

ISSN Online: 2456-1452
NAAS Rating (2026): 4.49
IJSAM 2026; 11(10): 89-102
www.mathsjournal.com
Received: 17-08-2026
Accepted: 26-09-2026
Published: 10-10-2026

Sarah Alessa

Department of Statistics,
North Dakota State University,
Fargo, North Dakota 58108,
USA

Rhonda Magel

Department of Statistics,
North Dakota State University,
Fargo, North Dakota 58108,
USA

Comparison of nonparametric tests for the simple tree alternative

Sarah Alessa¹, Rhonda Magel²

DOI: <https://www.doi.org/10.22271/math.2026.v11.i10b.2719>

ABSTRACT

Two new nonparametric tests are introduced and compared with existing tests and new proposed versions of these existing tests for testing differences in location or scale based on the simple tree alternative. There is a need to develop a general test to compare a control population with different treatment level populations to see whether applying any level of the treatment results in a difference of location and/or scale. Namely, this type of test would be applied if researchers wanted to know if any of the treatment levels studied had any effect at all on the outcome. It would be the first test applied to determine whether any difference exists. If the null hypothesis was rejected, follow up testing would need to be conducted in order to determine where the difference or differences occurred. Few nonparametric tests exist of this nature. This is particularly true in the case of skewed distributions or heavy tailed symmetric distributions. This research proposes two new tests for location and scale that combine the Mann-Whitney test for location with two nonparametric tests for scale that have not been considered in this setting. The two new tests are compared with existing tests (and slight variations of some of the existing tests) based on a simulation study. The simulation study is conducted to estimate the significance levels of the two proposed tests (in comparison to existing tests and their modified versions) in order to establish if significant levels are maintained under the assumption of independence of each of the components of the tests. This is done for both symmetric (normal, double exponential, logistic, Cauchy) and skewed (exponential) distributions. A problem in the past has been if the underlying distributions are skewed since the two test components are often dependent when this is the case. The assumption of being able to assume the two components of the test statistics are independent (or actually that the covariance is 0) will be checked by estimating the significance levels. If the significance levels are "close" to the stated significance levels, it will be assumed that the covariance of the two components making up the test is approximately 0. The significance levels will be considered "close" to the stated significance level of 0.05 if they are in the interval (0.04, 0.06). If the significance levels are maintained, the study also estimates the powers and compares them among the tests. Recommendations are given based on symmetric distributions, asymmetric distributions, and very heavy tailed symmetric distributions as to which test is better to use. Tests were found that would work for all types of distributions.

Keywords: Mann-Whitney, Moses, Conover Squared Rank, Hollander, Mood, Scale, Location, Control, Treatments

1. Introduction

Often when applying a treatment, it is common to observe a shift in both the location and scale simultaneously (Hussain (2026) [1]). In many situations, researchers would like to test for this effect, namely a location-scale shift. LePage (1970) proposed a nonparametric test based on independent samples from two populations to test for a difference in either location and/or scale when a treatment was applied. This test is the combination of the Mann-Whitney-Wilcoxon test for location (Mann *et al.*, 1947; Wilcoxon, 1945) [2, 3] and the Ansari-Bradley test for scale (Ansari *et al.*, 1960) [4]. The LePage test was found to be good for the location-scale problem when the underlying distributions were symmetric with medium tails (Hussain (2026) [1]). One example of the LePage test using real data may be found in Huang *et al.* (2014) [5]. The idea that LePage proposed by combining a nonparametric test for location with a nonparametric test for scale has been carried over by other researchers who have shown that a nonparametric test for location can be combined with a nonparametric scale in order to test for the location-scale problem (Marozzi, 2008) [6]. The LePage type test was extended to k populations in order to test whether there was a difference in either location and/or scale between any of the k populations (Rublik, 2005) [7]. This test was a general test for k populations. However, there are situations when researchers are interested in testing whether at least one treatment level has a higher location and/or higher scale than the control. The simple tree alternative is tested in this case for both location and scale. It is assumed that if there is a difference in location and/or scale be

Corresponding Author: Rhonda Magel

Department of Statistics,
North Dakota State
University, Fargo, North
Dakota 58108, USA

tween the control and a treatment level, the treatment level will have the higher location and/or scale. The simple tree alternative hypothesis for both location and scale is expressed as follows:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k, (1)$$

and $\sigma_1 = \sigma_2 = \dots = \sigma_k$, versus

$$H_a: \mu_1 \leq [\mu_2, \dots, \mu_k] \text{ and / or}$$

and $\leq [\sigma_2, \dots, \sigma_k]$ where at least one inequality is strict.

where population 1 is designated as the control. It is noted that μ_i represents the location parameter for population i (not necessarily the mean) and σ_i represents the scale parameter for population i . In this paper, we are interested in nonparametric statistical tests for the simple tree alternative. We will look at some that have been considered and we will be proposing other tests. All of the tests considered will be compared on the basis of how well they maintain their Type 1 error, and then on the basis of estimated power if their Type 1 errors are maintained. All of the null test statistic distributions considered are asymptotic distributions.

Mann-Whitney

The Mann-Whitney test (MW) is a nonparametric test, which tests the null hypothesis of equal location parameters for two populations (Mann *et al.*, 1947) [2]. The null and alternative hypothesis are expressed as follows:

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2, H_a: \mu_1 < \mu_2, H_a: \mu_1 > \mu_2$$

The problem with using the Mann-Whitney test alone is that it is a test to test for differences in location, not a test to test for differences in either location or scale. The Mann-Whitney test also assumes that scale parameters are equal when using this test. When scale parameters are not equal, the power of the Mann-Whitney test can be negatively affected even when the locations differ (Alsubie *et al.*, 2020a) [8].

To calculate the Mann-Whitney test statistic (MW), one has to assume that there are two independent samples of sizes of n_1 and n_2 from two different populations. A total of $N = n_1 + n_2$ observations are obtained by combining these samples and sorting them from smallest to largest. When more than one observation is equal, the average of the order of ranks is given to the matched observation. Each observation has a rank, and in Population 1, S_i represents the rank of the i th observation with the total sum of these ranks being $\sum S_i$. The Mann-Whitney MW test statistic for Population 1 is given by:

$$MW = \sum S_i - \frac{n_1(n_1 + 1)}{2}$$

The standardized form of the Mann-Whitney test has been provided by this expression:

$$MW^* = \frac{MW - E_0(MW)}{\sqrt{\text{var}_0(MW)}}$$

$$E_0(MW) = \frac{n_1 n_2}{2}$$

$$\text{var}_0(MW) = \frac{n_1 n_2 (N + 1)}{12}$$

The standardized Mann-Whitney test statistic MW^* approximates a standard normal distribution under the null hypothesis. For a one-sided alternative ($H_a: \mu_1 < \mu_2$), H_0 is rejected when $MW^* \leq -Z_\alpha$, where Z_α is the $(1 - \alpha)100$ percentile of the standard normal distribution (for example, $Z_\alpha = 1.645$ when $\alpha = 0.05$).

2. Moses test

The Moses test is a nonparametric test for testing the null hypothesis of equality of scale for two populations (Moses, 1963) [9]. The Moses test does not assume equality of location parameters (Azman *et al.*, 2021; Daniel, 2016) [10, 11]. It is a two-sample test for scales testing the following set of hypotheses.

