Communications for Statistical Applications and Methods
2023, Vol. 30, No. 4, 423–435

https://doi.org/10.29220/CSAM.2023.30.4.423
Print ISSN 2287-7843 / Online ISSN 2383-4757

A modified test for multivariate normality using second-power skewness and kurtosis

Namhyun Kim[1,a]

[a]Department of Science, Hongik University, Korea

Abstract

The Jarque and Bera (1980) statistic is one of the well known statistics to test univariate normality. It is based on the sample skewness and kurtosis which are the sample standardized third and fourth moments. Desgagné and de Micheaux (2018) proposed an alternative form of the Jarque-Bera statistic based on the sample second power skewness and kurtosis. In this paper, we generalize the statistic to a multivariate version by considering some data driven directions. They are directions given by the normalized standardized scaled residuals. The statistic is a modified multivariate version of Kim (2021), where the statistic is generalized using an empirical standardization of the scaled residuals of data. A simulation study reveals that the proposed statistic shows better power when the dimension of data is big.

Keywords: goodness-of-fit test, Jarque-Bera test, second power kurtosis, second power skewness, multivariate normality, power comparison

1. Introduction

Testing normality has been studied extensively for many statisticians, since normality is one of the most commonly made assumptions in the use of main statistical procedures. Consequently, there are numerous test procedures in the literature to assess the assumption. For a survey of normality tests, see D'Agostino and Stephens (1986) and Thode (2002).

Compared to the univariate normaity test, relatively less work has been done for multivariate normality (MVN) test, and it is a topic of ongoing interest. For a general review of several MVN tests, some references are Henze and Zirkler (1990), Henze (2002), Ebner and Henze (2020), and Srivastava and Mudhoklar (2003). Henze (2002) concentrated on affine invariant and consistent procedures, and Ebner and Henze (2020) emphasized on several classes of the weighted L^2-statistics. Comparative power study for MVN is done by Horswell and Looney (1992), Romeu and Ozturk (1993), Mecklin and Mundfrom (2005), Farrell *et al.* (2007), and Hanusz *et al.* (2018).

In univariate normality tests, one of the well known and widely used statistics is the Jarque and Bera (1980) test, which is based on the sample Pearson's skewness and kurtosis. The test is popular especially in econometrics. It is also known as the D'Agositno-Pearson or the Bowman-Shenton test in statistics (D'Agostino and Pearson (1973, 1974); Bowman and Shenton, 1975). A multivariate version of the test is proposed by Doornik and Hansen (2008) and Kim (2015, 2016).

While the Pearson's sample skewness and kurtosis are the sample standardized third and fourth moments, Desgagné and de Micheaux (2018) proposed the sample second power skewness and kurtosis. The combination of their measures could be an alternative to the Jarque-Bera statistic. Kim (2021)

[1] Corresponding author: Department of Science, Hongik University, 94 Wausan-Ro, Mapo-Gu, Seoul 04066, Korea. E-mail: nhkim@hongik.ac.kr

Published 31 July 2023 / journal homepage: http://csam.or.kr

generalized the Desgagné and de Micheaux's statistic to a multivariate version using orthogonalization or an empirical standardization of the data.

It is well known that a p-dimensional random vector $\boldsymbol{X}$ is multivariate normal if and only if $\boldsymbol{c}^T\boldsymbol{X}$ is univariate normal for all constant vectors $\boldsymbol{c}$. Therefore there exists a vector $\boldsymbol{c}$ such that $\boldsymbol{c}^T\boldsymbol{X}$ does not follow a normal if $\boldsymbol{X}$ is not normal. Malkovich and Afifi (1973) generalized some univariate normality test statistics such as skewness, kurtosis and Shapiro and Wilk (1965) statistic to a multivariate version using the idea. Fattorini (1986) improved the Malkovich and Afifi's multivariate Shapiro-Wilk statistic by considering some data-driven directions. Zhou and Shao (2014) suggested to pick a few data driven directions $\boldsymbol{c}$ to detect non-normality of data. In this paper, we apply their idea to the statistic of Desgagné and de Micheaux (2018) and generalize it to a MVN test statistic.

Section 2 describes the main idea and test statistics. Section 3 gives a data example and a comparison of power performances of the test statistics through a simulation study. Section 4 ends with concluding remarks.

2. Test statistics

2.1. Univariate test statistics

Consider independent observations $X_1, \ldots, X_n$ from X. We want to test

$$H_0 : X \text{ has a normal distribution } N\left(\mu, \sigma^2\right) \text{ for some } \mu \text{ and } \sigma^2.$$

The well known Jarque-Bera statistic for univariate normality is based on the Pearson's sample skewness $\sqrt{b_1}$ and kurtosis b_2. It is

$$\mathrm{JB} = \frac{n}{6}\left(\sqrt{b_1}\right)^2 + \frac{n}{24}\left(b_2 - 3\right)^2, \tag{2.1}$$

where

$$\sqrt{b_1} = \frac{1}{n}\sum_{i=1}^{n} Z_i^3 = \frac{1}{n}\sum_{i=1}^{n} |Z_i|^3 \operatorname{sign}(Z_i),$$

$$b_2 = \frac{1}{n}\sum_{i=1}^{n} Z_i^4,$$

with $Z_i = (X_i - \bar{X})/S$, $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$, $S^2 = n^{-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Desgagné and de Micheaux (2018) defined the sample second power skewness B_2 and the sample second power kurtosis K_2,

$$B_2 = \frac{1}{n}\sum_{i=1}^{n} |Z_i|^2 \operatorname{sign}(Z_i),$$

$$K_2 = \frac{1}{n}\sum_{i=1}^{n} Z_i^2 \log|Z_i|,$$

and they proposed a Jarque-Bera type normality test statistic based on B_2 and K_2,

$$\frac{nB_2^2}{3-8/\pi}+\frac{n\left(K_2-(2-\log 2-\gamma)/2\right)^2}{(3\pi^2-28)/8}. \tag{2.2}$$

The statistic comes from the limit distribution of (B_2, K_2),

$$\sqrt{n}\begin{pmatrix} -2B_2 \\ -2^{-1}\left[K_2-(2-\log 2-\gamma)/2\right] \end{pmatrix} \xrightarrow{d} N_2\left(\mathbf{0}, \begin{pmatrix} 4(3-8/\pi) & 0 \\ 0 & \left(3\pi^2-28\right)/32 \end{pmatrix}\right)$$

using the central limit theorem with $\gamma = -\Gamma'(1)/\Gamma(1) = 0.577215665\ldots$, $\Gamma(x)$ is the gamma function.

