Review Article

BHEP Class of Tests for Multivariate Normality: An Empirical Comparison

ABSTRACT

This paper compares the empirical power performances of eight tests for multivariate normality classified under BHEP class of tests. The tests are compared under eight different alternative distributions. The result shows that the eight statistics have good control over type-I-error. Also, some tests are more sensitive to distributional differences with respect to their power performances than others. Also, some tests are generally more powerful than others. The generally most powerful ones are therefore recommended.

Keywords: BHEP class, empirical characteristic function, empirical moment generating function, multivariate normality, power performance

Mathematics Subject Classification: 62E10, 62H25

1. INTRODUCTION

Consider a *d*-variate non-degenerate random vector, $\mathbf{x}$ which is defined by a distribution function $F(\mathbf{x})$ and probability density function $f(\mathbf{x})$, where d is a positive integer and let $F_0(\mathbf{x})$ be a completely specified distribution function on a *d*-dimensional Euclidean space, R^d of a multivariate normal population having mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, $N_d(\boldsymbol{\mu},\boldsymbol{\Sigma})$, with probability density function $f_0(\mathbf{x})$. Suppose a sample of n independent and identically distributed (i.i.d) observation vectors, $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n$ are available from the unknown continuous distribution function $F(\mathbf{x})$, it has been clearly stated in the literature that the problem of assessing multivariate normality (MVN) of the random vector $\mathbf{x}$ on the basis of the i.i.d. random observation vectors, (Fan [1]; Madukaife [2]), is that of testing the goodness-of-fit hypothesis:

$$H_0: F(\mathbf{x}) = F_0(\mathbf{x}) \text{ against } H_1: F(\mathbf{x}) \neq F_0(\mathbf{x}) \qquad (1)$$

Development of appropriate statistics for the test in (1) is an ongoing interest, arguably due to the importance of the multivariate normal distribution in classical statistical analysis. It is well known that most statistical techniques employed in multivariate analysis, and in fact generally in high dimensional data analysis, depend on MVN. Also, a number of datasets from other multivariate distributions, especially at large sample sizes, can be approximated by the multivariate normal distribution. The appropriate statistics (tests) for MVN are usually derived from different unique characterizations of the distribution. Such tests include those from the generating functions (Baringhaus & Henze [3]; Henze & Zirkler [4]; Pudelko [5]; Henze & Jimenez-Gamero [6]), quantile function (Madukaife & Okafor [7, 8]), distribution function (Kim & Park [9]), skewness and kurtosis (Mardia [10, 11]; Doornik & Hansen [12]; Enomoto, Hanusz, Hara & Seo [13]), spherical harmonics function (Manzott & Quiroz [14]) to mention but a few. In fact, Mecklin and Mundfrom [15] have stated that there are more than 50 different statistics for assessing MVN of *d*-dimensional datasets, $d \geq 1$. Since after their review article on this subject, dozens of yet new statistics have been developed, leaving the literature with well over 100 tests for MVN.

These scores of tests so far developed for assessing MVN of datasets do not have equal power performance. Power performance of a test, which is the ability of the test to reject null hypothesis of multivariate normality when the dataset is actually from a non-normal distribution, is a key performance indicator of the test and tests with comparatively high power performance are usually preferred to others.

Again, Chen and Genton [16] have stated that apart from power performance of a test, every good statistics should have desirable properties of affine invariance and consistency. A statistic is said to be affine invariant if it is closed with respect to full rank affine transformation. Algebraically, a statistic $T_n(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ is said to be affine invariant if for any non-zero vector of constants $\mathbf{b} \in R^d$ and any non-singular matrix of constants $A \in R^{dxd}$, $T_n(A\mathbf{x}_1 \pm \mathbf{b}, A\mathbf{x}_2 \pm \mathbf{b}, \ldots, A\mathbf{x}_n \pm \mathbf{b})$ = $T_n(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$. This property implies that under the null hypothesis of MVN, the statistic $T_n(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ does not depend on the parameters $\boldsymbol{\mu}$ and Σ of the underlying multivariate normal distribution. As a result, the power performance of any statistic which attains this property remains the same at all parameter values of $\boldsymbol{\mu}$ and Σ. On the other hand, a statistic is said to be consistent against all fixed alternatives if $\lim_{n \to \infty} \Pr\left(\text{rejecting } H_0 \middle| F(\mathbf{x}) \neq F_0(\mathbf{x})\right) = 1$ (Szekely & Rizzo [17]).

