Research Article GU J Sci 32(2): 718-735 (2019)

Gazi University
Journal of Science
http://dergipark.gov.tr/gujs

Test Statistic for Ordered Alternatives based on Wilcoxon Signed Rank

Hamza GAMGAM[1], Bulent ALTUNKAYNAK[1,*]

[1] *Gazi University, Department of Statistics, 06500, Ankara, Turkey*

Article Info

Received: 18/10/2018
Accepted: 14/02/2019

Keywords

Ordered alternatives
Wilcoxon signed rank
Power
Simulation

Abstract

This paper proposes a test statistic for ordered alternatives based on the Wilcoxon signed rank statistic. One of the classical tests, Jonckheere-Terpstra's *J* test, and the *R* test suggested by Chen et al. were used for type I error rate and power comparisons. For data generated from the normal distribution, all of the tests gave type I error rates close to nominal alpha. When the data were generated from chi-square distribution, the proposed *G* test and *J* test for type I error gave better results than the *R* test, but the error rates of the *J* test for Student's *t* distribution are better than those of the others. Power results of simulation study for normal distributions showed that the proposed *G* test was superior to all other considered tests. The *G* and *J* tests for the data generated from Student's t distributions performed well. When the data were generated from chi-square distributions, the proposed *G* test is more powerful than the others. The simulation showed that the *R* test was inferior to the other tests for all cases.

1. INTRODUCTION

Let $X_{i1}, X_{i2}, ..., X_{in_i}$, $i=1,...,k$ be random independent samples with size n_i from k populations with continuous cumulative distribution function $F_i(x)=F((x-\theta_i)/\sigma_i)$, where $-\infty<\theta_i<+\infty$ and $\sigma_i>0$ are location and scale parameters, respectively. The null hypothesis to identify whether the populations have common continuous cumulative distribution function can be expressed as

$$H_0: F_1(x)=F_2(x)=\cdots=F_k(x) \qquad \forall x. \tag{1}$$

A number of test statistics have been proposed to test the null hypothesis in equation 1 under certain assumptions and for different forms of H_1. To test the null hypothesis, $H_0:\mu_1=\cdots=\mu_k$, against H_1: *At least one of the means is different*, *F* statistic with $k-1$ and $n-k$ degrees of freedom is used under distribution normality and variance homogeneity. In above degree of freedom, n is sum of the sample sizes. Kruskal and Wallis [1] and Bhapkar [2] proposed some test statistics to test the null hypothesis in equation 1 against H_1: *At least one of the distribution functions* $F_i(x)$ *is different* under distributions non-normality. Without variance homogeneity, Bishop [3], Bishop and Dudewicz [4, 5], Chen and Chen [6], Chen [7], Gamage et al. [8], Lee and Ahn [9], Rice and Gaines [10], Weerahandi [11], Xu and Wang [12] proposed different test statistics to test H_0 against H_1.

The ordered alternative states that the distributions are stochastically ordered, i.e.,

$$H_1: F_1(x)\geq F_2(x)\geq ...\geq F_k(x) \quad \exists x: F_1(x)>F_k(x)\ . \tag{2}$$

Under H_1, X_i tends to be smaller than X_{i+1}, $i=1,2,...,k-1$, since $F_i(x)\geq F_{i+1}(x)$ implies that $P(X_i \leq X_{i+1})\geq 1/2$. For the special case of the location model, Equation 2 is equivalent to

*Corresponding author, e-mail: bulenta@gazi.edu.tr

$$H_1:\theta_1 \le \theta_2 \le ... \le \theta_k \quad (\theta_1 < \theta_k) \tag{3}$$

[13]. Similarly, the ordered alternative hypothesis

$$H_1: F_1(x) \le F_2(x) \le ... \le F_k(x) \;\; \exists x: F_1(x) < F_k(x) \tag{4}$$

states that X_i tends to be larger than X_{i+1}, $i = 1, 2, ..., k-1$, since $F_i(x) \le F_{i+1}(x)$ implies that $P(X_i \ge X_{i+1}) \ge 1/2$ under H_1 given in equation 4. For the location model, equation 4 is equivalent to

$$H_1:\theta_1 \ge \theta_2 \ge ... \ge \theta_k \quad (\theta_1 > \theta_k)\,. \tag{5}$$

To test H_0 against the ordered alternative hypothesis, Terpstra [14] and Jonckheere [15] independently proposed the same test. Subsequently, Bartholomew [16], Chacko [17], Puri [18], Odeh [19], Archambault et al. [20], Hettmansperger and Norton [21], Beier and Buning [22], Neuhauser et al. [23], Chen et al. [24], Shan et al. [25] and Gaur [26] studied the same subject.

This paper proposed a new test statistic. We compare its performance with the R and *J* tests for type I error rate and power. Section 2 describes Jonckheere-Terpstra's *J* test (*J* test), one of the classical tests, and the *R* test suggested by Chen et al. [24] using both to test the null hypothesis against the ordered alternative hypothesis. Section 3 proposes a test statistic called *G* test based upon the Wilcoxon signed rank statistic. Section 4 provides simulation results for type I error rate and power comparisons obtained for *G*, *J*, and *R* tests. As application, the data given by Jonckheere [15] are analyzed in section 5. Finally, section 6 concludes the paper.

2. CHANG'S *R* AND JONCKHEERE-TERPSTRA'S *J* TESTS

The *R* statistic suggested by Chen et al. [24] is based upon single stage sampling. Let $X_{i1}, X_{i2}, ..., X_{in_i}$, $i = 1, 2, ..., k$, be independent samples from *k* populations with mean μ_i and variance σ_i^2, size $n_i \ge 3$, without variance homogeneity. Let $2 \le n_0 < n_i$ for each sample. So, the sample mean and sample variance for *i*th sample with n_0 observation can be expressed as

$$\bar{X}_i = \frac{1}{n_0}\sum_{j=1}^{n_0} X_{ij} \text{ and } S_i^2 = \frac{1}{n_0 - 1}\sum_{j=1}^{n_0}(X_{ij} - \bar{X}_i)^2\,, \tag{6}$$

respectively. Let

$$z^* = \max\left\{\frac{S_1^2}{n_1}, \frac{S_2^2}{n_2}, ..., \frac{S_k^2}{n_k}\right\}, \tag{7}$$

and calculate the weights for observations in the *i*th sample as

$$u_i = \frac{1}{n_i} + \frac{1}{n_i}\sqrt{\frac{n_i - n_0}{n_0}(n_i z^*/S_i^2 - 1)} \tag{8}$$

and

$$v_i = \frac{1}{n_i} - \frac{1}{n_i}\sqrt{\frac{n_0}{n_i - n_0}(n_i z^*/S_i^2 - 1)}\,. \tag{9}$$

Then the final weighted sample mean using all observations is

$$\tilde{X}_i = \sum_{j=1}^{n_0} u_i X_{ij} + \sum_{j=n_0+1}^{n} v_i X_{ij} , \qquad i = 1,2,...,k , \tag{10}$$

where $\tilde{X}_i$ has conditional normal distribution and the transformations

$$\frac{\tilde{X}_i - \mu_i}{\sqrt{z*}} = t_i , \ i = 1,2,...,k \tag{11}$$

have i.i.d Student's *t* distributions with $v = n_0 - 1$ degrees of freedom.

