Centro de Estudios Económicos
www.colmex.mx
El Colegio de México, A.C.

Serie documentos de trabajo

TESTS FOR MULTIVARIATE NORMALITY OF OBSERVATIONS AND RESIDUALS

Carlos M. Urzúa

DOCUMENTO DE TRABAJO

Núm. III - 1989

TESTS FOR MULTIVARIATE NORMALITY OF OBSERVATIONS AND RESIDUALS

BY

Carlos M. Urzúa

DOCUMENTO DE TRABAJO

No. III - 89

February 1989
(Revision: April 1989)

Address correspondence to:

Centro de Estudios Económicos
El Colegio de México
Camino al Ajusco No. 20
México, D.F.10740
MEXICO

TESTS FOR MULTIVARIATE NORMALITY OF OBSERVATIONS AND RESIDUALS

Carlos M. Urzúa

El Colegio de Mexico
and
Woodrow Wilson School, Princeton University

This paper provides readily computable tests for multivariate normality of both observations and residuals of simultaneous equation models. They are derived by considering as the alternatives to the multinormal a class of "likely" multivariate distributions studied elsewhere by the author. The tests, being derived using the Lagrange multiplier procedure, have optimum asymptotic power. Furthermore, they include, in the case of a single variable, the popular Jarque-Bera (Bowman-Shenton) test for univariate normality.

1. INTRODUCTION

Since the pioneering work earlier this century by, among others, K. Pearson, R. Fisher and J. Wishart, the assumption of multivariate normality has played a key role in many methods of multivariate analysis. Handy as the assumption is, however, the consequences of departure from multinormality are documented to be quite serious for several multivariate methods (e.g., linear discriminant analysis). The judgment still awaits for more evidence in the case of other methods, simultaneous equation models being a case in point, but, in principle, the consequences could be serious as well. This can be surmised in cases such as simultaneous equation models, where the violation of the multinormality assumption may lead to inefficient estimators and invalid inferences.

Given the obvious importance of the multinormality assumption, it is thus somewhat surprising that for many years most researchers either ignored it, or were contented with the evaluation of marginal normality (which, of course, does not necessarily imply joint normality). One can argue that it was only until Mardia (1970) presented a simple test based on multivariate measures of skewness and kurtosis that the issue of testing for multinormality gained some favour among researchers. That this favour has grown since then can be attested by the burgeoning current literature on the subject (see the surveys by Mardia (1980) and Small (1985)).

The purpose of this paper is to provide readily computable tests for multivariate normality of both observations and residuals of simultaneous equation models. They are derived by considering as the alternatives to the multinormal a class of "likely" multivariate distributions introduced in Urzúa

(1988). The tests, by being derived using the Lagrange multiplier procedure, have optimum locally asymptotic power among those alternatives. Thus, they distinguish themselves from other ad-hoc tests in the literature that are simply patterned as the extensions of tests for univariate normality. This is not meant to deny the practical advantage of having multivariate tests with such property, for in fact the tests proposed here are the multivariate counterparts to the popular Jarque-Bera test for univariate normality (Jarque and Bera (1980, 1987)).

The paper is organized as follows: Section 2 reviews several of the properties that characterize the Q-exponential distributions, and presents some basic results for use in later sections. Section 3 derives the Lagrange multiplier (LM) test for multivariate normality under the premise that the alternatives to the multinormal are other Q-exponentials. It also provides, via a Montecarlo study, empirical significance points for the LM test statistics when the sample size is small. Section 4 then presents test statistics for multivariate normality of residuals of simultaneous equation models. Finally, Section 5 outlines several extensions to the results presented here (and that will appear elsewhere).

2. LIKELY ALTERNATIVES TO MULTIVARIATE NORMALITY

In his authoritative paper on significance tests written in the seventies, D. R. Cox complained about the inexistence of "a simple and general family of distributions to serve as alternatives [to the multinormal]" (1977, p. 56). This section reviews a class of multivariate distributions, taken from the multivariable exponential family and studied in Urzúa (1988), that could play that role.

