Journal of Statistical and Econometric Methods, vol.2, no.4, 2013, 81-98
ISSN: 2241-0384 (print), 2241-0376 (online)
Scienpress Ltd, 2013

Comparison of the Powers of the Kolmogorov-Smirnov Two-Sample Test and the Mann-Whitney Test for Different Kurtosis and Skewness Coefficients Using the Monte Carlo Simulation Method

M.Suphi Özçomak[1], Mahmut Kartal[2],

Ötüken Senger[3] and Ali Kemal Çelik[4]

Abstract

This study aims to compare the statistical powers of the Kolmogorov-Smirnov two-sample test and the Mann-Whitney test using the Monte Carlo simulation method, for specific sample sizes. The simulation results showed that the Mann-Whitney Test was more powerful for (5, 10) and (10, 5) sample sizes when the standard deviation rates were 2 and 1/2; for the (5, 20) sample size when the standard deviation rates were 2, 3, and 1/2; and for the (20, 5) sample size when the standard deviation rates were 1/2, 1/3, and 1/4. The Kolmogorov-Smirnov two-sample test was more powerful for (5, 10) and (10, 5) sample sizes when standard deviation rates were 3, 4, and 1/4; for the (10, 20) sample size for all standard deviation rates; for the (20, 10) sample size excepting the standard deviation of 1/2; and for the (20, 5) sample size when standard deviations were 2, 3, and 4.

---

[1] Atatürk University, Department of Econometrics.
[2] Cumhuriyet University, Department of Management Information Systems.
[3] Kafkas University, Deparment of Numerical Methods.
[4] Atatürk University, Department of Business Administration.

Article Info: *Received* : September 2, 2013. *Revised* : October 3, 2013
Published online : December 1, 2013

Mathematics Subject Classification: 62G08, 62GXX
Keywords: Kolmogorov-Smirnov two-sample test, Mann-Whitney test, Statistical power, Skewness, Kurtosis, Monte Carlo simulation

1 Introduction

Two-sample statistical comparison is one of the most important hypothesis tests used in the social sciences. Researchers are presented with a wide range of different parametric and non-parametric statistical procedures, varying according to the population distributions they take into account and the assumptions upon which they are based. After gathering data pertaining to his or her study, the researcher decides which hypothesis test he or she will use based on such criteria as the characteristics of the data, the population distributions, and normality. In circumstances where normality and variance homogeneity assumptions are infringed, it is wise to use non-parametric statistical tests. It is known that in the case of a non-normal distribution of data, non-parametric statistical procedures are more powerful than parametric ones. Among the non-parametric procedures used for testing data gathered from two independent samples, the most commonly used ones are the Kolmogorov-Smirnov two-sample test and the Mann-Whitney test.

When deciding on an analytical technique, researchers generally use one of these tests. Both the Kolmogorov-Smirnov two-sample test and the Mann-Whitney test are non-parametric procedures used for testing sequential data. Again, both these tests are applied in determining whether two independent samples are covered by the same population, and whether two populations are identical.

The aim of this study is to use Monte Carlo simulation to compare the statistical powers of the Kolmogorov-Smirnov two-sample test and the Mann-Whitney test for different skewness and kurtosis coefficients. We investigated the kurtosis coefficients when the skewness coefficients of both tests are equal, and the skewness coefficients when the kurtosis coefficients are equal. Our study is principally based on twelve populations with different skewness and kurtosis coefficients, derived from the normal distribution using Fleishman’s power function.

2 The Kolmogorov-Smirnov Two-sample Test

The Kolmogorov-Smirnov two-sample test is a general and comprehensive test for whether the populations of two independent samples are equivalent [1]. Conover (1999) [2] proposes that certain assumptions must be met for the Kolmogorov-Smirnov two-sample test to be applicable to a set of data. These assumptions are as follows:

- Every sample is chosen randomly from the population it represents;
- The measurement scale is sequential at least;
- The basic observed variable is a continuous variable;
- The two samples are independent from each other.

Conover (1999) [2] and Sheskin (2000) [3] define the data as follows:

Let $S_1(x)$ be the sample distribution function $x_1, x_2, \dots, x_{n1}$

Let $S_2(x)$ be the sample distribution function $y_1, y_2, \dots, y_{n2}$

Cumulative probabilities for $x_1, x_2, \dots, x_{n1}$ and $y_1, y_2, \dots, y_{n2}$ value are then identified.

Marascuilo and McSweeney (1977) [4] introduce a non-directional hypothesis for defining general disparities between the two populations. According to them, the zero and alternative hypotheses are as follows:

H_0: There is no difference between the two populations; or, for all values of x, H_0: $F(x) = G(x)$; from $-\infty$ to $+\infty$.

H_a: There are some differences between two populations; or, for at least one x, $F(x) \neq G(x)$.

According to Daniel (1990) [5], the test statistic for small and large samples in Kolmogorov-Smirnov two-sample test is as follows:

$$D = \max\left|S_1(x) - S_2(x)\right| \tag{1}$$

and the decision rule of the hypothesis is:

The zero hypothesis is rejected if the observed D value, for a specific significance level (α), is equal or greater than the D critical value ($D \geq D$ critical). In this case, one may say that there is a significant difference between the two populations.

3 The Mann-Whitney test

The Mann-Whitney test is one of the most powerful non-parametric tests and is used as an alternative to the parametric t-test [6]. Various assumptions were made in applying the Mann-Whitney test. For much of this study Sheskin's (2000) [3] data-regulating method, which computes rank sum, was used. Here are the assumptions that are in effect:

- Every single sample is chosen randomly from population it represents;
- The basic observed sample point is a continuous variable;
- Two sample points are randomly chosen and the point sequences are

independent from each other;
- The measure scale is sequential at least.

Let $x_1, x_2, \ldots, x_{n1}$ and $y_1, y_2, \ldots, y_{n2}$ be observation values chosen randomly from the first and second populations in turn; here the number of the xs is greater than the number of ys. Rank 1 is allocated according to $n_1 + n_2$, from minimum to maximum; let $N = n_1 + n_2$ be as follows:

$\sum R_1$ is the sum of the ranks of first sample group

$\sum R_2$ is the sum of the ranks of second sample group

Since our study is designed to test the alternative hypothesis (H_a), which suggests that there are differences between the two sample population distributions, we use the non-directional (double-ended test) hypothesis test developed by Conover (1999) [2] and Gibbons and Chakraborti (2003) [7].
For all values of x, the non-directional hypotheses are:

> Zero hypothesis H_0: $F(x) = G(x)$; or, there is no difference between the two populations.
>
> Alternative hypothesis H_a: For some values of x, $F(x) \neq G(x)$; or, there are some differences between the two populations.

