A Simple Chi-squared Test of Multivariate Normality Based on the Spherical Data[1]

Cheolyong Park[2]

Abstract

We provide a simple chi-squared test of multivariate normality based on rectangular cells on the spherical data. This test is simple since it is a direct extension of the univariate chi-squared test to multivariate case and the expected cell counts are easily computed. We derive the limiting distribution of the chi-squared statistic via the conditional limit theorems. We study the accuracy in finite samples of the limiting distribution and then compare the power of our test with those of other popular tests in an application to a real data.

Keywords : Conditional limit theorems, Geyser data, Rectangular cells.

1. Introduction

In the classical univariate chi-square test of normality, we first partition the real line into a fixed number of intervals and calculate the frequencies of observations falling into the intervals. We then compare the frequencies with those expected from the normal distribution with fixed mean and variance values. If mean and variance of normal distribution are not known in advance, we can first standardize the observation, so that the sample mean is zero and sample covariance is one, and then test the standardized observations are from the standard normal distribution. The boundaries of these intervals are fixed in terms of the standardized observations but are random in terms of the original variables. Watson (1957) considered these random intervals and derived the limiting distribution of the chi-squared test statistic.

We can easily extend the univariate chi-squared test for normality to p-dimensional observations. In this case, we can partition the Euclidean space of p-dimensional real numbers into rectangular cells, which are just products of p intervals in real line. Moore and

Spruill (1975) derived the limiting distributions of various chi-squared test statistics based on rectangular cells. However, expected cell counts are not easy to compute since the covariance matrix of multivariate normal distribution is not a diagonal matrix but can be arbitrary. We can overcome this difficulty by first spherizing the data, so that the resulting transformed data has sample mean vector zero and sample covariance matrix identity, and then testing if the transformed data is from the standard multivariate normal distribution. If we employ rectangular cells as a partition of the multivariate real space, the chi-squared statistic is essentially testing the independence of the variables of the transformed data under multivariate normality.

This procedure is easy to comprehend since it is a direct multivariate extension of the univariate case. This is easy to compute since all our procedures are spherizing the data and then counting the numbers of observations falling into rectangular cells. The expected probability that the transformed data belongs to a rectangular cell is easily computed since it is asymptotically the product of the probabilities that each variable of a standard multivariate normal vector belongs to the corresponding interval of the rectangular cell.

The rectangular cells on the transformed data are not rectangular in terms of the original data and so we can not apply the Moore and Spruill's (1975) result to derive the limiting distribution of our chi-squared test statistic. Instead, Pollard's (1979) result can be used to derive its limiting distribution since it can be applied to more general data dependent cells. However, it is not so easy to check the regularity conditions of Pollard (1979) and so we will provide an easier way of proof based on the conditional limit theorem by Holst (1981) and Park (1995) in deriving the limiting distribution.

In Section 2, we will explain our method more in detail and then provide main results with proofs. In Section 3, we will provide a simulation study to check the accuracy in finite samples of the limiting distribution and then provide an illustrative example of application to a real data, in which we compare the power of our method with those of other popular methods.

2. Our Methods and Main Results

First, some brief remarks on notation. We will use I, e, and 0 to denote an identity matrix, a column vector of ones, and a column vector or matrix of zeros respectively. The dimensions will usually be clear from context, but will be specified by subscripts if necessary. Unless otherwise noted vectors will be column vectors, but for convenience they will be written in text as row vectors.

We will denote the p-dimensional multivariate distribution of mean vector μ and covariance matrix Σ by $N_p(\mu, \Sigma)$ and the chi-squared distribution with the f degrees of freedom by $\chi^2(f)$. We will denote the probability density function and cumulative distribution function of the (univariate) standard normal distribution by ϕ and Φ, respectively.

