STATISTICS IN TRANSITION new series, June 2025
Vol. 26, No. 2, pp. xx–XX, https://doi.org/10.59139/stattrans-2025-xxx
Received – xx.xx.xxxx; accepted – xx.xx.xxxx

Testing for multinormality with goodness-of-fit tests based on phi divergence measures

Mbanefo S. Madukaife[1], Uchenna C. Nduka[2], Everestus O. Ossai[3]

Abstract

In this paper, a beta transform of multivariate normal datasets is obtained. The phi divergence measure, $D_{\Phi}(F,G)$ between two distributions F and G is used to obtain a goodness-of-fit test to multivariate normality (MVN) based on the theoretical density function of the beta transformed random variable and a window-size-spacing-based sample density function. Three versions of the statistic are derived from three known phi divergence measures that are based on a sum of squares. The empirical critical values of the statistics are obtained and the empirical type-one-error rates as well as powers of the statistics in comparison with those of other well-known competing statistics are computed through extensive simulation study. The study shows that the new statistics have good control over type-one-error and are highly competitive with the existing well-known ones in terms of power performance. The applicability of the new statistics is also carried out in comparison with three other efficient techniques using four different datasets, and all the competing statistics agreed perfectly in their decisions of rejection or otherwise of the multivariate normality of the datasets. As a result, they can be regarded as appropriate statistics for assessing multinormality of datasets especially, in large samples.

Key words: beta transform of multivariate normal observation, empirical critical value, entropy estimator, phi divergence measure, power of a test.

1. Introduction

The search for a more tractable, highly powerful and generally acceptable goodness-of-fit techniques for assessing the normality of a set of data has continued to receive the attraction of cross-generational researchers in the field of statistical methodology. Since the pioneer work of Pearson (1900), more than ten scores of such techniques at both univariate and multivariate spheres have been introduced in the literature from diverse unique characterizations of the normal distribution. These characterizations range from the distribution functions, generating functions (moment generating function, characteristic function and Laplace transform), skewness and kurtosis and entropy, to mention but a few, to other characterizations of various transformations of the normal distribution.

[1]University of Nigeria, Nsukka, Nigeria. E-mail: mbanefo.madukaife@unn.edu.ng. ORCID: https://orcid.org/0000-0003-2823-4223.
[2]University of Nigeria, Nsukka, Nigeria. E-mail: uchenna.nduka@unn.edu.ng. ORCID: https://orcid.org/0000-0001-5931-2840.
[3]University of Nigeria, Nsukka, Nigeria. E-mail: everestus.ossai@unn.edu.ng. ORCID: https://orcid.org/0000-0001-9742-2389.

Of particular attention are tests for multivariate normality (MVN). This is probably due to the fact that most classical multivariate statistical techniques, with diverse applications in many areas of study such as machine learning, econometrics and genomics, require multivariate normality. Suppose $\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n; \mathrm{x}_j \in R^d$, $j = 1, 2, \ldots, n$ is a sequence of n independent and identically distributed (iid) d-dimensional random vectors from an unknown distribution $F(\mathrm{x})$; where $d \geq 2$ is an integer. The problem of testing for MVN is that of testing the null hypothesis

$$H_0 : F(x) \in F_N \tag{1}$$

against an alternative that $F(x) \notin F_N$; where F_N is a class of nondegenerate d-dimensional multivariate normal distributions with mean vector μ and nondegenerate covariance matrix Σ. Examples of tests devoted to (1) in the literature include Healy (1968), Mardia (1970, 1974), Malkovich and Afifi (1973), Small (1978), Royston (1983), Srivastava (1984), Baringhaus and Henze (1988), Henze and Zirkler (1990), Singh (1993), Romeu and Ozturk (1993), Henze and Wagner (1997), Hwu et al. (2002), Szekeley and Rizzo (2005), Pudelko (2005), Doornik and Hansen (2008), Liang et al. (2009), Cardoso de Oliveira and Ferreira (2010), Hanusz and Tarasinska (2012), Zhou and Shao (2014), Thulin (2014), Korkmaz et al. (2014), Tenreiro (2017), Madukaife and Okafor (2018), Madukaife and Okafor (2019), Henze and Jimenez-Gamero (2019), Henze et al. (2019), Henze and Visagie (2020), Dorr et al. (2020a, 2020b). For extensive reviews on different tests for MVN in their various classes, see Henze (2002), Mecklin and Mundfrom (2004), Ebner and Henze (2020) as well as Chen and Genton (2023).

Some of the developed techniques are direct extension of tests for univariate normality to their multivariate counterparts. For instance, Epps and Pulley (1983) developed a test for univariate normality as an integral of the squared difference between the theoretical and empirical characteristic functions of the univariate normal distribution. They showed that the test was very consistent against all fixed alternatives and affine invariant (invariant with respect to changes in location and scale) with competitive high power performance. Because of its interesting properties, Baringhaus and Henze (1988) developed its multivariate counterpart. Since then, several versions of it have been developed and they are coined Baringhaus-Henze-Epps-Pulley (BHEP) class of tests for multivariate normality by Csorgo (1989). In a like manner, Shapiro and Wilks (1965) obtained an omnibus test for assessing univariate normality of a dataset, $x_1, x_2, \ldots, x_n; x_j \in R$, $j = 1, 2, \ldots, n$, which they defined as a ratio of two variance estimators obtained from the dataset and stated that if the dataset is drawn from a normal distribution, then the two estimators would amount to the same value, thereby approaching 1. With the intension of obtaining a multivariate test that inherits the good power performance of the Shapiro and Wilks (1965) test, Villasenor and Gonzalez-Estrada (2009) extended it to the multivariate sphere and the resultant statistic shows an appreciable good power performance. Recently, Tavakoli et al. (2020) applied the sample entropy measure of Vasicek (1976) to estimate phi divergence measures $D_\Phi(F,G)$ between a normal distribution, F and an unknown distribution, G, from where a random sample $x_1, x_2, \ldots, x_n; x_j \in R$, $j = 1, 2, \ldots, n$ is drawn. They argued that if G is also a normal distribution, then, $D_\Phi(F,G)$ will be a minimum. With this, they introduced consistent and affine invariant tests for univariate normality which are very tractable and have high power performances with good control over type-I-error. Since the search for more tractable tests for MVN with relatively high competitive performances is an open research in the literature, it suffices that extension of the Tavakoli et al. (2020) procedures to their multivariate counterparts with some one-to-one transformations would, no doubt, retain the properties and hence serve as highly competitive tests for MVN. This is the purpose of the present paper. The rest of the paper is presented as follows: the statistics are developed in Section 2, with their properties. Section 3 gives the empirical critical values as well as the empirical size and power comparisons. Section 4 gives some real-life applications of the new statistics in comparison with some other statistics while the paper is concluded in Section 5.