For a sample size which is equal to or smaller than 10 ($m \leq 10$ and $n \leq 10$), we use the W test statistics proposed for smaller samples. According to the W test statistics:

> $W_x = \sum R_1$ (i.e. the sum of ranks of multi-variables of x chosen from the 1st population)
>
> $W_y = \sum R_2$ (i.e. the sum of ranks of multi-variables of y chosen from the 2nd population)

$$W_x + W_y = \frac{N(N+1)}{2} \tag{2}$$

In this formula, $N = m + n$ may be used in lieu of N.
The smaller value of W_x and W_y is used as W test statistics. The decision rule for this test statistic is as follows:

> If the probability of the observed W value in the table is smaller than the specific significance level (α), the zero hypothesis (H_0: Mx = My) is rejected and one may conclude that there is a significant difference between the two populations.

If the sample size is larger than 10 ($m > 10$ or $n > 10$) the normal approach formula is used. This formula which is proposed for use with larger samples sizes can also be used in the case that one of the sample sizes is larger than 3 or 4 and another is larger than 12. The formula is as follows:

$$z = \frac{W_x \pm 0,5 - \frac{m(N+1)}{2}}{\sqrt{\frac{mn(N+1)}{12}}} \quad (3)$$

In this formula, $W_x = \sum R_1$ may be used in lieu of the W_x value.

The decision rule for the W test statistics which is used for larger sample sizes is as follows:

> If the calculated absolute value of z is larger than the z value of the $\alpha/2^{th}$ level, the zero hypothesis (H_0: Mx = My) is rejected and one may conclude that there is a significant difference between the two population medians.

In this study, the linked values are ignored for both the Mann-Whitney and the Kolmogorov-Smirnov two-sample tests.

4 Statistical Power, Variance Heterogenity, Skewness, and Kurtosis

The power of a hypothesis test relates to the rejection of a null hypothesis H_0. The power of the test is denoted by $1 - \beta$. In a hypothesis test, the lower the probability of a Type II error (β) the greater the power of the test. Another factor increasing the power of a test is growth in the sample size (n). As the sample size (n) increases, the probability of a Type II error (β) decreases, and the power of the test advances accordingly [8].

According to Vogt, variance homogeneity refers to the case in which the populations from which the samples are chosen have identical or equivalent variances [9]. Where the Monte Carlo simulation is applied, one could denote variance homogeneity in a number of ways. For Penfield, the ratio of variances of two populations, in other words the value $\frac{\sigma_1^2}{\sigma_2^2}$, can be used in lieu of the variance homogeneity index [10]. The symbol σ_1^2 refers to the population variance of the first sample group, while the symbol σ_2^2 refers to the second group's population variance. In addition, Zimmerman (2004) [11] uses the ratio of the standard deviations of the two populations as an indicator of variance homogeneity. When applied to the Monte Carlo simulation, both the ratio of variances and the ratio of standard deviations give same results. For this reason, in this study the ratio of standard deviations ($\frac{\sigma_1}{\sigma_2}$) was used as an index to determine infringements of the variance homogeneity assumption. Furthermore, by using the standard deviation rates of Zimmerman, i.e. 2, 3, 4 [12], and Gibbons and Chakraborti, i.e. 1/2, 1/3, and 1/4 [13], we make a comparison of test powers.

Skewness and kurtosis are used as distribution patterns by Balakrishnan and Nevzorov (2003) [14] and Joanest and Gill (1998) [15]. Skewness and kurtosis measures were first developed by Pearson in 1895 (quoted in [14]). Pearson formulates skewness as

$$\gamma_1 = \frac{\beta_3}{\sqrt{\beta_2^3}} \tag{4}$$

and kurtosis as

$$\gamma_2 = \frac{\beta_4}{\sqrt{\beta_2^2}} \tag{5}$$

In these formulas,

β_2 refers to 2nd central moment of the distribution function of the population,

β_3 refers to 3rd central moment of the distribution function of the population, and

β_4 refers to 4th central moment of the distribution function of the population.

In addition, Algina, Olejnik, and Ocanto (1989) [16] suggested the following formulas for computing skewness and kurtosis:

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \tag{6}$$

$$\gamma_2 = \frac{\mu_4}{\sigma^4} \tag{7}$$

According to Algina, Olejnik, and Ocanto (1989) [16], if $\gamma_1 = 0$ and $\gamma_2 = 3$, the distribution is normal. Furthermore, Balakrishnan and Nevzorov (2003) [14] identify $\gamma_1 > 0$ as a positive skewness distribution, $\gamma_1 < 0$ as a negative skewness distribution, $\gamma_2 = 3$ as a normal distribution, $\gamma_2 < 3$ as a platykurtosis distribution, and $\gamma_2 > 3$ as a leptokurtosis distribution.

5 Monte Carlo Simulation

Monte Carlo simulation is an alternative to analytical mathematics and investigates the distribution of a population via random samples. By using random samples gathered from simulated data regarding a known population, Monte Carlo simulation computes statistics experimentally [17], and thus obtains approximate solutions to mathematical or physical problems, providing a range of values, each of which has a calculated probability of being the solution.
In this study, we used the SAS 9.00 program to apply the Monte Carlo Simulation. To help researchers produce comprehensive and different distributions and simulate experimental distributions, Fleishman developed a power function which produces distributions. The RANNOR procedure in SAS is used for producing random numbers from the normal distribution, which has a mean of zero and a standard deviation of 1. In addition, this is a requirement for the power transformation that Fleishman used for producing normal distributions [18]. Fleishman's formula for the power function is as follows:

$$y = a + [(d \times x + c) \times x + b] \times x \tag{8}$$

In this formula, y is a distribution that depends on constants, and x is a random variable which has a mean of zero and standard deviation of 1. Coefficients a, b, c, and d are defined using standard deviations, skewness, and kurtosis. The values were first proposed by Fleishman, where a is constant and c takes the reverse sign to the constant a (i.e. c = –a). After establishing sample mechanisms, the PROC NPAR1WAY procedure is used for showing power simulations.