We will now explain our method in more details. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from $N_p(\mu, \Sigma)$ with nonsingular Σ. Let $\theta = (\mu, \Sigma)$ be the parameter of the distribution and let $\theta_0 = (0, I)$ be a special value of the parameter when $\mu = 0$, $\Sigma = I$. Then $\overline{Y} = \sum_{i=1}^{n} Y_i / n$ and $S = \sum_{i=1}^{n} (Y_i - \overline{Y})(Y_i - \overline{Y})^t / n$ are the maximum likelihood estimators of θ, where 't' denotes transpose. Here we assume that $n > p$, so that S is nonsingular.

We first spherize the original data so that the sample mean vector is zero and the sample covariance matrix is identity. In other words, the transformed data is obtained by

$$Z_i = R(S)(Y_i - \overline{Y}) \quad (i = 1, 2, \ldots, n)$$

where $R(S)$ is the lower triangular matrix with positive diagonal elements such that $R(S) S R^t(S) = I$. The quantity $R(S)$ is the function of the original data only through the sample covariance matrix S. There are many other choices of $R(S)$ that can also spherize the original data; for example, the rotation based on the principal components and the rotation using the square root matrix $R(S) = S^{-1/2}$ are two popular choices. We choose the method based on lower triangular matrix since the ancillarity of the transformed data based on low triangular matrix is easily verified whereas it is hard to verify for other procedures. In the proofs of main results, we can find that the ancillarity of the transformed data is essential for deriving the limiting distribution of the chi-squared test statistic via the conditional limit theorems.

We next obtain a grouped data by binning the transformed data based on the rectangular cells. In order to bin the transformed data based on the rectangular cells, we need to set the boundaries of intervals for each coordinate (or variable) of the transformed data. Here, different numbers of boundaries could be set for each coordinate but the boundaries must be set such that adjacent boundaries form equiprobable intervals; i.e., when the number of boundaries is $d + 1$, the boundaries are chosen as the quantiles of the standard normal distribution corresponding to the probabilities $0/d, \ldots, d/d$ and thus adjacent boundaries form intervals with equal probability $1/d$. Let d_i $(i = 1, \ldots, p)$ be the number of equiprobable intervals for the i-th coordinate and let $\xi_p = \Phi^{-1}(p)$ be the p-th quantile of the standard normal distribution. Then the intervals of the i-th coordinate are given by $(\xi_{(k-1)/d_i}, \xi_{k/d_i}]$ $(k = 1, 2, \ldots, d_i)$ and the k-th group will be assigned if the i-th coordinate of an transformed vector belongs to $(\xi_{(k-1)/d_i}, \xi_{k/d_i}]$.

We now form a contingency table from the grouped data. This contingency table contains $K = \prod_{i=1}^{p} d_i$ cells and n observations are distributed among these K cells. We use

$\pi = (\pi_1, \pi_2, \ldots, \pi_p)$, with $1 \le \pi_i \le d_i$ for each i, to denote a particular cell in our table. For each cell π, the cell count $u_{n\pi}$ is given by

$$u_{n\pi} = \sum_{k=1}^{n} I\{\xi_{(\pi_i - 1)/d_i} < z_{ki} \le \xi_{\pi_i/d_i}, \text{ for } 1 \le i \le p\},$$

where $I\{\cdot\}$ is an indicator function and z_{ki} is the i-th coordinate of the k-th transformed vector Z_k. Note that the cell count $u_{n\pi}$ is a function of the transformed data and will be explicitly denoted by $u_{n\pi}(Z_1, \ldots, Z_n)$ whenever necessary. As a measure for the degree of departure from multivariate normality, we use the chi-squared statistic X^2 defined by

$$X^2 = \sum_{\pi} \frac{(u_{n\pi} - np_0)^2}{np_0},$$

where $p_0 = K^{-1} = (\prod_{i=1}^{p} d_i)^{-1}$ is an asymptotic probability of belonging to a particular cell. Note that this is the usual chi-squared test statistic for testing total independence in a multi-way contingency table and that, under multivariate normality, the expected number of observations in any given cell is approximately np_0.

