JIRSS (2023)
Vol. 22, No. 01, pp 29-47
DOI: 10.22034/jirss.2024.713812

A Principal Components Based Moment Generating Function Test for Multivariate Normality

Mbanefo Solomon Madukaife [1], **Everestus Okafor Ossai** [1], **Uchenna Chinedu Nduka**[1], **Tobias Ejiofor Ugah** [1]

[1]Department of Statistics, University of Nigeria, Nsukka, Enugu State, Nigeria.

Received: 04/03/2022, Accepted: 07/09/2023, Published online: 15/06/2024

Abstract. In this paper, we propose an alternative generalization of a recent test for univariate normality which is based on the empirical moment generating function to the multivariate case. We show, among other properties, that the proposed weighted L^2-class of statistics is affine invariant and consistent. The empirical critical values of the proposed test are evaluated for different sample sizes, variable dimensions and values of the smoothing parameter through large scale simulations. The empirical power comparison of the test with a strong competitor shows that the test has a considerably high power performance, especially at large sample sizes as well as under heavy-tailed alternative distributions. The application of the statistic, together with its competitor, to six real-life datasets also supports the considerable good power performance of the proposed statistic as well as its ease of application.

Keywords. Empirical Critical Value, Moment Generating Function, Multivariate Normality, Principal Components, Weighted L^2 –Statistic.

MSC: 62E10, 62H15, 62H25.

1 Introduction

Testing the assumption of multivariate normality (MVN) of datasets has received quite commendable attention in the literature from different researchers, probably due to the

Corresponding Author: Mbanefo Solomon Madukaife (mbanefo.madukaife@unn.edu.ng)
Everestus Okafor Ossai (everestus.ossai@unn.edu.ng)
Uchenna Chinedu Nduka (uchenna.nduka@unn.edu.ng)
Tobias Ejiofor Ugah (tobias.ugah@unn.edu.ng).

obvious importance of the multivariate normal distribution both in statistical methodology and applications. Some of the goodness-of-fit statistics include Mardia (1970, 1974), Malkovic and Afifi (1973), Royston (1983), Baringhaus and Henze (1988), Henze and Zirkler (1990), Liang et al. (2009), Villasenor Alva and Estrada (2009), Cardoso de Oliveira and Ferreira (2010), Batsidis et al. (2013), Zhou and Shao (2014), Madukaife and Okafor (2019), Liang et al. (2022), Wang et al. (2022), and Ebner et al. (2022) to mention but a few. The widest review of several goodness-of-fit techniques devoted to this subject so far, yet without being exhaustive, can be seen in Henze (2002), Thode (2002), Mecklin and Mundfrom (2004), Joenssen and Vogal (2014), Ebner and Henze (2020), and Chen and Genton (2022).

From all the comparative reviews of these tests for MVN, the L^2-class of tests which is based on the L^2-distance between the empirical and theoretical moment generating function (or Laplace transform) as well as the characteristic function has figured out to be one of the most important classes. This is because tests in this class have been proved to attain all the desirable properties of good goodness-of-fit techniques for MVN such as good control over type-I-error rate, high power performance, affine invariance and consistency.

Suppose $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$; $\mathbf{x}_j \in \boldsymbol{R}^d$, $j = 1, 2, \ldots, n$ is a sequence of n independent and identically distributed (iid) d-dimensional random vectors from an unknown distribution $F(\mathbf{x})$; where $d \geq 2$ is an integer. The problem of testing for MVN is that of testing the null hypothesis

$$H_0 : F(\boldsymbol{x}) \in F_N(\boldsymbol{x}) \tag{1.1}$$

where $F_N(\boldsymbol{x})$ is a class of nondegenerate d-dimensional multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and nondegenerate covariance matrix $\boldsymbol{\Sigma}$. A test $T(.)$ of the problem in (1.1) is said to have a good control over type-I-error if when applied at any level of significance α, does not give a measure of power more than α under the null hypothesis of MVN. Also, it is said to be affine invariant if $T(\boldsymbol{A}\mathbf{x}_1 + \boldsymbol{b}, \boldsymbol{A}\mathbf{x}_2 + \boldsymbol{b}, \ldots, \boldsymbol{A}\mathbf{x}_n + \boldsymbol{b}) = T(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ for all $\boldsymbol{b} \in \boldsymbol{R}^d$ and all $d \times d$ nonsingular matrix $\boldsymbol{A} \in \boldsymbol{R}^{d \times d}$. This property of affine invariance is important because $F_N(\mathbf{x})$ is closed with respect to full rank affine transformation. As a result, every test of the problem in (1.1) is expected to attain the property in order to accommodate all the possible distributions in the class of $F_N(\boldsymbol{x})$ belonging to the null distribution of MVN. Again, $T(.)$ is said to be consistent against all fixed alternatives (Szekeley and Rizzo, 2005) if $\lim_{n \to \infty} Pr(rejecting\, H_0 \mid F(\mathbf{x}) \notin F_N(\boldsymbol{x})) = 1$.

Sequel to the foregoing, renewed interest has been given to the already defined L^2-class. Notable among the L^2-statistics for assessing MVN in this class include the Baringhaus and Henze (1988), Henze and Zirkler (1990), Henze and Wagner (1997), Tenreiro (2017), Henze, Jimenez-Gamero and Meintanis (2019), Henze and Jimenez-Gamero (2019), Henze and Visagie (2020) and Ebner et al. (2022). Each of the statistics is based on the distance function

$$T_n = \int_{R^d} \mid \xi_n(\boldsymbol{t}) - \xi(\boldsymbol{t}) \mid^2 \omega_\beta(\boldsymbol{t}) d\boldsymbol{t} \tag{1.2}$$

where $\xi(\boldsymbol{t})$ is a theoretical function of $F(\boldsymbol{x})$ which is defined in the function of $\boldsymbol{t}$; $\xi_n(\boldsymbol{t})$ is the corresponding empirical function and $\omega_\beta(\boldsymbol{t})$ is a weighting function with $\beta > 0$ as a smoothing parameter.

Of particular interest to us in this paper is Zghoul (2010) which is a test for univariate normality ($d = 1$) based on empirical moment generating function, where $\xi(\boldsymbol{t}) = \exp\left\{2^{-1}t^2\right\}$ which is the moment generating function of a univariate standard normal distribution, $\xi_n(\boldsymbol{t}) = n^{-1}\sum_{j=1}^{n}\exp\left\{tZ_j\right\}$ which is the corresponding empirical moment generating function of $Z_j = S_n^{-1}(X_j - \overline{X})$; $\overline{X} = n^{-1}\sum_{j=1}^{n} X_j$; $S_n^2 = (n-1)^{-1}\sum_{j=1}^{n}(X_j - \overline{X})^2$; $j = 1, 2, \ldots, n$ and the weight function $\omega_n(\boldsymbol{t}) = \exp\left\{-\beta t^2\right\}$; $\beta > 0$. The mathematical theory of the Zghoul (2010) test was provided by Henze and Koch (2020) while Henze and Jimenez-Gamero (2019) generalized the test to the d−dimensional case, where $d \geq 1$. In this generalized case,