One class of tests for MVN which attains the properties of affine invariance and universal consistency with considerably high power performance is the Baringhaus-Henze-Epps-Pulley (BHEP) class. In fact, Ebner and Henze [18] have stated that BHEP class of tests is the most thoroughly studied class of tests for MVN. Epps and Pulley [19] obtained a statistic for assessing univariate normality of univariate datasets. The statistic is based on the integral of the weighted difference between theoretical and empirical characteristic functions. The statistic is of the form:

$$T = \int_{-\infty}^{\infty} \left| \phi_n(t) - \hat{\phi}_0(t) \right|^2 g(t)dt \qquad (2)$$

where $\phi_n(t)$ is the empirical characteristic function and $\hat{\phi}_0(t) = \exp\left(it\bar{X} - \frac{1}{2}t^2 s^2\right)$; $\bar{X} = n^{-1}\sum_{j=1}^{n} X_j$; $s^2 = n^{-1}\sum_{j=1}^{n}(X_j - \bar{X})^2$ and $g(t)$ is an appropriate weight function which is symmetric about the origin. Through straightforward integration, they obtained the computational form of (2) as a function of $\alpha > 0$, given by:

$$T(\alpha) = n^{-2}\sum_{j=1}^{n}\sum_{k=1}^{n}\exp\left(-\frac{(X_j - X_k)^2}{2\alpha^2 s^2}\right) - 2n^{-1}(1+\alpha^{-2})^{-1/2}\sum_{j=1}^{n}\exp\left(-\frac{(X_j - \bar{X})^2}{2s^2(1+\alpha^2)}\right)$$

$$+ (1+2\alpha^{-2})^{-1/2}$$

and obtained the appropriate statistic for testing the univariate normality of a dataset as $T^*(\alpha) = -\log\{nT(\alpha)\}$. The test, which rejects normality for small values of $T^*(\alpha)$, was considered to have a high power performance. Baringhaus and Henze [3] generalized the statistic to the case of *d*-dimensional dataset, $d \geq 1$. Csorgo [20] coined tests for MVN belonging to this class as the BHEP class and proved their consistency. Several other works have emanated from this pioneer works either by the use the characteristic function, moment generating function or their functions and they are all regarded as members of this class.

Tests for MVN generally are developed independently with the use of several characterizations and as a result do not have the same power performance. In fact, some tests in the literature are known to have low power performance while others are known to have relatively moderate and high power performances. The importance of this has led many researchers to devote attention to comparing different tests for MVN. Such works include Henze [21], Thode [22], Mecklin and Mundfrom [15], Joenssen and Vogel [23] Chen and Genton [16] as well as Ebner and Henze [18]. None of such comparative works however has discussed only tests of the same class. Since it has been established that BHEP tests show interesting properties, the aim of this work is to compare their power performances at different sample sizes and different dimensions of the dataset under different alternative distributions. The tests that are considered in the study are described in section 2 while the simulation study for the power performances is presented in section 3. The paper is concluded in section 4.

2. DESCRIPTION OF THE BHEP TESTS FOR MULTIVARIATE NORMALITY

In this section, a vivid description of each of the tests belonging to the BHEP class of tests for the multinormality is presented.

Baringhaus and Henze test: Baringhaus and Henze [3] generalized the T statistic of Epps and Pulley [19] which is given in (2) to the *d*-dimensional case, where $d \geq 1$. The statistic is the integral of a weighted squared difference between the empirical characteristic function of a *d*-dimensional dataset of size *n* and the characteristic function of a *d*-dimensional standard multivariate normal distribution. It is given as:

$$T_n = n\int_{R^d}\left|\psi_n(\mathbf{t}) - \exp\left(-\tfrac{1}{2}|\mathbf{t}|^2\right)\right|^2 \varphi(\mathbf{t})d\mathbf{t} \qquad (3)$$

where $\psi_n(\mathbf{t})$ is the empirical characteristic function of the scaled observation vectors, $\mathbf{y}_j = S_n^{-1/2}(\mathbf{x}_j - \bar{X}_n);\ j = 1, 2, \ldots, n$, which is defined by $\psi_n(\mathbf{t}) = \sum_{j=1}^{n}\exp\left(i\mathbf{t}^T\mathbf{y}_j\right)$,

$\bar{X}_n = n^{-1}\sum_{j=1}^{n}\mathbf{x}_j$, $\quad S_n = n^{-1}\sum_{j=1}^{n}(\mathbf{x}_j - \bar{X}_n)(\mathbf{x}_j - \bar{X}_n)^T$ and

$\varphi(\mathbf{t}) = (2\pi)^{-d/2} \exp\left(-\frac{1}{2}|\mathbf{t}|^2\right)$ is an appropriate weight function. Through straightforward integration, they obtained the computational form of (3) as an appropriate statistic for assessing MVN of datasets. It is given as:

$$T_n = \frac{1}{n}\sum_{j,k=1}^{n} \exp\left(-\tfrac{1}{2} R_{jk}\right) - 2^{1-d/2}\sum_{j=1}^{n} \exp\left(-\tfrac{1}{4} R_j^2\right) + n3^{-d/2}$$

where $R_{jk} = (\mathbf{x}_j - \mathbf{x}_k)^T S_n^{-1} (\mathbf{x}_j - \mathbf{x}_k)$, $R_j^2 = (\mathbf{x}_j - \overline{\mathbf{x}}_n)^T S_n^{-1} (\mathbf{x}_j - \overline{\mathbf{x}}_n)$, $\mathbf{x}_j$ is the *j*th vector of observations in the dataset, $\overline{\mathbf{x}}_n$ is the sample mean vector and S_n is the sample covariance matrix. The test rejects MVN of datasets for large values of T_n.

