AN OVERALL TEST FOR MULTIVARIATE NORMALITY

C. RADHAKRISHNA RAO AND HYDAR ALI

ABSTRACT. There are a number of methods in the statistical literature for testing whether observed data come from a multivariate normal (MVN) distribution with an unknown mean vector and covariance matrix. Let $X_1, \ldots, X_n$ be an iid sample of size n from a p-variate normal distribution. Denote the sample mean and sample variance-covariance matrix by $\bar{X}$ and S respectively. Most of the tests of multivariate normality are based on the result that $Y_i = S^{-\frac{1}{2}}(X_i - \bar{X})$, $i = 1, \ldots, n$, are asymptotically iid as p-variate normal with zero mean vector and identity covariance matrix. Tests developed by Andrews et al., Mardia and others are direct functions of Y_i. We note that the $N = np$ components of the Y_i's put together can be considered as an asymptotically iid sample of size N from a univariate normal (UVN) distribution with zero mean and unit variance. We test this hypothesis using any well known test based on N independent observations for univariate normality. In particular we use univariate skewness and kurtosis tests, which are sensitive to deviations from normality.

1. INTRODUCTION

The assumption of multivariate normality (MVN) underlies many important techniques in multivariate analysis. When the sample size is large one can use test criteria based on (or optimal for) MVN but refer them to their asymptotic distributions which are independent of the underlying distribution of the observations. Thus if one computes Hotelling's T^2 to test the equality of means of p variables in two populations, a constant times T^2 can be considered as a chisquare on p degrees of freedom although the measurements are not MVN, provided the sample sizes for the two populations are large. In small samples such a procedure may give misleading results. However, simulation results show that the p-values based on the assumption of MVN may not be widely off if the underlying distribution is close to MVN. We suggest that in practice a test for MVN of observed data should generally precede any inferential analysis

Key words and phrases. Skewness and kurtosis tests. Univariate normality tests. Multivariate normality tests.

Research work of C.R. Rao was supported by the Army Research Office under Grant DAAH04-96-1-0082. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon. The work was completed during Hydar Ali's visit to the Center for Multivariate Analysis.

we wish to undertake using MVN as the underlying stochastic model for the observations. If non-normality is detected we may wish to see if a suitable transformation can restore multivariate normality or explore the possibility of using nonparametric procedures.

There are numerous tests available in the statistical literature for testing MVN. For general surveys of these tests, the reader is referred to papers by Mardia (1980), Andrews, Gnanadesikan and Warner (1971), Cox and Small (1978), Baringhaus and Henze (1988), Henze and Wagner (1997) and books by Siotani, Hayakawa and Fujikoshi (1985), Jobson (1992) and others. There appears to be no single test which is more powerful than all the other proposed tests. The choice of a test will depend on what kind of departures from normality need to be examined in the context of a particular problem under investigation as noted by Cox and Small (1978).

In this paper, we suggest an overall test which can be routinely implemented in the initial data analysis (or cross examination of data as discussed in Rao (1997)) for choosing a model for inferential data analysis on mean values, construction of discriminant functions and related problems. The proposed test appears to be more sensitive than some others in the data sets we have examined.

2. The proposed test

Let $X_1, \ldots, X_n$ be an iid sample of size n from a p-variate population and denote the partitioned $p \times n$ matrix $(X_1 : \cdots : X_n)$ by X. Then the sample mean and estimated variance-covariance matrix are

$$\begin{aligned} \bar{X} &= n^{-1}X1 \\ S &= (n-1)^{-1}X(I - n^{-1}11')X' \end{aligned} \tag{2.1}$$

where 1 is the n-vector of unities.