Chen et al. [24] proposed a test statistic based on one stage sampling to test the equality of means against $H_1 : \mu_1 \geq \mu_2 \geq ... \geq \mu_k$ with at least one strict inequality. Assume

$$\tilde{U} = \max_{1 \leq r \leq k} \frac{1}{r} \sum_{i=1}^{r} \frac{\tilde{X}_i}{\sqrt{z*}} \text{ and } \tilde{V} = \min_{1 \leq r \leq k} \frac{1}{k-r+1} \sum_{i=r}^{k} \frac{\tilde{X}_i}{\sqrt{z*}} , \tag{12}$$

where $r = 1,...,k$. Then the *R* test statistic can be expressed as

$$R = \tilde{U} - \tilde{V} \tag{13}$$

to test H_0 against the ordered alternative. Under H_0, the null distribution of *R* is

$$\tilde{Q} = \max_{1 \leq r \leq k} \frac{1}{r} \sum_{i=1}^{r} t_i - \max_{1 \leq r \leq k} \frac{1}{k-r+1} \sum_{i=r}^{k} t_i . \tag{14}$$

From the distribution of $\tilde{Q}$, when $P(\tilde{Q} > q_{\alpha,k,v}) = \alpha$, if $R_h > q_{\alpha,k,v}$, then H_0 is rejected; where R_h represents the value calculated from the sample, and $q_{\alpha,k,v}$ is the critical value. Chen et al. [24] calculated $q_{\alpha,k,v}$ for various combinations of v and k , using Monte Carlo simulation. Therefore, this paper also calculates critical $q_{\alpha,k,v}$ values using Monte Carlo simulation.

Let $X_{i1}, X_{i2}, ..., X_{in_i}$ be independent samples with size n_i from *k* populations and continuous cumulative distribution function $F_i(x)$, $i = 1,2,...,k$. For samples i' and *i*, the Mann-Whitney *U* statistic is defined as

$$U_{ii'} = \sum_{i'=1}^{n_{i'}} \sum_{i=1}^{n_i} D_{ii'} , \tag{15}$$

where Terpstra [14] and Jonckheere [15] proposed same test statistic based on the sum of Mann-Whitney statistics to test H0 against the ordered alternative

$$D_{ii'} = \begin{cases} 1, & X_{i'n_{i'}} > X_{in_i} \qquad i' = 1,2,...,n_{i'} \\ 0, & X_{i'n_{i'}} < X_{in_i} \qquad i = 1,2,...,n_i \end{cases} \tag{16}$$

[27, p.140]. Terpstra [14] and Jonckheere [15] proposed the same test statistic based on the sum of independent Mann-Whitney statistics to test H_0 against the ordered alternative as

$$J = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} U_{ij} . \tag{17}$$

Then, if

$$J_h > J'_\alpha , \tag{18}$$

where J_h is the value calculated from the sample and J'_α is the critical value for *J*, the H_0 against ordered alternative is rejected.

When $k = 3$, $n_1, n_2, n_3 = 2,3,...,8$, critical values derived from the exact *J* distributions can be found in existing tables [28, p.455-458]. However, the tables were very limited with respect to population and sample sizes, and were extended somewhat by Bucchianico [29] and Wiel [30]. Therefore, we employ the approach suggested by Wiel [30] in Section 4 to calculate critical values for *J*.

3. PROPOSED TEST STATISTIC

Let the i'th and $(i+i')$th samples be combined and ranked where $i = 1,...,k-1$ and $i' = 1,...,k-i$. For the $(i+i')$th sample, let the ranks of observations be represented by $R_{j(i+i')}$, where $j = 1,...,n_{(i+i')}$, and define

$$\bar{R}_t = \frac{\sum_{j=1}^{n_{(i+i')}} R_{j(i+i')}}{n_{(i+i')}} , \qquad t = 1,...,c , \tag{19}$$

where $\bar{R}_t$ is the mean of ranks for the $(i+i')$th sample, and $c = k(k-1)/2$. Algorithm 3.1 shows the process to obtain $\bar{R}_t$.

Algorithm 3.1: For a given *k* and $X_{i1}, X_{i2},..., X_{in_i}$, starting with initial value $t = 1$.

For $i = 1$ to $k-1$
    For $i' = 1$ to $k-i$
        The i'th and $(i+i')$th samples are combined and ranked
        Compute $\bar{R}_t$ from equation 19
        $t = t + 1$
    (end i' loop)
(end *i* loop)

Let the random variable D_l be

$$D_l = \bar{R}_t - \bar{R}_{t'} ,\ l = 1,...,d ;\ t = 1,...,c-1;\ t' = t+1,...,c , \tag{20}$$

where $d = c(c-1)/2$. If H_1 from equation 2 is true, then $\theta_i \le \theta_{i'}$, and the expectation of $\bar{R}_t$ is less than or equal to the expectation of $\bar{R}_{t'}$ for each $t = 1,...,c-1$; $t' = t+1,...,c$. Similarly, when H_1 from equation 4 is true, the expectation of $\bar{R}_t$ is greater than or equal to the expectation of $\bar{R}_{t'}$ for each $t = 1,...,c-1$; $t' = t+1,...,c$. Accordingly, for H_1 from equation 2, the random variable Z_l can be defined as

$$Z_l = \begin{cases} 1, & D_l < 0 \\ 0, & D_l > 0 \end{cases} , \tag{21}$$

and for H_1 from equation 4,

$$Z_l = \begin{cases} 1, & D_l > 0 \\ 0, & D_l < 0 \end{cases}. \tag{22}$$

The proposed test statistic to test the null hypothesis against the ordered alternative can be expressed as

$$G = \sum_{l=1}^{d} Z_l r(|D_l|), \tag{23}$$

where $r(|D_l|)$ is the rank order statistic for $|D_l|$. The G statistic takes values from the set $0,1,2,...,d(d+1)/2$. Although G is expected to have a value close or equal to $d(d+1)/4$ under H_0, it is expected to have larger values under H_1 from equation 2 or 4.

Substituting d for n in $E(T^+)$ and $V(T^+)$ of the Wilcoxon signed rank statistic T^+, the expected value and variance of G can be expressed as

$$E(G) = \frac{d(d+1)}{4} \tag{24}$$

and

$$V(G) = \frac{d(d+1)(2d+1)}{24} \tag{25}$$

[28, p.34]. In the definition of G, Z is a random variable, but $r(|D_l|)$, $l = 1,...,d$, are fixed numbers. Hence, the random variables are not independent, and the covariance between Z_t and $Z_{t'}$ is

$$Cov(Z_t, Z_{t'}) = \frac{-n_t n_{t'}}{n^2(n-1)}, \; t \neq t', \tag{26}$$

where $t = 1,...,c-1$, $t' = t+1,...,c$ and $n_t + n_{t'} = n$ [27, p.153]. The covariance terms effect on the variance of G is very small and tends to approximately zero as sample size increases. Therefore, we did not consider the effect on the variance of G.

The distribution of G is symmetrical around its own mean. Since the Wilcoxon signed rank statistic converges to the normal distribution, the critical values for G can be found through

$$\frac{G \pm 0.5 - E(G)}{(V(G))^{1/2}} = Z \sim N(0,1). \tag{27}$$

If there are tied values for $|D_l|$, the mean of the respective ranks is assigned to $r(|D_l|)$. In that case, the expected value of G remains the same, but the variance decreases. Thus, a correction term is defined as

$$CT = \frac{\sum_{i=1}^{s} t_i^3 - \sum_{i=1}^{s} t_i}{48}, \tag{28}$$

where s is the number of the groups with tied values in $|D_l|$, and t_i is the number of those in the ith group [28, p.35]. *CT* is used in the approximation Equation 27,

$$\frac{G \pm 0.5 - E(G)}{(V(G) - CT)^{1/2}} = Z \sim N(0,1)\ . \tag{29}$$

4. SIMULATION

The *J*, *G*, and *R* tests were compared for type I error rate and power. Nominal alpha and iteration number were 0.05 and 10,000 respectively, sample size $n_i = 3,4,5,6,7,8,9,10$, and population $k = 4,5,6,7,8,9,10$.

For comparison, the data were generated from normal, Student's t, and Chi-square distributions to evaluate test statistic the performances for light tailed symmetric, heavy tailed symmetric, and right skewed populations, respectively. The selected parameters for the simulation scenarios are shown in Table 1.