2. The test statistic

Suppose $x_1, x_2, \ldots, x_n; x_j \in R^d$, $j = 1, 2, \ldots, n$, and $d \geq 2$ is a d-dimensional random sample from a continuous distribution F. Healy (1968) obtained the sample Mahalanobis squared distances of the observations, which he defined as squared radii, as

$$y_j = (\mathbf{x}_j - \overline{\mathbf{x}}_n)^T S_n^{-1} (\mathbf{x}_j - \overline{\mathbf{x}}_n); \, j = 1, 2, \ldots, n \tag{2}$$

where $\overline{\mathbf{x}}_n = n^{-1} \sum_{j=1}^n \mathbf{x}_j$ is the sample mean vector and $S_n = (n-1)^{-1} \sum_{j=1}^n (\mathbf{x}_j - \overline{\mathbf{x}}_n)(\mathbf{x}_j - \overline{\mathbf{x}}_n)^T$ is the sample covariance matrix. Under the null distribution of multivariate normality of F, Healy (1968) stated that the squared radii are asymptotically distributed as chi-squared observations with d degrees of freedom. Gnanadesikan and Kettenring (1972) obtained a transform of the squared radii as

$$z_j = \frac{n}{(n-1)^2} (\mathbf{x}_j - \overline{\mathbf{x}}_n)^T S_n^{-1} (\mathbf{x}_j - \overline{\mathbf{x}}_n); \, j = 1, 2, \ldots, n \tag{3}$$

and stated that z_j's are exact independent univariate observations from the beta distribution of the first kind, $B(a,b)$ with parameters $a = d/2$ and $b = (n-d-1)/2$ under the MVN of F, where n and d have their usual meanings. The exactness of this assumption was proved by Bilodeau and Brenner (1999) and it has since been used in goodness-of-fit statistics such as Small (1978), Hanusz and Tarasinska (2012) and Madukaife (2017). It is interesting to note here that the transformations in (2) and (3) are functions of d-dimensional observations, $d > 1$. However, even when the observations emanate from a univariate distribution, $d = 1$, it is natural to still obtain z_j's as beta distributed independent observations, $B(a,b)$, with $a = 1/2$ and $b = (n-2)/2$. As a result, the statistics obtained in this study can also apply to univariate normality testing.

Now, the phi divergence measure between any two distributions F_X and G_X, with density functions $f(x)$ and $g(x)$ respectively, is defined by

$$D_\Phi(F_X, G_X) = \int_{-\infty}^{\infty} \Phi\left(\frac{g(x)}{f(x)}\right) f(x)\, dx \tag{4}$$

where $\Phi(x)$ is a convex function such that $\Phi(1) = 0$ and $\Phi''(1) > 1$. At different times, a number of works have independently obtained different convex functions satisfying the conditions of the $\Phi(x)$ in (4) and these works have led to different phi divergence functions. Some of them include the following, as listed in Tavakoli et al. (2019) and Tavakoli et al. (2020):

Kullback-Leibler divergence measure, $\Phi(x) = xlog(x)$.
Pearson divergence measure, $\Phi(x) = (x-1)^2$;
Hellinger divergence measure, $\Phi(x) = 1/2(\sqrt{x}-1)^2$;
Triangular divergence measure, $\Phi(x) = \frac{(1-x)^2}{1+x}$;
Lin-Wong divergence measure, $\Phi(x) = xlog\left(\frac{2}{1+x}\right)$;
Jeffreys divergence measure, $\Phi(x) = (x-1)log(x)$;
Total variation divergence measure, $\Phi(x) = |x-1|$; and
Balakrishnan-Sanghvi divergence measure, $\Phi(x) = \left(\frac{x-1}{x+1}\right)^2$. For more divergence measures and more details on them, readers are referred to Lin (1991).

Now, using the method of estimating the entropy of a random variable by Vasicek (1976), Tavakoli et al. (2020) obtained $D_\Phi(F,G)$ in (4) when G is normal with mean μ and variance σ^2 and F is unknown as:

$$D_\Phi(F_X,G_X) = \int_0^1 \Phi\left(\frac{(2\pi\sigma^2)^{-1/2}exp\left(-\left(F^{-1}(p)-\mu\right)^2/2\sigma^2\right)}{(dF^{-1}(p)/dp)-1}\right)^2 dp \qquad (5)$$

where $F(x) = p \implies F^{-1}(p) = inf\{x : F(x) = p\}; p \in (0,1)$. Replacing F in (5) with F_n (the empirical distribution function) and using the difference operator in place of differential operator, they obtained an estimator, V_Φ of $D_\Phi(F,G)$ as a generic statistic for testing the normality of a set of n observations. The statistic is given as:

$$V_\Phi = \frac{1}{n}\sum_{j=1}^{n}\Phi\left(\frac{n}{\sqrt{2\pi\hat{\sigma}^2}}exp\left\{-\frac{(X_{(j)}-\hat{\mu})^2}{2\hat{\sigma}^2}\right\}\frac{\left(X_{(j+m)}-X_{(j-m)}\right)}{2m}\right) \qquad (6)$$

where $X_{(j)}$ is the jth order statistic, $j = 1,2,,\ldots,n$, of the random sample, $X_1,X_2,\ldots,X_n$ such that $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$; $\hat{\mu} = \overline{X}$; $\hat{\sigma}^2 = n^{-1}\sum_{j=1}^{n}(X_j-\overline{X})^2$ and m, known as the window size or spacing, is an integer such that $m \leq \frac{n}{2}$. They proved that the statistic is consistent against fixed alternatives and that it is affine invariant. The test rejects the null hypothesis of normality for large values of the statistic and it is said to be generic because it is ammenable to any specific phi function in the class of phi divergence measures.

It is very clear from Vasicek (1976) as well as Tavakoli et al. (2020) that the development of the theory behind (5) and (6) does not depend on the normality of F_X and G_X. As a result, a plug-in method is possible for goodness-of-fit statistics to statistical distributions. Therefore, let $z_1,z_2,\ldots,z_n$ be the beta transforms of the random sample according to (3) and let G_Z be beta distributed with parameters a and b such that

$$g(z) = \frac{1}{B(a,b)}z^{a-1}(1-z)^{b-1}; 0 < z < 1 \qquad (7)$$

Then, replacing the normal density function in (5) with that of the beta in (7), $D_\Phi(F,G)$ can be presented as:

$$D_\Phi(F_Z,G_Z) = \int_0^1 \Phi\left(\frac{B(a,b)^{-1}\left[F^{-1}(p)\right]^{a-1}\left[1-F^{-1}(p)\right]^{b-1}}{\left(dF^{-1}(p)/dp\right)^{-1}}\right)^2 dp \tag{8}$$

This gives rise to a new generic goodness-of-fit statistic obtained similar to (6) by replacing F with F_n and using difference operator in place of differential operator. It is given as:

$$M_{n,\Phi} = \frac{1}{n}\sum_{j=1}^{n}\Phi\left(\frac{n\Gamma(a+b)(Z_{(j)})^{a-1}(1-Z_{(j)})^{b-1}\left(Z_{(j+m)}-Z_{(j-m)}\right)}{2m\Gamma(a)\Gamma(b)}\right) \tag{9}$$

where $a = \frac{d}{2}$; $b = \frac{(n-d-1)}{2}$ and m has its usual meaning. The test rejects the null hypothesis of MVN for large values of the statistic. Also, it is invariant with respect to changes in the scale and location of the observation vectors. This is because the transformations in (2) and (3) are standardized transformations that result in beta distributed observations with constant parameters for each n and d such that no matter the affine transformation in $\mathbf{x}_j$'s, $j = 1, 2, \ldots, n$, z_j's have a specified beta distribution with specified parameters.