Consider the spectral decomposition of S

$$S = \lambda_1^2 P_1 P' + \ldots + \lambda_p^2 P_p P_p' = P\Lambda^2 P' \tag{2.2}$$

where the λ_i^2's are the eigen values and the P_i's are the eigen vectors, $P = (P_1 : \cdots : P_p)$ and Λ^2 is a diagonal matrix with λ_i^2 as its i-th diagonal element. Then

$$S^{-\frac{1}{2}} = \lambda_1^{-1} P_1 P' + \ldots + \lambda_p^{-1} P_p P_p' = P\Lambda^{-1}P' \tag{2.3}$$

is a positive definite square root of S^{-1}. Further let

$$Y_i = S^{-\frac{1}{2}}(X_i - \bar{X}),\ i = 1, \ldots, n. \tag{2.4}$$

Then under the assumption that the X_i's are iid variables from $N_p(\mu, \Sigma)$, i.e., a p-variate normal distribution with mean μ and covariance matrix Σ, smooth functions of moments of Y_i tend to have the same distribution as the corresponding functions of $Z_i = \Sigma^{-\frac{1}{2}}(X_i - \mu)$ as $n \to \infty$. We consider

$$Y' = (Y_1' : \cdots : Y_n') \tag{2.5}$$

as approximately distributed as $N_{np}(0, I)$, i.e., treat all the $N = np$ components of Y as iid variables from $N_1(0,1)$, and apply a test for univariate normality based on N observations. Thus assessing MVN is reduced to a test for UVN (univariate normality). We believe such a test will be more sensitive as it is based on a large number of observations. Other authors cited in the introduction use different functions of Y_i in constructing tests of MVN. Andrews et al. (1971) use $Y_i'Y_i$, $i = 1, \ldots, n$, as independent chisquares on p degrees of freedom. Mardia (1980) uses $Y_i'Y_j$ and $Y_i'Y_i$ in defining multivariate skewness and kurtosis and suggests tests based on them. The test proposed by Baringhaus and Henz (1988) is a function of $\|Y_i\|^2$ and $\|Y_i - Y_j\|^2$.

In this paper we consider tests of UVN based on $\sqrt{b_1}$ and b_2, the measures of skewness and kurtosis computed from the N observations in (2.5). One could also use the Shapiro-Wilk test W by considering the normal approximation

$$z = [(1 - W)^{\lambda} - \mu]/\sigma \tag{2.6}$$

where λ, μ and σ can be obtained for given n from Table 2 of Royston (1982, p.119) or the normal approximation by using the Johnson (1949) S_b family

$$z = \hat{\gamma} + \hat{\delta}\log[(W - \hat{\epsilon})/(1 - W)]$$

where $\hat{\gamma}, \hat{\delta}$ and $\hat{\epsilon}$ are obtained from Table 1 of Shapiro and Wilk (1965) if $n \leq 50$. If $n > 50$, these estimates can be obtained using the results of Shapiro and Francia (1972) and Royston (1982). [See Srivastava and Hui (1987, p.16)]. Another possible test for univariate normality is the one proposed by Epps and Pulley (1983).

The tests based on skewness and kurtosis are described in D'Agostino et al. (1990). The computational steps are reproduced below. Denote the $N = np$ observations in (2.5) by $U_1, \ldots, U_N$ and denote by m_r the r-th corrected moment

$$m_r = \frac{1}{N}\sum_{i=1}^{N}(U_i - \bar{U})^r,$$

where $\bar{U}$ is the average of $U_1, \ldots, U_N$.

For skewness test, compute, $\sqrt{b_1} = m_2^{\frac{3}{2}}$,

$$Y = \sqrt{b_1}\left[\frac{(N+1)(N+3)}{6(N-2)}\right]^{\frac{1}{2}}, \ \beta_2(\sqrt{b_1}) = \frac{3(N^2+27N-70)(N+1)(N+3)}{(N-2)(N+5)(N+7)(N+9)},$$

$$W^2 = -1 + \left[2(\beta_2(\sqrt{b_1}) - 1)\right]^{-\frac{1}{2}}, \ \delta = (\ell n W)^{-\frac{1}{2}}, \ \alpha = \{2/(W^2-1)\}^{\frac{1}{2}},$$

$$Z(\sqrt{b_1}) = \delta \ell n\{Y/\alpha + [(Y/\alpha)^2 + 1]^{\frac{1}{2}}\} \sim N(0,1). \tag{2.7}$$

The test statistic for assessing skewness is (2.7), which can be referred to a normal table for obtaining the p-value.