Table 1. *Simulation scenario parameters*

	Sample			
Scenario	1	2	…	k
1	$N(100,10^2)$	$N(100,10^2)$	$\cdots$	$N(100,10^2)$
2	$N(100,10^2)$	$N(101,10^2)$	$\cdots$	$N(99+k,10^2)$
3	$N(100,10^2)$	$N(101,12^2)$	$\cdots$	$N(99+k,(8+2k)^2)$
4	t_5	t_5	$\cdots$	t_5
5	t_5	t_5+1	$\cdots$	t_5+k-1
6	t_{4+k}	$t_{3+k}+1$	$\cdots$	t_5+k-1
7	χ^2_5	χ^2_5	$\cdots$	χ^2_5
8	χ^2_5	χ^2_6	$\cdots$	χ^2_{4+k}

Scenarios 1, 4, and 7 were used to compare type I error rates under H_0. Scenarios 2 and 5 were used to compare power values when distribution means uniformly increased under H_1. Scenarios 3, 6, and 8 were used to compare power values when distribution means and variances uniformly increased under H_1. For scenarios 5 and 6, a constant that increased with increasing sample number was added to sample observations. Because the mean of Student's t distribution is zero, some constants as $1,...,k-1$ for power comparisons are added to the relevant observations so that the locations of the observations generated from this distribution are larger than zero.

Figures 1 and 2 show that type I error rates for all tests were close to nominal levels for all values of k and n when the data were generated from normal distributions under H_0.

Figure 3 shows type I error rates for the *J* test are generally closer to nominal compared to the other tests for all n and k when the data were generated from Student's t distribution under H_0.

Figure 4 shows that although type I error rates of the proposed test are not similar to nominal, these rates are approximately 0.06 when the nominal level = 0.05. Close type I error rates and nominal levels are considered sufficient if the type I error rates are within plus or minus 10% and 50% of nominal [31, 32].

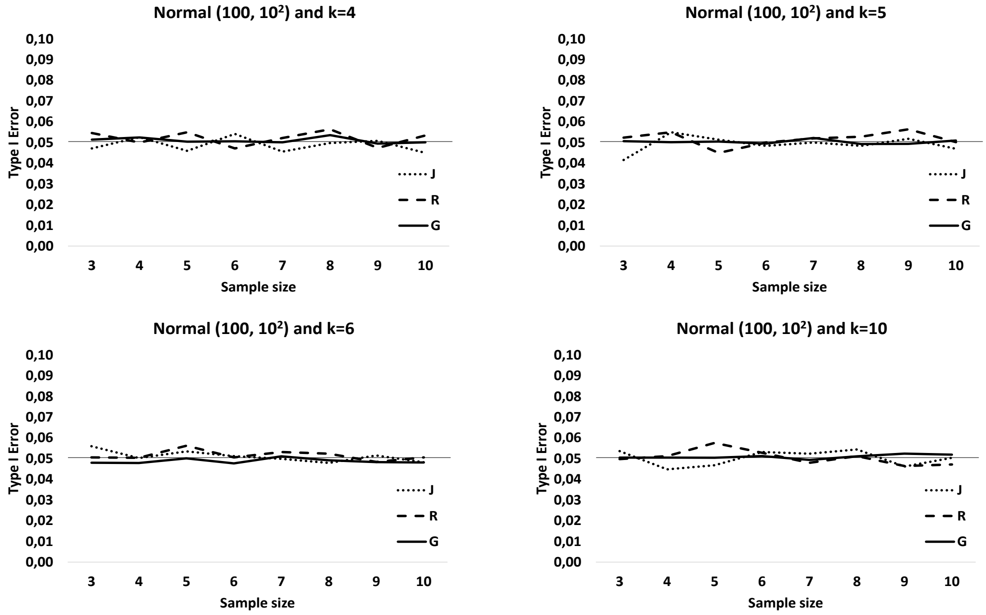


Figure 1. *Type I error rates for some k values for data generated from normal distributions with scenario1*

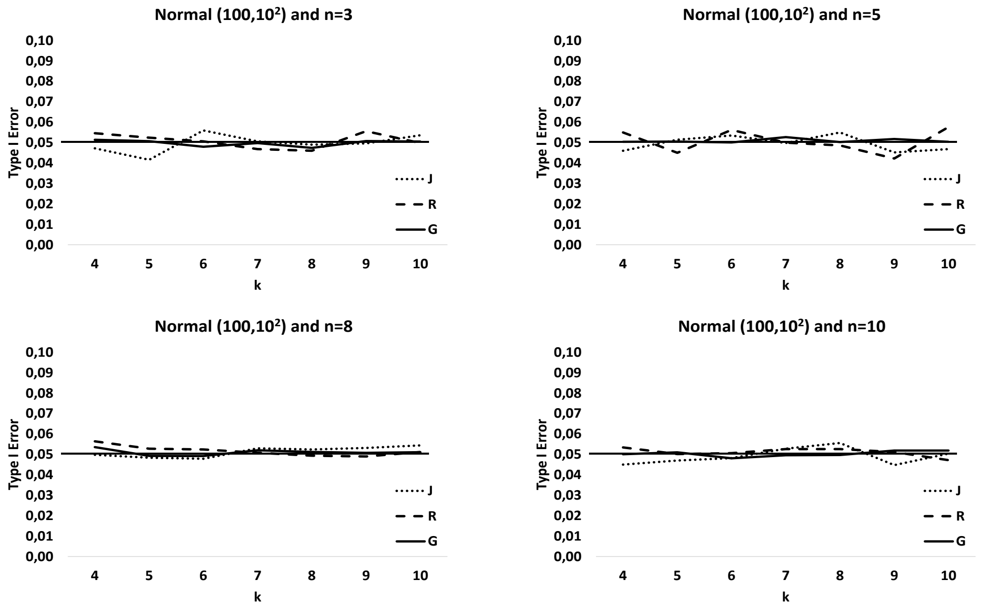


Figure 2. *Type I error rates for some n values for data generated from normal distributions with scenario1*

Figures 5 and 6 show that type I error rates for the *R* test are not close to nominal, being approximately 0.07 for data generated from Chi-square distributions. On the other hand, type I error rates for the proposed and *J* tests are close to nominal as *k* increases. Although the rates for the *R* and *J* tests are generally close to nominal, *J* tests rates are closer to nominal for small *k* and *n* values.

Figures 7 and 8 show that power values for the proposed test are superior to those for the *J* and *R* tests for all *k* and *n* values for data generated from normal distributions where the mean increases with increasing sample size.

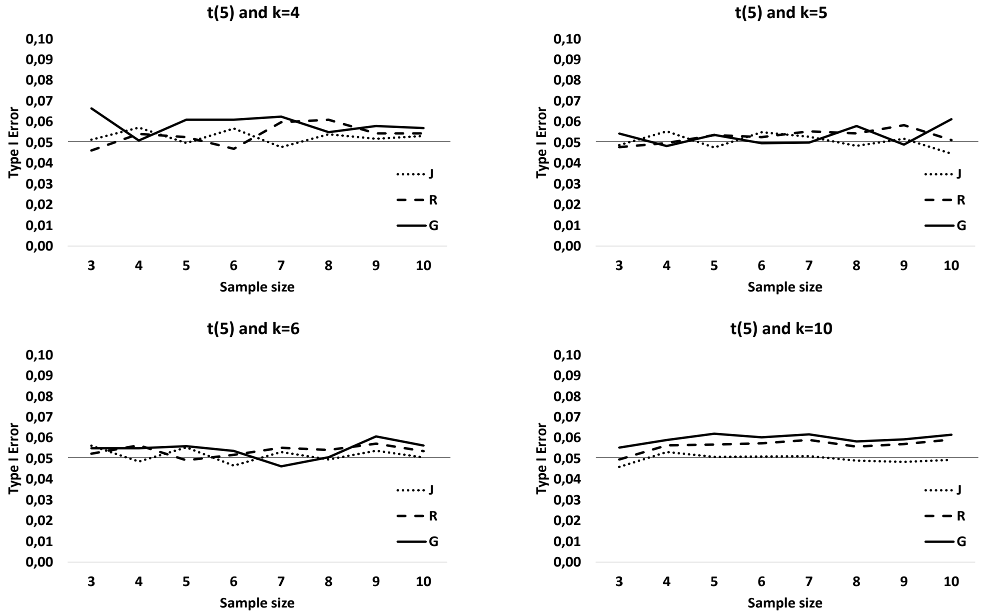


Figure 3. *Type I error rates for some k values for data generated from Student's t distributions with scenario 4*