$$\xi_n(\boldsymbol{t}) = \frac{1}{n}\sum_{j=1}^{n}\exp\left\{\boldsymbol{t}^T\boldsymbol{y}_j\right\} \text{ and } \xi(\boldsymbol{t}) = \exp\left\{\frac{\|\boldsymbol{t}\|^2}{2}\right\}$$

which are respectively the empirical moment generating function of $\boldsymbol{y}_j$; $j = 1, 2, \ldots, n$ and the theoretical moment generating function of d−dimensional multivariate standard normal distribution, where $\boldsymbol{y}_j$ is the standardized observation vector given by $\boldsymbol{y}_j = \boldsymbol{S}_n^{-1/2}(\mathbf{x}_j - \overline{\boldsymbol{X}}_n)$,

$$\overline{\boldsymbol{X}}_n = \frac{1}{n}\sum_{j=1}^{n}\mathbf{x}_j \text{ and } \boldsymbol{S}_n = \frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{x}_j - \overline{\boldsymbol{X}}_n)(\mathbf{x}_j - \overline{\boldsymbol{X}}_n)^T. \tag{1.3}$$

Now, instead of the generalization to the d−dimensional case, where $d \geq 1$, in the use of the class of statistics in (1.2) as employed in Zghoul (2010), the multivariate distribution can be transformed first to a univariate case before pursuit of a statistic in this class. One such transformation is the use of principal components, which has been employed in testing for MVN of datasets in two different classes of tests respectively by Srivastava (1984) and Madukaife and Okafor (2018). The purpose of this paper therefore is to obtain L^2−statistic in the class given in (1.2) for assessing MVN by using the principal components transformation of the multivariate datasets. The rest of the paper is organized as follows: the test is developed in Section 2 with its properties while the empirical critical values of the test are obtained in Section 3, empirical power studies of the test in comparison with some other competitive tests are obtained in Section 4 while the real-life application of the test in comparison with other competing technique is obtained in Section 5. The paper is concluded in Section 6.

2 The Test Statistic

Suppose a random vector $\mathbf{x} \in \boldsymbol{R}^d$ is defined by a d-dimensional multivariate distribution $F(\boldsymbol{x})$ with mean vector $\boldsymbol{\mu}$ and a nondegenerate covariance matrix $\boldsymbol{\Sigma}$. Let $\boldsymbol{\Gamma}$ be a dxd

orthogonal matrix such that $\boldsymbol{\Gamma}^T\boldsymbol{\Sigma}\boldsymbol{\Gamma} = \boldsymbol{\Lambda}$; where $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ with λ_i as the ith largest characteristic root of $\boldsymbol{\Sigma}$. The d-component vector of uncorrelated orthogonal linear transforms known as principal components of $\mathbf{x}$ is given by

$$\boldsymbol{y} = \boldsymbol{\Gamma}^T(\boldsymbol{x} - \boldsymbol{\mu}) \tag{2.1}$$

If $\mathbf{x} \in \boldsymbol{R}^d$ is a d-dimensional multivariate normal distribution, then $\boldsymbol{y}$ is a vector of d-independent univariate normal random variables with the mean of the ith variable $\mu_i = 0$ and variance $\sigma_i^2 = \gamma_i^T\boldsymbol{\Sigma}\gamma_i = \lambda_i$, where γ_i is the ith column of $\boldsymbol{\Gamma}$; $i = 1, 2, \ldots, d$. Hence,

$$\boldsymbol{y} = \begin{cases} Y_1 \sim N(0, \lambda_1) \\ Y_2 \sim N(0, \lambda_2) \\ \quad\cdots \\ Y_d \sim N(0, \lambda_d) \end{cases}. \tag{2.2}$$

It is therefore known from the theory of statistics that

$$X_i = \sqrt{\lambda_i}^{-1} Y_i \sim N(0,1);\ i = 1, 2, \ldots, d. \tag{2.3}$$

Now, the theoretical moment generating function of each of the new transformed independent random variables X_i is given by

$$m(t) = \exp\left\{\frac{t^2}{2}\right\};\ t \in \boldsymbol{R};\ i = 1, 2, \ldots, d \tag{2.4}$$

and the corresponding empirical moment generating function from a random sample of n observations $X_{i1}, X_{i2}, \ldots, X_{in}$ drawn from each X_i; $i = 1, 2, \ldots, d$ is given by

$$m_n(t) = \frac{1}{n}\sum_{j=1}^{n} \exp\left\{tX_{ij}\right\}. \tag{2.5}$$

From (2.4) and (2.5), we propose an L^2-statistic for testing for MVN of a multivariate dataset. The statistic is given by

$$T_{n,\beta} = \sum_{i=1}^{d}\left[\int_{-\infty}^{\infty}\left(\frac{1}{n}\sum_{j=1}^{n}\exp(tX_{ij}) - \exp\left\{\frac{t^2}{2}\right\}\right)^2 \exp(-\beta t^2)dt\right] \tag{2.6}$$

where $\exp(-\beta t^2)$ is the appropriate weight function with $\beta > 0$ as the smoothing parameter.

Following the development of the statistic in (2.6), it is observed that the random variable X_i which is defined in (2.3) is obtained from a population having a multivariate normal distribution with known parameters. This however is not attainable in real life, thereby making it imperative for each of the random samples $X_{i1}, X_{i2}, \ldots, X_{in}$ to be obtained as a function of estimated mean, variance and orthogonal matrix. Suppose a sample of n independent and identically distributed (iid) observation vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$; $x_j \in \boldsymbol{R}^d$, $j = 1, 2, \ldots, n$ is available from an unknown continuous multivariate distribution function $F(\boldsymbol{x})$. In order to test for the MVN of the dataset, the unbiased estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are obtained as $\overline{\boldsymbol{X}}_n$ and $\boldsymbol{S}_n$, defined in (1.3). Let $\boldsymbol{H}_n$ be a dxd orthogonal matrix defined as an estimator of $\boldsymbol{\Gamma}$ such that $\boldsymbol{H}_n^T \boldsymbol{S}_n \boldsymbol{H}_n = \boldsymbol{L}_n$; where $\boldsymbol{L}_n = \text{diag}(l_1, l_2, \ldots, l_d)$ with l_i as the ith largest characteristic root of $\boldsymbol{S}_n$. Then, the jth orthogonal linear transform of the sample observation vectors in the ith principal component is obtained as

$$Y_{ij} = \boldsymbol{h}_i^T (\mathbf{x}_j - \overline{\boldsymbol{X}}_n) \tag{2.7}$$

where $\boldsymbol{h}_i$ is the orthonormal vector in the ith column of $\boldsymbol{H}_n$. With (2.7), it is not difficult to obtain estimated normal observations with zero mean and unit variance as

$$X_{ij} = \frac{Y_{ij}\sqrt{n}}{\sqrt{(n+1)l_i}}; \; i = 1, 2, \ldots, d; \; j = 1, 2, \ldots, n. \tag{2.8}$$