For kurtosis test compute, $b_2 = m_4/m_2^2$

$$E(b_2) = \frac{3(N-1)}{N+1}, \ V(b_2) = \frac{24N(N-2)(N-3)}{(N+1)^2(N+3)(N+5)},$$

$$\sqrt{\beta_1(b_2)} = \frac{6(N^2-5N+2)}{(N+7)(N+9)}\left[\frac{6(N+3)(N+5)}{N(N-2)(N-3)}\right]^{\frac{1}{2}}$$

$$x = \frac{b_2 - E(b_2)}{\sqrt{V(b_2)}}, \ A = 6 + \frac{8}{\sqrt{\beta_1(b_2)}}\left[\frac{2}{\sqrt{\beta_1(b_2)}} + \left(1 + \frac{4}{\beta_1(b_2)}\right)^{\frac{1}{2}}\right],$$

$$Z(b_2) = \left[\left(1 - \frac{2}{9A}\right) - \left(\frac{1 - 2/A}{1 + x\sqrt{2/(A-4)}}\right)^{\frac{1}{3}}\right] / \sqrt{2/9A}. \tag{2.8}$$

Test for kurtosis can be carried out by referring $Z(b_2)$ to a normal table with mean zero and variance unity.

A combined test for skewness and kurtosis is to use

$$K^2 = Z^2(\sqrt{b_1}) + Z^2(b_2) \tag{2.9}$$

as chisquare on two degrees of freedom. The tests based on the N observations (2.5) are denoted by T_1 in the Tables 1, 2 and 3 in Section 3 of the paper.

3. Other transformations to independent variables

It is well known that if μ and Σ are known, the p-vector variable $X \sim N_p(\mu, \Sigma)$ can be transformed to $Y = N_p(0, I)$ by an affine transformation

$$Y = A(X - \mu) \tag{3.1}$$

where A is any $p \times p$ matrix satisfying the condition $A\Sigma A' = I$. The choice of A is not unique. One choice we made in Section 2 of the paper is $A = \Sigma^{-\frac{1}{2}}$, a symmetric inverse square root of Σ. Other choices can be made such as $A = \Lambda^{-1}P'$ where Λ and P are the matrices appearing in the spectral decomposition of $\Sigma = P\Lambda^2 P'$. Since μ and Σ are not known, we estimate μ and Σ by $\bar{X}$ and S from the sample and compute

$$Y_i = \Lambda^{-1}P'(X_i - \bar{X}) \tag{3.2}$$

where $P\Lambda^2 P'$ is the spectral decomposition of S, using the same notation for sample estimates of P and Λ. (See Srivastava and Hui (1987)). We may then consider the $N = np$ components of the vector

$$Y' = (Y_1' : \cdots : Y_n') \tag{3.3}$$

and apply the UVN test as indicated in Section 2. The test based on the N observations (3.3) is denoted by T_2 in Tables 1, 2 and 3.

Table 1 gives the p-values for the tests based on T_1 and T_2 on Fisher's data on Iris Setosa with $p = 4$ and $n = 50$. The tests were carried out on the original observations and also after transforming one of the measurements (petal width to logarithms).

Table 2 gives the p-values for tests based on T_1 and T_2 on the angular data on English and Naqada skulls. Only two angles are considered to test for bivariate normality. Table 3 referes to haemotology data considered by Royston (1983) in his discussion on tests of normality.