For the twelve population distributions we referred to the study by Algina, Olejnik, and Ocanto (1989) [16]. We investigated different kurtosis values with equal skewness, and different skewness values with equal kurtosis. There are eight distributions which have the same skewness and different kurtosis values. Six of these eight distributions have a skewness value of (0.00) and the remainder have a skewness value of (0.75). The distributions having a skewness value of (0.00) are: Normal, Platykurtic, Normal Platykurtic, Leptokurtic[1], Leptokurtic[2], and Leptokurtic[3]. The kurtosis values of these distributions are (0.00), (-0.50), (-1.00), (1.00), (2.00), and (3.75) respectively. Distributions having a skewness value of (0.75) are the Skewed, Skewed, and Leptokurtic[1]. The kurtosis values of these distributions are (0.00) and (3.75). For a double comparison we combined these distributions into two. Finally, we arrived at sixteen distributions having equal skewness and different kurtosis values.

There are ten distributions which have equal kurtosis values and different skewness values. Four of these distributions have a kurtosis of (3.75), two of them (-1.00), two of them (-0.50), and two of them (0.00). The distributions having kurtosis value of (3.75) are the Leptokurtic[3], Skewed and Leptokurtic[1], Skewed and Leptokurtic[2], and Skewed-Leptokurtic distributions. The skewness values of these distributions are (0.00), (0.75), (1.25), and (1.75) respectively. Distributions having a kurtosis of (-1.00) are the Normal Platykurtic and the Skewed and Platykurtic[2]. Skewness values of these distributions are (0.00) and (0.25). Distributions having a kurtosis value of (-0.50) are the Platykurtic and Skewed and Platykurtic[1] distributions. The skewness values of these are (0.00) and (0.50) respectively. The distributions having a kurtosis value of (0.00) are also the Normal and Skewed distributions, which have a skewness value of (0.00) and (0.75) respectively. Finally, we arrived at nine distributions that have equal kurtosis and different skewness values.

Table 1: Fleishman's Power function for μ=0 and σ=1

Distribution	**Skewness** (γ_1)	**Kurtosis** (γ_2)	**a**	**B**	**C**	**d**
Normal	0,00	0,00	0,00	1,0000000	0,00	0,00
Platykurtic	0,00	-0,50	0,00	1,0767327	0,00	-0,0262683
Normal Platykurtic	0,00	-1,00	0,00	1,2210010	0,00	-0,0801584
Leptokurtic[1]	0,00	1,00	0,00	0,9029766	0,00	0,0313565
Leptokurtic[2]	0,00	2,00	0,00	0,8356646	0,00	0,0520574
Leptokurtic[3]	0,00	3,75	0,00	0,7480208	0,00	0,0778727
Skewed	0,75	0,00	-0,1736300	1,1125146	0,1736300	-0,0503344
Skewed and Platykurtic[1]	0,50	-0,50	-0,1201561	1,1478491	0,1201561	-0,0575035
Skewed and Platykurtic[2]	0,25	-1,00	-0,0774624	1,2634128	0,0774624	-0,1000360
Skewed and Leptokurtic[1]	0,75	3,75	-0,0856306	0,7699520	0,0856306	0,0693486
Skewed and Leptokurtic[2]	1,25	3,75	-0,1606426	0,8188816	0,1606426	0,0491652
Skewed-Leptokurtic	1,75	3,75	-0,3994967	0,9296605	0,3994967	-0,0364670

Source: Lee (2007) [19], page 88.

In this study, we found a total of twenty-five distributions, sixteen of which have equal skewness and different kurtosis values and nine of which have equal kurtosis and different skewness values. Again, there are six distributions for standard deviation values of (2, 3, 4, 1/2, 1/3, and 1/4) and six distributions for sample sizes of (5, 10), (5, 20), (10, 5), (10, 20), (20, 5), and (20, 10). Therefore, in our study 25 × 6 × 6 = 900 syntaxes are drawn and investigated.

6 Simulation Results

6. 1 Results for Equal Skewness and Different Kurtosis Values

We undertook power comparisons for sixteen different populations. Very similar power characteristics are seen for the Normal and Platykurtic sample pairs (in which the first sample is taken from the Normal distribution while the other is taken from the Platykurtic distribution), the Normal and Normal Platykurtic sample pairs (first sample taken from the Normal distribution, the other from the Normal Platykurtic distribution), and the Platykurtic and Normal Platykurtic sample pairs (first sample taken from the Platykurtic distribution, the other from the Normal Platykurtic distribution). According to the sample pairs gathered from these distributions, the Mann-Whitney test is more powerful in the (5, 10) and (10, 5) sample sizes with standard deviations of 2 and 1/2; in the (5, 20) sample size with standard deviations of 2, 3, and 4; in the (10, 20) sample size with standard deviation of 2; and in the (20, 5) sample size with standard deviations of 2, 1/2, 1/3, and 1/4. For other sample sizes and standard deviations the Kolmogorov-Smirnov two-sample test is more powerful.

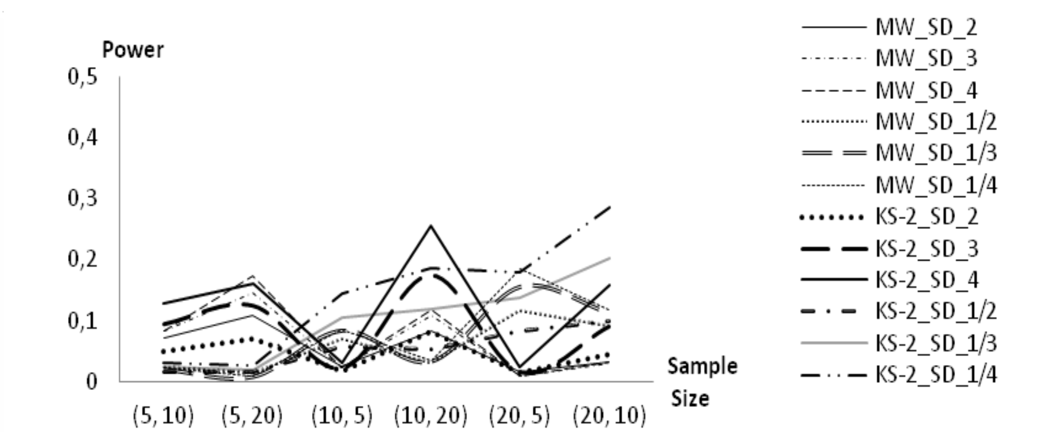

Figure 1: Powers of Mann-Whitney and Kolmogorov-Smirnov two-sample tests for Normal and Platykurtic, Normal and Normal Platykurtic, and Platykurtic and Normal Platykurtic distributions when standard deviation ratios are 2, 3, 4, 1/2, 1/3, and 1/4.

