Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 50 (3) (2021), 872 – 894
DOI : 10.15672/hujms.644516

Research Article

A new goodness of fit test for multivariate normality

Orhan Kesemen, Buğra Kaan Tiryaki*, Özge Tezel, Eda Özkul

Department of Statistics and Computer Sciences, Karadeniz Technical University, Trabzon, Turkey

Abstract

This paper presents a multivariate Kolmogorov-Smirnov (MVKS) goodness of fit test for multivariate normality. The proposed test is based on the difference between the empirical distribution function and the theoretical distribution function. While calculating them in multivariate case, the problem is that the variables cannot be distribution-free as in the univariate case. Firstly, the variables are made independent to solve this problem and the Rosenblatt transform is applied for independence of variates. Then the theoretical and empirical distribution values are calculated and the MVKS test statistic is computed. It provides an easy calculation for d-dimensional data by using the same algorithm and critical table values. This paper demonstrates the effectiveness of the MVKS for different dimensions with a simulation study which also includes the comparison of the MVKS critical tables with univariate Kolmogorov-Smirnov (KS) critical table and the power comparisons of the MVKS (bivariate case) against with the existing bivariate normality tests. Lastly, the MVKS is applied to two different multivariate data sets to confirm that it achieves consistent, accurate and correct results.

Mathematics Subject Classification (2020). 62H15, 62G10, 62Q05

Keywords. multivariate empirical distribution function, goodness of fit test, elliptical distributions

1. Introduction

The goodness of fit tests are used to test whether the data come from a given distribution. The most common goodness of fit test is normality test. The assumption of normality is based on many statistical procedures. The t-test, the linear regression analysis, the discriminant analysis and the variance analysis are the well-known statistical procedures. When the assumption of normality is violated, the interpretations and the inferences may not be reliable or valid [37].

The goodness of fit tests have been developed for the univariate distributions [23, 30]. However, there are several studies for the multivariate distributions in the literature. The adaptation of the chi-square test for multivariate data was performed by [9, 33]. A multivariate skewness-kurtosis measure was used by [5, 26, 27, 41, 49]. Furthermore, Székely

*Corresponding Author.
Email addresses: okesemen@gmail.com (O. Kesemen), bugrakaantiryaki@gmail.com (B.K. Tiryaki), ozge_tzl@hotmail.com (Ö. Tezel), eda.ozkul.gs@gmail.com (E. Özkul)
Received: 08.11.2019; Accepted: 08.02.2021

and Rizzo [46] and McAssey [31] used the Euclidean distance, Koziol [21] proposed the radius and angle test, Baringhaus and Henze [4] and Fan [13] presented the empirical characteristic function approaches for the multivariate goodness of fit test. The adaptation of the univariate Kolmogorov-Smirnov goodness of fit test to the multi-dimension was proposed by [6,14,18,25,36]. The other different studies were carried out by [1,2,8,11,15, 24,34,38,45,48].

The univariate empirical goodness of fit tests are based on the difference between the empirical distribution function and the theoretical distribution function. The distribution values must be independent from the distribution for the univariate empirical goodness of fit tests [7]. However, the distribution values are uniformly distributed in the interval of $[0,1]$ to satisfy the independence. Although these required conditions are sufficient for the univariate goodness of fit tests in the previous studies, they cannot be precisely applied to the multivariate tests. In addition, it is difficult to define the cumulative distribution function while adapting the univariate normality tests to the multivariate normality test in more than one dimension [25].

The empirical goodness of fit tests are generally based on the difference between the empirical distribution function and the theoretical distribution function. The well-known goodness of fit tests such as Kolmogorov-Smirnov [20,42], Lilliefors [23], Cramér-von Mises [10,47], Kuiper [22] methods are based on empirical distribution function. The proposed method can be applied to different empirical goodness of fit tests. Therefore, Kolmogorov-Smirnov test is choosen to show the applicability of it.

2. The Kolmogorov-Smirnov statistics

2.1. Univariate case

Let $x_1, x_2, \ldots, x_n$ is a random sample in $\mathbb{R}$ of independent and identically distributed (i.i.d.) random variables with the distribution function F.

The null hypothesis is $H_0 : F(x) = F_0(x)$ and against the alternative hypothesis is $H_1 : F(x) \neq F_0(x)$, where $F_0(x)$ is the desired theoretical distribution function. Also, F_0 is assumed to be a continuous distribution. In the univariate case, the Kolmogorov-Smirnov statistic is computed as follows

$$D_n = \sup_x |F_n(x) - F_0(x)|, \tag{2.1}$$

where $F_n(x)$ is the empirical distribution function. The basic approach used in the literature to calculate the empirical distribution function (F_n) is that the sample is divided by the sample size. Also, it is calculated as left points $(x \leq x_i)$ to the current point (x) by the total number of sample size.

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(x_i, x) \tag{2.2}$$

$I(.,.)$ is the indicator function and is calculated as in Equation (2.3).

$$I(a,b) = \begin{cases} 1, & a \leq b \\ 0, & otherwise \end{cases} \tag{2.3}$$

The empirical distribution values obtained from above approach will have an ideal uniform distribution in the range of $[0,1]$. Also, these values will be independent from the distribution. On the other hand, if random values x_i come from the distribution function F_0, the values u_i obtained from the theoretical distribution function are independent and uniformly distributed values in the range of $[0,1]$.

$$u_i = F_0(x_i), \quad (i = 1, 2, \ldots, n) \tag{2.4}$$

Regardless of the distribution, if obtained random values uniformly distributed in the interval $[0,1]$, it must satisfy that the theoretical distribution values are independent from the distribution.

2.2. Multivariate case

Let $\mathbf{x} = \{(x_{1i}, x_{2i}, \ldots, x_{pi}),\ i = 1, 2, \ldots, n\}$ a random sample in $\mathbb{R}^p$ of i.i.d random variables obtained from the p-variate joint distribution function F. The null hypothesis is $H_0 : F(\mathbf{x}) = F_0(\mathbf{x})$ and against the alternative hypothesis is $H_1 : F(\mathbf{x}) \neq F_0(\mathbf{x})$, where $F_0(\mathbf{x})$ is $N_p(\boldsymbol{\mu}, \Sigma)$, a multivariate normal distribution with a known mean $\boldsymbol{\mu}$ and a known covariance matrix Σ. The extension of the univariate Kolmogorov-Smirnov statistic to the multivariate case can be calculated in the following equation.

$$\tilde{D}_n^{(p)} = \sup_{\mathbf{x}} |F_n(\mathbf{x}) - F_0(\mathbf{x})| \tag{2.5}$$

Even though the empirical and the theoretical distribution function values in Equation (2.5) are in the range of [0,1], they cannot be independent of the distribution and do not also come from the uniform distribution. The methods [14, 18, 36] are improved to calculate the test statistic independently from the distribution either yield approximate results or have complex algorithms.

3. Computation of the test statistics

The problem in the multivariate case is that it cannot be independent from the distribution as in the univariate case. A simple transformation was proposed by [39] to solve this problem [18]. Accordingly, p-variate random values are transformed into random values which come from the univariate uniform distribution. This transformation is given in the Theorem 3.1.

Theorem 3.1 (Rosenblatt)**.** *If $X = (X_1, \ldots, X_p)$ are taken as a joint random density vector, the joint distribution function can be written as follows*

$$f_0(x_1, \ldots, x_p) = f_1(x_1)\, f_2(x_2|x_1) \ldots f_p(x_p|x_1, \ldots, x_{p-1}). \tag{3.1}$$

If $Y = T(X)$ transformation is applied, the following equation is obtained.

$$\begin{aligned} Y_1 &= F_1(X_1) \\ Y_k &= F_k(X_k|X_1, \ldots, X_{k-1}),\ k = 2, \ldots, p \end{aligned} \tag{3.2}$$

Y_k is calculated as Equation (3.2). Moreover, $Y_1, \ldots, Y_p$ have uniform distribution in the range of $[0,1]$ and they are independent variables by the help of conditional probability. However, if the conditional probability was not performed, the joint distribution function could be written as follows by using the marginal distribution instead of the conditional distribution function.

$$F_0(x_1, \ldots, x_p) = F_1(x_1)\, F_2(x_2) \cdots F_p(x_p) \tag{3.3}$$

A new random variable is obtained in Equation (3.4) by substituting a random variable from the marginal distribution function.

$$U_k = F_k(X_k), \quad (k = 1, \ldots, p) \tag{3.4}$$

3.1. Independence of variables for multivariate normal distributions

Covariance is a measure of the linear relationship between random variables [17]. It can be given by the following equation [19].

$$\begin{bmatrix} X_1 \\ X_2 \\ \dots \\ X_p \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \dots \\ \mu_p \end{bmatrix} + \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \dots & \dots & \dots & \dots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \\ \dots \\ Z_p \end{bmatrix} \tag{3.5}$$

Here, the variables Z_k come from a distribution with the expected value 0 and the variance 1. The variables X_k are linearly related with the variables Z_k. The covariance matrix of X_k is calculated as follows

$$\Sigma_X = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1p} \\ v_{21} & v_{22} & \dots & v_{2p} \\ \dots & \dots & \dots & \dots \\ v_{p1} & v_{p2} & \dots & v_{pp} \end{bmatrix}. \tag{3.6}$$

The calculation of the covariance matrix is performed by Equation (3.7).

$$v_{ij} = \sum_{k=1}^{p} \sigma_{ik}\sigma_{jk}, \quad (i, j = 1, 2, \dots, p) \tag{3.7}$$

The following equation can be written by using Equations (3.5)-(3.7) [3, 16].

$$\begin{bmatrix} Z_1 \\ Z_2 \\ \dots \\ Z_p \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1p} \\ v_{21} & v_{22} & \dots & v_{2p} \\ \dots & \dots & \dots & \dots \\ v_{p1} & v_{p2} & \dots & v_{pp} \end{bmatrix}^{-\frac{1}{2}} \begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \dots \\ X_p - \mu_p \end{bmatrix} \tag{3.8}$$

If $(X_1, X_2, \dots, X_p)$ has distribution F_0, then $Z_1, Z_2, \dots, Z_p$ in Equation (3.8) are independent $N(0, 1)$ random variables whose joint distribution will be denoted by F_0^*. Let $(z_{1i}, z_{2i}, \dots, z_{pi})$, $i = 1, \dots, n$ denote the result of transforming the observations $(x_{1i}, x_{2i}, \dots, x_{pi})$, $i = 1, \dots, n$ by Equation (3.8).

