The Annals of Statistics
1986, Vol. 14, No. 2, 708–723

TESTING FOR NORMALITY IN ARBITRARY DIMENSION

BY SÁNDOR CSÖRGŐ

Szeged University and University of California, San Diego

The univariate weak convergence theorem of Murota and Takeuchi (1981) is extended for the Mahalanobis transform of the d-variate empirical characteristic function, $d \geq 1$. Then a maximal deviation statistic is proposed for testing the composite hypothesis of d-variate normality. Fernique's inequality is used in conjunction with a combination of analytic, numerical analytic, and computer techniques to derive exact upper bounds for the asymptotic percentage points of the statistic. The resulting conservative large sample test is shown to be consistent against every alternative with components having a finite variance. (If $d = 1$ it is consistent against every alternative.) Monte Carlo experiments and the performance of the test on some well-known data sets are also discussed.

1. Introduction. Beside the permanent interest in testing univariate normality, recent years have witnessed a large increase of interest in the corresponding equally important but more intricate problem of testing for multivriate normality. The work of Weiss (1958), Anderson (1966), Cox (1968), Healy (1968), Wagle (1968), Wilk and Gnanadesikan (1968), Day (1969), Mardia (1970, 1974, 1975), Andrews, Gnanadesikan, and Warner (1971, 1972, 1973), Malkovich (1971), Aitkin (1972), Gnanadesikan and Kettenring (1972), Kessel and Fukunaga (1972), Dahiya and Gurland (1973), Malkovich and Afifi (1973), Mardia and Zemroch (1975), Giorgi and Fattorini (1976), and Hensler, Mehrota, and Michalek (1977) has been discussed in detail by Gnanadesikan (1977, pages 161–195), Cox and Small (1978), and Mardia (1980). The proposals by Sarkadi and Tusnády (1977), Small (1978, 1980), DeWet, Wenter, and van Wyk (1979), Pettitt (1979), Rincon-Gallardo, Quesenberry, and O'Reilly (1979), Hawkins (1981), Moore and Stubblebine (1981), Yang (1981), and Koziol (1982, 1983), not covered in the three surveys, are either continuations of earlier work or may be more or less fitted into one of the classification sections in Mardia (1980). The goodness-of-fit tests recently introduced by Bickel and Breiman (1983) [see also Schilling (1983a, b)] for a *simple* multidimensional hypothesis do not seem to lend themselves easily for adaptation to the composite case.

The approach of the present paper is based on the asymptotic behaviour of the multivariate "studentised" empirical characteristic function and is a multivariate extension of the recent approach of Murota and Takeuchi (1981) for testing univariate normality. Thus the basic weak convergence theorem in Section 2 extends the corresponding result, Theorem 6, of Murota and Takenchi (1981) to arbitrary dimension. The distributions of the most relevant functionals, such as

Received September 1984; revised April 1985.
AMS 1980 *subject classifications.* Primary 62H15; secondary 62F03, 62F05.
Key words and phrases. Empirical characteristic function, Mahalanobis transform, univariate and multivariate normality, weak convergence, maximal deviation, Fernique's and Borell's bounds on the absolute supremum of a Gaussian process.

the absolute supremum and the squared integral, of the limiting parameter-free Gaussian process are not known even in the univariate case. Therefore, Murota and Takeuchi (1981) have considered the simplest possible projection functional as a test statistic for testing univariate normality. [See also Murota (1981); related univariate tests are in Hall and Welsh (1983) and Welsh (1985).] Although we briefly mention multivariate extensions of the projection statistics in Section 3, our primary aim is to give a tight bound on the tail of the distribution of the absolute supremum of the limiting multivariate Gaussian process. This is achieved by applying to all possible limits a powerful inequality of Fernique (1975) in Section 4. The resulting formal conservative large sample Kolmogorov-type test is new even in the univariate case. The consistency of the test is also discussed in Section 4. Approximate computing formulae for our maximal deviation statistic are given in Section 5, a limited simulation study under the null hypothesis is discussed in Section 6, while in Section 7 the performance of the test is illustrated on the well-known Norton's bank data and Fisher's *Iris setosa* data. A related Cramér–von Mises type statistic is mentioned briefly in Section 8.

2. The basic weak convergence theorem. Let $d \geq 1$ be a fixed integer and let $X(1), \dots, X(n)$ be independent d-dimensional random column vectors identically distributed as $X' = (X_1, \dots, X_d)$, where the prime denotes transpose. Let

$$C_n(t) = U_n(t) + iV_n(t) = n^{-1} \sum_{j=1}^{n} \exp(i\langle t, X(j)\rangle), \qquad t' = (t_1, \dots, t_d) \in \mathbb{R}^d$$

be the sample characteristic function, where $\langle t, s\rangle = \sum_{k=1}^{d} t_k s_k$ with $s' = (s_1, \dots, s_d)$ stands for the inner product in $\mathbb{R}^d$, and let $S_n = (s_{km}(n))$ be the sample covariance matrix

$$s_{km}(n) = n^{-1} \sum_{j=1}^{n} (X_k(j) - \overline{X}_k(n))(X_m(j) - \overline{X}_m(n)), \qquad k, m = 1, \dots, d,$$

where $\overline{X}'(n) = (\overline{X}_1(n), \dots, \overline{X}_d(n)) = n^{-1}\sum_{j=1}^{n} X'(j)$ is the sample mean vector. Assuming that the underlying distribution function of X is continuous, we may almost surely consider the unique symmetric positive definite square root $S_n^{-1/2}$ of the inverse S_n^{-1} of S_n. The studentised empirical characteristic function, or, rather, the Mahalanobis transform of $C_n(t)$ is

$$\begin{aligned} C_n\big(S_n^{-1/2}t\big) &= n^{-1} \sum_{j=1}^{n} \exp\big(i\langle S_n^{-1/2}t, X(j)\rangle\big) \\ &= n^{-1} \sum_{j=1}^{n} \exp\big(i\langle t, X'(j)S_n^{-1/2}\rangle\big) \qquad (2.1) \\ &= n^{-1} \sum_{j=1}^{n} \exp(i\langle t, X(j)\rangle)\exp\big(i\langle t, X'(j)\big(S_n^{-1/2} - I\big)\rangle\big), \end{aligned}$$

where I is the $d \times d$ identity matrix. Since its squared modulus $|C_n(S_n^{-1/2}t)|^2$ is invariant under all nonsingular affine transformations of the sample $X(1),\dots, X(n)$ for any vector t, there will be no loss of generality while proceeding toward the main goal of this section, Theorem 2.2, if in the preliminary Theorem 2.1 below, we assume that $\mu = EX = 0 \in \mathbb{R}^d$ and $\Sigma = I$, or, in fact, that the components of X are independent.