Direct substitution of (2.8) into (2.6) with straightforward expansion and simplification gives:

$$T_{n,\beta} = \sum_{i=1}^{d} \left[\int_{-\infty}^{\infty} \left(\frac{1}{n} \sum_{j=1}^{n} \exp(tX_{ij}) - \exp\left\{\frac{t^2}{2}\right\} \right) \left(\frac{1}{n} \sum_{j=1}^{n} \exp(tX_{ij}) - \exp\left\{\frac{t^2}{2}\right\} \right) \exp(-\beta t^2) dt \right]$$

$$= \sum_{i=1}^{d} \left[\int_{-\infty}^{\infty} \frac{1}{n^2} \sum_{j,k=1}^{n} \exp\left\{t\left(X_{ij} + X_{ik} - \beta t\right)\right\} + \exp\left\{t^2(1-\beta)\right\} - \frac{2}{n} \sum_{j=1}^{n} \exp\left\{\frac{t}{2}\left(2X_{ij} + t - 2\beta t\right)\right\} dt \right]$$

which direct integration and simplification produces the statistic in a computational form as:

$$T_{n,\beta} = \sqrt{\frac{\pi}{\beta}} \sum_{i=1}^{d} \left[\frac{1}{n^2} \sum_{j,k=1}^{n} \exp\left\{\frac{(X_{ij} + X_{ik})^2}{4\beta}\right\} - \frac{2}{n}\sqrt{\frac{2\beta}{2\beta-1}} \sum_{j=1}^{n} \exp\left\{\frac{X_{ij}^2}{4\beta-2}\right\} + \sqrt{\frac{\beta}{\beta-1}} \right] \tag{2.9}$$

where $\beta \in R > 1$.

Considering the affine invariance of the statistic, Pudelko (2005) has stated that it is sufficient to show that a statistic, $T_n(\boldsymbol{y}_j)$ is affine invariant if the value of the statistic remains unchanged when $\boldsymbol{y}_j$, $j = 1, 2, \ldots, n$ is replaced with $\boldsymbol{z} = \boldsymbol{A}\boldsymbol{y}_j$ for any orthogonal matrix $\boldsymbol{A}$. This however is not the case with $T_{n,\beta}$ because it is not a function of observation vectors, $\boldsymbol{y}_j$, $j = 1, 2, \ldots, n$, but a function of standardized observations, X_{ij}, $i = 1, 2, \ldots, d$ $j = 1, 2, \ldots, n$. Hence, no matter the scale of the observation vectors, the value of $T_{n,\beta}$ remains unchanged because the observation vectors are transformed and standardized to the unit root prior to evaluation. Therefore, the statistic is obviously affine invariant. For the choice of $\exp\left\{-\beta t^2\right\}$ as an appropriate weight function, see Zghoul (2010).

Theorem 2.1. *Under the null distribution of multivariate normality, the expected value of $T_{n,\beta}$ is $E(T_{n,\beta}) = \frac{d\sqrt{\pi}}{n}\left(\frac{1}{\sqrt{\beta-2}} - \frac{1}{\sqrt{\beta-1}}\right)$; $\beta \in \boldsymbol{R} > 2$.*

Proof. From (2.6), $E(T_{n,\beta}) = \sum_{i=1}^{d}\int_{-\infty}^{\infty} E\left[\left(\frac{1}{n}\sum_{j=1}^{n}\exp\left\{tX_{ij}\right\} - \exp\left\{\frac{t^2}{2}\right\}\right)^2 \exp\left\{-\beta t^2\right\}\right]dt$
$= \sum_{i=1}^{d}\int_{-\infty}^{\infty}\exp\left\{-\beta t^2\right\}E\left[\frac{1}{n^2}\left(\sum_{j=k}^{n}\exp\left\{2tX_{ij}\right\} + \sum_{j\neq k}^{n}\exp\left\{t(X_{ij}+X_{ik})\right\}\right)\right.$
$\left. + \exp\left\{t^2\right\} - \frac{2}{n}\exp\left\{\frac{t^2}{2}\right\}\sum_{j=1}^{n}\exp\left\{tX_{ij}\right\}\right]dt$
$= \sum_{i=1}^{d}\int_{-\infty}^{\infty}\exp\left\{-\beta t^2\right\}\left[\frac{1}{n^2}\left(nE\exp\left\{2tX_{ij}\right\} + n(n-1)E\exp\left\{t(X_{ij}+X_{ik})\right\}\right)\right.$
$\left. + \exp\left\{t^2\right\} - 2\exp\left\{\frac{t^2}{2}\right\}E\exp\left\{tX_{ij}\right\}\right]dt$

But X_{ij}, X_{ik} are iid $N(0,1)$ for each i, $i = 1, 2, \ldots, d$. So $E\left(\exp\left\{tX_{ij}\right\}\right)$ is the moment generating function of $X_i \sim N(0,1)$ which is $\exp\left\{\frac{t^2}{2}\right\}$; $i = 1, 2, \ldots, d$. Similarly, $E\left(\exp\left\{2tX_{ij}\right\}\right)$ is $\exp\left\{2t^2\right\}$ for each i. Substituting appropriately and upon simplification gives the expected value of the statistic $T_{n,\beta}$ as $d\int_{-\infty}^{\infty}\exp\left\{-\beta t^2\right\}\frac{1}{n}\left(\exp\left\{2t^2\right\} - \exp\left\{t^2\right\}\right)dt$, which integrates explicitly to give the required result. □

Theorem 2.2. *Suppose $F_N(\boldsymbol{x})$ is a class of non-degenerate d-variate normal distributions, $d \geq 2$, with unknown parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Then, for any fixed value of $\beta > 1$, there exists $c_{\alpha,n,d}$ such that $\lim_{n\to\infty} T_{n,\beta} \geq c_{\alpha,n,d}$.*

Proof. Csorgo (1989) has shown the existence of a constant $c_0 > 0$ such that for every $t \in R^d$,

$$\lim_{n\to\infty} inf \left\| \left| \frac{1}{n}\sum_{j=1}^{n}\left\{itY_{ij}\right\}\right|^2 - exp\left\{-|t|^2\right\}\right| \geq c_0 \ a.s.$$

By the same principle, there exists $c_{\alpha,n,d}$ such that $\lim_{n\to\infty} T_{n,\beta} \geq c_{\alpha,n,d}$ where $c_{\alpha,n,d}$ is such that $P\left(T_{n,\beta} \geq c_{\alpha,n,d}\right) = \alpha$. □

Now, $\lim_{n\to\infty} T_{n,\beta} \geq 0$,with equality if and only if the sampled population is multinormal. That is, if the set of independent observation vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$; $\boldsymbol{x}_j \in \boldsymbol{R}^d$, $j = 1, 2, \ldots, n$ is actually from a nondegenerate d-variate normal distribution with unknown mean vector $\boldsymbol{\mu}$ and unknown covariance matrix $\boldsymbol{\Sigma}$, then, $X_{i1}, X_{i2}, \ldots, X_{in}$ for each i will have a normal distribution with zero mean and unit variance such that $E\left(m(t) - m_n(t)\right)^2 \to 0$ as $n \to \infty$, since if $X \geq Y$, then $E(X) \geq E(Y)$.