Theorem 2.1: Suppose $\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n$; $\mathrm{x}_j \in R^d$, $j = 1,2,\ldots,n$ *is a random sample from an unknown continuous distribution* $F(\mathbf{x})$ *with a probability density function* $f(\mathbf{x})$. *The statistic* $M_{n,\Phi}$ *obtained from the observation vectors is invariant with respect to changes in scale and location of the observation vectors.*

Proof:

Let C be defined as a dxd nonsingular matrix of constants and $\mathbf{u}$ a d-component vector of constants. The affine invariance of $M_{n,\Phi}$ stems from the affine invariance of the Mahalanobis squared distance. That is, for $\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n$, the sample mean vector is $\overline{X}_n$ and the sample covariance matrix is S_n. Also, for affine transformed observation vectors $C\mathrm{x}_1 \pm \mathbf{u}, C\mathrm{x}_2 \pm \mathbf{u}, \ldots, C\mathrm{x}_n \pm \mathbf{u}$, the sample mean vector is $C\overline{X}_n \pm u$ and the sample covariance matrix is CS_nC. Then the sample Mahalanobis squared distance of the affine transformed observation vectors is given by:

$$\left[C\mathbf{x}_j \pm u - C\overline{X}_n \pm u\right]^T \left(CS_nC\right)^{-1}\left[C\mathbf{x}_j \pm u - C\overline{X}_n \pm \mathbf{u}\right]$$

$$\left[C\left(\mathbf{x}_j - \overline{X}_n\right)\right]^T \left(CS_nC\right)^{-1}\left[C\left(\mathbf{x}_j - \overline{X}_n\right)\right]$$

$$\left(\mathbf{x}_j - \overline{X}_n\right)^T C^T \left(C^T\right)^{-1} S_n^{-1} C^{-1} C\left(\mathbf{x}_j - \overline{X}_n\right) = \left(\mathbf{x}_j - \overline{X}_n\right)^T S_n^{-1}\left(\mathbf{x}_j - \overline{X}_n\right)$$

Hence,

$$\frac{n}{(n-1)^2}\left(\mathbf{x}_j - \overline{X}_n\right)^T S_n^{-1}\left(\mathbf{x}_j - \overline{X}_n\right) = z_j$$

$$= \frac{n}{(n-1)^2}\left(C\mathbf{x}_j \pm \mathbf{u} - C\overline{X}_n \pm \mathbf{u}\right)^T \left(CS_nC\right)^{-1}\left(C\mathbf{x}_j \pm \mathbf{u} - C\overline{X}_n \pm \mathbf{u}\right)\ j = 1, 2, \ldots, n.$$

Since $Z \sim B(a,b)$, where $a = \frac{d}{2}$ and $b = \frac{(n-d-1)}{2}$ which do not depend on any sample observation vector $\mathbf{x}_j$, the invariance property is proved.

The invariance property of the $M_{n,\Phi}$ statistic, as proved in Theorem 2.1 is because the transformations in (2) and (3) are standardized transformations that result in beta distributed observations with constant parameters for each n and d such that no matter the affine transformation in $\mathbf{x}_j$'s, $j = 1,2,\ldots,n$, z_j's have a specified beta distribution (with specified parameters). As a result, under any null distribution of MVN, the value of the statistic is unaffected at any fixed sample size n and variable dimension d.

Theorem 2.2: Suppose $F(\mathbf{x})$ is an unknown continuous distribution in a d-dimensional real space, R^d, with a probability density function $f(\mathbf{x})$, having unknown mean vector and unknown covariance matrix. Then, the test based on $M_{n,\Phi}$ is consistent.

Proof:

Under the null distribution of multivariate normality, $Z \sim B(a,b)$, where Z is the random variable from where z_j in (3) is assumed to have come from. Hence, as $n, m \to \infty$ and $m/n \to 0$,

$$F_n\left(z_{(j+m)}\right) - F_n\left(z_{(j-m)}\right) \simeq F\left(z_{(j+m)}\right) - F\left(z_{(j-m)}\right)$$
$$\simeq \frac{f\left(z_{(j+m)}\right) + f\left(z_{(j-m)}\right)}{2}\left(z_{(j+m)} - z_{(j-m)}\right).$$

Now, it is obvious that the a and b in the distribution of Z are consistent since a is fixed and b is based on sample size.

Hence, $E\left(M_{n,\Phi}\right) = E\left(\frac{1}{n}\sum_{j=1}^{n}\Phi\left(\frac{n\Gamma(a+b)(Z_{(j)})^{a-1}(1-Z_{(j)})^{b-1}\left(Z_{(j+m)}-Z_{(j-m)}\right)}{2m\Gamma(a)\Gamma(b)}\right)\right)$

$$= E\left\{\Phi\left(\frac{n\Gamma(a+b)(Z_{(j)})^{a-1}(1-Z_{(j)})^{b-1}\left(Z_{(j+m)}-Z_{(j-m)}\right)}{2m\Gamma(a)\Gamma(b)}\right)\right\}.$$

Again, $Z_{(j-m)}$ and $Z_{(j+m)}$ belong to an interval where $f(z)$ is both positive and continuous. Then according to Vasicek (1976) and Tavakoli et al. (2020), there exists $z_j^* \in \left(Z_{(j-m)}, Z_{(j+m)}\right)$ such that

$$\frac{F\left(Z_{(j+m)}\right) - F\left(Z_{(j-m)}\right)}{Z_{(j+m)} - Z_{(j-m)}} = f(z_j^*).$$

Therefore, $M_{n,\Phi} \to D_\Phi\left(F_Z, G_Z\right)$ and hence, $M_{n,\Phi}$ is consistent.

3. Simulation study

In this section, extensive simulations are carried out to obtain the critical values of the proposed test as well as to determine their relative performance. For these purposes, it is important to first determine an appropriate window size, m for each sample size, n and number of variables, d in a multivariate dataset. Wieczorkowski and Grzegorzewski (1999) have proposed an optimal value of m for estimating the entropy of a distribution to be a function of the sample size as $m = [\sqrt{n}+0.5]$, where $[x]$ is the integer part of x. However, it has been shown that an appropriate m also depends on the underlying distribution in addition to the sample size as against the suggestion of Wieczorkowski and Grzegorzewski (1999). Again, one serious problem with the application of Tavakoli et al. (2020) statistic is lack of operational function for determining m.

Now, the statistic is based on the Vasicek (1976) estimator of the Shannon (1941) entropy of a random variable. Therefore, the empirical mean squared error (EMSE) of the Vasicek (1976) estimator was computed for all the possible values of m, $m \leq \frac{n}{2}$ under the beta distribution with parameters $a = \frac{d}{2}$; $b = \frac{(n-d-1)}{2}$. This is carried out for $n = 5(5)100(10)150$ and $d = 2, 5$, and, 10. In each combination of n and d, an appropriate m is obtained as the one with the smallest EMSE and we used the selected m values to obtain a linear trend equation of m for each combination of n and d as $m = 3.2349 + 0.0808n - 0.2929d$, with an R^2 value of 89 percent, see Figure 1. The Vasicek (1976) estimator is given by $H_{mn} = \frac{1}{n}\sum_{j=1}^{n} log\left\{\frac{n}{2m}\left(X_{(j+1)} - X_{(j-m)}\right)\right\}$ and the EMSE is based on 10,000 replications of each sample size, n drawn from the beta distribution with parameters $a = \frac{d}{2}$; $b = \frac{(n-d-1)}{2}$.