The choice of d_i's in our procedure is somewhat arbitrary. We generally prefer to have as many cells as possible without allowing the average cell count np_0 to be too small. If we wish to use the limiting distribution of the chi-squared statistic for testing purposes, the usual guidelines apply: the limiting distribution is fairly accurate when $np_0 \ge 5$. Our simulation study in next section shows that if the number of cells is large enough, it is reasonably good even for $np_0 = 1$. Since the number of cells grows rapidly with p, for high dimensional data sets we are often forced to use small values of d_i's in order to avoid extremely small average cell counts.

We now present the limiting distribution of the vector of cell counts $u_{n\pi}$ and of the chi-squared statistic X^2 under multivariate normality. First, we introduce various matrices which are need in the statements of our results. Let $U_n = (u_{n\pi})$ be the $K \times 1$ vector of cell counts. Unless otherwise specified, the vector is a function of the transformed data and will be explicitly denoted by $U_n(Z_1, \ldots, Z_n)$ whenever necessary. For easy representation of results, we assume the elements of the U_n are arranged in such a way that the corresponding cell vectors $\pi = (\pi_1, \pi_2, \ldots, \pi_p)$ are in a standard order; i.e. the first coordinate π_1 changes form 1 to d_1 the fastest, the second coordinate π_2 changes the second fastest, and so on.

Now, for $i=1,\ldots,p$; $j=1,\ldots,d_i$, define $\zeta_{ij}=\Phi^{-1}(j/d_i)$ and

$$\psi_{ij}=\phi(\zeta_{i(j-1)})-\phi(\zeta_{ij}), \qquad \omega_{ij}=\zeta_{i(j-1)}\phi(\zeta_{i(j-1)})-\zeta_{ij}\phi(\zeta_{ij}),$$

with the convention $\pm\infty\phi(\pm\infty)\equiv0$. We define D_1 to be a $K\times p$ matrix the i-th column of which is $\prod\limits_{j>i}^{p}d_j$ repetitions of the vector $'(d_i\psi_{i1}e_i,\ d_i\psi_{i2}e_i,\ \ldots,\ d_i\psi_{id_i}e_i)$, where e_i is the vector of $\prod\limits_{j<i}^{p}d_j$ ones, with the convention that $\prod\limits_{j>p}^{p}d_j=\prod\limits_{j<1}^{p}d_j\equiv1$. Also we define D_2 to be the same matrix as D_1 except that the ψ_{ij}'s are replaced by ω_{ij}'s. Let D_3 be the $K\times(p(p-1)/2)$ matrix obtained from D_1 such that the $p(p-1)/2$ columns of D_3 are all the possible products of two distinct columns from D_1.

Here we will present and derive the asymptotic joint distribution of the vector U_n of cell counts and the chi-squared test statistic X^2 in Theorems 1 and 2, respectively.

Theorem 1. If $Y_1,Y_2,\ldots,Y_n$ is a random sample from $N_p(\mu,\Sigma)$ where Σ is nonsingular, then

$$n^{-1/2}(U_n-np_0e)\ \xrightarrow{d}\ N_K(0,\Psi)\quad\text{as }n\to\infty$$

where

$$\Psi=p_0I-p_0^2ee^t-p_0^2D_1D_1^t-\frac{p_0^2}{2}D_2D_2^t-p_0^2D_3D_3^t.$$

Proof: When we obtain the transformed data $Z_1,\ldots,Z_n$ by spherizing the original data $Y_1,\ldots,Y_n$, we use the lower triangular matrix $R(S)$ with positive diagonal elements. Therefore, by Theorem 1 of Park (1998), the vector $U_n=U_n(Z_1,\ldots,Z_n)$ of cell counts are ancillary and thus independent of a sufficient statistic $(\overline{Y},S)$ for $\theta=(\mu,\Sigma)$. Therefore, for any θ, we have