From the foregoing, the test developed in this paper is affine invariant and consistent against any fixed alternative. It rejects MVN of datasets for large values of $T_{n,\beta}$. It is also important to note that the proposed statistic enjoys all the properties of the Zghoul (2010) statistic which has been presented by Henze and Koch (2020) in a similar form. This assertion stems from the fact that under the null distribution of MVN, $T_{n,\beta}$ is simply a sum of statistics from d-independent and identically distributed random variables. For instance, it shall be shown that the limit form of the proposed $T_{n,\beta}$ statistic corresponds to Theorem 4 of Henze and Koch (2020) for the Zghoul (2010) statistic.

Theorem 2.3: Suppose

$$\tau(\beta) = d\left(\frac{1}{\sqrt{\beta-1}} - \frac{2}{\sqrt{\beta-\frac{1}{2}}} - \frac{2}{(4\beta-2)\sqrt{\beta-\frac{1}{2}}} + \frac{1}{\sqrt{\beta}} + \frac{1}{2\beta^{3/2}} + \frac{3}{16\beta^{5/2}}\right)$$

and

$$b_{n,1}^{(i)} = \frac{\frac{1}{n}\sum_{j=1}^{n}(X_{ij}^{*}-\overline{X}_{in}^{*})^{3}}{S_{in}^{3}}$$

is the skewness of unstandardized forms of the values $X_{i1},\ X_{i2},\ \ldots,\ X_{in}$ where $S_{in}^{3} = \left(\frac{1}{n}\sum_{j=1}^{n}(X_{ij}^{*}-\overline{X}_{in}^{*})^{2}\right)^{3/2}$ with X_{ij}^{*} as an unstandardized form of the jth transformed observation in the ith principal component and $\overline{b}_{n,1}^{2}$ is the average of $b_{in,1}^{2}$ over d principal components. Then, the proposed $T_{n,\beta}$ statistic is such that

$$\lim_{\beta\to\infty}\frac{96}{5}\beta^{7/2}\left(\frac{T_{n,\beta}}{d\sqrt{\pi}} - \tau(\beta)\right) = \overline{b}_{n,1}^{2}.$$

Proof: From the standardized observations X_{ij}, we have

$$\sum_{j=1}^{n}X_{ij} = 0,\ \sum_{j=1}^{n}X_{ij}^{2} = n,\ \sum_{j=1}^{n}X_{ij}^{3} = nb_{n,1}^{(i)},\ \sum_{j=1}^{n}X_{ij}^{4} = nb_{n,2}^{(i)};\ \text{where}\ nb_{n,2}^{(i)} = \frac{\frac{1}{n}\sum_{j=1}^{n}(X_{ij}^{*}-\overline{X}_{in}^{*})^{4}}{S_{in}^{4}}$$

Expansion of the exponential terms in (2.9) for each ith term gives

$$\sum_{j,k=1}\exp\left\{\frac{(X_{ij}+X_{ik})2}{4\beta}\right\} = \sum_{j,k=1}\left(1 + \frac{(X_{ij}+X_{ik})2}{4\beta} + \frac{(X_{ij}+X_{ik})4}{32\beta^{2}} + \frac{(X_{ij}+X_{ik})6}{384\beta^{3}} + O(\beta^{-4})\right)$$

$$= \sum_{j,k=1}\left(1 + \frac{X_{ij}^{2}+X_{ik}^{2}+2X_{ij}X_{ik}}{4\beta} + \frac{X_{ij}^{4}+X_{ik}^{4}+4X_{ij}^{3}X_{ik}+4X_{ij}X_{ik}^{3}+6X_{ij}^{2}X_{ik}^{2}}{32\beta^{2}}\right.$$

$$\left.+\frac{X_{ij}^{6}+X_{ik}^{6}+6X_{ij}^{5}X_{ik}+6X_{ij}X_{ik}^{5}+15X_{ij}^{4}X_{ik}^{2}+15X_{ij}^{2}X_{ik}^{4}+20X_{ij}^{3}X_{ij}^{3}}{384\beta^{3}} + O(\beta^{-4})\right)$$

$$= n^{2} + \frac{n^{2}}{2\beta} + \frac{n^{2}b_{n,2}^{(i)}}{16\beta^{2}} + \frac{3n^{2}}{16\beta^{2}} + \frac{n}{192\beta^{3}}\sum_{j=1}^{n}X_{ij}^{6} + \frac{5n^{2}b_{n,2}^{(i)}}{64\beta^{3}} + \frac{5n^{2}b_{n,1}^{(i)2}}{96\beta^{3}} + O(\beta^{-4}) \qquad (2.10)$$

$$\text{and} \sum_{j=1}^{n} \exp\left\{\frac{X_{ij}^2}{4\beta-2}\right\} = \sum_{j=1}^{n}\left(1+\frac{X_{ij}^2}{4\beta-2}+\frac{X_{ij}^4}{2(4\beta-2)^2}+\frac{X_{ij}^6}{6(4\beta-2)^3}+O(\beta^{-4})\right)$$
$$= n+\frac{n}{4\beta-2}+\frac{nb_{n,2}^{(i)}}{2(4\beta-2)^2}+\frac{1}{6(4\beta-2)^3}\sum_{j=1}^{n}X_{ij}^6+O(\beta^{-4}) \quad (2.11)$$

Direct substitution of (2.10) and (2.11) into (2.9) with straightforward algebra gives

$$\frac{T_{n,\beta}}{\sqrt{\pi}}-\tau(\beta)=\frac{\sum_{i=1}^{d}b_{n,2}^{(i)}}{16\beta^{5/2}}+\frac{\sum_{i=1}^{d}\sum_{j=1}^{n}X_{ij}^6}{192n\beta^{7/2}}+\frac{5\sum_{i=1}^{d}b_{n,2}^{(i)}}{64\beta^{7/2}}+\frac{5\sum_{i=1}^{d}b_{n,1}^{(i)2}}{96\beta^{7/2}}-\frac{\sum_{i=1}^{d}b_{n,2}^{(i)}}{(4\beta-2)^2\sqrt{\beta-\frac{1}{2}}}-\frac{\sum_{i=1}^{d}\sum_{j=1}^{n}X_{ij}^6}{3n(4\beta-2)^3}. \quad (2.12)$$

Upon simplification of (2.12), we obtain

$$\frac{T_{n,\beta}}{\sqrt{\pi}}-\tau(\beta)=\frac{5\sum_{i=1}^{d}b_{n,1}^{(i)2}}{96\beta^{7/2}}+o(\beta^{-7/2}) \quad (2.13)$$

Note that the result in (2.13) is possible since $\frac{1}{3(4\beta-2)^3}=\frac{1}{192\beta^3}+O(\beta^{-4})$; $\frac{1}{192\beta^3\sqrt{\beta-\frac{1}{2}}}=\frac{1}{192\beta^{7/2}}+o(\beta^{-7/2})$ and $\frac{1}{16\beta^{5/2}}-\frac{1}{(4\beta-2)^2\sqrt{\beta-\frac{1}{2}}}=\frac{1}{16\beta^{5/2}}-\frac{1}{16\beta^2(1-\frac{1}{2\beta})^2\sqrt{\beta}\sqrt{1-\frac{1}{2\beta}}}=-\frac{5}{64\beta^{7/2}}+o(\beta^{-7/2})$ (Henze & Koch, 2020). Hence, the theorem is proved.