Very similar characteristics are seen for the Normal and Leptokurtic[1] sample pairs, the Normal and Leptokurtic[2] sample pairs, the Normal and Leptokurtic[3] sample pairs, the Normal Platykurtic and Leptokurtic[1] sample pairs, and the Leptokurtic[1] and Leptokurtic[2] sample pairs. According to the sample pairs gathered from these distributions, the Mann-Whitney test is more powerful in the (5, 10) sample size with standard deviations of 2, 1/2, and 1/3; in the (5, 20) sample size with standard deviations of 2, 3, and 1/2; in the (10, 5) sample size with standard deviations of 2 and 1/2; in the (20, 5) sample size with standard deviations of 1/2, 1/3, and 1/4; and in the (20,10) sample size with standard deviation of 1/2. For other sample sizes and standard deviations the Kolmogorov-Smirnov two-sample test is more powerful.

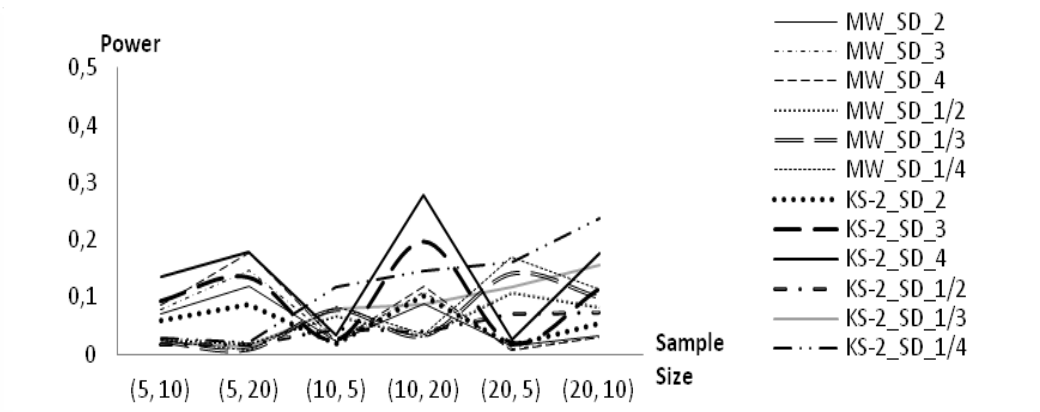

Figure 2: Powers of Mann-Whitney and Kolmogorov-Smirnov two-sample tests for Normal and Normal Leptokurtic[2], Normal and Leptokurtic[3], Normal Platykurtic and Leptkurtic[1] and Leptkurtic[1] and Leptkurtic[2] distributions when standard deviation ratios are 2, 3, 4, 1/2, 1/3, and 1/4.

Very similar power characteristics are shown by the Platykurtic and Leptokurtic[1] sample pairs, Platykurtic and Leptokurtic[2] sample pairs, and Leptokurtic[1] and Leptokurtic[3] sample pairs. According to the sample pairs gathered from these distributions, the Mann-Whitney test is more powerful in the (5, 10) and (10,5) sample sizes with standard deviations of 2, 1/2, and 1/3; in the (5, 20) sample size with standard deviations of 2, 3, 4, and 1/2; in the (20, 5) sample size with standard deviations of 1/2, 1/3, and 1/4; and in the (20,10) sample size with standard deviation of 1/2. For the other sample sizes and standard deviations the Kolmogorov-Smirnov two-sample test is more powerful.

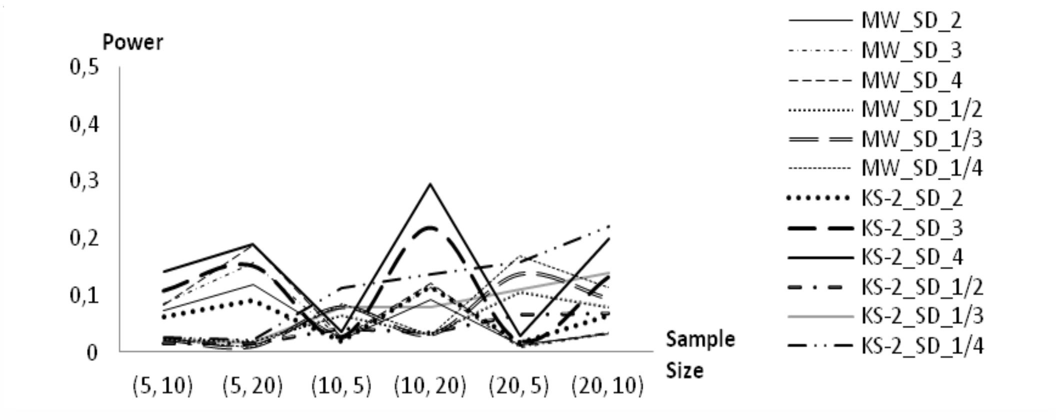

Figure 3: Powers of Mann-Whitney and Kolmogorov-Smirnov tests for Platykurtic and Leptkurtic[1], Platykurtic and Leptkurtic[2] and Leptkurtic[1] and Leptkurtic[3] distributions when standard deviations are 2, 3, 4, 1/2, 1/3, and 1/4.