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_{a1}: \sigma_1^2 \neq \sigma_2^2, H_{a2}: \sigma_1^2 < \sigma_2^2, H_{a3}: \sigma_1^2 > \sigma_2^2$$

The Moses test statistic is obtained by dividing two samples into m_1 and m_2 subsamples of equal size. For each of the subsamples, the mean of the subsample is calculated, followed by summing of squared deviations between each observation and the sample mean for each subsample. Furthermore, the sums of squared deviations for each of the m_1 subsets in the first sample are indicated as C_1 to C_{m_1} , and those of the second set of m_2 as D_1 to D_{m_2} .

The Mann-Whitney test (Mann *et al.*, 1947) [2] will be performed. The C values from the first sample and the D values from the second will be combined and then will be ranked from smallest to largest, and each value will be awarded a rank. The Moses test statistic (M) is the sum of the ranks associated with the C values from the first sample, which can be written as:

$$M = \sum S_i - \frac{m_1(m_1 + 1)}{2}$$

The standardized version of the Moses test is given by:

$$M^* = \frac{M - E_0(M)}{\sqrt{\text{var}_0(M)}}$$

$$E_0(M) = \frac{m_1 m_2}{2}$$

$$\text{var}_0(M) = \frac{m_1 m_2 (m_1 + m_2 + 1)}{12}$$

The asymptotic distribution of M^* under H_0 is a standard normal distribution. Alsubie *et al.* (2020b) [12] extended the use of the Moses test to the simple tree alternative.

One problem with the Moses test is that it transforms a scale test into a test for location and reduces the sample sizes in the process.

3. Conover squared ranks test

The Conover Squared Ranks test is a nonparametric test based on two independent samples selected from two populations for testing the null hypothesis of equal variances (Conover, 1999) [13]. It does not assume that location parameters are equal, but it does assume they are known which is often not the case.

To compute the Conover Squared Ranks test statistic (CSR), two independent random samples are selected from two populations $X_1, X_2, \dots, X_n$ and $Y_1, Y_2, \dots, Y_m$ with sizes of n and m , respectively. For each sample, the absolute deviations between each observation and their population mean will be computed as follows:

$$U_i = |X_i - \mu_1| \text{ for population 1 } (i = 1, \dots, n)$$

$$\text{and } V_j = |Y_j - \mu_2| \text{ for population 2 } (j = 1, \dots, m)$$

where, μ_1 is the population mean for population 1 and μ_2 is the population mean for population 2. If these population means are unknown, they can be substituted with sample means $\bar{X}$ and $\bar{Y}$, respectively and the test remains approximately accurate. It is noted that using the sample means is not an exact solution and these may not actually work for heavy tailed or skewed distributions. In this research, each value in a sample will also be adjusted by subtracting the sample median in addition to the version using the sample mean.

Next, the U_i 's and V_j 's are combined into a single sample and ranks are assigned from 1 to $N = n + m$. When more than one observation has equal values, the average of the order of ranks is given to the matched observations. $R(U_i)$ and $R(V_j)$ denote the ranks of U_i and V_j , respectively. The Conover Squared Ranks test statistic is defined by:

$$CSR = \sum_{i=1}^n (R(U_i))^2$$

The standardized version of the Conover Squared Ranks test is given by:

$$CSR^* = \frac{CSR - E_0(CSR)}{\sqrt{var_0(CSR)}}$$

$$E_0(CSR) = \frac{n(N+1)(2N+1)}{6}$$

$$var_0(CSR) = \frac{mn(N+1)(2N+1)(8N+11)}{180}$$

Where: $N = n + m$.

The Conover Squared Ranks test statistic CSR^* has approximately a standard normal distribution under the null hypothesis when the means are known. When the means are unknown, it is recommended that this approximation holds by using sample means. Again, in this research, we will also try using the sample medians. For a one-sided alternative ($H_{a2}: \sigma_1^2 < \sigma_2^2$), H_0 is rejected when $CSR^* \leq -Z_\alpha$.

4. Hollander test of extreme reaction

The Hollander test of extreme reaction is a nonparametric test designed to test for differences between the control and experimental subjects when some of the latter are predicted to behave in one way and the others in the opposite way (Hollander, 1963) [14]. The set of hypotheses is given below (Daniel, 2016) [11]. The Hollander test does assume each of the two populations is centered around the same location parameter.

H_0 : The two samples may be considered as having been drawn from the same population

H_a : One population consists of observations resulting from extreme reactions in both directions

The Hollander test statistic (H) is computed by combining size n_1 (control) and n_2 (experimental) samples from two populations. These samples are pooled and arranged from smallest to largest, yielding a total of $N = n_1 + n_2$ observations. Each observation is assigned a rank and r_i denotes the rank of the i th X value in the combined set of ranks. The Hollander test statistic (H), is calculated as: $H = \sum_{i=1}^{n_1} (r_i - \bar{r})^2$. The mean rank of the n_1 values in the first sample is denoted by $\bar{r}$.

$$\bar{r} = \frac{\sum_{i=1}^{n_1} r_i}{n_1}$$

$$MD^* = \frac{MD - E_0(MD)}{\sqrt{var_0(MD)}}$$

$$E_0(MD) = \frac{[n(n+m+1)(n+m-1)]}{12}$$

$$var_0(MD) = \frac{[nm(n+m+1)(n+m+2)(n+m-2)]}{180}$$

The standardized Mood test statistic MD^* has approximately a standard normal distribution under the null hypothesis. For a one-sided al-

The standardized version of Hollander statistic can be stated as:

$$H^* = \frac{H - E_0(H)}{\sqrt{var_0(H)}}$$

$$E_0(H) = \frac{(n_1 - 1)(N^2 + N)}{12}$$

$$var_0(H) = E(H^2) - [E(H)]^2$$

Where:

$$E(H^2) = \frac{(n_1 - 1)^2}{720} \left[\frac{6}{n_1} (N^4 + 2N^3 + N^2) + \left(\frac{n_1 + 1}{n_1 - 1} \right) (5N^4 + 6N^3 - 5N^2 - 6N) \right]$$

The standardized Hollander test statistic H^* has approximately a standard normal distribution under the null hypothesis. H_0 is rejected when $H^* \leq -Z_\alpha$. It is noted that in this research, a new modification using the median rank of the values in the first sample will also be used in addition to the version of using the sample mean.

5. Mood test

Mood's test is a nonparametric test, which tests the null hypothesis of equality of scale parameters for two populations (Gogoi et al., 2016) [15] assuming the location parameters are equal. The Mood test statistic is defined by:

$$MD = \sum_{i=1}^n \left(R_i - \frac{n+m+1}{2} \right)^2$$

where R_i is the rank of the i th value from the first sample when all observations from both samples are pooled and sorted from least to greatest. The element with the lowest value is ranked 1, the next lowest ranked 2, and so on, until the highest value element occupies rank $n + m$.

The standardized version of Mood statistic is expressed as follows:

ternative ($H_{a2}: \sigma_1^2 < \sigma_2^2$), H_0 is rejected when $MD^* \leq -Z_\alpha$, where Z_α is the $(1 - \alpha)100$ percentile of the standard normal distribution.