Their statistic in (2.2) follows the chi-squared distribution with 2 degrees of freedom χ_2^2 asymptotically. To make the distribution follow χ_2^2 even for small sample sizes, they proposed a modified version of the statistic in (2.2),

$$\mathrm{DX}=\frac{nB_2^2}{(3-8/\pi)(1-1.9/n)}+\frac{n\left[\left(K_2-B_2^2\right)^{\frac{1}{3}}-\left((2-\log 2-\gamma)/2\right)^{\frac{1}{3}}(1-1.026/n)\right]^2}{72^{-1}\left((2-\log 2-\gamma)/2\right)^{-\frac{4}{3}}(3\pi^2-28)(1-2.25/n^{0.8})}. \tag{2.3}$$

2.2. Multivariate test statistics

Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be independent observations from a p-variate random vector $\boldsymbol{X}$, and let $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ be a p-variate multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. We want to test the null hypothesis

$$H_0 : \boldsymbol{X} \text{ has } N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \text{ for some } \boldsymbol{\mu} \text{ and } \boldsymbol{\Sigma}.$$

Let

$$\bar{\boldsymbol{X}}=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i, \quad \boldsymbol{S}=\frac{1}{n}\sum_{i=1}^{n}\left(\boldsymbol{X}_i-\bar{\boldsymbol{X}}\right)\left(\boldsymbol{X}_i-\bar{\boldsymbol{X}}\right)^T \tag{2.4}$$

be a sample mean vector and a sample covariance matrix with T a transpose. If $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ are a sample from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ with

$$\boldsymbol{Z}_i=\boldsymbol{S}^{-\frac{1}{2}}\left(\boldsymbol{X}_i-\bar{\boldsymbol{X}}\right), \quad i=1,\ldots,n, \tag{2.5}$$

follow $N_p(\mathbf{0}, \boldsymbol{I})$ asymptotically, where $\boldsymbol{S}^{-1/2}$ is a symmetric positive definite square root of $\boldsymbol{S}^{-1}$, the inverse matrix of $\boldsymbol{S}$ in (2.4). The $\boldsymbol{Z}_i$'s are called the scaled residuals.

It is well known that $\boldsymbol{Z}$ is $N_p(\mathbf{0}, \boldsymbol{I})$ if and only if $\boldsymbol{c}^T\boldsymbol{Z}$ is a univariate normal for all

$$\boldsymbol{c} \in \boldsymbol{U}^p = \{\boldsymbol{c} \in \boldsymbol{R}^p : \|\boldsymbol{c}\| = 1\}.$$

Therefore we can find a vector $\boldsymbol{c}$ such that $\boldsymbol{c}^T\boldsymbol{Z}$ does not follow univariate normal if $\boldsymbol{Z}$ is not normal. Malkovich and Afifi (1973) generalized some univariate normality test statistics such as skewness, kurtosis and Shapiro-Wilk statistic to multivariate cases by use of Roy (1953)'s union-intersection principle. It is to investigate all the possible linear combinations that reduce to a standard normal distribution under the null hypothesis and find the direction $\boldsymbol{c}$ that gives the least normal projection. If the least normal projection is not significantly different from normal, then we cannot reject the joint normality of the variables. Kim and Bickel (2003) showed the null distribution of the generalized Shapiro-Wilk type statistic based on the idea. Kim (2006) generalized the Cramér-von Mises statistic to test multivariate normality using the principle, and showed the limit distribution of the proposed statistic.

One of good points of those statistics is that they are affine invariant. In other words they are invariant with respect to nonsingular matrix multiplication and vector addition. Hence the distributions of the statistics under the null hypothesis do not depend on the unknown parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. However a main drawback of those statistics is that they are very hard to compute especially when the dimension of data is large.

One possible way to overcome the drawback of the statistic is to consider some particular directions instead of all possible ones. The trivial direction is $\boldsymbol{e}_k = (0, \ldots, 0, 1, 0, \ldots, 0)^T$, the unit vector with its k^{th} component 1 and all the others 0. Then

$$\boldsymbol{e}_k^T \boldsymbol{Z}_i = (0, \ldots, 0, 1, 0, \ldots, 0) \cdot \left(Z_{1i}, Z_{2i}, \ldots, Z_{pi}\right)^T = Z_{ki}, \quad i = 1, \ldots, n,$$

is the k^{th} component of the scaled residuals $\boldsymbol{Z}_i$ in (2.5). Z_{ki}'s, $i = 1, \ldots, n$, follow approximately independent univariate standard normal $N(0, 1)$ under the null hypothesis. Srivastava and Hui (1987) suggested principal component approach for the directions. As a matter of fact, almost all univariate test statistics could be generalized to multivariate ones using orthogonalization. Several multivariate statistics are generated by considering the corresponding univariate statistic for each coordinate of the scaled residuals. Villasenor-Alva and González-Estrada (2009) generalized the Shapiro-Wilk's test to a multivariate version using the above idea. Kim (2015) used the idea to generalize the univariate skewness and kurtosis. Kim (2016) generalized the Jarque-Bera statistic in (2.1) to the multivariate statistic JB_M using the idea. Kim (2021) proposed the multivariate version of DX in (2.3), DX_M by the same way.

For a univariate statistic G, $G(\boldsymbol{c}) = G(\boldsymbol{c}^T \boldsymbol{Z}_1, \ldots, \boldsymbol{c}^T \boldsymbol{Z}_n)$ is a statistic based on projection of Z_i's in the direction $\boldsymbol{c}$. A way of picking directions $\boldsymbol{c}$ randomly among all possible directions is data driven directions $\{\boldsymbol{Z}_l/||\boldsymbol{Z}_l||, l = 1, \ldots, n\}$. In other words, we consider the normalized standardized scaled residuals $\{\boldsymbol{Z}_l/||\boldsymbol{Z}_l||\}$ whose empirical distribution function follows asymptotically uniform distribution in the unit p-dimensional sphere $\boldsymbol{U}^p$, instead of all vectors $\boldsymbol{c} \in \boldsymbol{U}^p$. The idea is used in several papers. Fattorini (1986) proposed a modified statistic to the Malkovich and Afifi's multivariate Shapiro-Wilk statistic based on the idea. Kim (2006) used the idea to compute the multivariate Cramér-von Mises statistic. The idea is closely related to the number-theoretic method (NTM) studied by Fang and Wang (1993). The main purpose of NTM is to find a set of points that is uniformly scatted over a p-dimensional unit cube.