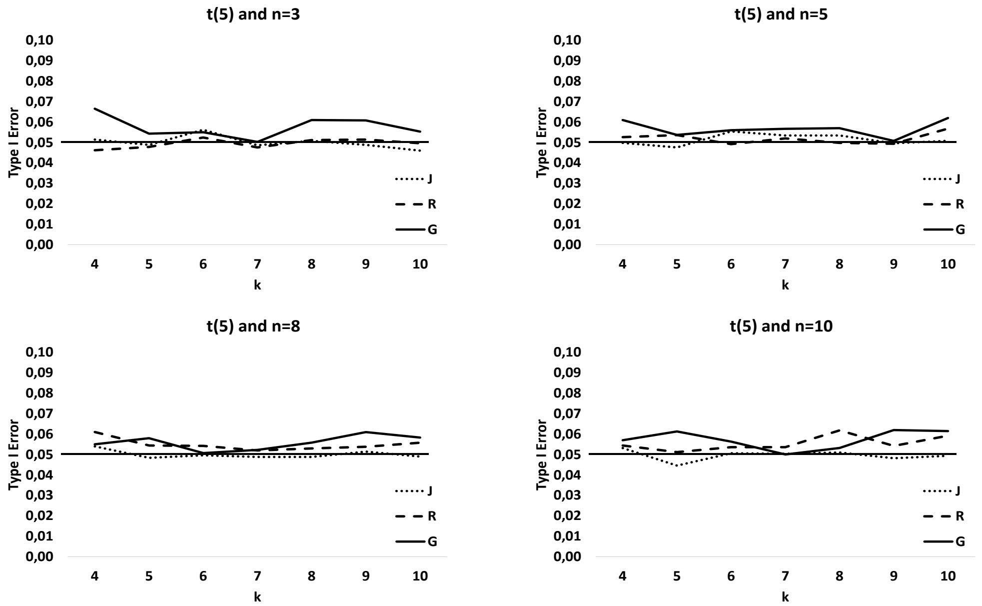


Figure 4. *Type I error rates for some n values for data generated from Student's t distributions with scenario 4*

Figures 9 and 10 show that power values for the proposed and *J* tests are almost similar and close to 1 when a constant that increased with increasing sample number was added to sample observations. Power values for the *R* test are always inferior to the other tests.

Figures 11 and 12 show that the proposed test is more powerful than the *J* and *R* tests for all *k* and *n* values for data generated from normal distributions whose mean and variance increase as the sample number increases.

Figures 13 and 14 show that power values of the proposed and *J* test are almost similar and close to 1 for data from Student's t distributions; power values for the test *R* are always inferior.

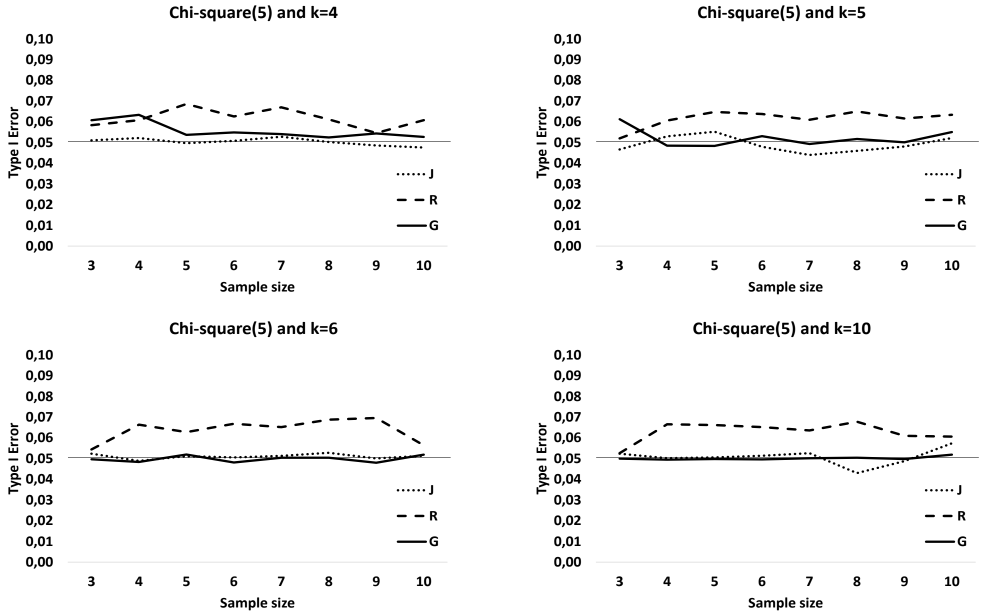


Figure 5. *Type I error rates for some k values for data generated from Chi-square distributions with scenario 7*

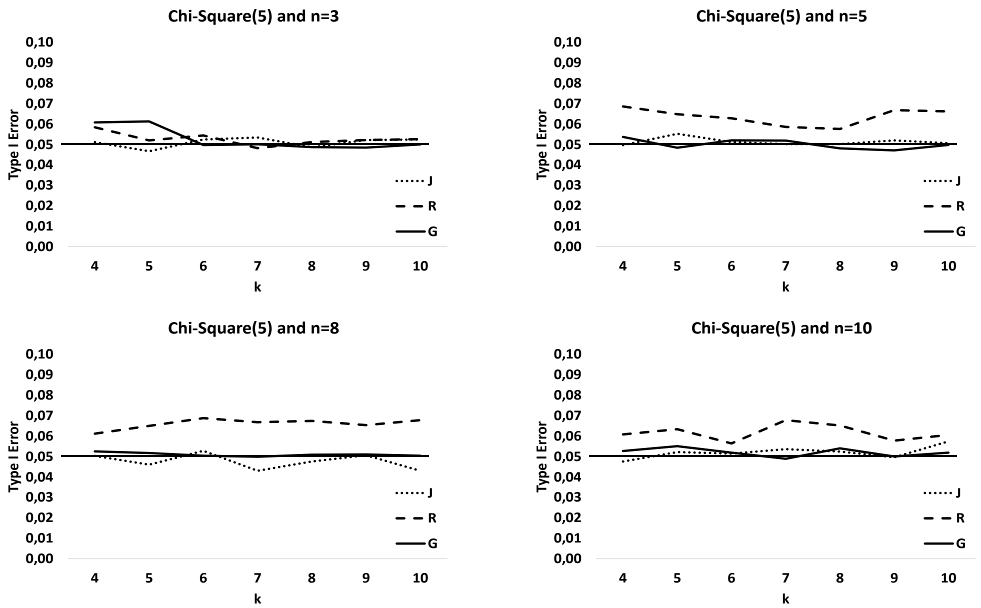


Figure 6. *Type I error rates for some n values for data generated from Chi-square distributions with scenario 7*

Figures 15 and 16 show that power values of the proposed test are superior to other tests for all k values except $k=4$ for data generated from Chi-square distributions whose mean and variance increase as the sample number increases.