Henze and Zirkler test: In a quest to obtain a test for MVN with an improved power performance, Henze and Zirkler [4] introduced a smoothing parameter, β, in the weight function of the Baringhaus and Henze [3] statistic to obtain the statistic:

$$T_{n,\beta} = n\left(4I\left\{S_n \text{ is singular}\right\} + D_{n,\beta} I\left\{S_n \text{ is nonsingular}\right\}\right) \quad (4)$$

where $D_{n,\beta} = \int_{R^d} \left|\psi_n(\mathbf{t}) - \exp\left\{-\tfrac{1}{2}\|\mathbf{t}\|^2\right\}\right|^2 \varphi_\beta(\mathbf{t})\, d\mathbf{t}$;

$\varphi_\beta(\mathbf{t}) = \left(2\pi\beta^2\right)^{-d/2} \exp\left\{-\frac{\|\mathbf{t}\|^2}{2\beta^2}\right\}$. In this test, $\beta = \beta_d(n) = \frac{1}{\sqrt{2}}\left(\frac{2d+1}{4}\right)^{\frac{1}{d+4}} n^{\frac{1}{d+4}}$ and straight forward integration gives $D_{n,\beta}$ as:

$$D_{n,\beta} =$$

$$\left(1+2\beta^2\right)^{-d/2} + \frac{1}{n^2}\sum_{j,k=1}^{n} \exp\left\{-\frac{\beta^2}{2}\left\|\mathbf{y}_j - \mathbf{y}_k\right\|^2\right\} - 2\left(1+\beta^2\right)^{-d/2} \frac{1}{n}\sum_{j=1}^{n} \exp\left\{-\frac{\beta^2\left\|\mathbf{y}_j\right\|^2}{2\left(1+\beta^2\right)}\right\}$$

The test rejects MVN of datasets for large values of the statistic.

Henze and Wagner test: Henze and Wagner [24] obtained a new approach to the class of BHEP tests as given by Henze and Zirkler [4]. By utilizing the theory of weak convergence in the Frechet space $C\left(R^d\right)$ of continuous functions on R^d, they obtained, among other things, a new representation given by:

$$T_{n,\beta} = \int_{R^d} Z_n^2(\mathbf{t})\varphi_\beta(\mathbf{t})\, d\mathbf{t} \quad (5)$$

where $Z_n(\mathbf{t}) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left[\cos\left(\mathbf{t}^T\mathbf{y}_j\right) + \sin\left(\mathbf{t}^T\mathbf{y}_j\right) - \exp\left\{-\tfrac{1}{2}\|\mathbf{t}\|^2\right\}\right]; \mathbf{t} \in R^d$, provided $n \geq d+1$.

Henze and Jimenez-Gamero test: Apart from the original BHEP statistics which employ the empirical and theoretical characteristic functions, Zgoul [25] obtained a test for univariate normality by replacing the characteristic functions in the Epps and Pulley [19] with their corresponding moment generating functions. Henze and Jimenez-Gamero [6] extended the statistic to the *d*-dimensional multivariate case. They obtained the statistic as:

$$T_{n,\beta} = n\int_{R^d} \left(M_n(\mathbf{t}) - m(\mathbf{t})\right)^2 \omega_\beta(\mathbf{t})d\mathbf{t} \tag{6}$$

where $M_n(\mathbf{t}); \mathbf{t} \in R^d$ is the empirical moment generating function of scaled vectors of observations, $\mathbf{y}_j;\ j = 1, 2, \ldots, n$, defined by $M_n(\mathbf{t}) = n^{-1}\sum_{j=1}^{n} \exp\left(\mathbf{t}^T \mathbf{y}_j\right)$, $m(\mathbf{t}) = \exp\left(\tfrac{1}{2}\mathbf{t}^T\mathbf{t}\right)$ and $\omega_\beta(\mathbf{t}) = \exp\left(-\beta\left\|\mathbf{t}\right\|^2\right); \beta > 1$. Straight forward integral gives the computational form of the statistics as:

$$T_{n,\beta} =$$

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

$$\beta > 1$$

The affine invariant and consistent statistic rejects null distribution of MVN for large values of the $T_{n,\beta}$.