Very similar power characteristics are shown by the Platykurtic and Leptokurtic[3] sample pairs, and the Normal Platykurtic and Leptokurtic[2] sample pairs. According to the sample pairs gathered from these distributions, the Mann-Whitney test is more powerful in the (5, 10) sample size with standard deviations of 2, 1/2, 1/3, and 1/4; in the (5, 20) sample size with standard deviations of 2 and 1/2; in the (10, 5) sample size with standard deviations of 2, 1/2, and 1/3; in the (10,20) sample size with standard deviation of 1/2; in the (20,5) sample size with standard deviations of 1/2, 1/3, and 1/4; and in the (20,10) sample size with standard deviation of 1/2. For other sample sizes and standard deviations the Kolmogorov-Smirnov two sample test is more powerful.

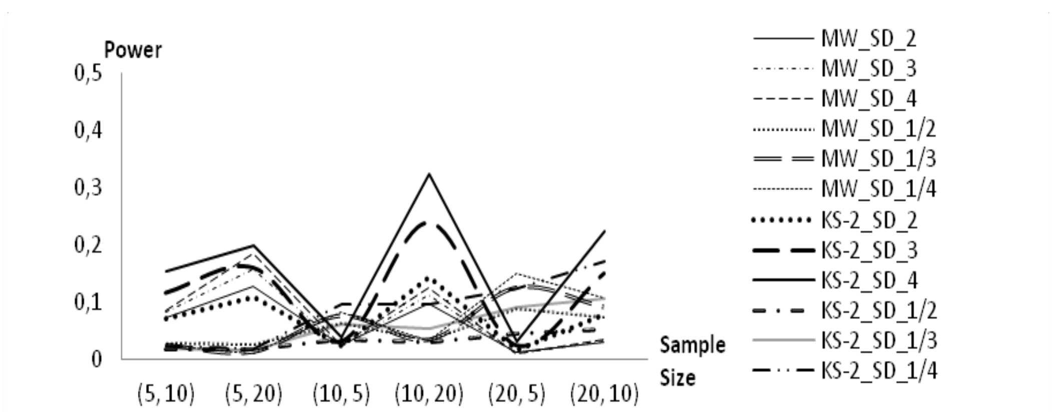

Figure 4: Powers of Mann-Whitney and Kolmogorov-Smirnov two-sample tests for Platykurtic and Leptkurtic[3] and Normal Platykurtic and Leptkurtic[2] distributions when standard ratios are 2, 3, 4, 1/2, 1/3, and 1/4.

Very similar power characteristics are shown by the Normal Platykurtic and Leptokurtic[3] sample pairs, the Skewed and Skewed and Leptokurtic[1] sample pairs. According to the sample pairs gathered from these distributions, the Mann-Whitney test is more powerful in the (5, 10) sample size with standard deviations of 1/2, 1/3, and 1/4; in the (5, 20) sample size with standard deviations of 2 and 1/2; in the (10, 5) sample size with standard deviations of 2, 1/2, and 1/3; in the (10,20) sample size with standard deviation of 1/2; in the (20,5) sample size with standard deviations of 1/2, 1/3, and 1/4; and in the (20,10) sample size with standard deviation of 1/2. For other sample sizes and standard deviations the Kolmogorov-Smirnov two-sample test is more powerful.

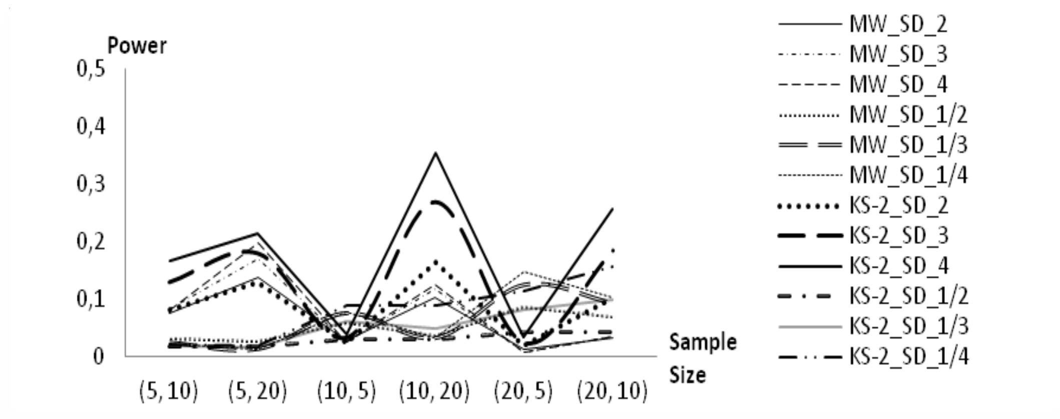

Figure 5: Powers of Mann-Whitney and Kolmogorov-Smirnov two-sample tests for Normal Platykurtic and Leptkurtic[3] and Skewed and Skewed and Leptokurtic[1] Distributions when standard deviation ratios are 2, 3, 4, 1/2, 1/3, and 1/4.

According to sample pairs in which first sample is taken from the Leptokurtic[2] and the other from the Leptokurtic[3] distribution, the Mann-Whitney test is more powerful in the (5, 10) sample size with standard deviations of 2, 1/2, and 1/3; in the (5, 20) sample size with standard deviations of 2, 3, and 1/2; in the (10, 5) sample size with standard deviations of 2, 3, 1/2, and 1/3; in the (10,20) sample size with standard deviation of 1/2; in the (20,5) sample size with standard deviations of 2, 1/2, 1/3, and 1/4; and in the (20,10) sample size with standard deviation of 1/2; for other sample sizes and standard deviations the Kolmogorov-Smirnov two-sample test is more powerful.