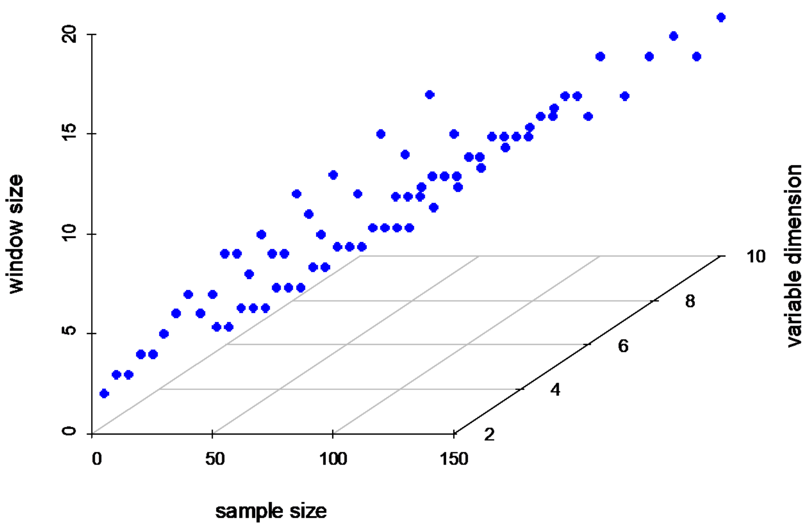

Figure 1: 3D scatter plot of n, d and associated m

3.1. Empirical critical values of the test

The statistic proposed in this work is generic in nature. Therefore, its critical value, application and performance depend on the specific phi divergence measure being used. We have stated that there are several phi divergence measures but our study in this section is limited to only three which are based on sum of squares. They are the Pearson; Hellinger; and Balakrishnan-Sanghvi divergence measures and the proposed statistic for them are $M(P)$, $M(H)$, and $M(BS)$ respectively.

For each combination of sample sizes n = 10 (5) 100 (10) 150 and random vector dimensions d = 2, 5, and 10, 5 percent level critical values were evaluated. To achieve this, a total of N = 100,000 samples for each combination of n and d from the standard multivariate normal distribution were generated and each generated sample was transformed into a beta sample according to (3). Then, each of the three sum of square versions of the statistic is computed from each of the beta transformed samples to arrive at N = 100,000 values of each statistic. The 5 percent level critical value is calculated as the 95 percentile of the N values in each version of the statistic. The critical values are presented in Table 1. The test then proposes to reject the MVN of a dataset with sample size n and number of variables d if the computed value of the statistic is greater than the corresponding critical value at 5 percent level of significance.

Table 1: Empirical critical values at $\alpha = 0.05$

	$M(P)$			$M(H)$			$M(BS)$		
n	$d=2$	$d=5$	$d=10$	$d=2$	$d=5$	$d=10$	$d=2$	$d=5$	$d=10$
10	0.5773	0.4318	-	0.1687	0.1695	-	0.1988	0.2195	-
15	0.6623	0.4764	0.8508	0.1604	0.1528	0.2220	0.1697	0.1853	0.2491
20	0.6173	0.4042	0.8188	0.1423	0.1250	0.2072	0.1422	0.1490	0.2281
25	0.5633	0.4044	0.4720	0.1262	0.1154	0.1311	0.1240	0.1324	0.1502
30	0.5281	0.3591	0.4713	0.1156	0.1019	0.1253	0.1112	0.1159	0.1402
35	0.4789	0.3473	0.3612	0.1047	0.0953	0.0997	0.1014	0.1067	0.1133
40	0.4464	0.3248	0.3557	0.0961	0.0884	0.0952	0.0932	0.0979	0.1064
45	0.4291	0.3141	0.3500	0.0914	0.0828	0.0918	0.0868	0.0916	0.1013
50	0.3956	0.2983	0.2953	0.0851	0.0790	0.0800	0.0812	0.0862	0.0896
55	0.3891	0.2848	0.2921	0.0826	0.0743	0.0775	0.0773	0.0813	0.0853
60	0.3626	0.2757	0.2622	0.0767	0.0714	0.0709	0.0728	0.0769	0.0784
65	0.3480	0.2656	0.2557	0.0728	0.0689	0.0683	0.0690	0.0741	0.0750
70	0.3434	0.2589	0.2518	0.0714	0.0653	0.0663	0.0667	0.0705	0.0725
75	0.3252	0.2490	0.2331	0.0682	0.0634	0.0622	0.0634	0.0677	0.0684
80	0.3243	0.2443	0.2284	0.0669	0.0617	0.0606	0.0619	0.0659	0.0660
85	0.3100	0.2368	0.2171	0.0642	0.0598	0.0581	0.0594	0.0632	0.0636
90	0.2980	0.2353	0.2144	0.0615	0.0587	0.0564	0.0569	0.0621	0.0615
95	0.2979	0.2306	0.2100	0.0608	0.0570	0.0549	0.0559	0.0600	0.0593
100	0.2876	0.2235	0.2041	0.0589	0.0550	0.0532	0.0537	0.0579	0.0580
110	0.2835	0.2163	0.1946	0.0565	0.0527	0.0509	0.0514	0.0553	0.0552
120	0.2741	0.2127	0.1880	0.0549	0.0510	0.0490	0.0495	0.0530	0.0528
130	0.2708	0.2093	0.1818	0.0536	0.0497	0.0468	0.0475	0.0512	0.0503
140	0.2542	0.2049	0.1786	0.0505	0.0483	0.0454	0.0450	0.0496	0.0485
150	0.2527	0.1975	0.1752	0.0495	0.0461	0.0442	0.0436	0.0474	0.0473

3.2. Description of the competing tests

Primarily, the three versions of our proposed test are according to the phi divergence measures due to Pearson, Hellinger, as well as Barakrishnan and Sanghvi. They are

$$M(P) = \frac{1}{n}\sum_{j=1}^{n}(U_j - 1)^2$$

$$M(H) = \frac{1}{2n}\sum_{j=1}^{n}\left(\sqrt{U_j} - 1\right)^2$$

$$M(BS) = \frac{1}{n}\sum_{j=1}^{n}\left(\frac{U_j - 1}{U_j + 1}\right)^2$$

where $U_j = \frac{n\Gamma(a+b)Z_{(j)}^{a-1}(1-Z_{(j)})^{b-1}\left(Z_{(j+m)}-Z_{(j-m)}\right)}{2m\Gamma(a)\Gamma(b)}$; $a = d/2$; $b = (n-d-1)/2$; $Z_{(j)}$ is the jth order statistic of the Z-transformed dataset such that $Z_{(j+m)} = Z_{(n)}$ for all $j+m \geq n$ and $Z_{(j-m)} = Z_{(1)}$ for all $j-m \leq 1$.

