1982

Nonparametric Analysis of Right Censored Data with Multiple Comparisons

Hwei-Weng Shih
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/gradreports
Part of the Applied Statistics Commons

Recommended Citation
Shih, Hwei-Weng, "Nonparametric Analysis of Right Censored Data with Multiple Comparisons" (1982). All Graduate Plan B and other Reports. 1170.
https://digitalcommons.usu.edu/gradreports/1170

This Report is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Plan B and other Reports by an authorized administrator of DigitalCommons@USU. For more information, please contact dylan.burns@usu.edu.
NONPARAMETRIC ANALYSIS OF RIGHT CENSORED
DATA WITH MULTIPLE COMPARISONS

by

Hwei-Weng Shih

A report submitted in partial fulfillment
of the requirements for the degree
of
MASTER OF SCIENCE
in
Applied Statistics
(Plan B)

UTAH STATE UNIVERSITY
Logan, Utah
1982
ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my major professor, David L. Turner, for his patient guidance, and help in organizing the paper. Thanks are also extended to my committee members, Dr. Ronald V. Canfield and Dr. Gregory Jones, for their help and criticism.

I would also like to express my sincere thanks to my parents for their encouragement and support in my graduate studies at Utah State University.

Hwei-Weng Shih
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td></td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td></td>
<td>iv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td></td>
<td>v</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I.</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II.</td>
<td>METHODOLOGY</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>A Special Comparison for Two Exponential Distributions</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>The Wilcoxon Two Sample Rank Sum Test and the Kruskal-Wallis Test</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>A Generalized Kruskal-Wallis Test for Comparing K Censored Samples</td>
<td>7</td>
</tr>
<tr>
<td>III.</td>
<td>MULTIPLE COMPARISONS</td>
<td>12</td>
</tr>
<tr>
<td>IV.</td>
<td>EXAMPLES</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>The Methods Used</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Example 1</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Example 2</td>
<td>21</td>
</tr>
<tr>
<td>V.</td>
<td>CONCLUSIONS</td>
<td>28</td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>APPENDIXES</td>
<td></td>
<td>31</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data and $\lambda_i$ for sorted Monte Carlo data</td>
<td>16</td>
</tr>
<tr>
<td>2. The results of example 1 using F test for censored data in Table 1</td>
<td>17</td>
</tr>
<tr>
<td>3. The ranks for data censored at 20 from Table 1</td>
<td>18</td>
</tr>
<tr>
<td>4. The results of example 1 using Kruskal-Wallis test and the ranking in Table 3</td>
<td>19</td>
</tr>
<tr>
<td>5. Statistics $S^{**}$, $S^*$ and $S$ and the results of example 1 for F test, Kruskal-Wallis test, Breslow's method and Bonferroni method</td>
<td>20</td>
</tr>
<tr>
<td>6. The values $n_i$, $\hat{X}_i$, $X_i$ and $\lambda_i$ of each treatment</td>
<td>24</td>
</tr>
<tr>
<td>7. The results for example 2 using F test</td>
<td>24</td>
</tr>
<tr>
<td>8. The results of example 2 using Kruskal-Wallis test</td>
<td>25</td>
</tr>
<tr>
<td>9. The results of example 2 using F test, Kruskal-Wallis test, Breslow's method and Bonferroni method</td>
<td>26</td>
</tr>
</tbody>
</table>
ABSTRACT

Nonparametric Analysis of Right Censored Data with Multiple Comparisons

by

Hwei-Weng Shih, Master of Science
Utah State University, 1982

Major Professor: Dr. David L. Turner
Department: Applied Statistics

This report demonstrates the use of a computer program written in FORTRAN for the Burroughs B6800 computer at Utah State University to perform Breslow's (1970) generalization of the Kruskal-Wallis test for right censored data. A pairwise multiple comparison procedure using Bonferroni's inequality is also introduced and demonstrated. Comparisons are also made with a parametric F test and the original Kruskal-Wallis test. Application of these techniques to two data sets indicate that there is little difference among the procedures with the F test being slightly more liberal (too many differences) and the Kruskal-Wallis test corrected for ties being slightly more conservative than Breslow's test statistic.

(30 pages)
CHAPTER I
INTRODUCTION

Statistical relationships between variables must sometimes be estimated from incomplete data or data which has been censored. Censoring occurs when an experiment is stopped before the event of interest occurs. When this happens the recorded data do not provide direct information about the event. In this paper we shall consider only samples censored on the right. This means that the only information about the censored observations is their total number and the fact that each is greater than some known value. For example, if we were studying survival time of a patient or animal under a set of experimental conditions, the data would be analyzed while some patients or animals are still alive. According to Lagakos (1979) the analysis of censored data can be used to obtain as much information as an uncensored experiment would yield.

A fundamental problem in many life testing problems is a comparison of the survival-time distributions from two or more samples of censored data. Norman Breslow (1970) reviews a generalization of Wilcoxon's statistic for comparing two populations as proposed by Gehan (1965) for use when the observations are subject to arbitrary right censorship. Breslow also discusses Mantel's (1967) further generalization to the case of arbitrarily restricted observations, or left and right censorship. Both Mantel and Gehan base their calculations on the permutation distribution of the statistic, conditional on the observed censoring pattern for the combined sample.
Breslow (1970) extended Gehan's generalization of Wilcoxon's test to allow for testing the equality of K continuous distribution functions when observations are subject to arbitrary right censorship. Breslow's generalization is an extension of the Kruskal-Wallis test, and is the "state of the art" nonparametric test of equality of K groups with possibly differing distributions for the censoring variables.

Breslow's development of this extended Kruskal-Wallis test involves some very complicated formulae. He gives two "easy" approximations but even these would be very laborious to compute.

This report demonstrates the use of a computer program written for the Burroughs B6800 computer which translates Breslow's formulae into a form which may actually be used. A pairwise multiple comparison procedure using Bonferroni's inequality is also developed and demonstrated. Two sets of data will be analyzed using Breslow's procedure and the Bonferroni multiple comparison procedure. Comparisons will also be made with the parametric (F test) procedure for the case of data from exponential distributions. The nonparametric Kruskal-Wallis test for uncensored data will also be applied using the modifications for tied data discussed in Ott (1977).
In this report the major method used to analyze the hypothesis is Breslow's generalization of the Kruskal-Wallis test. In addition to Breslow's method, the Kruskal-Wallis test is also used to test the hypothesis that \( K \geq 2 \) populations are identical using modifications when there are ties in the data. An F test for the case of two exponential distributions is also performed. Comparisons will then be made among the various methods when results are known using a set of generated or Monte Carlo Data. The methods are then applied to a real set of data.

A Special Comparison for Two Exponential Distributions

Let the two exponential distributions with parameters \( \lambda_1 \) and \( \lambda_2 \) have probability density function

\[
f(X_{ij}; \lambda_i) = \frac{1}{\lambda_i} \exp(-X_{ij}/\lambda_i), \quad i = 1, 2; \ j = 1, 2, \ldots, n_i
\]

Then

\[
\hat{\lambda}_1 = \frac{\sum_{j=1}^{n_i} X_{ij}}{\delta_{ij}} = \frac{X_{i.}}{\delta_{i.}},
\]

is the maximum likelihood estimate of \( \lambda_1, i = 1, 2 \), where \( X_{ij} \) equals the true value or censored value depending on whether \( \delta_{ij} \) equals 1 (uncensored) or 0 (censored). Then
is an $F$ distributed random variable with $2\delta_1$ and $2\delta_2$ degrees of freedom. This result may be used to test $H_0: \lambda_1 = \lambda_2$.

**The Wilcoxon Two Sample Rank Sum Test and the Kruskal-Wallis Test**

The Wilcoxon rank sum test provides a nonparametric test of the hypothesis that two populations are identical, since the experimenter has obtained two samples from possibly different populations, and we wish to use a statistical test to see if we can reject the null hypothesis that the two populations are identical. That is, we wish to detect differences between the two populations on the basis of random samples from those populations. An approach to the two-sample problem is to rank the combined data from lowest to highest. We let $R_1$ denote the sum of the ranks for sample 1. $R_1$ can take on values ranging from $n_1(n_1 + 1)/2$ to $(n_1 + n_2)(n_1 + n_2 + 1)/2 - n_2(n_2 + 1)/2$. Intuitively, if $R_1$ is close to either extreme, we would have evidence to reject the null hypothesis that the two populations are identical, since sample 1 would then be all close to the bottom or the top of the ranked distribution.

The concept of a rank sum test was extended to a comparison of more than two populations by Kruskal and Wallis (1952). The $K \geq 2$ random samples have been obtained from each of $K$ possibly different populations, and we want to test the null hypothesis that all of the populations are identical against the alternative that some of the populations tend to furnish greater observed values than other populations.
To perform the test, the \( K \geq 2 \) samples are combined into a single ordered sample, then ranks are assigned to the sample values from the smallest value to the largest, without regard to which population each value came from. Let \( N \) denote the total number of observations,

\[
N = \sum_{i=1}^{K} n_i
\]

where \( n_i \) is the number of observations from sample \( i \). Let \( R(X_{ij}) \) denote the rank assigned to \( X_{ij} \), \( R_i \) be the sum of the ranks assigned to the \( i \)th sample,

\[
R_i = \sum_{j=1}^{n_i} R(X_{ij}) \quad i=1, 2, \ldots, K.
\]

Note that

\[
\sum_{i=1}^{K} R_i = 1 + 2 + \ldots + N = \frac{N(N+1)}{2}.
\]

If there are several observations tied or equal to each other, the average of their ranks is assigned to each of the tied observations.

The large sample approximation for the test statistic \( T \) is based on the fact that \( R_i \) is the sum of \( n_i \) random variables. So the mean and variance of \( R_i \) are given by

\[
E(R_i) = \frac{n_i(N+1)}{2}
\]

and

\[
Var(R_i) = \frac{n_i(N+1)(N-n_i)}{2}
\]
Therefore

\[
\frac{R_j - E(R_j)}{\sqrt{\text{Var}(R_j)}}
\]

is approximately distributed as a standardized normal random variable when \( n_i \) is large enough. Thus

\[
\left( \frac{R_i - E(R_i)}{\sqrt{\text{Var}(R_i)}} \right)^2 = \frac{(R_i - [ni(N + 1)/2])^2}{n_i(N + 1)(N - n_i)/12}
\]

is approximately distributed as a chi-square random variable with one degree of freedom. If the \( R_i \) were independent of each other the distribution of the sum

\[
T = \sum_{i=1}^{K} \frac{(R_i - [ni(N + 1)/2])^2}{n_i(N + 1)(N - n_i)/12}
\]

could be approximated using the chi-square distribution with \( K \) degrees of freedom. However, since the sum of the \( n_i \)’s is \( N \), there is some dependence among the \( R_i \)’s. Kruskal (1952) showed that if the ith term in \( T \) is multiplied by \((N-n_i)/N\) for \( i = 1, 2, \ldots, K \), then the result

\[
T = \sum_{i=1}^{K} \frac{(R_i - [ni(N + 1)/2])^2}{n_i(N + 1)N/12}
\]

is asymptotically distributed as a chi-square random variable with \( K - 1 \) degrees of freedom. Since \( \sum_{i=1}^{K} R_i = N(N + 1)/2 \), \( T \) may be written as

\[
T = \sum_{i=1}^{K} \frac{(R_i - [ni(N + 1)/2])^2}{n_i(N + 1)N/12}
\]

\[
= \frac{12}{N(N + 1)} \sum_{i=1}^{K} \frac{1}{n_i} [R_i^2 - R_i n_i(N + 1) + \frac{1}{4} n_i^2(N + 1)^2]
\]
\[
= \frac{12}{N(N + 1)} \sum_{i=1}^{K} \frac{R_i^2}{n_i} - \frac{12}{N(N + 1)} \left( \frac{N(N + 1)}{2} \cdot (N + 1) - \frac{N}{4(N + 1)^2} \right) \\
= \frac{12}{N(N + 1)} \sum_{i=1}^{K} \frac{R_i^2}{n_i} - 3(N + 1),
\]

is an equivalent form for T, and is usually more convenient to use.