Let T be an arbitrary positive number and let $\mathscr{C} = \mathscr{C}([-T, T]^d)$ and $\mathscr{C}^2 = \mathscr{C}^2([-T, T]^d)$ be the separable Banach spaces of the d-variate continuous real and complex valued functions, respectively, defined on the cube $[-T, T]^d$ and endowed with the respective supremum norms. Let $C(t) = U(t) + iV(t) = E\exp(i\langle t, X\rangle)$, $t \in \mathbb{R}^d$, be the population characteristic function. Our basic stochastic process

$$Z_n(t) = n^{1/2}\{|C_n(S_n^{-1/2}t)|^2 - |C(t)|^2\}$$

is a random element of $\mathscr{C}$ for each $n = 1, 2, \dots$. Let us introduce the vector of partial derivatives of C

$$(\nabla C(t))' = \left(\frac{\partial C(t)}{\partial t_1}, \dots, \frac{\partial C(t)}{\partial t_d}\right)$$

and the corresponding $d \times d$ Laplacian matrix $\nabla^2 C(t)$ with elements

$$\frac{\partial^2 C(t)}{\partial t_k \partial t_m}, \qquad k, m = 1, \dots, d.$$

Assuming that the vector $(\mu^{(4)})' = (\mu_1^{(4)}, \dots, \mu_d^{(4)}) = (EX_1^4, \dots, EX_d^4)$ is finite, consider the d-variate complex Gaussian process $Y(t)$ satisfying $\overline{Y(t)} = \operatorname{Re} Y(t) - i \operatorname{Im} Y(t) = Y(-t)$ and $EY(t) = 0$ for each t and for each $s, t \in \mathbb{R}^d$,

$$\begin{aligned}
\rho(s,t) &= EY(s)Y(t) \\
&= C(s+t) - C(s)C(t) \\
&\quad + \frac{1}{2}\{s'(\nabla^2 C(t))\nabla C(s) + t'(\nabla^2 C(s))\nabla C(t) \\
&\qquad + C(s)\langle t, \nabla C(t)\rangle + C(t)\langle s, \nabla C(s)\rangle\} \\
&\quad + \frac{1}{4}\left\{\sum_{m=1}^{d}\left(\mu_m^{(4)} - 1\right)s_m t_m \frac{\partial C(s)}{\partial s_m}\frac{\partial C(t)}{\partial t_m}\right. \\
&\qquad \left. + \sum_{k,m=1,\,k\neq m}^{d}\left[s_k t_k \frac{\partial C(s)}{\partial s_m}\frac{\partial C(t)}{\partial t_m} + s_k t_m \frac{\partial C(s)}{\partial s_m}\frac{\partial C(t)}{\partial t_k}\right]\right\}.
\end{aligned} \tag{2.2}$$

That such a process exists, that is, that ρ indeed is a covariance function, may be seen by considering the random function

$$\begin{aligned}
R(t) &= R(t; X) \\
&= e^{i\langle t, X\rangle} - C(t) - \frac{1}{2}\sum_{m=1}^{d}\left\{(X_m^2 - 1)t_m + X_m \sum_{k=1,\,k\neq m}^{d} X_k t_k\right\}\frac{\partial C(t)}{\partial t_m}
\end{aligned} \tag{2.3}$$

in $\mathscr{C}^2$, where the components of $X' = (X_1, \dots, X_d)$ are independent with $EX_1 =$

$\cdots = EX_d = 0$ and $EX_1^2 = \cdots = EX_d^2 = 1$. Then a straightforward but somewhat lengthy computation shows that $\rho(s,t) \equiv ER(s)R(t)$. Moreover, since $\mu^{(4)}$ is finite, $C(t)$ has uniformly continuous partial derivatives of the fourth order over the whole space $\mathbb{R}^d$. Therefore a one-term d-variate Taylor expansion easily gives

$$E|Y(s) - Y(t)|^2 = E|R(s) - R(t)|^2 \leq K|s-t|^2, \tag{2.4}$$

with some constant $K > 0$, and this is more than enough to imply the sample continuity of the complex Gaussian process Y (Fernique, 1975, page 48). Hence Y may be considered as a random element of $\mathscr{C}^2$, and consequently the d-variate real Gaussian process

$$Z(t) = 2\{U(t)\operatorname{Re} Y(t) + V(t)\operatorname{Im} Y(t)\} = 2\operatorname{Re}\{C(-t)Y(t)\}, \tag{2.5}$$

with mean zero and covariance function

$$\begin{aligned} \sigma(s,t) &= EZ(s)Z(t) \\ &= 2\operatorname{Re}\{C(-s)C(-t)\rho(s,t) + C(s)C(-t)\rho(-s,t)\} \end{aligned} \tag{2.6}$$

is a random element of $\mathscr{C}$. Note that $Z(t) = Z(-t)$.

THEOREM 2.1. *If the components of $X' = (X_1, \ldots, X_d)$ are independent with $EX_1 = \cdots = EX_d = 0$, $EX_1^2 = \cdots = EX_d^2 = 1$, and the vector $\mu^{(4)}$ finite, then the sequence $\{Z_n(\cdot)\}$ converges weakly, as $n \to \infty$, in $\mathscr{C}([-T,T]^d)$ to the Gaussian process $Z(\cdot)$.*

The proof of Theorem 2.1 is in the Appendix.

