Comparison of Bootstrap with Other Tests for Several Distributions

Yu-Yu Wong
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/etd

Part of the Applied Statistics Commons

Recommended Citation
https://digitalcommons.usu.edu/etd/6995

This Thesis 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 Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.
COMPARISON OF BOOTSTRAP WITH OTHER TESTS
FOR SEVERAL DISTRIBUTIONS

by

Yu-Yu Wong

A thesis submitted in partial fulfillment of the requirement for the degree
of
MASTER OF SCIENCE
in
Applied Statistics

UTAH STATE UNIVERSITY
Logan, Utah

1988
ACKNOWLEDGEMENTS

I deeply appreciate Dr. David L. Turner, my major advisor, for his guidance, patience and encouragement through the development of this paper and during my three years at Utah State. I also wish to thank Dr. Donald V. Sisson and Dr. Robert Campbell for their participation on my thesis committee and for their help.

A special thanks goes to my parents who have always given me their encouragement and support in my graduate studies.

Yu-Yu Wong
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>The Purpose of This Paper</td>
<td>1</td>
</tr>
<tr>
<td>Definition of Testing Techniques</td>
<td>1</td>
</tr>
<tr>
<td>QQ Plots for P-values and Empirical Power</td>
<td>3</td>
</tr>
<tr>
<td>II. THE SIMULATION</td>
<td>5</td>
</tr>
<tr>
<td>Simulation Program</td>
<td>5</td>
</tr>
<tr>
<td>Generate Original Data</td>
<td>5</td>
</tr>
<tr>
<td>Procedure for The Simulation</td>
<td>7</td>
</tr>
<tr>
<td>III. RESULTS</td>
<td>8</td>
</tr>
<tr>
<td>Bootstrap Procedure Results</td>
<td>8</td>
</tr>
<tr>
<td>QQ Plots for P-values</td>
<td>10</td>
</tr>
<tr>
<td>Three-way Analysis for P-values</td>
<td>12</td>
</tr>
<tr>
<td>Three-way Analysis for Empirical Power</td>
<td>26</td>
</tr>
<tr>
<td>IV. SUMMARY &amp; CONCLUSION</td>
<td>29</td>
</tr>
<tr>
<td>Conclusion</td>
<td>29</td>
</tr>
<tr>
<td>Further Research</td>
<td>29</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>30</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>31</td>
</tr>
<tr>
<td>Appendix 1: Program Listing</td>
<td>32</td>
</tr>
<tr>
<td>Appendix 2: Data Table</td>
<td>44</td>
</tr>
</tbody>
</table>
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Empirical P-values Using the Bootstrap Method to Test ( H_0 = \mu = 0 ) When Sampling a Standard Normal Distribution for Different Sample Sizes (n) and Resample Times (t)</td>
<td>8</td>
</tr>
<tr>
<td>2. Analysis of Variance for the P-values in Table 1</td>
<td>10</td>
</tr>
<tr>
<td>3. ( \mu ) Values to Make the Indicated Power for n=10, ( \sigma = 1 ) and ( \alpha = 0.05 )</td>
<td>12</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Empirical P-values Using the Bootstrap Method to Test $H_0: \mu=0$ When Sampling a Standard Normal Distribution for Different Sample Sizes ($n$) and Resample Times ($t$)</td>
<td>9</td>
</tr>
<tr>
<td>2. QQ Plots of P-values against a Uniform Distribution for $n=10$, $n=20$ and $n=30$ for a T-test for Each of the Distributions, When $H_0$ Was True</td>
<td>13</td>
</tr>
<tr>
<td>3. QQ Plots of P-values against a Uniform Distribution for $n=10$, $n=20$ and $n=30$ for a Trimmed T-test for Each of the Distributions, When $H_0$ Was True</td>
<td>14</td>
</tr>
<tr>
<td>4. QQ Plots of P-values against a Uniform Distribution for $n=10$, $n=20$ and $n=30$ for a Signed-rank Test for Each of the Distributions, When the $H_0$ Was True</td>
<td>15</td>
</tr>
<tr>
<td>5. QQ Plots of P-values against a Uniform Distribution for $n=10$, $n=20$ and $n=30$ for the Bootstrap Test for Each of the Distributions, When $H_0$ Was True</td>
<td>16</td>
</tr>
<tr>
<td>6. QQ Plots of P-values against a Uniform Distribution for $n=10$, $n=20$ and $n=30$ for the Jackknife Test for Each of the Distributions, When $H_0$ Was True</td>
<td>17</td>
</tr>
<tr>
<td>7. Dotplots of P-values from the Normal Distribution, Sample Size $n=30$, Five Testing Methods, When the $H_0$ Was True</td>
<td>18</td>
</tr>
<tr>
<td>8. Dotplots of P-values from the Five Distributions, Sample Size $n=30$, Using T-test Method, When $H_0$ Was True</td>
<td>19</td>
</tr>
<tr>
<td>9. Plot of the Empirical P-values for the Normal Distribution as a Function of $n$ and the Different $\mu$ Values for Each of the Five Different Testing Procedures</td>
<td>20</td>
</tr>
<tr>
<td>10. Plot of the Distribution by Power by Test 3-way Average P-values for 2 Runs at Each of Value of $\mu$ from Table 3, Sample Size $n=30$</td>
<td>21</td>
</tr>
<tr>
<td>11. Plot of the Test by Power by Distribution 3-way Average P-values for 2 Runs at Each of Value of $\mu$ from Table 3, Sample Size $n=30$</td>
<td>22</td>
</tr>
<tr>
<td>12. Plot of the Distribution by Power by Test 3-way Average Empirical Power Values for 2 Runs at Each of Value of $\mu$ from Table 3, Sample Size $n=30$</td>
<td>23</td>
</tr>
<tr>
<td>13. Plot of the Test by Power by Distribution 3-way Average Empirical Power Values for 2 Runs at Each of Value of $\mu$ from Table 3, Sample Size $n=30$</td>
<td>24</td>
</tr>
</tbody>
</table>
Comparison of Bootstrap with Other Tests for Several Distributions

by

Yu-Yu Wong, Master of Science
Utah State University, 1988

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

This paper discusses results of a computer simulation to investigate several different tests when sampling several distributions. The hypothesis $H_0: \mu = 0$ was tested against $H_0: \mu \neq 0$, using the usual $t$-test, trimmed $t$-test, the Jackknife, the Bootstrap and signed-rank test. The p-values and empirical power show that the Bootstrap is as good as the $t$-test. The Jackknife procedure is too liberal, always obtaining small p-values. The signed-rank is a fairly good test if the data follows the Cauchy Distribution.

(53 pages)
CHAPTER I
INTRODUCTION

The Purpose of This Paper. An alternative to classical hypothesis testing is to use p-values which indicate the degree to which the observed data contradict the null hypothesis. P-values are related to traditional hypothesis testing since the hypothesis is rejected whenever $p \leq \alpha$.

Interest for this paper centers on testing the hypothesis $H_0: \mu = 0$, against $H_a: \mu \neq 0$. A t-test is the standard test for this hypothesis if the data follows a normal distribution. Several alternative tests and distributions are compared in this paper.

Definition of Testing Techniques. The specific techniques to be compared in this paper are as follows. The regular or usual t-test developed by W. S. Gosset (1908) is equivalent to the likelihood ratio test of $H_0: \mu = 0$ against $H_a: \mu \neq 0$, when sampling from a normal distribution. The test statistic is $t = (\bar{x} - \mu_0) \sqrt{n}/s$, which follows the t distribution with $(n-1)$ degrees of freedom. Given the t-value, the p-value for a two-sided alternative hypothesis is calculated as $p[|T| > |\text{observed value of } T|]$. If the null hypothesis is true, the p-values from repeated samples should be uniformly distributed on the interval $(0,1)$. If the hypothesis is not true, the p-values will tend to have values smaller than 0.5.