A modification proposed by Ott uses T' rather than T when there are groups of tied ranks. To do this we form the g groups composed of identical ranks, where the jth group contains \( t_j \) (\( j = 1, \ldots, g \)) ties.

The statistic T' is then close to a chi-square random variables with \( K - 1 \) degrees of freedom where

\[
T' = \frac{g}{1 - \left[ \sum_{j=1}^{g} \frac{(t_j^3 - t_j)}{N^3 - N} \right]}. 
\]

A Generalized Kruskal-Wallis Test for Comparing K Censored Samples

Although the Kruskal-Wallis test assumes only continuous underlying distributions, it does not do very well if there are large numbers of ties. This is especially so for censored data when the ties may lie among the upper values of the ranks.

To handle problems of right censored data, Breslow (1970) generalized the Kruskal-Wallis test. Let \( X_{ij}^0 \) be the true observation for the jth individual obtained from the ith population (\( j = 1, \ldots, N_i; i = 1, \ldots, K \)). Variable \( Z_{ij} \) is used to censor \( X_{ij}^0 \), so sometimes the true observation \( X_{ij}^0 \) may not be observed. The observed data which we can get from a real sample is \( X_{ij} = \min(X_{ij}^0, Z_{ij}) \). \( X_{ij} \) should indicate with a variable \( \delta_{ij} \) whether or not \( X_{ij} \) is in fact censored: i.e.,

\( \delta_{ij} = 1 \) when \( X_{ij} = X_{ij}^0 < Z_{ij} \) (uncensored); \( \delta_{ij} = 0 \) when \( X_{ij} = Z_{ij} < X_{ij}^0 \).
(censored). \( N = N_1 + \ldots + N_k \) is the total sample size and \( \lambda_i = N_i / N \) is the proportion of the \( i \)th sample size to the total sample size.

\( F_i \) is the \( i \)th cumulative distribution function. The null hypothesis to be tested is \( H_0: F_1 = \ldots = F_k \), which specified that \( K \) populations have equal distribution functions.

Breslow (1970) defined a scoring function \( x \) for comparing two observations \( X_{ij} \) and \( X_{i'j'} \) by

\[
x(X_{ij}, \delta_{ij}; X_{i'j'}, \delta_{i'j'}) = \begin{cases} 
-1 & X_{ij} < X_{i'j'}, \delta_{ij} = 1, \delta_{i'j'} = 1 \\
-1 & X_{ij} < X_{i'j'}, \delta_{ij} = 1, \delta_{i'j'} = 0 \\
+1 & X_{ij} > X_{i'j'}, \delta_{ij} = 1, \delta_{i'j'} = 1 \\
+1 & X_{ij} > X_{i'j'}, \delta_{ij} = 0, \delta_{i'j'} = 1 \\
0 & \text{otherwise.}
\end{cases}
\]

The \( x \) function is then used in computing a vector score statistic, \( W \). The \( i \)th component of this vector score statistic is defined to be the total score comparing the \( i \)th sample with the remaining \( K - 1 \) samples,

\[
W_i = \sum_{j=1}^{N_i} \sum_{i' = 1}^{K} \sum_{j' = 1}^{N_i} x(X_{ij}, \delta_{ij}, X_{i'j'}, \delta_{i'j'}). 
\]

For uncensored data sets, \( W_i = 2[R_i - (1/2)N_i(N + 1)] \). Large negative values of \( W_i \) mean that observations in the \( i \)th sample are smaller than those from other samples and large positive values of \( W_i \) would indicate that the \( i \)th sample had larger than average values. The total of \( W_i \) should be equal to 0.

Breslow (1970) goes on to use this \( W \) vector to form test statistics for testing the equality of \( K \) distribution functions. His first statistic refers to Rao (1965) which shows that the well-known large sample theory for chi-square statistics holds for the statistic.
\[ S^* = \sum_{i'=1}^{K} \sum_{j'=1}^{N_i} \delta_{i',j'} \left( \sum_{i=1}^{K} \sum_{j=1}^{N_i} e(x_{ij} - x_{i',j'}) \right)^2 \]

where
\[ e(x) = \begin{cases} 
1 & \text{if } x > 0 \\
0 & \text{if } x \leq 0.
\end{cases} \]

Under general regularity conditions, \( S^* \) can be shown to have an asymptotic chi-square distribution with \( K - 1 \) degrees of freedom. Breslow recommends evaluation of \( S^{**} \) in order to check on computational accuracy of \( S^* \). \( S^{**} \) is a lower bound for \( S^* \) and is easily computed as:
\[ S^{**} = 3N^{-2} \sum_{i=1}^{K} (W_i/\hat{N}_i)^2. \]

Breslow (1970) goes on to develop a statistic which is calculated as follows. A covariance matrix must be computed. Individual terms \( \sigma_{ii'} \) can be calculated from
\[
\begin{align*}
\sigma_{ii'} &= -\sum_{i'=1}^{K} \sum_{j'=1}^{N_{i''}} \sum_{j=1}^{N_{i''}} e(x_{ij} - x_{i'',j}) \times e(x_{i''j}) \times e(x_{ij} - x_{i'',j}) \\
\sigma_{ii''} &= -\sum_{i'' \neq i} \sigma_{ii''}\]
\end{align*}
\]

where \( e(x) = 1 \) if \( x > 0 \), 0 if \( x \leq 0 \).

The covariance matrix is then decomposed into \( K - 1 \) vectors
\[ \ell_i = (\ell_{ii}, \ldots, \ell_{ik})' \quad (i=1, \ldots, K-1) \] such that \( \ell_i' \ell_i = 1 \) and \( \ell_i' \ell_j = 0 \) \( (i \neq j) \). The vectors \( \ell_i \) may be easily found by using the Gram-Schmidt orthogonalization process. We can use the \( K - 1 \) vectors
\[ Y_1 = (1, 0, \ldots, 0)', \ Y_2 = (1, 1, 0, \ldots, 0)', \ldots, \]
\[ Y_{K-1} = (1, \ldots, 1, 0)' \]

as a starting point for the Gram-Schmidt process. We denote the inner product of two vectors by

\[(X, Y) = X' \cdot Y.\]

The Gram-Schmidt process proceeds as follows:

1. Calculate \( \xi_1 = \frac{Y_1}{\|Y_1\|} \) where \( \|Y_1\| = \sqrt{(Y_1, Y_1)} = \sqrt{Y_1' \cdot Y_1} \)

2. Use \( \xi_1 \) to calculate \( Z_2 = Y_2 - (Y_2, \xi_1)\xi_1 \) and \( \xi_2 = \frac{Z_2}{\|Z_2\|} \)

where \( \|Z_2\| = \sqrt{(Z_2, Z_2)} = \sqrt{Z_2' \cdot Z_2} \)

3. Use above information to calculate \( Z_3 = Y_3 - (Y_3, \xi_1)\xi_1 - (Y_3, \xi_2)\xi_2 \)

and \( \xi_3 = \frac{Z_3}{\|Z_3\|} \) where \( \|Z_3\| = \sqrt{(Z_3, Z_3)} = \sqrt{Z_3' \cdot Z_3} \)

\((K-1)\)st. Step. Use above information to calculate

\[ Z_{K-1} = Y_{K-1} - (Y_{K-1}, \xi_1)\xi_1 - (Y_{K-1}, \xi_2)\xi_2 - \cdots \]

\[ - (Y_{K-1}, \xi_{K-2})\xi_{K-2} \]

and

\[ \xi_{K-1} = \frac{Z_{K-1}}{\|Z_{K-1}\|} \]

where \( \|Z_{K-1}\| = \sqrt{(Z_{K-1}, Z_{K-1})} = \sqrt{Z_{K-1}' \cdot Z_{K-1}} \)

\[ = \frac{1}{\sqrt{Z_{K-1}' \cdot Z_{K-1}}}. \]

The statistics

\[ S_i = N^{-3/2} \xi_i' \frac{W}{i} \]

are easily found and a combined test statistic is calculated as
which is a chi-square random variable with \( K - 1 \) degrees of freedom.

Breslow suggests calculating \( S^{**} \) as a lower bound to \( S^* \). \( S \) and \( S^* \) are asymptotically equivalent statistics, but \( S^* \) is computationally far easier to compute. The "easier" \( S^* \) and \( S^{**} \) are needed only if a computer program is not available to calculate \( S \). A computer program is given in the Appendix which translates Breslow's formulae into a FORTRAN IV program for the Burroughs B6800 computer.
CHAPTER III
MULTIPLE COMPARISONS

The procedures described in Chapter II provide an overall test of the equality of \( K \geq 2 \) distributions. For \( K > 2 \), if the populations are declared significantly different, then a multiple comparison procedure is needed to isolate the differences.

The Bonferroni inequality provides one method of simultaneously estimating several confidence intervals. Let \( A_1 \) denote the first event, say a \( 1 - \alpha_1 \) confidence interval, and let \( A_2 \) denote the second event also a \( 1 - \alpha_2 \) confidence interval. We can then use the Bonferroni inequality to get the probability of both events of \( A_1 \) and \( A_2 \) occurring simultaneously. We already know that

\[
P(A_1 \cap A_2) = 1 - P(\bar{A}_1) - P(\bar{A}_2) + P(\bar{A}_1 \cap \bar{A}_2)
\]

and since \( P(A_1 \cap A_2) \geq 0 \), we obtain the Bonferroni inequality:

\[
P(A_1 \cap A_2) \geq 1 - P(A_1^c) - P(A_2^c).
\]

For this situation, the joint confidence is

\[
P(A_1 \cap A_2) \geq 1 - \alpha_1 - \alpha_2.
\]

The Bonferroni inequality can easily be extended to \( K \) simultaneous confidence intervals with family confidence coefficient \( 1 - \alpha \) by requiring \( P(\bar{A}_i) = \alpha_i \); and \( \sum \alpha_i = \alpha \) which then gives

\[
P(\bigcap_{i=1}^{K} A_i) \geq 1 - \alpha.
\]
For example, let

\[ A_{12} \] be a 99% confidence interval for \( \mu_1 - \mu_2 \),
\[ A_{13} \] be a 99% confidence interval for \( \mu_1 - \mu_3 \),
and \( A_{23} \) be a 99% confidence interval for \( \mu_2 - \mu_3 \).

The Bonferroni inequality then guarantees us a family or simultaneous confidence interval of at least 97 percent that the three intervals based on the same sample are simultaneously correct, i.e.,

\[ P(A_{12} \cap A_{13} \cap A_{23}) \geq .97. \]

If \( K \) interval estimates are desired with a family confidence coefficient \( 1 - \alpha \), constructing each interval estimate with statement confidence coefficient \( 1 - \alpha/K \) will suffice. The Bonferroni technique is ordinarily most useful when the number of simultaneous estimates is not too large. Note that different statement confidence coefficients also could be calculated, as long as \( \sum_{i=1}^{K} P(\bar{A}_i) = \alpha \).

For instance, the event \( A_1 \) may be a 98 percent confidence interval and the event \( A_2 \) could be a 97 percent confidence interval. The family confidence coefficient would then be at least 95 percent.
In this chapter two examples are given to illustrate the analysis of censored data. We will apply the F test for exponential data, the Kruskal-Wallis test for ranked data and Breslow's method for censored data. The Bonferroni method will then be used to test pairwise comparisons.