The existing statistics considered in this work for a proper comparison with these new ones include the Henze and Zirkler (HZ) test for MVN of Henze and Zirkler (1990); the Madukaife (M) test for MVN of Madukaife (2017); and the Henze and Jimenez-Gamero (HJG) test for MVN of Henze and Jimenez-Gamero (2019). The choice of the three competing tests is not completely arbitrary. First, they are all affine invariant and consistent L^2-type tests for MVN with good power performances. Secondly, in most comparative studies on powers of tests for MVN, the $HZ-$statistic has remained a reference point while the $HJG-$statistic is similar to it. In fact, any test for MVN that competes favourably with the $HZ-$statistic is generally regarded as a good statistic for assessing MVN of datasets. Again, the $M-$statistic is also based on beta transform of multivariate datasets. Also, since the choice of d presented in this work is $d = 2$, 5, 10, which represents multivariate datasets, comparison with good univariate tests for normality such as Jargue and Bera (1987) as well as Bayoud (2021) is not discussed here. It may be the interest of a future work. In what follows, therefore, the three competing statistics are described.

3.2.1 Henze and Zirkler HZ test

Henze and Zirkler (1990) introduced a smoothing parameter, β in the weight function of the consistent and affine invariant statistic due to Baringhaus and Henze (1988) to obtain a highly regarded test for MVN of multivariate datasets. The statistic is given as:

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

where $D_{n,\beta} = \left(1+2\beta^2\right) + n^{-2}\sum_{j,k=1}^{n} exp\left\{-\frac{\beta^2\left\|y_j - y_k\right\|^2}{2}\right\}$ $-2\left(1+\beta^2\right)^{-d/2}n^{-1}\sum_{j=1}^{n} exp\left\{-\frac{\beta^2\left\|y_j\right\|^2}{2\left(1+\beta^2\right)}\right\}$; $\beta > 0$ and $I\{.\}$ is an indicator function. The test is universally consistent, affine invariant and rejects MVN of datasets for large values of the statistic, with appropriate $\beta = \frac{((2d+1)n/4)/(1/(d+4))}{\sqrt{2}}$.

3.2.2 Madukaife *M* test

Madukaife (2017) obtained a statistic to formalize the graphical test of Small (1978), using the sum of squared differences between expected and sample order statistics according to Madukaife and Okafor (2018), who also formalized the geometric procedure of Hanusz and Tarasinska (2012) to a classical test procedure. The statistic, which is the sum of squared differences between observed and expected order statistics of beta transformed observations, is given as

$$M = \sum_{j=1}^{n} \left(z_{(j)} - c_j\right)^2$$

where $z_{(j)}$ is the jth order statistic of the beta transformed observations and c_j is the corresponding jth expected order statistic from the beta distribution with parameters $a = d/2$ and $b = (n-d-1)/2$. The consistent and affine invariant test rejects the null hypothesis of MVN for large values of the statistic.

3.2.3 Henze and Jimenez-Gamero *HJG* test

Henze and Jimenez-Gamero (2019) obtained a statistic for assessing MVN based on the empirical moment generating function. It is a weighted squared integral of the difference between the theoretical and empirical moment generating functions respectively of the standard multivariate normal distribution and a multivariate dataset. The statistic is given as

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

where $\beta > 1, Y_{n,j}$ is the jth d-dimensional standardized multivariate data point contained in the standardized sample of size n and $\|.\|$ is a vector norm. The *HJG* test rejects the null distribution of MVN for large values of the statistic.

3.3. Size and power comparison of the competing tests

The power of a test, which is the ability of the test to reject a wrong null hypothesis, and the size of a test, which is the maximum probability of rejecting a true null hypothesis, are among the most important properties of a test. Although they can be obtained theoretically when the true null distribution of the test statistic is known, the sizes and powers of the three specific versions of the $M(\Phi)$ statistic however are obtained empirically and compared with those of other well-known statistics in the literature. To achieve the objective of power comparison in this work, four different classes of distributions other than the multinormal distribution are identified. They are short-tailed symmetric distributions as group I; heavy-tailed symmetric distributions as group II; short-tailed asymmetric distributions as group III;

and heavy-tailed asymmetric distributions as group IV. Three distributions were considered from each of the four groups in this study and they include the following:

Group I

Standard multivariate Laplace distribution (MVL)

Products of the univariate Laplace distribution $\left(L^d(0,1)\right)$

Products of the univariate Laplace and the symmetric beta distribution $\left(L^p(0,1) \otimes B^{d-p}(1.5,1.5)\right)$

Group II

Multivariate Cauchy distribution (MVC)

Multivariate t distribution with 2 degrees of freedom $(MVt(2))$

Products of the univariate t with 5 degrees of freedom and the Cauchy distributions $\left(t^p(5) \otimes C^{d-p}(0,1)\right)$

Group III

Products of the standard exponential distribution $\left(Exp^d(1)\right)$

Products of the gamma distribution $\left(G^d(1,3)\right)$

Products of the gamma and Gumbel distributions $\left(Ga^p(1,3) \otimes Gu^{d-p}(0,1)\right)$

Group IV

Products of the Pareto distribution $\left(P^d(1,2)\right)$

Products of the standard lognormal distribution $\left(LN^d(0,1)\right)$

Products of the Weibull distribution $\left(W^d(1,2)\right)$

where p is an integar less than d.

A total of 10,000 datasets from each of the 12 distributions grouped into I-IV and the standard multivariate normal distribution were simulated in each combination of sample sizes $n = 10, 25, 50, \text{and} 100$ and variable dimensions $d = 2 \text{and} 5$. For each of the combinations of sample size and variable dimension, the values of each of the competing statistics were calculated and the estimated power performance of each statistic was obtained as the percentage of the 10,000 simulated samples that is rejected by the statistic at 5% level of significance. The null distribution is the multivariate normal distribution. Therefore the power performances of the statistics obtained from it are the empirical probabilities of committing the error of type one, also known as the size of a test, which in this work are expected to be equal to 5%. The type-one-error rates of the competing statistics are presented in Table 2. Also, their power performances are presented in Tables 3 and 4 for sample sizes $n = 10, 25, 50, \text{and} 100$ respectively.

From the results in Table 2, all the six tests considered showed very good control over type-one-error. This is because, none of them recorded a type-one-error of more than the 5% level of significance in any of the combinations of sample size, n and variable dimension, d. Again, while all the other tests, including the new techniques, maintained a type-one-error of 5% ($4.5\% \leq \alpha < 5.5\%$) in all the combinations of n and d considered, the HZ test maintained a conserved state (less than 5%) in all the variable dimensions considered at sample sizes up to 50. This, however, is not a disadvantage to the technique, it rather assures that the power performance of the statistic is completely devoid of the error of type-one. The

Table 2: Empirical type-I-error rate of the competing statistics

n	d	HZ	M	HJG	$M(P)$	$M(H)$	$M(BS)$
	2	2.5	4.9	5.1	4.6	5.1	5.0
10	5	0.9	5.0	5.2	4.7	5.0	4.9
	10	-	-	-	-	-	-
	2	4.2	5.2	5.2	4.8	5.0	5.2
25	5	3.3	4.9	4.5	4.9	5.0	5.2
	10	3.3	5.1	5.2	4.8	5.0	5.3
	2	4.4	5.0	5.2	5.1	4.8	4.8
50	5	4.1	4.9	5.2	5.2	4.8	4.6
	10	4.2	5.2	4.8	5.4	5.0	4.8
	2	4.7	5.0	5.2	5.0	4.9	5.0
100	5	4.7	4.9	4.7	4.6	4.9	4.8
	10	4.8	5.0	4.8	4.9	5.0	5.2
	2	4.8	5.0	5.0	5.0	4.8	4.9
150	5	4.8	5.1	4.9	4.9	5.0	5.0
	10	4.9	5.0	5.2	5.2	4.9	4.8

error rates of all the statistics considered at sample size, $n = 10$ and variable dimension, $d = 10$ are not obtained due to the fact that such a dataset is known to be singular. Based on the results in Table 2, the new phi-divergence statistics can be said to have a good control over type-one-error and hence can be recommended, at that instance, as a good technique for testing MVN of datasets.