An alternative to the traditional t-value for use with long tailed distributions due to Tukey and Mclaughlin (1963) is to trim the data by first sorting the data, and then deleting the smallest 5% and the largest 5% of the sorted values. The remaining middle 90% of
the data are by definition the trimmed data set. This trimmed sample is then used to test the hypothesis \( H_0 : \mu = 0 \) versus \( H_a : \mu \neq 0 \), using the \( t \)-statistic and \( p \)-value defined earlier but computed from the mean and standard deviation of the trimmed data.

The Jackknife procedure developed by Quenouille (1949) is a resampling technique where a positive integer \( k \) is selected such that \( n/k \) is an integer. The observations are then randomly divided into \( g \) subgroups each of size \( k \). Set \( d = n-k \). For \( i = 1, \ldots, g \), let \( x_{i1}, \ldots, x_{id} \) denote the \( d \) observations obtained by deleting the \( i \)th subgroup of \( k \) observations. In this report \( k \) was set to 1. This gives \( n \) sets of data each with sample size \( d = (n-1) \) from the original data set. The \( \bar{x}_i, s_i \) and then the \( t \) statistic, \( t_i \) are calculated for each data set. The \( p \)-value for Jackknife sampling is defined to be twice the minimum of the proportion of \( \#( \bar{x}'s < \mu_0 ), \#( \bar{x}'s > \mu_0 ) \).

Efron's (1979) Bootstrap procedure is another resampling technique where samples with replacement are drawn from the empirical distribution function where \( p(x) = 1/n \). For this paper 100 samples were drawn each with sample size equal to the original sample size. For each of the 100 bootstrap samples, \( \bar{x} \) and \( s \) were calculated and used to compute a \( t \)-value for each of the samples. The bootstrap \( p \)-value is defined to be twice the minimum of the proportion of

\[ \#( \bar{x}'s < \mu_0 ), \#( \bar{x}'s > \mu_0 ) \].

The last method considered in this paper is a signed-rank statistic due to Wilcoxon (1945) which modifies the observed data \( (x_1, \ldots, x_n) \), to \( d_i = (x_i - \mu_0) \), for \( i = 1, \ldots, n \). The absolute differences \( |d_1|, \ldots, |d_n| \), are ranked to get \( R_i \). The final signed rank variables, \( sr_i, i = 1, \ldots, n \), are defined as \( sr_i = -R_i \) if \( d_i < 0 \), \( R_i \) if \( d_i > 0 \). The
signed-rank data were then used to obtain the t statistic and the t distribution with (n-1) degrees of freedom was used to compute a p-value.

The t-test is uniformly most powerful if the data follows a normal distribution. Several other distributions are used in this paper to assess the robustness of the usual t distribution and to help compare the other test procedures. The negative exponential distribution just has one thick tail and is highly skewed to the right. The double exponential distribution is symmetric but has thicker tails than the normal distribution. The Cauchy distribution is a special heavy tailed distribution which has no moments. The uniform distribution has no tails.

QQ Plots for P-values and Empirical Power. Theoretical QQ plots as described in Chambers et. al. (1983) were used to compare the observed p-values with several distributions. A theoretical QQ (quantile-quantile) plot is obtained by plotting the empirical distribution function against the corresponding quantiles of the theoretical distribution. The empirical distribution function is defined to be the distribution obtained by assigning a probability of 1/n to each of the points \( x_1, \ldots, x_n \). This distribution function will be denoted by \( F_n(x) \). If \( f_X \) denotes the number of sample values that are less than or equal to \( x \), then \( F_n(x) = f_X/n \), so that \( F_n(x) \) gives the relative frequency of the event \( X \leq x \). If the theoretical distribution is a close approximation to the empirical distribution then the points on the plot will fall near the line \( y=x \). Random fluctuations in any particular data set will cause the points to drift away from the line, but if the theoretical distribution is
correct, the points will remain reasonably close to the line.

One other measure of the performance of these procedures was to examine the empirical power. Since the data was simulated, the null hypothesis was known to be true or false. If the null hypothesis was false, then the empirical power was computed as the proportion of runs resulting in rejections using $\alpha = 0.05$. For the normal distribution the theoretical power can be computed as a benchmark for comparing the different tests.
CHAPTER II
THE SIMULATION

Simulation Program. All the computations in this report were performed using programs written in Pascal for the VAX 8650 computer at Utah State University. Uniform random variables were generated using a Pascal program to generate uniform random numbers using the portable random number generator given by Wichmann and Hill (1987). This program generates very long cycles of uniform pseudo-random number sequences with very good statistical properties. In addition, it is 'portable' meaning it may be used on virtually any computer with a Pascal compiler.

Generate Original Data. To generate the normal distribution data, pairs of independent uniform(0,1) random numbers were generated and transformed using the inverse transform technique which uses the polar coordinate transformation attributed to Box and Muller (1958) to generate two independent standard normal random numbers. If \( u_1 \) and \( u_2 \) are independent uniform random numbers, then two standard normal random numbers may be generated as

\[
Z_1 = \left[ -2 \ln(u_1) \right]^{1/2} \cos(\pi u_2) \\
Z_2 = \left[ -2 \ln(u_1) \right]^{1/2} \sin(\pi u_2)
\]

For the other distributions, since explicit formulas for the distribution functions exist, the inverse transform technique was utilized. If the cdf, \( F(x) \), has a simple form such that its inverse, \( F^{-1} \), can be explicitly written, and if \( u \) is a uniform random variable then \( x = F^{-1}(u) \) follows the \( F(.) \) distribution.
The negative exponential distribution has density function \( f(x) = \theta e^{-\theta x} \) and distribution function \( F(x) = 1 - e^{-\theta x} \), if \( x > 0 \). If \( u \) has a uniform distribution over the interval \((0,1)\), then by the inverse transform method, \( x = F^{-1}(u) = -\ln(1-u)/\theta \). In order to get the mean equal to zero and the variance equal to one, \( \theta \) is set equal to 1 and \((x-1)\) is used.

The double exponential distribution has a density function of \( f(x) = 1/2/\beta \exp(-|x|/\beta), -\infty < x < \infty \). For this distribution the mean is equal to zero, the variance is equal to \( 2\beta^2 \) and the distribution function is given by \( F(x) = (1 - e^{-|x|/\beta})/2, -\infty < x < \infty \). Double exponential variates are generated from \( u \), a uniform random variate between \((0, 1)\) as follows. If \( u < 0.5 \) then \( x = F^{-1}(u) = \beta \ln(2u) \), if \( u = 0.5 \) then \( x = 0 \), and if \( u > 0.5, x = F^{-1}(u) = -\beta \ln(2(1-u)) \). In order to get the mean equal to zero and the variance equal to one, take \( \beta = (1/2)^{1/2} \).

The Cauchy distribution has density \( f(x) = \beta/\pi(\beta^2+(x-\alpha)^2) \), where \( \beta > 0, -\infty < x < \infty \) and cumulative distribution function \( F(x) = [\arctan((x-\alpha)/\beta+\pi/2)]/\pi \). By the inverse method if \( u \) is uniform\((0,1)\), then if \( x = \beta \tan((u*\pi)-(\pi/2)) + \alpha, 0 < u < 1 \), then \( x \) will follow the Cauchy distribution. The mean and variance of this distribution do not exist, yet it is symmetric about its median \( \alpha \), and the interquartile range is \( \beta \). In order to compare with 1.35 which is the interquartile range for the standard normal distribution, \( \beta \) is set equal to 0.675.