3 Critical Values of the Test

The exact distribution of the $T_{n,\beta}$ statistic is not explicitly known. This makes it difficult to obtain the theoretical critical values of the statistic and hence, empirical counterparts are obtained. They are computed for different combinations of the sample size n, number of variables d and smoothing parameter $\beta > 2$ through extensive simulation studies. Precisely, the critical values at the level of significance, $\alpha = 0.01$ and 0.05 for $n = 25$, 50 and 100 and $d = 2$, 3 and 5 are evaluated. $N = 100{,}000$ samples were generated from a d-dimensional standard multivariate normal distribution and the N values of the $T_{n,\beta}$ statistic were obtained from each generated set of N samples under each specified n, d and $\beta > 2$. The α-level critical value of the test for each n, d and $\beta > 2$ is then obtained as the $100(1-\alpha)$ percentile of the N values. The percentile values δ are presented in terms of τ in Table 1, where $\delta = 10^{-3}\tau$. It is important to note that no effort is made to obtain the critical values of the statistic for all the sample sizes n, number of variables d and smoothing parameter $\beta > 2$ which may be encountered in real-life applications as such will definitely be an effort in futility. It is indeed practicably impossible to have all of them computed and listed but it is expected that the statistic will be implemented in statistical software such as R in such a manner that for each data specification, the critical value is computed along with the statistic for an appropriate decision to be taken. As a result, the percentile values presented in Table 1 may be regarded as appropriate for demonstration purposes.

Table 1: Empirical critical values by 10^{-3}

			β				
d	n	α	2.5	5.0	10.0	20.0	30.0
2	25	0.01	15.859100	0.682109	0.051063	0.005096	0.001467
		0.05	40.359300	1.264105	0.078698	0.006876	0.001864
	50	0.01	13.528520	0.449740	0.027885	0.002375	0.000617
		0.05	41.914250	0.910553	0.046341	0.003558	0.000874
	100	0.01	10.045990	0.270579	0.015019	0.001159	0.000281
		0.05	35.132200	0.539123	0.025105	0.001799	0.000419
3	25	0.01	23.618320	0.976036	0.071641	0.007269	0.002101
		0.05	55.128690	1.651028	0.102935	0.009342	0.002545
	50	0.01	20.351630	0.623886	0.038226	0.003278	0.000859
		0.05	59.745860	1.187405	0.060408	0.004627	0.001146
	100	0.01	14.604610	0.360551	0.020157	0.001565	0.000384
		0.05	47.652580	0.703736	0.032161	0.002313	0.000548
5	25	0.01	37.661760	1.479510	0.110379	0.011415	0.003346
		0.05	80.858860	2.338584	0.146469	0.013769	0.003859
	50	0.01	32.298960	0.930090	0.056673	0.004959	0.001325
		0.05	85.156500	1.638061	0.082422	0.006449	0.001659
	100	0.01	23.162230	0.531502	0.029968	0.002314	0.000560
		0.05	68.536300	0.934430	0.043760	0.003208	0.000733

4 Empirical Power Studies

In this section, empirical power performance of the $T_{n,\beta}$ statistic is compared with the powers of Villasenor Alva and Estrada (2009) and the Henze and Jimenez – Gamero (2019) statistics through extensive simulation studies. The Shapiro-Wilks test due to Villasenor Alva and Estrada (2009) is considered as a competitor in this work due to its general acceptance as a test for MVN with a good power performance. Also, the Henze and Jimenez – Gamero (2019) statistic is considered in this work as a competitor because of its affinity with the $T_{n,\beta}$ test: both of them are extensions of the Zghoul (2010) statistic for univariate normality to multivariate sphere. However, while the Henze and Jimenez – Gamero (2020) extension statistic is a direct function of multivariate datasets, the $T_{n,\beta}$ extension statistic is a function of univariate transformed datasets. The generalized Shapiro-Wilks statistic of Villasenor Alva and Estrada (2009) is given by:

$$VE_n = \frac{1}{p}\sum_{i=1}^{p} W_{Z_i},$$

where W_{Z_i} is the Shapiro - Wilk's statistic (Shapiro & Wilk, 1965) evaluated on the ith coordinate of the transformed observations $Z_{i1}, Z_{i2}, \ldots, Z_{in}$; $i = 1, 2, \ldots, p$ and p is the number of variables. Also, the universally consistent and affine – invariant test due to Henze and Jimenez–Gamero (2020) is given by:

$$HJ - G_{n,\beta} =$$

$$\pi^{d/2}\left(\frac{1}{n}\sum_{j,k=1}^{n}\frac{1}{\beta^{d/2}}\exp\left\{\frac{\left\|Y_{n,j}+Y_{n,k}\right\|^2}{4\beta}\right\} + \frac{n}{(\beta-1)^{d/2}} - 2\sum_{j=1}^{n}\frac{1}{(\beta-\frac{1}{2})^{d/2}}\exp\left\{\frac{\left\|Y_{n,j}\right\|^2}{4\beta-2}\right\}\right); \beta > 1$$

where $Y_{n,j}$ is the jth d-dimensional standardized multivariate data point contained in the standardized sample of size n and $\|.\|$ is a vector norm. Like the $T_{n,\beta}$ statistic, the $HJ - G_{n,\beta}$ test rejects the null distribution of MVN for large values of the statistic.

In this comparison, four different classes of distributions alternative to the MVN are identified. They include the short-tailed symmetric, heavy-tailed symmetric, short-tailed non-symmetric and heavy-tailed non-symmetric distributions. A total of 10,000 data sets were generated in each combination of n = 25, 50 and 100 and d = 2, 3 and 5 from seven different multivariate distributions representing these classes. The distributions include the standard multivariate normal distribution (MVN); the standard multivariate Laplace distribution (MVL), for the short-tailed symmetric class of distributions; the standard multivariate t distribution with 2 degrees of freedom (MVt_2) and the standard multivariate Cauchy distribution (MVC), for the heavy-tailed symmetric class of distributions; the standard multivariate exponential distribution (MVexp) for the short-tailed non-symmetric class of distributions. Others are products of standard univariate Pareto distributions (Pareto) and products of standard lognormal distributions (LN), for heavy-tailed non-symmetric classes of distributions. The values of each of the three statistics ($T_{n,\beta}$, VE_n and $HJ - G_{n,\beta}$) are evaluated in each of the 10,000 simulated samples for each n and d from each of the distributions. The empirical power of each test statistic is obtained as the percentage of the 10,000 samples that are rejected by the statistic at $\alpha = 5\%$. The power performance is presented in Tables 2, 3 and 4 respectively for $n = 25, 50$ and 100.

Preliminary studies carried out on the $T_{n,\beta}$ and $HJ - G_{n,\beta}$ statistics show that they maintain consistent powers for $\beta \in (2, 30)$. However, the power behaviours of the two statistics across the interval are not exactly the same. Hence, each power comparison table for $\beta = 2.5$, 10 and 30 contains 4th power performance column for each of the statistics which has maximum power performances for each statistic under the conditions of n, d and the considered interval for β.