Henze, Jimenez- Gamero and Meintanis test: Based on certain characterization of the multivariate normal distribution concerning the product of the cosine transform and the moment generating function of the distribution, Henze, Jimenez-Gamero and Meintanis [26] obtained a test for MVN of datasets. The statistic is given by:

$$T_{n,\gamma} = \int_{R^d} U_n^2(\mathbf{t})\omega_\gamma(\mathbf{t})d\mathbf{t} \tag{7}$$

where $U_n(\mathbf{t}) = \sqrt{n}\left(R_n(\mathbf{t})M_n(\mathbf{t}) - 1\right); \mathbf{t} \in R^d$ and $R_n(\mathbf{t}) = n^{-1}\sum_{j=1}^{n}\cos(\mathbf{t}^T\mathbf{y}_j)$ is the empirical cosine transform of scaled vectors of observations and $M_n(\mathbf{t}) = n^{-1}\sum_{j=1}^{n}\exp(\mathbf{t}^T\mathbf{y}_j)$ is the empirical moment generating function of scaled vectors of observations, $\mathbf{y}_j;\ j = 1, 2, \ldots, n$. The test rejects null hypothesis of MVN for large values of the statistic. They gave the computational form of the statistic as:

$$T_{n,\gamma} = \left(\frac{\pi}{\gamma}\right)^{d/2} \left\{ \frac{1}{2n^3} \sum_{j,k,l,m=1}^{n} \left[\exp\left(\frac{\left\|\mathbf{y}_{j,k}^{+}\right\|^2 - \left\|\mathbf{y}_{l,m}^{-}\right\|^2}{4\gamma} \right) \cos\left(\frac{\mathbf{y}_{j,k}^{+\,T} \mathbf{y}_{l,m}^{-}}{2\gamma} \right) \right.\right.$$

$$\left.\left. + \exp\left(\frac{\left\|\mathbf{y}_{j,k}^{+}\right\|^2 - \left\|\mathbf{y}_{l,m}^{+}\right\|^2}{4\gamma} \right) \cos\left(\frac{\mathbf{y}_{j,k}^{+\,T} \mathbf{y}_{l,m}^{+}}{2\gamma} \right) \right] - \frac{2}{n} \sum_{j,k=1}^{n} \left[\exp\left(\frac{\left\|\mathbf{y}_{j}\right\|^2 - \left\|\mathbf{y}_{k}\right\|^2}{4\gamma} \right) \cos\left(\frac{\mathbf{y}_{j}^{T} \mathbf{y}_{k}}{2\gamma} \right) \right] \right\};$$

$$\gamma > 1$$

where $\mathbf{y}_{j,k}^{\pm} = \mathbf{y}_j \pm \mathbf{y}_k$.

Henze and Visagie test: Henze and Visagie [27] obtained the unique solution of the partial differential equation of a d-dimensional random vector, $\mathbf{x}$, to be $m(\mathbf{t}) = \exp\left\{\frac{\|\mathbf{t}\|^2}{2}\right\}; \mathbf{t} \in R^d$ which is the moment generating function of the *d*-dimensional standard multivariate normal distribution. Based on this unique property, they obtained a statistic for assessing MVN of datasets with a statistic given by:

$$T_{n,\gamma} = n\int_{R^d} \left\| M_n{}'(\mathbf{t}) - \mathbf{t}M_n(\mathbf{t}) \right\|^2 \omega_\gamma(\mathbf{t})\, d\mathbf{t} \qquad (8)$$

where $\omega_\gamma(\mathbf{t}) = \exp\left\{-\gamma\|\mathbf{t}\|^2\right\}$ is the weight function and $M_n(\mathbf{t}) = n^{-1}\sum_{j=1}^{n} \exp\left\{\mathbf{t}^T \mathbf{y}_j\right\}$ is the empirical moment generating function of the scaled vectors of observations. They stated that rejection of the null hypothesis is for large values of the statistic. Putting $\mathbf{y}_{j,k}^{\pm} = \mathbf{y}_j \pm \mathbf{y}_k$, they obtained the computational form of the statistic as:

$$T_{n,\gamma} = \frac{1}{n}\left(\frac{\pi}{\gamma}\right)^{d/2} \sum_{j,k=1}^{n} \exp\left(\frac{\left\|\mathbf{y}_{j,k}^{+}\right\|^2}{4\gamma} \right) \left(\mathbf{y}_j^{T}\mathbf{y}_k - \frac{\left\|\mathbf{y}_{j,k}^{+}\right\|^2}{2\gamma} + \frac{d}{2\gamma} + \frac{\left\|\mathbf{y}_{j,k}^{+}\right\|^2}{4\gamma^2} \right); \gamma > 0$$