The uniform distribution, \( f(x) = 1/(b-a), a < x < b \), has mean equal to \( (b-a)/2 \), variance equal to \( (b-a)^2/12 \) and a distribution function of \( F(x) = u = (x-a)/(b-a) \). To get the mean equal to zero and the variance equal to 1, the uniform random numbers should be distributed
on the interval \((-\sqrt{3}, \sqrt{3})\). If \(u\) is distributed as a uniform random variable on \((0,1)\) then \(x = (2u/\sqrt{3}) - \sqrt{3}\) will be uniformly distributed on the interval \((-\sqrt{3}, \sqrt{3})\) with mean zero and variance one.

**Procedure for The Simulation.** For a given run a value of \(\mu\) was read in along with the number of trials to be run, the sample size and the number of bootstrap samples to take. Random samples from each of the five different distributions were then generated and each of the five different tests were performed. For this paper, each combination was repeated 100 times and the empirical power or the proportion of rejects for \(\alpha=0.05\) as well as the average p-values were computed. These summary values provide the basis of the results discussed in the next chapter.
CHAPTER III
RESULTS

Bootstrap Procedure Results. As a preliminary step, the runs presented in Table 1 were made to decide on the number of resamples to take for the bootstrap procedure.

Table 1. Empirical P-values Using the Bootstrap Method to Test \( H_0: \mu = 0 \) When Sampling a Standard Normal Distribution for Different Sample Sizes (n) and Resample Times (t).

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>n = 10</th>
<th>n = 20</th>
<th>n = 30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resample Times</td>
<td>mean</td>
<td>std.</td>
<td>mean</td>
</tr>
<tr>
<td>t = 50</td>
<td>0.3640</td>
<td>0.2282</td>
<td>0.5720</td>
</tr>
<tr>
<td>t = 100</td>
<td>0.6040</td>
<td>0.3050</td>
<td>0.5380</td>
</tr>
<tr>
<td>t = 150</td>
<td>0.4720</td>
<td>0.2338</td>
<td>0.4573</td>
</tr>
<tr>
<td>t = 200</td>
<td>0.6300</td>
<td>0.2807</td>
<td>0.4740</td>
</tr>
<tr>
<td>t = 250</td>
<td>0.5512</td>
<td>0.3193</td>
<td>0.5216</td>
</tr>
<tr>
<td>t = 300</td>
<td>0.4960</td>
<td>0.3402</td>
<td>0.4673</td>
</tr>
<tr>
<td>Total</td>
<td>0.51953</td>
<td>0.50503</td>
<td>0.44910</td>
</tr>
<tr>
<td></td>
<td>0.09730</td>
<td>0.04585</td>
<td>0.07311</td>
</tr>
</tbody>
</table>

Samples of 10, 20 and 30 were generated from a standard normal distribution. These samples were then resampled 50, 100, 150, 200, 250 and 300 times. As displayed in Figure 1, the average p-values for the bootstrap bounce around 0.50 as the initial sample size increases. The number of resamples does not seem to have much effect. A two factor analysis of variance also confirmed this as shown in Table 2. There is no significant difference among the three initial
sample sizes, no significant difference among the resamples and no significant interaction.

Figure 1. Empirical P-values Using the Bootstrap Method to Test $H_0: \mu=0$ When Sampling a Standard Normal Distribution for Different Sample Sizes (n) and Resample Times (t).
Table 2. Analysis of Variance for the P-values in Table 1.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>ss</th>
<th>ms</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resample times</td>
<td>5</td>
<td>0.2090</td>
<td>0.0418</td>
<td>0.97532</td>
</tr>
<tr>
<td>Sample size</td>
<td>2</td>
<td>0.1660</td>
<td>0.0830</td>
<td>0.49117</td>
</tr>
<tr>
<td>Interaction</td>
<td>10</td>
<td>0.6367</td>
<td>0.0637</td>
<td>0.74853</td>
</tr>
<tr>
<td>Error</td>
<td>162</td>
<td>13.7904</td>
<td>0.0851</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>179</td>
<td>14.8021</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* F(2,162 ; 0.05)=3.68

As a result, 100 bootstrap resamples were judged as adequate for this study. All runs reported in this paper, therefore used 100 bootstrap resamples.

**QQ Plots for P-values.** QQ plots were done to check on the distribution of the p-values. Figures 2 through 6 show these plots for samples of sizes 10, 20 and 30 for each of the 5 testing methods and each of the different distributions when the null hypothesis was true. If the null hypothesis is true, the p-values for the regular t-test, the trimmed t-test, the signed-rank test and Bootstrap test perform very well, i.e. the QQ plots are reasonably close to a 45° line which indicates that the p-values for these tests are reasonably close to the expected uniform distribution. The shape of the QQ plots together with histograms of the p-values in figures 7 and 8, for the Jackknife method, indicates the p-values for this procedure do not follow the uniform distribution very well. Instead it appears a spike on zero point and has a heavy tail to the right. This indicates that the p-values do not follow a uniform distribution but instead looks like an exponential distribution shape as seen in the dotplots plotted in Figure 7 and 8. However the p-values from the Cauchy
distribution for the t-test are slightly away from the 45° line. This indicates that the t-test, the trimmed-t test, Bootstrap test and signed-rank test are fairly robust to normality departures, at least for the distributions used here.

From Figures 2 through 6, the QQ plots for the t-test, all distributions follow a straight line except for the Cauchy distribution. For the trimmed t-test, the lines from the exponential and Cauchy distributions are not too close to the 45° line, but p-values for the Cauchy distribution are closer to the 45° line for the trimmed t-test than for the regular t-test. For the signed-rank test, the p-values from the symmetric distributions all show a good shape but the negative exponential distribution does not do too well.

The Bootstrap test shows the desired 45° line for all five distributions for each of the different sample sizes. This indicates that the p-values all follow the expected uniform distribution. The Jackknife test produces p-values for each of these five distributions which tend to be smaller values than expected. In Figure 7, almost 50% of Jackknife p-values are seen to be equal to zero.

As shown in Figures 2-6, the QQ plots for the three different sample sizes are very similar. This is also seen in Figure 9, which plots the p-values against the different values of \( \mu \) which reflect the changing power values. This is done for each of the five tests for the normal distribution samples. The sample sizes used in this study do not seem to have much influence on the p-values. As a consequence, all the following graphs and discussion have been taken for the fixed sample size of 30.
Table 3. $\mu$ Values to Make the Indicated Power for $n=10$, $\sigma=1$ and $\alpha=0.05$.