1.6 Wald-Wolfowitz

The Wald-Wolfowitz test (Wald *et al.*, 1940) [16] is a test to see whether two independent samples come from identical distributions. The observations from both samples are arranged in order from smallest to largest keeping track of which sample the observation came from. If the two population distributions are identical, the observations from both samples should be well mixed. The null hypothesis of two identical population distributions is rejected based on a small number of runs where a run is a sequence of observations from the same sample. The power efficiency of this test is not very good (Daniel, 2016) [11]. The standardized runs test follows an asymptotic standard normal distribution when the null hypothesis is true.

1.7 Ansari-Bradley

The Ansari-Bradley test is a test to test whether two populations have the same scale parameter or if one scale parameter is greater than the other. Observations are arranged in order from smallest to largest and then ranked by assigning the smallest and largest value a rank of 1, the second smallest and second largest value a rank of 2, and this continues. The test statistic is the sum of the ranks in the second sample. If one is testing that the scale parameter of the first population is smaller than the scale parameter of the second population, the null hypothesis is rejected for small values of the test statistic. The asymptotic distribution of the Ansari-Bradley test statistic is normal under the null hypothesis.

2. Materials and methods

The simple tree alternative given in (1), was tested with the test statistics conducted on

$k = 3$ populations of equal and unequal sample sizes consisting of 9, 18, and 36. These sample sizes were selected since nonparametric test are often used with small sample sizes (Daniel, 2016) [11]. In conducting the tests, the treatment samples were combined into a single treatment group, and the alternative hypothesis stated that the location and/or scale parameter for the control population is less than the location and/or scale parameter for the combined treatment population. Random samples were generated from the following distributions: Normal, Exponential, Cauchy, Double Exponential, Laplace, and Logistic.

2.1. Tests to be compared

2.1.1 Test 1 (MW+MS):

(MW + MS) is a test originally proposed by Alsubie *et al.* (2020b) [12] which is the sum of standardized Mann-Whitney (MW^*) and Moses (M^*) test statistics, defined as follows:

$$MW + MS = \frac{MW^* + M^*}{\sqrt{2}}$$

Under the null hypothesis, MW^* and M^* both have an asymptotic standard normal distribution. The sum of these test statistics, $MW^* + M^*$, follows an asymptotic normal distribution with a mean of zero and a variance of 2 when H_0 is true, assuming asymptotic independence of MW^* and M^* .

The asymptotic distribution of the test statistic MW+MS approximates a standard normal distribution under the null hypothesis when both components are asymptotically independent. For a one-sided alternative, H_0 is rejected when $MW + MS \leq -Z_\alpha$, where Z_α is the $(1 - \alpha)100$ percentile of the standard normal distribution.

For the Moses test, Alsubie *et al.* (2020b) [12] used the sample mean for each subset in Moses test when calculating the sum of squared standard deviations. A proposed version using the sample median in place of the sample mean for each subgroup will also be examined.

2. Test 2 (proposed) (MW+CSR):

The second test being compared in this research is a newly proposed test, MW + CSR. This test statistic is the sum of the standardized

Mann-Whitney (MW^*) and Conover Squared Ranks (CSR^*) test statistics, defined as follows:

$$MW + CSR = \frac{MW^* + CSR^*}{\sqrt{2}}$$

Since under the null hypothesis, both MW^* and CSR^* have asymptotic standard normal distributions, the sum of these test statistics, $MW^* + CSR^*$, follows an asymptotic normal distribution with a mean of zero and a variance of 2 when H_0 is true, assuming the components are asymptotically independent. The asymptotic distribution of the test statistic MW+CSR approximates a standard normal distribution under the null hypothesis. For a one-sided left alternative, H_0 is rejected when $MW + CSR \leq -Z_\alpha$, where Z_α is the $(1 - \alpha)100$ percentile of the standard normal distribution. Two versions of this test will be examined. In the first version of the test, the population means in the CSR test will be replaced with sample means. In the second version of the test, the population means in the CSR test will be replaced with sample medians.

3. Test 3 (MW+HL)

(MW + HL) is the third test that will be compared. This was originally proposed by Nan *et al.* (2025) [17], which is the sum of standardized Mann-Whitney (MW^*) and Hollander (H^*) test statistics, defined as follows:

$$MW + HL = \frac{MW^* + H^*}{\sqrt{2}}$$

Nan *et al.* (2025) [17] used the sample mean when calculating the Hollander test statistic. The sample median will also be used in this research.

Under the null hypothesis, both MW^* and H^* have an asymptotic standard normal distribution. The sum follows a normal distribution with a mean of zero and a variance of 2, when H_0 is true and the two components are asymptotically independent.

The asymptotic distribution of the test statistic MW+HL approximates a standard normal distribution under the null hypothesis. For a one-sided alternative, H_0 is rejected when $MW + HL \leq -Z_\alpha$, where Z_α is the $(1 - \alpha)100$ percentile of the standard normal distribution.

4. Test 4 (proposed) (MW+MD)

The second test that is being proposed is MW + MD. This test is the sum of standardized Mann-Whitney (MW^*) and Mood (MD^*) test statistics, defined as follows:

$$MW + MD = \frac{MW^* + MD^*}{\sqrt{2}}$$

Under the null hypothesis, MW^* and MD^* both have an asymptotic standard normal distribution. The sum follows a normal distribution with a mean of zero and a variance of 2, when H_0 is true and the two components are asymptotically independent.

The asymptotic distribution of the test statistic MW+MD approximates a standard normal distribution under the null hypothesis. For a one-sided alternative, H_0 is rejected when $MW + MD \leq -Z_\alpha$, where Z_α is the $(1 - \alpha)100$ percentile of the standard normal distribution.

5. Test 5 (existing) (MW+AB)

Lepage (1971) [18] proposed combining the Mann-Whitney test for location and the Ansari-

Bradley test for scale together to test for differences in location and/or scale between two populations. Alsubie *et al.* (2020a) [8] extended this test to the simple tree alternative. Damperty *et al.* (2026) [19] analyzed the power of each of the two test components and the power of the combined test for the normal, t(3), and exponential distributions. They considered cases when the locations only were different, when the scales only were different, and when both the location and scale

parameters were different. Dampney *et al.* (2026) [19] found that the combined Alsubie and Magel test was the best test to use in the case of the normal and $t(3)$ when testing for differences in locations and/or scale. The combined test had higher powers than the individual components when there were differences in both, and sometimes, much higher powers. Dampney *et al.* (2026) [19] found the results were different for the exponential distribution. When both location and scale parameters differed, it was better to use the Mann-Whitney test instead of the combined Alsubie and Magel test. This test is included in our simulation study for comparison purposes. The test is often used in the two-sample case as a location scale test, but it does not do well for skewed distributions (such as the exponential) in testing for both location and scale. One goal of this research is to develop and/or recognize a location-scale test that performs well for skewed distributions.

6. Test 6 (existing) (MW+WW)

This test was first proposed by Schukr and Magel (2025) for the simple tree alternative.