Let $N_d(\mu, \Sigma)$ denote the d-dimensional normal distribution with mean vector μ and covariance matrix Σ, where $d \geq 1$. Our aim is to test the composite hypothesis

H_d: *The law of X is $N_d(\mu, \Sigma)$ with some μ and some nonsingular Σ.*

Note that when $C(t)$ is real then from (2.2) and (2.6) we obtain

$$\begin{aligned} \sigma(s,t) = 2C(s)C(t)\Bigg[& C(s+t) - C(s-t) - 2C(s)C(t) \\ & + \{s'(\nabla^2 C(t))\nabla C(s) + t'(\nabla^2 C(s))\nabla C(t) \\ & + C(s)\langle t, \nabla C(t)\rangle + C(t)\langle s, \nabla C(s)\rangle\} \\ & + \frac{1}{2}\Bigg\{\sum_{m=1}^{d} (\mu_m^{(4)} - 1) s_m t_m \frac{\partial C(s)}{\partial s_m}\frac{\partial C(t)}{\partial t_m} \\ & + \sum_{k,m=1,\,k\neq m}^{d} \left[s_k t_k \frac{\partial C(s)}{\partial s_m}\frac{\partial C(t)}{\partial t_m} + s_k t_m \frac{\partial C(s)}{\partial s_m}\frac{\partial C(t)}{\partial t_m} \right]\Bigg\}\Bigg]. \end{aligned} \tag{2.7}$$

THEOREM 2.2. *If H_d holds then the sequence of stochastic processes*

$$Z_n(t) = n^{1/2}\{|C_n(S_n^{-1/2}t)|^2 - e^{-\langle t,t\rangle}\}$$

converges weakly in $\mathscr{C}([-T, T]^d)$ *to the Gaussian process* $Z(t)$ *satisfying* $Z(t) = Z(-t)$, $EZ(t) = 0$ *and*

$$\sigma(s, t) = EZ(s)Z(t) = 4e^{-\langle s, s\rangle - \langle t, t\rangle}\left\{\cosh(\langle s, t\rangle) - 1 - \tfrac{1}{2}\langle s, t\rangle^2\right\}.$$

PROOF. Since the limiting process is the same for any μ and Σ under H_d, the theorem follows by substituting $C(t) = \exp(-\frac{1}{2}\langle t, t\rangle)$ and $\mu_m^{(4)} = 3$ into (2.7). Then a lengthy computation yields the indicated formula.

Note that $\cosh(x) - 1 - (x^2/2) = \mathscr{O}(x^4)$ as $x \to 0$, and the process $Z(\cdot)$ in Theorem 2.2 has the interesting property that $Z(s)$ and $Z(t)$ are uncorrelated and hence independent for any vectors s and t that are orthogonal to one another. For nonzero and nonorthogonal s and t, $\sigma(s, t) > 0$.

3. Simple projection statistics. The simplest possible such statistic is obtained if we consider the nonzero vectors $t_1, \dots, t_L$, somewhere in the vicinity of the origin, such that the $L \times L$ matrix $R = (\sigma(t_k, t_m))$ be nonsingular and form the quadratic form $Q_n = Q_n(t_1, \dots, t_L) = z_n' R^{-1} z_n$, where $z_n' = (Z_n(t_1), \dots, Z_n(t_L))$ with Z_n and σ of Theorem 2.2. Then under H_d the asymptotic distribution of Q_n is the chi-square distribution with L degrees of freedom. There is no theoretically justified ground, however, upon which the number L and the location of the vectors $t_1, \dots, t_L$ could reasonably be chosen.

Another, perhaps more appealing d-dimensional extension of the Murota and Takeuchi (1981) statistic is based on the observation at the end of the preceding section. Let $t_1, \dots, t_d$ be nonzero, pairwise orthogonal vectors from $\mathbb{R}^d$ and set $N_n^{(d)} = \max(|Z_n(t_1)|, \dots, |Z_n(t_d)|)$. Then under H_d,

$$\lim_{n\to\infty} \Pr\left\{N_n^{(d)} \le x\right\} = \prod_{k=1}^{d} \left\{2\Phi\left(\frac{x}{\sigma(|t_k|)}\right) - 1\right\}, \qquad 0 \le x < \infty,$$

where Φ is the standard normal distribution function and $\sigma^2(|t|) = \sigma(t, t) = 4\exp(-2|t|^2)\{\cosh(|t|^2) - 1 - \frac{1}{2}|t|^4\}$ depends only on the length of t. The standard-deviation function $\sigma(|t|)$ has a unique global maximum on $[0, \infty)$,

$$\sigma(|t_0|) = 0.23743 \quad \text{at } |t_0| = 1.4684924, \tag{3.1}$$

determined on the computer. So $N_n^{(d)} = N_n^{(d)}(t_1, \dots, t_d)$ is asymptotically "most variable" under H_d on the surface of the d-dimensional ball $r^2 = |t_0|^2$. For the sake of later comparison let us choose all the points $t_1, \dots, t_d$ on this surface and record the asymptotic α level significance points obtained from the equation $(2\Phi(x/0.23743) - 1)^d = 1 - \alpha$, $0 < \alpha < 1$.

4. The maximal deviation statistic. The natural extension of $N_n^{(d)}$ is

$$M_n^{(d)}(T) = \sup_{t\in[-T, T]^d} |Z_n(t)| = n^{1/2} \sup_{t\in[-T, T]^d} \left||C_n(S_n^{-1/2}t)|^2 - \exp(-|t|^2)\right|,$$

where T is some positive number. Note at the same time that the restriction of the supremum to a finite cube $[-T, T]^d$ is not a theoretical restriction for the

TABLE 1
Asymptotic $1-\alpha$ *percentage points of* $N_n^{(d)}(t_1,\dots,t_d)$ *with* $|t_l| = \cdots = |t_d| = |t_0|$

α \ d	1	2	3	4	5	6
0.1	0.3906	0.4630	0.5176	0.5276	0.5485	0.5893
0.05	0.4654	0.5318	0.5675	0.5912	0.6090	0.6481
0.01	0.6102	0.6672	0.6957	0.7194	0.7360	0.7669

problem at hand. Indeed, as a consequence of the corresponding well-known univariate result and the fact that the univariate normality of all the linear combinations of the components of a vector implies the multivariate normality of the vector, we have the following: If a d-variate characteristic function coincides with a given d-variate normal characteristic function in any small neighbourhood of the origin, then they coincide everywhere on $\mathbb{R}^d$. Therefore, the only considerations that should be made when choosing $T > 0$ are those that relate to finite sample behaviour and computational ease.