Example 1

In the first example, three groups of data were generated from known exponential distributions. The procedure to get the three data sets uses an integral transform, i.e., if \( F(x) \) is the distribution function for a random variable \( X \), and if \( X_1, \ldots, X_n \) is a random sample from \( F(\cdot) \) then \( U_i = F(X_i) \) for \( i = 1, \ldots, n \) will be a random sample of uniform random variables over the interval \((0, 1)\). It follows then that if \( U_1, \ldots, U_n \) is a random sample from a uniform distribution, then \( X_i = F^{-1}(U_i) \) for \( i = 1, \ldots, n \) will be a random sample from \( F(\cdot) \).

It is easy to use a computer to generate uniform random numbers, and then we may use the integral transformation technique for finding random numbers from a given distribution. For this example we got three groups of uniform random numbers from 0 to 1 using MINITAB. If \( F(\cdot) \) is a negative exponential distribution, \( F_X(x) = 1 - e^{-x/\lambda} = \mu \), then \( X = F_X^{-1}(\mu) = -\lambda \ln (1-\mu) \) has a negative exponential distribution.
with parameter λ. i.e., the density function of X is \( f_X(x) = \frac{1}{\lambda} e^{-x/\lambda} \), which is a negative exponential distribution. Table 1 presents such samples from three negative exponential distributions. Each sample has been sorted for ease in censoring at an arbitrary value of 20.

If we ignore the fact that the data is censored, we can get \( \hat{\lambda}_i = \frac{X_i}{\delta_i} \) for each group and then use \( R = \frac{\hat{\lambda}_i}{\hat{\lambda}_j} \) (\( i \neq j \)) to do an F test with \( 2 \delta_i \) and \( 2 \delta_j \) degree of freedom. All possible F tests are listed in Table 2.

If we ignore the censoring in Table 1, Table 3 then gives the \( R_i \)'s, the sum of the ranks for each group, and the Kruskal-Wallis tests are listed in Table 4.

Since Breslow's method includes many complicated formulae, the computer program listed in the Appendix was used to get the statistics \( S^{**}, S^* \) and \( S \). We used \( S \) to do the chi-square test and calculated \( S^* \) and \( S^{**} \) for illustrative purpose only. Using an experimentwise error rate of .05, the three pairwise comparisons are \( G_1 = G_2 \), \( G_1 = G_3 \) and \( G_2 = G_3 \). The Bonferroni procedure then uses \( 1 - .05/3 = .9833 \) as confidence coefficient for the individual intervals. We list the results of the F test, the Kruskal-Wallis test, Breslow's method and the Bonferroni method in Table 5.
Table 1. Data and $\lambda_i$ for sorted Monte Carlo data.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>True $\lambda_i$</td>
<td>12</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>$n_i$</td>
<td>15</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>Sorted Uncensored Data</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| | 1.6162 | 0.3113 | 1.0977 |
| | 3.3108 | 0.3193 | 1.5408 |
| | 6.4986 | 0.5793 | 1.7251 |
| | 8.7001 | 0.9393 | 2.1613 |
| | 9.5813 | 1.4585 | 2.8869 |
| | 10.3311 | 2.6174 | 3.0430 |
| | 11.3729 | 2.8732 | 3.1428 |
| | 11.6698 | 6.5377 | 3.5553 |
| | 14.7173 | 12.9082 | 3.9252 |
| | 20.1485 | 20.1756 | 4.2688 |
| | 28.3848 | 22.2597 | 4.3705 |
| | 32.3451 | 38.9166 | 6.9837 |
| | 34.7911 | | 7.7971 |
| | 51.9654 | | 7.9731 |
| | 66.7175 | | 8.2069 |
| | | | 9.2274 |
| | | | 11.2849 |
| | | | 20.2483 |

$X$ (Uncensored) | 20.81 | 9.158 | 5.7466 |

$\delta_i = \sum_{j=1}^{n_i} \delta_{ij}$ | 9 | 9 | 17 |

$X_i = \sum_{j=1}^{n_i} \lambda_{ij}$ for data censored at 20 | 197.798 | 88.544 | 103.191 |

$\lambda_i = X_i / \delta_i$. | 21.978 | 9.838 | 6.07 |
Table 2. The results of example 1 using F test for censored data in Table 1.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>$G_1 = G_2$</th>
<th>$G_1 = G_3$</th>
<th>$G_2 = G_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R = \hat{\lambda}_i / \hat{\lambda}_j$</td>
<td>2.234\textsuperscript{a}</td>
<td>3.621\textsuperscript{a}</td>
<td>1.621</td>
</tr>
<tr>
<td>Degrees of Freedom</td>
<td>18, 18</td>
<td>18, 34</td>
<td>18, 34</td>
</tr>
<tr>
<td>$P$ Value\textsuperscript{b}</td>
<td>0.04845</td>
<td>0.00060</td>
<td>0.11002</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Significant at $\alpha = 0.05$.

\textsuperscript{b}Run STATPAC/DIST to get the probability of an $F$ value larger than observed when the degrees of freedom are $2\hat{\alpha}_i$ and $2\hat{\alpha}_j$, respectively.
Table 3. The ranks for data censored at 20 from Table 1.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>$G_1 = G_2$</th>
<th>$G_1 = G_3$</th>
<th>$G_2 = G_3$</th>
<th>$G_1 = G_2 = G_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups</td>
<td>$G_1$</td>
<td>$G_2$</td>
<td>$G_1$</td>
<td>$G_3$</td>
</tr>
<tr>
<td>$R(X_{ij})$</td>
<td>6 1</td>
<td>3 1</td>
<td>1 5</td>
<td>8 1</td>
</tr>
<tr>
<td></td>
<td>9 2</td>
<td>9 2</td>
<td>2 7</td>
<td>16 2</td>
</tr>
<tr>
<td></td>
<td>10 3</td>
<td>14 4</td>
<td>3 8</td>
<td>21 3</td>
</tr>
<tr>
<td></td>
<td>12 4</td>
<td>19 5</td>
<td>4 9</td>
<td>27 4</td>
</tr>
<tr>
<td></td>
<td>13 5</td>
<td>21 6</td>
<td>6 12</td>
<td>29 6</td>
</tr>
<tr>
<td></td>
<td>14 7</td>
<td>22 7</td>
<td>10 13</td>
<td>30 11</td>
</tr>
<tr>
<td></td>
<td>15 8</td>
<td>24 8</td>
<td>11 14</td>
<td>32 12</td>
</tr>
<tr>
<td></td>
<td>16 11</td>
<td>25 10</td>
<td>19 15</td>
<td>33 22</td>
</tr>
<tr>
<td></td>
<td>18 17</td>
<td>26 11</td>
<td>26 6</td>
<td>35 34</td>
</tr>
<tr>
<td></td>
<td>23 23</td>
<td>30 12</td>
<td>28.5 17</td>
<td>40.5 19</td>
</tr>
<tr>
<td></td>
<td>23 23</td>
<td>30 13</td>
<td>28.5 18</td>
<td>40.5 20</td>
</tr>
<tr>
<td></td>
<td>23 23</td>
<td>30 15</td>
<td>28.5 20</td>
<td>40.5 23</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>30 16</td>
<td>21</td>
<td>40.5</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>30 17</td>
<td>22</td>
<td>40.5</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>30 18</td>
<td>23</td>
<td>40.5</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>20 24</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>23 25</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>30 28.5</td>
<td>40.5</td>
<td>40.5</td>
</tr>
</tbody>
</table>

$R_i = \frac{n_i}{\sum_{j=1}^{n_i} R(X_{ij})}$

| $R_i$ | 251 | 127 | 343 | 218 | 167.5 | 297.5 | 474 | 216.5 | 344.5 |
Table 4. The results of example 1 using Kruskal-Wallis test and the ranks given in Table 3.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>$G_1 = G_2$</th>
<th>$G_1 = G_3$</th>
<th>$G_2 = G_3$</th>
<th>$G_1 = G_2 = G_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>4.0024</td>
<td>10.1229</td>
<td>0.6134</td>
<td>9.70</td>
</tr>
<tr>
<td>T'</td>
<td>4.1546$^a$</td>
<td>10.2185$^a$</td>
<td>0.6148</td>
<td>9.8066$^a$</td>
</tr>
<tr>
<td>P Value$^b$</td>
<td>0.04152</td>
<td>0.00139</td>
<td>0.43299</td>
<td>0.00742</td>
</tr>
</tbody>
</table>

$^a$Significant at $\alpha = .05$.

$^b$Probability of a $X^2$ value larger than observed.
Table 5. Statistics S**, S* and S, and the results of example 1 for F test, Kruskal-Wallis test, Breslow's method and Bonferroni method.

<table>
<thead>
<tr>
<th>Method and Bonferroni</th>
<th>Hypothesis</th>
<th>G₁ = G₂</th>
<th>G₁ = G₃</th>
<th>G₂ = G₃</th>
<th>G₁ = G₂ = G₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breslow and Bonferroni</td>
<td>S**</td>
<td>4.1506</td>
<td>10.4296</td>
<td>0.6338</td>
<td>9.9127</td>
</tr>
<tr>
<td></td>
<td>S*</td>
<td>4.5410</td>
<td>11.0086a</td>
<td>0.6679</td>
<td>10.3524</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>5.0971a</td>
<td>11.7467b</td>
<td>0.7309</td>
<td>9.7876a</td>
</tr>
<tr>
<td>Kruskal-Wallis and</td>
<td>T'</td>
<td>4.1546a</td>
<td>10.2189b</td>
<td>0.6148</td>
<td>9.8066a</td>
</tr>
<tr>
<td>Bonferroni</td>
<td>X₂</td>
<td>3.84</td>
<td>3.84</td>
<td>3.84</td>
<td>5.991</td>
</tr>
<tr>
<td>F Test</td>
<td>R</td>
<td>2.234c</td>
<td>3.621c</td>
<td>1.621</td>
<td></td>
</tr>
</tbody>
</table>

P Values

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>0.02397</td>
</tr>
<tr>
<td>T'</td>
<td>0.04152</td>
</tr>
<tr>
<td>R</td>
<td>0.04845</td>
</tr>
</tbody>
</table>

\(a\) Significant for chi-square value with \(\alpha = .05\).

\(b\) Significant for chi-square value with \(\alpha = .05/3 = .01667\).

\(c\) Significant for F value with \(\alpha = .05\).
From Table 5, it is easy to see that the statistic \( S^{**} \) is a lower bound to \( S^* \) and it also is a lower bound to \( S \) except for the case \( G_1 = G_2 = G_3 \). \( T' \), the tie-corrected Kruskal-Wallis test statistic, is very close to \( S^{**} \) in this example. To allow easy comparison of the \( S \), \( T' \) and \( R \) test results, p-values were obtained by running STATPAC/DIST to get the probability of a large chi-square or F statistic. For this data set the statistic \( R \) always got the smallest probability except for the test of \( G_1 = G_3 \). This means that when we test the null hypothesis of equality of two groups, the \( R \) value is possibly too liberal, i.e., too easy to reject. The probability of \( S \) and \( T' \) listed in Table 5 show that the p-value for \( S \) is always smaller than the p-value of \( T' \) except the case \( G_1 = G_2 = G_3 \). In this example, the three statistics \( S \), \( T' \) and \( R \) yielded the same conclusions, i.e., the "significant" difference between group 1 and groups 2 and 3. These results are somewhat surprising since there is a relatively small difference between the \( \lambda_i \)'s for groups 1 and 2. Since this was Monte Carlo data, these differences may be ascribed to chance. Further Monte Carlo work would undoubtedly tend to "smooth" these unexpected differences.

