AN ABSTRACT OF THE THESIS OF

ANEK HIRUNRAKS for the Doctor of Philosophy
(Name) (Degree)

in Statistics presented on August 13, 1969
(Major) (Date)

Title: NONPARAMETRIC c-SAMPLE TESTS FOR ARBITRARILY CENSORED DATA

Abstract approved: Redacted for Privacy
David R. Thomas

Consider the c-sample testing problem with null hypothesis $H_o : F_1(x) = \cdots = F_c(x)$, where $F_j(x)$, $j = 1, \cdots, c$, represent absolutely continuous distribution functions for failure distributions. Furthermore, consider arbitrarily (right) censored samples, which may arise from putting the (censored) experimental units on test at different times.

The c-sample test that we shall be primarily concerned with is an extension of a generalized Savage test proposed by Thomas [25] for the two-sample case. This c-sample test statistic, $Q_{(r, N)}$, is shown to have an asymptotic chi-square distribution under H_o and an asymptotic non-central chi-square distribution locally under alternatives with proportional hazard functions. The $Q_{(r, N)}$ statistic is shown to give an asymptotic efficient test with respect to the likelihood ratio test for exponential failure distributions with

exponential censoring distributions. Conditional small sample power comparisons of $Q_{(r, N)}$ are also made with a c-sample extension of Gehan's [11] two-sample Wilcoxon type statistic.

Nonparametric c-Sample Tests for Arbitrarily Censored Data

by

Anek Hirunraks

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1970

APPROVED:

Redacted for Privacy

Assistant Professor of Statistics

in charge of major

Redacted for Privacy

Chairman of Department of Statistics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented August 13, 1969

Typed by Clover Redfern for Anek Hirunraks

ACKNOWLEDGMENTS

The author would like to express his sincere appreciations to Dr. David R. Thomas for all of his advice, understanding and patience for making this dissertation possible.

Thanks are also extended to "The Ford Foundation" for the sponsorship of this program. And special thanks are due to my wife, Wantana, for her understanding, encouragement and patience throughout this course of study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NONPARAMETRIC c-SAMPLE TESTS FOR ARBITRARILY CENSORED DATA

I. INTRODUCTION

A common problem arising in life testing and clinical trials is the comparison of failure time distributions, or recovery time distributions, for c-treatments. Frequently, the experimenter wishes to compare the treatments (drug, components, etc.) before all experimental units have failed, resulting in censored observations. Due to limited testing facilities or availability of experimental units, it may be impossible to put all units on test at the same time. Consequently, censored observations may not all be subjected to the same amount of test time, arbitrary censoring.

Motivated by clinical trials with arbitrarily singly-censored data, Gehan [11] developed a generalized Wilcoxon test for the null hypothesis $H_o : G(t) = F(t)$, where $G(t)$ and $F(t)$ represent the distribution functions for the two failure distributions.

Efron [7] has also proposed a test statistic $\hat{W}$ which is the maximum likelihood estimate for $Pr(X > Y)$ with arbitrarily censored data. He considers the case where the censoring distributions may be different for the two failure distributions. However, we will only be concerned with the case where all observations are censored according to the same censoring distribution. For an exponential censoring distribution, the Pitman asymptotic relative efficiency (A. R. E.) of the W test with respect to the likelihood ratio test is given in Table II of Efron [7]. For a particular censoring model, discussed in Chapter VI, the generalized Savage test statistic is shown to be asymptotically more efficient than the Efron statistic. The generalized Savage statistic has the additional advantage that its variance under H_o is easily evaluated, whereas the variance of the Efron statistics are difficult to estimate. On the other hand, $\hat{W}$ has the advantage of providing an estimate of $Pr(X > Y)$ under the alternative $F \not\equiv G$.

Thomas [25] has developed conditional locally most powerful rank tests (c. l. m. p. r. t. 's) for the two-sample problem with arbitrarily censored data. These test statistics are linear functions of rank-order statistics. The coefficients in these linear rank statistics can easily be evaluated for the case of proportional hazard function alternatives. The resulting statistic will be referred to as the "generalized Savage statistic." Thomas [25] has also discussed how the generalized Savage test may be modified for the case of grouped failures (ties).

Puri [20] has studied a class of test statistics $\mathcal{L}$ for the c-sample problem in the uncensored case. These $\mathcal{L}$-statistics are quadratic functions of linear rank statistics. Puri proves that these $\mathcal{L}$-statistics have asymptotic chi-square distributions by approximating the coefficients in the linear rank statistics by a continuous weighting function.

In this thesis, we consider a class of c-sample nonparametric tests for arbitrarily censored data. The null hypothesis is $H_o : F_1(t) \equiv F_2(t) \equiv \cdots \equiv F_c(t)$ against H_a: not all $F_j(t)$'s are equal. The test statistic $Q_{(r, N)}$, given in (4.3), is in the form of a Puri's $\mathcal{L}$-statistic. The linear rank statistics used in $Q_{(r, N)}$ are the generalized Savage statistics which will be derived as an extension of the Thomas c.l.m.p.r.t.'s for the two-sample case. The $Q_{(r, N)}$ tests are shown to be asymptotic efficient with respect to the parametric likelihood ratio test for exponential censoring distributions and exponential failure time distributions.

In Chapter II, the problem is formulated, assumptions are given, and order and rank-order statistics are defined for arbitrarily censored data. Distributions of order and rank-order statistics are derived in Chapter III. In Chapter IV, the vector of generalized Savage statistics is given for distributions with proportional hazard functions. The $Q_{(r, N)}$ statistic is also defined in Chapter IV.

Asymptotic normality of the generalized Savage statistic vector, $\underset{\sim}{T}_{(r, N)}$, is established in Chapter V under H_o and (locally) under H_a. For exponential failure time distributions with independent exponential censoring distributions, the $Q_{(r, N)}$-test is shown in Chapter VI to be asymptotic efficient with respect to the likelihood ratio test.

Numerical results are presented in Chapter VII. The exact and chi-square approximation of upper-tailed probabilities are given in Tables 2, 3 and 4 for three different censoring patterns in the three-sample case. Power curves of $Q_{(r, N)}$ tests are given in Figures 1 and 2 corresponding to test-sizes less than 0.14. Figures 3 and 4 give graphical power comparisons of $Q_{(r, N)}$ and a corresponding quadratic function $Q'_{(r, N)}$ of Gehan's generalized Wilcoxon statistics. At the end of Chapter VII, an example is given illustrating the computation of the test statistic $Q_{(r, N)}$ for the four-sample problem.

II. FORMULATION OF THE PROBLEM

Let there be random samples with $n_1, n_2, \ldots, n_c$ observations from the absolutely continuous distribution functions $F_1(t), F_2(t), \ldots, F_c(t)$. With arbitrarily censored data, we desire to test the null hypothesis,

$$H_o : F_1(t) \equiv F_2(t) \equiv \ldots \equiv F_c(t) \qquad (2.1)$$

against the alternative,

$$H_a : \text{not all } F_j(t) \text{ are equal; } j = 1, 2, \ldots, c. \qquad (2.2)$$

Denote the time on test for the failure observations and the censored observations from $F_1, F_2, \ldots, F_c$ respectively by $x^{(1)}, x^{(2)}, \ldots, x^{(c)}$ and $x^{(1)'}, x^{(2)'}, \ldots, x^{(c)'}$. Let there be r failures observed in the combined sample of $N = \sum_{j=1}^{c} n_j$ observations. Represent these r ordered failures by $t_1, t_2, \ldots, t_r$. Assuming no ties, then we have $0 \equiv t_o < t_1 < t_2 < \ldots < t_r < t_{r+1} \equiv \infty$. Define

c_i = total number of censored observations between the failure times t_i and t_{i+1}, $i = 0, 1, 2, \ldots, r$.

$u_i^{(j)}$ = number of censored $x^{(j)'}$ observations with $t_i < x^{(j)'} < t_{i+1}$, $i = 0, 1, 2, \ldots, r$.

$\xi_i^{(j)}$ = 1 (0) if t_i is an $x^{(j)}$ failure (if not), $i = 1, 2, \ldots, r$. (2.3)

Furthermore, let

$$\underset{\sim}{t}_{(i)} = (t_1, t_2, \ldots, t_i)$$

$$\underline{\xi}_{(i)} = \begin{bmatrix} \xi_1^{(1)} & \xi_2^{(1)} & \cdots & \xi_i^{(1)} \\ \xi_1^{(2)} & \xi_2^{(2)} & \cdots & \xi_i^{(2)} \\ \vdots & \vdots & & \vdots \\ \xi_1^{(c-1)} & \xi_2^{(c-1)} & \cdots & \xi_i^{(c-1)} \end{bmatrix},$$

$$i = 1, 2, \ldots, r, \qquad (2.4)$$

and

$$\underset{\sim}{C}_{(i, N)} = (c_o, c_1, c_2, \ldots, c_i)$$

$$\underline{u}_{(i)} = \begin{bmatrix} u_o^{(1)} & u_1^{(1)} & \cdots & u_i^{(1)} \\ u_o^{(2)} & u_1^{(2)} & \cdots & u_i^{(2)} \\ \vdots & \vdots & & \vdots \\ u_o^{(c-1)} & u_1^{(c-1)} & \cdots & u_i^{(c-1)} \end{bmatrix},$$

$$i = 0, 1, 2, \ldots, r. \qquad (2.5)$$

The random vector $\underset{\sim}{t} \equiv \underset{\sim}{t}_{(r)}$ and the random matrix $(\underline{\xi}, \underline{u}) \equiv (\underline{\xi}_{(r)}, \underline{u}_{(r)})$ will be called respectively "order" and "rank-order" statistics. The vector $\underset{\sim}{C}_{(r, N)}$ will be called the "censoring pattern" of the observations.

From (2.3), (2.4) and (2.5) we may note that

$$\xi_i^{(c)} = 1 - \sum_{j=1}^{c-1} \xi_i^{(j)}, \quad i = 1, 2, \ldots, r, \tag{2.6}$$

$$u_i^{(c)} = c_i - \sum_{j=1}^{c-1} u_i^{(j)}, \quad i = 0, 1, 2, \ldots, r, \tag{2.7}$$

$$u_o^{(j)} + \sum_{i=1}^{r} (\xi_i^{(j)} + u_i^{(j)}) = n_j, \; j = 1, 2, \cdots, c, \tag{2.8}$$

and

$$\sum_{j=1}^{c} u_o^{(j)} + \sum_{i=1}^{r} \sum_{j=1}^{c} (\xi_i^{(j)} + u_i^{(j)}) = c_o + \sum_{i=1}^{r} (c_i + 1)$$

$$= N. \tag{2.9}$$

The following assumptions are used in deriving the tests: (2.10)

A. Completely random allocation of treatments to experimental units.

B. Distributions of observations are independent of time put on test.

C. (**Only required for optimum properties.**) The observations are censored according to independent identically distributed random variables v_{ij}, $i = 1, \ldots, n_j$, $j = 1, \ldots, c$, that is,

$$x_i^{(j)} \text{ is a failure for } x_i^{(j)} \leq v_{ij}$$

$$x_i^{(j)'} = v_{ij} \text{ is a censored observation for } v_{ij} < x_i^{(j)}.$$

Furthermore, it is assumed that the probability density function $h(v)$ for v satisfies: $h(v) > 0$ for all v for which $f^{(j)}(v) > 0$, $j = 1, 2, \ldots, c$.

Note that time censored samples do not satisfy assumption C. Assumption C will only be required for optimum properties of the tests derived. Assumption B and C could be replaced by the assumption of progressively censored samples, that is, all units are put on test at the same time and after the i^{th} failure occurs a random c_i of the units are censored from the remaining $\sum_{j=i}^{r} (c_j+1) - 1$ units which have neither failed nor been censored at time t_i, $i = 1, 2, \ldots, r$.

III. CONDITIONAL DISTRIBUTIONS OF ORDER AND RANK-ORDER STATISTICS

In this chapter, the conditional distributions of order and rank-order statistics for a given censoring pattern, $\underset{\sim}{C}_{(r, N)}$, are derived. The following results are extensions of those obtained by Thomas [25] for the two-sample case.

Theorem 3.1. Under assumptions (2.10), the conditional density function of $(\underset{\sim}{t}, \underline{\xi}, \underline{u})$ is given by

$$f(\underset{\sim}{t}, \underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r, N)}, \underset{\sim}{F}) = \binom{N}{n_1, n_2, \ldots, n_{c-1}}^{-1} \prod_{i=0}^{r} \binom{c_i}{u_i^{(1)}, u_i^{(2)}, \ldots, u_i^{(c-1)}}$$
$$\times \left(\prod_{i=1}^{r} E_i \right) \left[\prod_{i=1}^{r} \prod_{j=1}^{c} \left\{ (1 - F_j(t_i))^{u_i^{(j)}} f_j(t_i)^{\xi_i^{(j)}} \right\} \right], \tag{3.1}$$

where

$$E_i = \sum_{j=i}^{r} (c_j + 1), \tag{3.2}$$

$$0 \equiv t_0 < t_1 < t_2 < \ldots < t_r < t_{r+1} \equiv \infty,$$

and $\underline{\xi}, \underline{u}$ are as on page 6.