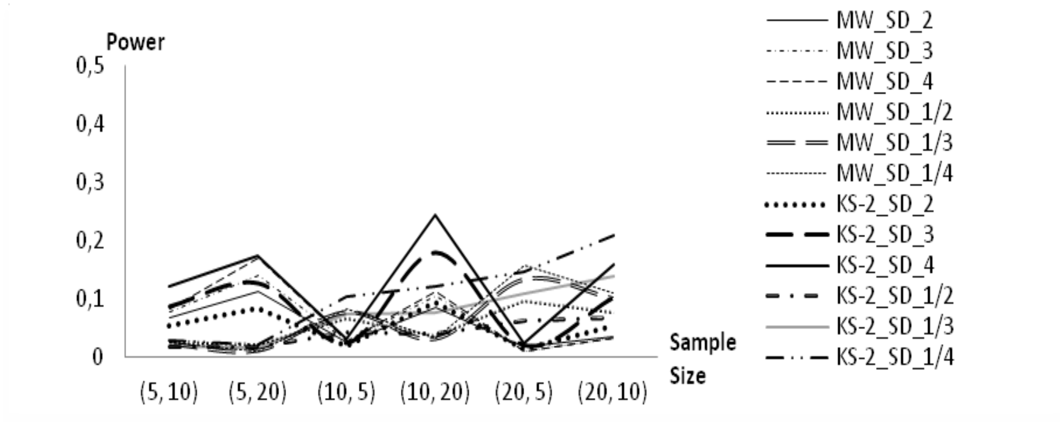

Figure 6: Powers of Mann-Whitney and Kolmogorov-Smirnov two-sample tests for Leptokurtic[2] and Leptokurtic[3] distributions when standard deviation ratios are 2, 3, 4, 1/2, 1/3, and 1/4.

6. 2 Results for Different Skewness and Equal Kurtosis Values

We made power comparisons for nine different populations. According to sample pairs in which first sample is taken from the Normal distribution and other from the Skewed distribution, the Mann-Whitney test is more powerful in the (5, 10) sample size with standard deviations of 2 and 1/2; in the (5, 20) sample size with standard deviations of 2, 3, and 4; in the (10, 5) sample size with standard deviations of 2, 3, and 1/2; and in the (20,5) sample size with standard deviations of 1/2 and 1/3; for other sample sizes and standard deviations the Kolmogorov-Smirnov two-sample test is more powerful.

Very similar power characteristics are shown by the Platykurtic and Skewed and Platykurtic[1] sample pairs, and the Normal Platykurtic and Skewed and Platykurtic[2] sample pairs. According to the sample pairs, the Mann-Whitney test is more powerful in the (5, 10) sample size with standard deviations of 1/2, 1/3, and 1/4; in the (5, 20) sample size with standard deviations of 2 and 1/2; in the (10, 5) sample size with standard deviations of 2, 1/3, and 1/2; in the (10, 20) sample size with standard deviation of 1/2; in the (20, 5) sample size with standard deviations of 1/2, 1/3, and 1/4; and in the (20, 10) sample size with standard deviation of 1/2; for other sample sizes and standard deviations the Kolmogorov-Smirnov two-sample test is more powerful.

Very similar power characteristics are shown by the Leptokurtic[3] and Skewed and Leptokurtic[1] sample pairs, the Leptokurtic[3] and Skewed and Leptokurtic[2] sample pairs, the Skewed and Leptokurtic[1] and Skewed and Leptokurtic[2] sample pairs. According to the sample pairs, the Mann-Whitney test is more powerful in the (5, 10) sample size with standard deviations of 2, 3, and 1/2; in the (5, 20) sample size with standard deviations of 2, 3, 4, and 1/2; in the (10, 5) sample size with standard deviations of 2, 3, and 1/2; in the (10, 20) sample size with standard deviation of 2; and in the (20, 5) sample size with standard deviations of 2, 1/2, and 1/3; for other sample sizes and standard deviations the Kolmogorov-Smirnov two-sample test is more powerful.

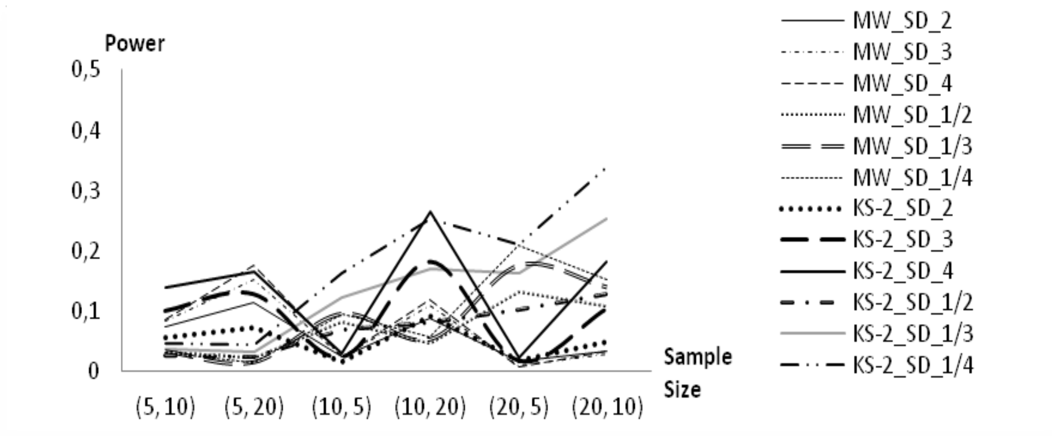

Figure 7: Powers of Mann-Whitney and Kolmogorov-Smirnov two-sample tests for Normal and Skewed distributions when standard deviation ratios are 2, 3, 4, 1/2, 1/3, and 1/4.

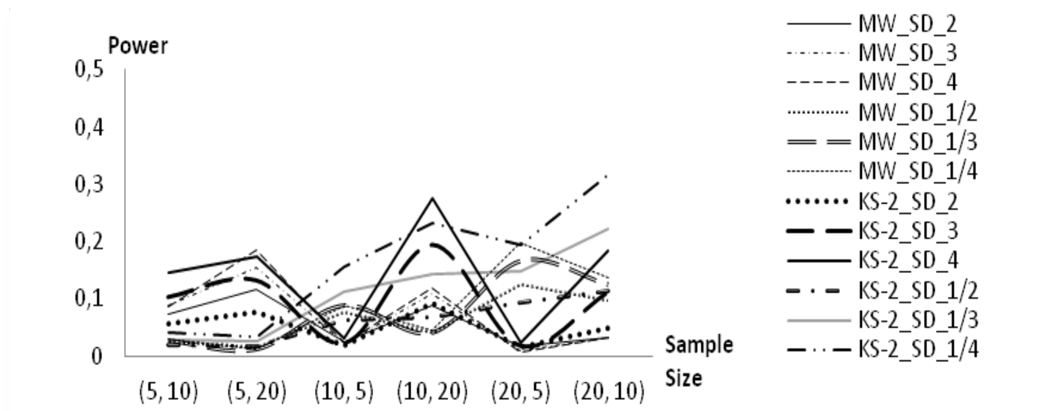

Figure 8: Powers of Mann-Whitney and Kolmogorov-Smirnov two-sample tests for Platykurtic and Skewed and Platykurtic[1] and Normal Platykurtic and Skewed and Platykurtic[2] distributions when standard deviation ratios are 2, 3, 4, 1/2, 1/3, and 1/4.