**Example 2**

The second example is a nutrition experiment conducted by Susan Collinge who was a graduate student in Nutrition Food Science Department, USU, in 1981. Susan looked at how many samples of meat products were bad each day when they were put in 27°C (80.6°F) temperature room. Each of the nine treatments contained twenty-five sealed bags of meat with different chemical additives. During each day a count of the number of swollen bags was made. The swelling indicated spoilage
of the contents. After 100 days the experiment was terminated, resulting in some treatments having censored data. The results of the nine treatments are compared below to see what kind of chemical combination added to the meat will keep the meat from spoilage for the longest period of time. The nine treatments were:

Treatment 1. Control - no chemicals
Treatment 2. Nitrite only
Treatment 3. Nitrite + 20 ppm FeCl₃
Treatment 4. Nitrite + Myoglobin
Treatment 5. Nitrite + 200 ppm EDTA + Myoglobin
Treatment 6. Nitrite + Nytrosylmyoglobin
Treatment 7. Nitrite + 200 ppm EDTA
Treatment 8. Nitrite + 200 ppm EDTA + 20 ppm FeCl₃
Treatment 9. Nitrite + 200 ppm EDTA + 40 ppm FeCl₃

We are interested in the following specific comparisons:

1. Treatment 1 vs. treatment 2 through 9.
2. Treatment 2 vs. treatment 3 through 9.
3. Treatment 3 vs. treatment 4.
4. Treatment 3 vs. treatment 8.
5. Treatment 4 vs. treatment 6.
6. Treatment 5 vs. treatment 7.
7. Treatment 7 vs. treatment 8.
8. Treatment 8 vs. treatment 9.

Table 6 gives the values $n_i$, $\delta_i$, $X_i$, and $\lambda_i$, and the results of the F test are listed in Table 7. The $R_i$'s and results of the Kruskal-Wallis test are shown in Table 8. Table 9 shows all the results of the
F test, the Kruskal-Wallis test, Breslow's method and the Bonferroni method. From Table 9 we find the different methods yield the same results, i.e., only treatment 3 and treatment 4 are homogeneous.
Table 6. The values of $n_i$, $\xi_i$, $X_i$, and $\lambda_i$ of each treatment.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>$T_1$</th>
<th>$T_2$</th>
<th>$T_3$</th>
<th>$T_4$</th>
<th>$T_5$</th>
<th>$T_6$</th>
<th>$T_7$</th>
<th>$T_8$</th>
<th>$T_2 - T_3$</th>
<th>$T_3 - T_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\xi_i$</td>
<td>25</td>
<td>25</td>
<td>23</td>
<td>25</td>
<td>14</td>
<td>13</td>
<td>11</td>
<td>11</td>
<td>113</td>
<td>88</td>
</tr>
<tr>
<td>$X_i$</td>
<td>100</td>
<td>559</td>
<td>587</td>
<td>473</td>
<td>1457</td>
<td>1893</td>
<td>2424</td>
<td>2424</td>
<td>2421</td>
<td>11620</td>
</tr>
<tr>
<td>$\lambda_i = \frac{X_i}{\xi_i}$</td>
<td>4.0</td>
<td>22.36</td>
<td>25.52</td>
<td>18.92</td>
<td>104.071</td>
<td>145.615</td>
<td>2424</td>
<td>164.182</td>
<td>2421</td>
<td>102.83</td>
</tr>
</tbody>
</table>

Table 7. The results for example 2 using $F$ test.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>$T_1$ vs $T_2 - T_3$</th>
<th>$T_2$ vs $T_4 - T_5$</th>
<th>$T_3$ vs $T_4$</th>
<th>$T_5$ vs $T_6$</th>
<th>$T_6$ vs $T_7$</th>
<th>$T_7$ vs $T_8$</th>
<th>$T_8$ vs $T_9$</th>
<th>$T_9$ vs $T_9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R = \frac{\lambda_i}{\lambda_j}$</td>
<td>25.7075$^a$</td>
<td>5.621$^d$</td>
<td>1.3488$^a$</td>
<td>6.433$^a$</td>
<td>7.696$^a$</td>
<td>23.222$^a$</td>
<td>14.764$^a$</td>
<td>14.746$^a$</td>
</tr>
<tr>
<td>Degrees of Freedom</td>
<td>226,50</td>
<td>175,50</td>
<td>46,50</td>
<td>22,46</td>
<td>26,50</td>
<td>2,28</td>
<td>2,22</td>
<td>2,22</td>
</tr>
<tr>
<td>P Values$^b$</td>
<td>&lt; .00001</td>
<td>&lt; .00001</td>
<td>.15046</td>
<td>0.00001</td>
<td>&lt; .00001</td>
<td>0.00001</td>
<td>0.00009</td>
<td>0.00009</td>
</tr>
</tbody>
</table>

$^a$Significant at $\alpha = .05$.

$^b$RUN STATPAC/DIST to get the p values.
Table 8. The results of example 2 using Kruskal-Wallis test.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>$T_1$ vs $T_2$</th>
<th>$T_2$ vs $T_3$</th>
<th>$T_3$ vs $T_4$</th>
<th>$T_4$ vs $T_6$</th>
<th>$T_5$ vs $T_7$</th>
<th>$T_7$ vs $T_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N = \Sigma n_i$</td>
<td>225</td>
<td>200</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>$R_i$</td>
<td>325 25100</td>
<td>1143.5 18956.5</td>
<td>584.5 690.5</td>
<td>399.5 875.5</td>
<td>364.5 910.5</td>
<td>476.5 798.5</td>
</tr>
<tr>
<td>$T'$</td>
<td>70.5512</td>
<td>27.8846</td>
<td>1.0617</td>
<td>22.1017</td>
<td>28.4833</td>
<td>14.8667</td>
</tr>
<tr>
<td>$P$ Values for $T'$</td>
<td>&lt; 0.0001</td>
<td>0.0001</td>
<td>0.3023</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0013</td>
</tr>
</tbody>
</table>

$^a$Significant at $\alpha = 0.05$. 
Table 9. The results of example 2 using F test, Kruskal-Wallis test, Breslow's method and Bonferroni method.

<table>
<thead>
<tr>
<th>Method</th>
<th>Hypothesis</th>
<th>$T_1$ vs $T_2$</th>
<th>$T_2$ vs $T_3$</th>
<th>$T_3$ vs $T_4$</th>
<th>$T_4$ vs $T_5$</th>
<th>$T_5$ vs $T_6$</th>
<th>$T_6$ vs $T_7$</th>
<th>$T_7$ vs $T_8$</th>
<th>$T_8$ vs $T_9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kruskal-</td>
<td>$T'$</td>
<td>70.5512</td>
<td>27.8846</td>
<td>1.0617</td>
<td>22.1017</td>
<td>28.4833</td>
<td>14.8667</td>
<td>10.2335</td>
<td>10.2335</td>
</tr>
<tr>
<td>Wallis and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonferroni</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P Values</td>
<td>$S$</td>
<td>&lt;.00001</td>
<td>&lt;.00001</td>
<td>0.26708</td>
<td>&lt;.00001</td>
<td>&lt;.00001</td>
<td>0.00005</td>
<td>0.00009</td>
<td>0.00095</td>
</tr>
<tr>
<td></td>
<td>$T'$</td>
<td>&lt;.00001</td>
<td>0.00001</td>
<td>0.30283</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00012</td>
<td>0.00138</td>
<td>0.00138</td>
</tr>
<tr>
<td></td>
<td>$R$</td>
<td>&lt;.00001</td>
<td>&lt;.00001</td>
<td>0.15046</td>
<td>&lt;.00001</td>
<td>&lt;.00001</td>
<td>0.00001</td>
<td>0.00009</td>
<td>0.00009</td>
</tr>
</tbody>
</table>

a Means significant for chi-square value with $\alpha = .05$.

b Means significant for chi-square value with $\alpha = .05/6 = .00833$.

c Means significant for F value with $\alpha = .05$. 

26
From Table 9 we see that for this example $S^{**}$ was a lower bound to $S^*$ and $S$. The statistic $T'$ was between $S^{**}$ and $S^*$, sometimes it was close to $S^{**}$ and sometimes close to $S^*$. In this example the $p$ value for $R$ was the smallest value for all of the cases, whereas the $T'$ value of Kruskal-Wallis test had the biggest probability. This suggests that the $R$ test may be too liberal (too easy to reject) while the Kruskal-Wallis test (corrected for ties) may be too conservative. Since the true population values are unknown for this case, it is impossible to say for certain.
CHAPTER V
CONCLUSIONS

A fundamental problem in many biological and medical investigations is a comparison of the survival distributions from two or more samples of censored data. The hypothesis of interest is the equality of survival time distribution functions across samples. In this paper we discussed this topic and analyzed censored data by using four different methods, namely: the $F$ test, the Kruskal-Wallis test, Breslow's generalization of the Kruskal-Wallis test and a Bonferroni multiple comparison method. The $F$ test is restricted to two exponential distributions, so it cannot be used widely. The Kruskal-Wallis test is suitable for two or more populations, but for censored data there will usually be a lot of ties. This violates the assumptions made in developing the Kruskal-Wallis test. For the Bonferroni method we use a given value $\alpha$ to do the $K$ multiple comparisons, then for each single case will only use $1 - \alpha/K$ to test the hypothesis. In this situation the given confidence interval is so large that it is hard to reject the null hypothesis. If a computer is available, we can translate the formulae of Breslow's method to a computer program as given in the Appendix. It will then be easy to analyze censored data. For all the reasons stated above, we prefer to use Breslow's method if a computer is available. If not, the Kruskal-Wallis procedure corrected for ties seemed to give almost the same results. There are only two examples in this paper; if
we want to get more information to tell the exact differences between the Kruskal-Wallis test and Breslow's method, more examples and a computer program for the Kruskal-Wallis test should be developed.

A Monte Carlo study could give enough different situations involving different distributions and different values of the parameters to help decide on the best overall procedure. A more exact multiple comparison procedure could also be developed using the asymptotical distribution of Breslow's vector of $W_i$'s rather than using the Bonferroni inequality. Since the Bonferroni method seems to work fairly well in these examples, further work might not be terribly worthwhile. Further Monte Carlo research could help in the decision on whether to pursue this matter in more detail.
REFERENCES


*THE IMS (INTERANTIONAL MATHEMATICAL ANO STATISTICAL
* LIBRARY CONSISTS OF A SUBSTANTIAL COLLECTION
* OF SUBROUTINES ANO FUNCTIONS SUBPROGRAMS IN THE AREAS OF
* MATHEMATICS AND STATISTICS,

**SET LINEINFO AUTOBIND
**BIND=FROM IMSL=

* THE DATA ARE ENTERED TO THE DATA FILE BY FOLLOWING STEPS:
* 1. Enter K, K IS THE NUMBER OF SAMPLES
* 2. Enter N(1), N(1) IS THE NUMBER OF OBSERVATIONS
* OF THE FIRST SAMPLE
* 3. Enter pairs data x(1,1),d(1,1);x(1,2),d(1,2);
* ...x(1,N(1)),d(1,N(1))
* 4. Enter N(2), N(2) IS THE NUMBER OF OBSERVATIONS
* OF THE SECOND SAMPLE
* 5. Enter pairs data x(2,1),d(2,1);x(2,2),d(2,2);
* ...x(2,N(2)),d(2,N(2))

DIMENSION N(10),X(10,200),D(10,200),SIGE(10,10),
**SIGE=SIGMA(X),X(1,1),D(1,1);
Y(10,10),X(10,200),D(10,200),N(10),N(10),
**N=K
READ(5,/) K
N=K
C**
DENOTE BY XO(I,J) THE TRUE OBSERVATION FOR THE (I)TH
INDIVIDUAL IN THE (J)TH SAMPLE (J=1,...,N(1)),I=1,...,K,
SINCE THIS OBSERVATION MAY BE CENSORED BY A VARIABLE
IT CANNOT ALWAYS BE OBSERVED, RATHER ONE
**C** OBSERVES X(I,J)=MIN(XO(I,J),Z(I,J))
**C** ALONG WITH THE INDICATOR VARIABLE