Kim (2004) approximated Malkovich and Afifi's multivariate skewness and kurtosis. The multivariate statistics are

$$b_{1,M}^{*} = \max_{1 \le l \le n} \left[\frac{1}{n} \sum_{i=1}^{n} \left(\frac{\boldsymbol{Z}_l'}{||\boldsymbol{Z}_l||} \boldsymbol{Z}_i \right)^3 \right]^2,$$

and

$$b_{2,M}^{2*} = \max_{1 \le l \le n} \left[\frac{1}{n} \sum_{i=1}^{n} \left(\frac{\boldsymbol{Z}_l'}{||\boldsymbol{Z}_l||} \boldsymbol{Z}_i \right)^4 - 3 \right]^2.$$

Table 1: Statistics and p-values for the Iris setosa data

Test	JB_M	DX_M	JB_{2p}	DX_{2p}	$JB_{\max}$	$DX_{\max}$	HZ	MN
Statistic	12.2712	13.2691	7.9933	7.2518	17.6064	12.6534	0.9488	30.4720
p-value	0.1395	0.1029	0.1529	0.0481	0.219	0.102	0.04995	0.0829

In that paper, it is explained that $\boldsymbol{c}_l = \boldsymbol{Z}_l/\|\boldsymbol{Z}_l\|$, $l = 1, \ldots, n$, approximately maximize $[(1/n)\sum_i(\boldsymbol{c}'\boldsymbol{Z}_i)^3]^2$. Upon the idea, we can consider the multivariate Jarque-Bera statistics such that

$$JB_{\max} = \max_{1\le l\le n} JB\left(\frac{\boldsymbol{Z}_l}{\|\boldsymbol{Z}_l\|}\right). \tag{2.6}$$

Likewise, the statistic $DX_{\max}$ can be defined

$$DX_{\max} = \max_{1\le l\le n} DX\left(\frac{\boldsymbol{Z}_l}{\|\boldsymbol{Z}_l\|}\right). \tag{2.7}$$

Meanwhile, Zhou and Shao (2014) considered the p least normal-like directions Θ_{p1} which are corresponding to p extreme G values evaluated at $\{\boldsymbol{Z}_l/\|\boldsymbol{Z}_l\|\}$ and the unit vector $\Theta_{p2} = \{\boldsymbol{e}_k, k = 1, \ldots, p\}$. Finally the multivariate test statistic becomes

$$\frac{1}{2p}\sum_{\boldsymbol{c}\in\Theta_{p1}\cup\Theta_{p2}} G(\boldsymbol{c}).$$

They constructed a kind of multivariate Shapiro-Wilk statistic combined with kurtosis by applying the idea. On the other hand, Fattorini (1986) statistic is considering all the directions $\{\boldsymbol{Z}_l/\|\boldsymbol{Z}_l\|, l = 1, \ldots, n\}$ and detecting the extreme non-normal-like direction.

As it is mentioned, Kim (2021) generalized DX statistic in (2.3) to a multivariate version using orthogonalization or an empirical standardization of data. That is just considering the direction Θ_{p2}. In this paper, we apply Zhou and Shao (2014)'s idea to generalize DX. Then the test statistic becomes

$$DX_{2p} = \frac{1}{2p}\sum_{\boldsymbol{c}\in\Theta_{p1}\cup\Theta_{p2}} DX(\boldsymbol{c}). \tag{2.8}$$

Here, $DX(\boldsymbol{c}) = DX(\boldsymbol{c}^T\boldsymbol{Z}_1, \ldots, \boldsymbol{c}^T\boldsymbol{Z}_n)$ is the statistic of DX evaluated with $(\boldsymbol{c}^T\boldsymbol{Z}_1, \ldots, \boldsymbol{c}^T\boldsymbol{Z}_n)$.

The statistic DX_{2p} in (2.8) is computed as follows.

1. Compute the standardized residuals $\boldsymbol{Z}_i = \boldsymbol{S}^{-1/2}(\boldsymbol{X}_i - \bar{\boldsymbol{X}}), i = 1, \ldots, n.$
2. Compute the DX statistic in (2.3) for each component of $(Z_{k1}, \ldots, Z_{kn}), k = 1, \ldots, p$, i.e. compute the statistic for Θ_{p2} directions.
3. Compute

$$DX(\boldsymbol{c}) = DX\left(\frac{\boldsymbol{Z}_l}{\|\boldsymbol{Z}_l\|}\right) = DX\left(\frac{\boldsymbol{Z}_l^T}{\|\boldsymbol{Z}_l\|}\boldsymbol{Z}_1, \ldots, \frac{\boldsymbol{Z}_l^T}{\|\boldsymbol{Z}_l\|}\boldsymbol{Z}_n\right)$$

   for $l = 1, \ldots, n$ and choose the biggest p values among n.
4. Average out the $2p$ values in the above procedures.

Table 2: Power comparison of the statistics ($\alpha = 0.05$, $p = 2$, $n = 20$)

Alternative	DX_M	JB_M	DX_{2p}	JB_{2p}	DX_{max}	JB_{max}	HZ	MN
$N(0,1)^2$	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.05
Cauchy$(0,1)^2$	0.97	0.94	0.98	0.96	0.97	0.95	0.97	0.97
Logistic$(0,1)^2$	0.14	0.16	0.15	0.17	0.14	0.16	0.10	0.17
$(t_2)^2$	0.71	0.72	0.74	0.71	0.70	0.69	0.64	0.74
$(t_5)^2$	0.25	0.25	0.27	0.27	0.25	0.28	0.17	0.30
Beta$(1,1)^2$	0.10	0.00	0.12	0.00	0.14	0.00	**0.18**	0.00
Beta$(2,2)^2$	0.04	0.01	0.05	0.01	0.04	0.01	0.06	0.00
Beta$(1,2)^2$	0.15	0.05	0.15	0.05	0.15	0.04	**0.28**	0.04
exp$(1)^2$	0.73	0.66	0.79	0.66	0.72	0.62	**0.86**	0.69
Lognormal$(0,0.5)^2$	0.51	0.48	0.51	0.51	0.46	0.46	**0.58**	0.50
Gamma$(0.5,1)^2$	0.92	0.87	0.97	0.87	0.95	0.83	0.99	0.90
Gamma$(5,1)^2$	0.21	0.20	0.21	0.22	0.19	0.21	0.24	0.21
$(\chi^2_5)^2$	0.38	0.34	0.40	0.36	0.35	0.35	**0.45**	0.37
$(\chi^2_{15})^2$	0.14	0.18	0.16	0.18	0.15	0.16	0.16	0.16
$N(0,1) * t_5$	0.16	0.17	0.17	0.17	0.15	0.17	0.12	0.17
$N(0,1) *$ Beta$(1,1)$	0.11	0.02	0.11	0.02	0.10	0.02	0.10	0.01
$N(0,1) *$ exp(1)	0.37	0.36	0.50	0.40	0.49	0.39	0.51	0.40
$N(0,1) * \chi^2_5$	0.27	0.24	0.26	0.23	0.20	0.20	0.22	0.20
$NMIX_2(0.5,4,0,0)$	**0.89**	0.04	0.86	0.02	0.81	0.02	0.52	0.01
$NMIX_2(0.5,0,0,0.9)$	0.19	0.17	0.20	0.19	0.18	0.18	0.14	0.18
$NMIX_2(0.5,4,0,0.9)$	0.82	0.22	0.80	0.23	0.69	0.25	0.82	0.23
$NMIX_2(0.9,4,0,0)$	**0.73**	0.66	**0.74**	0.58	0.67	0.53	0.64	0.55
$NMIX_2(0.9,0,0,0.9)$	0.06	0.07	0.07	0.07	0.05	0.06	0.05	0.06
$NMIX_2(0.9,4,0,0.9)$	0.71	0.65	**0.74**	0.59	0.67	0.53	0.62	0.54
Avaerage power	0.42	0.32	0.43	0.33	0.40	0.31	0.41	0.32