Proof. The density function $f(\underset{\sim}{t}, \underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r, N)}, \underset{\sim}{F})$ can be expressed as a product of conditional densities:

$$f(\underset{\sim}{t}, \underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r, N)}, \underset{\sim}{F}) = B_o \prod_{i=1}^{r} (A_i B_i) \tag{3.3}$$

with

$$A_i = f(t_i, \underset{\sim}{\xi}_i \mid \underset{\sim}{t}_{(i-1)}, \underline{\xi}_{(i-1)}, \underline{u}_{(i-1)} \mid \underset{\sim}{C}_{(r,N)}, \underset{\sim}{F})$$

$$B_i = p(\underset{\sim}{u}_i \mid \underset{\sim}{t}_{(i-1)}, \underline{\xi}_{(i-1)}, \underline{u}_{(i-1)} \mid \underset{\sim}{C}_{(r,N)}, \underset{\sim}{F}), \qquad (3.4)$$

where $\underset{\sim}{t}_{(i)}, \underline{\xi}_{(i)}, \underline{u}_{(i)}$ are given in (2.4) and (2.5) and $\underset{\sim}{\xi}_i, \underset{\sim}{u}_i$ are the $i^{\underline{th}}$ column vectors of the matrices of $\underline{\xi}$ and $\underline{u}$ respectively.

Let

$$n_{j1} = n_j - u_o^{(j)}$$

$$n_{j\alpha} = n_{j1} - \sum_{\beta=1}^{\alpha-1} (\xi_\beta^{(j)} + u_\beta^{(j)}),$$

$$\alpha = 2, 3, \ldots, r; \quad j = 1, 2, \ldots, c \qquad (3.5)$$

be the number of observations from the $j^{\underline{th}}$ population which have neither failed with $x^{(j)} < t_\alpha$ nor been censored with $x^{(j)'} < t_\alpha$.

Now, it is seen that the conditional density for the minimum t_i of n_{ji} $x^{(j)}$-observations, given $t_i > t_{i-1}$, is equal to

$$A_i = \prod_{j=1}^{c} \left[\left(n_{ji} \frac{f_j(t_i)}{1 - F_j(t_{i-1})} \right)^{\xi_i^{(j)}} \left(\frac{1 - F_j(t_i)}{1 - F_j(t_{i-1})} \right)^{n_{ji} - \xi_i^{(j)}} \right]. \qquad (3.6)$$

Hence, from (3.5) and (3.6) we obtain

$$\prod_{i=1}^{r} A_i = \prod_{i=1}^{r} \prod_{j=1}^{c} \left[(n_{ji} f_j(t_i))^{\xi_i^{(j)}} (1 - F_j(t_i))^{u_i^{(j)}} \right] . \tag{3.7}$$

Following the failure t_i there remain $n_{ji} - \xi_i^{(j)}$ $x^{(j)}$-observations $(j = 1, 2, \ldots, c)$ which have neither failed with $x^{(j)} \leq t_i$ nor been censored with $x^{(j)'} < t_i$. Under assumptions (2.10), the probability that a particular $u_i^{(j)}$ of these $n_{ji} - \xi_i^{(j)}$ $x^{(j)}$-observations are assigned to the c_i experimental units censored in the interval (t_i, t_{i+1}) is given by

$$B_i = \binom{\sum_{j=1}^{c} n_{ji} - 1}{c_i}^{-1} \prod_{j=1}^{c} \binom{n_{ji} - \xi_i^{(j)}}{u_i^{(j)}} \tag{3.8}$$

$$i = 1, 2, \ldots, r,$$

and

$$B_o = \binom{N}{c_o}^{-1} \prod_{j=1}^{c} \binom{n_j}{u_o^{(j)}} . \tag{3.9}$$

Since the generalized hypergeometric probabilities (3.8) and (3.9) may be rewritten respectively as

$$B_i = \left[\left(\sum_{j=1}^{c} n_{ji} - 1\right)^{[c_i]}\right]^{-1} \binom{c_i}{u_i^{(1)}, u_i^{(2)}, \ldots, u_i^{(c-1)}}$$

$$\times \left[\prod_{j=1}^{c} (n_{ji} - \xi_i^{(j)})^{[u_i^{(j)}]}\right], \quad i = 1, 2, \ldots, r, \tag{3.10}$$

and

$$B_o = \left[N^{[c_o]}\right]^{-1} \binom{c_o}{u_o^{(1)}, u_o^{(2)}, \ldots, u_o^{(c-1)}} \prod_{j=1}^{c} \left[(n_{ji})^{[u_o^{(j)}]}\right], \tag{3.11}$$

where $a^{[b]} = a(a-1)\ldots(a-b+1)$, it is easily verified that

$$\prod_{i=0}^{r} B_i = \binom{N}{n_1, n_2, \ldots, n_{c-1}}^{-1} \prod_{i=0}^{r} \binom{c_i}{u_i^{(1)}, u_i^{(2)}, \ldots, u_i^{(c-1)}}$$

$$\times \prod_{i=1}^{r} \left[E_i \prod_{j=1}^{c} n_{ji}^{-\xi_i^{(j)}}\right], \tag{3.12}$$

with

$$E_i = \sum_{\alpha=i}^{r} (c_\alpha + 1).$$

Hence, (3. 1) follows from the product of (3. 7) and (3. 12).

Corollary 3. 1. 1. Under $H_o : F_1 \equiv F_2 \equiv \ldots \equiv F_c$, the order statistic and the rank-order statistic are independently distributed with density function

$$f(\underset{\sim}{t} \mid \underset{\sim}{C}_{(r, N)}, F_c) = \prod_{i=1}^{r} [E_i(1-F_c(t_i))^{c_i} f_c(t_i)] \qquad (3.13)$$

and probability function

$$p(\underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r, N)}) = \binom{N}{n_1, n_2, \ldots, n_{c-1}}^{-1} \prod_{i=0}^{r} \binom{c_i}{u_i^{(1)}, u_i^{(2)}, \ldots, u_i^{(c-1)}}. \qquad (3.14)$$

Note that, this corollary still holds even when Assumption C of (2.10) does not, i.e., when the observations are time censored.

The distribution of the rank-order statistic under alternatives can be determined from (3.1) by integrating over the order statistic $\underset{\sim}{t}$ $(0 = t_o < t_1 < t_2 < \ldots < t_r < t_{r+1} = \infty)$. In general, the integration of $f(\underset{\sim}{t}, \underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r, N)}, \underset{\sim}{F})$ over $\underset{\sim}{t}$ is rather difficult. Under Lehmann alternatives of the form $(1-F_j) = (1-F_c)^{\theta_j}$, $j = 1, 2, \cdots, c-1$, $p(\underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r, N)}, \underset{\sim}{\theta})$ can be determined as follows:

Corollary 3.1.2. Under Lehmann alternatives $(1-F_j) = (1-F_c)^{\theta_j}$, $j = 1, 2, \cdots, c-1$, the conditional distribution of the rank-order statistic is given by

$$p(\underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r, N)}, \underset{\sim}{\theta}) = K(N, \underline{u}, \underset{\sim}{C}_{(r, N)}) \left(\prod_{j=1}^{c} \theta_j^{\xi_{\cdot}^{(j)}} \right)$$

$$\times \left(\prod_{i=1}^{r} E_i \right) \prod_{i=1}^{r} \left[\sum_{j=i}^{r} \sum_{a=1}^{c} \theta_a (u_j^{(a)} + \xi_j^{(a)}) \right]^{-1}, \qquad (3.15)$$

where

$$\xi_{\cdot}^{(j)} = \sum_{i=1}^{r} \xi_i^{(j)},$$

and

$$K(N, \underline{u}, C_{(r, N)}) = \binom{N}{n_1, n_2, \ldots, n_{c-1}}^{-1} \prod_{i=0}^{r} \binom{c_i}{u_i^{(1)}, u_i^{(2)}, \ldots, u_i^{(c-1)}}. \qquad (3.16)$$

Proof. Under Lehmann alternatives of the form $(1-F_j) = (1-F_c)^{\theta_j}$, $j = 1, 2, \cdots, c-1$ (3.1) can be rewritten as

$$f(\underset{\sim}{t}, \underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r,N)}, \underset{\sim}{\theta}, F_c) = K(N, \underline{u}, \underset{\sim}{C}_{(r, N)}) \left(\prod_{j=1}^{c} \theta_j^{\xi_{\cdot}^{(j)}} \right) \left(\prod_{i=1}^{r} E_i \right)$$

$$\times \prod_{i=1}^{r} \left[f_c(t_i)(1-F_c(t_i))^{\sum_{j=1}^{c} \theta_j (u_i^{(j)} + \xi_i^{(j)}) - 1} \right] \qquad (3.17)$$

By using the transformation $w_i = 1 - F_e(t_i)$, the marginal probability function of the rank-order statistic $(\underline{\xi}, \underline{u})$ can be obtained by integrating over $\underset{\sim}{w} : 1 > w_1 > w_2 > \ldots > w_r > 0$. Thus,

$$p(\underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r, N)}, \underset{\sim}{\theta})$$

$$= K(N, \underline{u}, \underset{\sim}{C}_{(r,N)}) \left(\prod_{i=1}^{r} E_i \right) \left(\prod_{j=1}^{c} \theta_j^{\xi_{\cdot}^{(j)}} \right) \int \cdots \int_{\underset{\sim}{w}} \prod_{i=1}^{r} \left[w_i^{\sum_{j=1}^{c} \theta_j (u_i^{(j)} + \xi_i^{(j)}) - 1} \, dw_i \right].$$

$$= K(N, \underline{u}, \underset{\sim}{C}_{(r,N)}) \left(\prod_{i=1}^{r} E_i \right) \left(\prod_{j=1}^{c} \theta_j^{\xi_{\cdot}^{(j)}} \right) \prod_{i=1}^{r} \left[\sum_{j=i}^{r} \sum_{a=1}^{c} \theta_a (u_j^{(a)} + \xi_j^{(a)}) \right]^{-1}$$

IV. THE STATISTICS $\underset{\sim}{T}_{(r, N)}$ AND $Q_{(r, N)}$

Puri [20] had proposed a class of c-sample tests for the uncensored case by considering test statistics of the form

$$\mathcal{L} = \sum_{j=1}^{c} m_j[(T_{N,j} - \mu_{N,j})/A_N]^2 . \qquad (4.1)$$

Here m_j is the size of the $j^{\underline{th}}$ sample, $\mu_{N,j}$ and A_N are normalizing constants, and

$$T_{N,j} = \frac{1}{m_j} \sum_{i=1}^{N} E_{N,i}\xi_{N,i}^{(j)} . \qquad (4.2)$$

$T_{N,j}$, $j = 1, 2, \cdots, c$, are linear rank statistics and $E_{N,i}$, $i = 1, 2, \cdots, N$, are constants which are often evaluated so as to give good tests for particular alternatives. Under suitable regularity conditions and assumptions, the $\mathcal{L}$ statistics have been shown to have limiting chi-square distributions.

The $Q_{(r, N)}$ test statistic considered in this chapter may be written in the form (4.1). The linear rank statistics $T_{N,j}$ will be replaced by $T^{(j)}_{(r, N)}$, $j = 1, 2, \cdots, c$, which give the locally most powerful rank test for the two-sample case with arbitrarily censored data. This statistic has been introduced and discussed by Thomas [25].

The statistic $T'_{(r,N)} = (T^{(1)}_{(r,N)}, T^{(2)}_{(r,N)}, \cdots, T^{(c-1)}_{(r,N)})$, defined below, reduces to the statistic T given in Theorem 2 of Thomas [25]. Since $\underset{\sim}{T}_{(r,N)}$ will be shown in Chapter V to have a limiting normal distribution, the statistic $Q_{(r,N)}$ will be defined as

$$Q_{(r,N)} = \underset{\sim}{T}'_{(r,N)} \underline{\Sigma}^{-1}_{(r,N)} \underset{\sim}{T}_{(r,N)} \tag{4.3}$$

where $\underline{\Sigma}_{(r,N)}$ is the variance-covariance matrix of $T_{(r,N)}$.

The statistics, $T^{(j)}_{(r,N)}$, $j = 1, 2, \cdots, c-1$, are defined by

$$T^{(j)}_{(r,N)} = \frac{1}{\sqrt{N}} \left[\frac{\partial \ln p(\underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r,N)}, \underset{\sim}{\theta})}{\partial \theta_j} \right]_{\underset{\sim}{\theta} = \underset{\sim}{1}} \tag{4.4}$$

where $p(\underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r,N)}, \underset{\sim}{\theta})$ is the probability function of rank-order statistic given by (3.15) for Lehmann alternatives.

Lemma 4.1. The statistic $T'_{(r,N)} = (T^{(1)}_{(r,N)}, T^{(2)}_{(r,N)}, \ldots, T^{(c-1)}_{(r,N)})$ reduces to

$$\underset{\sim}{T}_{(r,N)} = \frac{1}{\sqrt{N}} \begin{bmatrix} \sum_{i=1}^{r} [(1-b_i)\xi_i^{(1)} - b_i u_i^{(1)}] \\ \sum_{i=1}^{r} [(1-b_i)\xi_i^{(2)} - b_i u_i^{(2)}] \\ \vdots \\ \sum_{i=1}^{r} [(1-b_i)\xi_i^{(c-1)} - b_i u_i^{(c-1)}] \end{bmatrix} , \tag{4.5}$$

where

$$b_i = \sum_{a=1}^{i} \frac{1}{\sum_{j=a}^{r} (c_j+1)} . \tag{4.6}$$

The $T^{(j)}_{(r,N)}$ in expression (4.5) can also be written as

$$T^{(j)}_{(r,N)} = \frac{1}{\sqrt{N}} \left[\xi^{(j)}_{\bullet} - \sum_{i=1}^{r} \frac{1}{E_i} V_i^{(j)} \right], \tag{4.7}$$

$$j = 1, 2, \ldots, c-1 ,$$

where

$$V_i^{(j)} = \sum_{a=i}^{r} (u_a^{(j)} + \xi_a^{(j)}) , \tag{4.8}$$

$$i = 1, 2, \cdots, r; \quad j = 1, 2, \cdots, c-1 .$$

For general Lehmann alternatives $(1-F_j) = H_{\theta_j}((1-F_c), c)$, $(h_{\theta_c}(x) \equiv x)$, a statistic $\underset{\sim}{T}_{(r,N)}$ could be derived similarly as in (4.4).

