The Annals of Statistics
1992, Vol. 20, No. 4, 1889–1902

LIMIT DISTRIBUTIONS FOR MARDIA'S MEASURE OF MULTIVARIATE SKEWNESS

BY L. BARINGHAUS AND N. HENZE

Universität Hannover and Universität Karlsruhe

We study the asymptotic behavior of Mardia's measure of (sample) multivariate skewness. In the special case of an elliptically symmetric distribution, the limit law is a weighted sum of two independent χ^2-variates. A normal limit distribution arises if the population distribution has positive skewness. These results explain some curiosities in the power performance of a commonly proposed test for multivariate normality based on multivariate skewness.

1. Introduction and summary. Let $X_1, \dots, X_n$ be independent observations on a d-dimensional random column vector X with expectation $E[X] = \mu$ and nonsingular covariance matrix $\Sigma = E[(X-\mu)(X-\mu)']$, where the prime denotes transpose. Writing

$$\overline{X} = \frac{1}{n}\sum_{j=1}^{n} X_j, \qquad S = \frac{1}{n}\sum_{j=1}^{n}\left(X_j - \overline{X}\right)\left(X_j - \overline{X}\right)'$$

for the mean vector and the empirical covariance matrix of $X_1, \dots, X_n$, Mardia (1970, 1975) introduced

$$b_{1,d} = \frac{1}{n^2}\sum_{i,j=1}^{n}\left\{\left(X_i - \overline{X}\right)' S^{-1}\left(X_j - \overline{X}\right)\right\}^3$$

as an affine invariant measure of multivariate skewness and showed that it emerges in a natural way in connection with robustness studies on Hotelling's T^2-test [Mardia (1974)]. An algorithm for computing $b_{1,d}$ was given by Mardia and Zemroch (1975). For a survey on measures of multivariate skewness and kurtosis, see Schwager (1985). Obviously, the affine invariant "population counterpart" of $b_{1,d}$ is

$$\beta_{1,d} = E\left[\{(X_1 - \mu)'\Sigma^{-1}(X_2 - \mu)\}^3\right].$$

Within the class $\mathcal{N}_d$ of all nondegenerate d-dimensional normal distributions, $\beta_{1,d}$ takes the value 0. Denoting by P^X the distribution of X, Mardia proposed to use $b_{1,d}(X_1, \dots, X_n)$ as a statistic for testing the hypothesis H_0: $P^X \in \mathcal{N}_d$ of multivariate normality. Let

$$W = \Sigma^{-1/2}(X - \mu), \tag{1.1}$$

Received June 1990; revised February 1991.

AMS 1980 *subject classifications*. Primary 62H15; secondary 62H10.

Key words and phrases. Multivariate skewness, test for multivariate normality, consistency, elliptically symmetric distributions, V-statistics.

where $\Sigma^{-1/2}$ is a symmetric positive definite square root of Σ^{-1} so that $E[W] = 0$ and $E[WW'] = I_d$, the identity matrix of order d. Writing $W = (W_1, W_2, \dots, W_d)'$, we have

$$0 \le \beta_{1,d} = \sum_{i=1}^{d} \left(E\left[W_i^3\right]\right)^2 + 3\sum_{i\neq j}\left(E\left[W_i^2 W_j\right]\right)^2 + 6 \sum_{1\le i<j<k\le d}\left(E\left[W_i W_j W_k\right]\right)^2.$$

Consequently, rejection of H_0 is for large values of $b_{1,d}(X_1, \dots, X_n)$. Under H_0

$$nb_{1,d}(X_1, \dots, X_n) \to_{\mathscr{D}} 6\chi^2_{d(d+1)(d+2)/6} \qquad (1.2)$$

as $n \to \infty$ [Mardia (1970)], where "$\to_{\mathscr{D}}$" denotes convergence in distribution. Observe that $b_{1,d}$ is an estimator for the population parameter $\beta_{1,d}$ which is zero not only in case of normality but also within the wider class of all elliptically symmetric distributions (see Section 2). Therefore, it may be supposed that the test for multinormality based on $b_{1,d}$ is consistent only against alternative distributions having positive multivariate skewness.