<table>
<thead>
<tr>
<th>power $(1-\beta)$</th>
<th>$\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.00000</td>
</tr>
<tr>
<td>0.10</td>
<td>0.26400</td>
</tr>
<tr>
<td>0.25</td>
<td>0.49095</td>
</tr>
<tr>
<td>0.50</td>
<td>0.71536</td>
</tr>
<tr>
<td>0.75</td>
<td>0.93758</td>
</tr>
<tr>
<td>0.90</td>
<td>1.15271</td>
</tr>
<tr>
<td>0.95</td>
<td>1.29504</td>
</tr>
</tbody>
</table>

Three-way Analysis for P-values. Figures 10 and 11 plot the distribution by test by true $\mu$ value three way cell means for p-values computed on 2 runs for each value of the $\mu$'s presented in Table 3. From Figure 10, the p-values for the data from a normal distribution for the t-test, the trimmed t, the signed-rank and Bootstrap tests are virtually identical. The Jackknife is much more liberal for normal samples, with p-values much smaller than expected, even when the null hypothesis is true.
Figure 2. QQ Plots of P-values against a Uniform Distribution for \( n=10 \), \( n=20 \) and \( n=30 \) for a T-test for Each of the Distributions, When Ho Was True.
Figure 3. QQ Plots of P-values against a Uniform Distribution for n=10, n=20 and n=30 for a Trimmed T-test for Each of the Distributions, When H₀ Was True.
Figure 4. QQ Plots of P-values against a Uniform Distribution for n=10, n=20 and n=30 for a Signed-rank Test for Each of the Distributions, When the Ho Was True.
Figure 5. QQ Plots of P-values against a Uniform Distribution for n=10, n=20 and n=30 for the Bootstrap Test for Each of the Distributions, When H₀ Was True.
Figure 6. QQ Plots of P-values against a Uniform Distribution for n=10, n=20 and n=30 for the Jackknife Test for Each of the Distributions, When H0 Was True.
Figure 7. Dotplots of P-values from the Normal Distribution, Sample Size n=30, Five Testing Methods, When the Ho Was True.
Figure 8. Dotplots of P-values from the Five Distributions, Sample Size \( n=30 \), Using T-test Method, When \( H_0 \) Was True.
Figure 9. Plot of the Empirical P-values for the Normal Distribution as a Function of n and the Different $\mu$ Values for Each of the Five Different Testing Procedures.
Figure 10. Plot of the Distribution by Power by Test 3-way Average P-values for 2 Runs at Each of Value of \( \mu \) from Table 3, Sample Size \( n=30 \).
Figure 11. Plot of the Test by Power by Distribution 3-way Average P-values for 2 Runs at Each of Value of $\mu$ from Table 3, Sample Size $n=30$. 

- Normal Distribution
- Double exponential Distribution
- Uniform Distribution
- Exponential Distribution
- Cauchy Distribution
Figure 12. Plot of the Distribution by Power by Test 3-way Average Empirical Power Values for 2 Runs at Each of Value of $\mu$ from Table 3, Sample Size $n=30$. 

- Normal Distribution
- Double Exponential Distribution
- Uniform Distribution
- Negative Exponential Distribution
- Cauchy Distribution

Symbols:
- $t$-test
- Trimmed $t$
- Jackknife
- Bootstrap
- Signed-rank
Figure 13. Plot of the Test by Power by Distribution 3-way Average Empirical Power Values for 2 Runs at Each of Value of $\mu$ from Table 3, Sample Size $n=30$. 

- Normal Distribution
- Double Distribution
- Uniform Distribution
- Exponential Distribution
- Cauchy Distribution
For the double exponential distribution, the overall plot is very much like that for the normal distribution. The jackknife is again too liberal with p-values even smaller than the t-test's too small Jackknife p-values.

For the uniform distribution, the plot is virtually the same as the normal distribution. The Jackknife procedure is still too liberal when compared with the other testing methods.

For the negative exponential distribution, the t-test and Bootstrap are fairly close as the $|\mu - \mu_0|$ increases. The p-values for the signed-rank test do not consistently decrease as $|\mu - \mu_0|$ increases. This may be due to the fact that the negative exponential distribution is not symmetric. The Jackknife has about the same pattern as in the double exponential distribution, again having p-values which are far too small.

For the Cauchy distribution, all the p-values are larger than those from the other testing methods. The Jackknife has by far the smallest p-values. The t-test and Bootstrap are not too sensitive to increasing $|\mu - \mu_0|$. The trimmed-t test has a better power curve than t-test and Bootstrap. The signed-rank test has the best p-value curve among these five testing methods, since its p-values are more sensitive to change in $|\mu - \mu_0|$. This is logical since the Cauchy distribution has a median but no mean, and the signed-rank test is a test for medians.

From Figure 11 the p-values for the t-test, for the normal, negative exponential, double exponential and the uniform distribution are all very close indicating that the t-test is very robust to most normality departures. The p-values for the Cauchy distribution are
sightly too liberal if $H_0$ is true. If $H_0$ is false, then the p-values for the Cauchy samples are too large, indicating a very conservative test.

For the trimmed t-test, The plot is very similar to the plot for the t-test. For the Cauchy distribution, the p-values look better than those obtained using the usual t-test. This is likely due to the diminished effect of the very long tails when using the trimmed t-test.

For the Jackknife method, the p-values are all very small even when the hypothesis is true. This procedure has the smallest p-values among all the distribution and testing method combinations used in this paper. The p-values for the Cauchy distribution have a relatively flat response as the true $\mu$ values change, but are still much smaller than they should be.

For the Bootstrap method, the p-values are virtually identical to those obtained using the t-test. The p-values for the Cauchy distribution are again sightly too liberal if $H_0$ is true. If $H_0$ is false, then the p-values for the Cauchy samples are more conservative than the other distributions.

For the signed-rank test, the normal, double exponential and the uniform distribution are again very similar. The p-values from the negative exponential distribution are not consistent as $|\mu - \mu_0|$ increases. The signed-rank test does best among the tests considered when the data follows a Cauchy distribution.

Three-way Analysis for Empirical Power. Figures 12 and 13 plot the distribution by power by test three way cell means for the empirical power, defined to be the proportion of p-values $\leq \alpha = 0.05$. 
From Figure 12, each of the distributions show a typical power curve as the true $\mu$ value moves away from $\mu_0$. For all but the Cauchy distribution, the Jackknife test is very liberal, with the empirical power value always about 0.6 even when the null hypothesis is true. The other four tests have a very similar power curves. The trimmed-t test has slightly higher power than the other tests.

For the uniform distribution, the plot appears to be slightly steeper than the other distributions. The Jackknife test again has the highest power and there is little difference among the other four tests.

For the negative exponential distribution, the empirical power curves show more variability than the other tests. All but the signed-rank test are monotone increasing. The Jackknife again has the highest power.

For the Cauchy distribution, the signed-rank test performs best but is only slightly better than the trimmed-t test. The power curve is steeper than the others. The Jackknife test seems to be 'overpowered' when the null hypothesis is true, and 'underpowered' as the true $\mu$ value moves away from the $\mu_0$.

In Figure 13, the empirical powers for the five distributions are fairly close for the t-test but appear steeper for the other testing methods except for the Cauchy distribution. For the trimmed-t test, the empirical power values are more spread out than the t-test under these five distributions. For the Cauchy distribution, the trimmed t-test is better than the t-test, but still not as good as the signed-rank test. For the Jackknife, the empirical power under those true $\mu$'s are all above 0.5 for these five distributions even when the null
hypothesis is true. The Bootstrap performs remarkably well, and is very similar to the t-test. The signed-rank does best for the Cauchy, and the power curve for each distribution has about the right shape.
Conclusion. The Bootstrap procedure was found to be as good as the t-test. Both are fairly robust to different distributions for testing the hypothesis $H_0: \mu = 0$. The number of resamples and the sample size for the Bootstrap method did not make much difference if the number of resamples is at least 100 and if the original sample size is between 0 and 30. The Jackknife procedure has the smallest p-values among the five testing methods. There is not much difference between the regular t-test, the trimmed t-test or the signed-rank test for the five distributions considered in this paper. The signed-rank test was best when the data followed a Cauchy distribution.

Further Research. In this paper, p-values were calculated for five different testing methods using data generated from five different distributions. Each distribution was scaled so that the means (μ's) or medians were equal to zero and standard deviations were equal to one or the interquartile ranges were equal. For further study, different parameters values as well as different distributions might be tried.