The WW test does not have a lot of power to detect differences in location and/or scale. The researchers hypothesized that combining the WW test with the MW test, a test would be developed with significantly higher power than the WW test alone. Schukr and Magel (2025) found that this test did well when underlying distributions were exponential and both the location and scale parameters changed, but the significance levels were somewhat inflated. This test is included in this study to see how well the proposed tests do in comparison. Namely, is power gained by using one of the proposed tests, especially in the case of skewed distributions. A proposed modified version of this test will also be compared using the sample median instead of the sample mean to see the effects on the significance levels in skewed distributions.

1. Simulation for Simple Tree Alternative for Scale and Location Differences.

The simulation study was conducted in SAS version 9.4, where the random samples of each population in this study for each distribution were generated following the SAS syntax in 5000 replications. A random seed was used to start the generation of the random samples obtained from each of the distributions. All tests in one situation were conducted using the exact same sets of random samples. The results of this simulation study are reproducible within the maximum standard error for estimating a proportion based on a 95% confidence interval and 5000 sets of samples, which in this case would be 0.014. Random samples were obtained from the following distributions: Normal, Exponential, Cauchy, Double Exponential, and Logistic.

The type 1 errors were estimated for each differing condition. A random seed number was used to generate samples for each different condition. Each test was conducted 5000 times based on each of the 5000 sets of random samples generated for that condition. Every test was conducted on the same set of samples for comparison purposes. The number of times each test statistic rejected the null hypothesis when it is true was counted and divided by 5000. This resulted in the estimated type 1 error. All of the tests were conducted at the stated significance level of 0.05. In this study, it is assumed that a test statistic maintains its type 1 error if the estimated type 1 error is between 0.04 and 0.06. The reason for this is that the results of this study are reproducible to within ± 0.014 with 95% confidence using the maximum standard error. The power results in this study are also reproducible to within ± 0.014 with 95% confidence using the maximum standard error.

The powers were estimated by counting the number of times the null hypothesis was rejected when it was false divided by 5000. The performance of each test was evaluated by comparing their respective powers for different distributions based on $k = 3$ populations of equal and unequal sample sizes consisting of 9, 18, and 36. Powers of tests were only compared for those tests that maintained their significance levels.

3. Results

3.1 Results for the normal, double exponential, and logistic distributions

Selected tables are given for the Normal, Double Exponential, and Logistic distributions. The tables (Tables 1-7) give results for equal samples of size 18 and various unequal samples of sizes 36, 18, and 9. Tables 1-7 are given in terms of the sample mean, except for Table 4 when the underlying distributions are Double Exponential. The reader can compare the results using the sample means with the sample medians in this case (Table 3 and Table 4). The first 3 tests will have different versions (using mean or median). The tests MW+AB and MW+WW will be the same. One can see that using the sample means gives slightly better results in the case of the Double Exponential. The sample means also gave slightly better results for both the Normal and Logistic distributions. When three populations had equal location and equal scale parameters, all tests maintained estimated significance levels around 0.05, as shown in the first row of Tables 1-7. When the populations had unequal location parameters but equal scale parameters, the MW+HL test showed the highest estimated powers compared to the other tests, as observed in the second and third rows of Tables 1, 3, 4, and 6 or lines 4-6 of Tables 2, 5, and 7.

When the populations had equal location parameters but unequal scale parameters the MW+CSR test had the highest estimated powers, as shown in the fourth and fifth rows of Tables 1, 3, 4, and 6 or lines 7-9 of Tables 2, 5, and 7. When both location and scale parameters were unequal, the MW+CSR test again demonstrated the highest estimated powers, as shown in the sixth and seventh rows of Tables 1, 3, 4, and 6 or lines 10-12 of Tables 2, 5, and 7.

Overall, under non-heavy tailed symmetric distributions, the MW+CSR (a newly proposed test) test showed stronger performance than the other tests across various scenarios. The scenarios included when the locations were different and the scales were the same, when the scales were different and the locations were the same, and when both location and scales were different. The researcher would not know ahead of time whether the location parameters differed, the scale parameters differed, or both differed. This test ended up having the highest powers considering all situations. The versions using the sample mean produced slightly higher estimated powers than the versions using the median.

3.2 Results for Cauchy distribution

The results of the simulation study of the tests for three treatments under the heavy-tailed Cauchy distribution are given in Tables 8-11. Under the null hypothesis (when all populations have equal location and scale parameters), most of the tests maintained significance levels close to 0.05 (see first line of Tables 8-11). However, MW+CSR exceeded the nominal significance level when using the sample mean. This indicates that the test does not properly control Type I error in that case. Therefore, the MW+CSR test will not be compared with other tests when using the sample mean (see first line of Tables 8 and 10). When the populations had unequal location parameters but equal scale parameters, MW+HL showed the highest estimated powers using the sample mean, while both the MW+CSR and MW+HL tests performed best when using the sample median, as observed in the second and third rows of Tables 8 and 9 and lines 4-6 of Tables 10-11. The MW+MD test performed well against the MW+CSR test when using the sample median as shown in Table 11, but the MW+CSR using the sample median had the highest estimated powers in more cases in this study when all the test versions were compared that maintained their significance levels.

3.3 Results for exponential distribution

The results of the simulation study of the tests for three treatments under the Exponential distribution are shown in Tables 12-15. For this skewed distribution, when the three populations had equal location and scale parameters, all tests maintained their estimated significance levels around 0.05, except the MW+MS and MW+CSR tests. These tests had estimated significance levels higher than the stated level in

all scenarios. The estimated significance levels exceeded 0.06. Therefore, the MW+MS and MW+CSR tests will not be compared with the other tests. When the populations had unequal location parameters but equal scale parameters, the MW+WW and the MW+HL tests showed the highest estimated powers, as observed in the second and third rows of Tables 12-13, and lines 4-6 of Tables 14-15 with the MW+WW having the highest powers in this scenario. (It is noted that the estimated significance level for the MW+WW test was slightly inflated for equal sample sizes of 18). However, the combined Mann Whitney and Hollander test (MW+HL) achieved the highest estimated powers for all other scenarios. Using the sample mean for Hollander did give somewhat higher powers, but the significance level for using the mean was slightly more inflated than the significance level when using the median. For smaller sample sizes (9 and 18), one may want to consider using the sample median instead of the sample mean.

4. Discussion and conclusions

In this research, the significance level was first estimated for each test under each

different scenario and each estimated significance level was compared to the stated significance level of 0.05. A test was considered to have maintained its significance level if the estimated significance level was between 0.04 and 0.06. Only tests that maintained their significance levels were compared.

The findings for this research among the six tests indicate that, across all symmetric distributions (Normal, Double Exponential, and Logistic) all tests maintained estimated significance values around 0.05. In terms of powers, the MW+CSR had the highest powers in

more cases. The powers were the highest when the sample mean was used.

For the heavy-tailed Cauchy distribution, under the null hypothesis where all populations had equal location and scale parameters, all tests maintained estimated significance levels close to 0.05, except for the second test (MW+CSR) when the sample mean was used. The MW+CSR test when the mean was used was not further compared to the other tests. Considering the remaining tests that maintained their significance levels, the MW+CSR (using median) would be recommended overall since it had the highest powers in more cases. The powers of the MW+CSR test were fairly close to the highest powers in other cases.