DIMENSION N(10),X(10,200),D(10,200),SIGE(10,10),
**SIGE=SIGMA(X),X(1,1),D(1,1);
Y(10,10),X(10,200),D(10,200),N(10),N(10),
**N=K
READ(5,/) K
N=K
C**
DENOTE BY XO(I,J) THE TRUE OBSERVATION FOR THE (I)TH
INDIVIDUAL IN THE (J)TH SAMPLE (J=1,...,N(1)),I=1,...,K,
SINCE THIS OBSERVATION MAY BE CENSORED BY A VARIABLE
IT CANNOT ALWAYS BE OBSERVED, RATHER ONE
**C** OBSERVES X(I,J)=MIN(XO(I,J),Z(I,J))
**C** ALONG WITH THE INDICATOR VARIABLE

DIMENSION N(10),X(10,200),D(10,200),SIGE(10,10),
**SIGE=SIGMA(X),X(1,1),D(1,1);
Y(10,10),X(10,200),D(10,200),N(10),N(10),
**N=K
READ(5,/) K
N=K
C**
DENOTE BY XO(I,J) THE TRUE OBSERVATION FOR THE (I)TH
INDIVIDUAL IN THE (J)TH SAMPLE (J=1,...,N(1)),I=1,...,K,
SINCE THIS OBSERVATION MAY BE CENSORED BY A VARIABLE
IT CANNOT ALWAYS BE OBSERVED, RATHER ONE
**C** OBSERVES X(I,J)=MIN(XO(I,J),Z(I,J))
**C** ALONG WITH THE INDICATOR VARIABLE

DIMENSION N(10),X(10,200),D(10,200),SIGE(10,10),
**SIGE=SIGMA(X),X(1,1),D(1,1);
Y(10,10),X(10,200),D(10,200),N(10),N(10),
**N=K
READ(5,/) K
N=K
C**
DENOTE BY XO(I,J) THE TRUE OBSERVATION FOR THE (I)TH
INDIVIDUAL IN THE (J)TH SAMPLE (J=1,...,N(1)),I=1,...,K,
SINCE THIS OBSERVATION MAY BE CENSORED BY A VARIABLE
IT CANNOT ALWAYS BE OBSERVED, RATHER ONE
**C** OBSERVES X(I,J)=MIN(XO(I,J),Z(I,J))
**C** ALONG WITH THE INDICATOR VARIABLE

DIMENSION N(10),X(10,200),D(10,200),SIGE(10,10),
**SIGE=SIGMA(X),X(1,1),D(1,1);
Y(10,10),X(10,200),D(10,200),N(10),N(10),
**N=K
READ(5,/) K
N=K
C**
DENOTE BY XO(I,J) THE TRUE OBSERVATION FOR THE (I)TH
INDIVIDUAL IN THE (J)TH SAMPLE (J=1,...,N(1)),I=1,...,K,
SINCE THIS OBSERVATION MAY BE CENSORED BY A VARIABLE
IT CANNOT ALWAYS BE OBSERVED, RATHER ONE
**C** OBSERVES X(I,J)=MIN(XO(I,J),Z(I,J))
**C** ALONG WITH THE INDICATOR VARIABLE

DIMENSION N(10),X(10,200),D(10,200),SIGE(10,10),
**SIGE=SIGMA(X),X(1,1),D(1,1);
Y(10,10),X(10,200),D(10,200),N(10),N(10),
**N=K
READ(5,/) K
N=K
C**
DENOTE BY XO(I,J) THE TRUE OBSERVATION FOR THE (I)TH
INDIVIDUAL IN THE (J)TH SAMPLE (J=1,...,N(1)),I=1,...,K,
SINCE THIS OBSERVATION MAY BE CENSORED BY A VARIABLE
IT CANNOT ALWAYS BE OBSERVED, RATHER ONE
**C** OBSERVES X(I,J)=MIN(XO(I,J),Z(I,J))
**C** ALONG WITH THE INDICATOR VARIABLE
C** D(I,J)=1 IF X(I,J)=X0(I,J)
C** D(I,J)=0 IF X(I,J)≠X0(I,J)
C** N(I)=# OBSERVATION OF EACH SAMPLE
C** NT=NT+1
C** WRITE(6,101) NT,NT1
C** WRITE(6,100) NT,NT1
C** FORMAT(1X,10I4,1X)
C** N(I)=# OBSERVATION OF ACH SAMPLE
C** NT1=NT-..3
C** WRITE(o,101)
C** WRITE(6,100) NT,NT1
C** FOR HATC1X 1
C** NT=THE TOTAL SAMPLE SIZE=1,10)
C** NT1=THE CUBIC OF TOTAL SAMPLE SIZE=1,110)
C** FOR AT(1X,///,50(1H*),///)
C** FOR HATC1X,///)
C** HOCH IS AN INSL SUBROUTINE WHICH IS USED TO G
C** THE PRO A ILl TY
C** THE FORM IS CALL KMDCH(S ,K,I,PS3,IER)
C** INPUTE VALUE FOR WHzCH THE
C** COMPUTED,
C** INPUTE ~U R OF OGRES OF FREEDOM OF TH
C** WHICH FOLLOWS THE CHI~SQUAR DISTRIBUTION
C** WITH OF DEGREES OF FREEDOM IS LESS THAN OR
C** EQUAL TO CS,
C** IER - ERROR PARAMETER,
C** CALL KMDCH(S3,K1,P33,IER)
C** WRITE(6,102) S3
C** WRITE(6,103) K1,P33,IER
C** WRITE(6,104) S3
C** WRITE(6,105) K1,P3,IER
C** WRITE(6,106) K1,P3,IER
C** WRITE(6,107) K1,P3,IER
C** WRITE(6,108) K1,P3,IER
C** WRITE(6,109) K1,P3,IER
C** WRITE(6,1010) K1,P3,IER
C** WRITE(6,1011) K1,P3,IER
C** WRITE(6,1012) K1,P3,IER
C** WRITE(6,1013) K1,P3,IER
C** WRITE(6,1014) K1,P3,IER
C** WRITE(6,1015) K1,P3,IER
C** WRITE(6,1016) K1,P3,IER
C** WRITE(6,1017) K1,P3,IER
C** WRITE(6,1018) K1,P3,IER
C** WRITE(6,1019) K1,P3,IER
C** WRITE(6,1020) K1,P3,IER
C** WRITE(6,1021) K1,P3,IER
C** WRITE(6,1022) K1,P3,IER
C** WRITE(6,1023) K1,P3,IER
C** WRITE(6,1024) K1,P3,IER
C** WRITE(6,1025) K1,P3,IER
C** WRITE(6,1026) K1,P3,IER
C** WRITE(6,1027) K1,P3,IER
C** WRITE(6,1028) K1,P3,IER
C** WRITE(6,1029) K1,P3,IER
C** WRITE(6,1030) K1,P3,IER
C** WRITE(6,1031) K1,P3,IER
C** WRITE(6,1032) K1,P3,IER
C** WRITE(6,1033) K1,P3,IER
C** WRITE(6,1034) K1,P3,IER
C** WRITE(6,1035) K1,P3,IER
C** WRITE(6,1036) K1,P3,IER
C** WRITE(6,1037) K1,P3,IER
C** WRITE(6,1038) K1,P3,IER
C** WRITE(6,1039) K1,P3,IER
C** WRITE(6,1040) K1,P3,IER
C** WRITE(6,1041) K1,P3,IER
C** WRITE(6,1042) K1,P3,IER
C** WRITE(6,1043) K1,P3,IER
C** WRITE(6,1044) K1,P3,IER
C** WRITE(6,1045) K1,P3,IER
C** WRITE(6,1046) K1,P3,IER
C** WRITE(6,1047) K1,P3,IER
C** WRITE(6,1048) K1,P3,IER
C** WRITE(6,1049) K1,P3,IER
C** WRITE(6,1050) K1,P3,IER
C** WRITE(6,1051) K1,P3,IER
C** WRITE(6,1052) K1,P3,IER
C** WRITE(6,1053) K1,P3,IER
C** WRITE(6,1054) K1,P3,IER
C** WRITE(6,1055) K1,P3,IER
C** WRITE(6,1056) K1,P3,IER
C** WRITE(6,1057) K1,P3,IER
C** WRITE(6,1058) K1,P3,IER
C** WRITE(6,1059) K1,P3,IER
C** WRITE(6,1060) K1,P3,IER
C** WRITE(6,1061) K1,P3,IER
C** WRITE(6,1062) K1,P3,IER
C** WRITE(6,1063) K1,P3,IER
C** WRITE(6,1064) K1,P3,IER
C** WRITE(6,1065) K1,P3,IER
C** WRITE(6,1066) K1,P3,IER
C** WRITE(6,1067) K1,P3,IER
C** WRITE(6,1068) K1,P3,IER
C** WRITE(6,1069) K1,P3,IER
C** WRITE(6,1070) K1,P3,IER
1200 10b FORMAT(1x,///) S** AS A LOWER BOUND TO S, S AND S** WILL BE
1210 * 1 /// ASYMPTOTICALLY EQUIVALENT STATISTICS,**
1220 */// S** IS COMPUTATIONALLY SIMPLER THAN S,**
1230 * /// HOWEVER, ONLY S WILL BE AN ASYMPTOTICALLY**
1240 */// VALID STATISTIC UNDER HYPOTHESIS,**
1250 WRITE(6,101)
1260 IF(FL,T,3) GO TO 500
1270 C***
1280 C** THIS PART USES BONFERRONI MULTIPLE COMPARISON METHOD
1290 C** NPWC=NUMBER OF PAIRWISE COMPARISON
1300 C** ALPHA=CONFIDENT COEFFICIENT
1310 C** INDX(1) AND INDX(2) ARE THE TWO GROUPS WHICH
1320 C** WANT TO COMPARE
1330 C***
1340 READ(5,/) NPWC,ALPHA
1350 IF(300 KK=1,NPWC)
1360 READ(5,/) INDX(1),INDX(2)
1370 WRITE(*,119)
1380 114 FORMAT(1x,50(1x))
1390 WRITE(6,115) INDX(1),INDX(2)
1400 115 FORMAT(1x,///) V S 'I,II', V S 'II, ***///
1410 DO 200 IM=1,2
1420 ND(1)=N(INDX(1))
1430 DO 200 JM=1,N(INDX(1))
1440 ND(JM)=ND(JM)
1450 DX(1,J)=DX(INDX(1),J)
1460 NT=0
1470 DO 250 IM=1,2
1480 NT=NT+ND(INDX(1))
1490 WRITE(6,111)
1500 WRITE(6,102) NT
1510 WRITE(6,103) NT
1520 CALL P1(2,ND,DX,DD)
1530 CALL P2(2,ND,NT1,DD,DX,DD,SIGE,SIG)
1540 CALL P3(2,ND,NT1,DM,DL,DD,S3,S2)
1550 WRITE(6,111)
1560 CALL MCCH(S3,PS3,IER)
1570 PS3=1,PS3
1580 WRITE(6,102) S3
1590 WRITE(6,107) S3,PS3
1600 107 FORMAT(1x,///) S** AS A LOWER BOUND TO S, S AND S** WILL BE
1610 CALL MCCH(S3,PS3,IER)
1620 PS2=1,PS2
1630 WRITE(6,104) S2
1640 WRITE(6,107) S2,PS2
1650 CALL MCCH(TOT,1,PTOT,IER)
1660 PTOT=1,PTOT
1670 TEST=PHA/NPWC
1680 WRITE(6,105) TOT
1690 WRITE(6,107) TOT,PTOT
1700 WRITE(6,125) PTOT,PTOT,ALPHA,NPWC,TEST
1710 125 FORMAT(1x,///) S** AS A LOWER BOUND TO S, S AND S** WILL BE
1720 * /// USING BONFERRONI INEQUALITY,**
1740 * /// (ASSUMES ALPHAS ARE 'F8.o5, 'F8.o5, 'F8.o5, PAIRWISE COMPARISONS),
1760 C***
1770 C*** MCPCH IS AN IMSL SUBROUTINE WHICH IS USED TO GET
1780 C*** THE INVERSE VALUE OF A CHI-SQUARE DISTRIBUTION,
C* THE FORM IS1 CALL MDCH1(P,DF,X,IER)
C* P = INPUT PROBABILITY,
C* DF = INPUT NUMBER OF DEGREES OF FREEDOM,
C* X = OUTPUT CHI-SQUARE VALUE, SUCH THAT A RANDOM
C* VARIABLE, DISTRIBUTED AS CHI-SQUARE WITH DF
C* DEGREES OF FREEDOM, WILL BE LESS THAN OR
C* EQUAL TO X WITH PROBABILITY P,
C* IER = ERROR PARAMETER,
CALL MDCH1(P,1,CHI,IER)
WRITE(6,101)
500 CONTINUE
END
C* K = THE NUMBER OF SAMPLES
C* N(I) = THE NUMBER OF OBSERVATIONS OF THE (I)TH SAMPLE?
C* X(I,J) = THE VECTOR SCORE STATISTIC
C* PSI(X(I,J),DELTA(I,J)) = PSI FUNCTION
SUBROUTINE P1CK,N,w,x, )
THIS SUBROUTINE DEFINE A SCORING FUNCTION PSI, THEN
COMPUTE THE VECTOR w(I),
WE DEFINE THE SCORING FUNCTION PSI FROM EQUATION (3)
OF BRESLOW FOR COMPARING TWO OBSERVATION X(I,J)
AND X(I',J') BY
DIMENSION N(10),w(10),x(10,200),d(10,200)
WRITE(6,140)
DO 110 I=1,N
110 DO 120 J=1,IP
120 IF(DELTA(I,J) == 0) GO TO 10
DO 130 JP=1,IP
T1=DELTA(I,J)+DELTA(I,J)
T2=DELTA(I,J)+DELTA(I,J)
130 DELTA(I,J)=0, DELTA(I,J)=1
IF(T2==1) GO TO 16
IF(T1==0) GO TO 22
IF(T1==1) AND T2==1
DELTA(I,J)=1, DELTA(I,J)=1
16 DELTA(I,J)=0, DELTA(I',J')=1
22 DELTA(I,J)=0, DELTA(I',J')=1
1800 0 IF(X(I,J) < X(I',J')
PSI=1
2000 PSI=2
2100 PSI=3
2200 PSI=4
2300 PSI=5
2400 PSI=6
2500 PSI=7
2600 PSI=8
2700 PSI=9
2800 PSI=10
2900 PSI=11
3000 PSI=12
3100 PSI=13
3200 PSI=14
3300 PSI=15
3400 PSI=16
3500 PSI=17
3600 PSI=18
3700 PSI=19
3800 PSI=20
3900 PSI=21
4000 PSI=22
4100 PSI=23
4200 PSI=24
4300 PSI=25
4400 PSI=26
4500 PSI=27
4600 PSI=28
4700 PSI=29
4800 PSI=30
4900 PSI=31
END
SUBROUTINE P25K,N,NT1,W,X,O,SEG,SIG)