Very similar power characteristics are shown by the Leptokurtic[3] and Skewed-Leptokurtic sample pairs, and the Skewed and Leptokurtic[1] and Skewed-Leptokurtic sample pairs. According to sample pairs, the Mann-Whitney test is more powerful in the (5, 10) sample size with standard deviations of 2 and 1/2; in the (5, 20) sample size with standard deviations of 2, 3, and 4; in the (10, 5) sample size with standard deviations of 2, 3, and 1/2; and in the (20,5) sample size with standard deviation of 1/2; for other sample sizes and standard deviations Kolmogorov-Smirnov two-sample test is more powerful.

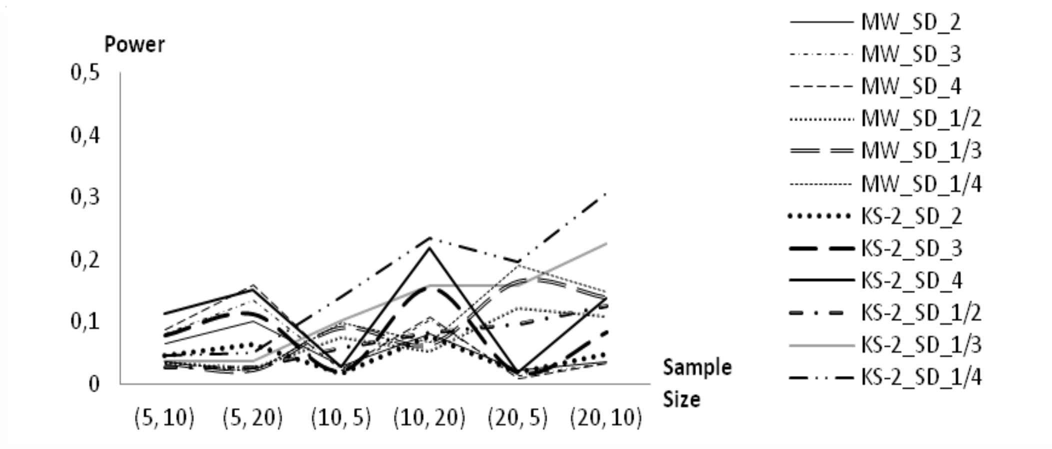

Figure 9: Powers of Mann-Whitney and Kolmogorov-Smirnov two-sample tests for Leptokurtic[3] and Skewed and Leptokurtic[1], Leptokurtic[3] and Skewed and Leptokurtic[2] and Skewed and Leptokurtic[1] and Skewed and Leptokurtic[2] distributions when standard deviation ratios are 2, 3, 4, 1/2, 1/3, and 1/4.

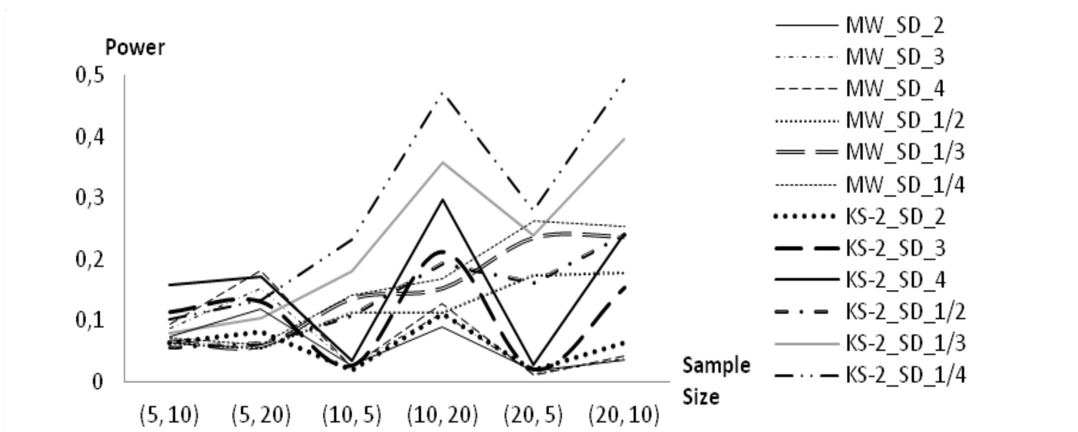

Figure 10: Powers of Mann-Whitney and Kolmogorov-Smirnov two-sample tests for Leptokurtic[3] and Skewed-Leptokurtic and Skewed and Leptokurtic[1] and Skewed-Leptokurtic distributions when standard deviation ratios are 2, 3, 4, 1/2, 1/3, and 1/4.

According to the Skewed and Leptokurtic[2] and Skewed-Leptokurtic sample pairs, the Mann-Whitney test is more powerful in the (5, 10) sample size with standard deviations of 2 and 1/2; in the (5, 20) sample size with standard deviations of 2, 3, and 4; in the (10, 5) sample size with standard deviations of 2 and 1/2; and in the (20,5) sample size with standard deviations of 2, 1/2, and 1/3. For other sample sizes and standard deviations the Kolmogorov-Smirnov two-sample test is more powerful.