The univariate statistic JB in (2.1) can be generalized to

$$JB_{2p} = \frac{1}{2p} \sum_{\boldsymbol{c} \in \Theta_{p1} \cup \Theta_{p2}} JB(\boldsymbol{c}) \tag{2.9}$$

by the exactly same way.

The multivariate statistics JB_M in Kim (2016) and DX_M in Kim (2021) can be written as follows.

$$JB_M = \sum_{\boldsymbol{c} \in \Theta_{p2}} JB(\boldsymbol{c}) = \sum_{k=1}^{p} JB(k). \tag{2.10}$$

$$DX_M = \sum_{\boldsymbol{c} \in \Theta_{p2}} DX(\boldsymbol{c}) = \sum_{k=1}^{p} DX(k), \tag{2.11}$$

where $JB(k)$, $DX(k)$ are the JB, DX statistics evaluated for each coordinate $(Z_{k1}, Z_{k2}, \ldots, Z_{kn})$, $k = 1, \ldots, p$, respectively.

The Henze and Zirkler (1990) statistic $T_{n,\beta}$ is frequently recommended as a formal test statistic for multivariate normality. It is as follows.

$$T_{n,\beta}(\mathbf{X}_1, \ldots, \mathbf{X}_n) = n\Big(4I(\mathbf{S} \text{ is singular}) + D_{n,\beta} I(\mathbf{S} \text{ is nonsingular})\Big), \tag{2.12}$$

Table 3: Power comparison of the statistics ($\alpha = 0.05$, $p = 2$, $n = 50$)

Alternative	DX_M	JB_M	DX_{2p}	JB_{2p}	$DX_{\max}$	$JB_{\max}$	HZ	MN
$N(0,1)^2$	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.05
Cauchy$(0,1)^2$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Logistic$(0,1)^2$	0.25	0.28	0.28	0.31	0.28	0.31	0.15	0.32
$(t_2)^2$	0.97	0.94	0.98	0.97	0.97	0.96	0.95	0.97
$(t_5)^2$	0.50	0.51	0.54	0.54	0.50	0.55	0.32	0.56
Beta$(1,1)^2$	0.32	0.00	0.54	0.00	0.58	0.00	**0.68**	0.00
Beta$(2,2)^2$	0.10	0.00	0.11	0.00	0.12	0.00	**0.17**	0.00
Beta$(1,2)^2$	0.45	0.10	0.63	0.05	0.63	0.02	**0.80**	0.12
exp$(1)^2$	0.98	0.94	1.00	0.98	1.00	0.96	1.00	1.00
Lognormal$(0,0.5)^2$	0.90	0.85	0.96	0.88	0.94	0.86	0.94	0.95
Gamma$(0.5,1)^2$	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00
Gamma$(5,1)^2$	0.53	0.46	0.58	0.49	0.53	0.46	0.53	0.56
$(\chi^2_5)^2$	0.80	0.74	0.88	0.75	0.85	0.70	0.87	0.86
$(\chi^2_{15})^2$	0.39	0.33	0.41	0.35	0.36	0.33	0.36	0.41
$N(0,1) * t_5$	0.32	0.34	0.35	0.35	0.31	0.33	0.17	0.32
$N(0,1) *$ Beta$(1,1)$	0.43	0.03	0.44	0.03	0.38	0.02	0.32	0.01
$N(0,1) *$ exp(1)	0.80	0.73	**0.96**	0.82	**0.97**	0.82	0.92	0.88
$N(0,1) * \chi^2_5$	**0.72**	0.59	0.69	0.53	0.64	0.49	0.52	0.54
$NMIX_2(0.5,4,0,0)$	1.00	0.08	1.00	0.03	1.00	0.02	1.00	0.01
$NMIX_2(0.5,0,0,0.9)$	**0.48**	0.41	**0.49**	0.36	0.45	0.33	0.32	0.28
$NMIX_2(0.5,4,0,0.9)$	1.00	0.55	1.00	0.55	1.00	0.51	1.00	0.90
$NMIX_2(0.9,4,0,0)$	0.99	0.99	0.99	0.97	0.99	0.95	0.95	0.97
$NMIX_2(0.9,0,0,0.9)$	0.07	0.06	0.08	0.08	0.08	0.07	0.05	0.08
$NMIX_2(0.9,4,0,0.9)$	0.99	0.98	0.98	0.98	0.98	0.96	0.94	0.98
Avaerage power	0.65	0.52	0.69	0.52	0.68	0.51	0.65	0.55

where

$$D_{n,\beta} = \frac{1}{n^2}\sum_{i,j=1}^{n}\exp\left(-\frac{\beta^2}{2}\left\|\mathbf{Y}_i - \mathbf{Y}_j\right\|^2\right) - 2\left(1+\beta^2\right)^{-\frac{p}{2}}\frac{1}{n}\sum_{j=1}^{n}\exp\left(-\frac{\beta^2}{2\left(1+\beta^2\right)}\left\|\mathbf{Y}_j\right\|^2\right) + \left(1+2\beta^2\right)^{-\frac{p}{2}}$$

with $\|\mathbf{Y}_i - \mathbf{Y}_j\|^2 = (\mathbf{X}_i - \mathbf{X}_j)'\mathbf{S}^{-1}(\mathbf{X}_i - \mathbf{X}_j)$ and $\|\mathbf{Y}_j\|^2 = (\mathbf{X}_j - \bar{\mathbf{X}})'\mathbf{S}^{-1}(\mathbf{X}_j - \bar{\mathbf{X}})$. β is defined as

$$\beta = \beta_p\left(n\right) = \frac{1}{\sqrt{2}}\left(\frac{2p+1}{4}\right)^{\frac{1}{p+4}} n^{\frac{1}{p+4}}.$$

