**x
af=ag-y
   adf=(alog(x)-alog(algk)-1.0/x+1.0-psi(x))*ag
xnew(i)=x-af/adf
if(abs(x-xnew(i)) .lt. 1e-6 .or. n .gt. 20) go to 4
n=n+1
x=xnew(i)
go to 3

4 continue
do 9 j=1,nth
do 8 i=1,nr
   if(ngn(i,j) .eq. 0)then
      est(i,j)=mstar
   else
      if(ngn(i,j) .gt. mstar)then
         est(i,j)=1
      else
         do 5 k=1,mstar
            if(ngn(i,j) .eq. k) go to 88
         5 continue
   end if
end do
end do

88 \text{est}(i,j)=\text{xnew}(k)

\text{endif}

\text{endif}

8 \text{ continue}

9 \text{ continue}

\text{write}(6,5001)

5001 \text{ format}(1x,'\theta',4x,'\text{Ave E.OC-Curve}',4x,'\text{N.OC-Curve}')

\text{write}(6,*)

\text{end}

8. The following program is used to calculate the distance between the average of 100 estimated QC-Curves and the nominal OC-Curve going through 1-\alpha at \theta_0.

dimension z(500),ir(1000),ngn(1000,500),xnew(4567),

dimension est(100,500),theta(500),xp(1000,500)

dimension stat(1000,500),a(15),sum(300)

dimension dis(324)

integer nr,iseed, nok

real xtheta,z,x

nth=17

do 2000 i=1,nth

\theta(i)=120.00-5.0*i
2000 continue
    m=15
    do 9000 nok=2,14
       ak=30.0+20.0*nok
       dis(nok)=0.0
    nr=100
    xnr=nr
    am=m
    algk=alog(ak)
    delta=4.0
    del=delta
dg=gamma(delta)
    xy=am*ak/algk
    write(6,*) ' '  
    write(6,5000) delta,m,ak
5000 format(2x,'delta=',f4.2,2x,'m=',i3,3x,'k=',f6.1)
    mstar=int(xy)+1
    do 850 j=1,nth
       v=theta(j)
       vv=(del**(del-1)/dg)
       z(j)=(100.0/v)**del*(am*ak/alog(ak)**del)* vv
    850 continue
    do 2 j=1,nth
iseed=0
xtheta=z(j)
call rnset(iseed)
call rnpoi(nr,xtheta,ir)
do 1 i=1,nr
   ngn(i,j)=ir(i)
1  continue
2  continue
do 4 i=1,mstar
   ai=i
   y=ai
   n=1
   x=1.20
3  aag=gamma(x)
   ag=((ak*am)/(aag*x))*(x/algk)**x
   af=ag-y
   adf=(alog(x)-alog(algk)-1.0/x+1.0-psi(x))*ag
   xnew(i)=x-af/adf
   if(abs(x-xnew(i)) .lt. 1e-6 .or. n .gt. 20) go to 4
   n=n+1
   x=xnew(i)
go to 3
4  continue
do 9 j=1,nth
 do 8 i=1,nr
   if(ngn(i,j) .eq. 0)then
     est(i,j)=mstar
   else
     if(ngn(i,j) .gt. mstar)then
       est(i,j)=1
     else
       do 5 k=1,mstar
         if(ngn(i,j) .eq. k) go to 88
       5 continue
     endif
   endif
  8 continue
 9 continue
a(1)=4.45
a(2)=2.871
a(3)=2.272
a(4)=1.937
a(5)=1.716
a(6)=1.557
a(7)=1.435
a(8)=1.338
a(9)=1.258
a(10)=1.191
a(11)=1.134
a(12)=1.084
a(13)=1.040
a(14)=1.001
a(15)=0.966

xbar=0.00
do 31 j=1,nth
  do 29 i=1,nr
    if (est(i,j) .ge. a(1)) then
      df=2
    else
      do 24 k=2,16
        if (est(i,j) .ge. a(k) .and. est(i,j) .lt. a(k-1)) then
          go to 25
        else
          continue
        endif
      24  continue
    endif
  df=2*(k)
25  endif
p=0.1
aa=chiin(p,df)
    bb=(100.00/theta(j))**est(i,j)*aa
p1=chidf(bb,df)
xp(i,j)=1-p1
29  continue
    sum(j)=0.00
    do 42 i=1,nr
        sum(j)=sum(j)+xp(i,j)
42  continue
    sum(j)=sum(j)/nr
    if (delta .ge. a(1))then
        dff=2
    else
        do 50 k=2,16
        if (delta .ge. a(k) .and. delta .lt. a(k-1))then
            go to 51
        else
            50 continue
41  continue
42  do 50 k=2,16
        if (delta .ge. a(k) .and. delta .lt. a(k-1))then
            go to 51
        else
            50 continue
51  dff=2*(k)
    endif
endif
dd=chiin(p,dff)
cc=(100.00/theta(j))**delta*dd
p3=chidf(cc,dff)
p4=1-p3
write(6,30)theta(j),sum(j),p4
30 format(f6.2,','f8.6,','f8.6)
dis(nok)=dis(nok)+(sum(j)-p4)**2*5.0
31 continue
9000 continue
write(6,9010)delta
9010 format(2x,fn6.2)
write(6,9013)
9013 format( 2x,'Dist of AEOC and NOC')
do 9011 ijk=2,14
write(6,9012)dis(ijk)
9012 format(2x,f8.3)
9011 continue
end
9. The following program is used to calculate the distance between the average of 100 estimated OC-Curves and the nominal OC-Curve going through \( \beta \) at \( \theta_1 \).