$$\begin{aligned}\mathcal{L}_\theta(U_n(Z_1,\ldots,Z_n))&=\mathcal{L}_{\theta_0}(U_n(Z_1,\ldots,Z_n))\\&=\mathcal{L}_{\theta_0}(U_n(Z_1,\ldots,Z_n)|\overline{Y}=0,S=I)=\mathcal{L}_{\theta_0}(U_n(Y_1,\ldots,Y_n)|\overline{Y}=0,S=I)\end{aligned} \tag{1}$$

where $\theta_0=(0,I)$ is a fixed parameter. The last equality of (1) holds since, by the relation $R(I)=I$, the transformed vector $Z_i=R(S)(Y_i-\overline{Y})$ is equal to Y_i for each i, given that $\overline{Y}=0,S=I$. In order to apply the conditional theorem by Park(1995), we need to express the condition $\{\overline{Y}=0,S=I\}$ in terms of the canonical sufficient statistic for $N_p(\mu,\Sigma)$. We now define some notations: For any p-vector $y=(y_1,\ldots,y_p)$, we define column vectors

$s(y) = (y, d(y), r(y))$ where $d(y) = (y_1^2, ..., y_p^2)$, $r(y) = (y_1 y_2, ..., y_{p-1} y_p)$. With these notations, the canonical sufficient statistic is given by $\sum_{i=1}^{n} s(Y_i)$ and the condition $\{\overline{Y} = 0, S = I\}$ is equivalent to $\{\sum_i s(X_i)/n = (0_p, e_p, 0_{p(p-1)/2})\}$. Therefore, the equation (1) becomes

$$\mathcal{L}_{\theta_0}\left(U_n(X_1, ..., X_n) \mid \sum_{i=1}^{n} s(Y_i)/n = (0_p, 1_p, 0_{p(p-1)/2})\right).$$

By applying almost the same arguments as in Theorem 2.1 of Park (1999), it is easy to verify the assumptions of Corollary 1 of Park (1995) and thus we have

$$n^{-1/2}(U_n - n p_0 e) \xrightarrow{d} N_K(0, A - BC^{-1}B^t) \quad \text{as } n \to \infty,$$

where

$$A = \text{Cov}_{\theta_0}(U_1(Y_1)), \quad B = \text{Cov}_{\theta_0}(U_1(Y_1), s(Y_1)), \quad C = \text{Cov}_{\theta_0}(s(Y_1)).$$

It is easy to see that

$$A = p_0 I - p_0^2 e e^t, \quad B = p_0(D_1, D_2, D_3), \quad C = \text{diag}(I_p, 2I_p, I_{p(p-1)/2}).$$

and thus

$$\Psi = A - BC^{-1}B^t = p_0 I - p_0^2 e e^t - p_0^2 D_1 D_1^t - \frac{p_0^2}{2} D_2 D_2^t - p_0^2 D_3 D_3^t,$$

which completes proof.   □

Theorem 2. Under the assumption of Theorem 1,

$$X^2 \xrightarrow{d} W_1 + \sum_{i=1}^{p}(1 - d_i a_i) W_{2i} + \sum_{i=1}^{p}(1 - d_i b_i/2) W_{3i} + \sum_{i=1}^{p}\sum_{j>i}^{p}(1 - d_i d_j a_i a_j) W_{4ij} \quad \text{as } n \to \infty$$

where W_1 is $\chi^2(K - 1 - 2p - p(p-1)/2)$ distributed, $\{W_{2i}\}, \{W_{3i}\}, \{W_{4ij}\}$ are $\chi^2(1)$ distributed, and all of the chi-squared variates are independent with

$$a_i = \sum_{j=1}^{d_i} \psi_{ij}^2, \quad b_i = \sum_{j=1}^{d_i} \omega_{ij}^2,$$

$K = \prod_{i=1}^{p} d_i$ and ψ_{ij}'s & ω_{ij}'s defined in Theorem 1.