Since the Jackknife performs so poorly here, investigation of its properties should be made for different tests such as oneway analysis of variance, regression, etc. Modification to its implementation might also be appropriate.
REFERENCES


Wilcoxon, Frank (1945): "Individual Comparisons by Ranking Methods," Biometrics 1:80-83. [Proposes the test statistic (3.6).].
Appendix 1: Program Listing

PROGRAM MAIN(INPUT, OUTPUT);
CONST
  PI = 3.1415927;

TYPE ONEDIM = ARRAY [1 .. 200] OF REAL;
VAR
  I, NUMBER, R, TR, K, TIMES, SUM, N, TTIMES : INTEGER;
  D, SS, PC, EXT1, EXT2, EXT3 : VARYING [5] OF CHAR;
  SEED : INTEGER;
  S, CHOOSE, AS : INTEGER;
  PCODE: INTEGER;
  NRAN_FLAG : BOOLEAN;
  FIRSTIME : [STATIC] BOOLEAN := TRUE;
  DF, TVVALUE, N1, N2, U, U1, VA, MED, MEAN, STD, T, ORGTVALUE, TRIM : REAL;
  DATA, SORTD, RAWP, TRIP, JP, BP, RANKP : ONEDIM;
  P, WP, RP: REAL;
  REMP, TEMP, JEMP, BEMP, SREMP : REAL;
  OUTFILE1, OUTFILE2, OUTFILE3 : TEXT;
  FILEN1, FILEN2, FILEN3 : PACKED ARRAY [1 .. 10] OF CHAR;

FUNCTION NOGE: REAL;
  [EXTERNAL, SYNCHRONOUS]
FUNCTION MTH$RANDOM(VAR SEED: INTEGER): REAL; EXTERN;
BEGIN
  IF FIRSTIME THEN
    BEGIN
      SEED := CLOCK * 2 + 1;
      FIRSTIME := FALSE
    END;
  NOGE := MTH$RANDOM(SEED)
END;

FUNCTION RAN: REAL;
VAR
  X, Y, Z, I : INTEGER;

FUNCTION RANDOM: REAL;
VAR
  TEMP: REAL;
BEGIN
  X := 171 * (X MOD 177) - 2 *(X DIV 177);
  IF X < 0 THEN
    X := X + 30269;
  Y := 172 *(Y MOD 176) - 35 *(Y DIV 176);
  IF Y < 0 THEN
    Y := Y + 30307;
  Z := 170 *(Z MOD 178) - 63 *(Z DIV 178);
  IF Z < 0 THEN
    Z := Z + 30323;
  TEMP := X / 30269.0 + Y / 30307.0 + Z / 30323.0;
  RANDOM := TEMP - TRUNC(TEMP);
END;
X := TRUNC (NOGE * 30000) + 1;
Y := TRUNC (NOGE * 30000) + 1;
Z := TRUNC (NOGE * 30000) + 1;
RAN := RANDOM;
END;

FUNCTION NRAN: REAL;
VAR
R1, R2: REAL;
BEGIN
IF (NRAN_FLAG) THEN
BEGIN
R1 := SQRT (-2.0 * LN(RAN));
R2 := RAN;
N1 := R1 * COS (2 * PI * R2);
N2 := R1 * SIN (2 * PI * R2);
NRAN_FLAG := FALSE;
NRAN := N1;
END
ELSE
BEGIN
NRAN_FLAG := TRUE;
NRAN := N2;
END;
END;

PROCEDURE RSORT (NUMBER: INTEGER; DATA: ONEDIM; VAR SORTD: ONEDIM);
VAR K, I, L: INTEGER;
T: REAL;
BEGIN
FOR K := NUMBER - 1 DOWNTO 1 DO
FOR I := 1 TO K DO
IF DATA[I] > DATA[I+1] THEN
BEGIN
T := DATA[I];
DATA[I] := DATA[I+1];
DATA[I+1] := T
END;
FOR I := 1 TO NUMBER DO
SORTD[I] := DATA[I];
END;

PROCEDURE TPROB (TVALUE, DF: REAL; VAR RP: REAL);
VAR
F, DF1, DF2: REAL;
P, A, B, TEMP, X, XC, AB, TOP, BOT, SUM, TERM, T, LNA, LNB, LNAB: REAL;
CON, SIGN, NTIMES, I: INTEGER;

FUNCTION LNGAM (W: REAL): REAL;
VAR
C1, C2, C3, C4, TEMP, W2: REAL;
I: INTEGER;
BEGIN
  C1 := 0.08333333333300002;
  C2 := 0.00277777777;
  C3 := 7.936507930000002E-04;
  C4 := 0.9189385330000002;
  TEMP:=0.0;
  IF (W <= 13)THEN
    BEGIN
      TEMP:=1.0;
      FOR I:= 1 TO (14 - ROUND(W)) DO
        BEGIN
          TEMP:=TEMP*W;
          W:=W+1;
        END;
      TEMP:=LN(TEMP);
    END;
  W2:=W*W;
  LNGAM:=(Cl -(C2 - C3/W2)/W2)/W + C4 - W + (W - 0.5) * LN(W) - TEMP;
END;
BEGIN
  F:=SQR(TVALUE);
  DF2:=DF;
  DF1:=1;
  IF (F<=O) OR (DF2<=0) THEN
    WRITELN('ILLEGAL VALUE FOR DF',F, DF2)
  ELSE
    BEGIN
      A:=DF1* 0.5;
      B:=DF2* 0.5;
      TEMP:=B+A*F;
      X:=A*F/TEMP;
      IF ((F > 0) AND (X > 0))
      THEN
        BEGIN
          XC:=B/TEMP;
          AB:=A+B;
          CON:=0;
          SIGN:=-1;
          IF (F < 1) THEN
            BEGIN
              TEMP:=A;
              A:=-B;
              B:=TEMP;
              TEMP:=-XC;
              XC:=X;
              X:=TEMP;
              CON:=1;
              SIGN:= -1;
            END;
          END;
          TOP:=AB;
          BOT:=B+1.0;
          SUM:=1.0;
        END;
TERM:=1.0;
NTIMES:=0;
REPEAT
BEGIN
  TEMP:=SUM;
  TERM:=TERM*(TOP/BOT)*XC;
  SUM:=SUM+TERM;
  TOP:=TOP+1;
  BOT:=BOT+1;
  NTIMES:=NTIMES+1;
END;
UNTIL ((SUM <= TEMP) OR (NTIMES >2000));
IF (NTIMES >2000) THEN
BEGIN
  WRITELN('NO CONVERGENCE AFTER ',NTIMES,' TERMS');
  WRITELN('PROBABILITY SET TO 1.00000');
  P:=1.0
END;
P:=CON*SIGN*EXP(A*LN(X)+B*LN(XC)+LNGAM(AB)-LNGAM(A)
   -LNGAM(B))*SUM/B;
RP:= P;
END;
END;
PROCEDURE TTEST(NUMBER:INTEGER; U,MEAN,STD:REAL; VAR T: REAL);
BEGIN
  T:=((MEAN - U)*SQRT(NUMBER))/STD;
END;
PROCEDURE AVE(DATA:ONEDIM; NUMBER:INTEGER; VAR MEAN, STD:REAL);
VAR SUM ,SSUM:REAL;
I: INTEGER;
BEGIN
  SUM:=0;
  SSUM:=0;
  FOR I:=1 TO NUMBER DO
    SUM:=DATA[I] + SUM;
  MEAN:= SUM/NUMBER;
  FOR I:=1 TO NUMBER DO
    SSUM:=SSUM +SQR(DATA[I]-MEAN);
  STD:= SQRT(SSUM/(NUMBER-1));
END;
PROCEDURE MEDI(DATA : ONEDIM; NUMBER:INTEGER; VAR MED: REAL);
VAR L :INTEGER;
BEGIN
  IF ODD(NUMBER) THEN
    BEGIN
      L:=TRUNC((NUMBER +1)/2);
      MED:=DATA[L];
    END
  ELSE
    BEGIN
      L := TRUNC(NUMBER/2);
    END