TABLE 1. p-values for tests T_1 and T_2 on Iris data
$n = 50, \ p = 4, \ N = 200$

	Original Measurements		After Transformation	
	T_1	T_2	T_1	T_2
Skewness $\sqrt{b_1}$	0.0674	0.2471	0.9937	0.8681
Kurtosis (b_2)	0.0311	0.0938	0.1453	0.1829
Combined	0.0184	0.1258	0.3462	0.4062

TABLE 2. p-values for tests T_1 and T_2 on skull measurements
Aitchinson (1986, p.385) $n = 51, \ p = 2, \ N = 102$

	English data		Naqada data	
	T_1	T_2	T_1	T_2
Skewness $\sqrt{b_1}$	0.8942	0.8077	0.1429	0.2124
Kurtosis (b_2)	0.4587	0.6751	0.5735	0.7863
Combined	0.7532	0.8892	0.2917	0.4430

TABLE 3. p-values for tests T_1 and T_2 on haemotology data
Royston (1983), $n = 103, \ p = 3, \ N = 309$

	Measurements on variables 1, 2, 5		Measurements on variables 2, 3, 5	
	T_1	T_2	T_1	T_2
Skewness $\sqrt{b_1}$	0.0090	0.0087	0.0404	0.0636
Kurtosis (b_2)	0.0632	0.0267	0.2773	0.3369
Combined	0.0059	0.0027	0.0678	0.1129

The test proposed in the paper, especially the one based on T_1 is simple to apply and is likely to be more sensitive to departures from MVN. The test can be included in the software for multivariate methods as an initial test to assess the suitability of MVN as a stochastic model for inferential analysis on observed data.

Studies on power comparison of T_1 with other tests for MVN and some theoretical investigations on the convergence of the sequence of random variables $Y_i, i = 1, \ldots, n$, defined in (2.5) will be reported in a future communication.

References

[1] Aitchinson, J. (1986). *The Statistical Analysis of Compositional Data*, Chapman and Hall.

[2] Andrews, D.F., Gnanadesikan, R. and Warner, J.L. (1971). Transformations of multivariate data. *Biometrics*, **27**, 825-840.

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

[4] Cox, D.R. and Small, N.J. (1978). Testing multivariate normality. *Biometrika*, **65**, 263-272.

[5] D'Agostino, Ralph R.B., Belanger, A. and D'Agostino, Ralph R.B., Jr. (1990). A suggestion for using powerful and informative tests of normality. *The American Statistician*, **44**, 316-321.

[6] Epps, T.W. and Pulley, L.B. (1983). A test for normality based on the empirical characteristic function. *Biometrika* **70**, 723-726.

[7] Henze, N. and Wagner, T. (1997). A new approach to the BHEP tests for multivariate normality. *J. Multivariate Analysis* **62**, 1-23.

[8] Jobson, J.D. (1991). *Applied Multivariate Data Analysis*, Vol II, Springer-Verlag, New York.

[9] Johnson, N.L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika* **36**, 149-176.

[10] Mardia, K.V. (1980). Tests of univariate and multivariate normality. In *Handbook of Statistics*, Vol 1, (Ed. P.R. Krishnaiah), 279-320.

[11] Rao, C.R. (1997). *Statistics and Truth: Putting Chance to Work* (Second Edition), World Scientific, Singapore and New Jersey.

[12] Royston, J.P. (1982). An extension of Shapiro and Wilk's W test for normality to large samples. *Applied Statistics* **31**, 115-124.

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

[14] Shapiro, S.S. and Francia, R.S. (1972). An approximate analysis of variance test for normality. *Journal of the American Statistical Association* **67**, 215-216.

[15] Shapiro, S.S. and Wilk, R.S. (1965). An analysis of variance test for normality (complete samples). *Biometrika* **52**, 591-611.

[16] Siotani, M., Hayakawa, T. and Fujikoshi, Y. (1985). *Modern Statistical Analysis: A Graduate Course and Handbook*. American Sciences Press, Inc., Columbus, Ohio.

[17] Srivastava, M.S. and Hui, T.K. (1987). On assessing multivariate normality based on Shapiro-Wilk statistic. *Statistics and Probability Letters* **5**, 11-18.

Department of Statistics, Pennsylvania State University, University Park, PA 16802

Department of Mathematics and Computer Science, The University of West Indies, St. Augustine, Trinidad and Tobago