From the power studies, it can be seen that all the statistics considered have very good control over type-I-error. This is because the power performances of the three statistics under the null distribution of standard multivariate normal distribution in all the sample sizes and variable dimensions are approximately equal to 5% (4.5 < power < 5.4) which is the level of significance. However, there is no clearly identified pattern of the power across n and d in all the statistics. On the other hand, the powers of the $T_{n,\beta}$ statistic were observed to be slightly higher than the VE_n statistic in most of the alternative distributions considered especially at smaller sample sizes. Also, they are observed to be slightly lower than those of the $HJ - G_{n,\beta}$ statistic in most of the alternative distributions considered at smaller sample sizes. In the case of comparison between the $T_{n,\beta}$ and the $HJ - G_{n,\beta}$, the result may be expected since $T_{n,\beta}$ procedure involves use of more estimated values than the $HJ - G_{n,\beta}$ procedure. That is, the new procedure first estimates the covariance matrix and the mean vector of the datasets which are in turn used to estimate the principal components datasets as well as their means and variances which are used to obtain the statistic. In contrast, the $HJ - G_{n,\beta}$ procedure only requires estimation of the covariance matrix and may be, the mean vector of the dataset to obtain the statistic. However, the slight differences between the power performances of the three statistics at small sample sizes are observed to approach zero with increasing sample size. In terms of power performance therefore, it is obvious that the three test procedures are almost at par in large samples, regardless of the distribution.

Besides power comparativeness of the new statistic, the relative ease of its computation can be considered. It has been stated earlier that it is a function of univariate datasets. No doubt, it is known that the tedium of computation involving univariate datasets is well less than that of multivariate counterparts. Since the observed power differences between the $T_{n,\beta}$ and the $HJ - G_{n,\beta}$ are insignificant, the superiority of the $T_{n,\beta}$ to the $HJ - G_{n,\beta}$ therefore stems from their degrees of complexity in application. Also, another intrinsic quality of the power performances of the new test in Tables 2 through 4 is a well defined pattern of power behaviour across β and d for different classes of distributions alternative to the MVN considered in the paper. Under the short-tailed and heavy-tailed symmetric distributions, the powers of the $T_{n,\beta}$ statistic decreased progressively from $\beta = 2.5$ to 30, but increased progressively from $d = 2$ to 5. On the other hand, the powers increased progressively for the same conditions of β under the short-tailed and heavy-tailed non-symmetric distributions. Under the short-tailed non-symmetric distributions, the powers decreased progressively from $d = 2$ to 5 while the reverse is the case under the heavy-tailed non-symmetric distributions.

5 Real-Life Applications

The applicability of the proposed $T_{n,\beta}$ statistics is presented in this section, in comparison with the $HJ - G_{n,\beta}$ statistic. It is done on a set of six different multivariate datasets. The datasets are described as follows:

Certificates of analysis dataset (Coa): This is a 4-component dataset of 122 observation vectors, obtained as measurements on four properties of an important powder raw material with the properties as impurity level, particle size between 1.80 and 2.40 microns, particle size between 4.00 and 5.70 microns and compressibility. It is well known that if a d- component random vector is multivariate normal, then, any r-component marginal random vector obtained from it , $r < d$, is also multivariate normal. As a result, only the first two components of the dataset are used in this study ($d = 2$, $n = 122$) and the dataset is retrieved from https://www.openmv.net/tag/multivariate.

Table 2: Empirical power comparison of the $T_{n,\beta}$, VE_n and $HJ - G_{n,\beta}$ tests at α= 0.05 and $n = 25$, presented in percentage

Distributions	d	$T_{n,\beta}$				VE_n	$HJ - G_{n,\beta}$			
		$\beta = 2.5$	$\beta = 10$	$\beta = 30$	max		$\beta = 2.5$	$\beta = 10$	$\beta = 30$	max
MVN	2	4.9	5.3	5.1	5.3	5.2	5.3	5.0	5.3	5.3
	3	4.7	4.7	5.0	5.0	4.9	5.1	5.2	4.5	5.2
	5	4.6	5.2	5.0	5.2	5.3	5.0	4.6	4.7	5.0
MVLaplace	2	40.8	36.6	35.4	40.8	39.7	49.9	48.0	46.1	49.9
	3	46.5	44.8	41.5	46.5	44.1	61.6	63.1	62.5	63.1
	5	55.0	53.1	50.8	55.0	46.0	79.8	86.1	85.6	86.1
MVt(2)	2	74.8	71.1	67.7	74.8	73.0	79.6	79.9	78.5	79.9
	3	81.2	78.2	75.6	81.2	75.7	89.5	89.1	89.2	89.5
	5	88.6	86.5	84.5	88.6	78.5	96.1	97.0	96.7	97.0
MVC	2	95.3	93.0	90.5	95.3	96.8	97.4	97.3	97.0	97.4
	3	97.4	96.2	94.6	97.4	97.3	99.3	99.4	99.3	99.4
	5	98.8	98.4	98.0	98.8	97.9	100.0	100.0	100.0	100.0
MVexp	2	72.4	81.7	82.9	82.9	82.2	76.9	89.1	90.7	90.7
	3	70.3	78.8	80.5	80.5	89.2	82.5	93.7	95.9	95.9
	5	68.7	75.2	76.9	76.9	90.1	83.2	94.2	96.5	96.5
Pareto	2	99.6	99.8	99.9	99.9	100.0	99.7	100.0	100.0	100.0
	3	99.8	99.9	99.9	99.9	100.0	99.9	100.0	100.0	100.0
	5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
LN(0,1)	2	92.2	95.3	95.3	95.3	99.9	92.2	97.9	98.2	98.2
	3	93.9	96.7	97.2	97.2	100.0	94.9	98.9	99.4	99.4
	5	95.0	97.0	97.0	97.0	100.0	97.0	99.5	99.8	99.8

Table 3: Empirical power comparison of the $T_{n,\beta}$, VE_n and $HJ-G_{n,\beta}$ tests at α= 0.05 and n = 50, presented in percentage