3.2. Computation of the empirical distribution values

Equation (3.9) is obtained when the approach used to calculate the univariate empirical distribution function in the Section 2.1 is extended for the empirical distribution function of the multivariate cases.

$$\tilde{F}_n(z_{1i}, z_{2i}, \dots, z_{pi}) = \frac{1}{n} \sum_{j=1}^{n} \prod_{k=1}^{p} I(z_{kj}, z_{ki}), \qquad (i = 1, 2, \dots, n) \tag{3.9}$$

The distribution value obtained from Equation (3.9) depends on the distribution. This eliminates the validity of the test. A different approach is proposed to make a more valid calculation in Equation (3.10).

$$F_n^*(\mathbf{z}) = \frac{1}{n} \sum_{j=1}^{n} I\Big(F_0^*(z_{1j}, z_{2j}, \dots, z_{pj}), F_0^*(\mathbf{z})\Big) \tag{3.10}$$

In this approach, the calculation is made according to the distribution values instead of the points. The obtained values have a joint distribution which does not depend on F_0, so they are distribution-free. They have a uniform distribution in the range $[0, 1]$. Since the distributions are monotone increasing function, the approach given in Equation (3.10) yields the same result as Equation (3.4) in the case of the univariate empirical distribution function.

3.3. Computation of the theoretical distribution values

The theoretical distribution function, which is calculated with the help of being the independent variables in Equation (3.8), is obtained by using definitions of Equations (3.3)-(3.4) as follows

$$F_0^*(\mathbf{z}) = \prod_{k=1}^{p} U_k. \tag{3.11}$$

U_k are the independent random variables which are uniformly distributed in the range of $[0, 1]$. The probability density function of the multiplication of U_k is defined as in Equation (3.12).

$$f_U^{(p)}(u) = \frac{(-1)^{p-1}}{(p-1)!} \log^{p-1}(u) \tag{3.12}$$

In addition, the distribution function is calculated as below

$$F_U^{(p)}(u) = \sum_{k=1}^{p} \frac{(-1)^{k-1}}{(k-1)!} u \log^{k-1}(u). \tag{3.13}$$

If the value of F_0^* is substituted in the function $F_U^{(p)}(u)$ as follows, the theoretical distribution function of the independent random values from the uniform distribution in the range of $[0, 1]$ is obtained.

$$T_0(\mathbf{z}) = F_U^{(p)}\Big(F_0^*(\mathbf{z})\Big) \tag{3.14}$$

3.4. The proposed multivariate Kolmogorov-Smirnov statistic

After the theoretical distribution value and the empirical distribution value are calculated, the proposed multivariate Kolmogorov-Smirnov test statistic is obtained as follows

$$D_n^{(p)} = \sup_{\mathbf{z}} \Big| F_n^*(\mathbf{z}) - T_0(\mathbf{z}) \Big|. \tag{3.15}$$

$F_n^*(\mathbf{z})$ is calculated as Equation (3.10) and $T_0(\mathbf{z})$ is calculated as Equation (3.14).

3.5. Algorithm of the proposed multivariate Kolmogorov-Smirnov test

In this study, a new algorithm is developed for the multivariate Kolmogorov-Smirnov test. It calculates the multivariate KS (MVKS) test statistics for d-dimensional data. The decision rule is performed for the null hypothesis (H_0) by comparing the calculated KS test statistic $\left(D_n^{(p)}\right)$ with the KS table value ($K_{n,1-\alpha}$). Algorithm of the MVKS test is introduced in the Algorithm 1.

4. Experimental results

Two different simulation studies are performed to test the accuracy and validity of the proposed multivariate KS test statistic. In the first part of the simulation, the critical table values of the proposed multivariate KS test method are obtained by the Monte Carlo approach. Lastly, several bivariate KS tests in the literature and the proposed multivariate KS test (bivariate case) statistics are compared in terms of type I error and power. The simulation study is made on a computer that has Intel® Core (TM) i7-4740 CPU, 16 GB of RAM. Also, Matlab® R17b software is used to make the power calculations.

Algorithm 1 Algorithm of the MVKS Test

Determination of the initial parameter

Step 1. Determine the data (x) and the significant value (α),

Step 2. If the proposed Kolmogorov-Smirnov test is used, determine the mean vector and the covariance matrix parameters of the distribution $F_0(x;\theta,\ldots)$, Else, estimate the mean vector and the covariance matrix parameters from the data (x) for the Lilliefors test.

Independence of variables

Step 3. Subtract the mean vector from x,

Step 4. Multiply the zero-mean data by the inverse of the square root of the covariance matrix.

Calculation of the KS test statistics

Step 5. Evaluate $F_0^*(z_{1i}, z_{2i}, \ldots, z_{pi})$ for $i = 1, 2, \ldots, n$.

Step 6. Calculate the empirical distribution value via Equation (3.10),

Step 7. Calculate the theoretical distribution value by Equation (3.14),

Step 8. Calculate the Kolmogorov-Smirnov statistic in Equation (3.15).

Comparison of the test statistics with the table value

Step 9. IF $D_n^{(p)} < K_{n,1-\alpha}$, accept the H_0 hypothesis, ELSE reject the H_0 hypothesis.

4.1. Calculation of the proposed KS test tables for multivariate case

In this section, the critical table values for the proposed multivariate KS test method are obtained by using the Monte Carlo approach. Each of the tables is created with different dimensions ($d = 1, 5, 10$) and different quantiles ($0.8, 0.9, 0.95, 0.975, 0.99, 0.995, 0.999$). While the critical table values are calculated, the samples are generated from a normal distribution considering that the samples are dependent. The mean vector and covariance matrix of the multivariate normal distribution are selected randomly. Likewise, the same critical table values can be obtained for different mean vectors and covariance matrices. Also, the mean absolute percentage error (MAPE) which is widely known for performance criteria is calculated for each table with Equation (4.1) [12, 32, 35].

$$MAPE = \frac{1}{N_C} \sum_{i=1}^{N_C} \frac{|T_i^S - T_i^C|}{T_i^C} \tag{4.1}$$

N_C is the number of total cells. T_i^S is i^{th} critical table value calculated from the simulation and T_i^C is critical value of the univariate KS test using R program. While the MAPE is calculated, the univariate KS test table is computed by the function "ks.test" in R is used.

Firstly, the proposed multivariate KS test statistics is calculated by generating samples at different sample sizes from the univariate normal distribution with the mean ($\mu_1 = 3$) and the variance ($\sigma_1^2 = 25$) for the univariate case ($d = 1$). The univariate KS test table values obtained from 100,000 replications are given in Table 1. The MAPE for Table 1 is found as 0.0021. The similarity of Table 1 with critical table of the univariate KS test is 99.79%.

The quantiles of the univariate Kolmogorov-Smirnov statistics obtained with 100,000 replications are calculated with different means and variances. The critical table values are obtained by selecting randomly the mean (μ_1) from a uniform distribution in the interval of $[-1000,\ 1000]$ and the variance (σ_1^2) from a uniform distribution in the interval of [0,10000]. The average MAPE values at the end of 100 trials is calculated as 0.007.

However, comparisons of critical table values obtained from the simulation (T^S) and the univariate KS test using R program (T^C) with different sample sizes for $d = 1$ is given

in Figure 1 according to different quantiles. When the graphs are examined, it is observed that the T^S and T^C have almost the same values for the given quantiles.

Table 1. Monte-Carlo approximation to the quantiles of the univariate Kolmogorov-Smirnov statistics with 100,000 replications ($d = 1$)

n \ $1-\alpha$	0.8	0.9	0.95	0.975	0.99	0.995	0.999
10	0.3221	0.3678	0.4093	0.4453	0.4874	0.5177	0.5829
11	0.3081	0.3525	0.3915	0.4270	0.4680	0.4969	0.5598
12	0.2949	0.3378	0.3748	0.4083	0.4491	0.4753	0.5354
13	0.2855	0.3261	0.3618	0.3950	0.4343	0.4603	0.5180
14	0.2748	0.3147	0.3486	0.3802	0.4155	0.4432	0.4945
15	0.2656	0.3032	0.3365	0.3675	0.4045	0.4296	0.4832
20	0.2312	0.2647	0.2945	0.3211	0.3527	0.3749	0.4229
25	0.2079	0.2373	0.2637	0.2877	0.3168	0.3367	0.3747
30	0.1899	0.2171	0.2415	0.2637	0.2901	0.3079	0.3498
40	0.1657	0.1893	0.2103	0.2297	0.2526	0.2678	0.3004
50	0.1486	0.1699	0.1883	0.2054	0.2253	0.2402	0.2720
60	0.1356	0.1548	0.1721	0.1877	0.2065	0.2193	0.2465
80	0.1180	0.1347	0.1497	0.1631	0.1794	0.1920	0.2140
100	0.1058	0.1208	0.1341	0.1464	0.1614	0.1709	0.1911
150	0.0864	0.0986	0.1096	0.1195	0.1311	0.1393	0.1566
200	0.0751	0.0857	0.0951	0.1034	0.1142	0.1217	0.1372
300	0.0613	0.0701	0.0779	0.0850	0.0936	0.0997	0.1125