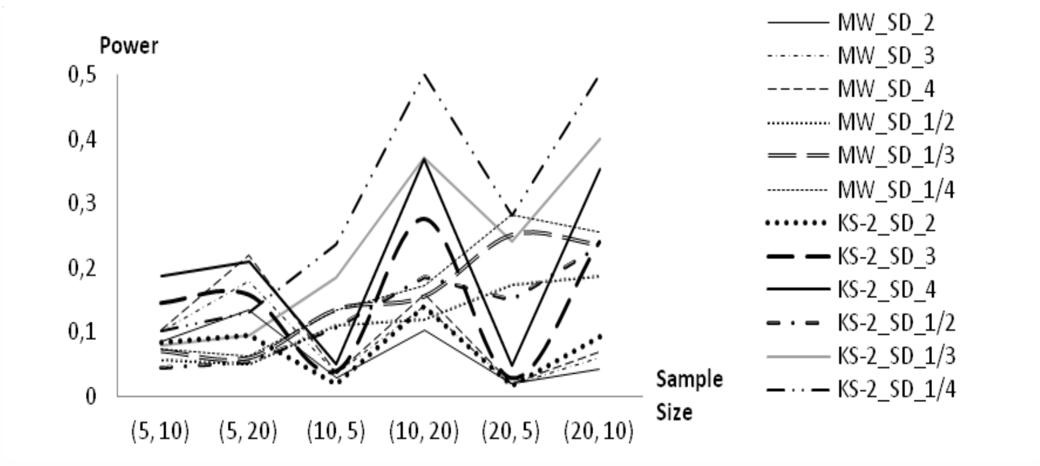

Figure 11:. Powers of Mann-Whitney and Kolmogorov-Smirnov two-sample tests for Skewed and Leptokurtic[2] and Skewed-Leptokurtic Distributions when standard deviation ratios are 2, 3, 4, 1/2, 1/3, and 1/4.

7 Results

According to sample pairs gathered from twenty-five different population distributions, the Mann-Whitney test is observed to be more powerful in the (5, 10) and (10, 5) sample sizes with standard deviations of 2 and 1/2; in the (5, 20) sample size with standard deviations of 2, 3, and 4; and in the (20, 5) sample size with standard deviations of 1/2, 1/3, and 1/4. However, the Kolmogorov-Smirnov two-sample test is more powerful in the (5, 10) and (10, 5) sample sizes with standard deviations of 3, 4, 1/3, and 1/4; in the (5, 20) sample size with standard deviations of 1/2, 1/3, and 1/4; in all (10, 20) and (20, 10) sizes; and in the (20, 5) sample sizes with standard deviations of 2, 3, and 4.

In all distributions, the largest power values are observed in standard deviation values of 4 or 1/4. For most distributions, when the standard deviations increase from 2 to 4 or decrease from 1/2 to 1/4 it is observed that both tests increased their powers. The Mann-Whitney test is more powerful in the case where the first sample's size is larger than the second's but the larger sample's standard deviation is smaller than the other's (e.g. in the (20, 5) sample size with standard deviation rates of 1/2, 1/3, and 1/4). Where the standard deviation of the larger sample is bigger than the others, the Kolmogorov-Smirnov two-sample test is more powerful. Similarly, the Mann-Whitney test is more powerful if the first sample's size is smaller but the standard deviation is larger (e.g. in the (5, 20) sample size with standard deviation rates of 2, 3, and 4); whereas if the standard deviation of the smaller sample is smaller than the others, the Kolmogorov-Smirnov two-sample test is more powerful.

In terms of power comparisons of equal kurtosis and different skewness and different skewness and equal kurtosis, the largest power is observed for the (20,10) sample size with standard deviation of 1/4 in the Kolmogorov-Smirnov two-sample test (0.502). The smallest power value in this study is observed for same pairs gathered from the Normal Platykurtic and Leptokurtic[1] distributions where the standard deviation of the Mann-Whitney test is 4 (0.007).

References

[1] J.J. Higgins, *Introduction to Modern Statistics*, Pacific Grove, Thomson Learning, Inc., 2004.

[2] W.J. Conover, *Parametric Nonparametric Statistics*, Pacific Grove, John Wiley & Sons, Inc., New York, 1999.

[3] D.J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC, Baca Raton, 2000.

[4] L. A. Marascuilo and M. McSweeney, *Nonparametric and Distribution-free Methods for Social Science*, Brooks/Cole Publishing Company, Monterey, 1977.

[5] W.W. Daniel, *Applied Nonparametric Statistics*, PWS-Kent Publishing Company, Boston, 1990.

[6] S. Siegel and J. N. J. Castellan, *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill, Massachusetts, 1988.

[7] J.D. Gibbons and S. Chakraborti, *Nonparametric Statistical Inference*, NY: Marcel Dekker, Inc., New York, 2003.

[8] M. Kartal, *Hypothesis Tests for Scientific Researches*, Nobel Press, Ankara, 2006.

[9] W.P. Vogt, *Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Sciences*, SAGE Publications, Thousand Oaks, 2005.

[10] D.A. Penfield, Choosing a Two-sample Location Test, *Journal of Experimental Education*, **62**(4), (1994), 343-350.

[11] D.W. Zimmerman, Inflation of Type I Error Rates by Unequal Variance Associated with Parametric, Nonparametric, and Rank-transformation Tests, *Psicologica*, **25**, (2004), 103-133.

[12] D.W. Zimmerman, Invalidation of Parametric and Nonparametric Statistical Tests by Concurrent Violation of Two Assumptions, *Journal of Experimental Education*, **67**(1), (1998), 55-68.

[13] J.D. Gibbons and S. Chakraborti, Comparisons of Mann-Whitney, Student's t, and Alternative t Test for Means of Normal Distributions, *Journal of Experimental Education*, **59**(3), (1991), 258-267.
[14] N. Balakrishnan and V.B. Nevzerov, *A Primer on Statistical Distributions*, John & Wiley Sons Inc., New York, 2003.
[15] D.N. Joanest and G. A. Gill, Comparing Measures of Sample skewness and Kurtosis, *The Statistician*, **47**(1), (1998), 183-189.
[16] J. Algina, S. Olejnik and R. Ocanto, Type I Error Rates and Power Estimates for Selected Two-sample Tests of Scale, *Journal of Educational Statistics*, **14**(4), (1989), 373-384.
[17] C.Z. Mooney, *Monte Carlo Simulation*, SAGE Publication Inc., Thousand Oaks, 1997.
[18] X. Fan, A. Felsovalyi, S.A. Sivo and S.C. Keenan, *SAS for Monte Carlo Studies: A Guide for Quantitative Researchers*, SAS Publishing, Ary, NC, 2003.
[19] C.H. Lee, *A Monte Carlo Study of Two Nonparametric Statistics with Comparisons of Type I Error Rates and Power*, Unpublished PhD Dissertation, Graduate College of the Oklahoma State University, Stillwater.