MED := (DATA[L] + DATA[L+1])/2;
END;

PROCEDURE TRIDATA(DATA: ONEDIM; NUMBER: INTEGER;
TRIM: REAL; VAR TRIP: ONEDIM);
VAR L, J, K : INTEGER;
V : REAL;
TRIDATA : ONEDIM;
BEGIN
  V := NUMBER * TRIM;
  L := TRUNC(V);
  IF L = 0 THEN
    L := 1;
  K := NUMBER - 2*L;
  RSORT(NUMBER, DATA, SORTD);
  FOR J := 1 TO K DO
    TRIDATA[J] := SORTD[J + L];
  AVE(TRIDATA, K, MEAN, STD);
  TTEST(K, U, MEAN, STD, T);
  IF T = 0 THEN
    TRIP[TTIMES] := 1
  ELSE
    BEGIN
      TPROB(T, K-1, RP);
      TRIP[TTIMES] := RP;
    END;
END;

PROCEDURE JACK(DATA: ONEDIM; K, NUMBER: INTEGER; U: REAL; VAR JP: ONEDIM);
VAR G, C, BB, SB, I, J: INTEGER;
GG: REAL;
JDATA: ONEDIM;
BEGIN
  SB := 0;
  BB := 0;
  GG := NUMBER/K;
  IF NUMBER MOD K = 0 THEN
    BEGIN
      G := TRUNC(GG);
      FOR I := 1 TO G DO
        BEGIN
          C := (I-1)*K + 1;
          FOR J := 1 TO NUMBER-K DO
            IF J < C THEN
              JDATA[J] := DATA[J]
            ELSE
              JDATA[J] := DATA[J+K];
          IF MEAN < U THEN
            SB := SB+1
        END
    END;
  END;
END;
ELSE
    BB:= BB+1;
END;

IF (SB/G) < (BB/G) THEN
    IF (SB/G) > 0.5 THEN
        JP[TTIMES] := 2*(1-(SB/G))
    ELSE
        JP[TTIMES] := 2*(SB/G)
    END;
ELSE
    IF (BB/G) > 0.5 THEN
        JP[TTIMES] := 2*(1-(BB/G))
    ELSE
        JP[TTIMES] := 2*(BB/G);
    END;
END;

PROCEDURE BOOTS(DATA:ONEDIM; R,T,R,NUMBER:INTEGER; U:REAL;
    VAR BP:ONEDIM);
VAR SB,BB,J,I,RANO :INTEGER;
    BTVALUE ,BDATA:ONEDIM;
BEGIN
    SB:=0;
    BB:=0;
    FOR J := 1 TO TR DO
        BEGIN
            FOR I:= 1 TO R DO
                BEGIN
                    RANO:=TRUNC (RAN * NUMBER) +1;
                    BDATA[I] := DATA[RANO];
                END;
                AVE(BDATA,R,MEAN,STD);
                IF MEAN < U THEN
                    SB:= SB+1
                ELSE
                    BB:= BB+1;
                END;
            IF (SB/TR) < (BB/TR) THEN
                IF (SB/TR) > 0.5 THEN
                    BP[TTIMES] := 2*(1-(SB/TR))
                ELSE
                    BP[TTIMES] := 2*(SB/TR)
                END;
            ELSE
                IF (BB/TR) > 0.5 THEN
                    BP[TTIMES] := 2*(1-(BB/TR))
                ELSE
                    BP[TTIMES] := 2*(BB/TR);
                END;
            END;
        PROCEDURE RANK(DATA:ONEDIM; NUMBER:INTEGER; VAR RANKP: ONEDIM);
VAR I,J :INTEGER;
    ABSDATA,DDATA,R,SR :ONEDIM;
BEGIN
    FOR I := 1 TO NUMBER DO
        BEGIN
DDATA[I] := DATA[I] - U;
ABSDDATA[I] := ABS(DDATA[I]);
END;
RSORT(NUMBER, ABSDDATA, SORTD);
FOR I := 1 TO NUMBER DO
BEGIN
  SUM := SUM + I;
  N := N + 1;
  IF SORTD[I] <> SORTD[I + 1] THEN
  BEGIN
    FOR J := I - N + 1 TO I DO
      R[J] := SUM/N;
    SUM := 0;
    N := 0;
  END;
END;
FOR J := 1 TO NUMBER DO
FOR I := 1 TO NUMBER DO
  IF (SORTD[I] = ABS(DDATA[J]))
  THEN
    IF DDATA[J] < 0
    THEN
      SR[J] := -R[I]
    ELSE
      SR[J] := R[I];
AVE(SR, NUMBER, MEAN, STD);
TTEST(NUMBER, U, MEAN, STD, T);
IF T = 0 THEN
  RANKP[TTIMES] := 1
ELSE
  BEGIN
    TPROB(T, NUMBER - 1, RP);
    RANKP[TTIMES] := RP;
  END;
END;

(*-------------------------------------------------------------*)

PROCEDURE NORM_PARA(VAR U, U1, VA: REAL; VAR NUMBER: INTEGER);
BEGIN
  WRITELN ('ENTER MEAN , VARIANCE AND SAMPLE SIZE---> ');
  READLN (U1, VA, NUMBER);
  WRITELN ('ENTER THE U FOR HO: ');
  READLN(U);
END;

PROCEDURE GENERATE_NORM(U1, VA: REAL; NUMBER: INTEGER; VAR DATA: ONEDIM);
BEGIN
  NRAN_FLAG := TRUE;
  FOR I := 1 TO NUMBER DO
    DATA[I] := NRAN * SQRT(VA) + U1;

(*----------------------------------------------------------------*)
PROCEDURE EXP_PARA(VAR U, U1: REAL; VAR NUMBER: INTEGER);
BEGIN
  WRITELN('ENTER THE MEAN FOR NEGATIVE EXP. DIS');
  WRITELN('AND SAMPLE SIZE');
  READLN(U1, NUMBER);
  WRITELN('ENTER THE U FOR HO:');
  READLN(U);
END;

PROCEDURE GENERATE_EXP(U, U1: REAL; NUMBER: INTEGER; VAR DATA: ONEDIM);
BEGIN
  FOR I := 1 TO NUMBER DO
    DATA[I] := (-LN(1 - RAN)) - 1 + U1;
END;

PROCEDURE DEXP_PARA(VAR U, U1: REAL; VAR NUMBER: INTEGER);
BEGIN
  WRITELN('ENTER THE MEAN ');
  WRITELN('ENTER THE SAMPLE SIZE');
  READLN(U1, NUMBER);
  WRITELN('ENTER THE U FOR HO:');
  READLN(U);
END;

PROCEDURE GENERATE_DEXP(U, U1: REAL; NUMBER: INTEGER; VAR DATA: ONEDIM);
VAR B, R: REAL;
BEGIN
  B := SQRT(1/2);
  FOR I := 1 TO NUMBER DO
    BEGIN
      R := RAN;
      IF (R) < 0.5
      THEN
        DATA[I] := B * LN(2*R) + U1
      ELSE
        IF R = 0.5
        THEN
          DATA[I] := 0 + U1
        ELSE
          DATA[I] := -B * LN(2*(1 - R)) + U1;
      END;
    END;
END;
PROCEDURE CAUCHY_PARA(VAR U, U1: REAL; VAR NUMBER: INTEGER);
BEGIN
   WRITELN('ENTER THE U FOR CAUCHY DISTRIBUTION');
   WRITELN('ENTER THE SAMPLE SIZE');
   READLN(U1, NUMBER);
   U := 0;
END;