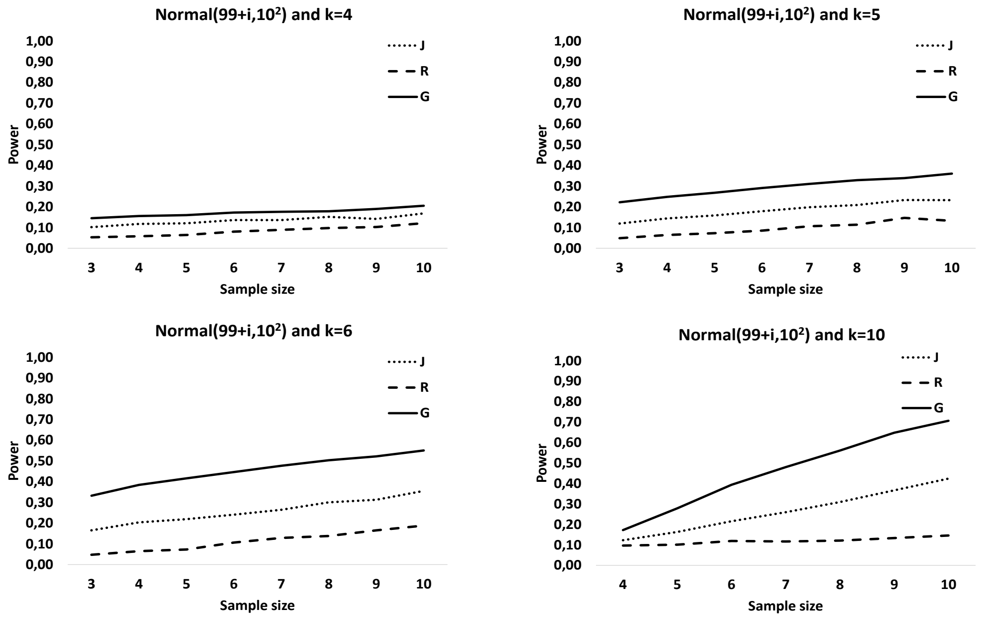


Figure 7. *Power values for some k values for data generated from normal distributions with scenario 2*

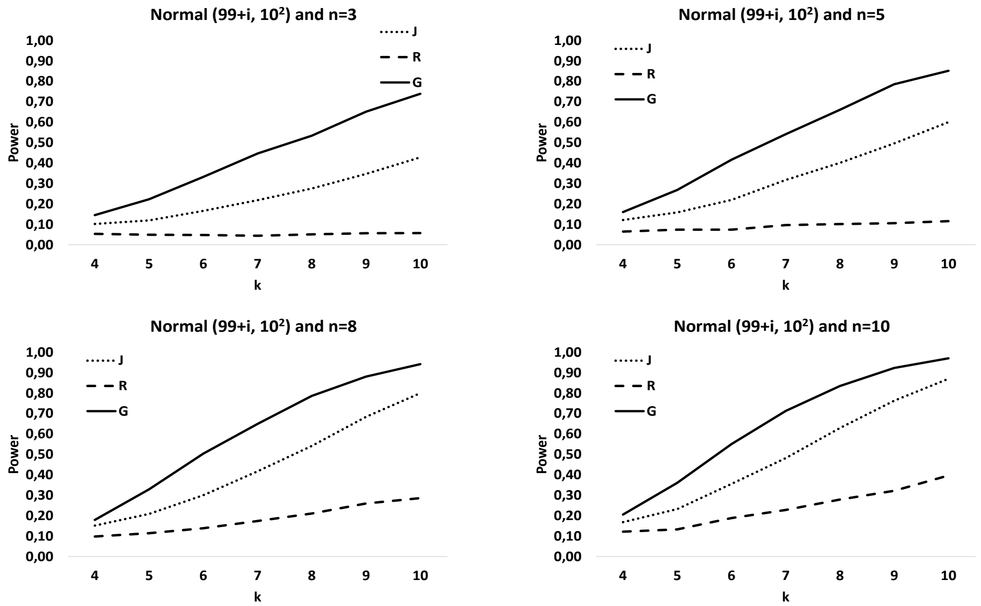


Figure 8. *Power values for some n values for data generated from normal distributions with scenario 2*

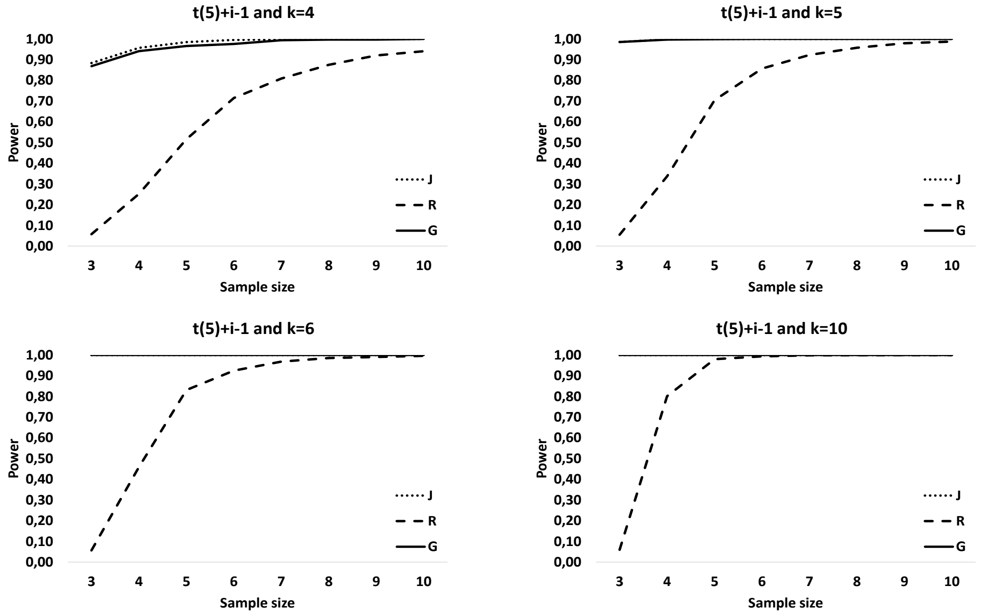


Figure 9. *Power values for some k values when a constant that increased with increasing sample number was added to sample observations for the data generated from Student's t distribution with scenario 5*

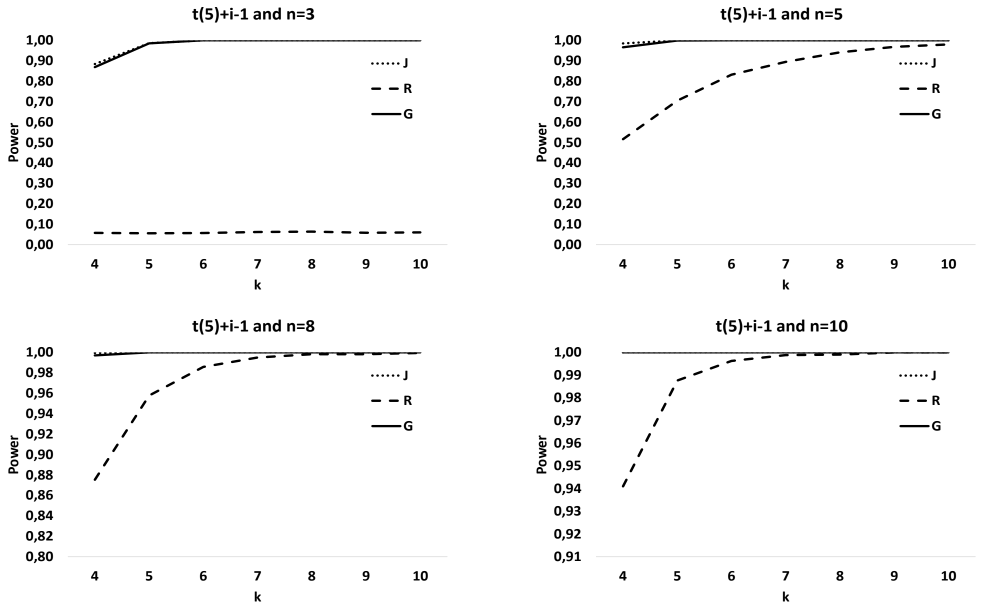
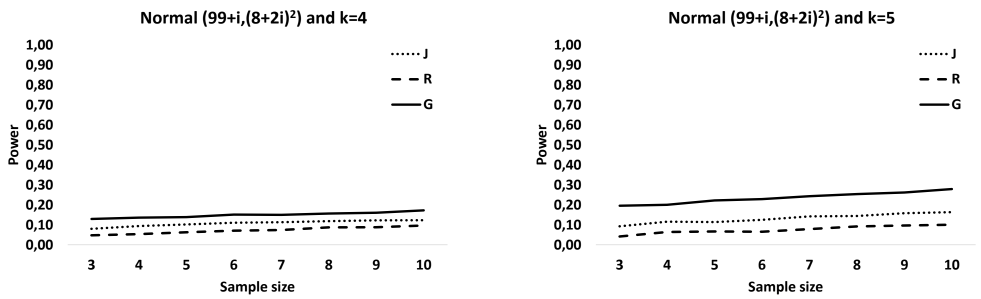