Distribution	d	$T_{n,\beta}$				VE_n	$HJ-G_{n,\beta}$			
		$\beta=2.5$	$\beta=10$	$\beta=30$	max		$\beta=2.5$	$\beta=10$	$\beta=30$	max
MVN	2	5.1	5.2	5.3	5.3	5.2	4.7	5.1	4.7	5.1
	3	4.7	4.7	5.1	5.1	4.7	5.2	5.1	5.1	5.2
	5	5.2	5.3	4.9	5.3	4.8	4.6	5.3	5.0	5.3
MVLaplace	2	61.5	51.6	46.7	61.5	68.9	71.1	66.6	62.1	71.1
	3	70.9	61.0	55.6	70.9	78.8	83.5	84.8	81.3	84.8
	5	80.8	73.6	67.0	80.8	86.6	96.1	98.1	97.5	98.1
MVt(2)	2	93.3	89.3	85.2	93.3	95.7	96.6	94.8	93.1	96.6
	3	97.0	94.8	91.8	97.0	97.5	98.9	99.0	98.3	99.0
	5	99.0	98.0	96.3	99.0	99.0	100.0	100.0	99.9	100.0
MVC	2	99.8	99.6	98.2	99.8	99.9	99.9	99.9	99.8	99.9
	3	100.0	100.0	99.4	100.0	100.0	100.0	100.0	100.0	100.0
	5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
MVexp	2	91.5	97.2	97.8	97.8	90.2	95.5	99.8	99.9	99.9
	3	89.4	95.4	96.1	96.1	96.8	97.3	100.0	100.0	100.0
	5	89.2	93.2	94.3	94.3	99.1	98.1	100.0	100.0	100.0
Pareto	2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	3	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
LN(0,1)	2	99.2	99.7	100.0	100.0	100.0	99.7	100.0	100.0	100.0
	3	99.6	99.8	100.0	100.0	100.0	99.9	100.0	100.0	100.0
	5	99.8	100.0	100.0	100.0	100.0	99.9	100.0	100.0	100.0

Wine dataset (Wn): This is a 14-component dataset with 178 observation vectors, retrieved from https://archive.ics.uci.edu/ml/datasets. It was generated as the result of a chemical analysis of wines grown in the same region in Italy from three different cultivars (variable one). The analysis determined the quantities of 13 different constituents found in each of the wines. For the same obvious reason, only data on two of the constituents (ash and alkalinity) are used in this study (d = 2, n = 178).

Room temperature dataset (Rt): The dataset is temperature measurements, in Kelvin, taken from four corners of a room. It consists of four components representing the four corners of a room with 144 measured observation vectors, retrieved from https://archive.ics.uci.edu/ml/datasets. Only the first three components (d = 3, n = 144) of the dataset are used in this study.

Film thickness dataset (Fth): The dataset consists of four components representing measurements taken at four positions of 160 plastic films after being cut. The positions of measurement included top right, top left, bottom right and bottom left and the data were retrieved from https://archive.ics.uci.edu/ml/datasets, (d = 4, n = 160).

Table 4: Empirical power comparison of the $T_{n,\beta}$, VE_n and $HJ-G_{n,\beta}$ tests at α= 0.05 and n = 100, presented in percentage

Distribution	d	$T_{n,\beta}$				VE_n	$HJ-G_{n,\beta}$			
		$\beta=2.5$	$\beta=10$	$\beta=30$	max		$\beta=2.5$	$\beta=10$	$\beta=30$	max
MVN	2	4.6	5.0	5.4	5.4	4.8	4.6	4.8	4.7	4.8
	3	4.8	5.2	5.1	5.2	4.8	4.7	4.7	5.2	5.2
	5	5.0	5.0	4.9	5.0	4.6	5.1	4.9	5.4	5.4
MVLaplace	2	82.2	68.1	56.8	82.2	94.4	89.3	84.4	77.5	89.3
	3	89.7	76.4	67.8	89.7	98.1	97.0	97.0	93.6	97.0
	5	96.3	82.8	70.5	96.3	99.6	99.8	99.9	99.8	99.9
MVt(2)	2	99.7	98.6	96.1	99.7	99.9	99.9	99.7	99.2	99.9
	3	100.0	99.8	98.7	100.0	100.0	100.0	100.0	100.0	100.0
	5	100.0	100.0	99.4	100.0	100.0	100.0	100.0	100.0	100.0
MVC	2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	3	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
MVexp	2	99.1	100.0	100.0	100.0	99.2	99.7	100.0	100.0	100.0
	3	98.5	99.8	100.0	100.0	100.0	97.3	100.0	100.0	100.0
	5	95.0	96.2	99.4	99.4	100.0	96.1	100.0	100.0	100.0
Pareto	2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	3	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
LN(0,1)	2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	3	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Food texture dataset (Fte): The dataset is a 5-component texture measurement of a pastry-type food, retrieved from https://archive.ics.uci.edu/ml/datasets. It consists of 50 observation vectors, with vector having measurements on oil percentage, product's density, crispiness measurement, the angle, in degrees, through which the pasty can be slowly bent before it fractures, and hardness measurement, (d = 5, n = 50).

Blender efficiency dataset (Be): The dataset is the effect of four factors on blending efficiency, with the factors as particle size, mixer diameter, mixer rotational speed and blending time, giving rise to a 5-component dataset consisting of the four factors and the blending efficiency. It is retrieved from https://archive.ics.uci.edu/ml/datasets, with only 18 observation vectors, (d = 5, n = 50).

Now, each of the six datasets, as used in this study, is tested for multivariate normality at 5% level of significance using the $T_{n,\beta}$, VE_n and $HJ-G_{n,\beta}$ statistics. In the cases of $T_{n,\beta}$ and $HJ-G_{n,\beta}$, the statistics are applied at three different smoothing parameters, namely: $\beta = 2.5$, 10 and 30 and multinormality of each dataset is rejected if the result leads to rejection in at least one of the three instances of β. The results are presented in Table 5 with critical values in parentheses, 'DNR' being decision not to reject the null hypothesis of multinormality and 'Reject' being decion to reject the null hypothesis of multinormality.

Table 5: Test for multivariate normality of the real-life datasets

Data	d	n	$T_{n,\beta}$ test statistic values			VE_n	$HJ-G_{n,\beta}$ test statistic values			Decision		
			$\beta=2.50$	$\beta=10.0$	$\beta=30.0$		$\beta=2.50$	$\beta=10.0$	$\beta=30.0$	$T_{n,\beta}$	VE_n	$HJ-G_{n,\beta}$
Coa	2	122	0.0044	0.00001	0.00000	0.97992	0.29675	0.00031	0.00000	DNR	DNR	DNR
			(0.0079)	(0.0000)	(0.0000)	(0.9788)	(0.9497)	(0.0004)	(0.0000)			
Wn	2	178	0.0131	0.00004	0.00000	0.92045	1.20449	0.00128	0.00001	Reject	Reject	Reject
			(0.0058)	(0.0000)	(0.0000)	(0.9790)	(0.9454)	(0.0004)	(0.0000)			
Rt	3	144	0.0719	0.00012	0.00000	0.96632	4.55140	0.00114	0.00001	Reject	Reject	Reject
			(0.0121)	(0.0000)	(0.0000)	(0.9762)	(1.8618)	(0.0003)	(0.0000)			
Fth	4	160	0.0224	0.00002	0.00000	0.94467	2.47755	0.00012	0.00000	Reject	Reject	DNR
			(0.0130)	(0.0000)	(0.0000)	(0.9628)	(3.1092)	(0.0001)	(0.0000)			
Fte	5	50	0.0780	0.00006	0.00000	0.92278	39.70505	0.00019	0.00000	Reject	Reject	Reject
			(0.0104)	(0.0000)	(0.0000)	(0.9660)	(2.4021)	(0.0001)	(0.0000)			
Be	5	18	0.0312	0.00013	0.00001	0.94389	0.59031	0.00003	0.00000	DNR	DNR	DNR
			(0.0334)	(0.0003)	(0.0000)	(0.9316)	(0.6485)	(0.0000)	(0.0000)			

From the results in Table 5, the $T_{n,\beta}$, VE_n and $HJ - G_{n,\beta}$ tests presented similar decisions in all the six datasets except in the Film thickness dataset where the proposed $T_{n,\beta}$ and VE_n statistics rejected its multivariate normality which the later failed to do. The results showed the $T_{n,\beta}$ statistic as being 'tighter' than the $HJ - G_{n,\beta}$ statistic but at par with the VE_n statistic. Hence, it is a good statistic for multinormality of datasets.