The distributions are the multivariate generalization of the still now relatively unknown distributions introduced by R. A. Fisher (1922). Defined over the real line, Fisher's univariate densities are of the form

$$(2.1) \qquad f(x) = \tau(\alpha)\exp(-Q(x)), \quad Q(x) = \alpha_1 x+\alpha_2 x^2+\ldots+\alpha_k x^k,$$

where k is an even number, $\alpha_k > 0$, and $\tau(\alpha)$ is the constant of normalization given the vector of parameters α. Aside of course from the normal (obtained when k = 2), the densities in (2.1) were considered to be of little interest for many years. More recently, however, there has been an increasing interest on them since they play a key role in Cobb's stochastic catastrophe theory (see Urzúa (1989a) and references therein). Furthermore, as Zellner and Highfield (1988) have strikingly illustrated in the case of the quartic exponential (obtained setting k = 4 above), Fisher's distributions are flexible enough, and simple enough, to act as <u>bona fide</u> approximations to other univariate distributions.

We now turn to their multivariable counterparts. Let x denote the real column vector $(x_1,\ldots,x_p)'$. If Q(x) is a polynomial of degree k in the p variables, then it can always be written, ignoring the constant term, as

$$(2.2) \qquad Q(x) = \sum_{q=1}^{k} Q^{(q)}(x),$$

where $Q^{(q)}(x)$ is a homogeneous polynomial (a form) of degree q. Namely,

$$(2.3) \qquad Q^{(q)}(x) = \Sigma\, \alpha^{(q)}_{j_1\cdots j_p} \prod_{l=1}^{p} x_l^{j_l},$$

with the summation taken over all nonnegative integer p-tuples $(j_1,\ldots,j_p)$ such that $j_1+\ldots+j_p = q$. The polynomial Q will be assumed to be such that

$g(x) = \exp(-Q(x))$ is integrable on the entire Euclidean space R^p (a necessary condition for this to happen is that the degree of $Q(x)$ relative to each x_i is an even integer).

Following Urzúa (1988), the continuous random vector $X = (X_1,\ldots,X_p)$ is said to have a p-variate Q-exponential distribution with support R^p if its density is given by

$$f(x) = \tau(\alpha)\exp(-Q(x)), \quad -\infty < x_i < \infty, \qquad i = 1,\ldots,p, \tag{2.4}$$

where $\tau(\alpha)$ is the constant of normalization.

For the sake of simplicity, it will be often implicitly assumed below that the polynomial Q is of degree k relative to all of its components. In such a case, several important distributions emerge: If $k = 2$, then the p-variate normal is obtained; while when $k = 4$ and $k = 6$ the p-variate quartic and sextic exponentials are obtained.

Note also that as k is increased the number of coefficients required by the corresponding Q-exponential increases at an increasing rate. In fact, as can be readily shown (see Urzúa (1988, p. 4042)), if $K(p,k)$ denotes the maximum possible number of parameters of a p-variate Q-exponential, then

$$K(p,k) = C(p+k,k)-1, \tag{2.5}$$

where $C(p+k,k)$ is the binomial coefficient $(p+k)!/(p!k!)$. In particular, the number of possible coefficients in the homogeneous polynomial of degree q given in (2.3) above is $C(p+q-1,q)$.

It is now time to introduce a key characteristic of the Q-exponential distributions. Consider all densities f relative to Lebesgue measure that have support $S = R^p$ and have finite population moments of some predetermined orders. That is, they satisfy constraints of the form

$$(2.6) \qquad E\{\prod_{l=1}^{p} X_l^{j_{lm}}\} = c_m, \ m = 1,\ldots,r,$$

where each j_{lm} is a predetermined nonnegative integer, and $c_1,\ldots,c_r$ is a sequence of real numbers. For each density we define, following Shannon (1948), the entropy of f as

$$(2.7) \qquad H(f) = -\int_S f(x)\log[f(x)]dx.$$

It can be shown (see Urzúa (1988, Proposition 2)) that, among the densities satisfying (2.6), if there is a density that maximizes Shannon's entropy, it is necessarily a Q-exponential of the form

$$f(\chi) = \tau(\alpha)\exp(-Q(x)), \text{ with } Q(x) = \sum_{m=1}^{r} \alpha_m \prod_{l=1}^{p} X_l^{j_{lm}} .$$

For instance, the p-variate quartic exponential maximizes the entropy among the distributions with support R^p that are known to have finite moments up to order four; likewise, as Shannon (1948) in his influential paper first proved, the multinormal maximizes the entropy among the distributions that have second order moments.