Mardia (1970, 1974) defined the well known multivariate measure of skewness $b_{1,p}$ and kurtosis $b_{2,p}$,

$$b_{1,p} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left\{\left(\boldsymbol{X}_i - \bar{\boldsymbol{X}}\right)'\boldsymbol{S}^{-1}\left(\boldsymbol{X}_j - \bar{\boldsymbol{X}}\right)\right\}^3,$$

and

$$b_{2,p} = \frac{1}{n}\sum_{j=1}^{n}\left\{\left(\boldsymbol{X}_j - \bar{\boldsymbol{X}}\right)'\boldsymbol{S}^{-1}\left(\boldsymbol{X}_j - \bar{\boldsymbol{X}}\right)\right\}^2.$$

Table 4: Power comparison of the statistics ($\alpha = 0.05$, $p = 5$, $n = 20$)

Alternative	DX_M	JB_M	DX_{2p}	JB_{2p}	$DX_{\max}$	$JB_{\max}$	HZ	MN
$N(0,1)^5$	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.05
Cauchy$(0,1)^5$	1.00	0.99	1.00	1.00	0.99	0.99	0.99	1.00
Logistic$(0,1)^5$	0.12	0.14	0.16	0.18	0.14	0.15	0.08	0.18
$(t_2)^5$	0.84	0.84	0.90	0.88	0.81	0.80	0.69	0.88
$(t_5)^5$	0.26	0.28	0.33	0.34	0.29	0.28	0.14	0.34
Beta$(1,1)^5$	0.05	0.02	0.01	0.01	0.01	0.01	**0.10**	0.00
Beta$(2,2)^5$	0.04	0.02	0.01	0.01	0.02	0.01	0.06	0.00
Beta$(1,2)^5$	0.08	0.05	0.04	0.03	0.04	0.03	**0.18**	0.03
exp$(1)^5$	0.66	0.61	0.79	0.74	0.55	0.58	0.81	0.79
Lognormal$(0,0.5)^5$	0.49	0.49	0.58	0.56	0.42	0.46	0.49	0.60
Gamma$(0.5,1)^5$	0.91	0.86	0.98	0.94	0.81	0.80	0.99	0.97
Gamma$(5,1)^5$	0.17	0.17	0.19	0.21	0.16	0.18	0.17	0.21
$(\chi^2_5)^5$	0.30	0.30	0.38	0.38	0.26	0.29	0.34	0.39
$(\chi^2_{15})^5$	0.12	0.14	0.14	0.15	0.12	0.14	0.12	0.15
$N(0,1)^4 * t_5$	0.10	0.12	0.11	0.11	0.10	0.10	0.06	0.10
$N(0,1)^4 *$ Beta$(1,1)$	0.05	0.03	0.04	0.03	0.04	0.04	0.06	0.03
$N(0,1)^4 *$ exp(1)	0.14	0.17	0.17	0.19	0.18	0.19	0.13	0.18
$N(0,1)^4 * \chi^2_5$	**0.19**	**0.18**	0.14	0.12	0.09	0.11	0.08	0.11
$NMIX_5(0.5,4,0,0)$	**1.00**	0.04	0.74	0.04	0.05	0.04	0.17	0.03
$NMIX_5(0.5,0,0,0.9)$	0.52	0.43	**0.69**	0.58	0.42	0.40	0.54	0.65
$NMIX_5(0.5,4,0,0.9)$	0.95	0.39	0.96	0.71	0.53	0.51	0.96	0.82
$NMIX_5(0.9,4,0,0)$	**0.86**	0.70	0.71	0.52	0.39	0.34	0.32	0.40
$NMIX_5(0.9,0,0,0.9)$	0.09	0.10	0.11	0.11	0.08	0.09	0.07	0.10
$NMIX_5(0.9,4,0,0.9)$	0.84	0.75	0.82	0.59	0.58	0.35	0.40	0.43
Avaerage power	0.43	0.34	0.43	0.37	0.31	0.30	0.35	0.37

The statistic

$$MN = \frac{nb_{1,p}}{6} + \left(\frac{b_{2,p} - ((n-1)/(n+1))\, p\,(p+2)}{\sqrt{\left[((n-3)(n-p-1)(n-p+1))/\left((n+1)^2(n+3)(n+5)\right)\right]\,(8p\,(p+2))}} \right)^2 \tag{2.13}$$

is considered in Kim (2020). The MN statistic is almost the same statistic proposed by Doornik and Hansen (2008).

In the next section, we compare the statistics DX_M in (2.11), JB_M in (2.10), DX_{2p} in (2.8), JB_{2p} in (2.9), $DX_{\max}$ in (2.7), $JB_{\max}$ in (2.6) through a simulation in terms of power. The $T_{n,\beta}$ in (2.12) and MN in (2.13) are also compared.

3. Example and simulation study

3.1. Example

We consider the well known Fisher's iris data introduced by Fisher (1936). The data set consists of 50 samples from each of three species of Iris setosa, Iris versicolor, and Iris virginica. Four features, sepal length, sepal width, petal length, and petal width were measured from each of the three species.

We examine the data on the variety Iris setosa. The statistics JB_M, DX_M, JB_{2p}, DX_{2p}, $JB_{\max}$, $DX_{\max}$, HZ and MN in (2.6) to (2.13) are computed in Table 1. The corresponding p-values are also

Table 5: Power comparison of the statistics ($\alpha = 0.05$, $p = 5$, $n = 50$)