Under H_d, $\lim_{n\to\infty} \Pr\{M_n^{(d)}(T) \geq y\} = F_{d,T}(y)$ for any $y > 0$, where with Z as in Theorem 2.2,

$$F_{d,T}(y) = \Pr\Big\{ \sup_{t\in[-T,T]^d} |Z(t)| \geq y \Big\}.$$

Of course, this function is not known. We wish to give an upper bound for it. By the inequality of Fernique [(1975), page 51], for any integer $p \geq 2$, we have

$$F_{d,T}(xK_{d,T}(p)) \leq 5\left(\frac{\pi}{2}\right)^{1/2} p^{2d}(1-\Phi(x)) \tag{4.1}$$

for any

$$x \geq (1 + 4d\log p)^{1/2}, \tag{4.2}$$

where

$$K_{d,T}(p) = \sup_{s,t\in[-T,T]^d} \sigma^{1/2}(s,t) + (2+\sqrt{2})\int_1^\infty \phi_{d,T}(Tp^{-u^2})\,du, \tag{4.3}$$

with

$$\phi_{d,T}(h) = \sup_{s,t\in[-T,T]^d,\ \|s-t\|\leq h} (\sigma(s,s) + \sigma(t,t) - 2\sigma(s,t))^{1/2}, \tag{4.4}$$

where $\|s-t\| = \max\{|s_k - t_k| : 1 \leq k \leq d\}$ is the maximum norm.

By the Cauchy–Schwarz inequality we have

$$\sup_{s,t\in[-T,T]^d} \sigma^{1/2}(s,t) = \sup_{0\leq|t|\leq Td^{1/2}} \sigma(|t|), \tag{4.5}$$

where $\sigma^2(|t|) = \sigma(t,t)$ is the variance function. As noted, $\sigma(|t|)$ has a global maximum at $|t_0|$ of (3.1). For the sake of definiteness and in order to include the surface of the ball where $Z(t)$ is most variable, we choose

$$T = 1.47/d^{1/2} \tag{4.6}$$

and suppress henceforth this T in the notation. The first term of $K_d(p)$ is then given by (3.1), and the basic problem is to determine the function $\phi_d(h)$ in (4.4). We shall see that the main advantage of the choice in (4.6) is that $K_d(p)$ will not depend on the dimension if it is higher than one. [See (4.10) and (4.11) below]. Set

$$A_d = \sup_{s,\, t \in [-1.47/d^{1/2},\, 1.47/d^{1/2}]^d} \frac{\sigma(s, s) + \sigma(t, t) - 2\sigma(s, t)}{|s - t|^2}$$

and

$$B = \sup_{0 \le u \le 1.47} 4e^{-2u^2} g'(u^2), \tag{4.7}$$

where $g'(x)$ is the derivative of the function

$$g(x) = \cosh(x) - 1 - \frac{x^2}{2}, \qquad x \ge 0,$$

and note that

$$A_1 = \sup_{0 \le u,\, v \le 1.47} 4 \frac{e^{-2u^2} g(u^2) + e^{-2v^2} g(v^2) - 2e^{-u^2 - v^2} g(uv)}{(u - v)^2}. \tag{4.8}$$

The key step is the following lemma, also proved in the Appendix.

LEMMA 4.1. *For any* $d \ge 2$, $A_d = \max(A_1, B)$.

Of course, B can be computed on the computer easily, and we obtain

$$B = 0.1265243. \tag{4.9}$$

A combination of careful numerical analysis and some computer work also gives that $A_1 \le 0.1085898$. Consequently, for $\phi_d(h)$ in (4.4), we have

$$\phi_d(h) \le A_d^{1/2} d^{1/2} h, \qquad 0 \le h < \infty, \tag{4.10}$$

where by Lemma 4.1 and (4.9),

$$A_d^{1/2} \begin{cases} \le 0.3295297, & \text{if } d = 1, \\ = 0.3557025, & \text{if } d \ge 2. \end{cases} \tag{4.11}$$

Now putting $T = 1.47\, d^{-1/2}$ in $K_d(p)$ of (4.3), by (4.5), (3.1), (4.10), (4.11), and a simple substitution in the integral, we obtain

$$K_d(p) \le L_d(p) = 0.23743 + V_d (\log p)^{-1/2} \left(1 - \Phi\left((2 \log p)^{1/2}\right)\right), \tag{4.12}$$

where

$$1.47(2 + \sqrt{2})(A_d \pi)^{1/2} \le V_d = \begin{cases} 2.9314164, & \text{if } d = 1, \\ 3.1642433, & \text{if } d \ge 2. \end{cases}$$

We may now return to (4.1) and (4.2). Fix an α, $0 < \alpha < 1$, and introduce the set

$$E_d(\alpha) = \left\{ p\colon p \text{ integer}, \ p \geq 2, \ \Phi^{-1}\left(1 - \frac{\alpha}{5(\pi/2)^{1/2} p^{2d}}\right) \geq (1 + 4d \log p)^{1/2} \right\},$$

where Φ^{-1} is the inverse function to Φ. Then, with $L_d(p)$ as in (4.12), for $M_n^{(d)} = M_n^{(d)}(1.47/d^{1/2})$ we have the following result.

THEOREM 4.2. *If H_d holds then*

$$\lim_{n \to \infty} \Pr\{M_n^{(d)} \geq z_d(\alpha)\} \leq \alpha,$$

where

$$z_d(\alpha) = \inf_{p \in E_d(\alpha)} \Phi^{-1}\left(1 - \frac{\alpha}{5(\pi/2)^{1/2} p^{2d}}\right) L_d(p).$$

The differences $z_d(\alpha) - y_d(\alpha) \geq 0$, where $y_d(\alpha)$ is the real asymptotic $1 - \alpha$ percentage point, i.e., $F_d(y_d(\alpha)) = \alpha$, are not known. However, the following table of the $z_d(\alpha)$ values suggests in comparison with Table 1 that the Fernique inequality is quite powerful on our process $Z(t)$ and, therefore, the unknown differences $z_d(\alpha) - y_d(\alpha) \geq 0$ are hopefully not too large.

The computation of this table required a table of $\Phi(x)$ with x in the interval [2.65, 8.35] and with 17 precise decimals. The p values at which the corresponding infima were taken ranged from 7 to 12.

As to the consistency of the test, we can prove the following

THEOREM 4.3. *If the components of X are linearly independent with finite variances but H_d does not hold, then $M_n^{(d)} \to \infty$ almost surely as $n \to \infty$. Hence the test is consistent against all such alternatives.*

PROOF. Since $S_n^{-1/2} \to \Sigma$ almost surely, where Σ is a symmetric positive definite $d \times d$ matrix, and $C_n(\cdot)$ converges to the characteristic function $C(\cdot)$ of X almost surely uniformly on compact sets, $C_n(S_n^{-1/2}t) \to C_0(t) = C(\Sigma t)$ almost surely uniformly on $K_d = [-T, T]^d$, where T is as in (4.6). This implies $|C_n(S_n^{-1/2}t)|^2 \to |C_0(t)|^2$ uniformly on K_d.