From Tables 3 and 4, it is observed that the new statistics are generally slightly more powerful than the other competing techniques considered in this work under the alternative symmetric distributions in Table 3, especially at large sample sizes of $n > 25$. The only exception, however, is the products of the univariate Laplace and symmetric beta distributions where the HZ statistic is observed to be slightly more powerful than the rest of the techniques considered, including the new statistics. Among the three new statistics obtained from the sample phi-divergence measure statistic in (9), the $M(BS)$ generally recorded least power performance at small sample size but most powerful, together with the $M(H)$, at large sample sizes of $n \geq 25$ under these alternative symmetric distributions.

Conversely, under the asymmetric alternative distributions in groups III and IV as presented in Table 4, it is observed that the new statistics are generally slightly less powerful than the other three L^2-type statistics, especially at large sample sizes. The only exception is the products of the univariate gamma and Gumbel distributions where the new statistics are as good as the other statistics. It is, however, expected that at large sample sizes of $n > 100$, the power performances of all the competing statistics would be equal. Again, it can be seen that under these alternative distributions in groups III and IV, the $M(H)$ performed better in their powers than the other two versions of the new phi-divergence technique.

Table 3: Empirical power comparisons of competing tests for MVN under alternative symmetric distributions in groups I and II, $\alpha = 0.05$

Group	Distributions	n	d	HZ	M	HJG	$M(P)$	$M(H)$	$M(BS)$
I	MVL	10	2	16.6	27.0	24.1	26.0	24.8	21.7
	MVL		5	14.6	39.4	38.9	35.8	33.6	29.7
	MVL	25	2	50.2	57.4	49.0	52.0	60.7	62.3
	MVL		5	85.6	93.0	80.1	83.3	92.7	93.0
	MVL	50	2	82.8	82.2	71.3	74.6	86.1	91.0
	MVL		5	99.7	99.7	95.4	98.8	99.9	100.0
	MVL	100	2	98.8	92.3	89.2	91.5	98.5	92.8
	MVL		5	100.0	100.0	99.9	100.0	100.0	100.0
	$L^d(0,1)$	10	2	11.3	21.5	17.6	19.5	19.3	16.2
	$L^d(0,1)$		5	3.4	16.3	15.3	14.3	12.6	11.4
	$L^d(0,1)$	25	2	34.4	45.1	39.8	40.4	45.9	46.4
	$L^d(0,1)$		5	28.8	52.7	43.3	38.6	48.8	47.2
	$L^d(0,1)$	50	2	63.3	68.6	58.2	59.4	71.2	76.1
	$L^d(0,1)$		5	64.3	82.9	67.5	64.1	80.2	83.7
	$L^d(0,1)$	100	2	91.1	90.4	78.6	77.7	91.8	96.3
	$L^d(0,1)$		5	94.6	98.1	87.2	84.5	97.6	98.8
	$L(0,1)\otimes B(1.5,1.5)$	10	2	5.4	8.3	10.1	7.0	7.3	5.5
	$L^3(0,1)\otimes B^2(1.5,1.5)$		5	1.8	7.5	8.6	6.9	6.6	6.8
	$L(0,1)\otimes B(1.5,1.5)$	25	2	19.5	17.7	20.0	16.1	15.2	9.6
	$L^3(0,1)\otimes B^2(1.5,1.5)$		5	15.1	24.5	25.4	18.1	19.8	14.1
	$L(0,1)\otimes B(1.5,1.5)$	50	2	43.8	29.7	33.9	32.3	30.1	19.4
	$L^3(0,1)\otimes B^2(1.5,1.5)$		5	40.8	44.6	43.2	36.0	38.1	32.6
	$L(0,1)\otimes B(1.5,1.5)$	100	2	81.2	45.5	49.8	52.5	53.0	42.0
	$L^3(0,1)\otimes B^2(1.5,1.5)$		5	80.7	69.0	66.0	51.6	61.6	58.8
II	MVC	10	2	70.6	75.3	72.2	74.4	73.5	67.9
	MVC		5	50.9	73.1	82.0	67.6	70.9	67.8
	MVC	25	2	98.7	98.8	97.6	98.5	99.0	99.0
	MVC		5	99.9	100.0	99.9	99.9	100.0	100.0
	MVC	50	2	100.0	100.0	100.0	100.0	100.0	100.0
	MVC		5	100.0	100.0	100.0	100.0	100.0	100.0
	MVC	100	2	100.0	100.0	100.0	100.0	100.0	100.0
	MVC		5	100.0	100.0	100.0	100.0	100.0	100.0
	$MVt(2)$	10	2	29.4	38.7	43.8	36.7	36.3	36.1
	$MVt(2)$		5	13.7	34.5	50.8	30.5	30.0	33.3
	$MVt(2)$	25	2	71.1	81.2	79.8	78.1	79.6	83.9
	$MVt(2)$		5	89.0	96.5	96.2	92.1	95.2	96.6
	$MVt(2)$	50	2	93.4	96.7	95.8	95.3	96.7	98.7
	$MVt(2)$		5	99.6	100.0	99.9	99.8	99.9	100.0
	$MVt(2)$	100	2	99.8	99.9	99.9	99.7	99.9	100.0
	$MVt(2)$		5	100.0	100.0	100.0	100.0	100.0	100.0
	$t(5)\otimes C(0,1)$	10	2	41.9	51.0	50.6	47.3	47.4	38.8
	$t^3(5)\otimes C^2(0,1)$		5	12.3	30.3	44.1	26.6	30.5	26.1
	$t(5)\otimes C(0,1)$	25	2	88.5	88.9	86.3	86.5	88.3	87.1
	$t^3(5)\otimes C^2(0,1)$		5	90.3	95.5	94.2	90.7	93.5	90.7
	$t(5)\otimes C(0,1)$	50	2	99.2	99.2	98.4	98.6	99.1	99.3
	$t^3(5)\otimes C^2(0,1)$		5	99.9	100.0	99.8	99.6	99.9	99.9
	$t(5)\otimes C(0,1)$	100	2	100.0	100.0	99.0	100.0	100.0	100.0
	$t^3(5)\otimes C^2(0,1)$		5	100.0	100.0	100.0	100.0	100.0	100.0

Table 4: Empirical power comparisons of competing tests for MVN under alternative skewed distributions in groups III and IV, $\alpha = 0.05$