It is the purpose of this paper to provide the asymptotic behavior of multivariate skewness for a wide class of multivariate distributions. If the underlying distribution is elliptically symmetric, it turns out that

$$nb_{1,d}(X_1, \dots, X_n) \to_{\mathscr{D}} \alpha_1\chi^2_d + \alpha_2\chi^2_{d(d-1)(d+4)/6},$$

where the coefficients α_i depend on $E[|W|^4]$ and $E[|W|^6]$, with W given in (1.1). Here and in what follows, $|x|$ denotes the Euclidean norm of a vector x. On the other hand, if $E[(x'W)^3]$ is not constant P^W-almost surely, we have

$$\sqrt{n}\left(b_{1,d}(X_1, \dots, X_n) - \beta_{1,d}\right) \to_{\mathscr{D}} \mathscr{N}(0, \sigma^2)$$

with σ^2 depending on the distribution of W.

The main results will be presented in Sections 2 and 3. In Section 4 we clarify some curiosities in the power performance of Mardia's test for multivariate normality based on $b_{1,d}$ observed in Monte Carlo studies. Moreover, it will be seen that this test is consistent for a fixed alternative distribution if and only if $\beta_{1,d} > 0$.

2. The limit distribution of $b_{1,d}$ for elliptically symmetric distributions. A random $(d \times 1)$ vector X is said to have a *spherically symmetric distribution* (or simply *spherical distribution*) if

$$X =_{\mathscr{D}} HX \quad \text{for every orthogonal } (d\times d) \text{ matrix } H,$$

where the symbol "$=_{\mathscr{D}}$" denotes equality in distribution. A random $(d \times 1)$ vector X is said to have an *elliptically symmetric distribution* (or simply *elliptical distribution*) *with center* $\mu \in \mathbb{R}^d$ *and ellipticity matrix* Δ if there is a random $(k \times 1)$ vector Y having spherically symmetric distribution and a $(k \times d)$ matrix A of rank k such that $\Delta = A'A$ and

$$X =_{\mathscr{D}} \mu + A'Y.$$

In the following we assume that Δ is positive definite and that $P(X = \mu) = 0$. Since Mardia's coefficient of skewness $b_{1,d}$ is invariant with respect to affine linear transformations of $X_1, \dots, X_n$, that is, $b_{1,d}(X_1, \dots, X_n) = b_{1,d}(X_1^*, \dots, X_n^*)$, where $X_i^* = b + BX_i$ for a nonsingular $(d \times d)$ matrix B and a vector $b \in \mathbb{R}^d$, we can (and do) assume without loss of generality that

$$E[X] = 0 \quad \text{and} \quad E[XX'] = I_d.$$

In other words, the distribution of X is spherical with

$$E\left[|X|^2\right] = d. \tag{2.1}$$

We need the further assumption

$$E\left[|X|^6\right] < \infty, \tag{2.2}$$

which guarantees that mixed moments of sufficiently high order exist.

Note that $|X|$ and $X/|X|$ are independent and that $X/|X|$ is uniformly distributed on the surface of the unit d-sphere [see Fang, Kotz and Ng (1989), page 30]. Then $\beta_{1,d} = 0$ because this is true for the uniform distribution on the surface of the unit d-sphere [see Fang, Kotz and Ng (1989), page 72]. Letting

$$V_n(X_1, \dots, X_n) = \frac{1}{n^2} \sum_{i,j=1}^{n} h(X_i, X_j) \tag{2.3}$$

be the V-statistic with kernel

$$h(x, y) = (x'y)^3, \qquad x, y \in \mathbb{R}^d \tag{2.4}$$

[see Serfling (1980), page 174], we have

$$b_{1,d}(X_1, \dots, X_n) = V_n\left(S^{-1/2}(X_1 - \overline{X}), \dots, S^{-1/2}(X_n - \overline{X})\right),$$

where $S^{-1/2}$ is a symmetric positive definite square root of S^{-1} which exists almost surely if $n \geq d + 1$ [see Eaton and Perlman (1973)].