Suppose that $C_0(t)C_0(-t) = |C_0(t)|^2 = \exp(-|t|^2)$ on K_d. Since $\exp(-|t|^2)$ is a d-variate normal characteristic function, and since $|C_0(t)|^2$ is a characteristic

TABLE 2
The upper bounds $z_d(\alpha)$ of the asymptotic $1 - \alpha$ percentage points of $M_n^{(d)}$

α \ d	1	2	3	4	5	6
0.1	0.9648	1.2613	1.4963	1.6985	1.8804	2.0466
0.05	1.0101	1.2998	1.5294	1.7296	1.9024	2.0730
0.01	1.1087	1.3822	1.6034	1.7973	1.9719	2.1257

function, the observation made at the beginning of this section implies that the equality $C_0(t)C_0(-t) = \exp(-|t|^2)$ must hold on the whole space $\mathbb{R}^d$. By the obvious multivariate version of Cramér's theorem [Lukacs (1970), Theorem 8.2.1; obtained again from the univariate result by taking linear combinations] the latter identity implies that the component $C_0(t)$ is a normal characteristic function. This, in turn, implies that $C(t)$ itself is a normal characteristic function, which contradicts to the assumption that H_d is not satisfied. Therefore,

$$\lim_{n\to\infty} \sup_{t\in K_d} ||C_n(S_n^{-1/2}t)|^2 - e^{-|t|^2}| = \sup_{t\in K_d} ||C_0(t)|^2 - e^{-|t|^2}| > 0$$

almost surely, and since the $M_n^{(d)}$ are $n^{1/2}$ times these suprema, $n = 1, 2, \dots,$ the assertion follows.

If one of the components of X has an infinite variance, then one feels that the test is "all the more consistent" that is, the natural conjecture is that it is consistent against every fixed alternative. However, the behaviour of $S_n^{-1/2}$ is unclear with infinite variances and I do not have a formal proof for this case, except when $S_n^{-1/2}$ converges to the zero matrix. This is the situation if $d = 1$, and so *the test is indeed consistent against every alternative in the univariate case*. Note that if $d = 1$ and $EX^2 = \infty$, then almost surely as $n \to \infty$,

$$\sup_{-1.47\le t\le 1.47} |C_n(t/S_n^{1/2}) - e^{-t^2}| \to \sup_{0\le t\le 1.47} |1 - e^{-t^2}| = 1 - e^{-(1.47)^2} \approx 0.8848.$$

5. Computing formulae. Writing $Y^j = (Y_1^j, \dots, Y_d^j) = a_d(L)X'(j)S_n^{-1/2}$, $j = 1, \dots, n$, with $a_d(L) = 1.47/(d^{1/2}10^L)$ and $c_k(x) = \cos kx$, $s_k(x) = \sin kx$, where k and m will denote integers and $L > 0$, and using that $Z_n(t) = Z_n(-t)$, we have

$$M_n^{(1)} \approx n^{1/2} \max_{1\le k\le 10^L} \left| \left(\frac{1}{n} \sum_{j=1}^n c_k(Y^j) \right)^2 + \left(\frac{1}{n} \sum_{j=1}^n s_k(Y^j) \right)^2 - e^{-(ka_1(L))^2} \right|,$$

and using the sine and cosine addition formulae,

$$\begin{aligned} M_n^{(2)} \approx n^{1/2} \max_{-10^L\le m\le 10^L, 1\le k\le 10^L} \Bigg| & \left(\frac{1}{n} \sum_{j=1}^n \{c_m(Y_1^j)c_k(Y_2^j) - s_m(Y_1^j)s_k(Y_2^j)\} \right)^2 \\ & + \left(\frac{1}{n} \sum_{j=1}^n \{s_m(Y_1^j)c_k(Y_2^j) + c_m(Y_1^j)s_k(Y_2^j)\} \right)^2 \\ & - \exp\left(-(ka_2(L))^2 - (ma_2(L))^2\right) \Bigg|, \end{aligned}$$

where the larger L is, the more precise are the approximate equalities $\approx$. Analogously, increasingly more complex formulae can be written down for $M_n^{(d)}$, $d \ge 3$. The point is that many computers compute sine and cosine slowly, and if we use the recursion

$$c_2(\cdot) = 2c_1(\cdot)c_1(\cdot) - 1, s_2(\cdot) = 2c_1(\cdot)s_1(\cdot),$$
$$c_k(\cdot) = 2c_1(\cdot)c_{k-1}(\cdot) - c_{k-2}(\cdot), s_k(\cdot) = 2c_1(\cdot)s_{k-1}(\cdot) - s_{k-2}(\cdot),$$
$$k=3,\dots,10^L,$$

then in d dimensions we need to compute only nd sines and nd cosines. In practice, $L = 2$ is usually sufficient.

6. Simulation. In the univariate case, we conducted a very limited Monte Carlo experiment to determine empirically approximate values of the unknown limiting percentage points $y_1(\alpha)$, $\alpha = 0.1, 0.05, 0.01$. Normal $(0,1)$ random numbers with sample sizes 50 and 100 were generated 500 times in both cases, and $M_{50}^{(1)}$ and $M_{100}^{(1)}$ were computed by the above formula with $L = 2$. The obtained percentage points of the 500 samples for $M_{50}^{(1)}$ and $M_{100}^{(1)}$ are the following:

α \ n	50	100
0.1	0.5069	0.5160
0.05	0.6044	0.6173
0.01	0.9678	0.8455

These should be compared with the first columns of Tables 1 and 2.

Of course large order statistics of $M_n^{(1)}$ in 500 samples are more unstable than smaller ones.

In general ($d \geq 1$), the very sharp inequality of Borell (1975) says instead of (4.1) that

$$F_{d,T}(x\sigma + m_d) \leq 1 - \Phi(x), \qquad x > 0,$$

where σ is the supremum of the pointwise standard deviation of $Z(t)$ on $[-T, T]^d$, and m_d is the median of the distribution of $M^{(d)} = \sup\{|Z(t)|: t \in [-T, T]^d\}$. The problem is, of course, that we do not know m_d. With the choice of T as in (4.6) (or larger), $\sigma = 0.23743$ according to (3.1). (If we use Fernique's inequality first to give an upper bound for m_d and *then* Borell's inequality with this upper bound, then we obtain larger values than those given in Table 2.)