Group	Distributions	n	d	HZ	M	HJG	M(P)	$M(H)$	$M(BS)$
III	$Exp^d(1)$	10	2	35.0	32.3	37.0	32.0	30.4	25.6
	$Exp^d(1)$		5	11.4	19.3	29.6	18.8	17.7	16.6
	$Exp^d(1)$	25	2	92.4	66.0	72.4	69.9	67.8	58.5
	$Exp^d(1)$		5	92.2	77.2	76.8	69.4	71.0	66.4
	$Exp^d(1)$	50	2	96.4	88.6	94.6	90.3	91.3	88.5
	$Exp^d(1)$		5	99.9	96.6	95.1	92.9	94.6	93.9
	$Exp^d(1)$	100	2	100.0	98.9	98.0	98.8	99.4	99.4
	$Exp^d(1)$		5	100.0	100.0	99.8	99.6	99.9	99.9
	$Ga^d(1,3)$	10	2	34.6	33.1	35.5	31.5	30.3	25.0
	$Ga^d(1,3)$		5	12.2	19.1	29.0	19.3	17.3	16.1
	$Ga^d(1,3)$	25	2	92.6	65.1	71.8	69.4	68.2	59.2
	$Ga^d(1,3)$		5	92.2	76.9	75.3	68.3	71.3	65.0
	$Ga^d(1,3)$	50	2	99.9	89.0	94.2	90.9	91.1	88.1
	$Ga^d(1,3)$		5	100.0	96.5	94.8	92.5	94.4	93.6
	$Ga^d(1,3)$	100	2	100.0	99.0	99.0	99.0	99.4	99.3
	$Ga^d(1,3)$		5	100.0	99.9	99.9	99.6	99.9	99.9
	$Ga(1,3) \otimes Gu(0,1)$	10	2	10.0	23.7	26.4	20.9	22.1	16.7
	$Ga^3(1,3) \otimes Gu^2(0,1)$		5	2.3	14.1	21.7	13.7	12.7	11.3
	$Ga(1,3) \otimes Gu(0,1)$	25	2	36.4	51.0	57.5	51.5	49.0	39.5
	$Ga^3(1,3) \otimes Gu^2(0,1)$		5	28.8	63.2	63.5	53.9	56.5	49.4
	$Ga(1,3) \otimes Gu(0,1)$	50	2	67.0	74.7	84.6	75.0	75.5	67.8
	$Ga^3(1,3) \otimes Gu^2(0,1)$		5	67.7	90.6	88.5	80.4	86.1	82.3
	$Ga(1,3) \otimes Gu(0,1)$	100	2	93.2	93.5	98.7	91.3	93.4	93.0
	$Ga^3(1,3) \otimes Gu^2(0,1)$		5	95.8	99.4	99.0	96.3	98.5	98.3
IV	$Pa^d(1,2)$	10	2	73.4	67.5	68.7	65.9	65.7	57.7
	$Pa^d(1,2)$		5	50.8	57.3	74.3	55.8	56.1	52.3
	$Pa^d(1,2)$	25	2	99.9	95.7	97.8	96.9	96.6	95.2
	$Pa^d(1,2)$		5	100.0	99.5	99.5	99.1	99.4	99.2
	$Pa^d(1,2)$	50	2	100.0	99.7	100.0	99.9	99.9	99.9
	$Pa^d(1,2)$		5	100.0	100.0	100.0	100.0	100.0	100.0
	$Pa^d(1,2)$	100	2	100.0	100.0	100.0	100.0	100.0	100.0
	$Pa^d(1,2)$		5	100.0	100.0	100.0	100.0	100.0	100.0
	$LN^d(0,1)$	10	2	57.7	54.3	56.8	52.3	52.8	44.5
	$LN^d(0,1)$		5	31.0	40.1	56.2	38.9	37.4	35.9
	$LN^d(0,1)$	25	2	98.9	90.0	92.4	90.8	90.5	86.5
	$LN^d(0,1)$		5	99.7	97.6	96.7	95.5	96.5	95.5
	$LN^d(0,1)$	50	2	100.0	99.3	99.9	99.4	99.4	99.1
	$LN^d(0,1)$		5	100.0	100.0	100.0	100.0	100.0	100.0
	$LN^d(0,1)$	100	2	100.0	100.0	100.0	100.0	100.0	100.0
	$LN^d(0,1)$		5	100.0	100.0	100.0	100.0	100.0	100.0
	$W^d(1,2)$	10	2	33.7	33.0	35.8	31.3	30.2	24.9
	$W^d(1,2)$		5	11.7	19.9	28.1	19.0	17.4	15.9
	$W^d(1,2)$	25	2	72.9	65.2	73.7	69.9	68.1	58.7
	$W^d(1,2)$		5	92.4	76.5	75.6	67.4	70.8	65.7
	$W^d(1,2)$	50	2	100.0	89.1	94.8	90.6	91.4	87.9
	$W^d(1,2)$		5	100.0	96.5	95.0	92.3	94.8	93.4
	$W^d(1,2)$	100	2	100.0	98.7	99.9	98.9	99.5	99.6
	$W^d(1,2)$		5	100.0	100.0	99.9	99.6	99.9	99.9

Table 5: Mean empirical powers of competing tests for MVN, under alternative distributions, $n \geq 25$ and $\alpha = 0.05$

Distribution group	HZ	M	HJG	$M(P)$	$M(H)$	$M(BS)$
I	65.2611	66.3000	60.9889	59.5278	66.1778	64.6722
II	96.0778	97.5944	97.0444	96.6000	97.2833	97.5111
III	86.3611	84.7833	86.8611	82.7222	83.7389	80.1111
IV	97.9889	94.8778	95.8444	94.4611	94.8222	93.3667

In order to give a clearer picture of the competitive nature of the new tests, the mean empirical power performances of the six competing statistics are presented in Table 5 and it is evident from the table that the new statistics can be recommended as good tests for assessing MVN of datasets, especially at large samples as well as when the dataset is known to be symmetric.

4. Data application

In this section, the applicability of the new statistics is presented in comparison with those of the other three competing techniques. This is carried out on a set of four multivariate datasets, which are retrieved from https://openmv.net/tag/multivariate. The datasets are as follows:

Brittleness index dataset: This is a 3-component dataset, comprising of 18 observation vectors. It is obtained as measures of brittleness of plastic products produced in three parallel reactors, TK104, TK105 and TK107 as the components.

Film thickness dataset: This is a 4-component dataset obtained as thickness measurements taken at four different positions of 160 plastic films after being cut. The measurement positions which make up the data components included top right, top left, bottom right and bottom left.

Room temperature dataset: This is another 4-component dataset which is obtained as temperature measurements, in Kelvin, taken at four corners of a room. The measurement corners which form the data components included front left, front right, back left and back right and the measurements were taken 144 times, giving rise to a 144 rows by 4 columns dataset.

Solvents dataset: The solvents dataset is a 9-component dataset which consists of physical properties of a sample of 103 chemical solvents. The properties which form the data components included melting point, boiling point, dielectric, dipole moment, refractive index, ET30, density, logP and solubility.

The four datasets are tested independently for MVN using the six competing techniques considered in this study, each at 5% $\alpha-$level. The results comprise of their test statistics, critical values and decisions of either rejection or otherwise of their MVN reached by comparing the test statistics values with their corresponding critical values. They are presented in Table 6.