Alternative	DX_M	JB_M	DX_{2p}	JB_{2p}	$DX_{\max}$	$JB_{\max}$	HZ	MN
$N(0,1)^5$	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04
Cauchy$(0,1)^5$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Logistic$(0,1)^5$	0.24	0.26	0.38	0.35	0.30	0.29	0.14	**0.41**
$(t_2)^5$	1.00	0.99	1.00	1.00	1.00	0.99	1.00	1.00
$(t_5)^5$	0.55	0.56	0.71	0.70	0.62	0.61	0.30	0.68
Beta$(1,1)^5$	0.07	0.01	0.01	0.00	0.01	0.00	**0.50**	0.00
Beta$(2,2)^5$	0.05	0.01	0.01	0.00	0.01	0.00	**0.14**	0.00
Beta$(1,2)^5$	0.19	0.03	0.09	0.01	0.03	0.02	**0.66**	0.05
exp$(1)^5$	0.98	0.95	1.00	0.99	0.98	0.93	1.00	1.00
Lognormal$(0,0.5)^5$	0.89	0.87	0.98	0.94	0.90	0.86	0.96	0.99
Gamma$(0.5,1)^5$	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00
Gamma$(5,1)^5$	0.44	0.43	0.58	0.49	0.40	0.40	0.47	**0.67**
$(\chi^2_5)^5$	0.75	0.69	0.89	0.78	0.70	0.64	0.88	0.95
$(\chi^2_{15})^5$	0.30	0.33	0.39	0.34	0.28	0.29	0.28	**0.49**
$N(0,1)^4 * t_5$	0.22	0.25	0.24	0.21	0.20	0.20	0.09	0.20
$N(0,1)^4 *$ Beta$(1,1)$	**0.15**	0.06	0.06	0.04	0.04	0.03	0.10	0.02
$N(0,1)^4 *$ exp(1)	0.35	0.35	**0.59**	0.49	0.56	0.43	0.39	0.55
$N(0,1)^4 * \chi^2_5$	**0.54**	0.42	0.39	0.27	0.24	0.21	0.16	0.25
$NMIX_5(0.5,4,0,0)$	1.00	0.11	1.00	0.03	0.19	0.03	0.70	0.03
$NMIX_5(0.5,0,0,0.9)$	0.96	0.84	0.98	0.85	0.85	0.64	0.98	0.90
$NMIX_5(0.5,4,0,0.9)$	1.00	0.86	1.00	0.92	0.98	0.75	1.00	1.00
$NMIX_5(0.9,4,0,0)$	1.00	0.98	1.00	0.84	0.93	0.56	0.92	0.88
$NMIX_5(0.9,0,0,0.9)$	0.16	0.15	**0.21**	0.18	0.15	0.13	0.11	0.17
$NMIX_5(0.9,4,0,0.9)$	0.99	1.00	0.99	0.96	1.00	0.74	0.91	0.99
Avaerage power	0.60	0.53	0.63	0.54	0.54	0.47	0.60	0.57

presented. We can use the R-package 'MVN' (see Korkmaz *et al*., 2014) or 'mvnormalTest' for the statistic HZ.

According to the result, the multivariate normality of the Iris setosa data is rejected at the significance level 0.05 when we use the statistics DX_{2p} or HZ. The p-values of these statistics are slightly less than 0.05. The other statistics do not reject it. The data set is also analyzed in Looney (1995) and Small (1980).

3.2. Power comparison

We compare the power of DX_M, JB_M, DX_{2p}, JB_{2p}, $DX_{\max}$, $JB_{\max}$, the Henze and Zirkler (HZ) statistic and the statistic MN in (2.13) through a simulation. The simulation is conducted at the significance level $\alpha = 0.05$ for dimensions $p = 2, 5, 10$. The sample sizes are $n = 20, 50$ for $p = 2, 5$ and $n = 50, 100$ for $p = 10$. The result in Kim (2021) is used for the power of DX_M, JB_M and HZ for $p = 2, 5$. The critical values from the simulation are used except DX_M. As for DX_M, the critical values from χ_{2p} are used. Regarding the HZ statistic, we used the simulated critical values given in Henze and Zirkler (1990). They also gave an approximation to the quantile of HZ for given β and the dimension p.

$N = 5,000$ samples are generated from various alternative distributions. The same alternatives in Kim (2021) are selected for comparison. They are the distributions with independent marginals and mixtures of normal distributions. $F_1 * F_2$ stands for the distribution with independent marginal distributions F_1 and F_2. F_1^p denotes the product of p independent copies of F_1. Here Cauchy$(0,1)$, Logistic$(0,1)$, and t-distribution t_k are symmetric long-tailed marginals. *Beta*$(1,1)$, *Beta*$(2,2)$ are symmetric short-tailed marginals. As for asymmetric marginals *Beta*$(1,2)$, *exp*(1), Lognormal$(0,0.5)$,

Table 6: Power comparison of the statistics ($\alpha = 0.05$, $p = 10$, $n = 50$)

Alternative	DX_M	JB_M	DX_{2p}	JB_{2p}	$DX_{\max}$	$JB_{\max}$	HZ	MN
$N(0,1)^{10}$	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Cauchy$(0,1)^{10}$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Logistic$(0,1)^{10}$	0.20	0.22	0.41	0.41	0.25	0.24	0.08	0.40
$(t_2)^{10}$	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00
$(t_5)^{10}$	0.54	0.56	**0.80**	**0.80**	0.62	0.61	0.22	0.77
Beta$(1,1)^{10}$	0.05	0.01	0.00	0.00	0.00	0.00	**0.28**	0.00
Beta$(2,2)^{10}$	0.03	0.02	0.00	0.00	0.01	0.00	**0.11**	0.00
Beta$(1,2)^{10}$	0.07	0.03	0.02	0.01	0.02	0.02	**0.36**	0.01
exp$(1)^{10}$	0.95	0.92	1.00	0.99	0.92	0.88	1.00	1.00
Lognormal$(0,0.5)^{10}$	0.87	0.84	0.98	0.97	0.82	0.81	0.86	1.00
Gamma$(0.5,1)^{10}$	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00
Gamma$(5,1)^{10}$	0.31	0.31	0.51	0.49	0.31	0.30	0.26	**0.61**
$(\chi^2_5)^{10}$	0.60	0.55	0.87	0.81	0.54	0.52	0.65	**0.94**
$(\chi^2_{15})^{10}$	0.21	0.21	0.35	0.32	0.22	0.22	0.17	**0.41**
$N(0,1)^9 * t_5$	0.16	0.17	0.15	0.14	0.13	0.12	0.05	0.12
$N(0,1)^9 *$ Beta$(1,1)$	**0.09**	0.05	0.04	0.04	0.04	0.04	0.06	0.03
$N(0,1)^9 *$ exp(1)	0.17	0.17	0.25	0.25	0.25	0.22	0.10	0.25
$N(0,1)^9 * \chi^2_5$	**0.40**	0.37	0.22	0.15	0.12	0.12	0.07	0.12
$NMIX_{10}(0.5,4,0,0)$	1.00	0.09	1.00	0.04	0.04	0.05	0.15	0.03
$NMIX_{10}(0.5,0,0,0.9)$	1.00	0.96	1.00	1.00	0.94	0.87	1.00	1.00
$NMIX_{10}(0.5,4,0,0.9)$	1.00	0.96	1.00	1.00	0.96	0.92	1.00	1.00
$NMIX_{10}(0.9,4,0,0)$	1.00	0.96	0.98	0.56	0.38	0.25	0.33	0.53
$NMIX_{10}(0.9,0,0,0.9)$	0.29	0.30	**0.44**	0.29	0.20	0.15	0.37	0.29
$NMIX_{10}(0.9,4,0,0.9)$	0.99	0.99	0.99	0.80	0.90	0.33	0.86	0.82
Avaerage power	0.56	0.51	0.61	0.52	0.46	0.42	0.48	0.53

Gamma and χ^2-distributions are used.