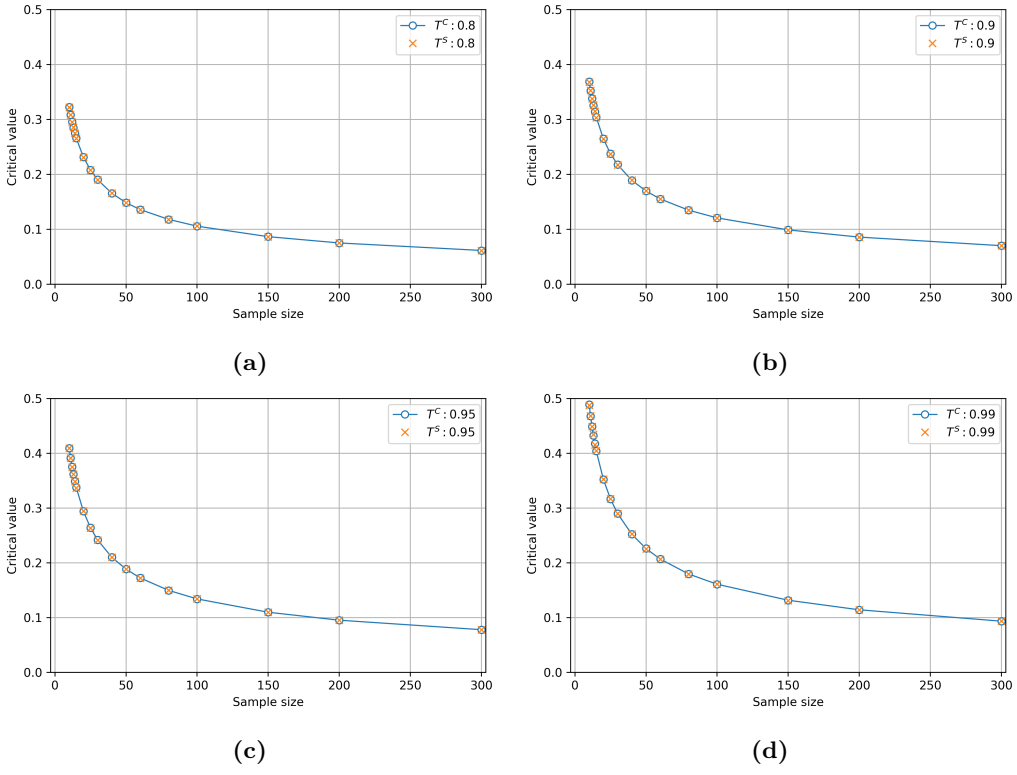

Figure 1. Comparisons of critical table values obtained from the simulation (T^S) and the univariate KS test using R program (T^C) with different sample sizes for $d = 1$; (a) $(1-\alpha) = 0.80$, (b) $(1-\alpha) = 0.90$, (c) $(1-\alpha) = 0.95$, (d) $(1-\alpha) = 0.99$

In case the proposed multivariate KS test statistic is five-variate ($d = 5$), the test statistics is calculated from the five-variate normal distribution with the mean vector (μ_2) and covariance matrix (Σ_2) respectively as follows

$$\mu_2 = \begin{bmatrix} -1 & 2 & 1 & 3 & 4 \end{bmatrix},$$

$$\Sigma_2 = \begin{bmatrix} 3 & 0.7 & -1 & 0.3 & 0.2 \\ 0.7 & 1 & -0.4 & 0.5 & 0.8 \\ -1 & -0.4 & 2 & -0.2 & 0.9 \\ 0.3 & 0.5 & -0.2 & 5 & 0.1 \\ 0.2 & 0.8 & 0.9 & 0.1 & 4 \end{bmatrix}.$$

The five-variate KS test critical table values calculated with 100,000 replications are obtained as in Table 2. The MAPE for Table 2 is found as 0.0066. The similarity of Table 2 with critical table of the univariate KS test is 99.34%.

Table 2. Monte-Carlo approximation to the quantiles of the five-variate Kolmogorov-Smirnov statistics with 100,000 replications ($d = 5$)

n \ $1-\alpha$	0.8	0.9	0.95	0.975	0.99	0.995	0.999
10	0.3231	0.3693	0.4082	0.4429	0.4870	0.5178	0.5832
11	0.3087	0.3555	0.3927	0.4288	0.4763	0.5024	0.5748
12	0.2976	0.3398	0.3793	0.4144	0.4585	0.4891	0.5461
13	0.2844	0.3260	0.3648	0.3951	0.4343	0.4689	0.5211
14	0.2753	0.3147	0.3511	0.3846	0.4206	0.4430	0.4827
15	0.2652	0.3049	0.3379	0.3710	0.4074	0.4307	0.4775
20	0.2314	0.2653	0.2952	0.3221	0.3567	0.3796	0.4228
25	0.2090	0.2392	0.2676	0.2921	0.3207	0.3389	0.3800
30	0.1901	0.2169	0.2420	0.2641	0.2908	0.3084	0.3448
40	0.1652	0.1901	0.2116	0.2322	0.2524	0.2672	0.3067
50	0.1479	0.1699	0.1898	0.2077	0.2269	0.2395	0.2717
60	0.1362	0.1554	0.1709	0.1871	0.2079	0.2200	0.2475
80	0.1175	0.1339	0.1473	0.1617	0.1788	0.1902	0.2061
100	0.1052	0.1202	0.1337	0.1460	0.1609	0.1676	0.1875
150	0.0872	0.0986	0.1103	0.1196	0.1311	0.1401	0.1600
200	0.0755	0.0859	0.0953	0.1039	0.1151	0.1217	0.1359
300	0.0614	0.0700	0.0781	0.0852	0.0926	0.0975	0.1085

The quantiles of the five-variate Kolmogorov-Smirnov statistics are also calculated with different mean vectors and covariance matrices. The mean vector (μ_2) is randomly selected from the uniform distribution in the range of $[-1000,\ 1000]$. In addition, the table values are obtained by randomly selecting the diagonal elements (variance values) of the covariance matrix (Σ_2) from the uniform distribution in the range of $[0,\ 100]$ and the correlation values from the uniform distribution in the range of $[-1,\ 1]$. The average MAPE values at the end of 100 trials is calculated as 0.0221.

However, comparisons of critical table values obtained from the simulation (T^S) and the univariate KS test using R program (T^C) with different sample sizes for $d = 5$ is given in Figure 2 according to different quantiles. When the graphs are examined, it is observed that the T^S and T^C have almost the same values for the given quantiles like $d = 1$.

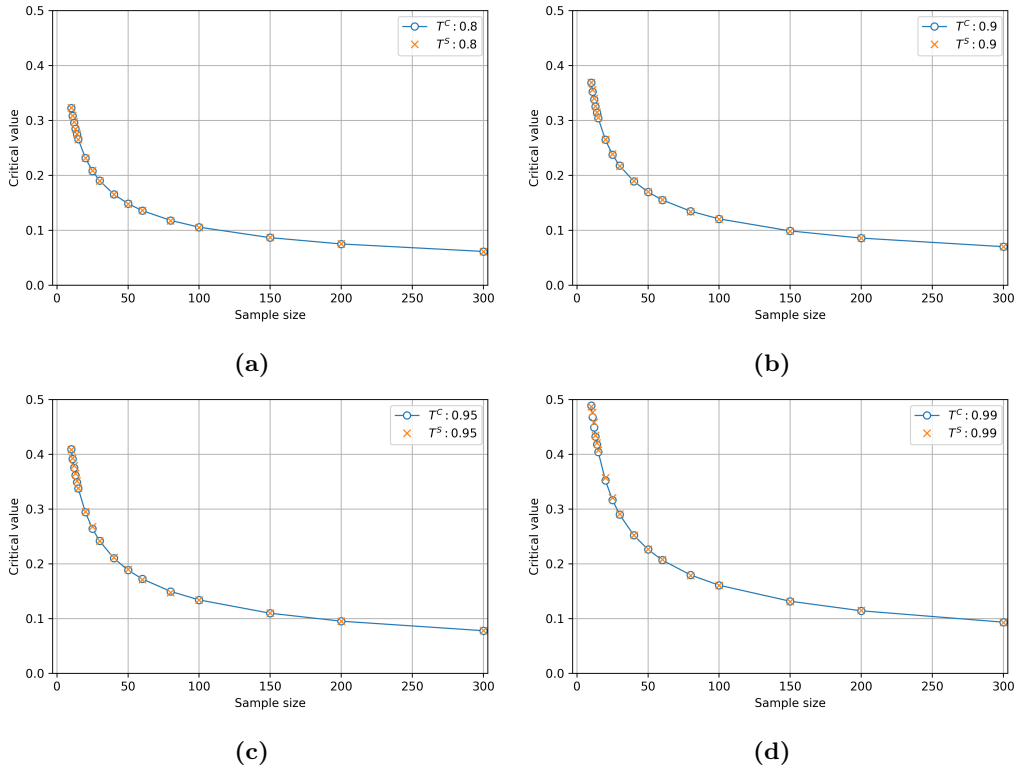

Figure 2. Comparisons of critical table values obtained from the simulation (T^S) and the univariate KS test using R program (T^C) with different sample sizes for $d = 5$; (a) $(1-\alpha) = 0.80$, (b) $(1-\alpha) = 0.90$, (c) $(1-\alpha) = 0.95$, (d) $(1-\alpha) = 0.99$