From the results in Table 6, all the six tests show perfect agreement in their decisions. Specifically, none of the six competing techniques could reject MVN of the brittleness in-

Table 6: Results of tests for MVN of some real-life datasets, $\alpha = 0.05$

Dataset	Test Components	HZ	M	HJG	$M(P)$	$M(H)$	$M(BS)$
Brittleness Index	Computed value of statistic	0.428886	0.018040	0.113696	0.255089	0.079972	0.095587
	Critical value	0.804912	0.055778	0.465075	0.495144	0.137537	0.154695
	Decision	Do not reject	Do not reject	Do not reject	Do not reject	Do not reject	Do not reject
Film Thickness	Computed value of statistic	0.992884	0.000945	2.477551	0.168497	0.036121	0.032316
	Critical value	1.009868	0.002093	3.076987	0.202478	0.045625	0.045660
	Decision	Do not reject	Do not reject	Do not reject	Do not reject	Do not reject	Do not reject
Room Temperature	Computed value of statistic	2.029179	0.025055	2134.348	0.507064	0.079452	0.053684
	Critical value	1.008710	0.002518	3.051355	0.212408	0.047989	0.048165
	Decision	Reject	Reject	Reject	Reject	Reject	Reject
Solvents	Computed value of statistic	1.750361	0.108639	8604.996	0.351212	0.104571	0.107069
	Critical value	0.997587	0.006947	9.441462	0.203694	0.053017	0.056799
	Decision	Reject	Reject	Reject	Reject	Reject	Reject

dex and film thickness data while, on the other hand, they all rejected the MVN of room temperature and solvents datasets. The result in this section further shows that the three new statistics can be regarded as good statistics for testing MVN of datasets.

5. Conclusion

The plug-in techniques developed in this study for assessing MVN of multivariate datasets have shown, through their size and power performances, that they can be regarded as good statistics. Their affine invariance and consistency properties have been proved. Also, the statistics can be adapted for goodness-of-fit test to other continuous distributions. Besides good power performances, the new statistics are computationally less tedious since they are based on univariate transform of multivariate datasets. Finally, it is not difficult to implement the new statistics developed in this paper to statistical software such as R so that users can access them for applicability to real-life situations. As a result, they are recommended as good techniques for testing normality of d-dimensional datasets, $d \geq 1$.

Acknowledgements

The authors wish to thank the editorial team and the anonymous reviewers for their useful suggestions, which have improved the quality of the paper.

References

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

Bayoud, H. A., (2021). Tests of normality: new test and comparative study. *Communications in Statistics - Simulation and Computation*, 50, pp. 4442–4463.

Bilodeau, M., Brenner, D., (1999). *Theory of multivariate statistics*. New York: Springer-Verlag.

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

Chen, W., Genton, M. G., (2023). Are you all normal? It depends! *International Statistical Review*, 91, pp. 114–139.

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

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

Dorr, P., Ebner, B. and Henze, N., (2020). Testing multivariate normality by zeros of the harmonic oscillator in characteristic function spaces. *Scandinavian Journal of Statistics*, doi.org/10.1111/sjos.12477.

Dorr, P., Ebner, B. and Henze, N., (2020). A new test of multivariate normality by a double estimation in a characterizing PDE. *Metrika*, doi.org/10.1007/s00184-020-00795-x.

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

Epps, T. W., Pulley, L. B., (1983). A test for normality based on the empirical characteristic function. *Biometrika*, 70, pp. 723–726.

Gnanadesikan, R., Kettenring, J. R., (1972). Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics*, 28, pp. 81–124. doi:10.2307/2528963.

Hanusz, Z., Tarasinska, J., (2012). New tests for multivariate normality based on Small's and Srivastava's graphical methods. *Journal of Statistical Computation and Simulation*, 82, pp. 1743–1752.

Healy, M. J. R., (1968). Multivariate normal plotting. *Applied Statistics*, 17 (2), pp. 157–161. doi:10.2307/2985678.

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

Henze, N., Jimenez-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, pp. 499–521.

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

Henze, N., 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, pp. 1109–1136.

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

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

Hwu, T., Han, C. and Rogers, K. J., (2002). The combination test for multivariate normality. *Journal of Statistical Computation and Simulation*, 72, pp. 379–390.

Jarque, C. M., Bera, A. K., (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55, pp. 163–172.

Korkmaz, S., Göksülük, D. and Zararsiz, G., (2014). MVN: An R package for assessing multivariate normality. *R Journal*, 6(2), pp. 151–162.

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

Lin, J., (1991). Divergence measures based on the Shannon entropy. *Information Theory–IEEE Transactions on Reliability*, 37 (1), pp. 145–151.

Madukaife, M. S., (2017). A new affine invariant test for multivariate normality based on beta probability plots. *Journal of the Nigerian Statistical Association*, 29, pp. 58–70.

Madukaife, M. S., 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, pp. 1264–1275.

Madukaife, M. S., 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), pp. 1651–1664.

Malkovich, J. F., Afifi, A. A., (1973). On tests for multivariate normality. *Journal of the American Statistical Association*, 68(341), pp. 176–179.

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

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

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

Pearson, K., (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling,". *Philosophical Magazine*, 5th Series 50, pp. 157–175.

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

Romeu, J. L., Ozturk, A., (1993). A comparative study of goodness-of-fit tests for multivariate normality. *Journal of Multivariate Analysis*, 46(2), pp. 309–334.

Royston, J. P., (1983). Some techniques for assessing multivariate normality based on the Shapiro-Wilk W. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 32(2), pp. 121–133.

Shannon, C. E., (1948). A mathematical theory of communications. *Bell System Technical Journal*, 27, pp. 379–423, 623–656. doi:10.1002/bltj.1948.27.issue-3.

Shapiro, S. S., Wilk, M. B., (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52 (3 and 4), pp. 591–611. doi:10.1093/biomet/52.3-4.591.

Singh, A., (1993). Omnibus robust procedures for assessment of multivariate normality and detection of multivariate outlier, in Multivariate Environmental Statistics, G.P. Patil and C.R. Rao eds., Amsterdam: North Holland.

Small, N. J. H., (1978). Plotting squared radii. *Biometrika*, 65 (3), pp. 657–658.

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

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

Tavakoli, M., Alizadeh Noughabi, H. and Borzadaran, G. R. M., (2020). An estimation of phi divergence and its application in testing normality. *Hacettepe Journal of Mathematics and Statistics*, 49 (6), pp. 2104-2118.

Tavakoli, M., Arghami, N. and Abbasnejad, M., (2019). A goodness of fit test for normality based on Balakrishnan-Sanghvi information. *Journal of the Iranian Statistical Society*, 18 (1), pp. 177–190.

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

Thulin, M., (2014). Tests for multivariate normality based on canonical correlations. *Statistical Methods & Applications*, doi:10.1007/s10260-013-0252-5.

Vasicek, O., (1976). A test for normality based on sample entropy. *Journal of the Royal Statistical Society B*, 38, pp. 54–59.

Villasenor, J. A., Gonzalez-Estrada, E., (2009). A generalization of Shapiro-Wilks test for multivariate normality. *Communications in Statistics-Theory and Methods*, 38 (11), pp. 1870–1883.

Wieczorkowski, R., Grzegorzewsky, P., (1999). Entropy estimators improvements and comparisons. *Communication in Statistics Simulation and Computation*, 28(2), pp. 541–567. doi:10.1080/03610919908813564.

Zhou, M., Shao, Y., (2014). A powerful test for multivariate normality. *Journal of Applied Statistics*, 41, pp. 1–13.