Thus, when the only known information about a distribution is the existence of population moments of some orders, the Q-exponentials can be considered to be the "most likely to be true". This at least according to the maximum entropy principle, which states that "in making inferences on the basis of partial information we must use that probability distribution which has maximum entropy subject to whatever is known" (Jaynes (1957,p. 623)).

There is a second characteristic of the Q-exponentials that is also relevant for our purposes. It can be shown that, near the multinormal, already the quartic exponential is capable of approximating as close as needed the van Uven-Steyn multivariate Pearson family (see Urzúa (1988)). This is interesting because the latter family, although made of distributions more complex (and suspect) than the Q-exponentials, could be thought by some to constitute a class of possible alternatives to multinormality (in fact, Bera and John (1983) have used such a family to derive tests for multinormality).

Before concluding this section, it is worth briefly mentioning other interesting properties exhibited by the Q-exponential distributions that, although not directly relevant for this paper, help to illustrate furthermore the generality of the distributions (see Urzúa 1988 for details): First, they can exhibit several modes, and they do so with a relatively small number of parameters (as compared to mixtures of multinormals). Second, they are the stationary distributions of certain multivariate diffusion processes. Third, the Maximum Likelihood (ML) estimators of their population moments are the sample moments (as can be directly seen from equation (3.2) below). And fourth, using the method of moments one can easily obtain consistent estimators for the parameters of the Q-exponential distributions.

3. TESTS FOR NORMALITY OF OBSERVATIONS

Let X be a p×1 random vector following a Q-exponential distribution. Consider a set of n observations $\{x_1,\ldots,x_n\}$ on X. The corresponding log-likelihood function $L(\alpha)$ can then be easily shown to be

$$(3.1) \qquad L(\alpha) = -n\log[\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp(-Q(x))dx] - \sum_{r=1}^{n} Q(x_r).$$

Moreover, the components of the gradient (score) of $L(\alpha)$ are of the form

$$(3.2) \qquad \partial L/\partial\alpha^{(q)}_{j_1\ldots j_p} = nE\{\prod_{i=1}^{p} X_i^{j_i}\} - \sum_{r=1}^{n}\prod_{i=1}^{p} x_{ir}^{j_i},$$

while the elements of the Hessian of $L(\alpha)$ are of the form

$$(3.3) \qquad \partial^2 L/\partial\alpha^{(q)}_{j_1\ldots j_p}\,\partial\alpha^{(r)}_{k_1\ldots k_p} = -n[E\{\prod_{i=1}^{p} X_i^{j_i+k_i}\} - E\{\prod_{i=1}^{p} X_i^{j_i}\}E\{\prod_{i=1}^{p} X_i^{k_i}\}];$$

that is, Fisher's information matrix is simply made of covariances of products of the random components.

It will prove useful to transform the random vector X to a random vector Y having zero mean and the identity matrix as the covariance. Let μ and Σ be the mean vector and the covariance matrix of X. Let Γ denote the orthogonal matrix whose columns are the standardized eigenvectors of Σ, and Λ denote the diagonal matrix of the respective eigenvalues of Σ. Define $\Sigma^{-1/2}$ as the inverse of the square root decomposition of Σ; that is, $\Sigma^{-1/2} = \Gamma\Lambda^{-1/2}\Gamma'$. Then the random vector

$$(3.4) \qquad Y = \Sigma^{-1/2}(X-\mu)$$

follows a p-variate Q-exponential, with $Q(y)$ as in (2.2) and $Q^{(q)}(y)$ as in (2.3). It has a zero mean vector, and an identity matrix as its covariance matrix.