However, an explicit form of the distribution of the rank-order statistic $(\underline{u}, \underline{\xi})$ is not required. But some suitable regularity conditions must be imposed on the alternatives which allow orders of differentiation and integration to be interchanged.

The following regularity conditions on the functionals $h_{\theta_j}(w)$, $0 < w < 1$, would be sufficient

i) $h_{\theta_j}(w)$, $h'_{\theta_j}(w) \equiv \frac{\partial h_{\theta_j}(w)}{\partial w}$, $\frac{\partial h_{\theta_j}(w)}{\partial \theta_j}$ and $\frac{\partial h'_{\theta_j}(w)}{\partial \theta_j}$ are continuous with respect to θ_j in an ϵ-neighborhood of 1 for all $j = 1, 2, \ldots, c-1$.

ii) There exist integrable functions over $(0, 1)$, $M_1(w)$, $M_2(w)$, $M_3(w)$, and $M_4(w)$, such that

$$|h_{\theta_j}(w)| \leq M_1(w), \quad |h'_{\theta_j}(w)| \leq M_2(w)$$

$$\left| \frac{\partial \ln h_{\theta_j}(w)}{\partial \theta_j} \right| \leq M_3(w), \quad \left| \frac{\partial \ln h'_{\theta_j}(w)}{\partial \theta_j} \right| \leq M_4(w) \tag{4.9}$$

in an ϵ-neighborhood of 1 for all $j = 1, 2, \ldots, c-1$.

<u>Lemma 4.2.</u> Under the null hypothesis, the mean and variance of $\underset{\sim}{T}_{(r, N)}$ is given respectively by

$$E\,(\underset{\sim}{T}_{(r, N)}) = \underset{\sim}{0} \tag{4.10}$$

and

$$\underline{\Sigma}_{(r,N)} = [\sigma_{ij}(r,N)],$$

where

$$\sigma_{ij}(r,N) = \begin{cases} \dfrac{(r-b_r)n_i(N-n_i)}{N^2(N-1)}, & i = j \\[2ex] -\dfrac{(r-b_r)n_in_j}{N^2(N-1)}, & i \neq j \end{cases} \tag{4.11}$$

$$i, j = 1, 2, \ldots, c-1,$$

with b_r given by (4.6).

Proof. From (4.7), for $j = 1, 2, \ldots, c-1$, we have

$$T^{(j)}_{(r,N)} = \frac{1}{\sqrt{N}}\left[\xi^{(j)}_{\bullet} - \sum_{i=1}^{r} \frac{1}{E_i} V^{(j)}_i\right].$$

The expected value of $T^{(j)}_{(r,N)}$ under H_o is then

$$E(T^{(j)}_{(r,N)}) = \frac{1}{\sqrt{N}}\left[\frac{rn_j}{N} - \sum_{i=1}^{r}\frac{1}{E_i}\sum_{a=i}^{r}(c_a+1)\frac{n_j}{N}\right]$$

$$= \frac{1}{\sqrt{N}}\left[\frac{rn_j}{N} - \frac{n_j}{N}\sum_{i=1}^{r}\frac{E_i}{E_i}\right]$$

$$= 0 .$$

Hence,

$$E\,(\underset{\sim}{T}_{(r,\,N)}) = \underset{\sim}{0}\ .$$

Due to Thomas [25], $T^{(j)}_{(r,\,N)}$ can always be expressed as

$$T^{(j)}_{(r,\,N)} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} d_i v_i^{(j)} \equiv \frac{1}{\sqrt{N}} \sum_{i=1}^{r} [(1-b_i)\xi_i^{(j)} - b_i u_i^{(j)}]\ , \tag{4.12}$$

where $v_i^{(j)} = 1(0)$ if $i^{\underline{th}}$ observation is from the $j^{\underline{th}}$ population (if not). And the expected value of $T^{(j)^2}_{(r,\,N)}$ can easily be shown to reduce to

$$E\,(T^{(j)^2}_{(r,\,N)}) = \sum_{i=1}^{r} [(1-b_i)^2 + b_i^2 c_i]\, \frac{n_j(N-n_j)}{N^2(N-1)}\ .$$

By using (4.6), $\sum\limits^{r} [(1-b_i)^2 + b_i^2 c_i]$ can be reduced to $r - b_r$. Hence, the variance of $T^{(j)}_{(r,\,N)}$ is $\frac{(r-b_r)n_j(N-n_j)}{N^2(N-1)}$. Similarly,

$$E\,(T^{(j)}_{(r,\,N)} T^{(j')}_{(r,\,N)}) = -\frac{n_j n_j'}{N^2(N-1)}\,(r-b_r)\ . \tag{4.13}$$

<u>Lemma 4.3.</u> The statistic $Q_{(r,\,N)}$ given by (4.3), can be expressed in a computing form

$$Q_{(r,N)} = \frac{N(N-1)}{(r-b_r)} \sum_{j=1}^{c} \frac{1}{n_j} T_{(r,N)}^{(j)^2} , \tag{4.14}$$

where

$$T_{(r,N)}^{(c)} = - \sum_{j=1}^{c-1} T_{(r,N)}^{(j)} . \tag{4.15}$$

Proof. For $\underline{\Sigma}_{(r,N)}$ as given in (4.11), it is readily seen that

$$\underline{\Sigma}_{(r,N)}^{-1} = [\sigma^{ij}(r,N)] ,$$

where

$$\sigma^{ij}(r,N) = \begin{cases} \dfrac{N(N-1)}{(r-b_r)n_c}\left(1+\dfrac{n_c}{n_i}\right), & i = j \\[2ex] \dfrac{N(N-1)}{(r-b_r)n_c} , & i \neq j \end{cases} . \tag{4.16}$$

Then,

$$Q_{(r,N)} = \underset{\sim}{T}'_{(r,N)} \underline{\Sigma}_{(r,N)}^{-1} \underset{\sim}{T}_{(r,N)}$$

$$= \sum_{i=1}^{c-1} \sum_{j=1}^{c-1} T_{(r,N)}^{(i)} \sigma^{ij}(r,N) T_{(r,N)}^{(j)}$$

$$= \frac{N(N-1)}{(r-b_r)n_c} \left[\sum_{j=1}^{c-1} T_{(r,N)}^{(j)} \right]^2 + \frac{N(N-1)}{(r-b_r)} \sum_{j=1}^{c-1} \frac{1}{n_j} T_{(r,N)}^{(j)^2}$$

$$= \frac{N(N-1)}{(r-b_r)} \sum_{j=1}^{c} \frac{1}{n_j} T_{(r,N)}^{(j)^2} \quad .$$

It should be noted that $Q_{(r,N)}$ in (4. 14) has the same form as the Puri's $\mathcal{L}$-statistic (4. 1).

V. LIMITING DISTRIBUTIONS OF $\underset{\sim}{T}_{(r,N)}$ AND $Q_{(r,N)}$

Under the assumptions

$$0 < \lambda_o \leq \lambda_i \leq 1 - \lambda_o < 1$$

$$\lim_{N\to\infty} \frac{r}{N} = \nu > 0 , \tag{5.1}$$

where λ_o is a positive constant with

$$\lambda_i = \lim_{N\to\infty} \frac{n_i}{N} \quad ,$$

the asymptotic distributions of $\underset{\sim}{T}_{(r,N)}$ and $Q_{(r,N)}$ as defined by (4.5) and (4.3) will be derived under H_o and locally under H_a.

<u>Theorem 5.1.</u> Under $H_o:\theta_1=\theta_2=\ldots=\theta_{c-1}=1$, the statistic $\underset{\sim}{T}_{(r,N)}$ has a limiting (c-1)-variate normal distribution with mean 0 and variance-covariance $\underset{=}{\Sigma}$, where

$$\underset{=}{\Sigma} = [\sigma_{ij}] \ ,$$

with

$$\sigma_{ij} = \begin{cases} \nu\lambda_i(1-\lambda_i), & i = j \\ -\nu\lambda_i\lambda_j, & i \neq j \end{cases} \quad . \tag{5.2}$$

<u>Lemma 5.1.1.</u> For each $k \geq 1$ and $r \leq N$, define

$$b_r^{(k)} = \sum_{i=1}^{r} \frac{1}{E_i^k} , \tag{5.3}$$

where

$$E_i = \sum_{\alpha=i}^{r} (c_\alpha+1) ,$$

then

$$\lim_{N\to\infty} \frac{b_r^{(k)}}{N} = 0 . \tag{5.4}$$

Proof. For a given pattern $\underset{\sim}{C}_{(r,N)}$, $r \leq N$ with $\sum_{j=1}^{r} (c_j+1) = N$, it is seen that

$$b_r^{(k)} \leq b_r^{(1)} = b_r = \sum_{i=1}^{r} \frac{1}{\sum_{j=i}^{r} (c_j+1)} \leq \sum_{i=1}^{N} \frac{1}{i} . \tag{5.5}$$

Recall the Euler's constant, (see [16])

$$\gamma = \lim_{N\to\infty} [\sum_{i=1}^{N} \frac{1}{i} - \ln N]$$

$$= .5772156649. \tag{5.6}$$

It follows that

$$\lim_{N\to\infty} \frac{1}{N} \left[\sum_{i=1}^{N} \frac{1}{i} - \ln N\right] = 0 \tag{5.7}$$

Combining the results of (5. 5) and (5. 7), we get

$$\lim_{N\to\infty} \frac{b_r^{(k)}}{N} \leq \lim_{N\to\infty} \frac{b_r}{N}$$

$$\leq \lim_{N\to\infty} \frac{\ln N}{N}$$

$$= 0, \quad \text{for} \quad k \geq 1.$$

By applying (5. 1) and (5. 4) to (4. 11), we have

$$\lim_{N\to\infty} \Sigma_{(r,N)} = \underline{\Sigma} \tag{5.8}$$

Proof of Theorem 5.1. From (4. 12), the statistic $T_{(r,N)}^{(j)}$, $j = 1, 2, \cdots, c-1$ is given by

$$T_{(r,N)}^{(j)} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} d_i v_i^{(j)} \equiv \frac{1}{\sqrt{N}} \sum_{i=1}^{r} [(1-b_i) z_i^{(j)} - b_i u_i^{(j)}] .$$

Thomas has shown that $\{d_\alpha\}$ satisfies the condition

$$\frac{N^{-1} \sum_{\alpha=1}^{N} d_\alpha^k}{(N^{-1} \sum_{\alpha=1}^{N} d_\alpha^2)^{k/2}} = \frac{N^{-1} \sum_{i=1}^{r} (a_i^k + b_i^k e_i)}{[N^{-1} \sum_{i=1}^{r} (a_i^2 + b_i^2 e_i)]^{k/2}} = 0(1) \tag{5.9}$$

for $k = 3, 4, \cdots$

of the Wald-Wolfowitz-Noether theorem for permutation distributions (see page 236 of Fraser [9]). Theorem 5.1 then follows from (5.8), (5.9) and Fraser's [10] vector form of the Wald-Wolfowitz-Hoeffding theorem.

Theorem 5.2. Under the sequence of alternatives $\{\theta_j = 1 + \delta_j/\sqrt{N}\}$, where δ_j is a fixed constant for each $j = 1, 2, \ldots, c-1$, $\delta_c = 0$, the limiting distribution of $\underset{\sim}{T}_{(r,N)}$, given $\underset{\sim}{C}_{(r,N)}$, is $(c-1)$-variate normal with mean $\underline{\Sigma}\,\underset{\sim}{\delta}$ and variance-covariance $\underline{\Sigma}$.

Lemma 5.2.1. For each $j = 1, 2, \ldots, c-1$, $B_{N,j}$ converges in probability to σ_{jj}, where

$$B_{N,j} = \frac{1}{N} \sum_{i=1}^{r} \frac{V_i^{(j)}}{E_i} \left[1 - \frac{V_i^{(j)}}{E_i}\right] \tag{5.10}$$

with

$$V_i^{(j)} = \sum_{a=i}^{r} (u_a^{(j)} + \xi_a^{(j)}),$$

E_i given by (3.2) and $\sigma_{jj} = \nu\lambda_j(1-\lambda_j)$, the $j^{\underline{th}}$ diagonal element of $\underline{\Sigma}$.

Lemma 5. 2. 2. For each $j \neq j' = 1, 2, \ldots, c-1$, $B_{N, jj'}$ converges in probability to $\sigma_{jj'}$, where

$$B_{N, jj'} = -\frac{1}{N} \sum_{i=1}^{r} \frac{1}{E_i^2} V_i^{(j)} V_i^{(j')} \tag{5.11}$$

and $\sigma_{jj'} = -\nu\lambda_j\lambda_{j'}$ the $j, j'^{\underline{th}}$ element of $\underline{\Sigma}$.