Proof: Note that $X^2 = (U_n - n p_0 e)^t (U_n - n p_0 e)/(n p_0)$. Therefore the limiting distribution of X^2 is given by the form $\sum_{i=1}^{K} \lambda_i W_i$ where λ_i's are the eigenvalues of

$$E = \Psi/p_0 = I - p_0 e e^t - p_0 D_1 D_1^t - \frac{p_0}{2} D_2 D_2^t - p_0 D_3 D_3^t$$

and W_i's are a random sample from $\chi^2(1)$ distribution.

By simple algebra, we can easily show that

$$\sum_{j=1}^{d_i} \psi_{ij} = \sum_{j=1}^{d_i} \omega_{ij} = \sum_{j=1}^{d_i} \psi_{ij} \omega_{ij} = 0 \tag{2}$$

for each i, where ψ_{ij}'s and ω_{ij}'s are defined in Theorem 1. From (2), we have

$$e^t D_1 = 0, \quad e^t D_2 = 0, \quad e^t D_3 = 0, \quad D_1^t D_2 = 0, \quad D_1^t D_3 = 0, \quad D_2^t D_3 = 0$$

and thus $ee^t, D_1 D_1^t, D_2 D_2^t,$ and $D_3 D_3^t$ are orthogonal to each other.

Also from (2), it is easy to show that

$$e^t e = K = 1/p_0, \quad D_1^t D_1 = \mathrm{diag}(d_1 a_1, \dots, d_p a_p)/p_0,$$

and

$$D_2^t D_2 = \mathrm{diag}(d_1 b_1, \dots, d_p b_p)/p_0, \quad D_3^t D_3 = \mathrm{diag}(d_1 d_2 a_1 a_2, \dots, d_{p-1} d_p a_{p-1} a_p)/p_0.$$

Therefore the eigenvalues of E are 1 with multiplicity $K - 1 - 2p - p(p-1)/2$,

$$d_1 a_1, \dots, d_p, \quad d_1 a_1/2, \dots, d_p a_p/2, \quad d_1 d_2 a_1 a_2, \dots, d_{p-1} d_p a_{p-1} a_p, \quad \text{and } 0$$

with multiplicity 1. This completes the proof. $\qquad\square$

When our method is applied to real data, the grouped data is usually obtained by using the same grouping rule for each coordinate of the transformed data unless some prior information is known that some coordinates need closer examination than others. In this case, we have $d = d_1 = \dots = d_p$ and we can further simplify the limiting distribution of X^2 as follows:

Corollary 1. If $d = d_1 = \dots = d_p$ is satisfied in addition to the assumption of Theorem 1, then

$$X^2 \xrightarrow{d} W_1 + (1 - da)W_2 + (1 - db/2)W_3 + (1 - d^2 a^2)W_4 \quad \text{as } n \to \infty$$

where $W_1, W_2, W_3,$ and W_4 are independent chi-squared variates with degrees of freedom

$$\nu_1 = K - 1 - 2p - p(p-1)/2, \quad \nu_2 = p, \quad \nu_3 = p, \quad \text{and } \nu_4 = p(p-1)/2$$

respectively, with $a = \sum_{j=1}^{d} \psi_{1j}^2$ and $b = \sum_{j=1}^{d} \omega_{1j}^2$.

Proof: If $d = d_1 = \dots = d_p$, then we have

$$a = a_1 = \dots = a_p = \sum_{j=1}^{d} \psi_{1j}^2, \quad b = b_1 = \dots = b_p = \sum_{j=1}^{d} \omega_{1j}^2.$$

Since $\{W_{2i}\}, \{W_{3i}\}, \{W_{4ij}\}$ are independent and $\chi^2(1)$ distributed, the variates

$$W_2 = \sum_{i=1}^{p} W_{2i}, \quad W_3 = \sum_{i=1}^{p} W_{3i}, \quad \text{and } W_4 = \sum_{i=1}^{p} \sum_{j>i}^{p} W_{4ij}$$

are chi-squared distributed with degrees of freedom p, p, and $p(p-1)/2$, respectively. This completes the proof. $\qquad\square$