The median of the 500 samples for $M_{50}^{(1)}$ was 0.2258, and that of for $M_{100}^{(1)}$ was 0.2273. Arguing that simulation of middle percentage points is more stable, let us accept that 0.23 is a close upper bound for m_1. Then Borell's inequality gives the following tentative close upper bounds for $y_1(\alpha)$:

α	
0.1	0.5344
0.05	0.6206
0.01	0.7832

A thorough simulation study of the properties of the $M_n^{(d)}$ test would be desirable.

7. Examples. Our first example is testing the bivariate normality of *Norton's rate of discount and ratio of reserves to deposits* data of size 780 as given on page 205 of Yule and Kendall (1950). The histogram of Yule and Kendall shows clearly that the distribution is not normal. Indeed, Mardia (1970) rejects normality by his two tests and Rincon-Gallardo et al. (1979) also reject normality by both the tests they use after their transformation. In our case $M_{780}^{(2)} \approx 19.9785$ (with $L = 2$), which in comparison with the second column of Table 2 shows an extremely significant departure from bivariate normality.

The second example is the well-known *Iris setosa* data originally analysed by Fisher (1936). The data consists of 50 observations on each of four variables (sepal length, sepal width, petal length, petal width), and it is commonly believed that, in some form or other, it is from a quadrivariate normal distribution. Rincon-Gallardo et al. (1979) accept this hypothesis for the original data by both of their tests. In our case $M_{50}^{(4)} \geq 5.6967$ (we computed the latter value with only $L = 1$ in the approximate four-variate formula) which in comparison with the fourth column of Table 2 shows a highly significant departure from quadrivariate normality. Contradicting Rincon-Gallardo et al. (1979), Koziol (1982, 1983) accepts the four-variate normality of the *logarithms* of the original *Iris setosa* data as given in Gnanadesikan [(1977), page 219]. For these logarithms, we obtained $M_{50}^{(4)} \geq 5.8845$ (again with $L = 1$), so that the significance of departure from normality is even higher than for the original data. Hence we reject the hypothesis of quadrivariate normality of both the original and the logarithmic data.

The limited experience in the present and preceding sections suggests that the test based on $M_n^{(d)}$ may be highly sensitive. The only arbitrary element in our test $M_n^{(d)}(T)$ is the choice $T = 1.47/\sqrt{d}$. It is conceivable that the larger is T, the more powerful will be the test. Of course, the A_d and hence the V_d constants belonging to another choice $T = T_0/\sqrt{d} > 1.47/\sqrt{d}$ can be easily recomputed, and then the subsequent bounds $z_d(\alpha)$ can also be obtained. However, the larger is T_0, the slower is the convergence of $M_n^{(d)}(T_0/\sqrt{d})$ and also more computer time is needed for the computation of the statistic. We believe that our choice $T_0 = 1.47$, with the given motivation, is a reasonable compromise.

8. The Cramér–von Mises type statistic. Another plausible statistic based on the process $Z_n(\cdot)$ would be to consider

$$\int_{[-T,T]^d} Z_n^2(t)\,dt,$$

with some $T > 0$, which by Theorem 2.2 converges in distribution to $\sum_{j=1}^{\infty}\lambda_j W_j^2$, where $W_1, W_2, \ldots$ are independent standard normal random variables and $\lambda_j = \lambda_j(d, T)$ are the solutions of the eigenvalue–eigenfunction equation

$$\int_{[-T,T]^d} \sigma(s,t)\phi(s)\,ds = \lambda\phi(t),$$

with the covariance function $\sigma(\cdot,\cdot)$ given in Theorem 2.2. It would be desirable to compute numerically a sufficient set of the largest eigenvalules to approximate adequately the limiting distribution. A referee suggested a discrete approximation to the above equation using the covariance matrix $\sigma(s_k, t_k)$ on some appropriate grid $\{(s_k, t_k)\}$. [See, for example, Schilling (1983b).]

APPENDIX

PROOF OF THEOREM 2.1. We have

$$Z_n(t) = n^{1/2}\{U_n(S_n^{-1/2}t) - U(t)\}\{U_n(S_n^{-1/2}t) + U(t)\}$$
$$+n^{1/2}\{V_n(S_n^{-1/2}t) - V(t)\}\{V_n(S_n^{-1/2}t) + V(t)\}.$$

Hence, on account of the fact that $S_n^{-1/2} \to I$ almost surely and the triviality (Csörgő, 1981) that C_n converges almost surely uniformly to C on any bounded set in $\mathbb{R}^d$, the theorem will follow once we have shown that the complex valued processes

$$Y_n(t) = n^{1/2}\{C_n(S_n^{-1/2}t) - C(t)\}$$

converge weakly in $\mathscr{C}^2$ to the complex Gaussian process $Y(t)$. We proceed to prove this.

Applying the one-term d-variate Taylor formula to the second exponential function in the third line at (2.1), we obtain

$$Y_n(t) = n^{-1/2} \sum_{j=1}^{n} \{\exp(i\langle t, X(j)\rangle) - C(t)$$
$$+\langle (S_n^{-1/2} - I)t, \nabla C_n(t)\rangle\} + A_n(t),$$

where

$$\sup_{t\in[-T,T]^d} |A_n(t)| \le \sup_{t\in[-T,T]^d} \frac{1}{2} n^{1/2} \sum_{j=1}^{n} \langle (S_n^{-1/2} - I)t, X(j)\rangle^2$$
$$\le (2n)^{-1} \sum_{j=1}^{n} |X(j)|^2 \sup_{t\in[-T,T]^d} n^{1/2}|(S_n^{-1/2} - I)t|^2$$
$$= O(n^{-1/2}\log\log n)$$

almost surely by the law of the large numbers and the law of the iterated logarithm, the latter being applied to the elements of $(S_n - I)$ after the rearrangement

$$S_n^{-1/2} - I = -S_n^{-1/2}(I + S_n^{1/2})^{-1}(S_n - I). \tag{A.1}$$

A result in Csörgő (1981), conveniently formulated for the present purpose in Theorem 2.1 of Ledoux (1982), implies that

$$\sup_{t\in[-T,T]^d} |\nabla C_n(t) - \nabla C(t)| = O\left(n^{-1/2}(\log\log n)^{1/2}\right)$$

almost surely. Whence, using (A.1), we see that

$$\sup_{t\in[-T,T]^d} |\langle n^{1/2}(S_n^{-1/2} - I)t, \nabla C_n(t)\rangle + \tfrac{1}{2}\langle n^{1/2}(S_n - I)t, \nabla C(t)\rangle|$$
$$= O(n^{-1/2}\log\log n)$$

almost surely. Let $1 \le m \le d$. A simple computation justifies that the mth component of the vector $n^{1/2}(S_n - I)t$ is

$$n^{-1/2}\sum_{j=1}^{d}\left\{(X_m^2(j) - 1)t_m + X_m(j)\sum_{k=1,\,k\ne m}^{d} X_k(j)t_k\right\} - n^{1/2}\overline{X}_m(n)\langle\overline{X}(n), t\rangle.$$