Dorr, Ebner and Henze tests: Based on a unique characterization of the standard *d*-dimensional multivariate normal distribution as the unique solution of an initial value problem of a partial differential equation by the harmonic operator, Dorr, Ebner and Henze [28] derived a new test for MVN. The statistic is given by:

$$T_{n,a} = n\int_{R^d} \left| \Delta\psi_n(\mathbf{t}) - \Delta\psi(\mathbf{t}) \right|^2 \omega_a(\mathbf{t})\, d\mathbf{t} \qquad (9)$$

$$= n\int_{R^d} \left| \frac{1}{n}\sum_{j=1}^{n} \left\|\mathbf{y}_j\right\|^2 \exp\left\{ i\mathbf{t}^T\mathbf{y}_j \right\} + \left(\|\mathbf{t}\|^2 - d \right) \exp\left\{ -\frac{\|\mathbf{t}\|^2}{2} \right\} \right|^2 \omega_a(\mathbf{t})\, d\mathbf{t}$$

With $\omega_a(\mathbf{t}) = \exp\left\{-a\|\mathbf{t}\|^2\right\}; a > 0, \mathbf{t} \in R^d$, they obtained the computationally amenable form of the statistic as:

$$T_{n,a} = \left(\frac{\pi}{a}\right)^{d/2} \frac{1}{n} \sum_{j,k=1}^{n} \|\mathbf{y}_j\|^2 \|\mathbf{y}_k\|^2 \exp\left\{-(4a)^{-1}\|\mathbf{y}_j - \mathbf{y}_k\|^2\right\}$$

$$- \frac{2(2\pi)^{d/2}}{(2a+1)^{2+d/2}} \sum_{j=1}^{n} \|\mathbf{y}_j\|^2 \left(\|\mathbf{y}_j\|^2 + 2ad(2a+1)\right) \exp\left(\frac{\|\mathbf{y}_j\|^2}{2(2a+1)}\right)$$

$$+ \frac{n\pi^{d/2}}{(a+1)^{2+d/2}} \left(ad^2(a+1) + \frac{d(d+2)}{4}\right)$$

The affine invariant and consistent test rejects the H_0 of MVN for large values of $T_{n,a}$.

Also, Dorr et al [29] established as a theorem that the characteristic function of a *d*-variate normal distribution is the only solution of the partial differential equation $\Delta f(\mathbf{t}) = \left(\|\mathbf{t}\|^2 - d\right) f(\mathbf{t}), \mathbf{t} \in R^d$, subject to the condition $f(0) = 1$. Based on this characterization, Dorr et al [29] obtained a statistic for assessing MVN of datasets. The statistic is given by:

$$U_{n,a} = n\int_{R^d} \left|\Delta \psi_n(\mathbf{t}) - \left(\|\mathbf{t}\|^2 - d\right)\psi_n(\mathbf{t})\right|^2 \omega_a(\mathbf{t})\, d\mathbf{t} \qquad (10)$$

They stated that large values of the statistic will lead to rejection of the null hypothesis of MVM and gave a straightforward integration (10) to obtain a computationally amenable form as $U_{n,a}$

$$= \left(\frac{\pi}{a}\right)^{d/2} \frac{1}{n} \text{x}$$

$$\sum_{j,k=1}^{n} \exp\left\{-\frac{\|\mathbf{y}_j - \mathbf{y}_k\|^2}{4a}\right\} \Bigg[\|\mathbf{y}_j\|^2 \|\mathbf{y}_k\|^2 - \left(\|\mathbf{y}_j\|^2 + \|\mathbf{y}_k\|^2\right) \frac{1}{4a^2} \left(\|\mathbf{y}_j - \mathbf{y}_k\|^2 + 2ad(2a-1)\right)$$

$$+ \frac{1}{16a^4} \left(16a^3 d^2 (a-1) + 4a^2 d(d+2) + \|\mathbf{y}_j - \mathbf{y}_k\|^4 + \left(8a^2 d - 4a(d+2)\right)\|\mathbf{y}_j - \mathbf{y}_k\|^2\right) \Bigg];$$

$$a > 0$$

.

The Ebner, Henze and Strieder test: Ebner, Henze and Strieder [30] proved a theorem that the characteristic function, $\psi(\mathbf{t})$, of the *d*-variate standard normal distribution is the only characteristic function satisfying $\nabla \psi(\mathbf{t}) = -\mathbf{t}\psi(\mathbf{t})$, ∇ is the gradient operator. Based on this characterization, they obtained a test for MVN of datasets. The statistic is given by:

$$T_{n,a} = n\int_{R^d} \left\| \Delta\psi_n(\mathbf{t}) - \Delta\psi(\mathbf{t}) \right\|_C^2 \omega_a(\mathbf{t})\, d\mathbf{t} \tag{11}$$

where $\omega_a(\mathbf{t}) = \exp\left\{-a\|t\|^2\right\};\ a > 0$ and $\|.\|_C$ is the complex Euclidean vector norm. They stated that rejection of the null hypothesis of MVN is for large values of the statistic and went further obtain the computational form of the statistic as:

$$T_{n,a} = n\left(\frac{\pi}{a+1}\right)^{d/2} \frac{d}{2(a+1)} - 2\left(\frac{2\pi}{2a+1}\right)^{d/2} \sum_{j=1}^{n} \frac{\|\mathbf{y}_j\|^2}{2a+1} \exp\left\{-\frac{\|\mathbf{y}_j\|^2}{2(2a+1)}\right\}$$

$$+ \frac{1}{n}\left(\frac{\pi}{a}\right)^{d/2} \sum_{j,k=1}^{n} \mathbf{y}_j{}^T \mathbf{y}_k \exp\left\{-\frac{\|\mathbf{y}_j - \mathbf{y}_k\|^2}{4a}\right\}$$

3. SIMULATION STUDIES

In this section, empirical power performances of the eight statistics in the class are compared through extensive simulation studies. In this comparison, four different classes of distributions are employed. They include the null distribution, which is presented as the standard multivariate normal distribution; *d*-variate distributions with symmetric marginals ($t(2)^d$, $beta(0.5, 0.5)^d$ and $t(2)^q \otimes beta(0.5, 0.5)^{d-q}$); *d*-variate distributions with both symmetric and non-symmetric marginals ($t(2)^q \otimes gamma(5,1)^{d-q}$ and $beta(2, 2)^q \otimes \chi^2(10)^{d-q}$) and *d*-variate distributions with non-symmetric marginals ($\chi^2(10)^q \otimes beta(3, 2)^{d-q}$ and $gamma(5,1)^q \otimes Weibull(2,1)^{d-q}$) where $q < d$. A total of 10,000 data sets were generated via Monte Carlo simulation in each combination of n = 25 and 50 and d = 2 and 5 from the resultant eight different multivariate distributions.

The values of each of the eight statistics being compared in this study are evaluated in each of the 10,000 simulated samples for each n and d from each of the distributions. The statistics are denoted by HZ for Henze and Zirkler test, HW for Henze and Wagner test, HJG for Henze and Jimenez-Gamero test, HJM for the Henze, Jimenez-Gamero and Meintanis test, HV for Henze and Visagie test, DEHT for the Dorr, Ebner and Henze *T* test, DEHU for the Dorr, Ebner and Henze *U* test while EHS denotes the Ebner, Henze and Strieder test. The empirical power of each test statistic is obtained as the percentage of the 10,000 samples that is rejected by the statistic at α of 5%. The power performance is presented in Tables 1 and 2 respectively for n = 25 and 50. All the simulations and computations are carried out with the use of the R – statistical package, mnt. Preliminary studies carried out on these statistics in this class show that most of them maintain maximum power performances in all the distributions considered for the smoothing (tuning) parameter $\beta \in (1, 10]$. As a result, the power of each statistic was obtained as the maximum value within the interval of the smoothing parameter.

From the powers in Tables 1 and 2, it can be seen that all the statistics have very good control over type-I-error. This is because the power performances of the statistics under the null distribution of standard multivariate normal distribution in all the sample sizes and variable dimensions are approximately equal to 5%. Also as expected, all the statistics show improvement in their powers with increasing sample size from 25 to 50. This however is with the exception of the HJG, HV, DEHT and DEHU under the symmetric distribution with Beta(0.5, 0.5) marginal. Under this alternative distribution, these statistics showed almost a zero power performance. Under the symmetric alternative distributions, the HZ and HW tests generally outperformed all the other statistics in the class, except under the symmetric with *t*(2) marginal where the DEHT, DEHU and EHS tests appear to have more power. Under the alternative distributions with mixed symmetric and non-symmetric marginal, the DEHU and EHS statistics clearly show themselves to be most powerful at both n = 25 and n = 50. Under the asymmetric distributions, the HZ, DEHU and EHS recorded the highest power performances. Very importantly, the HJG, HV, DEHT and the DEHU statistics showed themselves to be very sensitive to distributional differences while the HZ, HW and EHS statistics are more robust statistics with respect to distributional differences.

Table 1. Empirical power performance (in %) of BHEP class of tests for MVN, n = 25, alpha = 0.05