PROCEDURE GENERATE_CAUCHY(U, U1: REAL; NUMBER: INTEGER; VAR DATA: ONEDIM);
VAR R, A, B: REAL;
BEGIN
   FOR I := 1 TO NUMBER DO
      BEGIN
         R := RAN;
         A := SIN((R * PI) - (PI/2));
         B := COS((R * PI) - (PI/2));
         DATA[I] := 0.675 * (A/B) + U1;
      END;
END;

PROCEDURE UNI_PARA(VAR U, U1: REAL; VAR NUMBER: INTEGER);
BEGIN
   WRITELN('ENTER SAMPLE SIZE, AND MEAN');
   READLN(NUMBER, U1);
   U := 0;
END;

PROCEDURE GENERATE_UNI(U, U1: REAL; NUMBER: INTEGER; VAR DATA: ONEDIM);
BEGIN
   FOR I := 1 TO NUMBER DO
      DATA[I] := (RAN * SQRT(12)) - SQRT(3) + U1;
END;

(*---------------------------------------------*)

BEGIN
   WRITELN('CHOOSE THE DISTRIBUTION-- YOU WANT TO SIMULATE');
   WRITELN('NORMAL--1, EXP--2, DOUBLE EXP--3, CAUCHY--4, UNIFORM--5');
   READLN(CHOICE);
   WRITELN('ENTER NUMBER OF TRAIL--');
   READLN(TIMES);
   IF CHOICE = 1 THEN
      BEGIN
         NORM_PARA(U, U1, VA, NUMBER);
         D := 'N';
         AS := ROUND(NUMBER/10 + 48);
         SS := CHR(AS);
      END;
IF CHOOSE = 2 THEN
BEGIN
EXP_PARA(U,U1,NUMBER);
D:='E';
AS:=ROUND( NUMBER/10 +48);
SS:=CHR(AS);
END;

IF CHOOSE = 3 THEN
BEGIN
DEXP_PARA(U,U1,NUMBER);
D:='D';
AS := ROUND(NUMBER/10 +48);
SS:=CHR(AS);
END;

IF CHOOSE = 4 THEN
BEGIN
CAUCHY_PARA(U,U1,NUMBER);
D:='C';
AS := ROUND(NUMBER/10 +48);
SS:=CHR(AS);
END;

IF CHOOSE = 5 THEN
BEGIN
UNI_PARA(U,U1,NUMBER);
D:='U';
AS := ROUND(NUMBER/10 +48);
SS:=CHR(AS);
END;

WRITELN('ENTER POWER CODE TO TEST HO:');
READLN(PCODE);
PC:=CHR(PCODE + 48);
WRITELN('ENTER THE SAMPLE SIZE--USE BOOTSTRAP METHOD');
READLN(R);
WRITELN('ENTER HOW MANY TIMES YOUR WANT TO RESAMPLE--');
READLN(TR);
WRITELN;
WRITELN('ENTER THE SUBGROUP K--USE JACKNIFE METHOD ');
READLN (K);
WRITELN('ENTER THE % TO TRIM---EACH TAIL');
READLN(TRIM);
S:=0;
REPEAT

FILEN1:=D+SS+PC+EXT1;
FILEN2:=D+SS+PC+EXT2;
FILEN3:=D+SS+PC+EXT3;
OPEN(OUTFILE1, FILEN1, NEW);
REWRITE(OUTFILE1);
OPEN(OUTFILE2, FILEN2, NEW);
REWRITE(OUTFILE2);
OPEN (OUTFILE3, FILEN3, NEW);
REWRITE(OUTFILE3);
S:=0;
REPEAT
BEGIN
  S:=S+1;
  TTIMES:=0;
  REMP:=0;
  TEMP:=0;
  JEMP:=0;
  BEMP:=0;
  SREMP:=0;
REPEAT
  BEGIN
    TTIMES:= TTIMES + 1;
    IF CHOOSE = 1 THEN
      GENERATE_NORM(U1,VA,NUMBER,DATA);
    END;
    IF CHOOSE = 2 THEN
      GENERATE_EXP(U,U1,NUMBER,DATA);
    END;
    IF CHOOSE = 3 THEN
      GENERATE_DEXP(U,U1,NUMBER,DATA);
    END;
    IF CHOOSE = 4 THEN
      GENERATE_CAUCHY(U,U1,NUMBER,DATA);
    END;
    IF CHOOSE = 5 THEN
      GENERATE_UNI(U,U1,NUMBER,DATA);
      AVE(DATA,NUMBER,MEAN,STD);
      TTEST(NUMBER,U,MEAN,STD,T);
      ORGTVALUE:= T;
      IF ORGTVALUE =0 THEN
        RAWP[TTIMES) :=1
      ELSE
        TPROB( ORGTVALUE,NUMBER-1,RP);
        RAWP[TTIMES] :=RP;
        IF RAWP[TTIMES] <0.05 THEN
          REMP := REMP +1;
          TRIDATA(DATA,NUMBER,TRIM,TRIP);
          IF TRIP[TTIMES] < 0.05 THEN
            TEMP := TEMP +1;
            JACK(DATA,K,NUMBER,U,JP);
            IF JP[TTIMES] < 0.05 THEN
              JEMP := JEMP +1;
              BOOTS(DATA,R,TR,NUMBER,U,BP);
              IF BP[TTIMES] < 0.05 THEN
                BEMP := BEMP +1;
                RANK(DATA,NUMBER,RANKP);
                IF RANKP[TTIMES] < 0.05 THEN
                  SREMP:= SREMP + 1;
          END;
      END;
      WRITELN(OUTFILE2,RAWP[TTIMES]:10:6,TRIP[TTIMES]:10:6,
               JP[TTIMES]:10:6,BP[TTIMES]:10:6,RANKP[TTIMES]:10:6);
  END;
UNTIL TTIMES=TIMES;
  WRITELN(OUTFILE3,REMP/TIMES:10:6,TEMP/TIMES:10:6,
JEMP/TIMES:10:6, BEMP/TIMES:10:6, SREMP/TIMES:10:6);
WRITE(OUTFILE1, TIMES:3, CHOOSE:3, NUMBER:3, PCODE:2);
AVE(RAWP, TIMES, MEAN, STD);
RSORT(TIMES, RAWP, SORTD);
MEDI(SORTD, TIMES, MED);
WRITE(OUTFILE1, MEAN:8:4, STD:7:4, MED:7:4);

AVE(TRIP, TIMES, MEAN, STD);
RSORT(TIMES, TRIP, SORTD);
MEDI(SORTD, TIMES, MED);
WRITE(OUTFILE1, MEAN:8:4, STD:7:4, MED:7:4);

AVE(JP, TIMES, MEAN, STD);
RSORT(TIMES, JP, SORTD);
MEDI(SORTD, TIMES, MED);
WRITE(OUTFILE1, MEAN:8:4, STD:7:4, MED:7:4);

AVE(BP, TIMES, MEAN, STD);
RSORT(TIMES, BP, SORTD);
MEDI(SORTD, TIMES, MED);
WRITE(OUTFILE1, MEAN:8:4, STD:7:4, MED:7:4);

AVE(RANKP, TIMES, MEAN, STD);
RSORT(TIMES, RANKP, SORTD);
MEDI(SORTD, TIMES, MED);
WRITE(OUTFILE1, MEAN:8:4, STD:7:4, MED:7:4);
END;
UNTIL S=2;
CLOSE(OUTFILE1);
CLOSE(OUTFILE2);
CLOSE(OUTFILE3);
END.
## Appendix 2: Data Table

### Distribution Code
1: Normal Distribution  
2: Double Exponential Distribution  
3: Uniform Distribution  
4: Exponential Distribution  
5: Cauchy Distribution