Let the $K(p,k)\times 1$ vector of parameters of $Q(y)$ be denoted as α where $K(p,k)$ is given in (2.5) above. Suppose now that α is partitioned as $\alpha = (\theta_1',\theta_2')'$, where θ_1 is the $C(p+1,2)\times 1$ vector of parameters of the homogeneous polynomial $Q^{(2)}(y)$. The hypothesis of multinormality can be then assessed by testing the null hypothesis H_0: $\theta_2 = 0$. There are several asymptotic tests available for that purpose (see, e.g., the survey by Godfrey 1988). Given the complexity of the alternatives considered here, the Lagrange Multiplier (LM) test (Rao 1948, and Aitchison and Silvey 1958) will be used below, for it only requires the estimation of the restricted model under the null hypothesis.

In order to give an expression for the LM statistic, it is necessary to introduce some notation. Let $s(\alpha)$ be the gradient (score) of the log-likelihood function, and let I be the information matrix. Given the partition of α as $(\theta_1',\theta_2')'$ the score can be written as $s(\alpha) = (s_1',s_2')'$, with $s_j = \partial L(\alpha)/\partial\theta_j$, $j = 1,2$; while the information matrix can be partitioned into four submatrices of the form $I_{ij} = E\{-\partial^2 L(\alpha)/\partial\theta_i\partial\theta_j'\}$, $i,j = 1,2$.

Let now $(\tilde{\theta}_1,0)$ denote the restricted maximum likelihood estimator for $\alpha = (\theta_1',\theta_2')'$; that is, $\tilde{\theta}_1$ is the maximum likelihood estimator for θ_1 after imposing the constraint $\theta_2 = 0$. Let also $\tilde{s} = s(\tilde{\theta}_1,0)$ and $\tilde{I} = I(\tilde{\theta}_1,0)$. Then the Lagrange multiplier statistic is defined as $LM = \tilde{s}\tilde{I}^{-1}\tilde{s}/n$, or, taking advantage of the fact that $\tilde{s}_1 = 0$,

$$(3.5) \qquad LM = \tilde{s}_2'(\tilde{I}_{22} - \tilde{I}_{21}\tilde{I}_{11}^{-1}\tilde{I}_{12})^{-1}\tilde{s}_2/n.$$

LM is under H_0 asymptotically distributed as a χ^2_ν, a Chi-square with degrees of freedom ν equal to the dimension of the vector θ_2.

Given the complexity of the alternatives to multinormality considered here, the computation of the LM statistic would appear to be a daunting task. However, as Gart and Tarone (1983) has shown, the LM test can be trivially found in the case of distributions coming from an exponential family (as the Q-exponentials do). Following their reasoning, it can be shown that the statistic can be constructed as follows. (For simplicity we will assume from now on that the quartic exponential is the alternative distribution with the highest degree, although the result can be trivially extended to the general case.)

First transform the original observations on X: Let $\bar{x}$ and S be the sample mean vector and the sample variance-covariance matrix found using the set of observations $\{x_1,\dots,x_n\}$. Let G denote the orthogonal matrix whose columns are the standardized eigenvectors of S, and D denote the diagonal matrix of the eigenvalues. Using $S^{-1/2} = GD^{-1/2}G'$, transform the observations as follows:

$$(3.6) \qquad y_r = S^{-1/2}(x_r-\bar{x}) \qquad r=1,\dots,n.$$

Define next

$$Q_{ijk} = \sum_{r=1}^{n} y_{ri}y_{rj}y_{rk}/n \text{ and } R_{ijkq} = \sum_{r=1}^{n} y_{ri}y_{rj}y_{rk}y_{rq}/n.$$

Then the LM test statistic is simply given by:

$$(3.7) \quad LM_p = n[\sum_{i=1}^{p} Q_{iii}^2/6 + \sum_{\substack{i,j=1\\ i\neq j}}^{p} Q_{iij}^2/2 + \sum_{\substack{i,j,k=1\\ i<j<k}}^{p} Q_{ijk}^2 + \sum_{i=1}^{p} (R_{iiii}-3)^2/24 +$$

$$\sum_{\substack{i,j=1\\ i<j}}^{p} (R_{iijj}-1)^2/4 + \sum_{\substack{i,j=1\\ i\neq j}}^{p} R_{iiij}^2/6 + \sum_{\substack{i,j,k=1\\ i\neq j,\, i\neq k,\, j<k}}^{p} R_{iijk}^2/2 + \sum_{\substack{i,j,k,q=1\\ i<j<k<q}}^{p} R_{ijkq}^2],$$

where the statistic LM_p is asymptotically distributed as a χ^2_ν, with degrees of freedom $\nu = p(p+1)(p+2)(p+7)/24$. The Appendix presents both a GAUSS program and a FORTRAN program that compute the LM test statistic for any dimension p.