The limiting distributions in Theorem 1 and Corollary 1 are not exact chi-squared distributions since we use the estimators $\overline{Y}$, S, based on the original data, for μ, Σ. This have been well known since the work of Chernoff and Lehmann (1954). In order to achieve an exact chi-squared limiting distribution, we might use the Rao-Robson chi-squared test statistic of the form $(U-np_0e)'\Psi^-(U-np_0e)$ (see, Rao and Robson (1974) and Nikulin (1973) for details) but we will not pursue this topic further in this paper.

3. Simulation and Application

We first provide a small simulation study to check accuracy in finite samples of the limiting distribution of our chi-squared test statistic. We consider the case where $d=d_1=\cdots=d_p$ since it will be usually employed in practice as noted just prior to Corollary 1. We have tried many configurations but almost all of them show the identical results unless the number K of cells is small. Therefore we present the results for the case where $p=3$, $d=3$ so that the limiting distribution of the chi-square statistic is

$$\chi^2(17)+0.207\chi^2(3)+0.279\chi^2(3)+0.790\chi^2(3).$$

We consider four different sample sizes, $n=27,54,125$ and 270, which have an average of 1, 2, 5 and 10 observations per cell, respectively.

For each sample size n, we generate 500 samples of size n from $N_3(0,I)$ and then calculate the chi-squared statistics for each of them. These 500 values are plotted against the "expected order statistics" of a sample of size 500 from the limiting distribution. The expected order statistics are generated from the limiting distribution as follows: We generate 100 samples of size 500 from the weighted chi-squared distribution and average the order statistics of these 100 samples. The resulting quantile-quantile plots are displayed in Figure 1.

Each plot displays the reference line with slope 1 and intercept 0, which corresponds to the ideal case where empirical and theoretical distributions coincide. Examining the plots, we see that the limiting distribution is a good approximation for the cases where average cell counts np_0 is 2, 5 and 10. However the discreteness of the chi-squared values are apparent in the case where np_0 is 1 but the points do not deviate much from the reference line even in this case.

We next provide an illustrative example of application to geyser data, in which we compare the power of our method with those of other popular methods. The three competitors we use for testing multivariate normality are the skewness and kurtosis tests of Mardia (1970) and the Q_n test (with Cholesky implementation) of Ozturk and Romeu (1992). We have chosen these tests since they performed quite well in an extensive simulation study of Romeu and

Figure 1. Quantile-quantile plots for four different sample sizes
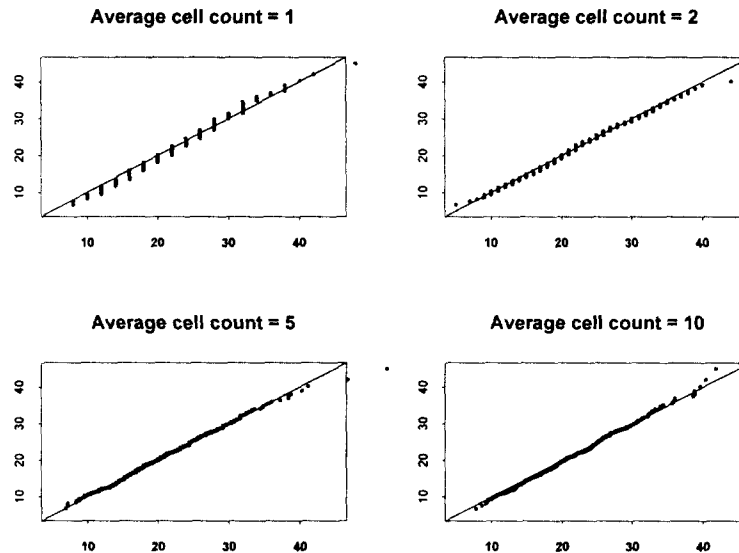
Ozturk (1993). There are two time series in geyser data and the waiting time between eruptions is used for our example. We will examine the normality of the residuals from fitting a time series model to the waiting time.