Figure 10. *Power values for some n values when a constant that increased with increasing sample number was added to sample observations for the data generated from Student's t distribution with scenario 5*



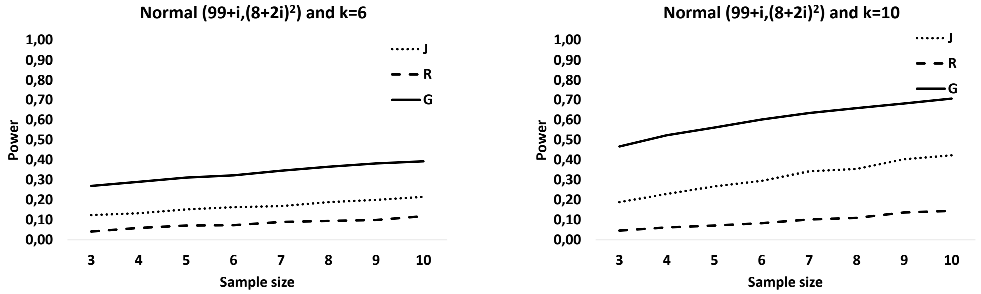


Figure 11. *Power values for some k values for data generated from normal distributions with scenario 3*

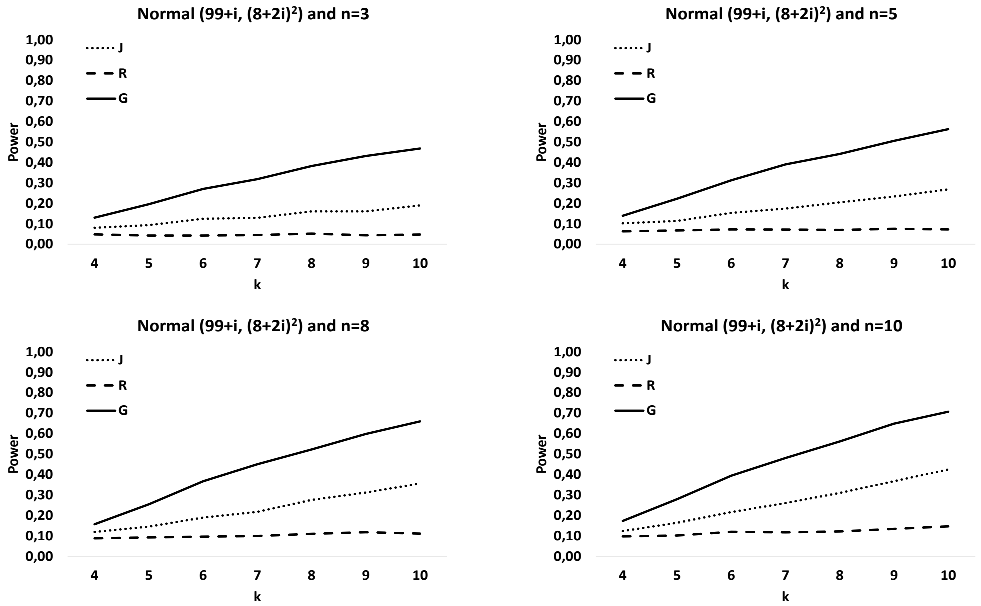
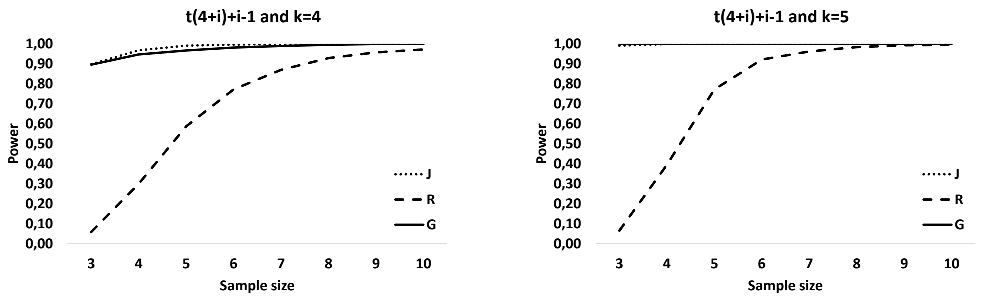


Figure 12. *Power values for some n values for data generated from normal distributions with scenario 3*



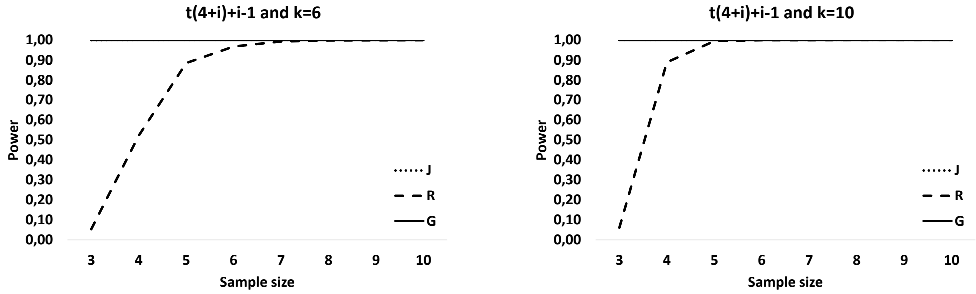


Figure 13. *Power values for some k values when a constant that increased with increasing sample number was added to a sample observations generated from Student's t distributions with scenario 6*