For the asymmetric distributions, the first two tests (MW+MS and MW+CSR) did not maintain their significance values. These two tests were not further compared to the other tests. Regarding powers of the remaining tests, when the populations had unequal scale parameters only or when both location and scale parameters were unequal, the third test (MW+HL) achieved the highest estimated powers. The MW+WW test did well for location changes only, but not in other situations. The MW+WW test also had an inflated significance level for equal samples of size 18. Overall, the MW+HL test is recommended since it had the highest powers in more cases, and the powers were fairly close to the highest in others.

Based on the results from the simulation study, the MW+CSR is recommended with symmetric populations. If the populations have heavy tails or if this is unknown, the sample median should be used. Otherwise, the sample mean should be used. For asymmetric distributions, the recommended test is the MW+HL with the sample mean or sample median for small sample sizes.

80-83.

Table 1. Estimated rejection percentages for $k = 3$ populations; normal distribution using the sample mean, $n_1 = n_2 = n_3 = 18$ and subsample size 3.

$n_1 = 2 = 3$	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
18	0	1	0	1	0	1	0.0476	0.0528	0.0550	0.0432	0.0426	0.0650*
18	0	1	1	1	1.25	1	0.7956	0.8086	0.9360	0.6782	0.6880	0.9388
18	0	1	1.25	1	1.75	1	0.9286	0.9414	0.9952	0.8698	0.8600	0.9948
18	0	1	0	2 ²	0	2.5 ²	0.4944	0.7146	0.6264	0.6852	0.6484	0.1836
18	0	1	0	2.5 ²	0	3 ²	0.6078	0.8376	0.7662	0.8112	0.7830	0.2812
18	0	1	0.25	2 ²	0.5	2.5 ²	0.7124	0.8718	0.8158	0.8570	0.8228	0.3508
18	0	1	0.5	2.5 ²	0.75	3 ²	0.8778	0.9710	0.9430	0.9600	0.9432	0.5666

*Slightly inflated

(continued on next page)

Table 2. Estimated rejection percentages for $k = 3$ populations under the normal distribution, using the sample mean, with varying sample sizes and a subsample size of 3 for Moses test.

n_1	n_2	n_3	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
36	36	18	0	1	0	1	0	1	0.0524	0.0592	0.0618*	0.0536	0.0546	0.0552
36	18	18	0	1	0	1	0	1	0.0550	0.0522	0.0564	0.0500	0.0518	0.0572
36	18	9	0	1	0	1	0	1	0.0502	0.0560	0.0648	0.0566	0.0566	0.0608
36	36	18	0	1	1	1	1.25	1	0.9454	0.9478	0.9924	0.9400	0.9348	0.9940
36	18	18	0	1	1	1	1.25	1	0.9290	0.9350	0.9906	0.9536	0.9464	0.9882
36	18	9	0	1	1	1	1.25	1	0.8602	0.8622	0.9704	0.9198	0.9136	0.9666
36	36	18	0	1	0	2 ²	0	2.5 ²	0.7090	0.8806	0.8584	0.8630	0.8192	0.2748
36	18	18	0	1	0	2 ²	0	2.5 ²	0.6496	0.8464	0.8224	0.8334	0.7740	0.3112
36	18	9	0	1	0	2 ²	0	2.5 ²	0.5590	0.7736	0.7366	0.7554	0.6790	0.2826
36	36	18	0	1	0.25	2 ²	0.5	2.5 ²	0.8914	0.9696	0.9612	0.9670	0.9466	0.5062
36	18	18	0	1	0.25	2 ²	0.5	2.5 ²	0.8522	0.9444	0.9312	0.9402	0.9106	0.5354
36	18	9	0	1	0.25	2 ²	0.5	2.5 ²	0.7546	0.8930	0.8716	0.8858	0.8364	0.4820

*Slightly inflated

Table 3. Estimated rejection percentages for $k = 3$ populations; double exponential distribution using the sample mean, $n_1 = n_2 = n_3 = 18$ and subsample size 3.

$n_1 = n_2 = n_3$	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
18	0	2(1) ²	0	2(1) ²	0	2(1) ²	0.0448	0.0502	0.0560	0.0470	0.0484	0.0582
18	0	2(1) ²	1	2(1) ²	1.25	2(1) ²	0.6682	0.6736	0.8084	0.5248	0.4946	0.8576
18	0	2(1) ²	1.25	2(1) ²	1.75	2(1) ²	0.8262	0.8376	0.9346	0.6720	0.6290	0.9668
18	0	2(1) ²	0	2(2) ²	0	2(2.5) ²	0.3824	0.5250	0.4320	0.4930	0.4742	0.1184
18	0	2(1) ²	0	2(2.5) ²	0	2(3) ²	0.5104	0.6974	0.5918	0.6600	0.6382	0.1716
18	0	2(1) ²	0.25	2(2) ²	0.5	2(2.5) ²	0.6028	0.7330	0.6516	0.7006	0.6698	0.2688
18	0	2(1) ²	0.5	2(2.5) ²	0.75	2(3) ²	0.8094	0.9082	0.8644	0.8916	0.8748	0.4426

(continued on next page)

Table 4. Estimated rejection percentages for $k = 3$ populations; double exponential distribution using the sample median, $n_1 = n_2 = n_3 = 18$ and subsample size 3.

$n_1 = 2 = 3$	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
18	0	$2(1)^2$	0	$2(1)^2$	0	$2(1)^2$	0.0440	0.0504	0.0506	0.0470	0.0484	0.0582
18	0	$2(1)^2$	1	$2(1)^2$	1.25	$2(1)^2$	0.6708	0.6950	0.7440	0.5248	0.4946	0.8576
18	0	$2(1)^2$	1.25	$2(1)^2$	1.75	$2(1)^2$	0.8280	0.8562	0.9028	0.6720	0.6290	0.9668
18	0	$2(1)^2$	0	$2(2)^2$	0	$2(2.5)^2$	0.3818	0.4968	0.4156	0.4930	0.4742	0.1184
18	0	$2(1)^2$	0	$2(2.5)^2$	0	$2(3)^2$	0.5058	0.6686	0.5770	0.6600	0.6382	0.1716
18	0	$2(1)^2$	0.25	$2(2)^2$	0.5	$2(2.5)^2$	0.5992	0.7046	0.6298	0.7006	0.6698	0.2688
18	0	$2(1)^2$	0.5		0.75	$2(3)^2$	0.8056	0.8990	0.8524	0.8916	0.8748	0.4426
				$2(2.5)^2$								

Table 5. Estimated rejection percentages for $k = 3$ populations under the double exponential distribution, using the sample mean, with varying sample sizes and a subsample size of 3 for Moses test.