THIS SUBROUTINE COMPUTE THE INDIVIDUAL TERMS

OF A COVARIANCE MATRIX SIGMA

TERMS SIG(I,I), IN COVARIANCE MATRIX SIGMA

MAY BE FOUND FROM THE FORMULAE OF EQUATION (8)

OF BESLHD

I UNEQUAL I

SIG(I,1) = SUM I' UNEQUAL I OF SIG(I,I')

DIMENSION N(10),W(10),X(10,200),D(10,200),SIGE(10,10)

Do 100 I=1,K
Do 210 IP=1,I-1
If (EQQ,IP) Go To 210
100 CONTINUE
210 CONTINUE
$\text{EPS2} = 0$

$\text{DO 240 JP1=1,N(IP)}$

$\text{IF} \left( X(IP,JP) \leq X(IP,JP) \right) \text{GO TO 240}$

$\text{EPS2} = \text{EPS2} + 1$

$\text{CONTINUE}$

$\text{END}$

$\text{DO 250} \text{ I1=1,K}$

$\text{SIG1(I,1)} = \text{SIG1(I,1)}$

$\text{CONTINUE}$

$\text{WRITE}(6,204)$

$\text{WRITE}(6,205) \text{SIG1(J), J=1,K)}$

$\text{RETURN}$

$\text{DIMENSION W(10), W1(10), D(10,1), D(10,2)}, \text{DIMENSION X(10,200), D(10,200)}$

$\text{REAL N(10), NT}$

$\text{DO 310 JP1=1,N(IP)}$

$\text{IF} \left( X(IP,JP) \leq X(IP,JP) \right) \text{GO TO 310}$

$\text{EPS} = \text{EPS} + 1$

$\text{CONTINUE}$

$\text{CONTINUE}$
C* K = THE NUMBER OF SAMPLES
C* SIGMA MATIX HAS RANK K=1 PROVIDED THAT EACH OF
C* THE K SAMPLES CONTAINS AT LEAST ONE UNCENORED
C* OBSERVATION, FOR SUCH SIGMA THERE EXIST K=1 VECTOR
C* X(1)=XX(I),..., XX(I,K)) (I=1,..., K=1) SUCH
C* AS I AND J ARE EQUAL OR UNEQUAL, CONSEQUENTLY THE
C* STATISTICS
C* WILL BE ASYMPTOTICALLY UNCORRELATED WITH MEAN 0
C* AND UNIT VARIANCE, N=(3/2) HAS ASYMPTOTICALLY
C* A MULTIDIMENSIONAL NORMAL DISTRIBUTION, FROM
C* THIS IT FOLLOWS THAT
C* IS SYMMETRIC, K=1 OF (S(I)**2)
C* IS ASYMPTOTICALLY DISTRIBUTED IN A CHI-SQUARED
C* DISTRIBUTION WITH K=1 DEGREES OF FREEDOM, THE
C* STATISTICS S WILL BE USED TO TEST THE HYPOTHESIS.
C* IN PRACTICE, THE X(I) ARE FOUND BY MEANS OF THE
C* GRAM-SCHMIDT ORTHOMOGRAMIZATION PROCESS WITH THE
C* INNER PRODUCT OF TWO VECTORS A AND B BY
C* A.apsulation, THE K=1 VECTORS
C* MAY BE USED AS A STARTING POINT FOR THE GRAM-
C* SCHMIDT PROCESS.
DIMENSION YV(10,10),XX(10,10),SIGE(10,10),X1(10),Y1(10),

REAL NT1

AN=SQRT(NT1)

WRITE(6,470)

470 FORMAT(1X,15X) _GE_ GRAM-SCHmidt ORTHOGONALIZATION,_

= 1 12 1 VECTORS ===

DO 410 I=1,K

410 410 Y1(I)=Y1(I)_=0

CONTINUE

DO 420 J=1,K

420 420 V(J)=VA

CONTINUE

DO 430 L=1,K

430 430 SUM=0

CONTINUE

S(I)=SUM

CONTINUE

WRITE(6,465)(I,I=1,K)

465 FORMAT(1X,15X)  GE_ A X

DO 465 J=1,K

465 XX(I,J)=YY(I,J)/DEN

CONTINUE

DO 470 J=1,K

470 YY(I,J)=Y1(J)*V(J)*XX(I,J)

CONTINUE

DO 475 J=1,K

475 Y2(J)=YY(I,J)

CONTINUE

VA=VAL(K,Y1,Y2,ANS,SIGE)

DEN=SQRT(VA)

SUM=0

CONTINUE

S(I)=SUM

CONTINUE

WRITE(6,460)(I,I=1,K)

460 FORMAT(1X,15X)  GE_ A X

DO 460 J=1,K

460 XX(I,J)=YY(I,J)/DEN

CONTINUE

RETURN

FUNCTION VAL(K,Y1,Y2,ANS,SIGE)

CALCULATE THE INNER PRODUCT OF TWO VECTORS A AND B

BY < A, B > = A^T (SIGMA) B

DIMENSION Y1(10),Y2(10),ANS(10),SIGE(10,10)
48600  486 ANS(M)=ANS(M)+Y1(J)*SIGE(J,M)
48700  VAL=0
48800  DO 470 I=1,K
48900  470 VAL=VAL+ANS(I)*Y2(I)
49000  RETURN
49100  END
49200  C* NT1 = THE CUBIC OF TOTAL OBSERVATIONS
<table>
<thead>
<tr>
<th>Time</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>200</td>
<td>15</td>
</tr>
<tr>
<td>300</td>
<td>1.61e2</td>
</tr>
<tr>
<td>400</td>
<td>3.3101</td>
</tr>
<tr>
<td>500</td>
<td>8.8988</td>
</tr>
<tr>
<td>600</td>
<td>8.7001</td>
</tr>
<tr>
<td>700</td>
<td>9.5813</td>
</tr>
<tr>
<td>800</td>
<td>10.3311</td>
</tr>
<tr>
<td>900</td>
<td>11.3729</td>
</tr>
<tr>
<td>1000</td>
<td>11.6698</td>
</tr>
<tr>
<td>1100</td>
<td>14.7173</td>
</tr>
<tr>
<td>1200</td>
<td>20.0</td>
</tr>
<tr>
<td>1300</td>
<td>20.0</td>
</tr>
<tr>
<td>1400</td>
<td>20.0</td>
</tr>
<tr>
<td>1500</td>
<td>20.0</td>
</tr>
<tr>
<td>1600</td>
<td>20.0</td>
</tr>
<tr>
<td>1700</td>
<td>20.0</td>
</tr>
<tr>
<td>1800</td>
<td>12</td>
</tr>
<tr>
<td>1900</td>
<td>0.3113</td>
</tr>
<tr>
<td>2000</td>
<td>0.3193</td>
</tr>
<tr>
<td>2100</td>
<td>0.5793</td>
</tr>
<tr>
<td>2200</td>
<td>0.9393</td>
</tr>
<tr>
<td>2300</td>
<td>1.5585</td>
</tr>
<tr>
<td>2400</td>
<td>2.6174</td>
</tr>
<tr>
<td>2500</td>
<td>3.8732</td>
</tr>
<tr>
<td>2600</td>
<td>6.5777</td>
</tr>
<tr>
<td>2700</td>
<td>12.9082</td>
</tr>
<tr>
<td>2800</td>
<td>20.0</td>
</tr>
<tr>
<td>2900</td>
<td>20.0</td>
</tr>
<tr>
<td>3000</td>
<td>20.0</td>
</tr>
<tr>
<td>3100</td>
<td>18</td>
</tr>
<tr>
<td>3200</td>
<td>1.0977</td>
</tr>
<tr>
<td>3300</td>
<td>1.5406</td>
</tr>
<tr>
<td>3400</td>
<td>1.7251</td>
</tr>
<tr>
<td>3500</td>
<td>2.0613</td>
</tr>
<tr>
<td>3600</td>
<td>2.8649</td>
</tr>
<tr>
<td>3700</td>
<td>3.0430</td>
</tr>
<tr>
<td>3800</td>
<td>3.1428</td>
</tr>
<tr>
<td>3900</td>
<td>3.3553</td>
</tr>
<tr>
<td>4000</td>
<td>3.9252</td>
</tr>
<tr>
<td>4100</td>
<td>4.2688</td>
</tr>
<tr>
<td>4200</td>
<td>4.3705</td>
</tr>
<tr>
<td>4300</td>
<td>6.7837</td>
</tr>
<tr>
<td>4400</td>
<td>7.7971</td>
</tr>
<tr>
<td>4500</td>
<td>7.9731</td>
</tr>
<tr>
<td>4600</td>
<td>8.2049</td>
</tr>
<tr>
<td>4700</td>
<td>9.2274</td>
</tr>
<tr>
<td>4800</td>
<td>11.2809</td>
</tr>
<tr>
<td>4900</td>
<td>20.0</td>
</tr>
<tr>
<td>5000</td>
<td>5.0,05</td>
</tr>
<tr>
<td>5100</td>
<td>1.2</td>
</tr>
<tr>
<td>5200</td>
<td>1.3</td>
</tr>
<tr>
<td>5300</td>
<td>2.3</td>
</tr>
</tbody>
</table>
NT = THE TOTAL SAMPLE SIZE = 45
NI = THE CUBIC OF TOTAL SAMPLE SIZE = 9125

VECTOR SCORE STATISTIC
W(1) = 258,
W(2) = -119,
W(3) = -139,
The sum of W(i) equal 0.