For large samples, the hypothesis of p-variate normality of observations is rejected at some significant level (usually taken to be 10%) if the value of LM_p exceeds the corresponding critical value of the χ^2_ν. For small samples, however, the critical value derived from a Chi-square is far from being adequate, and a Montecarlo study is called forth. This was implemented for the case $p = 2$ using 10000 replications. The results thus obtained are reported in Table I. (Montecarlo studies for higher dimensions are currently being undertaken.)

Note as a final point that in the univariate case the LM statistic can be expressed in terms of the standardized third and fourth moments, given by $\sqrt{b_1} = m_3/m_2^{3/2}$ and $b_2 = m_4/m_2^2$ where the i-th central moment m_i equals $\Sigma(x_j-\bar{x})^i/n$, as:

$$(3.8) \qquad LM_1 = n[(\sqrt{b_1})^2/6 + (b_2-3)^2/24].$$

This statistic has been proposed by Bowman and Shenton (1975), and by Jarque and Bera (1980 and 1987). The former authors suggested the use of this statistic as the simplest possible omnibus test for normality since, under the null, the asymptotic means of $\sqrt{b_1}$ and b_2 are respectively 0 and 3, their asymptotic variances are 6/n and 24/n, and their asymptotic covariance is zero. While the latter authors found (3.8) to be the LM test statistic obtained when the alternatives to normality are in the Pearson family.

4. MULTIVARIATE NORMALITY OF RESIDUALS OF SIMULTANEOUS EQUATION MODELS.

Consider now the case of the linear structural model

$$(4.1) \qquad By_r + \Gamma z_r = u_r, \qquad r=1,\ldots,n,$$

where y_r is a p×1 vector of observed endogenous variables, z_r is a k×1 vector of observed predetermined variables, u_r is a p×1 vector of unobserved disturbances, B is a p×p nonsingular matrix of coefficients with ones in its diagonal, and Γ is a p×k matrix of coefficients. All identities are assumed to be substituted out, and the system is assumed to be identified through exclusions in B and Γ. Assume furthermore that the alternative to the possible p-variate normal distribution of u_r is a p-variate quartic exponential.

Suppose first that the system is estimated using full information maximum likelihood (FIML) under the assumption of multivariate normality. Following the reasoning in the last section, one can construct an LM test for multivariate normality of the residuals as follows: Let $\hat{u}_r$ denote the estimated FIML residuals of the structural equations (one could similarly use the estimated residuals of the reduced form). Using the transformation $e_r = S^{-1/2}\hat{u}_r$, where $S^{-1/2}$ is given as in (3.6) last section, define next

$$V_{ijk} = \sum_{r=1}^{n} e_{ri}e_{rj}e_{rk}/n \text{ and } W_{ijkq} = \sum_{r=1}^{n} e_{ri}e_{rj}e_{rk}e_{rq}/n.$$

Then the LM test statistic for multinormality of the residuals is given by:

$$LM_{N,p} = n[\sum_{i=1}^{p} V_{iii}^2/6 + \sum_{\substack{i,j=1\\ i\neq j}}^{p} V_{iij}^2/2 + \sum_{\substack{i,j,k=1\\ i<j<k}}^{p} V_{ijk}^2 + \sum_{i=1}^{p} (W_{iiii}-3)^2/24 +$$

$$\sum_{\substack{i,j=1\\ i<j}}^{p} (W_{iijj}-1)^2/4 + \sum_{\substack{i,j=1\\ i\neq j}}^{p} W_{iiij}^2/6 + \sum_{\substack{i,j,k=1\\ i\neq j,\, i\neq k,\, j<k}}^{p} W_{iijk}^2/2 + \sum_{\substack{i,j,k,q=1\\ i<j<k<q}}^{p} W_{ijkq}^2],$$

where the statistic $LM_{N,p}$ is asymptotically distributed as a χ_ν^2, with $\nu = p(p+1)(p+2)(p+7)/24$.