dimension z(500),ir(1000),ngn(1000,500),xnew(4567)
dimension est(100,500),theta(500),xp(1000,500)
dimension stat(1000,500),a(15),sum(300)
dimension dis(324)
integer nr,iseed, nok
real xtheta,z,x

nth=17

do 2000 i=1,nth
    theta(i)=120.00-5.0*i

2000 continue
m=15

do 9000 nok=2,14
    ak=30.0+20.0*nok
    dis(nok)=0.0

nr=100
xnr=nr
am=m
algk=alog(ak)
delta=4.0
del=delta
dg = gamma(d, l)
xy = am * ak / al(gk)
write(6, *) ' ' 
write(6, 5000) delta, m, ak
5000  format(2x, 'delta=', f4.2, 2x, 'm=', i3, 3x, 'k=', f6.1)
mstar = int(xy) + 1
do 850 j = 1, nth
   v = theta(j)
vv = (d**(d-1)/dg)
z(j) = (100.0/v)**d*(am*ak ALOG(ak)**d) * vv
850  continue
do 2 j = 1, nth
   iseed = 0
   xtheta = z(j)
call rnset(iseed)
call mpoi(nr, xtheta, ir)
do 1 i = 1, nr
   ngn(i,j) = ir(i)
1  continue
2  continue
do 4 i = 1, mstar
   ai = i
   y = ai
n=1
x=1.20
3  aag=gamma(x)
   ag=((ak*am)/(aag*x))*(x/algl)**x
af=ag-y
   adf=(alog(x)-alog(algl)-1.0/x+1.0-psi(x))*ag
   xnew(i)=x-af/adf
   if(abs(x-xnew(i)) .lt. 1e-6 .or. n .gt. 20) go to 4
n=n+1
   x=xnew(i)
go to 3
4  continue
do 9 j=1,nth
do 8 i=1,nr
   if(ngn(i,j) .eq. 0)then
      est(i,j)=mstar
   else
      if(ngn(i,j) .gt. mstar)then
         est(i,j)=1
      else
         do 5 k=1,mstar
            if(ngn(i,j) .eq. k) go to 88
   continue
est(i,j)=xnew(k)
endif
endif
continue
continue
a(1)=4.45
a(2)=2.871
a(3)=2.272
a(4)=1.937
a(5)=1.716
a(6)=1.557
a(7)=1.435
a(8)=1.338
a(9)=1.258
a(10)=1.191
a(11)=1.134
a(12)=1.084
a(13)=1.040
a(14)=1.001
a(15)=0.966
xbar=0.00
do 31 j=1,nth
do 29 i=1,nr
   if (est(i,j) .ge. a(1)) then
      df=2
   else
      do 24 k=2,16
         if (est(i,j) .ge. a(k) .and. est(i,j) .lt. a(k-1)) then
            go to 25
         else
            24 continue
         endif
      25 df=2*(k)
   endif
   endif
   p=0.1
   aa=chiin(p,df)
   bb=(100.00/theta(j))**est(i,j)*aa
   p1=chidf(bb,df)
   xp(i,j)=1-p1
29 continue
sum(j)=0.00
do 42 i=1,nr
   sum(j)=sum(j)+xp(i,j)
42 continue
sum(j)=sum(j)/nr
if(delta .ge. a(1)) then
  dff=2
else
  do 50 k=2,16
    if(delta .ge. a(k) .and. delta .lt. a(k-1)) then
      go to 51
    else
      50 continue
    endif
  endif
  dff=2*(k)
endif
dd=chiin(p,dff)
cc=(100.00/theta(j))**delta*dd
p3=chidf(cc,dff)
p4=1-p3
write(6,30)theta(j),sum(j),p4
30  format(f6.2,",","f8.6","f8.6")
dis(nok)=dis(nok)+(sum(j)-p4)**2*5.0
31  continue
9000  continue
write(6,9010)delta
9010  format(2x,fn6.2)
write(6,9013)
9013 format( 2x,'Dist of AEOC and NOC')
   do 9011 ijk=2,14
      write(6,9012)dis(ijk)
9012 format(2x,f8.3)
9011 continue
   end
Appendix B

Figures

Figure 4. Asymptotic Relative Efficiency of MAD w.r.t. MSD
Figure 5. Levels and powers of Sign test and Wilcoxon test
1 = level of S-test, 2 = level of W-test, 3 = Power of S-test, 4 = Power of W-test
Figure 6. Significant levels, Lambda=0.3, Lambda=1.2, Lambda=3.0.
Figure 7. Minimal and ideal power, $\Lambda_0=3.0$, $\Lambda_1=2.0$
Broken=Minimal Power, Solid=Ideal Power
PLEASE NOTE:

Page(s) not included with original material and unavailable from author or university. Filmed as received.
Figure 10. Average of 100 Estimated OC-Curves and Nominal OC-Curve, delta=2.0
Figure 11. Average of 100 Estimated OC-Curves and Nominal OC-Curve, \( \delta = 3.0 \)
Figure 12. Average of 100 Estimated OC-Curves and Nominal OC-Curve, delta=4.0
LIST OF REFERENCES