SIGMA MATRIX
0.0761 -0.0272 -0.0489
-0.0272 0.0504 -0.0232
-0.0489 -0.0232 0.0721

GRAM SCHMIDT ORTHOGONALIZATION VECTORS
1
2
3 0.2437 1.7737
0.0000 0.9573
0.0000 0.0000

TESTING HYPOTHESIS
S** = 9.912700
P( Chi-Square(2) >= 9.912700 | 1 = 0.007039
S** = 10.352355
P( Chi-Square(2) >= 10.352355 | 1 = 0.005650
S = 9.787615
\[ \text{\textsc{plchisquare} (2) = 9,707615 } \Rightarrow 0,007493 \]

**s** as a lower bound to **s**, **s** and **s** will be asymptotically equivalent statistics. **s** is computationally simpler than **s**, however, only **s** will be an asymptotically valid statistic under hypothesis.

**************************************************

**************************************************

----- 1 VS 2 -----

\( n_1 = \text{the total sample size} = 27 \)
\( n_1 = \text{the cubic of total sample size} = 19683 \)

----- vector score statistic -----
\( w(1) = 02, \)
\( w(2) = -02, \)

the sum of \( w(i) \) equal 0.

----- sigma matrix -----
\[ 
\begin{pmatrix} 
0,0670 & -0,0670 \\
-0,0670 & 0,0670 
\end{pmatrix} 
\]

----- gram-schmidt orthogonalization vectors -----
\[ 
1 \\
3,6299 \\
0,00000 
\]

----- testing hypothesis -----
\[ s** = 4,150617 \]
\[ \text{p(chi-square}(1) \Rightarrow 4,150617 = 0,041619 \]
\( P(\text{Chi-Square}(1) \geq 4.540970) = 0.033093 \)

\( S = 5.097801 \)

\( P(\text{Chi-Square}(1) \geq 5.097801) = 0.023956 \)

Using Bonferroni inequality, reject 0.023956 if 0.023956 < 0.05/3 = 0.016667

\( s^* \) as a lower bound to \( S^* \), \( S \) and \( S^* \) will be asymptotically equivalent statistics. \( S^* \) is computationally simpler than \( S \), however, only \( S \) will be an asymptotically valid statistic under hypothesis.

Bonferroni critical value
(Assumes \( \alpha = 0.05 \), 3 pairwise comparisons)

\( \chi^2 \) square for 1DF, \((1 - 0.05/3)^{100(TH)}\)

\( P(\text{Chi-Square}(1) < 5.737029) = 0.983333 \)

--------------------

****

--------------------

---- 1 VS 3 ----

\( n = \) the total sample size = 33
\( n^1 = \) the cubic of total sample size = 35937

---- Vector score statistic ----

\( \mathbf{W}(1) = \begin{pmatrix} 176 \\ \end{pmatrix} \)

\( \mathbf{W}(2) = \begin{pmatrix} -176 \\ \end{pmatrix} \)

The sum of \( \mathbf{W}(i) \) equal 0.

---- Sigma matrix ----

\( \begin{pmatrix} 0.0734 & 0.0734 \\ 0.0734 & 0.0734 & \end{pmatrix} \)
----- GRAM-SCHMIDT ORTHOGONALIZATION VECTORS ----- 

1
3.69161
0.00000

----- TESTING HYPOTHESIS ----- 

S** = 10,429b30
P(Chi-Square(1) >= 10,429630 ) = 0,001240

S** = 11,008594
P(Chi-Square(1) >= 11,008594 ) = 0,000907

S = 11,746682
P(Chi-Square(1) >= 11,746682 ) = 0,000610

By Bonferroni inequality reject 0,000610 if 0,000610 < 0,05/3=0,01667

S** as a lower bound to S*, S and S* will be asymptotically equivalent statistics.
S* is computationally simpler than S, however, only S will be an asymptotically valid statistic under hypothesis.

Bonferroni critical value
(assumes alpha = 0.05, 3 pairwise comparisons)

P(Chi-Square(1) <= 5,737029 ) = 0,98333

*******************************

*******************************

----- 2 VS 3 ----- 

N = THE TOTAL SAMPLE SIZE = 30
N1 = THE CUBIC OF TOTAL SAMPLE SIZE = 27000
VECTOR SCORE STATISTIC

W(1) = -37,
W(2) = 37,
THE SUM OF W(1) EQUAL 0.

SIGMA MATRIX

\[
\begin{pmatrix}
0.0694 & 0.0694 \\
0.0694 & 0.0694
\end{pmatrix}
\]

GRAM-SCHMIDT ORTHOGONALIZATION VECTORS

1
3.79676
0.00000

TESTING HYPOTHESIS

\[
P(\chi^2(1)) = 0.633796 \geq 0.05
\]

\[
P(\chi^2(1)) = 0.667857 \geq 0.913800
\]

\[
P(\chi^2(1)) = 0.730913 \geq 0.99
\]

USING BONFERRONI INEQUALITY

REJECT 0.392587 IF 0.392587 < 0.05/3 = 0.016667

AS A LOWER BOUND TO S*, S AND S* WILL BE
ASYMPTOTICALLY EQUIVALENT STATISTICS.
S* IS COMPUTATIONALLY SIMPLER THAN S,
HOWEVER, ONLY S WILL BE AN ASYMPTOTICALLY
VALID STATISTIC UNDER HYPOTHESIS.

BONFERRONI CRITICAL VALUE

ASSUMES ALPHA = 0.05, 3 PAIRWISE COMPARISONS
\[
\chi^2\text{FOR 1DF, } (1 = 0.05/3) = 10.83
\]

\[
P(\chi^2(1)) = 5.37029 \geq 0.983333
\]
***THE TOTAL SAMPLE SIZE = 225
NT = THE CUBIC OF TOTAL SAMPLE SIZE = 11390625

----- VECTOR SCORE STATISTIC -----
\( W(1) = 5000 \),
\( W(2) = 5000 \),

THE SUM OF \( W(i) \) EQUAL 0.

----- SIGMA MATRIX -----
\[
\begin{pmatrix}
0 & 0.0008 \\
0.0008 & 0.0008
\end{pmatrix}
\]

----- GRAM-SCHMIDT ORTHOGONALIZATION VECTORS -----
\[
\begin{pmatrix}
1 \\
34.61047 \\
0.00000
\end{pmatrix}
\]

----- TESTING HYPOTHESIS -----
\( S** = 64.666667 \)
\( P(\text{CHI-SQUARE}(1) = 64.666667) \approx 0.000000 \)

\( S* = 74,692909 \)
\( P(\text{CHI-SQUARE}(1) = 74,692909) \approx 0.000000 \)

\( S = 2659.574468 \)
\( P(\text{CHI-SQUARE}(1) = 2659.574468) \approx 0.000000 \)

\( S** \) AS A LOWER BOUND TO \( S* \) AND \( S* \) WILL BE
ASYMPTOTICALLY EQUIVALENT STATISTICS,
\( S* \) IS COMPUTATIONALLY SIMPLER THAN \( S* \).
HOWEVER, ONLY S WILL BE AN ASYMPTOTICALLY VALID STATISTIC UNDER HYPOTHESIS.
NT: THE TOTAL SAMPLE SIZE = 200
NT1: THE CUBIC OF TOTAL SAMPLE SIZE = 8000000

VELOCITY SCORE STATISTIC

W(1) = 2738,
W(2) = 2738

THE SUM OF W(I) EQUAL 0,

SIGMA MATRIX

\[
\begin{pmatrix}
0.0189 & 0.0169 \\
0.0189 & 0.0169
\end{pmatrix}
\]

GRAM-SCHMIDT ORTHOGONALIZATION VECTORS

\[
\begin{pmatrix}
1 \\
7.27161 \\
0.00000
\end{pmatrix}
\]

TESTING HYPOTHESIS

\[
P(\text{CHI-SQUARE}(1) \geq 25.70277) = 0.000000
\]

\[
P(\text{CHI-SQUARE}(1) \geq 29.037435) = 0.000000
\]

\[
P(\text{CHI-SQUARE}(1) \geq 49.552139) = 0.000000
\]

**ASA** AS A LOWER BOUND TO S*, S AND S* WILL BE ASYMPTOTICALLY EQUIVALENT STATISTICS, S* IS COMPUTATIONALLY SIMPLER THAN S,
HOWEVER, ONLY S WILL BE AN ASYMPTOTICALLY VALID STATISTIC UNDER HYPOTHESIS.
**THE TOTAL SAMPLE SIZE** = 225  
**NT**  
**THE CUBIC OF TOTAL SAMPLE SIZE** = 11390625

--- VECTOR SCORE STATISTIC ---

\[ W(1) = 5000, \]
\[ W(2) = 2113, \]
\[ W(3) = 1944, \]
\[ W(4) = 1945, \]
\[ W(5) = 1055, \]
\[ W(6) = 1694, \]
\[ W(7) = 3272, \]
\[ W(8) = 1727, \]
\[ W(9) = 3254. \]

**THE SUM OF W(I) EQUAL 0.**

--- SIGMA MATRIX ---

\[
\begin{pmatrix}
0.0008 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 \\
0.0001 & 0.0230 & 0.0025 & 0.0027 & 0.0034 & 0.0035 & 0.0036 & 0.0035 & 0.0036 \\
0.0001 & 0.0025 & 0.0234 & 0.0029 & 0.0035 & 0.0035 & 0.0037 & 0.0035 & 0.0037 \\
0.0001 & 0.0027 & 0.0029 & 0.0251 & 0.0038 & 0.0038 & 0.0039 & 0.0038 & 0.0039 \\
0.0001 & 0.0034 & 0.0035 & 0.0038 & 0.0347 & 0.0058 & 0.0062 & 0.0057 & 0.0062 \\
0.0001 & 0.0035 & 0.0035 & 0.0038 & 0.0058 & 0.0364 & 0.0068 & 0.0062 & 0.0067 \\
0.0001 & 0.0036 & 0.0037 & 0.0039 & 0.0062 & 0.0062 & 0.0382 & 0.0066 & 0.0073 \\
0.0001 & 0.0035 & 0.0035 & 0.0038 & 0.0057 & 0.0062 & 0.0066 & 0.0360 & 0.0066 \\
0.0001 & 0.0036 & 0.0037 & 0.0039 & 0.0062 & 0.0067 & 0.0073 & 0.0066 & 0.0382
\end{pmatrix}
\]

--- GRAM-SCHMIDT ORTHOGONALIZATION VECTORS ---
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>14.81047</td>
<td>0.82469</td>
<td>0.93110</td>
<td>1.01961</td>
<td>1.09893</td>
<td>1.37721</td>
<td>1.94955</td>
<td>3.29697</td>
</tr>
<tr>
<td></td>
<td>0.00000</td>
<td>6.59750</td>
<td>0.72631</td>
<td>0.85743</td>
<td>1.07714</td>
<td>1.36545</td>
<td>1.96068</td>
<td>3.27845</td>
</tr>
<tr>
<td></td>
<td>0.00000</td>
<td>0.00000</td>
<td>6.58051</td>
<td>0.82522</td>
<td>1.08137</td>
<td>1.36211</td>
<td>1.95909</td>
<td>3.28122</td>
</tr>
<tr>
<td></td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>6.41692</td>
<td>1.09559</td>
<td>1.36753</td>
<td>1.95578</td>
<td>3.28701</td>
</tr>
<tr>
<td></td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>5.54332</td>
<td>1.35728</td>
<td>1.97038</td>
<td>3.26213</td>
</tr>
<tr>
<td></td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>5.56225</td>
<td>1.97524</td>
<td>3.25345</td>
</tr>
<tr>
<td></td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>5.77519</td>
<td>3.24642</td>
</tr>
<tr>
<td></td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
</tr>
</tbody>
</table>

----- TESTING HYPOTHESIS -----  

$*** 154,753470$

\[ P(\text{CHI-SQUARE}(8) \geq 154,753470) = 0.000000 \]

$** 173,849062$

\[ P(\text{CHI-SQUARE}(8) \geq 173,849062) = 0.000000 \]

$= 2766,360697$

\[ P(\text{CHI-SQUARE}(8) \geq 2766,360697) = 0.000000 \]

$** AS A LOWER BOUND TO S$, $ S AND S$ WILL BE ASYMPTOTICALLY EQUIVALENT STATISTICS, $ S IS COMPUTATIONALLY SIMPLER THAN S$, HOWEVER, ONLY $S WILL BE AN ASYMPTOTICALLY VALID STATISTIC UNDER HYPOTHESIS.