Lemma 2.1. *If the distribution of X is elliptical we have*

$$\frac{1}{n} \sum_{i,j=1}^{n} \left[(X_i - \overline{X})' S^{-1}(X_j - \overline{X})\right]^3 = \frac{1}{n} \sum_{i,j=1}^{n} h_*(X_i, X_j) + o_P(1), \tag{2.5}$$

where the kernel h_* is given by

$$h_*(x, y) = (x'y)^3 - 3(|x|^2 + |y|^2)x'y + 3(d + 2)x'y \tag{2.6}$$

and satisfies

$$E[h_*(x, X)] = 0, \qquad x \in \mathbb{R}^d. \tag{2.7}$$

Proof. For spherically distributed X we have

$$E[h(x, X)] = E\left[(x'X)^3\right] = 0, \qquad x \in \mathbb{R}^d,$$

which shows that the kernel h figuring in (2.4) is degenerate. Since (2.2) entails $E[h^2] < \infty$, standard results [see, e.g., Gregory (1977)] yield that

$$nV_n = \frac{1}{n} \sum_{i,j=1}^{n} (X_i' X_j)^3$$

has a nondegenerate limit distribution. Using this fact, the idea is to expand

$$\frac{1}{n} \sum_{i,j=1}^{n} \left[(X_i - \overline{X}) S^{-1} (X_j - \overline{X}) \right]^3 \tag{2.8}$$

by neglecting terms which are of order $o_P(1)$ as $n \to \infty$.

To this end, observe that

$$\frac{1}{n} \sum_{i=1}^{n} |X_i|^2 X_i = E\left[|X|^2 X \right] + o_P(1) = o_P(1)$$

because $E[|X|^2 X] = 0$ for spherically distributed X satisfying (2.2). Furthermore, letting

$$\mathbf{A}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i X_i' - I_d), \tag{2.9}$$

the multivariate central limit theorem yields $\mathbf{A}_n = O_P(1)$ and $\sqrt{n}\,\overline{X} = O_P(1)$. Consequently,

$$S = I_d + \frac{1}{\sqrt{n}} \mathbf{A}_n + O_P\left(\frac{1}{n} \right)$$

and thus

$$S^{-1} = I_d - \frac{1}{\sqrt{n}} \mathbf{A}_n + O_P\left(\frac{1}{n} \right). \tag{2.10}$$

A tedious but straightforward evaluation of the squared bracket occurring in (2.8) leads to 64 terms, most of which are of order $o_P(1)$ and thus are asymptotically negligible. The only asymptotically nonvanishing terms (ignoring symmetric cases) are

$$\frac{1}{n} \sum_{i,j} \left(X_i' S^{-1} X_j \right)^3 = \frac{1}{n} \sum_{i,j} (X_i' X_j)^3 + o_P(1),$$

$$\frac{1}{n} \sum_{i,j} \left(X_i' S^{-1} X_j \right)^2 X_j' S^{-1} \overline{X} = \frac{1}{n} \sum_{i,j} |X_j|^2 X_j' X_i + o_P(1),$$

$$\frac{1}{n} \sum_{i,j} \left(X_i' S^{-1} X_j \right)^2 \overline{X}' S^{-1} \overline{X} = E\left[|X|^2 \right] \frac{1}{n} \sum_{i,j} X_i' X_j + o_P(1)$$

and

$$\frac{1}{n} \sum_{i,j} X_i' S^{-1} X_j X_j' S^{-1} \overline{X}\,\overline{X}' S^{-1} X_i = \frac{1}{n} \sum_{i,j} X_i' X_j + o_P(1).$$

Counting the number of symmetric cases, we finally obtain

$$\frac{1}{n}\sum_{i,j=1}^{n}\left[(X_i-\bar{X})'S^{-1}(X_j-\bar{X})\right]^3$$
$$=\frac{1}{n}\sum_{i,j=1}^{n}\Big[(X_i'X_j)^3-3\left(|X_i|^2+|X_j|^2\right)X_i'X_j$$
$$+3\left(E\left[|X|^2\right]+2\right)X_i'X_j\Big]+o_P(1).$$

Recalling (2.1), the result (2.5) follows; (2.7) is an immediate consequence of the spherical symmetry of X. □