$NMIX_p(\kappa,\mu,\rho_1,\rho_2)$ denotes for the normal mixture

$$\kappa N_p(\mathbf{0},\boldsymbol{R}_1) + (1-\kappa)N_p(\mu\mathbf{1},\boldsymbol{R}_2), \tag{3.1}$$

where $\boldsymbol{R}_i$ is a matrix with diagonal elements equal to 1 and off-diagonal equal to ρ_i, $0 \le \rho_i < 1$, $i = 1,2$. According to Mecklin and Mundfrom (2005, Table 2), the mixture of normal with $\kappa = 0.5$ is severely contaminated, symmetric and short-tailed distribution. The mixture with $\kappa = 0.9$ is mildly contaminated, skewed and long-tailed.

Tables 2–7 represent the empirical power of the statistics. The average power of all alternatives are given in the last row of each table just for references. Kim (2021) mentioned that the power of the statistic DX_M shows better than or comparable to JB_M against almost all the alternatives. We can see the same phenomena for DX_{2p} and JB_{2p}, $DX_{\max}$ and$JB_{\max}$, respectively. The power of DX_{2p}, $DX_{\max}$ is better than or comparable to JB_{2p}, $JB_{\max}$, respectively. In addition, DX_{2p}, $DX_{\max}$ show very good power against some normal mixtures with the mean vector $\boldsymbol{\mu}$ in (3.1) away from 0, as it is like DX_M.

By comparing power of DX_M, DX_{2p}, and $DX_{\max}$, the statistics show similar power overall for the dimensions $d = 2$ or $d = 5$, and DX_{2p} shows the best power for some distributions with asymmetric marginals. In the meantime, DX_{2p} is superior than the other statistics DX_M, $DX_{\max}$ in power for dimension $p = 10$, and $DX_{\max}$ shows poor power for $p = 10$. DX_{2p} shows a lot better power against symmetric marginal alternatives such as Logistic$(0,1)$, t_5 and asymmetric marginal alternatives $Gamma(5,1)$, χ^2_5, and χ^2_{15} for $p = 10$. We can see almost the same phenomena for the statistics JB_M, JB_{2p}, and $JB_{\max}$.

The statistics DX_M, JB_M, DX_{2p}, JB_{2p}, $DX_{\max}$, $JB_{\max}$ that are composed of skewness and kurto-

Table 7: Power comparison of the statistics ($\alpha = 0.05$, $p = 10$, $n = 100$)

Alternative	DX_M	JB_M	DX_{2p}	JB_{2p}	$DX_{\max}$	$JB_{\max}$	HZ	MN
$N(0,1)^{10}$	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Cauchy$(0,1)^{10}$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Logistic$(0,1)^{10}$	0.28	0.32	0.66	0.67	0.42	0.42	0.16	0.64
$(t_2)^{10}$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$(t_5)^{10}$	0.78	0.80	0.97	0.97	0.89	0.86	0.51	0.96
Beta$(1,1)^{10}$	0.07	0.01	0.00	0.00	0.00	0.00	**0.75**	0.00
Beta$(2,2)^{10}$	0.04	0.01	0.00	0.00	0.00	0.00	**0.23**	0.00
Beta$(1,2)^{10}$	0.16	0.04	0.03	0.00	0.01	0.01	**0.84**	0.11
$\exp(1)^{10}$	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00
Lognormal$(0,0.5)^{10}$	0.99	0.98	1.00	1.00	1.00	0.98	1.00	1.00
Gamma$(0.5,1)^{10}$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Gamma$(5,1)^{10}$	0.58	0.53	0.89	0.81	0.57	0.54	0.65	**0.99**
$(\chi^2_5)^{10}$	0.88	0.85	1.00	0.98	0.90	0.83	0.98	1.00
$(\chi^2_{15})^{10}$	0.38	0.36	0.70	0.60	0.39	0.39	0.40	**0.88**
$N(0,1)^9 * t_5$	0.27	**0.30**	0.25	0.25	0.23	0.21	0.06	0.17
$N(0,1)^9 *$ Beta$(1,1)$	**0.28**	0.05	0.07	0.04	0.04	0.04	0.08	0.03
$N(0,1)^9 * \exp(1)$	0.29	0.30	**0.64**	0.51	0.59	0.43	0.21	0.61
$N(0,1)^9 * \chi^2_5$	**0.80**	0.67	0.46	0.29	0.23	0.18	0.10	0.27
$NMIX_{10}(0.5,4,0,0)$	1.00	0.25	1.00	0.04	0.04	0.04	0.37	0.02
$NMIX_{10}(0.5,0,0,0.9)$	1.00	1.00	1.00	1.00	1.00	0.97	1.00	1.00
$NMIX_{10}(0.5,4,0,0.9)$	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00
$NMIX_{10}(0.9,4,0,0)$	1.00	1.00	1.00	0.81	0.79	0.29	0.84	0.91
$NMIX_{10}(0.9,0,0,0.9)$	0.54	0.55	0.69	0.53	0.44	0.27	**0.74**	0.36
$NMIX_{10}(0.9,4,0,0.9)$	1.00	1.00	1.00	0.99	1.00	0.61	1.00	1.00
Avaerage power	0.67	0.61	0.71	0.63	0.59	0.52	0.65	0.65

sis have terribly poor power for the beta marginals that are symmetric shorter tailed or asymmetric, whereas HZ shows relatively good power for these alternatives.

As with HZ statistic, overall it has comparable or better power than DX_{2p} and JB_{2p} for $p = 2$ or $p = 5$. However for the dimension $p = 10$, HZ shows poor power than DX_{2p}, JB_{2p} except for the beta marginals. Although HZ is recommended as an omnibus formal test statistics for multivariate normality, it seems that the decrease in power for high dimensions is severe than it is expected.

The power of the statistic MN is very similar to that of JB_{2p} and JB_M. It shows good power for some asymmetric $Gamma(5,1)$ or χ^2_{15} marginal alternatives when $p = 10$. The remarkably good power values among all the statistics are written in bold.

4. Conclusions

The Jarque-Bera statistic is one of the well known statistics for univariate normality. That is based on the Pearson's sample skewness and kurtosis, which are the sample standardized third and fourth moments. Desgagné and de Micheaux (2018) defined the sample second power skewness and kurtosis, and proposed a normality test statistic based on them. It can be an alternative to the Jarque-Bera statistic.