******************************************************************************

******************************************************************************

----- 3 VS 4 -----  

NTTHE TOTAL SAMPLE SIZE So  
NT1 THE CUBIC OF TOTAL SAMPLE SIZE = 125000
VECTOR SCORE STATISTIC

\[ \begin{align*}
W(1) &= 106, \\
W(2) &= 106,
\end{align*} \]

The sum of \( W(1) \) equals 0.

SIGMA MATRIX

\[
\begin{pmatrix}
0.0730 & 0.0730 \\
0.0730 & 0.0730
\end{pmatrix}
\]

GRAM-SCHMIDT ORTHOGONALIZATION VECTORS

\[
\begin{pmatrix}
1 \\
3.70177 \\
0.00000
\end{pmatrix}
\]

TESTING HYPOTHESIS

\[
\begin{align*}
S^* &= 1.078654 \\
P(\text{Chi-Square}(1) \geq 1.078654) &= 0.240946 \\
S^* &= 1.200299 \\
P(\text{Chi-Square}(1) \geq 1.200299) &= 0.273262 \\
S &= 1.231747 \\
P(\text{Chi-Square}(1) \geq 1.231747) &= 0.267067
\end{align*}
\]

Using Bonferroni inequality, reject 0.267067 if \( 0.05/6 = 0.008333 \).

\( S^* \) as a lower bound to \( S \), \( S \) and \( S^* \) will be asymptotically equivalent statistics, \( S \) is computationally simpler than \( S^* \), however, only \( S \) will be an asymptotically valid statistic under hypothesis.

Bonferroni critical value

(assuming \( \alpha = 0.05 \), 6 pairwise comparisons)

\[
\begin{align*}
\text{Chi-Square for 10f, } (1 - 0.05/6)100(1+m)
\end{align*}
\]

\[
P(\text{Chi-Square}(1) \leq 6.961694) = 0.991667
\]
****THE TOTAL SAMPLE SIZE = 50****

**NT1 = THE CUBIC OF TOTAL SAMPLE SIZE = 125000**

----- VECTOR SCORE STATISTIC ----- 

\[ W(1) = 476, \]
\[ W(2) = 476, \]

THE SUM OF W(I) EQUAL 0.

----- SIGMA MATRIX ----- 

\[
\begin{pmatrix}
0.0634 & 0.0634 \\
0.0634 & 0.0634
\end{pmatrix}
\]

----- GRAM-SCHMIDT ORTHOGONALIZATION VECTORS ----- 

1

3,97251

0,00000

----- TESTING HYPOTHESIS ----- 

**S** = 21,751296

**P(Chi-Square(1) >= 21.751296) = 0.000003**

**S** = 24,494703

**P(Chi-Square(1) >= 24.494703) = 0.000001**

**S** = 28,60449
I

**USING BONFERRONI INEQUALITY**

REJECT $0.000000$ IF $0.000000 < 0.05/6 = 0.008333$

$S^*$ AS A LOWER BOUND TO $S^*$, $S$ AND $S^*$ WILL BE
ASYMPTOTICALLY EQUIVALENT STATISTICS.

$S^*$ IS COMPUTATIONALLY SIMPLER THAN $S$.

HOWEVER, ONLY $S$ WILL BE AN ASYMPTOTICALLY
VALID STATISTIC UNDER HYPOTHESIS.

**BONFERRONI CRITICAL VALUE**

(ASSUMES $\alpha = 0.05$, 6 PAIRWISE COMPARISONS)

$\chi^2 = \chi^2$ FOR $10$, $(1 = 0.05/6)100(\text{TH})%$

$\chi^2 = \chi^2$ FOR $10$, $0.9816941 = 0.991667$

-----------------------------

-----------------------------

----- 4 VS 6 -----}

**NT** THE TOTAL SAMPLE SIZE = 50
**NT** THE CUBIC OF TOTAL SAMPLE SIZE = 125000

----- VECTOR SCORE STATISTIC -----}

$w(1) = 546$

$w(2) = 596$

THE SUM OF $w(i)$ EQUAL 0.

----- SIGMA MATRIX -----}

$0.0622 -0.0622$

$-0.0622 0.0622$

----- GRAM SCHMIDT ORTHOGONALIZATION VECTORS -----}

1

$0.01067$
Testing Hypothesis

\[ S^* = 28.619136 \]

\[ \text{Pr}(\chi^2(1) \geq 28.619136) = 0.000000 \]

\[ S^* = 30.006655 \]

\[ \text{Pr}(\chi^2(1) \geq 30.006655) = 0.000000 \]

\[ S = 38.362630 \]

\[ \text{Pr}(\chi^2(1) \geq 38.362630) = 0.000000 \]

Using Bonferroni inequality, reject 0.000000 if 0.000000 < 0.05/6 = 0.008333

As a lower bound to \( S^* \), \( S \) and \( S^* \) will be asymptotically equivalent statistics. \( S^* \) is computationally simpler than \( S \).

However, only \( S^* \) will be an asymptotically valid statistic under hypothesis.

Bonferroni critical value (assumes \( \alpha = 0.05 \), 6 pairwise comparisons)

\[ \chi^2 = \text{squared for 1 DF, } (1 = 0.05/6) \times 100(\text{TH}) \]

\[ \text{Pr}(\chi^2(1) \leq 6.981694) = 0.991667 \]

-----------------------------

\( \chi^2 \) vs 7-----------------------------

Total sample size = 50

Cubic of total sample size = 125000

Vector score statistic

\( \begin{align*}
\mathbf{w}(1) &= 322, \\
\mathbf{w}(2) &= 322,
\end{align*} \)
THE SUM OF $w(i)$ EQUAL 0.

====== SIGMA MATRIX ======

\[ 0.0497 \quad 0.0497 \]
\[ 0.0497 \quad 0.0497 \]

====== GRAM-SCHMIDT ORTHOGONALIZATION VECTORS ======

1

\[ 0.48688 \]
\[ 0.00000 \]

====== TESTING HYPOTHESIS ======

**S** = 9.953664

$\text{P}[\text{Chi-Square}(1) \geq 9.953664] = 0.001605$

**S** = 10.058234

$\text{P}[\text{Chi-Square}(1) \geq 10.058234] = 0.000061$

**S** = 10.698985

$\text{P}[\text{Chi-Square}(1) \geq 10.698985] = 0.000044$

USING BONFERRONI INEQUALITY
REJECT 0.000044 IF 0.000044 < 0.05/6 = 0.008333

**S** AS A LOWER BOUND TO S*, S AND S* WILL BE ASYMPTOTICALLY EQUIVALENT STATISTICS.
S* IS COMPUTATIONALLY SIMPLER THAN S, HOWEVER, ONLY S WILL BE AN ASYMPTOTICALLY VALID STATISTIC UNDER HYPOTHESIS.

BONFERRONI CRITICAL VALUE
(ASSUMES ALPHA = 0.05, 6 PAIRWISE COMPARISONS)

$\text{CHI}^2\text{-SQUARE FOR 10F, } (1 = 0.05/6)100(\text{TH})$

$\text{P}[\text{Chi-Square}(1) \leq 6.981694] = 0.991667$
-----------

----- 7 VS 8 -----

NT= THE TOTAL SAMPLE SIZE= 50
NT1= THE CUBIC OF TOTAL SAMPLE SIZE= 125000

----- VECTOR SCORE STATISTIC -----

W(1)= 247,
W(2)= -247,
THE SUM OF W(1) EQUAL 0,

----- SIGMA MATRIX -----

0,0447 0,0447
0,0447 0,0447

----- GRAM-SCHMIDT ORTHOGONALIZATION VECTORS -----

1
4,73090
0,00000

----- TESTING HYPOTHESIS -----

S* = 5,856864
P(Chi-Square(1) >= 5,856864) = 0,015516

S = 10,724500
P(Chi-Square(1) >= 10,724500) = 0,001057

S = 10,923724
P(Chi-Square(1) >= 10,923724) = 0,000949
USING BONFERRONI INEQUALITY
REJECT 0,000949 IF 0,000949 < 0,05/6 = 0,008333

S* AS A LOWER BOUND TO S*, S AND S* WILL BE
ASYMPTOTICALLY EQUIVALENT STATISTICS.
S* IS COMPUTATIONALLY SIMpler THAN S,
HOWEVER, ONLY S WILL BE AN ASYMPTOTICALLY
VALID STATISTIC UNDER HYPOTHESIS.

BONFERRONI CRITICAL VALUE
(ASSUMES ALPHA = 0.05, 6 PAIRWISE COMPARISONS)
\( \chi^2 = \text{SQUARE FOR 1DF, } (1 = 0.05/6)100(\text{TH})a \)

\( P(\chi^2 = \text{SQUARE}(1) \leq 0.981694) = 0.991667 \)

***************

***************

----- 8 VS 9 ----- 

\( N^2 = \text{THE TOTAL SAMPLE SIZE} = 50 \)
\( N^1 = \text{THE CUBIC OF TOTAL SAMPLE SIZE} = 125000 \)

----- VECTOR SCORE STATISTIC ----- 
\( w(1) = 247, \)
\( w(2) = 247, \)

THE SUM OF \( w(i) \) EQUAL 0.

----- SIGMA MATRIX ----- 
\( \begin{bmatrix}
0.0447 & 0.0447 \\
0.0447 & 0.0447 \\
\end{bmatrix} 
\)

----- GRAM-SCHMIDT ORTHOGONALIZATION VECTORS ----- 

\( \begin{bmatrix}
1 \\
0.73090 \\
0.00000 \\
\end{bmatrix} 
\)

----- TESTING HYPOTHESIS ----- 
\( S** = 5.856864 \)
$\chi^2$-SQUARE($l$) \geq 10,724500 \quad p = 0.001057

$s = 10,92374$

$\chi^2$-SQUARE($l$) \geq 10,923748 \quad p = 0.000949

Using Bonferroni inequality:

Reject 0.000949 if 0.000949 < 0.05/6 = 0.008333

$s^2$ as a lower bound to $s$, $s$ and $s^2$ will be asymptotically equivalent statistics.
$s^2$ is computationally simpler than $s$.
However, only $s$ will be an asymptotically valid statistic under hypothesis.

Bonferroni critical value (assumes $\alpha = 0.05$, 6 pairwise comparisons):

$\chi^2$-SQUARE for 10F, $1 = 0.05/6 \times 100$(TH)

$\chi^2$-SQUARE($l$) \leq 6,981694 \quad p = 0.991667

*******************************************************************************