The supremum of the second term here, over $[-T, T]^d$, is again $O(n^{-1/2}\log\log n)$ almost surely by the classical log log law. Hence

$$Y_n(t) = n^{-1/2}\sum_{j=1}^{n} R_j(t) + B_n(t), \tag{A.2}$$

where $R_j(t) = R(t; X(j))$ with $R(t; X)$ as in (2.3), and

$$\sup_{t\in[-T,T]^d} |B_n(t)| = O(n^{-1/2}\log\log n)$$

almost surely. Since the random functions $R_j(t)$ in the representation (A.2) are independent and identically distributed with mean zero and covariance $\rho(s, t)$, the multidimensional central limit theorem implies that the finite-dimensional distributions of $Y_n(\cdot)$ converge to those of $Y(\cdot)$. The tightness of $\{Y_n(\cdot)\}$ follows from (2.4), and hence the theorem.

PROOF OF LEMMA 4.1. First we fix the lengths $u = |t|$, $v = |s|$ of the vectors and let the inner product $x = \langle s, t\rangle$ vary in its range $0 \le x \le uv \le 2^{-1}(u^2 + v^2)$. We have

$$A_d = \sup_{0\le u,\,v\le 1.47}\ \sup_{0\le x\le uv} f_{u,v}(x),$$

where

$$f_{u,v}(x) = 4\frac{e^{-2u^2}g(u^2) + e^{-2v^2}g(v^2) - 2e^{-u^2-v^2}g(x)}{(u^2 + v^2 - 2x)^2}$$

with derivative

$$f'_{u,v}(x) = 4\left[\frac{2\{e^{-2u^2}g(u^2) + e^{-2v^2}g(v^2) - 2e^{-u^2-v^2}g(x)\}}{(u^2 + v^2 - 2x)^2}\right.$$
$$\left.- \frac{2e^{-u^2-v^2}g'(x)(u^2 + v^2 - 2x)}{(u^2 + v^2 - 2x)^2}\right].$$

Clearly,

$$A_d \ge \sup_{0\le u,\,v\le 1.47} f_{u,v}(uv) = A_1,$$

and

$$A_d \geq \sup_{0 \leq u \leq 1.47} \sup_{0 \leq x \leq u^2} f_{u=v}(x)$$
$$= \sup_{0 \leq u \leq 1.47} 4e^{-2u^2} \sup_{0 \leq x \leq u^2} \frac{g(u^2) - g(x)}{u^2 - x}$$
$$= B$$

since g is a convex function. Hence $A_d \geq \max(A_1, B)$.

To prove the reverse inequality, consider the alternative:

$$e^{-2u^2} g(u^2) + e^{-2v^2} g(v^2) \leq \text{ (or } >) \, 2e^{-u^2-v^2} g\left(\tfrac{1}{2}(u^2+v^2)\right).$$

If " $\leq$ ", then

$$f_{u,v}(x) \leq 4e^{-u^2-v^2} \frac{g\left(\frac{1}{2}(u^2+v^2)\right) - g(x)}{\frac{1}{2}(u^2+v^2) - x}$$
$$\leq 4e^{-u^2-v^2} g'\left(\tfrac{1}{2}(u^2+v^2)\right)$$

by the convexity of g. In the opposite case the numerator of $f'_{u,v}(x)$ is larger than

$$16e^{-u^2-v^2}\left\{g\left(\tfrac{1}{2}(u^2+v^2)\right) - g(x) - g'(x)\left\{\tfrac{1}{2}(u^2+v^2) - x\right\}\right\}$$

and this lower bound is nonnegative again by the convexity of g. So if " $>$ ", then $f_{u,v}(x) \leq f_{u,v}(uv)$. Hence $A_d \leq \max(A_1, B)$ and the lemma is proved.

Acknowledgments. I am grateful to David Mason and Vilmos Totik for their comments and to Pál Tőke for programming the computations in Sections 6 and 7. The comments of two referees are also appreciated.

REFERENCES

AITKIN, M. A. (1972). A class of tests for multivariate normality based on linear functions of order statistics. Unpublished manuscript.

ANDERSON, T. W. (1966). Some nonparametric multivariate procedures based on statistically equivalent blocks. In *Multivariate Analysis* (P. R. Krishnaiah, ed.) 5–27. Academic, New York.

ANDREWS, D. F., GNANADESIKAN, R. and WARNER, J. L. (1971). Transformations of multivariate data. *Biometrics* **27** 825–840.

ANDREWS, D. F., GNANADESIKAN, R. and WARNER, J. L. (1972). Methods for assessing multivariate normality. Bell Laboratories Memorandum.

ANDREWS, D. F., GNANADESIKAN, R. and WARNER, J. L. (1973). Methods for assessing multivariate normality. In *Multivariate Analysis* **3** (P. R. Krishnaiah, ed.) 95–116. Academic, New York.

BICKEL, P. J. and BREIMAN, L. (1983). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. *Ann. Probab.* **11** 185–214.

BORELL, C. (1975). The Brunn–Minkowski inequality in Gauss space. *Invent. Math.* **30** 207–216.

COX, D. R. (1968). Notes on some aspects of regression analysis. *J. Roy Statist. Soc. Ser A* **131** 265–279.

COX, D. R. and SMALL, N. J. H. (1978). Testing multivariate normality. *Biometrika* **65** 263–272.

CSÖRGŐ, S. (1981). Multivariate empirical characteristic functions. *Z. Wahrsch. verw. Gebiete* **55** 203–229.