n_1	n_2	n_3	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
36	36	18	0	$2(1)^2$	0	$2(1)^2$	0	$2(1)^2$	0.0512	0.0474	0.0548	0.0476	0.0456	0.0560
36	18	18	0	$2(1)^2$	0	$2(1)^2$	0	$2(1)^2$	0.0454	0.0530	0.0602	0.0504	0.0490	0.0554
36	18	9	0	$2(1)^2$	0	$2(1)^2$	0	$2(1)^2$	0.0426	0.0492	0.0524	0.0476	0.0502	0.0584
36	36	18	0	$2(1)^2$	1.25	$2(1)^2$	1.75	$2(1)^2$	0.9530	0.9602	0.9914	0.9266	0.9102	0.9954
36	18	18	0	$2(1)^2$	1.25	$2(1)^2$	1.75	$2(1)^2$	0.9498	0.9570	0.9934	0.9574	0.9530	0.9970
36	18	9	0	$2(1)^2$	1.25	$2(1)^2$	1.75	$2(1)^2$	0.8944	0.9066	0.9800	0.9364	0.9382	0.9790
36	36	18	0	$2(1)^2$	0	$2(2.5)^2$	0	$2(3)^2$	0.7320	0.8812	0.8488	0.8530	0.8270	0.2624
36	18	18	0	$2(1)^2$	0	$2(2.5)^2$	0	$2(3)^2$	0.6582	0.8396	0.7950	0.8040	0.7628	0.2638
36	18	9	0	$2(1)^2$	0	$2(2.5)^2$	0	$2(3)^2$	0.5742	0.7694	0.7152	0.7336	0.6868	0.2632
36	36	18	0	$2(1)^2$	0.5		0.75	$2(3)^2$	0.9546	0.9858	0.9794	0.9812	0.9760	0.6184
						$2(2.5)^2$								
36	18	18	0	$2(1)^2$	0.5		0.75	$2(3)^2$	0.9148	0.9736	0.9630	0.9690	0.9570	0.6186
						$2(2.5)^2$								
36	18	9	0	$2(1)^2$	0.5		0.75	$2(3)^2$	0.8494	0.9386	0.9160	0.9262	0.9038	0.5506
						$2(2.5)^2$								

Table 6. Estimated rejection percentages for $k = 3$ populations; logistic distribution using the sample mean, $n_1 = n_2 = n_3 = 18$ and subsample size 3.

$n_1 = 2 = 3$	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
18	0	$\frac{1^2\pi^2}{3}$	0	$\frac{1^2\pi^2}{3}$	0	$\frac{1^2\pi^2}{3}$	0.0510	0.0582	0.0612*	0.0504	0.0486	0.0588
18	0	$\frac{1^2\pi^2}{3}$	1	$\frac{1^2\pi^2}{3}$	1.25	$\frac{1^2\pi^2}{3}$	0.4616	0.4750	0.6064	0.3772	0.3792	0.6042
18	0	$\frac{1^2\pi^2}{3}$	1.25	$\frac{1^2\pi^2}{3}$	1.75	$\frac{1^2\pi^2}{3}$	0.6268	0.6518	0.7946	0.5124	0.5150	0.7950
18	0	$\frac{1^2\pi^2}{3}$	0	$\frac{2^2\pi^2}{3}$	0	$\frac{2.5^2\pi^2}{3}$	0.4364	0.6374	0.5486	0.6186	0.5852	0.1558
18	0	$\frac{1^2\pi^2}{3}$	0	$\frac{2.5^2\pi^2}{3}$	0	$\frac{3^2\pi^2}{3}$	0.5586	0.7786	0.6992	0.7472	0.7288	0.2370
18	0	$\frac{1^2\pi^2}{3}$	0.25	$\frac{2^2\pi^2}{3}$	0.5	$\frac{2.5^2\pi^2}{3}$	0.5912	0.7728	0.6902	0.7502	0.7188	0.2498

(continued on next page)

$n_1 = 2 = 3$	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
18	0	$\frac{1^2\pi^2}{3}$	0.5	$\frac{2.5^2\pi^2}{3}$	0.75	$\frac{3^2\pi^2}{3}$	0.7638	0.9110	0.8638	0.8972	0.8790	0.3976

*Slighted inflated

Table 7. Estimated rejection percentages for $k = 3$ populations under the logistic distribution, using the sample mean, with varying sample sizes and a subsample size of 3 for Moses test.

n_1	n_2	n_3	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
36	36	18	0	$\frac{1^2\pi^2}{3}$	0	$\frac{1^2\pi^2}{3}$	0	$\frac{1^2\pi^2}{3}$	0.0510	0.0526	0.0566	0.0504	0.0528	0.0582
36	18	18	0	$\frac{1^2\pi^2}{3}$	0	$\frac{1^2\pi^2}{3}$	0	$\frac{1^2\pi^2}{3}$	0.0466	0.0486	0.0550	0.0470	0.0478	0.0544
36	18	9	0	$\frac{1^2\pi^2}{3}$	0	$\frac{1^2\pi^2}{3}$	0	$\frac{1^2\pi^2}{3}$	0.0492	0.0486	0.0558	0.0494	0.0502	0.0542
36	36	18	0	$\frac{1^2\pi^2}{3}$	1.25	$\frac{1^2\pi^2}{3}$	1.75	$\frac{1^2\pi^2}{3}$	0.8154	0.8256	0.9236	0.7868	0.7830	0.9308
36	18	18	0	$\frac{1^2\pi^2}{3}$	1.25	$\frac{1^2\pi^2}{3}$	1.75	$\frac{1^2\pi^2}{3}$	0.7834	0.7970	0.9166	0.8058	0.8004	0.9136
36	18	9	0	$\frac{1^2\pi^2}{3}$	1.25	$\frac{1^2\pi^2}{3}$	1.75	$\frac{1^2\pi^2}{3}$	0.6874	0.7010	0.8436	0.7462	0.7446	0.8354
36	36	18	0	$\frac{1^2\pi^2}{3}$	0	$\frac{2.5^2\pi^2}{3}$	0	$\frac{3^2\pi^2}{3}$	0.7918	0.9402	0.9290	0.9218	0.8990	0.3700
36	18	18	0	$\frac{1^2\pi^2}{3}$	0	$\frac{2.5^2\pi^2}{3}$	0	$\frac{3^2\pi^2}{3}$	0.7262	0.9072	0.8914	0.8896	0.8514	0.3886
36	18	9	0	$\frac{1^2\pi^2}{3}$	0	$\frac{2.5^2\pi^2}{3}$	0	$\frac{3^2\pi^2}{3}$	0.6496	0.8562	0.8326	0.8354	0.7796	0.3696
36	36	18	0	$\frac{1^2\pi^2}{3}$	0.5	$\frac{2.5^2\pi^2}{3}$	0.75	$\frac{3^2\pi^2}{3}$	0.9384	0.9900	0.9842	0.9850	0.9774	0.6012
36	18	18	0	$\frac{1^2\pi^2}{3}$	0.5	$\frac{2.5^2\pi^2}{3}$	0.75	$\frac{3^2\pi^2}{3}$	0.8832	0.9698	0.9644	0.9636	0.9450	0.5996
36	18	9	0	$\frac{1^2\pi^2}{3}$	0.5	$\frac{2.5^2\pi^2}{3}$	0.75	$\frac{3^2\pi^2}{3}$	0.8100	0.9300	0.9178	0.9242	0.8916	0.5362

Table 8. Estimated rejection percentages for $k = 3$ populations; Cauchy distribution using the sample mean, $n_1 = n_2 = n_3 = 18$ and subsample size 3.