Lemma 2.1 implies that $nb_{1,d}$ and nV_n^*, where

$$V_n^*(X_1,\dots,X_n)=\frac{1}{n^2}\sum_{i,j=1}^{n}h_*(X_i,X_j),$$

have the same limit distribution. However, from standard theorems on the asymptotic behavior of V-statistics in the degenerate case [recall (2.7)], we have

$$nV_n^*\to_{\mathscr{D}}\sum_{k\geq 1}\lambda_k U_k^2,$$

where $U_1,U_2,\dots$ are independent unit normal random variables. The λ_k, $k\geq 1$, are the nonzero eigenvalues corresponding to the integral operator

$$g\to Ag(x)=\int h_*(x,y)g(y)P^X(dy) \tag{2.11}$$

on the Hilbert space $L_2(P^X)$ of measurable functions g on $\mathbb{R}^d$ that are square-integrable with respect to P^X associated with the kernel h_* [see, e.g., Gregory (1977)]. More precisely, we have the following result.

THEOREM 2.2. *Let X have an elliptical distribution with expectation μ and nonsingular covariance matrix Σ such that $E[\{(X-\mu)'\Sigma^{-1}(X-\mu)\}^3]<\infty$. Then*

$$nb_{1,d}(X_1,\dots,X_n)\to_{\mathscr{D}}\alpha_1\chi^2{}_d+\alpha_2\chi^2{}_{d(d-1)(d+4)/6},$$

where

$$\alpha_1=\frac{3}{d}\left[\frac{m_6}{d+2}-2m_4+d(d+2)\right],$$
$$\alpha_2=\frac{6m_6}{d(d+2)(d+4)}$$

and

$$m_{2j}=E\left[\{(X-\mu)'\Sigma^{-1}(X-\mu)\}^j\right],\qquad j=2,3.$$

$\chi^2{}_d$ and $\chi^2{}_{d(d-1)(d+4)/6}$ are independent χ^2-variables with d and $d(d-1)(d+4)/6$ degrees of freedom.

PROOF. We may assume that $d \geq 2$ since in the univariate case the assertion of Theorem 2.2 follows from Theorem 1 of Gastwirth and Owens (1977). Let X have a spherical distribution with $P(X=0)=0$, $E[|X|^2]=d$ and $E[|X|^6]<\infty$. To determine the eigenvalues of the integral operator A given in (2.11), we put $R=|X|$ and $Z=X/R$. Remember that R and Z are independent, with Z being uniformly distributed on the unit sphere $S^{d-1}=\{x \in \mathbb{R}^d: |x|=1\}$. Let ω denote the uniform distribution on S^{d-1}, and let F be the distribution of R. Note that (2.11) takes the special form

$$\iint\left[r^3s^3(z'w)^3 - 3(r^2+s^2)rs(z'w) + 3(d+2)rs(z'w)\right] \times g(sw)\,d\omega(w)\,dF(s) = \lambda g(rz), \tag{2.12}$$

which holds for $F \otimes \omega$-almost all $(r,z) \in [0,\infty) \times S^{d-1}$.

We now treat the cases $d \geq 3$ and $d=2$ separately. First, let $d \geq 3$. Denoting by $C_q^\gamma(t)$ the Gegenbauer polynomial of degree q and order $\gamma = (d-2)/2$ [see Erdélyi, Magnus, Oberhettinger and Tricomi (1953), Section 10.9], we have

$$t = \frac{1}{2\gamma}C_1^\gamma(t) \quad \text{and} \quad t^3 = \frac{3}{4(2+\gamma)(1+\gamma)\gamma}C_3^\gamma(t) + \frac{3}{4(2+\gamma)\gamma}C_1^\gamma(t).$$

Thus (2.12) can be rewritten as

$$\frac{3}{d-2}\iint\left[\frac{2C_3^\gamma(z'w)r^3s^3}{d(d+2)} + \left(\frac{r^3s^3}{d+2} - (r^2+s^2)rs + (d+2)rs\right)C_1^\gamma(z'w)\right]g(sw)\,d\omega(w)\,dF(s) = \lambda g(rz). \tag{2.13}$$

Putting

$$\nu(q) = \binom{d-3+q}{d-2} + \binom{d-2+q}{d-2},$$

there is a complete system $\{\varphi_{q,k}: k=1,\ldots,\nu(q);\ q=0,1,\ldots\}$ of orthonormal continuous functions $\varphi_{q,k} \in L_2(\omega)$ such that

$$C_q^\gamma(z'w) = \left(1+\frac{q}{\gamma}\right)^{-1}\sum_{k=1}^{\nu(q)}\varphi_{q,k}(z)\varphi_{q,k}(w), \qquad z,w \in S^{d-1}$$