The proofs of Lemma 5. 2. 1 and 5. 2. 2. will be presented in Appendix I.

Proof of Theorem 5. 2. Recall the probability density function of the rank-order statistic from (3. 15),

$$p(\underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r,N)}, \underset{\sim}{\theta}) = K(N, \underline{u}, \underset{\sim}{C}_{(r,N)}) \left[\prod_{j=1}^{c} \theta_j^{\xi_{\cdot}^{(j)}} \right] \left[\prod_{i=1}^{r} \frac{E_i}{E_i(\underset{\sim}{\theta})} \right] \tag{5.12}$$

where

$$E_i(\underset{\sim}{\theta}) = \sum_{j=i}^{r} \sum_{\alpha=1}^{c} \theta_\alpha (u_j^{(\alpha)} + \xi_j^{(\alpha)})$$

$$= \sum_{\alpha=1}^{c} \theta_\alpha V_i^{(\alpha)} . \tag{5.13}$$

Define the probability ratio $L(\underset{\sim}{\theta})$ as $p_{\underset{\sim}{\theta}} / p_{\underset{\sim}{1}}$, then

$$L(\underset{\sim}{\theta}) = \left[\prod_{j=1}^{c} \theta_j^{\xi_{\cdot}^{(j)}}\right] \left[\prod_{i=1}^{r} \frac{E_i}{E_i(\underset{\sim}{\theta})}\right] \tag{5.14}$$

Replacing θ_j by $\theta_{j,N} = \delta_j/\sqrt{N}$, $j = 1, 2, \cdots, c-1$, and using Taylor's expansion of $\ln(1+x)$, we get

$$\ln L(\underset{\sim}{\theta}_N) = \sum_{j=1}^{c-1} \xi_{\cdot}^{(j)} \left[\frac{\delta_j}{\sqrt{N}} - \frac{\delta_j^2}{2N}\right] - \sum_{i=1}^{r}\left[\sum_{\alpha=1}^{c-1} \frac{\delta_\alpha V_i^{(\alpha)}}{E_i\sqrt{N}} - \frac{1}{2E_i^2 N}\left\{\sum_{\alpha=1}^{c-1} \delta_\alpha V_i^{(\alpha)}\right\}^2\right] + R_N \tag{5.15}$$

where R_N is the remainder term which will be shown to be $o(1)$ in Appendix II. Then,

$$\ln L(\underset{\sim}{\theta}_N) = \sum_{j=1}^{c-1} \delta_j T_{(r,N)}^{(j)}$$

$$- \frac{1}{2\sqrt{N}} \sum_{j=1}^{c-1} \delta_j^2 \left[T_{(r,N)}^{(j)} + \frac{1}{\sqrt{N}} \sum_{i=1}^{r} \frac{V_i^{(j)}}{E_i} \left\{1 - \frac{V_i^{(j)}}{E_i}\right\}\right]$$

$$+ \frac{1}{2N} \sum\sum_{j \neq j'=1}^{c-1} \left[\sum_{i=1}^{r} \frac{\delta_i \delta_j V_i^{(j)} V_i^{(j')}}{E_i^2}\right] + 0(1) .$$

Substituting $B_{N,j}$ and $B_{N,jj'}$ from Lemma 5. 2. 1 and 5. 2. 2, we have

$$\ln L(\underset{\sim}{\theta}_N) = \sum_{j=1}^{c-1} \delta_j T_{(r,N)}^{(j)} - \frac{1}{2\sqrt{N}} \sum_{j=1}^{c-1} \delta_j^2 [T_{(r,N)}^{(j)} + \sqrt{N} B_{N,j}] - \frac{1}{2} \sum\sum_{j \neq j'=1}^{c-1} \delta_j \delta_{j'} B_{N,jj'} + 0(1) \tag{5.16}$$

The moment generating function of $\underset{\sim}{T}_{(r,N)}$ and $\underset{\sim}{B}_{N,j}$, $\underset{\sim}{B}_{N,jj'}$ can be obtained by taking expected value of

$$\exp[\underset{\sim}{\tau}'\underset{\sim}{T}_{(r,N)} + \underset{\sim}{a}'\underset{\sim}{B}_N + \underset{\sim}{\gamma}'\underset{\sim}{\tilde{B}}_N]$$

where $\underset{\sim}{\tau}$, $\underset{\sim}{a}$, $\underset{\sim}{\gamma}$ are vectors of arguments of dimension (c-1), (c-1) and (c-1)(c-2) respectively, and $\underset{\sim}{B}_N$, $\underset{\sim}{\tilde{B}}_N$ are vectors of (c-1) and (c-1)(c-2) dimension having $B_{N,j}$ and $B_{N,jj'}$ as their elements.

$$M_{\underset{\sim}{T}_{(r,N)}, \underset{\sim}{B}_N, \underset{\sim}{\tilde{B}}_N; \underset{\sim}{\delta}}(\underset{\sim}{\tau}, \underset{\sim}{a}, \underset{\sim}{\gamma})$$

$$= E_{\underset{\sim}{\delta}}[\exp(\underset{\sim}{\tau}'\underset{\sim}{T}_{(r,N)} + \underset{\sim}{a}'\underset{\sim}{B}_N + \underset{\sim}{\gamma}'\underset{\sim}{\tilde{B}}_N)]$$

$$= \sum_{\underline{\xi}} \sum_{\underline{u}} K(N, \underline{u}, \underset{\sim}{C}_{(r,N)})$$

$$\times \exp\left[\sum_{j=1}^{c-1}(\delta_j - \frac{1}{2\sqrt{N}}\delta_j^2 + \tau_j)T_{(r,N)}^{(j)} - \sum_{j=1}^{c-1}(\frac{\delta_j^2}{2} - a_j)B_{N,j} - \sum_{j\neq j'=1}^{c-1}\sum(\frac{\delta_j\delta_{j'}}{2} - \gamma_{jj'})B_{N,jj'}\right] + 0(1) \qquad (5.17)$$

where

$$K(N, \underline{u}, \underset{\sim}{C}_{(r,N)}) = p(\underline{\xi}, \underline{u} \mid \underset{\sim}{C}_{(r,N)}, \underset{\sim}{\theta} = \underset{\sim}{1})$$

as defined in (3. 16). It is then seen that

$$M_{\underset{\sim}{T}_{(r,N)}, \underset{\sim}{B}_N, \underset{\sim}{\tilde{B}}_N; \underset{\sim}{\delta}}(\underset{\sim}{\tau}, \underset{\sim}{a}, \underset{\sim}{\gamma}) = M_{\underset{\sim}{T}_{(r,N)}, \underset{\sim}{B}_N, \underset{\sim}{\tilde{B}}_N; \underset{\sim}{0}}(\underset{\sim}{\tilde{\tau}}, \underset{\sim}{\tilde{a}}, \underset{\sim}{\tilde{\gamma}})\, e^{0(1)}, \qquad (5.18)$$

where

$$\tilde{T}_j = \delta_j - \frac{1}{2\sqrt{N}}\delta_j^2 + T_j$$

$$\tilde{\alpha}_j = \alpha_j - \frac{1}{2}\delta_j^2 \qquad (5.19)$$

$$\tilde{\gamma}_{jj'} = \gamma_{jj'} - \delta_j\delta_{j'} \quad .$$

Then,

$$\lim_{N\to\infty} M_{\underset{\sim}{T}(r,N),\, \underset{\sim}{B}_N,\, \underset{\sim}{\tilde{B}}_N;\, \underset{\sim}{\delta}}(\underset{\sim}{T}, \underset{\sim}{\alpha}, \underset{\sim}{\gamma}) = M_{\underset{\sim}{T},\, \underset{\sim}{B},\, \underset{\sim}{\tilde{B}};\, \underset{\sim}{\delta}}(\underset{\sim}{T}, \underset{\sim}{\alpha}, \underset{\sim}{\gamma})$$

$$= M_{\underset{\sim}{T},\, \underset{\sim}{B},\, \underset{\sim}{\tilde{B}};\, \underset{\sim}{\delta}}(\underset{\sim}{\delta}+\underset{\sim}{T}, \underset{\sim}{\tilde{\alpha}}, \underset{\sim}{\tilde{\gamma}}) \ . \qquad (5.20)$$

But,

$$M_{\underset{\sim}{T};\, \underset{\sim}{\delta}}(\underset{\sim}{T}) = M_{\underset{\sim}{T},\, \underset{\sim}{B},\, \underset{\sim}{\tilde{B}};\, \underset{\sim}{\delta}}(\underset{\sim}{T}, \underset{\sim}{0}, \underset{\sim}{0})$$

$$= M_{\underset{\sim}{T},\, \underset{\sim}{B},\, \underset{\sim}{\tilde{B}};\, \underset{\sim}{0}}(\underset{\sim}{\tilde{T}}, \underset{\sim}{\tilde{\alpha}}, \underset{\sim}{\tilde{\gamma}})\Big|_{\substack{\underset{\sim}{\alpha}=0 \\ \underset{\sim}{\gamma}=0}} \ .$$

Applying the results of Theorem 5.1, Lemma 5.2.1 and Lemma 5.2.2, we get

$$M_{\underset{\sim}{T};\, \underset{\sim}{\delta}}(\underset{\sim}{T}) = \exp[\tfrac{1}{2}(\underset{\sim}{\delta}+\underset{\sim}{T})' \ \underline{\Sigma}\ (\underset{\sim}{\delta}+\underset{\sim}{T}) - \tfrac{1}{2}\underset{\sim}{\delta}'\ \underline{\Sigma}\ \underset{\sim}{\delta}]$$

$$= \exp[\underset{\sim}{\delta}'\ \underline{\Sigma}\ \underset{\sim}{T} + \tfrac{1}{2}\underset{\sim}{T}'\ \underline{\Sigma}\ \underset{\sim}{T}] \qquad (5.21)$$

which is the moment generating function of (c-1)-variate normal distribution with mean $\underline{\Sigma}\ \underset{\sim}{\delta}$ and variance-covariance matrix $\underline{\Sigma}$.

Lemma 5.2.3. The limiting distribution of $Q_{(r,N)}$ under the sequence of alternatives $\{\theta_{j,N} = 1 + \delta_j/\sqrt{N}\}$ is non-central chi-square with c-1 degree of freedom and non-centrality parameter $\underset{\sim}{\delta}'\underline{\Sigma}\,\underset{\sim}{\delta}$.

Proof. Since $Q_{(r,N)}$ is defined as

$$Q_{(r,N)} = \underset{\sim}{T}'_{(r,N)}\,\underline{\Sigma}^{-1}_{(r,N)}\,\underset{\sim}{T}_{(r,N)} \tag{5.22}$$

and $\underline{\Sigma}^{-1}_{(r,N)}$ has $\underline{\Sigma}^{-1}$ as its limit, then by the result of Theorem 5.2, we have $Q_{(r,N)}$ converging in distribution to a non-central chi-square variable with (c-1) degree of freedom. The non-centrality parameter then will be

$$(\underline{\Sigma}\,\underset{\sim}{\delta})'\underline{\Sigma}^{-1}(\underline{\Sigma}\,\underset{\sim}{\delta}) = \underset{\sim}{\delta}'\underline{\Sigma}\,\underset{\sim}{\delta}\,. \tag{5.23}$$

The non-centrality parameter (5.23) may also be expressed as

$$\underset{\sim}{\delta}'\underline{\Sigma}\,\underset{\sim}{\delta} = \sum_{i=1}^{c-1}\sum_{j=1}^{c-1} \delta_i \sigma_{ij} \delta_j\,,$$

which by use of (5.2) reduces to

$$\nu \sum_{i=1}^{c-1} \lambda_i(1-\lambda_i)\delta_i^2 - \nu \sum\sum_{i\neq j=1}^{c-1} \lambda_i\lambda_j\delta_i\delta_j$$

$$= \nu \sum_{i=1}^{c-1} \lambda_i \delta_i^2 - \nu \left[\sum_{i=1}^{c-1} \lambda_i \delta_i \right]^2$$

$$= \nu \sum_{i=1}^{c} \lambda_i (\delta_i - \overline{\delta})^2 , \qquad (5.24)$$

where

$$\delta_c = 0, \quad \overline{\delta} = \sum_{i=1}^{c} \lambda_i \delta_i .$$

VI. ASYMPTOTIC EFFICIENCY OF $Q_{(r, N)}$

The statistic $Q_{(r, N)}$ will be shown to be asymptotic efficient relative to a parametric likelihood ratio test. Exponential failure time distributions and an exponential censoring distribution will be assumed. The experimental units used during experimentation are randomly censored according to an exponential distribution. That is, a failure arises if $x_i^{(j)} \leq y_i^{(j)}$ and a censoring arises if $x_i^{(j)} > y_i^{(j)}$,

$$f_j(x_i^{(j)}) = \theta_j e^{-\theta_j x_i^{(j)}}$$

$$h(y) = \alpha e^{-\alpha y}, \qquad (6.1)$$

where $x_i^{(j)}$ and $y_i^{(j)}$ are the $i^{\underline{th}}$ observation of failure time and censoring time from the $j^{\underline{th}}$ population respectively, $i = 1,2,\ldots, n_j$ and $j = 1, 2, \ldots, c$. Let

$\epsilon_i^{(j)} = 1$ if the $i^{\underline{th}}$ observation of the $j^{\underline{th}}$ population is a failure