In case the proposed multivariate KS test statistic is ten-variate ($d = 10$), the test statistics is calculated from the ten-variate normal distribution with the mean vector (μ_3) and variance matrix (Σ_3) receptively as follows

$$\mu_3 = \begin{bmatrix} 680 & 250 & -489 & 809 & 534 & 125 & 794 & -263 & -324 & 237 \end{bmatrix},$$

$$\Sigma_3 = \begin{bmatrix} 9.60 & -0.81 & 0.03 & 0.90 & 0.69 & 0.87 & 0.95 & -0.25 & -0.36 & 0.57 \\ -0.81 & 38.17 & -0.58 & 0.92 & 0.86 & -0.27 & -0.31 & 0.39 & 0.43 & 0.60 \\ 0.03 & -0.58 & 19.55 & 0.19 & -0.81 & 0.12 & -0.91 & 0.52 & 0.32 & 0.20 \\ 0.90 & 0.92 & 0.19 & 81.24 & -0.55 & 0.43 & 0.45 & -0.31 & -0.93 & 0.69 \\ 0.69 & 0.86 & -0.81 & -0.55 & 72.6 & 0.54 & 0.02 & 0.41 & 0.58 & -0.34 \\ 0.87 & -0.27 & 0.12 & 0.43 & 0.54 & 67.79 & 0.83 & 0.15 & -0.42 & -0.47 \\ 0.95 & -0.31 & -0.91 & 0.45 & 0.02 & 0.83 & 91.48 & 0.78 & -0.27 & -0.70 \\ -0.25 & 0.39 & 0.52 & -0.31 & 0.41 & 0.15 & 0.78 & 82.28 & 0.89 & 0.30 \\ -0.36 & 0.43 & 0.32 & -0.93 & 0.58 & -0.42 & -0.27 & 0.89 & 55.32 & 0.96 \\ 0.57 & 0.60 & 0.20 & 0.69 & -0.34 & -0.47 & -0.70 & 0.30 & 0.96 & 2.54 \end{bmatrix}.$$

The ten-variate KS test critical table values calculated with 100,000 replications are obtained as in Table 3. The MAPE for Table 3 is found as 0.0215. The similarity of Table 3 with critical table of the univariate KS test is 97.85%.

The quantiles of the ten-variate Kolmogorov-Smirnov statistics are also calculated with different mean vectors and covariance matrices. The mean vector (μ_3) is randomly selected from the uniform distribution in the range of $[-1000,\ 1000]$. In addition, the table values are obtained by randomly selecting the diagonal elements (variance values) of the

covariance matrix (Σ_3) from the uniform distribution in the range of $[0,\ 100]$ and the correlation values from the uniform distribution in the range of $[-1,\ 1]$. The average MAPE values at the end of 100 trials is calculated as 0.02.

Table 3. Monte-Carlo approximation to the quantiles of the ten-variate Kolmogorov-Smirnov statistics with 100,000 replications ($d = 10$)

n \ $1-\alpha$	0.8	0.9	0.95	0.975	0.99	0.995	0.999
10	0.3229	0.3659	0.4096	0.4562	0.4980	0.5275	0.5780
11	0.3062	0.3458	0.3808	0.4260	0.4623	0.4775	0.4993
12	0.2855	0.3337	0.3676	0.4071	0.4714	0.5053	0.5652
13	0.2826	0.3241	0.3584	0.3927	0.4325	0.4769	0.5334
14	0.2842	0.3220	0.3542	0.3786	0.4236	0.4524	0.4874
15	0.2657	0.3068	0.3474	0.3709	0.3978	0.4200	0.5044
20	0.2335	0.2629	0.2919	0.3120	0.3535	0.3913	0.4158
25	0.2129	0.2395	0.2629	0.2871	0.3082	0.3346	0.3619
30	0.1913	0.2227	0.2446	0.2668	0.2905	0.3079	0.3242
40	0.1665	0.1939	0.2139	0.2315	0.2475	0.2573	0.2760
50	0.1473	0.1697	0.1920	0.2046	0.2267	0.2351	0.2540
60	0.1332	0.1548	0.1695	0.1857	0.2112	0.2159	0.2666
80	0.1182	0.1345	0.1464	0.1566	0.1717	0.1871	0.1991
100	0.1072	0.1245	0.1347	0.1476	0.1557	0.1642	0.1923
150	0.0891	0.1015	0.1129	0.1243	0.1377	0.1467	0.1602
200	0.0768	0.0856	0.0935	0.1008	0.1117	0.1179	0.1278
300	0.0609	0.0696	0.0766	0.0835	0.0918	0.1036	0.1118

However, comparisons of critical table values obtained from the simulation (T^S) and the univariate KS test using R program (T^C) with different sample sizes for $d = 10$ is given in Figure 3 according to different quantiles. When the graphs are examined, it is observed that the T^S and T^C have almost the same values for the given quantiles like $d = 1$ and $d = 5$.

On the other hand, if the mean vector and covariance matrix are not given in the null hypothesis and they are calculated from the sample, the Kolmogorov-Smirnov test statistic becomes the Lilliefors test. In this case, it is necessary to look at the Lilliefors table instead of the KS table. In this study, the related critical table values are calculated only if the data is bivariate. The bivariate Lilliefors test table at different quantiles is shown in Table 4. The mean vector and the variance matrix of the normal distribution are utilized for obtaining Table 4 respectively as follows

$$\mu_4 = [1 \quad 2], \quad \Sigma_4 = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}.$$

Similarly, when the number of variables is more than 2, the table values can be calculated as in the univariate and the bivariate cases via the given algorithm in the Section 3.5.

The MAPE for Table 4 is found as 0.0172. The similarity of the univariate Lilliefors test with Table 4 is 98.28%. The MAPE is calculated by using critical table of the univariate Lilliefors test [23] in the literature.

However, comparisons of critical table values obtained from the simulation (T^S) and the univariate KS test using R program (T^C) with different sample sizes for $d = 2$ is given in Figure 4 according to different quantiles. When the graphs are examined, it is observed that the T^S and T^C have almost the same values for the given quantiles.

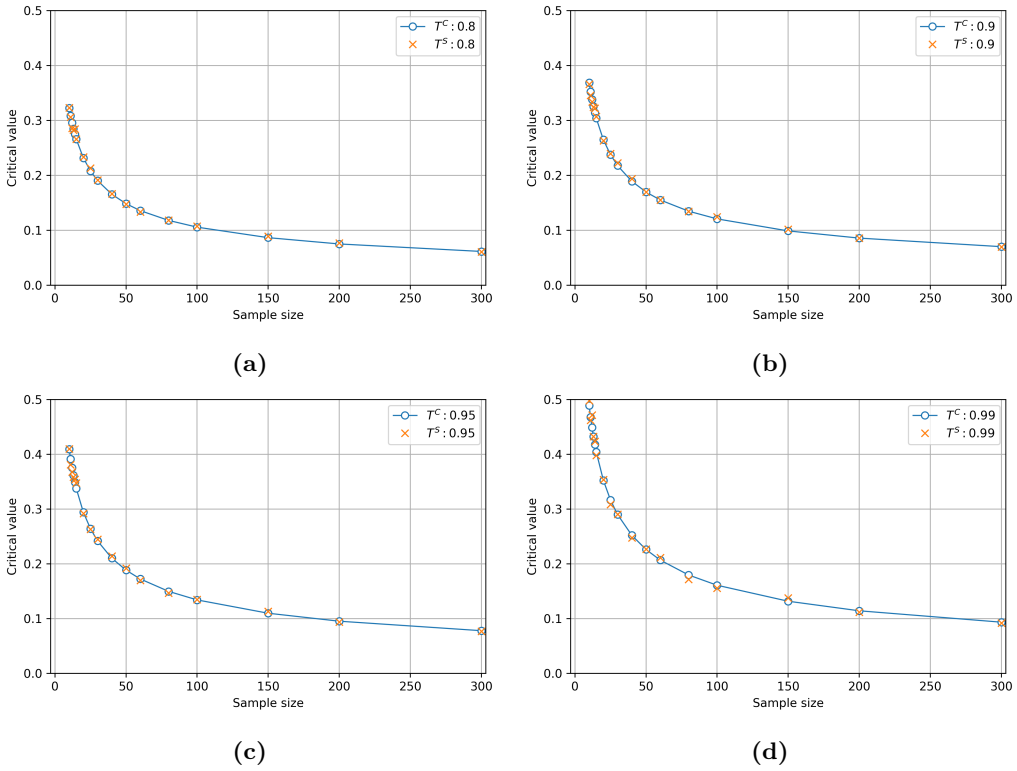

Figure 3. Comparisons of critical table values obtained from the simulation (T^S) and the univariate KS test using R program (T^C) with different sample sizes for $d = 10$; (a) $(1-\alpha) = 0.80$, (b) $(1-\alpha) = 0.90$, (c) $(1-\alpha) = 0.95$, (d) $(1-\alpha) = 0.99$

Table 4. Monte-Carlo approximation to the quantiles of the bivariate Lilliefors statistics with 100,000 replications ($d = 2$)

n \ $1-\alpha$	0.8	0.9	0.95	0.975	0.99	0.995	0.999
10	0.2189	0.2430	0.2637	0.2829	0.3053	0.3208	0.3519
11	0.2102	0.2334	0.2532	0.2713	0.2922	0.3077	0.3381
12	0.2023	0.2240	0.2434	0.2613	0.2823	0.2957	0.3269
13	0.1952	0.2167	0.2352	0.2528	0.2730	0.2879	0.3162
14	0.1893	0.2099	0.2280	0.2441	0.2643	0.2774	0.3062
15	0.1830	0.2033	0.2211	0.2373	0.2563	0.2692	0.2966
20	0.1607	0.1784	0.1943	0.2082	0.2252	0.2368	0.2612
25	0.1449	0.1610	0.1749	0.1881	0.2034	0.2147	0.2363
30	0.1330	0.1477	0.1606	0.1725	0.1866	0.1972	0.2186
40	0.1159	0.1286	0.1399	0.1506	0.1635	0.1724	0.1916
50	0.1040	0.1155	0.1257	0.1355	0.1463	0.1543	0.1693
60	0.0954	0.1060	0.1152	0.1237	0.1338	0.1407	0.1581
80	0.0831	0.0924	0.1007	0.1081	0.1174	0.1234	0.1357
100	0.0746	0.0828	0.0901	0.0968	0.1050	0.1108	0.1235
150	0.0613	0.0680	0.0741	0.0799	0.0865	0.0912	0.1013
200	0.0532	0.0590	0.0642	0.0690	0.0749	0.0791	0.0873
300	0.0435	0.0483	0.0526	0.0564	0.0615	0.0650	0.0725