d	Distributions	HZ	HW	HJG	HJM	HV	DEHT	DEHU	EHS
2	$N_d(0, 1)$	5.2	4.6	4.7	4.5	5.4	5.0	5.4	4.7
5	$N_d(0, 1)$	4.8	4.9	5.0	4.6	5.1	5.1	5.0	4.9
2	$t(2)^2$	73.2	71.1	76.4	70.3	77.2	81.7	80.9	77.9
5	$t(2)^5$	82.3	75.2	88.6	80.9	89.6	94.9	94.9	93.2
2	$\text{beta}(0.5, 0.5)^2$	78.1	81.3	0.1	47.7	0.1	0.2	0.7	1.8
5	$\text{beta}(0.5, 0.5)^5$	45.3	46.9	0.1	40.8	0.2	0.0	0.1	1.8
2	$t(2) \otimes \text{beta}(0.5, 0.5)$	73.5	74.8	50.4	69.3	50.6	50.1	51.9	59.9
5	$t(2)^3 \otimes \text{beta}(0.5, 0.5)^2$	70.9	65.6	70.2	70.1	72.2	73.6	75.1	75.5
2	$\text{Chisquare}(10) \otimes \text{beta}(3,2)$	17.1	17.1	12.5	16.9	12.5	16.2	16.9	19.3
5	$\text{Chisquare}(10)^3 \otimes \text{beta}(3,2)^2$	15.6	10.5	14.5	14.4	14.3	13.2	15.7	18.6
2	$\text{Gamma}(5, 1) \otimes \text{Weibull}(2,1)$	22.8	20.5	18.6	19.3	17.3	23.6	27.3	26.7
5	$\text{Gamma}(5, 1) \otimes \text{Weibull}(2,1)$	18.8	13.7	18.3	19.0	17.4	20.0	22.8	24.5
2	$t(2) \otimes \text{gamma}(5,1)$	55.3	52.6	55.9	52.3	57.5	62.6	63.6	64.7
5	$t(2)^3 \otimes \text{gamma}(5,1)^2$	63.6	53.3	75.5	70.1	75.6	83.8	83.2	82.6
2	$\text{beta}(2, 2) \otimes \text{Chisquare}(5)$	29.7	26.8	20.7	20.3	18.9	24.7	27.2	30.4
5	$\text{beta}(2,2)^3 \otimes \text{Chisquare}(10)$	17.2	13.7	16.7	17.0	16.4	13.1	15.6	17.9

Table 2. Empirical power performance (in %) of BHEP class of tests for MVN, *n* = 50, alpha = 0.05

d	Distributions	HZ	HW	HJG	HJM	HV	DEHT	DEHU	EHS
2	$N_d(0, 1)$	4.8	5.0	5.1	4.9	4.7	4.7	5.1	5.2
5	$N_d(0, 1)$	4.8	5.0	5.0	4.8	4.8	5.0	5.0	4.9
2	$t(2)^2$	95.1	96.6	79.8	98.8	89.4	97.7	97.4	96.4
5	$t(2)^5$	99.1	98.6	99.1	99.9	99.2	100.0	100.0	99.7
2	beta$(0.5, 0.5)^2$	99.6	99.6	0.0	62.8	0.0	47.0	66.9	35.1
5	beta$(0.5, 0.5)^5$	97.8	96.5	0.0	70.3	0.0	0.0	3.7	5.6
2	$t(2) \otimes$ beta(0.5, 0.5)	98.3	98.1	74.6	95.8	75.4	78.1	80.7	89.1
5	$t(2)^3 \otimes$ beta$(0.5, 0.5)^2$	98.2	97.2	93.0	96.5	93.1	96.0	97.1	98.0
2	Chisquare(10) $\otimes$ beta(3,2)	36.4	35.1	20.9	37.8	19.1	35.4	38.6	44.1
5	Chisquare$(10)^3 \otimes$ beta$(3,2)^2$	31.9	24.9	22.9	35.5	24.4	32.0	36.7	48.3
2	Gamma(5, 1) $\otimes$ Weibull(2,1)	44.0	40.7	27.6	59.2	25.1	52.8	55.0	56.9
5	Gamma(5, 1) $\otimes$ Weibull(2,1)	38.0	28.1	28.0	60.8	28.2	46.9	52.1	64.0
2	$t(2) \otimes$ gamma(5,1)	99.3	83.0	80.8	92.8	79.6	89.5	89.8	90.0
5	$t(2)^3 \otimes$ gamma$(5,1)^2$	94.6	91.8	95.3	98.8	95.1	99.1	99.1	98.8
2	beta(2, 2) $\otimes$ Chisquare(5)	58.8	57.2	31.6	61.3	29.8	56.8	63.9	69.3
5	beta$(2,2)^3 \otimes$ Chisquare(10)	40.9	33.3	29.4	40.2	29.1	32.6	35.7	53.7

4. CONCLUSION

In this study, eight different statistics in the BHEP class of tests for assessing MVN of multivariate datasets are compared via empirical power performances. The results show that the statistics are good tools for testing for MVN. Both at small sample sizes and large sample sizes as well as low and high variable dimensions, the HZ, HW, DEHU and EHS statistics distinguished themselves with generally superior power performances in relation to the rest of the statistics. As a result, they are recommended for testing the MVN of datasets.