$= 0$ if it is a censored observation. (6.2)

Lemma 6.1.

$$m_j/n_j \xrightarrow{p} \frac{\theta}{\theta+\alpha} \qquad (6.3)$$

and

$$r/N \xrightarrow{p} \frac{\theta}{\theta+\alpha} \qquad (6.4)$$

where θ is a parameter in the sequence of alternatives $\{\theta_{jN} = \theta(1+\delta_j/\sqrt{N})\}$, m_j is the number of failure observations from the $j^{\underline{th}}$ population, i.e.,

$$m_j = \sum_{i=1}^{n_j} \epsilon_i^{(j)} \quad . \tag{6.5}$$

Proof. Due to Efron [7], the $\epsilon_i^{(j)}$'s are statistically independent for each $j = 1, 2, \ldots, c$, with

$$P(\epsilon_i^{(j)} = 1) = \frac{\theta_{jN}}{\theta_{jN}+a}$$

$$P(\epsilon_i^{(j)} = 0) = \frac{a}{\theta_{jN}+a} \quad . \tag{6.6}$$

The mean of independent Bernoulli trials,

$$\frac{1}{n_j} \sum_{i=1}^{n_j} \epsilon_i^{(j)} = m_j/n_j \ ,$$

then has expected value $\frac{\theta_{jN}}{\theta_{jN}+a}$ and variance $\frac{1}{n_j} \frac{a\theta_{jN}}{(\theta_{jN}+a)^2}$. Since the variance of m_j/n_j is $\frac{1}{n_j} \cdot \frac{a\theta_{jN}}{(\theta_{jN}+a)^2}$, which tends to zero when $N \to \infty$, then

$$m_j/n_j \xrightarrow{p} \lim_{n\to\infty} \frac{\theta_{jN}}{\theta_{jN}+a}$$
$$= \frac{\theta}{\theta+a} \quad . \qquad (6.8)$$

Moreover, we have immediately from (6.8) that

$$r/N = \sum_{j=1}^{c} m_j/N \xrightarrow{p} \sum_{j=1}^{c} \frac{n_j}{N} \frac{\theta}{\theta+a}$$
$$= \frac{\theta}{\theta+a} \quad . \qquad (6.9)$$

Lemma 6.2. If $X^{(j)}$ is distributed as the exponential with parameter $\theta_{jN} = \theta(1+\delta_j/\sqrt{N})$, $j = 1, 2, \ldots, c$, then under (6.1) $R = -2\ln$ L. R. (L. R. = likelihood ratio test statistic) has a limiting chi-square with $c-1$ degree of freedom and non-centrality parameter

$$\lambda(R, \delta, \theta) \approx \frac{\theta}{\theta+a} \sum_{j=1}^{c} \lambda_j(\delta_j-\bar{\delta})^2 \ , \qquad (6.10)$$

where

$$\lambda_j = \lim_{n\to\infty} n_j/N$$

and

$$\bar{\delta} = \sum_{j=1}^{c} \lambda_i \delta_j \ .$$

Proof. From (6. 1), we have

$$1 - H(y) = e^{-\alpha y}$$

$$1 - F_j(x) = e^{-\theta_j x} \qquad (6.11)$$

The likelihood function can be written as

$$f(\underline{x}, \underline{y} | \underline{\epsilon}, \underset{\sim}{\theta}, \alpha)$$

$$= K \prod_{j=1}^{c} \prod_{i=1}^{n_j} \left[f_j(x_i^{(j)})(1-H(x_i)) \right]^{\epsilon_i^{(j)}} \left[h(y_i^{(j)})(1-F_j(y_i^{(j)})) \right]^{1-\epsilon_i^{(j)}}$$

$$= K \prod_{j=1}^{c} \prod_{i=1}^{n_j} \left[\theta_j e^{-x_i^{(j)}(\theta_j+\alpha)} \right]^{\epsilon_i^{(j)}} \left[\alpha e^{-y_i^{(j)}(\theta_j+\alpha)} \right]^{1-\epsilon_i^{(j)}} . \qquad (6.12)$$

Thus,

$$\ln f = \ln K + \sum_{j=1}^{c} \sum_{i=1}^{n_j} [\epsilon_i^{(j)}(\ln \theta_j - x_i^{(j)}(\theta_j+\alpha)) + (1-\epsilon_i^{(j)})(\ln \alpha - y_i^{(j)}(\theta_j+\alpha))],$$

and

$$\frac{\partial \ln f}{\partial \theta_j} = \frac{m_j}{\theta_j} - \sum_{i=1}^{n_j} [\epsilon_i^{(j)} x_i^{(j)} + (1-\epsilon_i^{(j)}) y_i^{(j)}].$$

Consequently, the unrestricted maximum likelihood estimators for θ_j are given by

$$\hat{\theta}_j = \frac{m_j}{\sum_{i=1}^{n_j} [\epsilon_i^{(j)} x_i^{(j)} + (1-\epsilon_i^{(j)}) y_i^{(j)}]} , \quad j = 1, \ldots, c . \qquad (6.13)$$

And under $H_o : \theta_1 \equiv \theta_2 \equiv \cdots \equiv \theta_c \equiv \theta$, the maximum likelihood estimator is

$$\hat{\theta} = \frac{r}{\sum_{j=1}^{c} \sum_{i=1}^{n_j} [\epsilon_i^{(j)} x_i^{(j)} + (1-\epsilon_i^{(j)}) y_i^{(j)}]} \quad . \tag{6.14}$$

Substituting $\hat{\theta}_j$ and $\hat{\theta}$ from (6. 13) and (6. 14) in (6. 12), we obtain

$$R = -2 \ln L.R. = -2 \sum_{j=1}^{c} m_j [\ln \hat{\theta} - \ln \hat{\theta}_j] \quad .$$

Since, this is the well-known (see Kendall and Stuart [17]) likelihood ratio statistic with $\hat{\theta}_j$ and $\hat{\theta}$ are the reciprocal of the averages of independent observations, then

$$R \approx \sum_{j=1}^{c} m_j (\theta_j - \bar{\theta})^2 / \bar{\theta}^2 \tag{6.15}$$

has limiting central chi-square distribution under H_o. Under H_a, R will have limiting non-central chi-square with non-centrality parameter conditional on $\underset{\sim}{m}$,

$$\lambda(R, \underset{\sim}{\delta}, \theta \,|\, \underset{\sim}{m}) = \sum_{j=1}^{c} m_j (\theta_j - \bar{\theta})^2 / \bar{\theta}^2 \quad , \tag{6.16}$$

where

$$\theta_j = \theta(1 + \delta_j / \sqrt{N})$$

and

$$\bar{\theta} = \frac{1}{r} \sum_{j=1}^{c} m_j \theta_j$$

$$= \theta(1+\bar{\delta}/\sqrt{N}) \ . \tag{6.17}$$

By using the result of Lemma 6.1, the unconditional non-centrality parameter of R is

$$\lambda(R, \underset{\sim}{\delta}, \theta) = \sum_{j=1}^{c} \frac{\theta n_j}{\theta+\alpha} \cdot \frac{(\delta_j - \bar{\delta})^2}{N(1+\bar{\delta}/\sqrt{N})^2}$$

$$\approx \frac{\theta}{\theta+\alpha} \sum_{j=1}^{c} \lambda_j (\delta_j - \bar{\delta})^2 \ . \tag{6.18}$$

<u>Lemma 6.3.</u> The unconditional non-centrality parameter of $Q_{(r,N)}$ is approximately equal to

$$\frac{\theta}{\theta+\alpha} \sum_{j=1}^{c} \lambda_j (\delta_j - \bar{\delta})^2 \ . \tag{6.19}$$

<u>Proof.</u> The conditional non-centrality parameter of $Q_{(r,N)}$ as given in (5.35) is

$$\lambda(Q, \underset{\sim}{\delta} \mid r) \approx \frac{r}{N} \sum_{i=1}^{c} \lambda_i (\delta_i - \bar{\delta})^2 \ .$$

From (6.9), it follows that the non-centrality parameter of the unconditional limiting distribution of $Q_{(r, N)}$ is equal to

$$\frac{\theta}{\theta+\alpha} \sum_{i=1}^{c} \lambda_i (\delta_i - \bar{\delta})^2 . \tag{6.20}$$

<u>Theorem 6.1.</u> Under the exponential alternatives of failure observations and exponential censoring time, the asymptotic relative efficiency of $Q_{(r, N)}$ with respect to R is 1.

This result follows from the fact that the non-centrality parameters of $Q_{(r, N)}$ in (6.20) and that of R in (6.18) are equal.

Comments on a time censored model: For uniform random arrivals over the internal $(0, T)$ with all units on test at time T being censored, we have not been able to evaluate, for arbitrary T, the A. R. E. of the $Q_{(r, N)}$-test with respect to the likelihood ratio test for the case of exponential failure distributions. However, for the limiting case $T \to \infty$, the A. R. E. is equal to 1 and $T \to 0$ is equal to 8/9. The A. R. E. for $T \to \infty$ is well-known (e. g., see Puri [20]). For $T \to 0$ the Q statistic and the corresponding c-sample statistic using Gehan's $W^{(j)}$ statistics in place of the $T^{(j)}$ statistics (see Gehan [11]), Q', can be shown to be asymtotically equivalent. Since the A. R. E. of Q' with respect to the likelihood ratio test is independent of c, the evaluation, by Gehan, of the A. R. E. of Q' with respect to the likelihood ratio test for $c = 2$ may be used.

VII. NUMERICAL RESULTS AND ILLUSTRATIVE EXAMPLE

In the first part of this chapter, comparisons of exact and chi-square approximation upper-tail probabilities of $Q_{(r,N)}$ are made for three different censoring patterns. For each pattern, the upper-tail probabilities corresponding to six different values of $\underset{\sim}{\theta}$ are tabulated below in Tables 2, 3, and 4. $\hat{P}(1, 1)$ denotes the chi-square approximation probability under H_o, i.e., $\theta_1 = 1$ and $\theta_2 = 1$, and $P(\theta_1, \theta_2)$ denotes the exact probabilities for six pairs of values for (θ_1, θ_2) with $\theta_3 = 1$. $P(1, 1)$ and $P(\theta_1, \theta_2)$ are computed according to (3. 14) and (3. 15) respectively.

Due to the large number of points in the sample space of the rank-order statistic, it was necessary to choose small sample sizes and small numbers of failures. For sample sizes $n_1 = n_2 = n_3 = 4$ and $r = 3, 4, 5$ in censoring patterns given in Table 1, there were respectively 648, 1230, and 1890 points in the sample spaces of the rank-order statistics. Consequently, the $Q_{(r, N)}$-test has small power. In pattern I, the largest power for test size $\alpha = 0.10$ occurs when $\underset{\sim}{\theta} = (1, 3)$, which is about 0. 20. The largest powers in patterns II and III for the corresponding test size are slightly larger, which are about 0. 23 and 0. 25 respectively. It is interesting to note that in all three patterns, the maximum powers occur at the same $\underset{\sim}{\theta} = (1, 3)$. Figures 1 and 2 show the relations of power and size α of the test for various values of $\underset{\sim}{\theta}$.

The corresponding generalized Gehan's statistics, $Q'_{(r, N)}$ (see Appendix III), are also computed and ordered. The upper-tail probabilities of $Q'_{(r, N)}$ can be obtained by cumulating the probabilities of the corresponding rank-order statistics. Power comparisons of $Q_{(r, N)}$ and $Q'_{(r, N)}$ are presented graphically in Figures 3 and 4. In pattern II, $Q_{(r, N)}$ and $Q'_{(r, N)}$ have nearly the same power. In pattern III, the powers of $Q_{(r, N)}$ are slightly larger than those of $Q'_{(r, N)}$ for the case of $\theta = (2, 2)$ and substantially larger for the case of $\underset{\sim}{\theta} = (1, 3)$.

Table 1. Arbitrary censoring patterns I, II and III.
$N = 12,\ n_1 = n_2 = n_3 = 4$

Pattern I ($r = 3$)		Pattern II ($r = 4$)		Pattern III ($r = 5$)	
i	c_i	i	c_i	i	c_i
1	1	1	1	1	1
2	3	2	0	2	0
3	5	3	2	3	1
		4	5	4	0
				5	5

Table 2. Exact and chi-square approximation cumulative probabilities of $Q_{(r,N)}$ for arbitrary censoring pattern I.

$Q_{(r,N)}$	$\hat{P}(1,1)$	$P(1,1)$	$P(1,2)$	$P(1,3)$	$P(2,2)$	$P(2,3)$	$P(3,3)$
8.3642	.01530	.00087	.00146	.00262	.00117	.00160	.00150
7.7208	.02107	.00866	.01433	.02522	.01152	.01550	.01442
7.5362	.02305	.02424	.03898	.06670	.03169	.04194	.03911
6.7407	.03439	.02511	.04024	.06860	.03280	.04334	.04050
6.4409	.03996	.05455	.08246	.13248	.06936	.08864	.08413
4.2709	.11765	.05714	.08560	.13646	.07237	.09208	.08765
3.9734	.13807	.05974	.08862	.14009	.07536	.09542	.09114
3.5353	.17033	.09177	.12552	.18414	.11106	.13466	.13102
3.3277	.19014	.12121	.15804	.22042	.14318	.16907	.16627
3.1640	.20598	.12727	.16450	.22702	.19466	.22330	.17417
2.8526	.24171	.20519	.24806	.31647	.23328	.26371	.26366
2.6935	.25924	.28139	.32736	.39744	.31370	.34664	.34938

Table 3. Exact and chi-square approximation cumulative probabilities of $Q_{(r,N)}$ and $Q'_{(r,N)}$ for arbitrary censoring pattern II.