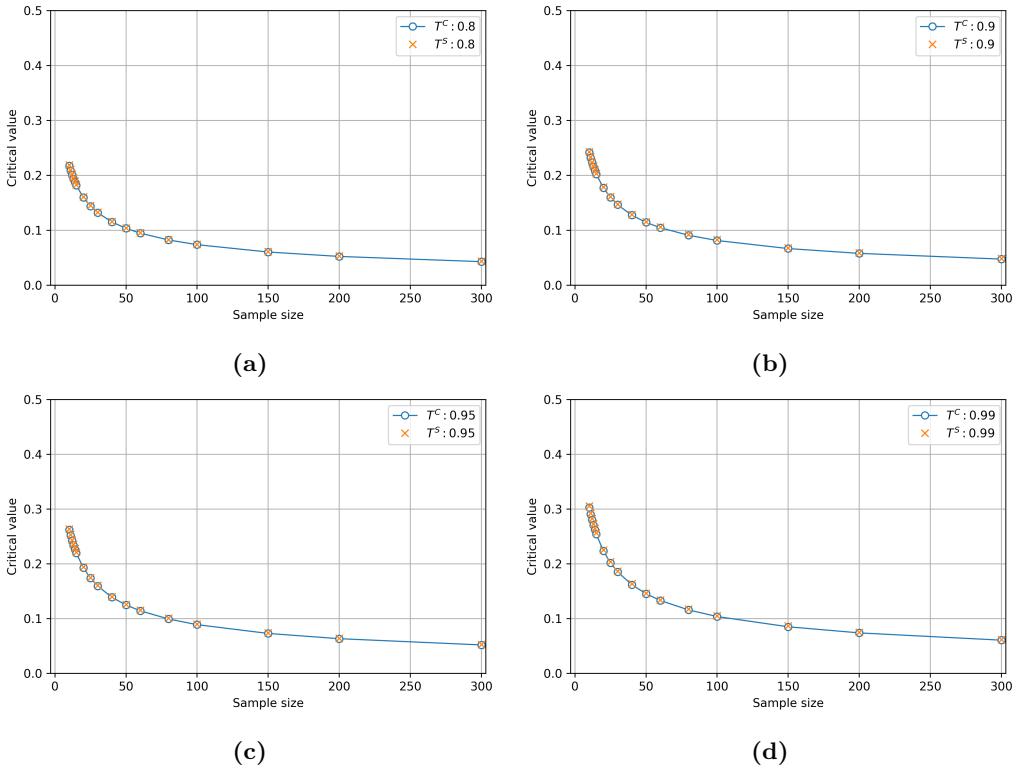

Figure 4. Comparisons of critical table values obtained from the simulation (T^S) and the univariate Lilliefors test using R program (T^C) with different sample sizes for $d = 2$; (a)$(1-\alpha) = 0.80$, (b) $(1-\alpha) = 0.90$, (c) $(1-\alpha) = 0.95$, (d) $(1-\alpha) = 0.99$

4.2. The power comparisons of the tests

In this section, the bivariate KS test which are improved from Peacock [36], Fasano and Franceschini [14], and Justel et al. [18] are compared with the proposed multivariate KS (bivariate case) test in terms of type I error and the power. Simulated critical values are utilized for the all tests. In the comparisons, the bivariate KS test are represented by P (Peacock), FF (Fasano and Franceschini), J (Justel), and MVKS (the proposed method). The Type I error (α) and the power values ($1 - \beta$) are calculated with a simulation performed 10,000 times.

Furthermore, the simulation of the power comparisons consists of three parts. The first part of the simulation, the type I errors are calculated at different sample sizes as $n = 15, 30, 50, 100, 200, 500$ for the four methods. In the second part of the simulation, the null hypothesis is determined as the bivariate standard normal distribution. The alternative hypothesis is chosen from the multivariate symmetric (elliptical) distributions and the powers of the tests are calculated for the four methods. In the last part, the null hypothesis is determined as in the bivariate standard normal distribution and the alternative hypothesis is chosen from the multivariate asymmetric distributions. The power of test is also calculated for the four methods.

4.2.1. Type-I error comparisons for the P, FF, J, and MVKS tests. In the first part of simulation, the type I error are calculated at different sample sizes as $n = 15, 30, 50, 100, 200, 500$ for the four methods. Also, the null hypothesis is defined as a bivariate normal distribution with the mean vector μ. The covariance matrices Σ_1 and Σ_2 are determined respectively as follows

$$\mu = \begin{bmatrix} 0 & 0 \end{bmatrix}, \quad \Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Sigma_2 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}.$$

If the data comes from a bivariate normal distribution, it is expected that each test will reject the null hypothesis at the 5% level. The Type I error of the bivariate KS statistics for Σ_1 and Σ_2 are given in Table 5.

Table 5. Type I error of the bivariate KS statistics for four methods

Covariance Matrix	Method	Sample Size					
		15	**30**	**50**	**100**	**200**	**500**
$\boldsymbol{\Sigma_1}$	**P**	0.0066	0.005	0.0078	0.0093	0.0118	0.016
	FF	0.0546	0.0425	0.0452	0.0508	0.0463	0.0586
	J	0.2505	0.2138	0.1902	0.1656	0.1625	0.1514
	MVKS	0.0479	0.0493	0.0499	0.0501	0.0507	0.0486
$\boldsymbol{\Sigma_2}$	**P**	0.0058	0.0062	0.0079	0.0093	0.0091	0.0135
	FF	0.0500	0.0397	0.0416	0.0463	0.0493	0.0552
	J	0.2516	0.2127	0.1925	0.1723	0.1575	0.1434
	MVKS	0.0508	0.0478	0.0478	0.0507	0.0487	0.0465

As the sample size increases, the Type I error of the P and the J methods increase for Σ_1 and Σ_2. However, while Type I error for the P method does not exceed 2% for each sample size, Type I error for the J method is quite more than 5%. When the Type I errors of the P and the J methods are examined for covariance matrices Σ_1 and Σ_2, it is seen that these two tests do not perform consistently. On the other hand, the Type I error of the FF and the MVKS change at 5% level for Σ_1 and Σ_2. In this case, these methods perform more consistent and correct than both the P and the J methods.

4.2.2. Power comparisons for multivariate symmetric distributions. In the second part of the simulation, random numbers are generated from the multivariate elliptical distributions such as Normal, Laplace, Student-t, Cauchy, Logistic, and Triangular distributions and the power $(1-\beta)$ of four tests are calculated. The null hypothesis is defined as a bivariate standard normal distribution.

The mean vectors $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5)$ and covariance matrix (Σ_1) of the normal distribution are utilized for obtaining Table 8 as follows, respectively. The power comparisons of the bivariate KS statistics for Σ_1 are given in Table 6.

μ_1		μ_2		μ_3		μ_4		μ_5	
$[-1$	$-1]$	$[-0.5$	$-0.5]$	$[0$	$0]$	$[0.5$	$0.5]$	$[1$	$1]$

$$\Sigma_1 = \begin{bmatrix} 1 & 0.2 \\ 0.2 & 1 \end{bmatrix}$$

As the change in the mean vector increases or decreases according to $\mu_3 = [0 \quad 0]$, the powers of the P, FF, J, and MVKS methods increase, provided that the covariance matrix Σ_1 of the bivariate normal distribution remains constant. Although the power of the P method is low for the small samples, the power of test increases for the four methods as the sample size increases under the same conditions.

On the other hand, as the number of sample size increases in case of the mean vector is $\mu_3 = [0 \quad 0]$, the power of test is low for the P method and high for the J method according to significant level (0.05). The power of test is quite close to significant level (0.05), regardless of the number of sample sizes for the FF and MVKS methods. Furthermore, the correlation value is 0.2. As the correlation degree of covariance matrix increase, the FF, J, and MVKS methods will be more powerful. Also, the MVKS method has higher

power than the FF method in all samples.

Table 6. Power comparisons of the bivariate KS statistics for Σ_1

		Method			
Sample Size	**Mean Vector**	**P**	**FF**	**J**	**MVKS**
15	$\boldsymbol{\mu_1}$	0.8565	0.9757	0.9993	0.987
	$\boldsymbol{\mu_2}$	0.2065	0.4898	0.8635	0.537
	$\boldsymbol{\mu_3}$	0.0075	0.0546	0.2493	0.0628
	$\boldsymbol{\mu_4}$	0.1949	0.4781	0.7555	0.6198
	$\boldsymbol{\mu_5}$	0.8617	0.977	0.9984	0.9917
30	$\boldsymbol{\mu_1}$	0.9988	0.9999	1	1
	$\boldsymbol{\mu_2}$	0.5582	0.8002	0.9758	0.8308
	$\boldsymbol{\mu_3}$	0.0088	0.06	0.227	0.0636
	$\boldsymbol{\mu_4}$	0.5427	0.7942	0.9242	0.8921
	$\boldsymbol{\mu_5}$	0.999	1	1	1
50	$\boldsymbol{\mu_1}$	1	1	1	1
	$\boldsymbol{\mu_2}$	0.8784	0.9642	0.9973	0.9643
	$\boldsymbol{\mu_3}$	0.013	0.0691	0.2197	0.0736
	$\boldsymbol{\mu_4}$	0.8734	0.9628	0.9866	0.9889
	$\boldsymbol{\mu_5}$	1	1	1	1
100	$\boldsymbol{\mu_1}$	1	1	1	1
	$\boldsymbol{\mu_2}$	0.9976	0.9998	1	0.9999
	$\boldsymbol{\mu_3}$	0.0242	0.0937	0.2145	0.0885
	$\boldsymbol{\mu_4}$	0.9984	1	1	1
	$\boldsymbol{\mu_5}$	1	1	1	1

The power values obtained from the test statistic by changing the covariance matrix as $\Sigma_2 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ in the alternative hypothesis are given in Table 7, provided that the null hypothesis remains the same.