REFERENCES

[1] Fan Y. Goodness of fit tests for a multivariate distribution by the empirical characteristic function. Journal of Multivariate Analysis 1997; 62: 36-63.
[2] Madukaife MS. A new affine invariant test for multivariate normality based on beta probability plots, Journal of the Nigerian Statistical Association. 2017; 29:58-70.
[3] Baringhaus L, Henze N. A consistent test for multivariate normality based on the empirical characteristic function. Metrika. 1988; 35:339–348.
[4] Henze N, Zirkler B. A class of invariant consistent tests for multivariate normality. Communications in Statistics-Theory and Methods. 1990; 19:3595 – 3618.
[5] Pudelko J. On a new affine invariant and consistent test for multivariate normality. Probability and Mathematical Statistics. 2005; 25:43–54.
[6] Henze N, Jiménez-Gamero MD. A new class of tests for multinormality with i.i.d. and garch data based on the empirical moment generating function. Test. 2019; 28:499 – 521.

[7] Madukaife MS, Okafor FC. A powerful affine invariant test for multivariate normality based on interpoint distances of principal components. Communications in Statistics - Simulation and Computation. 2018; 47:1264 –1275.
[8] Madukaife MS, Okafor FC. A new large sample goodness of fit test for multivariate normality based on chi squared probability plots, Communications in Statistics - Simulation and Computation. 2019; 48(6): 1651-1664.
[9] Kim I, Park S. Likelihood ratio tests for multivariate normality. Communications in Statistics-Theory and Methods. 2018; 47(8): 1923–1934.
[10] Mardia KV. Measures of multivariate skewness and kurtosis with applications. Biometrika. 1970; 573:519-530.
[11] Mardia KV. Applications of some measures of multivariate skewness and kurtosis for testing normality and robustness studies. Sankhya. 1974; 36:115–128.
[12] Doornik JA, Hansen H. An omnibus test for univariate and multivariate normality. Oxford Bulletin of Economics and Statistics. 2008; 70:927–939.
[13] Enomoto R, Hanusz Z, Hara A, Seo T. Multivariate normality test using normalizing transformation for Mardia's multivariate kurtosis. Communications in Statistics-Simulation and Computation. 2020; 49:684–698.
[14] Manzotti A, Quiroz AJ. Spherical harmonics in quadratic forms for testing multivariate normality. Test. 2001; 10:87–104
[15] Mecklin CJ, Mundfrom DJ. An appraisal and bibliography of tests for multivariate normality. International Statistical Review. 2004; 72:123–138.
[16] Chen W, Genton MG. Are you all normal? It depends! Preprint. 2020; arXiv:2008.10957v1
[17] Székely GJ, Rizzo ML. A new test for multivariate normality. Journal of Multivariate Analysis. 2005; 93:58–80.
[18] Ebner B, Henze N. Tests for multivariate normality – a critical review with emphasis on weighted L2 –statistics. Preprint. 2020; arXiv:2004.07332vl
[19] Epps TW, Pulley LB. A test for normality based on the empirical characteristic function. Biometrika. 1983; 70(3):723–726.
[20] Csörgo S. Consistency of some tests for multivariate normality. Metrika. 1989; 36:107–116.
[21] Henze N. Invariant tests for multivariate normality: a critical review. Statistical Papers. 2002; 43:467–506.
[22] Thode HC. Testing for Normality. New York: Marcel Dekker; 2002.
[23] Joenssen DW, Vogel J. A power study of goodness-of-fit tests for multivariate normality implemented in R. Journal of Statistical Computation and Simulation. 2014; 84:1055–1078.
[24] Henze N, Wagner T. A new approach to the BHEP tests for multivariate normality. Journal of Multivariate Analysis. 1997; 62:1–23.
[25] Zghoul AA. A goodness-of-fit test for normality based on the empirical moment generating function. Communications in Statistics-Simulation and Computation. 2010; 39:1292–1304.
[26] Henze N, Jiménez-Gamero MD, Meintanis SG. Characterizations of multinormality and corresponding tests of fit, including for Garch models. Economic Theory. 2019; 35:510 – 546.
[27] Henze N, Visagie J. Testing for normality in any dimension based on a partial differential equation involving the moment generating function. Annals of the Institute of Statistical Mathematics. 2020; 72:1109–1136.
[28] Dörr Ph, Ebner B, Henze N. Testing multivariate normality by zeros of the harmonic oscillator in characteristic function spaces. Scandinavian Journal of Statistics. 2020; https://doi.org/10.1111/sjos.12477
[29] Dörr Ph, Ebner B, Henze N. A new test of multivariate normality by a double estimation in a characterizing PDE. Metrika. 2020; https://doi.org/10.1007/s00184-020-00795-x

[30] Ebner B, Henze N, Strieder D. Testing normality in any dimension by Fourier methods in a multivariate Stein equation. The Canadian Journal of Statistics. 2021; doi.org/10.1002/cjs.11670.