$Q_{(r,N)}$	$\hat{P}(1,1)$	P(1,1)	P(1,2)	P(1,3)	P(2,2)	P(2,3)	P(3,3)	$Q'_{(r,N)}$	P(1,1)	P(1,2)	P(3,3)
10.5322	.00519	.00087	.00204	.00462	.00140	.00226	.00197	4.3874	.00087	.00204	.00197
10.3365	.00570	.00606	.01424	.03232	.00954	.01538	.01304	9.2855	.00606	.01424	.01304
7.1583	.02788	.00693	.01576	.03516	.01076	.01710	.01466	7.6050	.00693	.01576	.01466
6.8561	.03288	.01039	.03181	.04650	.01542	.02356	.02048	7.4012	.01212	.02483	.02322
6.6403	.03615	.01212	.02483	.05216	.01769	.02660	.02322	6.5864	.01386	.02751	.02636
6.2871	.04328	.01472	.02887	.05866	.02126	.03137	.02795	6.2299	.02082	.03809	.02706
6.0407	.04979	.01905	.03552	.06922	.02696	.03873	.03513	6.1281	.02514	.04457	.04346
5.8546	.05340	.02424	.04351	.08211	.03365	.04735	.04332	5.8225	.03726	.06121	.06076
5.6021	.06081	.03463	.05865	.10483	.04708	.06419	.06010	5.6528	.05196	.08139	.08032
5.4409	.06588	.03810	.06390	.11334	.05132	.06958	.06497	5.5170	.05892	.09145	.08846
5.3517	.06925	.04156	.06870	.12006	.05577	.07501	.07056	5.3642	.06066	.09411	.09106
5.1934	.07502	.05628	.08892	.14804	.07449	.09771	.09479	4.9568	.06324	.09779	.09568
4.4833	.08291	.06494	.10059	.16397	.08522	.11056	.10677	4.6682	.07594	.11550	.11702
4.9056	.08629	.07619	.11562	.18430	.09908	.12706	.12342	4.3966	.08770	.13266	.13660
4.7099	.09537	.09697	.14317	.22150	.12403	.15650	.15243	4.1929	.09814	.14616	.15146
4.5919	.10127	.10476	.15326	.23464	.13350	.16756	.16371	4.0232	.10852	.15930	.16750
4.3933	.11080	.12035	.17266	.25858	.15233	.18914	.18622	3.8144	.11284	.16504	.17438
4.2572	.11884	.12900	.18348	.27220	.16263	.20099	.19835	3.5309	.12586	.18160	.19352
4.1669	.12400	.13939	.19590	.28669	.17514	.21506	.21344	3.2083	.13740	.19461	.20926
4.0759	.13100	.14199	.19890	.28995	.17844	.21879	.21777	2.8688	.14088	.19887	.21396

Table 4. Exact and chi-square approximation cumulative probabilities of $Q_{(r,N)}$ and $Q'_{(r,N)}$ for arbitrary censoring pattern III.

$Q_{(r,N)}$	$\hat{P}(1,1)$	P(1,1)	P(1,2)	P(1,3)	P(2,2)	P(2,3)	P(3,3)	$Q'_{(r,N)}$	P(1,1)	P(2,2)	P(1,3)
9.5017	.00865	.00087	.00209	.00480	.00154	.00255	.00241	9.3771	.00087	.00154	.00236
8.7884	.01240	.00346	.00808	.01820	.00584	.00941	.00866	9.0766	.00260	.00432	.00472
8.1350	.01725	.00952	.02165	.04738	.01530	.02420	.02159	8.3552	.00692	.01108	.01148
7.5236	.02328	.01558	.03373	.07042	.02447	.03765	.03404	7.9946	.00866	.01376	.01870
7.1778	.02901	.01818	.03850	.07876	.02845	.04324	.03964	7.6339	.01214	.01880	.03312
6.7082	.03508	.02684	.05380	.10498	.04134	.06112	.05745	7.0328	.01556	.02428	.04476
6.4758	.03916	.02857	.05654	.10921	.04416	.06495	.06191	6.4918	.02162	.03246	.06418
6.2853	.04505	.03463	.06680	.12650	.05283	.07673	.07355	6.0110	.02762	.04162	.08254
5.9684	.05079	.04069	.07542	.13780	.06260	.08924	.08924	5.7400	.03368	.05144	.09898
5.7698	.05558	.04848	.08685	.15387	.07429	.10471	.10620	5.4700	.04316	.06120	.11864
5.5699	.06142	.05455	.09465	.16252	.08395	.11594	.12157	5.2446	.05078	.07448	.13360
5.5039	.06393	.05887	.10059	.17019	.09053	.12416	.13151	5.0492	.06206	.08370	.16296
5.2775	.07136	.06667	.11187	.18655	.10164	.13822	.14707	4.8689	.06890	.09957	.17732
5.1626	.07577	.07013	.11751	.19647	.10602	.14401	.15220	4.7036	.07316	.10597	.18518
4.7603	.09255	.08052	.13128	.21391	.12027	.16124	.17164	4.5683	.08090	.11725	.19680
4.5117	.10435	.09437	.14925	.23659	.13856	.18283	.19556	4.3730	.09206	.13293	.21848
4.3543	.11304	.11169	.17147	.26463	.16089	.20907	.22420	4.1626	.10250	.14627	.23426
4.2485	.12003	.12208	.18410	.27912	.17452	.22470	.24218	3.8921	.12338	.17347	.26194
4.1482	.12619	.13420	.19838	.29461	.19051	.24284	.26362	3.6670	.13904	.19463	.28026
3.9368	.13946	.14805	.21613	.31734	.20709	.26197	.28280	3.1257	.15296	.21211	.29958

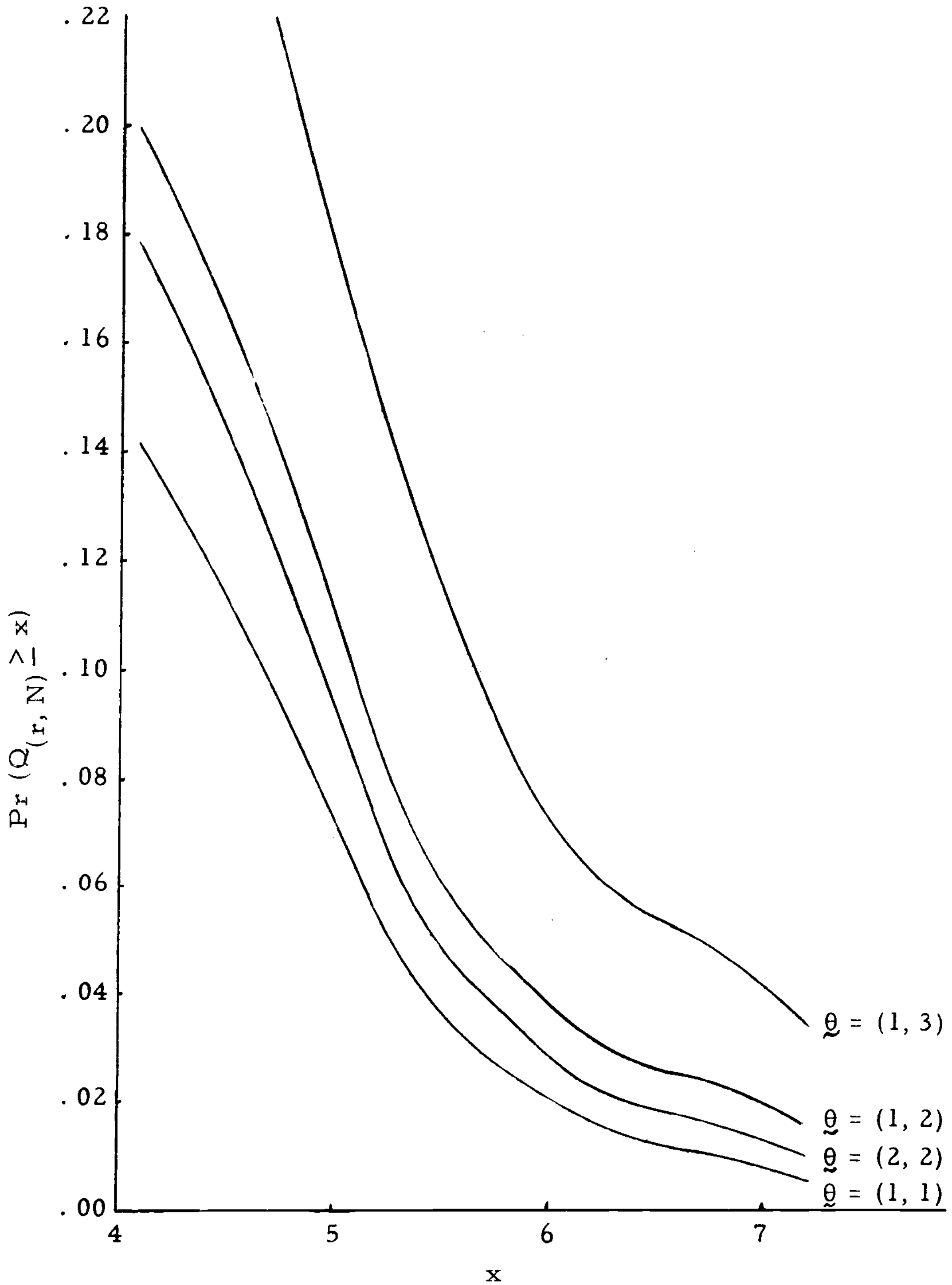

Figure 1. Upper-tailed probabilities of $Q_{(r,\ N)}$ for different values of $\underset{\sim}{\theta}$ (pattern II).

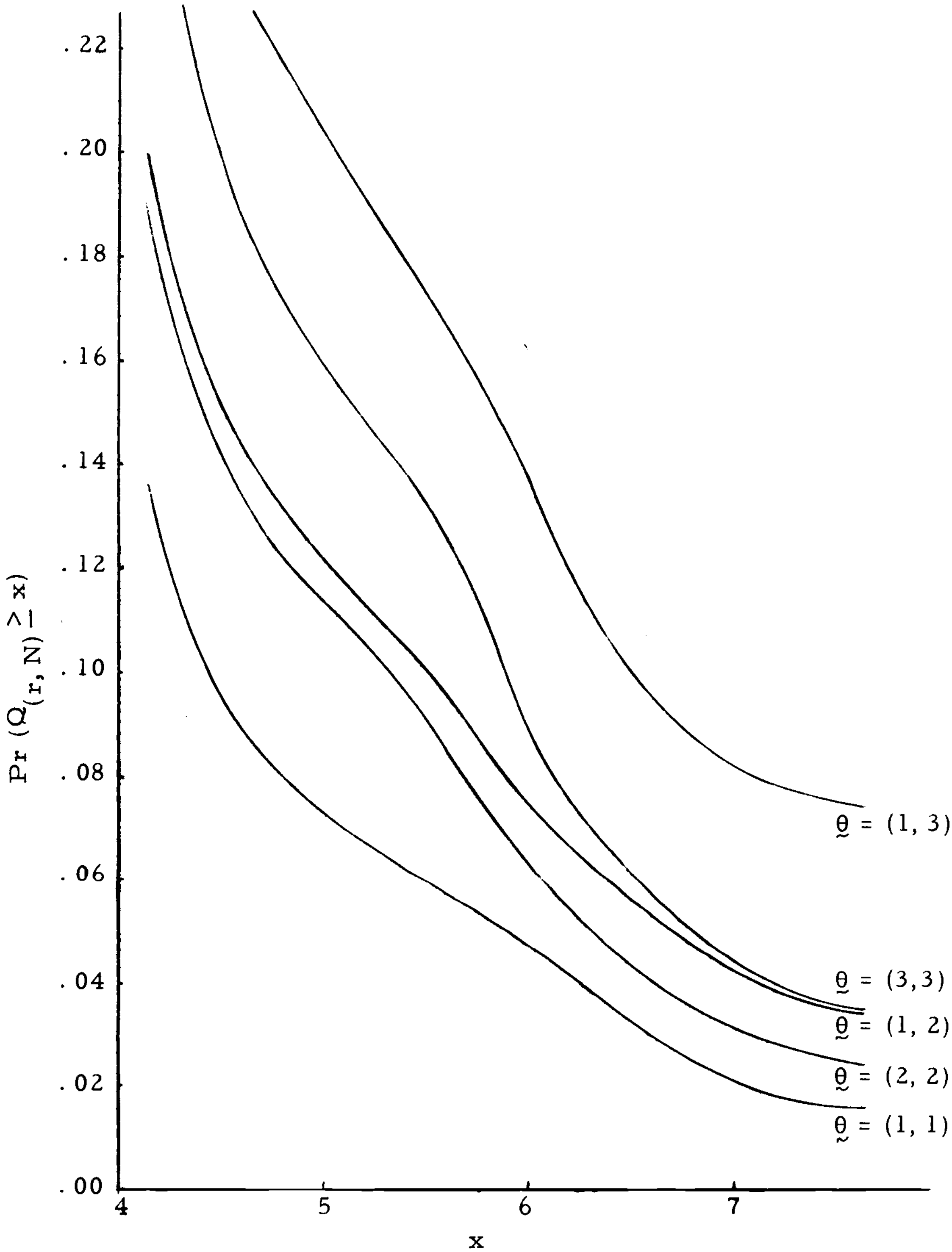

Figure 2. Upper-tailed probabilities of $Q_{(r, N)}$ for different values of $\underset{\sim}{\theta}$ (pattern III).