DAHIYA, R. C. and GURLAND, J. (1973). A test of fit for bivariate distributions. *J. Roy. Statist. Soc. Ser. B* **35** 452–465.

DAY, N. R. (1969). Divisive cluster analysis and a test for multivariate normality. *Bull. Inst. Internat. Statist.* **43** 110–112.

DEWET, T., VENTER, J. H. and VAN WYK, J. W. J. (1979). The null distributions of some test criteria of multivariate normality. *South African Statist. J.* **13** 153–176.

FERNIQUE, X. (1975). Régularité des trajectoires des fonctions aléatoires gaussiennes. In *École d'Été de Probabilités de Saint-Flour, IV-1974. Lecture Notes in Math.* **480** 1–96. Springer, Berlin.

FISHER, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics* **7** 179–188. (Also paper 32 in *Contributions to Mathematical Statistics, R. A. Fisher.* Wiley, New York.)

GIORGI, G. M. and FATTORINI, L. (1976). An empirical study of some tests for multivariate normality. *Quaderni dell'Instituto di Statistica* No. 20 (8 pages). Siena.

GNANADESIKAN, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations.* Wiley, New York.

GNANADESIKAN, R. and KETTENRING, J. R. (1972). Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics* **28** 81–124.

HALL, P. and WELSH, A. H. (1983). A test for normality based on the empirical characteristic function. *Biometrika* **70** 485–489.

HAWKINS, D. M. (1981). A new test for multivariate normality and homoscedasticity. *Technometrics* **23** 105–110.

HEALY, M. J. R. (1968). Multivariate normal plotting. *Appl. Statist.* **17** 157–161.

HENSLER, G. L., MEHROTA, K. G. and MICHALEK, J. E. (1977). A goodness of fit test for multivariate normality. *Comm. Statist. A—Theory Methods* **6** 33–41.

KESSEL, D. L. and FUKUNAGA, K. (1972). A test for multivariate normality with unspecified parameters. Unpublished report, Purdue Univ.

KOZIOL, J. A. (1982). A class of invariant procedures for assessing multivariate normality. *Biometrika* **69** 423–427.

KOZIOL, J. A. (1983). On assessing multivariate normality. *J. Roy. Statist. Soc. Ser B.* **45** 358–361.

LEDOUX, M. (1982). Loi du logarithme itéré dans $\mathscr{C}(S)$ et fonction caracteristique empirique. *Z. Wahrsch verw. Gebiete* **60** 425–435.

LUKACS, E. (1970). *Characteristic Functions.* Griffin, London.

MALKOVICH, J. F. (1971). Tests for multivariate normality. Unpublished Ph.D. thesis, Univ. of California, Los Angeles.

MALKOVICH, J. F. and AFIFI, A. A. (1973). On tests for multivariate normality. *J. Amer. Statist. Assoc.* **68** 176–179.

MARDIA, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika* **57** 519–530.

MARDIA, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis for testing normality and robustness studies. *Sankhyā Ser. A* **36** 115–128.

MARDIA, K. V. (1975). Assessment of multinormality and the robustness of Hotelling's T^2 test. *J. Roy. Statist. Soc. Ser. C* **24** 163–171.

MARDIA, K. V. (1980). Tests of univariate and multivariate normality. In *Handbook of Statistics 1. Analysis of Variance* (P. R. Krishnaiah, ed.) 279–320. North-Holland, Amsterdam.

MARDIA, K. V. and ZEMROCH, P. J. (1975). Algorithm AS84. Measures of multivariate skewness and kurtosis. *J. Roy. Statist. Soc. Ser. C* **24** 262–265.

MOORE, D. S. and STUBBLEBINE, J. B. (1981). Chi-square tests for multivariate normality with applications to common stock prices. *Comm. Statist. A—Theory Methods* **10** 713–738.

MUROTA, K. (1981). Test for normality based on the empirical characteristic function. *Rep. Statist. Appl. Res. Un. Japan Sci. Engrs.* **28** 1–17.

MUROTA, K. and TAKEUCHI, K. (1981). The studentized empirical characteristic function and its application to test for the shape of distribution. *Biometrika* **68** 55–65.

PETTITT, A. N. (1979). Testing for bivariate normality using the empirical distribution function. *Comm. Statist. A—Theory Methods* **8** 699–712.

RINCON-GALLARDO, S., QUESENBERRY, C. P. and O'REILLY, F. J. (1979). Conditional probability integral transformations and goodness-of-fit tests for multivariate normal distributions. *Ann. Statist.* **7** 1052–1057.

SARKADI, K. and TUSNÁDY, G. (1977). Testing for normality and for the exponential distribution. In *Proc. Fifth Conference on Probability Theory 1974, Brasov* (B. Bereanu, M. Josifescu, and G. Popescu, eds.) 99–118. Editura Academici, Bucharest.

SCHILLING, M. F. (1983a). Goodness of fit testing in $\mathbb{R}^m$ based on the weighted empirical distribution of certain nearest neighbor statistics. *Ann. Statist.* **11** 1–12.

SCHILLING, M. F. (1983b). An infinite-dimensional approximation for nearest neighbor goodness of fit tests. *Ann. Statist.* **11** 13–24.

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

SMALL, N. J. H. (1980). Marginal skewness and kurtosis in testing multivariate normality. *Appl. Statist.* **29** 85–87.

WAGLE, B. (1968). Multivariate beta distribution and a test for multivariate normality. *J. Roy. Statist. Soc. Ser. B* **30** 511–536.

WELSH, A. H. (1985). A note on scale estimates based on the empirical characteristic function and their application to test for normality. *Statist. Probab. Letters* **3**.

WEISS, L. (1958). A test of fit for multivariate distributions. *Ann. Math. Statist.* **29** 595–599.

WILK, M. B. and GNANADESIKAN, R. (1968). Probability plotting methods for the analysis of data. *Biometrika* **55** 1–17.

YANG, S.-S. (1981). Linear combinations of concomitants of order statistics with applications to testing and estimation. *Ann. Inst. Statist. Math.* **33** 463–470.

YULE, G. U. and KENDALL, M. G. (1950). *An Introduction to the Theory of Statistics.* 14th ed. Hafner, New York.

BOLYAI INSTITUTE
SZEGED UNIVERSITY
ARADI VÉRTANÚK TERE 1
H-6720 SZEGED
HUNGARY