Table 7. Power comparisons of the bivariate KS statistics for Σ_2

		Method			
Sample Size	**Mean Vector**	**P**	**FF**	**J**	**MVKS**
15	$\boldsymbol{\mu_1}$	0.8873	0.9758	0.9994	0.9638
	$\boldsymbol{\mu_2}$	0.2768	0.5379	0.886	0.4319
	$\boldsymbol{\mu_3}$	0.02	0.0984	0.2954	0.0911
	$\boldsymbol{\mu_4}$	0.2945	0.5522	0.6396	0.6781
	$\boldsymbol{\mu_5}$	0.8861	0.9734	0.9892	0.9914
30	$\boldsymbol{\mu_1}$	0.9989	0.9999	1	0.9999
	$\boldsymbol{\mu_2}$	0.6818	0.8605	0.9828	0.7254
	$\boldsymbol{\mu_3}$	0.0324	0.1453	0.327	0.1202
	$\boldsymbol{\mu_4}$	0.6847	0.8593	0.8394	0.9237
	$\boldsymbol{\mu_5}$	0.9989	1	1	1
50	$\boldsymbol{\mu_1}$	1	1	1	1
	$\boldsymbol{\mu_2}$	0.9377	0.9807	0.9991	0.9112
	$\boldsymbol{\mu_3}$	0.0693	0.2304	0.3763	0.1518
	$\boldsymbol{\mu_4}$	0.9386	0.9826	0.9519	0.9929
	$\boldsymbol{\mu_5}$	1	1	1	1
100	$\boldsymbol{\mu_1}$	1	1	1	1
	$\boldsymbol{\mu_2}$	0.9996	0.9999	1	0.9966
	$\boldsymbol{\mu_3}$	0.1962	0.4839	0.511	0.2782
	$\boldsymbol{\mu_4}$	1	1	0.9992	1
	$\boldsymbol{\mu_5}$	1	1	1	1

As the change in the mean vector increases or decreases according to $\mu_3 = [0 \quad 0]$, the power of the P, FF, J, and MVKS methods increases, provided that the covariance matrix Σ_2 of the bivariate normal distribution remains constant. Although the power of the P method is low for smaller samples, the power of test increases for the four methods as the sample size increases under the same conditions as Table 6.

On the other hand, when the mean vector is $\mu_3 = [0 \quad 0]$, the distribution in the null and alternative hypothesis are nearly the same. The power of test is low for the P method and high for the J method according to significant level (0.05) in case the mean vector is $\mu_3 = [0 \quad 0]$. The powers of the FF and MVKS methods increase according to significant level (0.05) in case the mean vector is $\mu_3 = [0 \quad 0]$ as the number of samples size increases.

The power values obtained from the test statistic by changing the covariance matrix as $\Sigma_3 = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$ in the alternative hypothesis are given in Table 8, provided that the null hypothesis remains the same.

Table 8. Power comparisons of the bivariate KS statistics for Σ_3

		Method			
Sample Size	**Mean Vector**	**P**	**FF**	**J**	**MVKS**
15	$\boldsymbol{\mu_1}$	0.9376	0.9853	0.9993	0.9274
	$\boldsymbol{\mu_2}$	0.4295	0.68	0.9196	0.3719
	$\boldsymbol{\mu_3}$	0.0611	0.2237	0.397	0.143
	$\boldsymbol{\mu_4}$	0.4381	0.6834	0.5424	0.7296
	$\boldsymbol{\mu_5}$	0.9391	0.9866	0.969	0.9929
30	$\boldsymbol{\mu_1}$	0.9996	0.9999	1	0.9974
	$\boldsymbol{\mu_2}$	0.8518	0.9516	0.9931	0.6362
	$\boldsymbol{\mu_3}$	0.1532	0.4388	0.5318	0.2247
	$\boldsymbol{\mu_4}$	0.8515	0.9507	0.7541	0.9515
	$\boldsymbol{\mu_5}$	0.9998	1	0.9994	1
50	$\boldsymbol{\mu_1}$	1	1	1	1
	$\boldsymbol{\mu_2}$	0.9895	0.9989	0.9998	0.8471
	$\boldsymbol{\mu_3}$	0.3574	0.7548	0.6851	0.3406
	$\boldsymbol{\mu_4}$	0.987	0.9975	0.9044	0.9976
	$\boldsymbol{\mu_5}$	1	1	1	1
100	$\boldsymbol{\mu_1}$	1	1	1	1
	$\boldsymbol{\mu_2}$	1	1	1	0.9889
	$\boldsymbol{\mu_3}$	0.9036	0.9975	0.8977	0.6201
	$\boldsymbol{\mu_4}$	1	1	0.9963	1
	$\boldsymbol{\mu_5}$	1	1	1	1

As the change of the mean vector of the bivariate standard normal distribution increases, the power of the P, the FF, J, and MVKS methods also rises in Table 8 as Table 6 and Table 7. As the relationship between the two variables increases, the power of the tests are higher than according to Table 6 and Table 7 under the same conditions. As the number of sample size increases, the power of the tests rises for the four methods. Although the P method has lower power in the case of small samples with the same mean vector and covariance matrix as Table 6 and Table 7.

The power of the P, FF, J, and the MVKS methods are calculated using bivariate Laplace distribution. The power values performed at the $n = 10, 30, 50, 100, 200, 300$ sample sizes are given in Table 9.

As the number of sample size increases, the powers of four methods increase for the bivariate Laplace distribution. The P method has the lower power than the others. The J method has the highest power in all sample sizes. However, it is more accurate to compare the powers of the test between the FF and MVKS, considering that the P and J methods do not perform consistently and correctly in the Type I error comparisons. Then, MVKS

has a higher power in case of small sample sizes. As the number of sample size increases, the powers of both the FF and MVKS methods increase to 1.

Table 9. Power test comparisons for the bivariate Laplace distribution

	Sample Size					
Method	**10**	**30**	**50**	**100**	**200**	**300**
P	0.0181	0.0919	0.3166	0.8905	0.9998	1
FF	0.1522	0.3775	0.6802	0.9839	1	1
J	0.4624	0.7272	0.89	0.9947	1	1
MVKS	0.1862	0.4363	0.6757	0.9526	0.9999	1

The power of the P, FF, J, and MVKS methods are calculated using the bivariate Student-t distribution with $\upsilon = 10$ degrees of freedom. The power values performed at the $n = 10, 30, 50, 100, 200, 300$ sample sizes are given in Table 10.

Table 10. Power comparisons for the bivariate Student-t distribution

	Sample Size					
Method	**10**	**30**	**50**	**100**	**200**	**300**
P	0.0074	0.0065	0.0061	0.009	0.0145	0.0201
FF	0.0669	0.0495	0.0472	0.057	0.0707	0.0891
J	0.2441	0.229	0.206	0.1887	0.1991	0.204
MVKS	0.0537	0.0582	0.0588	0.0636	0.0751	0.0791

As the number of sample size increases, the powers of four methods increase for the bivariate Student-t distribution with $\upsilon = 10$ degrees of freedom. The P method has the lower power than the others. The J method has the higher power in all sample sizes than other methods. However, it is more accurate to compare the powers of the test between the FF and the MVKS, considering that the P and J methods do not perform consistently and correctly in Type I error comparisons. Then, the MVKS has higher power value except for sample size 10 and 300. Also, as the number of sample size increases, the powers of both the FF and MVKS methods increase.

The power of test values of the P, FF, J, and MVKS methods are calculated using the bivariate Cauchy distribution. The power values performed at $n = 10, 30, 50, 100, 200, 300$ sample sizes are given in Table 11.

Table 11. Power comparisons for the bivariate Cauchy distribution

	Sample Size					
Method	**10**	**30**	**50**	**100**	**200**	**300**
P	0.0204	0.1505	0.4787	0.9782	1	1
FF	0.1725	0.4917	0.8191	0.9987	1	1
J	0.4945	0.8298	0.9524	0.9997	1	1
MVKS	0.2236	0.5961	0.8458	0.9954	1	1

As the number of sample size increases, the power of four methods increases for the bivariate Cauchy distribution. The P method has the lower power than the others. The J method has a higher power in all sample sizes than other methods. However, it is more accurate to compare the power of the test between FF and MVKS, considering that the P and J methods do not perform consistently and correctly in Type I error comparisons. Then, the MVKS has the same power or higher power than FF method. Also, as the number of sample size increases, the powers of both the FF and MVKS methods increase.

The power of test values of the P, FF, J, and MVKS methods are calculated using the bivariate Logistic distribution. The power values performed at $n = 10, 30, 50, 100, 200, 300$ sample sizes are given in Table 12.