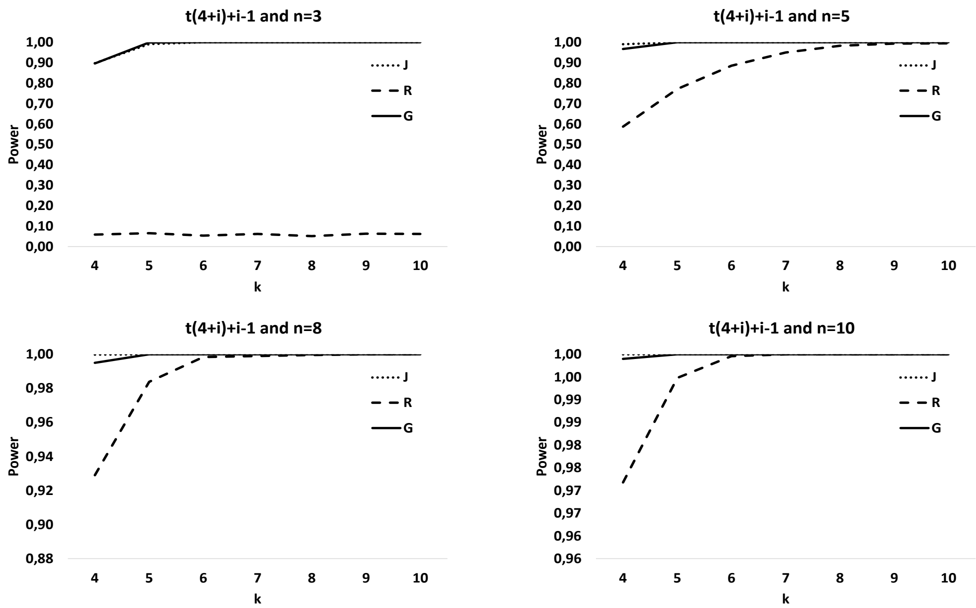
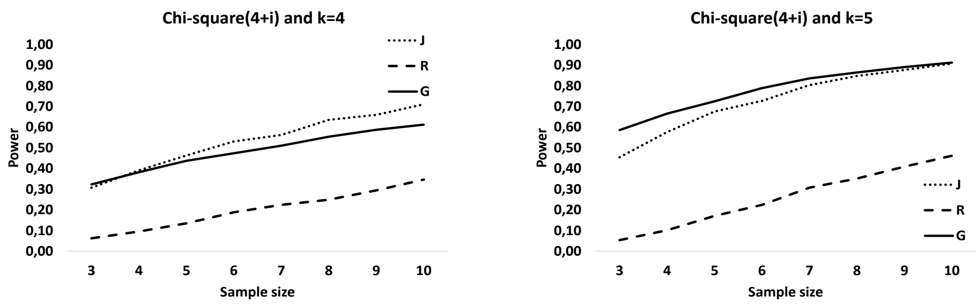


Figure 14. *Power values for some n values when a constant that increased with increasing sample number was added to a sample observations generated from Student's t distributions with scenario 6*



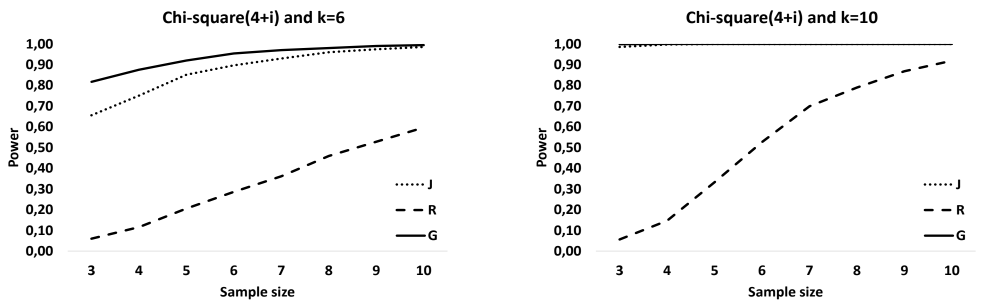


Figure 15. *Power values for some k values for data generated from Chi-square distributions with scenario 8*

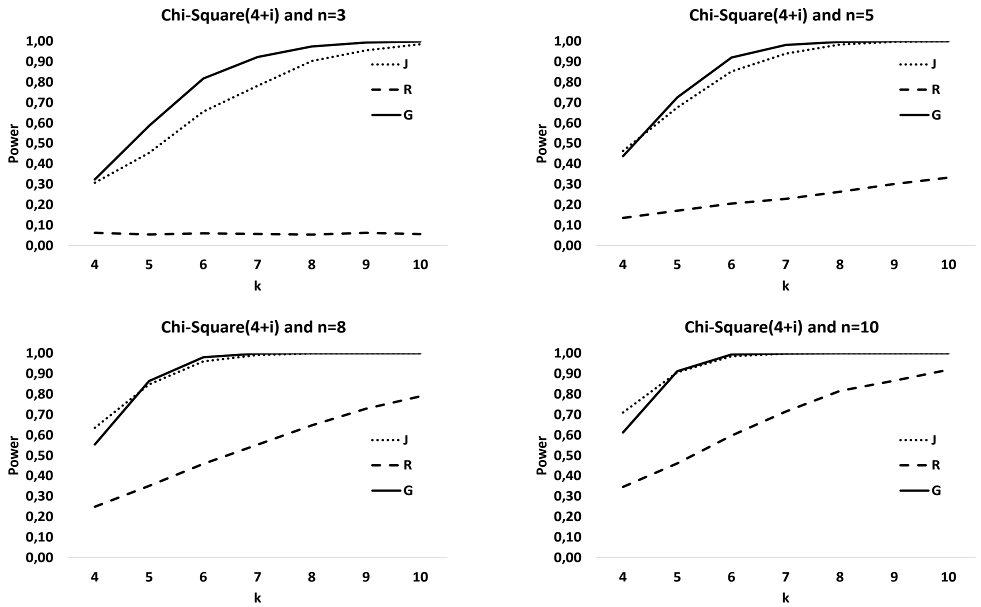


Figure 16. *Power values for some n values for data generated from Chi-square with scenario 8*

5. NUMERIC EXAMPLE

As an application, data used by Jonckheere [15] was considered to illustrate the performances of the proposed G test and the J test. In the simulation study in section 4, the R test has been found to be quite poor in terms of power compared to the other tests. Therefore, the R test was not included in this application. This data given by Jonckheere [15] is shown below.

Table 2. *Jonckheere's Data*

Sample			
1	2	3	4
19 20 60 130	21 61 80 129	40 99 100 149	49 110 151 160

Because of $k = 4$, the value of c is obtained as $k(k-1)/2 = 4(3)/2 = 6$. By using equation 19, the values of the $\overline{R}_t$ statistic are calculated as follows.

Table 3. *The values of the* $\bar{R}_t$ *statistic*

t	i	i'	$i+i'$	$\bar{R}_t$
1	1	1	2	5.25
2	1	2	3	5.25
3	1	3	4	5.50
4	2	1	3	5.50
5	2	2	4	5.50
6	3	1	4	5.75

In the above table, for example, common ranks are first assigned to the observations from samples 1 and 2, then the $\bar{R}_1 = 5.25$ shows the mean of the ranks corresponding to the observations in sample 2. Because the value of c is 6, d is calculated as $c(c-1)/2 = 6(5)/2 = 15$. For these cases, the values of the D_l, Z_l, $r(|D_l|)$ and $Z_l r(|D_l|)$ statistics are obtained as follows for $l = 1,...,d$.

Table 4. *The values of the* D_l, Z_l, $r(|D_l|)$ *and* $Z_l r(|D_l|)$ *statistics*

l	t	t'	D_l	Z_l	$\|D_l\|$	$r(\|D_l\|)$	$Z_l r(\|D_l\|)$
1	1	2	0	-	-	-	-
2	1	3	-0.25	1	0.25	5	5
3	1	4	-0.25	1	0.25	5	5
4	1	5	-0.25	1	0.25	5	5
5	1	6	-0.50	1	0.50	10.5	10.5
6	2	3	-0.25	1	0.25	5	5
7	2	4	-0.25	1	0.25	5	5
8	2	5	-0.25	1	0.25	5	5
9	2	6	-0.50	1	0.50	10.5	10.5
10	3	4	0	-	-	-	-
11	3	5	0	-	-	-	-
12	3	6	-0.25	1	0.25	5	5
13	4	5	0	-	-	-	-
14	4	6	-0.25	1	0.25	5	5
15	5	6	-0.25	1	0.25	5	5
							66

Based on the results above, the value of the proposed G statistic using Equation 23 is calculated as 66. As shown in the table above, when the values of D_l are examined, it is seen that 4 of them are zero. As is known, the relevant observations that cause this situation are removed from the data set. Since there are 4 observations that cause this cases, the value of d is taken as 11, not 15.