$n_1 = 2 = 3$	μ_1	γ_1	μ_2	γ_2	μ_3	γ_3	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
18	0	γ	0	γ	0	γ	0.0474	0.2220	0.0534	0.0420	0.0450	0.0584
18	0	γ	1	γ	2	γ	0.5090		0.6574	0.4284	0.4038	0.7640
18	0	γ	1.25	γ	2.75	γ	0.6440		0.7804	0.5300	0.4858	0.8904
18	0	γ	0	4γ	0	5γ	0.4554		0.6160	0.6736	0.7296	0.2936

(continued on next page)

$n_1 = n_2 = n_3$	μ_1	γ_1	μ_2	γ_2	μ_3	γ_3	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
18	0	γ	0	5γ	0	6γ	0.5136		0.7056	0.7582	0.7992	0.3800
18	0	γ	0.5	3γ	1	3γ	0.5354		0.6538	0.6934	0.7300	0.3300
18	0	γ	1	3γ	2	5γ	0.7278		0.8406	0.8656	0.8834	0.5440

Table 9. Estimated rejection percentages for $k = 3$ populations; Cauchy* distribution using the sample median, $n_1 = n_2 = n_3 = 18$ and subsample size 3.

$n_1 = n_2 = n_3$	μ_1	γ_1	μ_2	γ_2	μ_3	γ_3	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
18	0	γ	0	γ	0	γ	0.0450	0.0468	0.0486	0.0420	0.0450	0.0584
18	0	γ	1	γ	2	γ	0.5114	0.5896	0.5802	0.4284	0.4038	0.7640
18	0	γ	1.25	γ	2.75	γ	0.6448	0.7486	0.7198	0.5300	0.4858	0.8904
18	0	γ	0	4γ	0	5γ	0.4456	0.6870	0.6030	0.6736	0.7296	0.2936
18	0	γ	0	5γ	0	6γ	0.5044	0.7648	0.6906	0.7582	0.7992	0.3800
18	0	γ	0.5	3γ	1	3γ	0.5306	0.7158	0.6360	0.6934	0.7300	0.3300
18	0	γ	1	3γ	2	5γ	0.7204	0.8824	0.8240	0.8656	0.8834	0.5440

μ_i and γ_i represent location and scale shifts. The Cauchy distribution does not have a mean or variance.

Table 10. Estimated rejection percentages for $k = 3$ populations under the Cauchy* distribution, using the sample mean, with varying sample sizes and a subsample size of 3 for Moses test.

n_1	n_2	n_3	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
36	36	18	0	γ	0	γ	0	γ	0.0426	0.2608	0.0482	0.0426	0.0416	0.0580
36	18	18	0	γ	0	γ	0	γ	0.0442		0.0520	0.0460	0.0474	0.0580
36	18	9	0	γ	0	γ	0	γ	0.0480		0.0552	0.0488	0.0484	0.0572

(continued on next page)

n_1	n_2	n_3	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
36	36	18	0	γ	1.25	γ	2.75	γ	0.8074		0.9058	0.7582	0.7330	0.9592
36	18	18	0	γ	1.25	γ	2.75	γ	0.8030		0.9280	0.8260	0.8280	0.9696
36	18	9	0	γ	1.25	γ	2.75	γ	0.6904		0.8636	0.7712	0.7852	0.9064
36	36	18	0	γ	0	5γ	0	6γ	0.7518		0.9384	0.9330	0.9528	0.5354
36	18	18	0	γ	0	5γ	0	6γ	0.6784		0.9098	0.9050	0.9174	0.5080
36	18	9	0	γ	0	5γ	0	6γ	0.5952		0.8580	0.8630	0.8730	0.4422
36	36	18	0	γ	1	3γ	2	5γ	0.9000		0.9710	0.9764	0.9826	0.7110
36	18	18	0	γ	1	3γ	2	5γ	0.8614		0.9626	0.9694	0.9714	0.6890
36	18	9	0	γ	1	3γ	2	5γ	0.7624		0.9078	0.9204	0.9208	0.5818

* μ_i and σ_i^2 represent location and scale shifts. A mean and variance does not exist for the Cauchy

Table 11. Estimated rejection percentages for $k = 3$ populations under the Cauchy* distribution, using the sample median, with varying sample sizes and a subsample size of 3 for Moses test.

n_1	n_2	n_3	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
36	36	18	0	γ	0	γ	0	γ	0.0434	0.0456	0.0422	0.0426	0.0416	0.0580
36	18	18	0	γ	0	γ	0	γ	0.0444	0.0442	0.0466	0.0460	0.0474	0.0580
36	18	9	0	γ	0	γ	0	γ	0.0484	0.0454	0.0502	0.0488	0.0484	0.0572
36	36	18	0	γ	1.25	γ	2.75	γ	0.8074	0.8868	0.8434	0.7582	0.7330	0.9592
36	18	18	0	γ	1.25	γ	2.75	γ	0.8034	0.8938	0.8838	0.8260	0.8280	0.9696
36	18	9	0	γ	1.25	γ	2.75	γ	0.6886	0.7930	0.8124	0.7712	0.7852	0.9064
36	36	18	0	γ	0	5γ	0	6γ	0.7382	0.9420	0.9352	0.9330	0.9528	0.5354

(continued on next page)

n_1	n_2	n_3	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
36	18	18	0	γ	0	5γ	0	6γ	0.6664	0.9118	0.9066	0.9050	0.9174	0.5080
36	18	9	0	γ	0	5γ	0	6γ	0.5852	0.8636	0.8544	0.8630	0.8730	0.4422
36	36	18	0	γ	1	3γ	2	5γ	0.8966	0.9792	0.9664	0.9764	0.9826	0.7110
36	18	18	0	γ	1	3γ	2	5γ	0.8552	0.9710	0.9588	0.9694	0.9714	0.6890
36	18	9	0	γ	1	3γ	2	5γ	0.7556	0.9194	0.9006	0.9204	0.9208	0.5818

* μ_i and σ_i^2 represent location and scale shifts. A mean and variance does not exist for the Cauchy

Table 12. Estimated rejection percentages for $k = 3$ populations; exponential distribution using the sample mean, $n_1 = n_2 = n_3 = 18$ and sub-sample size.

$n_1 = 2 = 3$	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
18	1	1	1	1	1	1	0.0950	0.1756	0.0626*	0.0518	0.0542	0.0652*
18	1	1	2.25	1	2.5	1			0.8764	0.2264	0.2030	0.9998
18	1	1	2.5	1	3	1			0.9434	0.3374	0.2790	0.9998
18	1	1	1		1				0.5290	0.4760	0.5424	0.3046
				3.5^2		6^2						
18	1	1	1		1				0.5460	0.4856	0.5540	0.3210
				3.75^2		6.5^2						
18	1	1	3		5				0.9694	0.8956	0.8868	0.8746
				3^2		5^2						
18	1	1	3.5		6				0.9922	0.9520	0.9468	0.9288
				3.5^2		6^2						

*Slightly inflated

Table 13. Estimated rejection percentages for $k = 3$ populations; exponential distribution using the sample median, $n_1 = n_2 = n_3 = 18$ and sub-sample size 3.