Table 12. Power comparisons for the bivariate Logistic distribution

	Sample Size					
Method	**10**	**30**	**50**	**100**	**200**	**300**
P	0.0391	0.1713	0.4963	0.9744	1	1
FF	0.2641	0.55	0.8408	0.9983	1	1
J	0.6542	0.8444	0.951	0.9993	1	1
MVKS	0.2874	0.5048	0.7251	0.9716	1	1

As the number of sample size increases, the powers of four methods increase for the bivariate Logistic distribution. The P method has the lower power than the others and the J method has a higher power in each sample size than the other methods such as the simulation results for the above distributions. However, it is more accurate to compare the power of the test between the FF and MVKS, considering that the P and J methods do not perform consistently and correctly in the Type I error comparisons. Hence, the MVKS has a lower power within 30, 50, 100 sample sizes than the FF method, but the MVKS method has the same power or higher power in the other sample sizes than the FF method. Also, as the number of sample size increases, the powers of four methods increase to 1.

The power of the P, FF, J, and MVKS methods are calculated using the bivariate Triangular distribution. The power values performed at $n = 10, 30, 50, 100, 200, 300$ sample sizes are given in Table 13.

Table 13. Power comparisons for the bivariate Triangular distribution

	Sample Size					
Method	**10**	**30**	**50**	**100**	**200**	**300**
P	0.0083	0.0231	0.0612	0.2599	0.789	0.9757
FF	0.0957	0.1526	0.271	0.593	0.9513	0.9976
J	0.3499	0.4682	0.5544	0.7586	0.9639	0.9964
MVKS	0.1033	0.1632	0.2397	0.4348	0.7706	0.9299

As the number of sample size increases, the powers of four methods increase for the bivariate Triangular distribution. The P method has the lower power than the others and the J method has a higher power in each sample size than the other methods such as the simulation results for the above distributions. However, it is more accurate to compare the powers between the FF and MVKS, considering that the P and J methods do not perform consistently and correctly in Type I error comparisons. Hence, the proposed method (MVKS) has a lower power within 30, 50, 100 sample sizes than the FF method, but the MVKS method has the same power or higher power in the other sample sizes than the FF method. Also, as the number of sample size increases, the powers of four methods increase to 1.

4.2.3. Power comparisons for multivariate asymmetric distributions. In the last part of the simulation, random numbers are generated from the multivariate asymmetric (non-elliptical) distributions such as Uniform, Exponential, and Chi-Square distributions. The null hypothesis is defined as a bivariate standard normal distribution.

Firstly, the alternative hypothesis is determined by choosing the different parameters of the bivariate uniform distribution. The different parameters of the bivariate uniform distribution are not defined in any interval. The mean vectors $(\mu_1, \mu_2, \mu_3, \mu_4)$ and covariance matrix (Σ_4) for the bivariate uniform distribution are determined as follows

μ_1	μ_2	μ_3	μ_4
[0.5 0.5]	[0.5 1]	[1 1]	[−1 −1]

$$\Sigma_4 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}.$$

The power comparisons of the P, FF, J, and MVKS methods for Σ_4 are given in Table 14.

Table 14. Power comparisons for bivariate Uniform distribution

		Mean Vector			
Sample Size	**Method**	μ_1	μ_2	μ_3	μ_4
15	**P**	0.2154	0.564	0.8391	0.8419
	FF	0.4666	0.8337	0.9673	0.9714
	J	0.6746	0.9239	0.9813	0.9998
	MVKS	0.5913	0.8658	0.9875	0.958
25	**P**	0.4382	0.8837	0.9915	0.9917
	FF	0.6883	0.978	0.9995	0.9995
	J	0.8133	0.9886	0.9982	1
	MVKS	0.8087	0.9785	0.9999	0.9972
50	**P**	0.8583	0.9992	1	1
	FF	0.9566	1	1	1
	J	0.959	0.9999	1	1
	MVKS	0.9783	0.9997	1	1
100	**P**	0.9976	1	1	1
	FF	0.9999	1	1	1
	J	0.9985	1	1	1
	MVKS	0.9999	1	1	1

As the change of the mean vectors increases, the power also increases in the P, FF, J, and MVKS methods. As the number of sample size increases, the powers of four methods increase for the bivariate Uniform distribution. Although, the P and FF methods have lower power than the J and MVKS methods under the same conditions. However, it is more accurate to compare the powers between the FF and MVKS, considering that the P and J methods do not perform consistently and correctly in the Type I error comparisons. When Table 14 is examined, the power of the FF method is only higher than the MVKS method for μ_4 mean vector and small sample sizes ($n = 15$ and $n = 25$). In all other cases, the MVKS method is either higher or has the same power values than the other methods.

The power of the P, FF, J, and MVKS methods are calculated using the bivariate Exponential distribution. The power values performed at $n = 10, 11, 12, 13, 14, 15$ sample sizes are given in Table 15.

Table 15. Power comparisons for bivariate Exponential distribution

	Sample Size					
Method	**10**	**11**	**12**	**13**	**14**	**15**
P	0.703	0.802	0.8557	0.9072	0.937	0.9657
FF	0.9885	0.9966	0.9988	0.9993	0.9999	1
J	1	1	1	1	1	1
MVKS	1	1	1	1	1	1

The power of test is calculated in small sample sizes for the bivariate Exponential distribution. Because the powers of four methods are '1' in the case of the number of sample size larger than the number of sample size given in Table 15. Unlike the other power comparisons, the powers of four methods are too high. One reason for this is that

the distribution in the alternative hypothesis is chosen from a non-symmetrical distribution rather than an elliptical (symmetrical) distribution.

The P method has the lower power than the others and the J method has a higher power in each sample size than the other methods such as the simulation results for the above distributions. However, it is more accurate to compare the powers between the FF and MVKS, considering that the P and J methods do not perform consistently and correctly in the Type I error comparisons. Hence, the MVKS has higher power than the FF method. Also, as the number of sample size increases, the powers of two methods increase to '1'.

The power of the P, FF, J, and MVKS methods are calculated using the bivariate Chi-Square distribution with $\upsilon = 1$ degrees of freedom. The power values performed at $n = 10, 11, 12, 13, 14, 15$ sample sizes are given in Table 16.

Table 16. Power comparisons for bivariate Chi-Square distribution

	Sample Size					
Method	**10**	**11**	**12**	**13**	**14**	**15**
P	0.5793	0.601	0.7198	0.7862	0.8477	0.8663
FF	0.9524	0.9667	0.987	0.9933	0.9978	0.9982
J	1	1	1	1	1	1
MVKS	1	1	1	1	1	1

The power of test is calculated in small sample sizes for the bivariate Chi-Square distribution. Since the powers are '1' in all four methods in large samples, the small sample sizes are examined to reveal the awareness of this distribution. The power of test values are too high percentages like power comparisons for the bivariate Exponential distribution. One reason for this is that the distribution in the alternative hypothesis is chosen from a non-symmetrical distribution rather than an elliptical (symmetrical) distribution.

The P method has the lower power in all sample sizes than the other methods and the power of the J method are calculated as '1' in all sample sizes. However, it is more accurate to compare the powers between the FF and MVKS, considering that the P and J methods do not perform consistently and correctly in the Type I error comparisons. Then, the MVKS has higher power than the FF method. Furthermore, the power of the MVKS is '1' in all sample sizes.

5. Real-life examples

Two different real-life data are chosen to demonstrate the performance of the proposed method. These are Fisher's Iris [28] and Royston's Hematology [40] data sets.

Firstly, the normality test is performed for the Fisher's Iris data set. Fisher's Iris data consists of three plant species. These are 'Setosa', 'Versicolor', and 'Virginica' species and four different attributes of each species (sepal length, sepal width, petal length, petal width) are available. There are 50 observation data for three plant species. The marginal normality test results of each attributes of species and the multivariate normality test of each species are obtained as in Table 17. Since the mean vector and the covariance matrix are not given in the example, they are estimated from the data. Therefore, the Lilliefors test table is used instead of the KS test table. The critical table value of 0.95% in Table 4 is 0.1257. '*' denotes significance at 5% level.

The 'Setosa', 'Versicolor', and 'Virginica' species have multivariate normal distribution as a result of Table 17. Similar results are obtained by [24, 29]. Looney [24] used the methods developed by [41, 43, 44]. These methods reveal that the Fisher Iris data set species have a multivariate normal distribution. The distribution of the Petal Length, Petal Width variables of the Setosa species, the Petal Width variable of the Versicolor species, and the Sepal Width variable of the Virginica species indicate a violation of the

univariate normality assumption. In addition, Looney [24] performed a normality test on the variables of the Setosa species. He showed that only Petal Width variable has not univariate normal distribution. In another study, Sürücü [45] showed that Petal Length, Petal Width variables of Setosa species have not the univariate normal distribution.

Table 17. The proposed multivariate normality test statistics for Fishers Iris dataset

Species	MVKS	Attributes	MVKS
Setosa	0.0779	Sepal Length	0.1149
		Sepal Width	0.1047
		Petal Length	0.1534*
		Petal Width	0.3488
Versicolor	0.0637	Sepal Length	0.0962
		Sepal Width	0.1207
		Petal Length	0.1171
		Petal Width	0.1477*
Virginica	0.0708	Sepal Length	0.1150
		Sepal Width	0.1279*
		Petal Length	0.1136
		Petal Width	0.1208

The other real-life data is the Royston Hematology data set which has six variables [40]. These are hemoglobin concentration, packed cell volume, white blood cell count, lymphocyte count, neutrophil count, and serum lead concentration variables. There are 103 observations in this data set. Since the mean vector and the covariance matrix are not given in the example, they are estimated from the data. Therefore, the Lilliefors test table is used instead of the KS test table. The multivariate Kolmogorov-Smirnov test statistic for this dataset is calculated as 0.1870. The critical table value can be obtained from Table 4 as approximately 0.0901 or the approximate value can also be calculated from the critical test table in [23]. Since the calculated test statistic is greater than the critical table value, it can be said that this data set does not provide the multivariate normality assumption at 5% significance level. The same result is shown by [40, 41, 43, 44] used in [24].