But what if, as is usually the case, the system is not estimated by FIML, but rather by some other method (e.g., 2SLS)? Provided the method renders consistent estimators, one can use the corresponding $LM_{N,p}$ constructed using the estimated residuals of the structural equations (or, equivalently, the estimated residuals of the reduced form equations). This is so because, following White and MacDonald (1980), one can show that the statistics constructed using the estimated residuals are consistent estimators of the true statistics.

As a final point, it is important to realize that, in the case of small samples, one cannot use the empirical significance points given in Table I. The reason is that the small sample properties of $LM_{N,p}$ can be shown to depend in general on the particular design matrices of each simultaneous equation system. In any case, it is quite straightforward to write a program to estimate, via Montecarlo simulation, the significance points for any $LM_{N,p}$.

5. FURTHER EXTENSIONS

Three extensions to the results presented above are reported here. First, using the same LM test procedure, and maintaining the hypothesis of Q-exponentials as the alternative distributions, tests for the multinormality of residuals of simultaneous limited dependent variable models (including as particular cases the typical univariate models) are given in Urzúa (1989b). Second, following the literature on multivariate generalized linear models, Urzúa (1989c) considers simultaneous equation models in which residuals are now explictly allowed to follow general Q-exponentials. And third, some initial steps (in the univariate case) are currently being undertaken by the author to derive algorithms to transform from Q-exponentials to the normal distribution.

APPENDIX: GAUSS AND FORTRAN PROGRAMS TO COMPUTE THE LM TEST STATISTIC

The following procedure written in GAUSS 2.0 returns the value of the LM test statistic. The argument of the procedure is the matrix of observations x with size n×p (instead of p×n as in the text), for any p ≧ 1.

```
proc lmnstat(x);
  local g,i,j,k,lm,q,s,u,v,va,ve,y;
  s=moment(x-meanc(x)',0)/rows(x);
  {va,ve}=eigrs2(s);
  g=ve*diagrv(eye(cols(x)),1/sqrt(va))*ve';
  y=(x-meanc(x)')*g;
  lm=sumc(meanc(y^3)^2)/6+sumc((meanc(y^4)-3)^2)/24;
  i=1;
  do while i<cols(x);
    j=i+1;
    do while j<=cols(x);
      u=y[.,i].*y[.,j];
      lm=lm+(meanc(u.*y[.,i])^2+meanc(u.*y[.,j])^2)/2;
      lm=lm+(meanc(u.*y[.,i]^2)^2+meanc(u.*y[.,j]^2)^2)/6;
      lm=lm+(meanc(u.*u)-1)^2/4;
      k=j+1;
      do while k<=cols(x);
        v=u.*y[.,k];
        lm=lm+meanc(v)^2+meanc(v.*y[.,i])^2/2;
        lm=lm+(meanc(v.*y[.,j])^2+meanc(v.*y[.,k])^2)/2;
        q=k+1;
        do while q<=cols(x);
          lm=lm+meanc(v.*y[.,q])^2;
          q=q+1;
        endo;
        k=k+1;
      endo;
      j=j+1;
    endo;
    i=i+1;
  endo;
  retp(lm*rows(x));
endp;
```