$n_1 = 2 = 3$	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
18	1	1	1	1	1	1	0.0942	0.1430	0.0562	0.0518	0.0542	0.0652 *
18	1	1	2.25	1	2.5	1			0.7750	0.2264	0.2030	0.9998
18	1	1	2.5	1	3	1			0.8924	0.3374	0.2790	0.9998
18	1	1	1		1				0.5144	0.4760	0.5424	0.3046
				3.5^2		6^2						
18	1	1	1		1				0.5316	0.4856	0.5540	0.3210
				3.75^2		6.5^2						

(continued on next page)

$n_1 = 2 = 3$	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
18	1	1	3	3^2	5	5^2			0.9626	0.8956	0.8868	0.8746
18	1	1	3.5	3.5^2	6	6^2			0.9868	0.9520	9468. 9468.	9288. 9288.

*Slightly inflated

Table 14. Estimated rejection percentages for $k = 3$ populations under the exponential distribution, using the sample mean, with varying sample sizes and a subsample size of 3 for Moses test.

n_1	n_2	n_3	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
36	36	18	1	1	1	1	1	1	0.0912	0.1764	0.0558	0.0496	0.0494	0.0554
36	18	18	1	1	1	1	1	1			0.0596	0.0504	0.0490	0.0554
36	18	9	1	1	1	1	1	1			0.0598	0.0532	0.0516	0.0580
36	36	18	1	1	2.25	1	2.5	1			0.9776	0.6020	0.5186	1.0000
36	18	18	1	1	2.25	1	2.5	1			0.9730	0.7648	0.6844	1.0000
36	18	9	1	1	2.25	1	2.5	1			0.9686	0.8372	0.7692	1.0000
36	36	18	1	1	1		1	5^2			0.8686	0.6988	0.7294	0.4360
						3^2								
36	18	18	1	1	1		1	5^2			0.8130	0.6852	0.6742	0.4300
						3^2								
36	18	9	1	1	1		1	5^2			0.7746	0.6688	0.6276	0.4366
						3^2								
36	36	18	1	1	3		5	5^2			0.9998	0.9966	0.9948	0.9814
						3^2								
36	18	18	1	1	3		5	5^2			0.9948	0.9890	0.9842	0.9662
						3^2								
36	18	9	1	1	3		5	5^2			0.9894	0.9830	0.9776	0.9438
						3^2								

Table 15. Estimated rejection percentages for $k = 3$ populations under the exponential distribution, using the sample median, with varying sample sizes and a subsample size of 3 for Moses test.

n_1	n_2	n_3	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
36	36	18	1	1	1	1	1	1	0.0892	0.1406	0.0504	0.0496	0.0494	0.0554
36	18	18	1	1	1	1	1	1			0.0530	0.0504	0.0490	0.0554
36	18	9	1	1	1	1	1	1			0.0544	0.0532	0.0516	0.0580
36	36	18	1	1	2.25	1	2.5	1			0.9164	0.6020	0.5186	1.0000
36	18	18	1	1	2.25	1	2.5	1			0.9156	0.7648	0.6844	1.0000
36	18	9	1	1	2.25	1	2.5	1			0.9246	0.8372	0.7692	1.0000
36	36	18	1	1	1		1	5^2			0.8580	0.6988	0.7294	0.4360
						3^2								

(continued on next page)

n_1	n_2	n_3	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	MW+MS	MW+CSR	MW+HL	MW+MD	MW+AB	MW+WW
36	18	18	1	1	1		1				0.8056	0.6852	0.6742	0.4300
						3^2		5^2						
36	18	9	1	1	1		1			0.7636		0.6688	0.6276	0.4366
						3^2		5^2						
36	36	18	1	1	3		5			0.9994		0.9966	0.9948	0.9814
						3^2		5^2						
36	18	18	1	1	3		5			0.9930		0.9890	0.9842	0.9662
						3^2		5^2						
36	18	9	1	1	3		5			0.9872		0.9830	0.9776	9438.
						3^2		5^2						9438.

References

- Hussain A. An Improved Lepage-type test for the location-scale problem. *Communications in Statistics-Theory & Methods*. 2026;55(5):1537-1555.
- Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*. 1947;18(1):50-60.
- Wilcoxon F. Individual Comparisons by Ranking Methods. *Biometrics*. 1945;8.
- Ansari AR, Bradley RA. Rank-Sum tests for Dispersions. *The Annals of Mathematical Statistics*. 1960;31(4):1174-1189.
- Huang F, Xia ZQ, Guo LD, Yang FC. Effects of reservoirs on seasonal discharge of Irtysh River measured by LePage Test. *Water Science & Engineering*. 2014;7(4):363-372.
- Marozzi M. The LePage Location-Scale Test Revisited. *Far East Journal of Theoretical Statistics*. 2008;24(2):137-155.
- Rublik F. The Multisample Version of LePage Test. *Kybernetika*. 2005;41(6):713-733.
- Alsubie A, Magel R. Proposed Nonparametric Tests for the Simple Tree Alternative for Location and Scale Testing. *International Journal of Statistics and Applied Mathematics*. 2020;5(2):26-32.
- Moses LE. Rank tests of dispersion. *Annals of Mathematical Statistics*. 1963;34:973-983.
- Azman NA, Ahad NA, Okwonu FZ. Comparative analysis of classical and modified Moses' test. *Journal of Statistical Modeling and Analytics*. 2021;3(1):47-54.
- Daniel WW. *Applied Nonparametric Statistics* (Second Edition). Boston, MA: Cole Cengage Learning; 2016.
- Alsubie A, Magel R. Proposed nonparametric tests using Moses test for location and scale testing. *Journal of Progressive Research in Mathematics*. 2020;16(2):2877-2887.
- Conover WJ. *Practical nonparametric statistics* (3rd ed.). John Wiley & Sons; 1999.
- Hollander M. A nonparametric test for the two-sample problem. *Psychometrika*. 1963;28(4):395-403.
- Gogoi P, Gogoi B. Performance of two sample scale tests: An empirical study. *International Advanced Research Journal in Science, Engineering and Technology*. 2016;3(11):7.
- Wald A, Wolfowitz J. On a Test Whether Two Samples are from the same Population. *Annals of Mathematical Statistics*. 1940;11:147-162.
- Nan Y, Magel R. Nonparametric Tests for the Simple Tree Alternative on Both the Mean and Variance. *International Journal of Statistics and Applied Mathematics*. 2025;10(2):61-69.
- Lepage Y. A Combination of Wilcoxon's and Ansari-Bradley's Statistics. *Biometrika*. 1971;58(1):213-217.
- Dampney D, Magel R. Analysis of LePage's test applied to the Simple Tree Alternative. *International Journal of Statistics and Applied Mathematics*. 2026;11(3):97-102.
- Bailer AJ. *Statistical programming in SAS*. Cary, NC: SAS Institute Inc.; 2010.
- Brown MB, Forsythe AB. Robust tests for the equality of variances. *Journal of the American Statistical Association*. 1974;69(346):364-367.
- Shukr B, Magel R. Nonparametric tests for testing the control versus effects of treatment levels. *International Journal of Statistics and Applied Mathematics*. 2024;9(6):185-190.

How to cite this article: Alessa S, Magel R. Comparison of non-parametric tests for the simple tree alternative. *International Journal of Statistics and Applied Mathematics*. 2026;11(10):89-102. doi: <https://www.doi.org/10.22271/mathsjournal.com>