6 Conclusion

The proposed $T_{n,\beta}$ statistic maintains similar theoretical properties with the Zghoul (2010) statistic as obtained by Henze and Koch (2020). Also, it shows a considerably high power performance when compared with the Henze and Jimenez-Gamero (2019) statistic at large sample sizes. Since both the proposed and the Henze and Jimenez-Gamero (2019) statistics are different multivariate extensions of the Zghoul (2010) statistic for assessing univariate normality of datasets and since the Henze and Jimenez-Gamero (2019) statistic has been proved to be powerful affine-invariant consistent test for MVN, it is appropriate to compare only the two. However, the Shapiro-Wilks test due to Villasenor Alva and Estrada (2009) was included in the comparison since it is another powerful statistic for assessing MVN which is not in the BHEP class of tests. With the relative ease of computation of the proposed statistic and its considerable high power performance especially at large sample sizes, it can be recommended as a good statistic for testing the MVN of multivariate datasets. This is further supported with a real-life investigation which gave rise to the $T_{n,\beta}$ statistic being a rather more powerful technique than the $HJ - G_{n,\beta}$ statistic. However, the overall behaviour of the proposed statistic with respect to optimal power performances under different smoothing parameters is not considered in this paper. Its determination will, no doubt, further give credence to its applicability.

Acknowledgements

The authors wish to thank Prof. A. A. Zghoul of the Department of Mathematics, the University of Jordan, for his helpful input during the preparation of the manuscript.

References

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

Batsidis, A., Martin, N., Pardo, L., and Zografos, K. (2013). A necessary power divergence type family tests of multivariate normality. *Communications in Statistics - Simulation and Computation*, **42**(10), 2253–2271.

Cardoso de Oliveira, I. R. C., and Ferreira, D. F. (2010). Multivariate extension of chi-squared univariate normality. *Journal of Statistical Computation and Simulation*, **80**(5), 513–525.

Chen, W., and Genton, M. G. (2022). Are you all normal? It depends! *International Statistical Review*, https://doi.org/10.1111/insr.12512.

Csörgo, S. (1989). Consistency of some tests for multivariate normality. *Metrika*, **36**, 107–116.

Ebner, B., and Henze, N. (2020). Tests for multivariate normality – a critical review with emphasis on weighted L^2-statistics. *Test*, **29**, 847–892.

Ebner, B., Henze, N., and Strieder, D. (2022). Testing normality in any dimension by Fourier methods in a multivariate Stein equation. *Canadian Journal of Statistics*, **50**(3), 992–1033.

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

Henze, N., and Jiménez-Gamero, M. D. (2019). A new class of tests for multinormality with i.i.d. and garch data based on the empirical moment generating function. *Test*, **28**, 499–521.

Henze, N., Jiménez-Gamero, M. D., and Meintanis, S. G. (2019). Characterizations of multinormality and and corresponding tests of fit, including for Garch models. *Economic Theory*, **35**, 510–546.

Henze, N., and Koch, S. (2020). On a test of normality based on the empirical moment generating function. *Statistical Papers*, **61**, 17–29.

Henze, N., and Visagie, J. (2020). Testing for normality in any dimension based on a partial differential equation involving the moment generating function. *Annals of the Institute of Statistical Mathematics*, **72**, 1109–1136.

Henze, N., Wagner, T. (1997). A new approach to the BHEP tests for multivariate normality. *Journal of Multivariate Analysis*, **62**, 1–23.

Henze, N., and Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality. *Communications in Statistics - Theory and Methods*, **19**, 3595–3618.

Joenssen, D. W., and Vogel, J. (2014). A power study of goodness-of-fit tests for multivariate normality implemented in R. *Journal of Statistical Computation and Simulation*, **84**, 1055–1078.

Liang, J., Fang, M. L., and Chang, P. S. (2009). A generalized Shapiro-Wilk W statistic for testing high-dimensional normality. *Computational Statistics and Data Analysis*, **53**, 3883–3891.

Liang, J., He, P., and Yang, J. (2022). Testing multivariate normality based on t-representative points. *Axioms*, **11**(11), 587.

Madukaife, M. S., and Okafor, F. C. (2018). A powerful affine invariant test for multivariate normality based on interpoint distances of principal components. *Communications in Statistics - Simulation and Computation*, **47**, 1264–1275.

Madukaife, M. S., and Okafor, F. C. (2019). A new large sample goodness of fit test for multivariate normality based on chi squared probability plots. *Communications in Statistics - Simulation and Computation*, **48**(6), 1651–1664.

Malkovich, J. F., and Afifi, A. A. (1973). On tests for multivariate normality. *Journal of the American Statistical Association*, **68**, 176–179.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, **573**, 519–530.

Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis for testing normality and robustness studies. *Sankhya*, **36**, 115–128.

Mecklin, C. J., and Mundfrom, D. J. (2004). An appraisal and bibliography of tests for multivariate normality. *International Statistical Review*, **72**, 123–138.

Pudelko, J. (2005). On a new affine invariant and consistent test for multivariate normality. *Probability and Mathematical Statistics*, **25**, 43–54.

Royston, J. P. (1983). Some techniques for assessing multivariate normality based on Shapiro-Wilk W. *Applied Statistics*, **32**(2), 121–133.

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, **52**(3 and 4), 591–611.

Srivastava, M. S. (1984). A measure of skewness and kurtosis and a graphical method for assessing multivariate normality. *Statistics and Probability Letters*, **2**, 263–267.

Székely, G. J., and Rizzo, M. L. (2005). A new test for multivariate normality. *Journal of Multivariate Analysis*, **93**, 58–80.

Tenreiro, C. (2017). A new test for multivariate normality by combining extreme and nonextreme BHEP tests. *Communications in Statistics - Theory and Methods*, **46**, 1746–1759.

Thode, H. C. (2002). *Testing for Normality*. New York: Marcel Dekker.

Villasenor Alva, J. A., and Estrada, E. G. (2009). A generalization of Shapiro-Wilk's test for multivariate normality. *Communications in Statistics - Theory and Methods*, **38**(11), 1870–1883.

Wang, S., Liang, J., Zhou, M., and Ye, H. (2022). Testing multivariate normality based on F-representative points. *Mathematics*, **10**(22), 4300.

Zghoul, A. A. (2010). A goodness-of-fit test for normality based on the empirical moment generating function. *Communications in Statistics - Simulation and Computation*, **39**, 1292–1304.

Zhou, M., and Shao, Y. (2014). A powerful test for multivariate normality. *Journal of Applied Statistics*, **41**(2), 351–363.