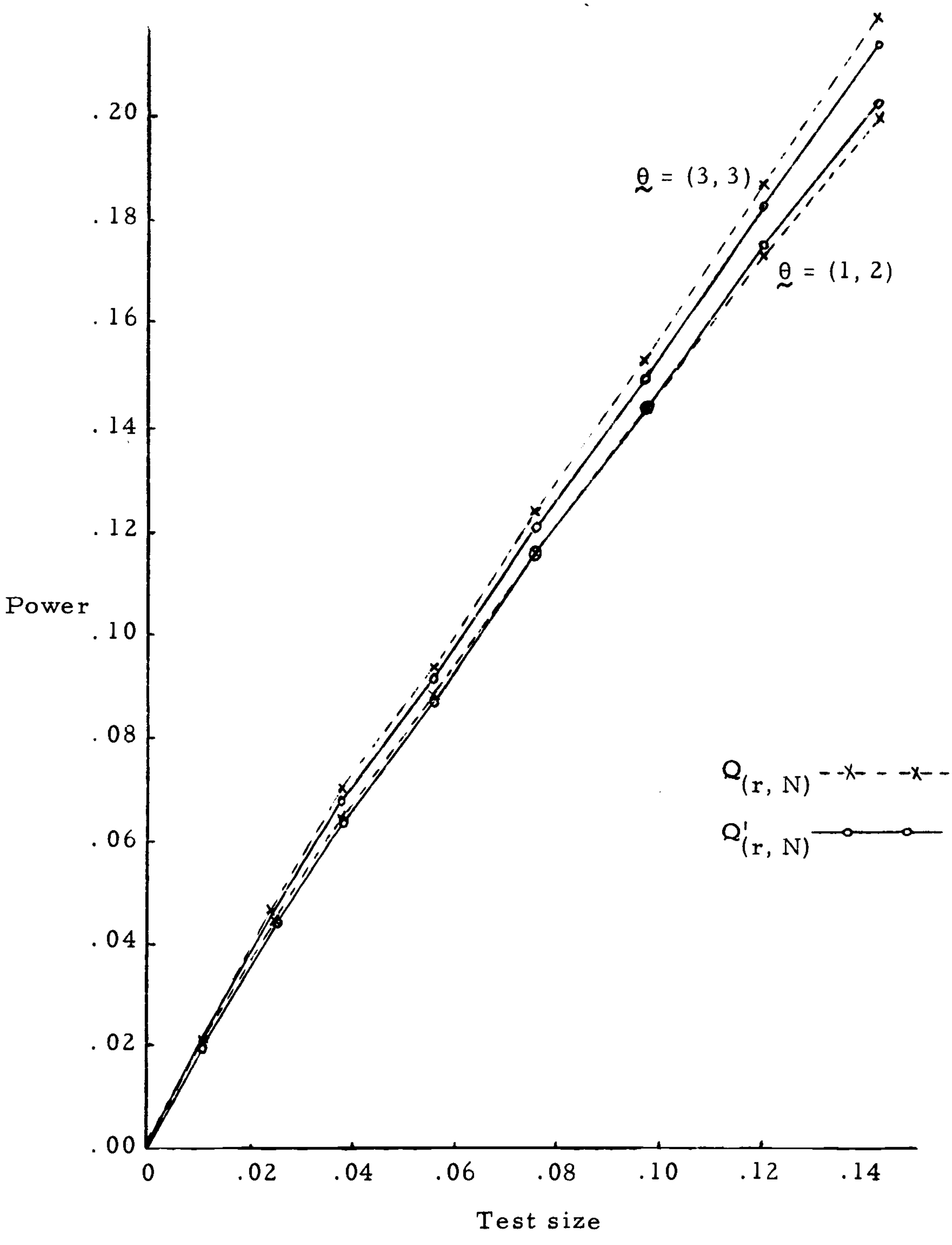

Figure 3. Power comparisons of $Q_{(r, N)}$ and $Q'_{(r, N)}$ for $\underset{\sim}{\theta} = (1, 2)$ and (3, 3) (pattern II).

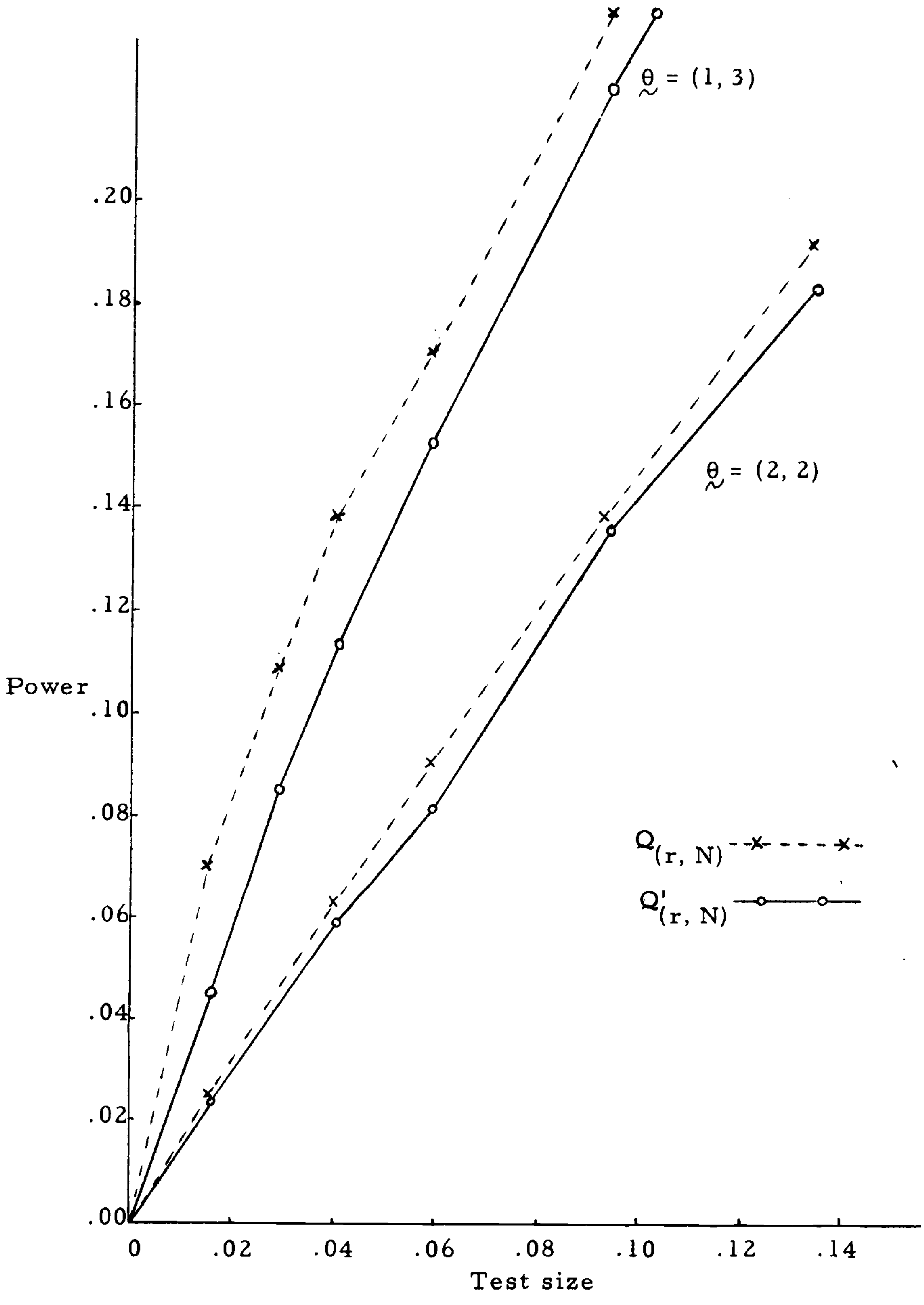

Figure 4. Power comparison of $Q_{(r, N)}$ and $Q'_{(r, N)}$ for $\underset{\sim}{\theta}$ = (2, 2) and (1, 3) (pattern III).

As an illustrative example, we consider data for a 4-sample problem. For $n_1 = n_2 = 5$, $n_3 = 6$, $n_4 = 9$, $r = 10$ and $N = 25$, the data is presented as follows:

$x^{(1)}$: 2.4^+, 3.1, 7.2, 7.2^+, 7.2^+

$x^{(2)}$: 2.4^+, 2.7, 4.6, 5.7

$x^{(3)}$: 1.9^+, 2.3, 2.5^+, 3.9, 7.2^+, 7.2^+

$x^{(4)}$: 1.7^+, 3.2^+, 3.3^+, 3.7, 4.8^+, 7.2^+, 7.2^+, 7.2^+, 7.2^+,

where $x^{(\alpha)+}$ denotes the censoring time of an observation from $\alpha^{\underline{th}}$ sample, and $x^{(\alpha)}$ denotes the failure time of an observation from $\alpha^{\underline{th}}$ sample. In combining these 4 samples, the values of rank-order are obtained in Table 5. With the computed values b_i's, the linear rank statistics can easily be computed by formula,

$$T^{(\alpha)}_{(r, N)} = \frac{1}{\sqrt{N}}[\xi^{(\alpha)}_{\bullet} - \sum_{i=1}^{r} b_i(u_i^{(\alpha)}+\xi_i^{(\alpha)})] ,$$

and the $Q_{(r, N)}$ computed according to the formula,

$$Q_{(r, N)} = \frac{N(N-1)}{(r-b_r)} \sum_{i=1}^{c} T^{(i)^2}_{(r, N)}/n_i \quad .$$

$$T_{(10,25)} = \frac{1}{\sqrt{25}} \begin{bmatrix} \xi_{\cdot}^{(1)} - \sum_{i=1}^{10} b_i(u_i^{(1)}+\xi_i^{(1)}) \\ \xi_{\cdot}^{(2)} - \sum_{i=1}^{10} b_i(u_i^{(2)}+\xi_i^{(2)}) \\ \xi_{\cdot}^{(3)} - \sum_{i=1}^{10} b_i(u_i^{(3)}+\xi_i^{(3)}) \end{bmatrix}$$

$$T_{(10,25)}^{(1)} = \frac{1}{\sqrt{25}}(2-2.42151064) = \frac{1}{\sqrt{25}}(-.42151064)$$

$$T_{(10,25)}^{(2)} = \frac{1}{\sqrt{25}}(4-1.39616851) = \frac{1}{\sqrt{25}}(2.60383149)$$

$$T_{(10,25)}^{(3)} = \frac{1}{\sqrt{25}}(2-2.09435379) = \frac{1}{\sqrt{25}}(-.09435379)$$

$$T_{(10,25)}^{(4)} = -(T_{(10,25)}^{(1)} + T_{(10,25)}^{(2)} + T_{(10,25)}^{(3)})$$

$$= \frac{1}{\sqrt{25}}(-2.08796706)$$

Computational form of $Q_{(r,N)}$ from (4.14) is

$$Q_{(r,N)} = \frac{N-1}{(r-b_r)\lambda_c} \sum_{i=1}^{c} \frac{\lambda_c}{\lambda_i} T_{(r,N)}^{(i)^2}$$

$$\therefore\ Q_{(10,\,25)} = \frac{24}{(10-.68429638)9}\left[\frac{9}{5}(-.42151064)^2 + \frac{9}{5}(2.60383149)^2 + \frac{9}{6}(-.09435379)^2 + (-2.08796706)^2\right]$$

$$= \frac{24}{8.38413326}\left[\frac{9}{5}(.17767122) + \frac{9}{5}(6.08796706)^2 + \frac{9}{6}(.00890264) + (4.35960644)\right]$$

$$= 48.36752635$$

This indicates a significance at level $< .0001$.

Table 5. Rank-order and arbitrary censoring pattern.

i	$\xi_i^{(1)}$	$\xi_i^{(2)}$	$\xi_i^{(3)}$	$u_i^{(1)}$	$u_i^{(2)}$	$u_i^{(3)}$	c_i	E_i	b_i
1	0	1	0	0	0	0	0	25	.04000000
2	0	0	0	0	0	1	1	24	.08166667
3	0	0	1	1	1	1	3	22	.12712121
4	0	1	0	0	0	0	0	18	.18267676
5	1	0	0	0	0	0	2	17	.24150029
6	0	0	0	0	0	0	0	14	.31292886
7	0	0	1	0	0	0	0	13	.38985194
8	0	1	0	0	0	0	1	12	.47318527
9	0	1	0	0	0	0	0	10	.57318527
10	1	0	0	2	0	2	8	9	.68429638

BIBLIOGRAPHY

1. Andrews, F. C. Asymptotic behavior of some rank tests for analysis of variance. The Annals of Mathematical Statistics 25:724-735. 1954.

2. Basu, A. P. On a generalized Savage statistic with applications to life testing. The Annals of Mathematical Statistics 39:1591-1604. 1968.

3. Basu, A. P. On two-sample and k-sample rank tests for censored data. The Annals of Mathematical Statistics 38:1520-1535. 1967.

4. Capon, J. Asymptotic efficiency of certain locally most powerful rank tests. The Annals of Mathematical Statistics 32:88-100. 1961.