In addition, there are tied values among the $|D_l|$ values. For this reason, the mean ranks are assigned to them. By taking by $t_1 = 9$ and $t_2 = 2$, the correction term given in Equation 28 is calculated as follows.

$$CT = \frac{(9^3 + 2^3) - (9+2)}{48} = 15.125 \tag{29}$$

The expected value and variance of the proposed statistic by Equations 24 and 25 are found as follows, respectively.

$$E(G) = \frac{11(12)}{4} = 33 \tag{30}$$

and

$$V(G) = \frac{11(12)(23)}{24} = 126.5 \tag{31}$$

Using these results, the following value for standard normal Z variable by Equation 29 is calculated.

$$z = \frac{66 - 0.5 - 33}{(126.5 - 15.125)^{1/2}} = 3.174 \tag{32}$$

Finally, we find that $p\text{-}value = P(Z > 3.174)$ is 0.001035. On the other hand, the *p-value* of the *J* test for the above data by Jonckheere [15] was obtained as 0.0168. These results support the simulation findings. When the significance level is taken as 0.01, the null hypothesis is rejected by the proposed *G* test, but this hypothesis is not rejected by the *J* test.

6. CONCLUSION

We proposed a new *G* test based on the Wilcoxon test for ordered alternative hypothesis and compared the *G* test with *J* and *R* tests proposed by Jonckheere [15] and Chen et al. [24], respectively. Simulation results for type I error rate showed that the proposed *G* test and *J* test are generally similar with type I error rates close to nominal for data generated from normal and Chi-square distributions. On the other hand, *J* test type I error rates are close to nominal for data generated from Student's *t* distributions. However, the proposed *G* and *R* test type 1 error rates were approximately 0.06. Since this result is within the acceptable range, they may be considered sufficiently close to nominal type I error rates [31,32].

Power simulations showed that the proposed *G* test was superior to all other considered tests for normal distributions. Adding constants that increase as sample number increases for data generated from Student's *t* distributions showed that the proposed *G* test was still superior, although very similar to the *J* test. Two other Student's *t* distribution scenarios showed the proposed *G* and *J* tests performed well but the *R* showed inferior performance in terms of power.

When data were generated from Chi-square distributions, the proposed *G* test was more powerful than the *J* test, except for $k = 4$, and the *R* test was inferior for all *k*.
Thus, the proposed *G* test produces generally good type I error rates and power for all considered distributions. We think that the proposed test may be applied to real data in future studies.

ACKNOWLEDGMENTS

The authors are grateful to the reviewers for their suggestions and valuable comments on the paper.

CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

[1] Kruskal, W.H., Wallis, W.A., "Use of ranks in one-criterion variance analysis", J Amer Statist Assoc., 47: 583–621, (1952).

[2] Bhapkar, V.P., "A nonparametric test for the problem of several samples", Annals of Mathematical Statistics, 32: 1108–1117, (1961).

[3] Bishop, T.A., "Heteroscedastic ANOVA, MANOVA and multiple comparisons", Phd.Thesis, The Ohio State University, Ohio, 15-35 (1976).

[4] Bishop, T.A., Dudewicz, E.J., "Exact analysis of variance with unequal variances:test procedures and tables", Technometrics, 20: 419–430, (1978).

[5] Bishop, T.A., Dudewicz, E.J. "Heteroscedastic ANOVA", Sankhya, 43: 40–57, (1981).

[6] Chen, S., Chen, J.H., "Single-stage analysis of variance under heteroscedasticity", Comm Statist Simulation and Computation. 27(3): 641–666, (1998).

[7] Chen, S., "One-stage and two-stage statistical inference under heteroscedasticity", Comm Statist Simulation and Computation, 30(4): 991–1009, (2001).

[8] Gamage, J., Mathew, T., Weerahandi, S., "Generalized p-values and generalized confidence regions for the multivariate Behrens–Fisher problem and MANOVA", J Multivariate Anal., 88: 177–189, 2004.

[9] Lee, S., Ahn, C.H., "Modified ANOVA for unequal variances", Comm Statist Simulation and Computation, 32: 987–1004, (2003).

[10] Rice, W.R, Gaines, S.D., "One-way analysis of variance with unequal variances", Proc Nat Acad Sci., 86: 8183–8184, (1989).

[11] Weerahandi, S., "ANOVA under unequal error variances", Biometrics, 51: 589–599, (1995).

[12] Xu, L., Wang, S.A., "A new generalized p-value for ANOVA under heteroscedasticity", Statistics & Probability Letters, 78(8): 963–969, (2008).

[13] Terpstra, T.J., Chang, C.H., Magel, R.C., "On the use of Spearman's correlation coefficient for testing ordered alternatives", Journal of Statistical Computation and Simulation, 81(11): 1381–1392, (2011).

[14] Terpstra, T.J., "The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking", Indigationes Mathematicae, 14: 327-333, (1952).

[15] Jonckheere, A.R., "A distribution-free k-sample test against ordered alternatives", Biometrika, 41: 133–145, (1954).

[16] Bartholomew, D.J., "Ordered tests in the analysis of variance", Biometrika, 48: 325–332, (1961).

[17] Chacko, V.J., "Testing homogeneity against ordered alternatives", Ann Math Statist., 34: 945–956, (1963).

[18] Puri, M.L., "Some distribution-free k-sample rank tests of homogeneity against ordered alternatives", Comm Pure Appl Math., 18: 51–63, (1965).

[19] Odeh, R.E., “On Jonckheere's k-sample test against ordered alternatives”, Technometrics, 13: 912–918, (1971).

[20] Archambault, W.A.T., Mack, G.A., Wolfe, D.A., “K-sample rank tests using pair-specific scoring functions”, Canadian Journal of Statistics, 5: 195–207, (1977).

[21] Hettmansperger, T.P., Norton, R.M., “Tests for patterned alternatives in k-sample problems”, J Amer Statist Assoc., 82: 292–299, (1987).

[22] Beier, F., Buning, H., “An adaptive test against ordered alternatives”, Computational Statistics & Data Analysis, 25(4): 441-452, (1997).

[23] Neuhauser, M., Liu, P.Y., Hothorn, L.A., “Nonparametric test for trend: Jonckheere’s test, a modification and a maximum test”, Biometrical Journal, 40: 899–909, (1998).

[24] Chen, S., Chen, J.H., Chang, H.F., “A one-stage procedure for testing homogeneity of means against an ordered alternative under unequal variances”, Comm Statist Simulation and Computation, 33(1): 49–67, (2004).

[25] Shan, G., Young, D., Kang, L., “A new powerful nonparametric rank test for ordered alternative problem”, Plos One, 9(11): 1–10, (2014).

[26] Gaur, A., “A class of k-sample distribution-free test for location against ordered alternatives”, Comm Stat Theory and Methods, 46(5): 2343–2353, (2017).

[27] Gibbons, J.D., Nonparametric Statistical Inference, McGraw-Hill, New York, (1971).

[28] Daniel, W.W., Applied Nonparametric Statistics, Houghton Mifflin, Boston, (1978).

[29] Bucchianico, A.D., “Computer algebra, combinatorics, and the Wilcoxon-Mann-Whitney statistic”, J Stat Plan Inf., 79: 349–364, (1999).

[30] Wiel, M.A., “Exact distributions of distribution-free test statistics” PhD.Thesis, Eindhoven University of Technology, The Netherlands, 27-43 (2000).

[31] Bradley, J.V., “Robustness?”, Br J Math Stat Psychol. 31: 144–152, (1978).

[32] Fagerland, M.W., Sandvik, L., “Performance of five two-sample location tests for skewed distributions with unequal variances”, Contemporary Clinical Trials, 30: 490-496, (2009).