The next subroutine written in FORTRAN 77 returns the value of the LM test statistic as ST. The subroutine requests the number of dimensions P, the sample size N, and the N×P matrix of observations X. It also requests the largest expected P and N as MP and MN. It requires another subroutine to compute the eigenvalues and the normalized eigenvectors of a $P \times P$ real symmetric matrix.

```
      SUBROUTINE LMNTEST(P,N,X,MP,MN,ST)
      INTEGER P
      DIMENSION X(MN,MP),Y(MN,MP),S(MP,MP),G(MP,MP),D(MP)
C
C  Transformation of the data matrix
      DO 3 I=1,IP
        SUM=0.0
        DO 1 J=1,N
1       SUM=SUM+X(J,I)
        DO 2 J=1,N
2       X(J,I)=X(J,I)-SUM/N
3     CONTINUE
      DO 5 I=1,IP
      DO 5 K=I,IP
        SUM=0.0
        DO 4 J=1,N
4       SUM=SUM+X(J,I)*X(J,K)
        S(I,K)=SUM/N
5     S(K,I)=S(I,K)
      IF (IP.GT.1) THEN
C
C  The next line calls a subroutine (based on, say, the Jacobi method) that
C  returns the eigenvalues and the normalized eigenvectors of the sample
C  covariance P×P matrix S.  The eigenvalues are returned by the vector D,
C  and the eigenvectors by the matrix G.
        CALL JACOBI(S,IP,MP,D,G)
        DO 7 I=1,IP
        DO 7 K=I,IP
          SUM=0.0
          DO 6 J=1,IP
6         SUM=SUM+G(I,J)*G(K,J)/SQRT(D(J))
          S(I,K)=SUM
7       S(K,I)=S(I,K)
      ELSE
        S(1,1)=1/SQRT(S(1,1))
      ENDIF
      DO 9 J=1,N
      DO 9 I=1,IP
        SUM=0.0
        DO 8 K=1,IP
8       SUM=SUM+X(J,K)*S(K,I)
9     Y(J,I)=SUM
C
C  Computation of the test statistic
      ST=0.0
      DO 17 I=1,IP
```

```
      SUM1=0.0
      SUM2=0.0
      DO 10 J=1,N
        SUM1=SUM1+Y(J,I)**3
10    SUM2=SUM2+Y(J,I)**4
      ST=ST+((SUM1/N)**2)/6.0
      ST=ST+((SUM2/N-3.0)**2)/24.0
      DO 16 K=I+1,IP
        SUM1=0.0
        SUM2=0.0
        SUM3=0.0
        SUM4=0.0
        SUM5=0.0
        DO 11 J=1,N
          PR1=Y(J,I)*Y(J,K)
          SUM1=SUM1+PR1*Y(J,I)
          SUM2=SUM2+PR1*Y(J,K)
          SUM3=SUM3+PR1**2
          SUM4=SUM4+PR1*Y(J,I)**2
11      SUM5=SUM5+PR1*Y(J,K)**2
        ST=ST+((SUM1/N)**2)/2.0
        ST=ST+((SUM2/N)**2)/2.0
        ST=ST+((SUM3/N-1.0)**2)/4.0
        ST=ST+((SUM4/N)**2)/6.0
        ST=ST+((SUM5/N)**2)/6.0
        DO 15 L=K+1,IP
          SUM1=0.0
          SUM2=0.0
          SUM3=0.0
          SUM4=0.0
          DO 12 J=1,N
            PR2=PR1*Y(J,L)
            SUM1=SUM1+PR2
            SUM2=SUM2+PR2*Y(J,I)
            SUM3=SUM3+PR2*Y(J,K)
12        SUM4=SUM4+PR2*Y(J,L)
          ST=ST+(SUM1/N)**2
          ST=ST+((SUM2/N)**2)/2.0
          ST=ST+((SUM3/N)**2)/2.0
          ST=ST+((SUM4/N)**2)/2.0
          DO 14 M=L+1,IP
            SUM1=0.0
            DO 13 J=1,N
13          SUM1=SUM1+PR2*Y(J,M)
14        ST=ST+(SUM1/N)**2
15      CONTINUE
16    CONTINUE
17  CONTINUE
    ST=N*ST
    RETURN
    END
```

TABLE I

SIGNIFICANCE POINTS OF THE LM TEST STATISTIC FOR BIVARIATE NORMALITY

Sample Size	$\alpha = .10$	Ratio[a]	Sample Size	$\alpha = .10$	Ratio
10	7.47	0.48	65	13.75	0.88
15	9.54	0.59	80	14.11	0.89
20	10.61	0.66	100	14.15	0.91
25	11.54	0.72	150	14.66	0.94
30	12.08	0.76	200	14.66	0.96
40	12.90	0.81	300	14.69	0.98
50	13.18	0.84	∞	14.68	1.00

[a]Ratio of empirical to asymptotic mean values.