### Test Code
1: T-test  
2: Trimmed t-test  
3: Jackknife  
4: Bootstrap  
5: Signed-rank

<table>
<thead>
<tr>
<th>Distribution Code</th>
<th>Test Code</th>
<th>μ value Code</th>
<th>Average P-value</th>
<th>Empirical Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.5159</td>
<td>0.04</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0.2435</td>
<td>0.32</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0.06475</td>
<td>0.72</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td>0.007</td>
<td>0.965</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>5</td>
<td>0.00035</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>6</td>
<td>0.00005</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0.51895</td>
<td>0.04</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0.2712</td>
<td>0.3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0.05805</td>
<td>0.8</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4</td>
<td>0.0111</td>
<td>0.935</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>5</td>
<td>0.0012</td>
<td>0.99</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>6</td>
<td>0.00015</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0.5202</td>
<td>0.045</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0.30975</td>
<td>0.215</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
<td>0.0529</td>
<td>0.755</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4</td>
<td>0.00485</td>
<td>0.98</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>5</td>
<td>0.0007</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0.50685</td>
<td>0.045</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>0.2644</td>
<td>0.205</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>3</td>
<td>0.02935</td>
<td>0.855</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
<td>0.00135</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>5</td>
<td>0.0001</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0.44490</td>
<td>0.015</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>0.44250</td>
<td>0.030</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>3</td>
<td>0.37255</td>
<td>0.145</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>4</td>
<td>0.32290</td>
<td>0.200</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>5</td>
<td>0.27925</td>
<td>0.290</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>6</td>
<td>0.25420</td>
<td>0.385</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>7</td>
<td>0.22050</td>
<td>0.420</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0.48295</td>
<td>0.08</td>
</tr>
<tr>
<td>Distribution Code</td>
<td>Test Code</td>
<td>(\mu) value Code</td>
<td>Average P-value</td>
<td>Empirical Power</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------</td>
<td>---------------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0.2153</td>
<td>0.39</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0.05145</td>
<td>0.81</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>4</td>
<td>0.00465</td>
<td>0.98</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>5</td>
<td>0.00015</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0.47125</td>
<td>0.07</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.2049</td>
<td>0.395</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0.03805</td>
<td>0.865</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>4</td>
<td>0.00339</td>
<td>0.985</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>5</td>
<td>0.0003</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0.49395</td>
<td>0.075</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0.3028</td>
<td>0.245</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>3</td>
<td>0.0559</td>
<td>0.755</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>4</td>
<td>0.00525</td>
<td>0.98</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>5</td>
<td>0.00085</td>
<td>0.995</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>1</td>
<td>0.42925</td>
<td>0.155</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2</td>
<td>0.33815</td>
<td>0.17</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>3</td>
<td>0.03865</td>
<td>0.83</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>4</td>
<td>0.0008</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>1</td>
<td>0.42475</td>
<td>0.060</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>2</td>
<td>0.38345</td>
<td>0.155</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>3</td>
<td>0.25730</td>
<td>0.370</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>4</td>
<td>0.17015</td>
<td>0.545</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>5</td>
<td>0.15880</td>
<td>0.645</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>6</td>
<td>0.09965</td>
<td>0.740</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>7</td>
<td>0.07240</td>
<td>0.760</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0.1493</td>
<td>0.64</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>2</td>
<td>0.033</td>
<td>0.895</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
<td>0.0017</td>
<td>0.99</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0.104</td>
<td>0.605</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
<td>0.0463</td>
<td>0.845</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0.002</td>
<td>0.99</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0.14335</td>
<td>0.74</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>0.0677</td>
<td>0.875</td>
</tr>
<tr>
<td>Distribution Code</td>
<td>Test Code</td>
<td>$\mu$ value Code</td>
<td>Average P-value</td>
<td>Empirical Power</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------</td>
<td>------------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>1</td>
<td>0.119</td>
<td>0.655</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>0.03935</td>
<td>0.84</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>1</td>
<td>0.05965</td>
<td>0.555</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>2</td>
<td>0.06370</td>
<td>0.535</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>3</td>
<td>0.06130</td>
<td>0.610</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>4</td>
<td>0.02900</td>
<td>0.705</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>5</td>
<td>0.03500</td>
<td>0.740</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>6</td>
<td>0.02670</td>
<td>0.745</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>7</td>
<td>0.01735</td>
<td>0.845</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0.4937</td>
<td>0.085</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>2</td>
<td>0.2272</td>
<td>0.36</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>3</td>
<td>0.0616</td>
<td>0.765</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>4</td>
<td>0.0048</td>
<td>0.955</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>5</td>
<td>0.0001</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>1</td>
<td>0.5083</td>
<td>0.06</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>2</td>
<td>0.2484</td>
<td>0.34</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>3</td>
<td>0.053</td>
<td>0.8</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>4</td>
<td>0.0103</td>
<td>0.935</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>5</td>
<td>0.0012</td>
<td>0.99</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6</td>
<td>0.0001</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>1</td>
<td>0.5113</td>
<td>0.075</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>2</td>
<td>0.2926</td>
<td>0.27</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>3</td>
<td>0.0473</td>
<td>0.8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>4</td>
<td>0.0023</td>
<td>0.995</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>0.0003</td>
<td>0.995</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>1</td>
<td>0.489</td>
<td>0.085</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2</td>
<td>0.2437</td>
<td>0.315</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>0.0162</td>
<td>0.925</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>1</td>
<td>0.43620</td>
<td>0.095</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>2</td>
<td>0.43520</td>
<td>0.120</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>3</td>
<td>0.35930</td>
<td>0.290</td>
</tr>
<tr>
<td>Distribution Code</td>
<td>Test Code</td>
<td>μ value Code</td>
<td>Average P-value</td>
<td>Empirical Power</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------</td>
<td>--------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>4</td>
<td>0.30710</td>
<td>0.320</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>5</td>
<td>0.25840</td>
<td>0.435</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>6</td>
<td>0.24090</td>
<td>0.510</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>7</td>
<td>0.20120</td>
<td>0.545</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>1</td>
<td>0.51090</td>
<td>0.045</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>2</td>
<td>0.25630</td>
<td>0.295</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>3</td>
<td>0.07175</td>
<td>0.695</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>4</td>
<td>0.00815</td>
<td>0.955</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>5</td>
<td>0.00035</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>6</td>
<td>0.00005</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>1</td>
<td>0.5191</td>
<td>0.04</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>2</td>
<td>0.21445</td>
<td>0.37</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>3</td>
<td>0.03825</td>
<td>0.86</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>4</td>
<td>0.00455</td>
<td>0.965</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>5</td>
<td>0.00005</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>6</td>
<td>0.00005</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>1</td>
<td>0.51615</td>
<td>0.045</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>2</td>
<td>0.327</td>
<td>0.205</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>3</td>
<td>0.08235</td>
<td>0.685</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>4</td>
<td>0.01195</td>
<td>0.945</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>5</td>
<td>0.00215</td>
<td>0.99</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>1</td>
<td>0.3819</td>
<td>0.175</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>2</td>
<td>0.45585</td>
<td>0.07</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>3</td>
<td>0.10815</td>
<td>0.615</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>4</td>
<td>0.001</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>1</td>
<td>0.52290</td>
<td>0.045</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>2</td>
<td>0.35640</td>
<td>0.135</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>3</td>
<td>0.14630</td>
<td>0.500</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>4</td>
<td>0.06390</td>
<td>0.790</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>0.03265</td>
<td>0.850</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>6</td>
<td>0.01850</td>
<td>0.915</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>7</td>
<td>0.00880</td>
<td>0.945</td>
</tr>
</tbody>
</table>