We use an automatic procedure called **AR** in S-Plus to select one of the best autoregressive models based on the Akaike information criterion. By using Yule-Walker equations to estimate the autoregression coefficients, the procedure have chosen AR(2) model. Both autocorrelation and partial autocorrelation functions are well inside the error bars up to 25 lags and the value of Shapiro and Wilk (1965) W test statistic is 0.9822 with p-value 0.3654. Therefore we could not find unusual pattern by the usual time series diagnostics and a univariate test of normality.

We now examine the residuals by our method and other competitors for multivariate normality. To obtain multivariate observations, we divide the residuals into subseries of three consecutive residuals and take each subseries as an observation. In this way, we obtain 99 multivariate observations. Our method with $d=3$ leads to chi-squared value 33.27 with z-score 1.96 and $d=4$ leads to chi-squared value 87.83 with z-score 2.92. Therefore, our method signals there are some deviations from normality in the residuals. The 'lag 1' plot reveals some tendency that the variance of the residual at the present time t increases with the residual at lag 1 time $t-1$. However, other competitors could not detect this structure: Marida's skewness and kurtosis have (asymptotic) p-values 0.160 and 0.569, respectively and the test based on Q_n has (asymptotic) p-value 0.601. Therefore, in this example, our test has more power for detecting deviations from normality than other tests.

References

[1] Chernoff, H., and Lehmann, E.L. (1954). The Use of Maximum Likelihood Estimates in χ^2 Tests for Goodness of Fit. *Annals of Mathematical Statistics* 25, 579-586.

[2] Holst, L. (1981). Some Conditional Limit Theorems in Exponential Families. *Annals of Probability* 9, 818-30.

[3] Mardia, K.V. (1970). Measures of Multivariate Skewness and Kurtosis with Applications. *Biometrika* 57, 519-530.

[4] Moore, D.S, and Spruill, M.C. (1975). Unified Large-sample Theory of General Chi-squared Statistics for Tests of Fit. *Annals of Statistics* 3, 599-616.

[5] Nikulin, M.S. (1973). Chi-Square Test for Continuous Distributions with Shift and Scale Parameters. *Theory of Probability and Its Applications* 18, 559-568.

[6] Ozturk, A. and Romeu, J.L. (1992). A New Method for Assessing Multivariate Normality with Graphical Applications. *Communications in Statistics - Simulations and Computation* 21, 15-34.

[7] Park, C. (1995). Some Remarks on the Chi-squared Test with Both Margins Fixed. *Communications in Statistics - Theory and Methods* 24, 653-61.

[8] Park, C. (1998). On the Distributions of the Scaled Residuals under Multivariate Normal Distributions. *The Korean Communications in Statistics* 5, 591-97.

[9] Park, C. (1999). A Note on the Chi-Square Test for Multivariate Normality Based on the Sample Mahalanobis Distances. *Journal of the Korean Statistical Society* 28, 479-488.

[10] Pollard, D. (1979). General Chi-square Goodness-of-fit Tests with Data Dependent Cells. *Z. Wahrsch. verw. Gebiete* 50, 317-32.

[11] Rao, C.R., and Robson, D.S. (1974). A Chi-Square Statistic for Goodness-of-Fit Tests within the Exponential Family. *Communications in Statistics* 3, 1139-1153.

[12] Romeu, J.L. and Ozturk, A. (1993). A Comparative Study of Goodness-of-fit Tests for Multivariate Normality. *Journal of Multivariate Analysis* 46, 309-34.

[13] Shapiro, S.S. and Wilk, M.B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* 52, 591-611.

[14] Watson, G.S. (1957). The χ^2 Goodness-of-fit Test for Normal Distributions. *Biometrika* 44, 336-48.