6. Conclusion

The proposed method is easily applicable to the multi-dimensional data and the multivariate KS test statistic can be used with the univariate KS test table. There is no need more than one table for each dimension. These results show that the proposed method is more useful and advantageous than the other methods. Furthermore, the understanding the complex algorithms of other methods and the calculation difficulties of the test statistic in case of more than two variables reveals the advantages of the proposed method in this study.

In the continuation of the simulation study, some methods in the literature for the multivariate KS tests have been compared with the proposed method in terms of power and Type I error. The P and J methods do not perform consistently and correctly in the Type I error comparisons. Hence, it is more accurate to compare the powers between the FF and MVKS methods. The MVKS has either better or same power values according to the FF method in both symmetric and non-symmetric distributions power comparisons. It also has high power especially in non-symmetrical distributions.

The effectiveness of the proposed method is investigated in known and easily accessible two data sets. Then, the multivariate normality test results of two data sets in this study are the same as literature.

Consequently, the simulation study shows that the proposed multivariate Kolmogorov-Smirnov test is a simple and consistent method. It can be easily applied to all data sizes. Also, it is useful for practitioners regarding the applicability.

References

[1] K. Alam, R. Abernathy and C.L. Williams, *Multivariate goodness-of-fit tests based on statistically equivalent blocks*, Comm. Statist. Theory Methods **22** (6), 1515-1533, 1993.

[2] K. Alam and C.L. Williams, *A multivariate goodness-of-fit test for stochastically ordered distributions*, Biom. J. **37** (8), 945-956, 1995.

[3] F.R. Bach and M.I. Jordan, *Kernel independent component analysis*, J. Mach. Learn. Res. **3** (Jul), 1-48, 2002.

[4] L. Baringhaus and N. Henze, *A consistent test for multivariate normality based on the empirical characteristic function*, Metrika **35** (1), 339-348, 1988.

[5] H. Bozdogan and D.E. Ramirez, *Testing for model fit: Assessing the Box–Cox transformations of multivariate data to nearnormality*, Comput. Stat. Q. **3**, 127-150, 1986.

[6] A. Cabaña and E.M. Cabaña, *Transformed empirical processes and modified Kolmogorov-Smirnov tests for multivariate distributions*, Ann. Statist. **25** (6), 2388-2409, 1997.

[7] R. Chicheportiche and J.P. Bouchaud, *Goodness-of-fit tests with dependent observations*, J. Stat. Mech. Theory Exp. **9**, P09003, 2011.

[8] S.N. Chiu and K.I. Liu, *Generalized Cramér-von Mises goodness-of-fit tests for multivariate distributions*, Comput. Statist. Data Anal. **53** (11), 3817-3834, 2009.

[9] W.G. Cochran, *The χ^2 test of goodness of fit*, Ann. Math. Statist. **23** (3), 315-345, 1952.

[10] H. Cramér, *On the composition of elementary errors: First paper: Mathematical deductions*, Scand. Actuar. J. **1928** (1), 13-74, 1928.

[11] S. Csorgo, *Testing for normality in arbitrary dimension*, Ann. Statist. **14** (2), 708-723, 1986.

[12] S. Facchinetti, *A procedure to find exact critical values of Kolmogorov-Smirnov test*, Statistica Applicata – Italian Journal of Applied Statistics **21** (34), 337–359, 2009.

[13] Y. Fan, *Goodness-of-fit tests for a multivariate distribution by the empirical characteristic function*, J. Multivariate Anal. **62** (1), 36-63, 1997.

[14] G. Fasano and A. Franceschini, *A multidimensional version of the Kolmogorov-Smirnov test*, Mon. Notices Royal Astron. Soc. **225** (1), 155-170, 1987.

[15] S. Ghosh and F.H. Ruymgaart, *Applications of empirical characteristic functions in some multivariate problems*, Canad. J. Statist. **20** (4), 429-440, 1992.

[16] A. Hyvärinen, J. Karhunen and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2004.

[17] R.A. Johnson and D.W. Wichern, *Applied Multivariate Statistical Analysis*, 5th ed., Prentice Hall, New Jersey, 2001.

[18] A. Justel, D. Peña and R. Zamar, *A multivariate Kolmogorov-Smirnov test of goodness of fit*, Statist. Probab. Lett. **35** (3), 251-259, 1997.

[19] J.F. Kenneys, *Mathematics of Statistics*, D. Van Nostrand Company Inc, 2013.

[20] A. Kolmogorov, *Sulla determinazione empirica di una lgge di distribuzione (in Italian)*, G. Inst. Ital. Attuari **4** (1933), 83-91, 1933.

[21] J.A. Koziol, *Assessing multivariate normality: A compendium*, Comm. Statist. Theory Methods **15** (9), 2763-2783, 1986.

[22] N.H. Kuiper, *Tests concerning random points on a circle*, Nederl. Akad. Wetensch. Proc. Ser. A **63** (1), 38-47, 1960.

[23] H.W. Lilliefors, *On the Kolmogorov-Smirnov test for normality with mean and variance unknown*, J. Amer. Statist. Assoc. **62** (318), 399-402, 1967.
[24] S.W. Looney, *How to use tests for univariate normality to assess multivariate normality*, Amer. Statist. **49** (1), 64-70, 1995.
[25] R. Lopes, I. Reid and P. Hobson, *The two-dimensional Kolmogorov-Smirnov test*, Proceedings of XI International Workshop on Advanced Computing and Analysis Techniques in Physics Research, Amsterdam, The Netherlands, 2007.
[26] J.F. Malkovich and A.A. Afifi, *On tests for multivariate normality*, J. Amer. Statist. Assoc. **68** (341), 176-179, 1973.
[27] K.V. Mardia, *Measures of multivariate skewness and kurtosis with applications*, Biometrika **57** (3), 519-530, 1970.
[28] K.V Mardia and J.T. Kent, *Multivariate Analysis*, Academic Press, London, 1979.
[29] K.V. Mardia and J.T. Kent, *Rao score tests for goodness of fit and independence*, Biometrika **78** (2), 355-363, 1991.
[30] J.R. Massey and J. Frank, *The Kolmogorov-Smirnov test for goodness of fit*, J. Amer. Statist. Assoc. **46** (253), 68-78, 1951.
[31] M.P. McAssey, *An empirical goodness-of-fit test for multivariate distributions*, J. Appl. Stat. **40** (5), 1120-1131, 2013.
[32] M.S. Mecibah, T.E. Boukelia, R. Tahtah and K. Gairaa, *Introducing the best model for estimation the monthly mean daily global solar radiation on a horizontal surface (Case study: Algeria)*, Renew. Sust. Energ. Rev. **36**, 194-202, 2014.
[33] D.S. Moore and J.B. Stubblebine, *Chi-square tests for multivariate normality with application to common stock prices*, Comm. Statist. Theory Methods **10** (8), 713-738, 1981.
[34] G.S. Mudholkar, M. McDermott and D.K. Srivastava, *A test of p-variate normality*, Biometrika **79** (4), 850-854, 1992.
[35] A. Novobilski and F. Kamangar, *Absolute Percent Error Based Fitness Functions for Evolving Forecast Models*, Proceedings of FLAIRS Conference, FL, USA, 591-595, 2001.
[36] J.A. Peacock, *Two-dimensional goodness-of-fit testing in astronomy*, Mon. Notices Royal Astron. Soc. **202** (3), 615-627, 1983.
[37] N.M. Razali and Y.B. Wah, *Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests*, Journal of Statistical Modeling and Analytics **2** (1), 21-33, 2011.
[38] J.L. Romeu and A. Ozturk, *A comparative study of goodness-of-fit tests for multivariate normality*, J. Multivariate Anal. **46** (2), 309-334, 1993.
[39] M. Rosenblatt, *Remarks on a multivariate transformation*, Ann. Math. Statist. **23** (3), 470-472, 1952.
[40] J.P. Royston, *Some techniques for assessing multivarate normality based on the Shapiro-Wilk W*, J. R. Stat. Soc. Ser. C. Appl. Stat. **32** (2), 121-133, 1983.
[41] N.J. Small, *Marginal skewness and kurtosis in testing multivariate normality*, J. R. Stat. Soc. Ser. C. Appl. Stat. **29**, 85-87, 1980.
[42] N.V. Smirnov, *On the estimation of the discrepancy between empirical curves of distribution for two independent samples*, Bull. Math. Univ. Moscou **2** (2), 3-14, 1939.
[43] M.S. Srivastava, *A measure of skewness and kurtosis and a graphical method for assessing multivariate normality*, Statist. Probab. Lett. **2** (5), 263-267, 1984.
[44] M.S. Srivastava and T.K. Hui, *On assessing multivariate normality based on Shapiro-Wilk W statistic*, Statist. Probab. Lett. **5** (1), 15-18, 1987.
[45] B. Sürücü, *Goodness-of-fit tests for multivariate distributions*, Comm. Statist. Theory Methods **35** (7), 1319-1331, 2006.
[46] G.J. Székely and M.L. Rizzo, *A new test for multivariate normality*, J. Multivariate Anal. **93** (1), 58-80, 2005.

[47] R. Von Mises, *Wahrscheinlichkeitsrechnung und ihre anwendung in der statistik und theorestischen physik*, Leipzig, Deutsche, 1931.
[48] C. Zhang, Y. Xiang and X. Shen, *Some multivariate goodness-of-fit tests based on data depth*, J. Appl. Stat. **39** (2), 385-397, 2012.
[49] M. Zhou and Y. Shao, *A powerful test for multivariate normality*, J. Appl. Stat. **41** (2), 351-363, 2014.