5. Cox, D. R. Some simple approximate tests for poisson variates. Biometrika 40:354-360. 1953.

6. Cramer, H. Mathematical methods of statistics. New Jersey, Princeton University, 1956. 575 p.

7. Efron, B. The two sample problem with censored data. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California, 1965. Berkeley, University of California 1967. p. 831-853.

8. Feller, W. An introduction to probability and its applications. Vol. 2. New York, Wiley, 1966. 626 p.

9. Fraser, D. A. S. Nonparametric methods in statistics. New York, Wiley, 1957. 299 p.

10. Fraser, D. A. S. A vector form of the Wald-Wolfowitz-Hoeffding theorem. The Annals of Mathematical Statistics 27:540-543. 1956.

11. Gehan, E. A. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. Biometrika 52:203-223. 1965.

12. Gehan, E. A. and D. G. Thomas. The performance of some two-sample tests in small samples with and without censoring. Biometrika 56:127-132. 1969.

13. Hajek, J. Asymptotic normality of simple linear rank statistics under alternatives. The Annals of Mathematical Statistics 39: 325-346. 1968.

14. Hajek, J. Asymptotically most powerful rank-order tests. The Annals of Mathematical Statistics 33:1124-1147. 1962.

15. Hajek, J. and Z. Sidak. Theory of rank tests. Prague, Academic, 1967. 297 p.

16. Jolly, L. B. W. Summation of series. New York, Dover, 1961. 251 p.

17. Kendall, M. G. and A. Stuart. The advanced theory of statistics. Vol 2. New York, Hafner, 1967. 690 p.

18. Kruskal, W. H. A nonparametric test for the several sample problem. The Annals of Mathematical Statistics 23:525-540. 1952.

19. Lehmann, E. L. Testing statistical hypotheses. New York, 1959. 369 p.

20. Puri, M. L. Asymptotic efficiency of a class of c-sample tests. The Annals of Mathematical Statistics 35:102-121. 1964.

21. Rao, C. R. Linear statistical inference and its applications. New York, Wiley, 1965. 522 p.

22. Rao, U. V. R., I. R. Savage and M. Sobel. Contributions to the theory of rank order statistics: The two-sample censored case. The Annals of Mathematical Statistics 31:415-426. 1960.

23. Sarhan, A. E. and B. G. Greenberg. Contributions to order statistics. New York, Wiley, 1962. 487 p.

24. Savage, I. R. Contributions to the theory of rank order statistics: The two sample case. The Annals of Mathematical Statistics 27:590-615. 1956.

25. Thomas, D. R. Conditional locally most powerful rank tests for the two-sample problems with arbitrarily censored data. Corvallis, 1969. 36 p. (Oregon State University. Dept. of Statistics. Technical Report no. 7)

APPENDICES

APPENDIX I

Proofs of Lemmas 5. 2. 1 and 5. 2. 2

Lemma 5. 2. 1.

$$B_{N,j} = \frac{1}{N}\sum_{i=1}^{r} \frac{V_i^{(j)}}{E_i}\left(1-\frac{V_i^{(j)}}{E_i}\right) \xrightarrow{p} \nu\lambda_j(1-\lambda_j) \qquad \text{(A. 1)}$$

Proof.

$$E\ (V_i^{(j)}) = \sum_{\alpha=i}^{r} E\ (u_\alpha^{(j)}+\xi_\alpha^{(j)})$$

$$= E_i\lambda_j \qquad \text{(A. 2)}$$

and

$$E_{\underset{\sim}{1}}(V_i^{(j)^2}) = E_{\underset{\sim}{1}}\left[\sum_{\alpha=i}^{r}(u_\alpha^{(j)}+\xi_\alpha^{(j)})^2+\sum\sum_{\alpha\neq\alpha'=i}^{r}(u_\alpha^{(j)}+\xi_\alpha^{(j)})(u_{\alpha'}^{(j)}+\xi_{\alpha'}^{(j)})\right]$$

$$= \sum_{\alpha=i}^{r}[\lambda_j(c_\alpha+1)+\lambda_j^2c_\alpha(c_\alpha+1)] + \lambda_j^2\sum\sum_{\alpha\neq\alpha'=i}^{r}(c_\alpha+1)(c_\alpha+1) + o(N)$$

$$= \lambda_jE_i + \lambda_j^2E_i^2 - \lambda_j^2E_i + o(N) \qquad \text{(A. 3)}$$

Combining (A. 2) and (A. 3), we have

$$E\ (B_{N,j}) = \frac{1}{N}\lambda_j(1-\lambda_j)(r-b_r) + o(1) \qquad \text{(A. 4)}$$

Hence,

$$\lim_{N \to \infty} E\,(B_{N,j}) = \nu\lambda_j(1-\lambda_j). \tag{A.5}$$

By expanding the square of $B_{N,j}$ and taking expected value, we get

$$E\,[B_{N,j}^2] = \frac{1}{N^2}\, r(r-1)\lambda_j^2(1-\lambda_j)^2 + o(1)\,. \tag{A.6}$$

From (A. 5) and (A. 6), it follows that

$$\begin{aligned} \mathrm{Var}\ (B_{N,j}) &= o(1) \\ &= 0 \quad \text{as} \quad N \to \infty\,. \end{aligned}$$

Then,

$$B_{N,j} \xrightarrow{p} \nu\lambda_j(1-\lambda_j).$$

Lemma 5. 2. 2.

$$B_{N,jj'} = -\frac{1}{N}\sum_{i=1}^{r} \frac{V_i^{(j)}V_i^{(j')}}{E_i^2} \xrightarrow{p} -\nu\lambda_j\lambda_{j'}\,. \tag{A.7}$$

Proof. Consider $V_i^{(j)}V_i^{(j')}$,

$$V_i^{(j)}V_i^{(j')} = \sum_{a=i}^{r}(u_a^{(j)}+\xi_a^{(j)}) \sum_{a=i}^{r}(u_a^{(j')}+\xi_a^{(j')})$$

$$= \sum_{a=i}^{r}(u_a^{(j)}u_a^{(j')}+\xi_a^{(j)}u_a^{(j')}+\xi_a^{(j')}u_a^{(j)}) \quad +$$

$$+ \sum_{\alpha \neq \alpha' = i}^{r}\sum (u_{\alpha}^{(j)}+\xi_{\alpha}^{(j)})(u_{\alpha'}^{(j')}+\xi_{\alpha'}^{(j')}).$$

$$E\ (V_i^{(j)}V_i^{(j')}) = \sum_{\alpha=i}^{r} [c_{\alpha}(c_{\alpha}-1)+c_{\alpha}+c_{\alpha}]\lambda_j\lambda_{j'}$$

$$+ \sum_{\alpha \neq \alpha' = i}^{r}\sum (c_{\alpha}+1)(c_{\alpha'}+1)\lambda_j\lambda_{j'}$$

$$= \lambda_j\lambda_{j'}E_i^2 - \lambda_j\lambda_{j'}E_i. \tag{A. 8}$$

Hence,

$$E\ (B_{N,jj'}) = -\frac{r}{N}\lambda_j\lambda_{j'} + \lambda_j\lambda_{j'}\frac{b_r}{N}.$$

From results of Lemma 5. 2. 2, we get

$$\lim_{N\to\infty} E\ (B_{N,jj'}) = -\nu\lambda_j\lambda_{j'}. \tag{A. 9}$$

By taking the expected value of sum squares and cross-products of $B^2_{N,jj'}$,

$$E\ [B^2_{N,jj'}] = \frac{1}{N^2} r(r-1)\lambda_j\lambda_{j'} + o(1) \tag{A.10}$$

From the results of (A. 9) and (A. 10), we have

$$\lim_{N\to\infty} \mathrm{Var}\ [B^2_{N,jj'}] = 0.$$

APPENDIX II

Limitation of R_N in Theorem 5. 2

By using Taylor's expansion of $\ln(1+x)$ to expand $\ln L(\underset{\sim}{\theta})$ of (5. 25), we have

$$\ln L(\underset{\sim}{\theta}) = \sum_{j=1}^{c-1} \xi_{\bullet}^{(j)} \ln(1+\delta_j/\sqrt{N}) - \sum_{a=1}^{c-1} \ln\left[1+\sum_{a=1}^{c-1} \frac{\delta_a V_i^{(a)}}{E_i\sqrt{N}}\right]$$

$$= \sum_{j=1}^{c-1} \xi_{\bullet}^{(j)} \left[\frac{\delta_j}{\sqrt{N}} - \frac{\delta_j^2}{2N} + \frac{\delta_j'^3}{3N\sqrt{N}}\right]$$

$$-\sum_{i=1}^{r}\left[\sum_{a=1}^{c-1} \frac{\delta_a V_i^{(a)}}{E_i\sqrt{N}} - \frac{1}{2NE_i^2}\left(\sum_{a=1}^{c-1} \delta_a V_i^{(a)}\right)^2 + \frac{1}{3N\sqrt{N}E_i^3}\left(\sum_{a=1}^{c-1} \delta_a' V_i^{(a)}\right)^3\right] \tag{A.11}$$

where $-\delta_j \leq \delta_j' \leq \delta_j$ for $j = 1, 2, \ldots, c-1$. Then the remainder term R_N in (5. 26) is

$$\sum_{j=1}^{c-1} \xi_{\bullet}^{(j)} \frac{\delta_j'^3}{3N\sqrt{N}} - \sum_{i=1}^{r} \frac{1}{3N\sqrt{N}\, E_i^3}\left[\sum_{a=1}^{c-1} \delta_a' V_i^{(a)}\right]^3$$

Since

$$E_i = \sum_{a=1}^{c} V_i^{(a)}, \quad \text{then} \quad \left|\sum_{a=1}^{c-1} \delta_a' V_i^{(a)}\right|^3 \leq \tilde{\delta}^3 E_i^3 \,,$$

where $\tilde{\delta} = \max(|\delta_1|, |\delta_2|, \ldots, |\delta_c|)$, it follows that

$$|R_N| \leq \frac{1}{3\sqrt{N}}\tilde{\delta}^3 + \sum_{i=1}^{r} \frac{\tilde{\delta}^3}{3N\sqrt{N}}$$

$$\leq \frac{2}{3}\frac{\tilde{\delta}^3}{\sqrt{N}} = o(1). \qquad (A.12)$$

APPENDIX III

c-Sample Generalization of Gehan's Statistics

Define

$$W_\alpha = \sum_{i=1}^{n_\alpha} \sum_{j=1}^{N-n_\alpha} U_{ij}^{(\alpha)} \quad , \tag{A.13}$$

with

$$U_{ij}^{(\alpha)} = \begin{cases} -1 & x_i^{(\alpha)} < y_j^{(\alpha)} \quad \text{or} \quad x_i^{(\alpha)} < y_j^{(\alpha)'} \\ 0 & x_i^{(\alpha)} = y_j^{(\alpha)} \quad \text{or} \quad x_i^{(\alpha)'} < y_j^{(\alpha)} \quad \text{or} \quad y_j^{(\alpha)'} < x_i^{(\alpha)} \\ 1 & x_i^{(\alpha)} > y_j^{(\alpha)} \quad \text{or} \quad x_i^{(\alpha)'} > y_j^{(\alpha)} \quad , \end{cases} \tag{A.14}$$

where $x_i^{(\alpha)}$ and $x_i^{(\alpha)'}$ are the $i^{\underline{th}}$ failure and the $i^{\underline{th}}$ censored observations within the $\alpha^{\underline{th}}$ sample, and $y_j^{(\alpha)}$ and $y_j^{(\alpha)'}$ are the $j^{\underline{th}}$ failure and the $j^{\underline{th}}$ censored observations within the $\alpha^{\underline{th}}$ sample respectively.

Assuming no ties in observations, we may express W_α as a linear function of $\underset{\sim}{\xi}^{(\alpha)}$ and $\underset{\sim}{u}^{(\alpha)}$,

$$W_\alpha = \sum_{i=1}^{r} \left[\left(i - \sum_{j=i}^{r} (c_j+1) \right) \xi_i^{(\alpha)} + i\, u_i^{(\alpha)} \right] \tag{A.15}$$

$\alpha = 1, 2, \ldots, c.$

It can be easily verified, by using the similar techniques for the case of $T^{(j)}_{(r, N)}$, that

$$E_{\underset{\sim}{1}}(W_\alpha) = 0$$

$$V_{\underset{\sim}{1}}(W_\alpha) = \frac{n_\alpha(N-n_\alpha)}{N(N-1)} \sum_{i=1}^{r} [(i-E_i)^2 + i^2 c_i]$$

and

$$\mathrm{Cov}_{\underset{\sim}{1}}(W_\alpha, W_{\alpha'}) = -\frac{n_\alpha n_{\alpha'}}{N(N-1)} \sum_{i=1}^{r} [(i-E_i)^2 + i^2 c_i] \tag{A.16}$$

$$(\alpha \neq \alpha' = 1, 2, \ldots, c).$$

The modified Gehan's statistic for c-sample is then defined as

$$\begin{aligned} Q'_{(r, N)} &= \underset{\sim}{W}' \Sigma^{-1}_{\underset{\sim}{W}} \underset{\sim}{W} \\ &= \frac{N-1}{\sum_{i=1}^{r} [(i-E_i)^2 + i^2 c_i]} \sum_{i=1}^{c} \frac{W_i^2}{n_i} . \end{aligned} \tag{A.17}$$