Kim (2021) generalized the statistic to a multivariate version using an orthogonalization or an empirical standardization of data. In this paper, the statistic is modified by considering a few data driven directions, which are the normalized standardized scaled residuals $\{\boldsymbol{Z}_l/\|\boldsymbol{Z}_l\|, l = 1,\ldots,n\}$ with $\boldsymbol{Z}_i$ in (2.5). We consider the p least normal-like directions among $\{\boldsymbol{Z}_l/\|\boldsymbol{Z}_l\|, l = 1,\ldots,n\}$ and the principal component directions. A simulation study shows that the multivariate Desgagné and de

Micheaux (2018) statistics are better than or comparable to the multivariate Jarque-Bera statistics. And the modified Desgagné and de Micheaux (2018) statistic proposed in this paper shows better power for a big dimension.

Acknowledgement

This work was supported by 2022 Hongik University Research Fund.

References

Bowman KO and Shenton LR (1975). Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and b_2, *Biometrika*, **62**, 243–250.

D'Agostino RB and Pearson ES (1973). Tests for departure from normality: Empirical results for the distributions of b_2 and $\sqrt{b_1}$, *Biometrika*, **60**, 613–622.

D'Agostino RB and Pearson ES (1974). Correction and amendment: Tests for departure from normality: Empirical results for the distributions of b_2 and $\sqrt{b_1}$, *Biometrika*, **61**, 647–647.

D'Agostino RB and Stephens MA (1986). *Goodness-of-Fit Techniques*, Marcel Dekker, New York.

Desgagné A and de Micheaux PL (2018). A powerful and interpretable alternative to the Jarque-Bera test of normality based on 2nd-power skewness and kurtosis, using the Rao's score test on the APD family, *Journal of Applied Statistics*, **45**, 2307–2327.

Doornik JA and Hansen H (2008). An omnibus test for univariate and multivariate normality, *Oxford Bulletin of Economics and Statistics*, **70**, 927–939.

Ebner B and Henze N (2020). Tests for multivariate normality-a critcal review with emphasis on weighted L^2-statistics, *Test*, **29**, 845–892.

Farrell PJ, Salibian-Barrera M, and Naczk K (2007). On tests for multivariate normality and associated simulation studies, *Journal of Statistical Computation and Simulation*, **77**, 1065–1080.

Fang, K-T and Wang Y (1993). *Number-Theoretic Methods in Statistics*, Chapman & Hall, London.

Fattorini L (1986). Remarks on the use of the Shapiro-Wilk statistic for testing multivariate normality, *Statistica*, **46**, 209–217.

Fisher RA (1936). The use of multiple measurements in taxonomic problems, *Annals of Eugenics*, **7**, 179–188.

Hanusz Z, Enomoto R, Seo T, and Koizumi K (2018). A Monte Carlo comparison of Jarque-Bera type tests and Henze-Zirkler test of multivariate normality, *Communications in Statistics - Simulation and Computation*, **47**, 1439–1452.

Henze N (2002) Invariant tests for multivariate normality: A critical review, *Statistical Papers*, **43**, 467–506.

Henze N and Zirkler B (1990). A class of invariant consistent tests for multivariate normality, *Communications in Statistics-Theory and Methods*, **19**, 3539–3617.

Horswell RL and Looney SW (1992). A comparison of tests for multivariate normality that are based on measures of multivariate skewness and kurtosis, *Journal of Statistical Computation and Simulation*, **42**, 21–38.

Jarque C and Bera A (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals, *Economics Letters*, **6**, 255–259.

Kim N (2004). Remarks on the use of multivariate skewness and kurtosis for testing multivariate normality, *The Korean Journal of Applied Statistics*, **17**, 507–518.

Kim N (2006). Testing multivariate normality based on EDF statistics, *The Korean Journal of Applied Statistics*, **19**, 39–54.

Kim N (2015). Tests based on skewness and kurtosis for multivariate normality, *Communications for Statistical Applications and Methods*, **22**, 361–375.
Kim N (2016). A robustified Jarque-Bera test for multivariate normality, *Economics Letters*, **140**, 48–52.
Kim N (2020). Omnibus tests for multivariate normality based on Mardia's skewness and kurtosis using normalizing transformation, *Communications for Statistical Applications and Methods*, **27**, 501–510.
Kim N (2021). A Jarque-Bera type test for multivariate normality based on second-power skewness and kurtosis, *Communications for Statistical Applications and Methods*, **28**, 463–475.
Kim N and Bickel PJ (2003). The limit distribution of a test statistic for bivariate normality, *Statistica Sinica*, **13**, 327–349.
Korkmaz S, Goksuluk D, and Zararsiz G (2014). MVN : An R package for assessing multivariate normality, *The R Journal*, **6**, 151–162.
Looney SW (1995). How to use tests for univariate normality to assess multivariate normality, *The American Statistician*, **49**, 64–70.
Malkovich JF and Afifi AA (1973). On tests for multivariate normality, *Journal of the American Statistical Association*, **68**, 176–179.
Mardia KV (1970). Measures of multivariate skewness and kurtosis with applications, *Biometrika*, **57**, 519–530.
Mardia KV (1974). Applications of some measures of multivariate skewness and kurtosis for testing normality and robustness studies, *Sankhya A*, **36**, 115–128.
Mecklin CJ and Mundfrom DJ (2005). A Monte Carlo comparison of the type I and type II error rates of tests of multivariate normality, *Journal of Statistical Computation and Simulation*, **75**, 93–107.
Romeu JL and Ozturk A (1993). A comparative study of goodness-of-fit tests for multivariate normality, *Journal of Multivariate Analysis*, **46**, 309–334.
Roy SN (1953). On a heuristic method of test construction and its use in multivariate analysis, *Annals of Mathematical Statistics*, **24**, 220–238.
Shapiro SS and Wilk MB (1965). An analysis of variance test for normality (complete samples), *Biometrika*, **52**, 591–611.
Small N (1980). Marginal skewness and kurtosis in testing multivariate normality, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **29**, 85–87.
Srivastava MS and Hui TK (1987). On assessing multivariate normality based on Shapiro-Wilk *W* statistic, *Statistics & Probability Letters*, **5**, 15–18.
Srivastava DK and Mudholkar GS (2003). Goodness-of-Fit tests for univariate and multivariate normal models, In R Khattree, and CR Rao (Eds), *Handbook of Statistics 22: Statistics in Industry* (pp. 869–906), Elsevier, North Holland.
Thode Jr. HC (2002). *Testing for Normality*, Marcel Dekker, New York.
Villasenor-Alva JA and González-Estrada E (2009). A generalization of Shapiro-Wilk's test for multivariate normality, *Communications in Statistics-Theory and Methods*, **38**, 1870–1883.
Zhou M and Shao Y (2014). A powerful test for multivariate normality, *Journal of Applied Statistics*, **41**, 351–363.

Received February 7, 2023; Revised May 12, 2023; Accepted June 9, 2023