Note: A rough approximation for critical values at a significance level other than $\alpha = .10$ can be found multiplying "Ratio" times the critical value of the corresponding Chi-square [i.e., $\chi^2_{\nu,\alpha}$ with $\nu = p(p+1)(p+2)(p+7)/24$].

REFERENCES

Aitchison, J., and S. D. Silvey (1958): "Maximum Likelihood Estimation of Parameters Subject to Restraints," Annals of Mathematical Statistics, 29, 813-828.

Bera, A. K., and S. John (1983): "Tests for Multivariate Normality with Pearson Alternatives," Communications in Statistics, Theory and Methods, 12, 103-117.

Bowman, K. O., and L. R. Shenton (1975): "Omnibus Contours for Departures from Normality Based on $\sqrt{b_1}$ and b_2," Biometrika, 62, 243-250.

Cox, D. R. (1977): "The Role of Significance Tests," Scandinavian Journal of Statistics, 4, 49-70.

Fisher, R. A. (1922): "On the Mathematical Foundations of Theoretical Statistics," Philosophical Transactions of the Royal Society of London, Ser. A, 222, 309-368.

Gart, J. J., and R. E. Tarone (1983): "The Relation Between Score Tests and Approximate UMPU Tests in Exponential Models Common in Biometry," Biometrics, 39, 781-786.

Godfrey, L. G. (1988): Mispecification Tests in Econometrics: The Lagrange Multiplier Principle and other Approaches, Cambridge: Cambridge University Press.

Jarque, C. M., and A. K. Bera (1980): "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals," Economics Letters, 6, 255-259.

______ (1987): "A Test for Normality of Observations and Regression Residuals," International Statistical Review, 55, 163-172.

Jarque, C. M., and C. R. McKenzie (1983): "Testing for Multivariate Normality in Simultaneous Equations Models," Working Papers in Economics and Econometrics No. 82, Australian National University.

Jaynes, E. T. (1957): "Information Theory and Statistical Mechanics," Physical Review, 106, 620-630.

Mardia, K. V. (1970): "Measures of Multivariate Skewness and Kurtosis with Applications," Biometrika, 57, 519-530.

Mardia, K. V. (1980): "Tests of Univariate and Multivariate Normality," in Handbook of Statistics, Vol. I, ed. by P. R. Krishnaiah. Amsterdam: North-Holland.

Rao, C. R. (1948): "Large Sample Tests of Statistical Hypotheses Concerning Several Parameters with Applications to Problems of Estimation," Proceedings of the Cambridge Philosophical Society, 44, 49-57.

Shannon, C. E. (1948): "A Mathematical Theory of Communication," Bell System Technical Journal, 27, 379-423 and 623-656. Reprinted in The Mathematical Theory of Communication (1949), C. E. Shannon and W. Weaver, Urbana: University of Illinois Press.

Small, N. J. H. (1985): "Multivariate Normality, Testing For," in Encyclopedia of Statistical Sciences, Vol. 6, ed. by S. Kotz and N. L. Johnson. New York: John Wiley and Sons.

Urzúa, C. M. (1988): "A Class of Maximum-Entropy Multivariate Distributions," Communications in Statistics, Theory and Methods, 17, 4039-4057.

______ (1989a): "Lending to Sovereign Borrowers: Snapshots of the Eurocurrency Market Using a Thomian Camera," forthcoming in International Economic Review.

______ (1989b): "Testing for Normality in Limited Dependent Variable Models," manuscript, Princeton University.

______ (1989c): "Simultaneous Equations with Nonnormal Residuals," manuscript, El Colegio de Mexico.

White, H., and G. M. MacDonald (1980): "Some Large-Sample Tests for Nonnormality in the Linear Regression Model," Journal of the American Statistical Association, 75, 16-28.

Zellner, A., and R. A. Highfield (1988): "Calculation of Maximum Entropy Distributions and Approximation of Marginal Posterior Distributions," Journal of Econometrics, 37, 195-209.