[see Erdélyi, Magnus, Oberhettinger and Tricomi (1953), page 243, and Stein and Weiss (1971)]. It is easily verified that the functions

$$g_{1,k}(x) = \frac{\left[r^3-(d+2)r\right]\varphi_{1,k}(z)}{\left(m_6 - 2(d+2)m_4 + (d+2)^2d\right)^{1/2}}, \tag{2.14}$$

$x \in \mathbb{R}^d$, $r = |x|$, $z = x/r$, $k = 1, \dots, \nu(1)$, and

$$g_{3,k}(x) = \frac{r^3 \varphi_{3,k}(z)}{m_6^{1/2}}, \tag{2.15}$$

$x \in \mathbb{R}^d$, $r = |x|$, $z = x/r$, $k = 1, \dots, \nu(3)$, are orthonormal eigenfunctions with associated nonzero eigenvalues

$$\lambda_{1,k} = \frac{3}{d}\left[\frac{m_6}{d+2} - 2m_4 + (d+2)d\right] \tag{2.16}$$

and

$$\lambda_{3,k} = \frac{6m_6}{d(d+2)(d+4)}. \tag{2.17}$$

Addition of $\lambda_{1,k}$ and $\lambda_{3,k}$ according to their multiplicities $\nu(1) = d$ and $\nu(3) = d(d-1)(d+4)/6$ yields

$$\sum_{k=1}^{\nu(1)} \lambda_{1,k} + \sum_{k=1}^{\nu(3)} \lambda_{3,k} = m_6 - 6m_4 + 3(d+2)d = E[h_*(X,X)].$$

This shows that we have obtained all nonzero eigenvalues of the kernel h_* [see, e.g., Serfling (1980)], so that the assertion of Theorem 2.2 follows for $d \geq 3$. The case $d = 2$ can be treated in a similar manner by considering Chebyshev polynomials instead of Gegenbauer polynomials. Denoting by $C_q(t)$ the Chebyshev polynomial of the first kind and degree q [see Erdélyi, Magnus, Oberhettinger and Tricomi (1953), Section 10.11], we have

$$t = C_1(t) \quad \text{and} \quad t^3 = \tfrac{1}{4}C_3(t) + \tfrac{3}{4}C_1(t).$$

Then, starting from (2.12) and expressing the powers of $(z'w)$ in terms of Chebyshev polynomials, one gets the similar expression to (2.13). Now there is a complete system $\{1\} \cup \{\varphi_{q,k}: k = 1, 2;\ q = 1, 2, \dots\}$ of orthonormal continuous functions $\varphi_{q,k} \in L_2(\omega)$ such that

$$C_q(z'w) = \tfrac{1}{2}\left[\varphi_{q,1}(z)\varphi_{q,1}(w) + \varphi_{q,2}(z)\varphi_{q,2}(w)\right], \qquad z, w, \in S^1.$$

Keeping these new notations in mind, the functions given in (2.14) and (2.15) turn out to be orthonormal eigenfunctions with associated nonzero eigenvalues given in (2.16) and (2.17) with $d = 2$. Since these eigenvalues have multiplicity 2, the assertion of Theorem 2.2 is shown to be true also for $d = 2$. □

EXAMPLE 2.3 (Normal distribution). If X has a normal distribution with expectation μ and nonsingular covariance matrix Σ, the quantities m_4 and m_6 figuring in the statement of Theorem 2.2 are

$$m_4 = d(d+2), \qquad m_6 = d(d+2)(d+4).$$

From this it follows that $\alpha_1 = \alpha_2 = 6$. Since $d + d(d-1)(d+4)/6 = d(d+1)(d+2)/6$, we obtain Mardia's result (1.2).

EXAMPLE 2.4 (Symmetric multivariate Pearson Type II distribution). The random vector X is said to have a *symmetric multivariate Pearson Type II distribution* [denoted by $X \sim \mathrm{MPII}_d(\kappa, \mu, \Delta)$] if X has the density

$$f(x) = \frac{\Gamma(d/2 + \kappa + 1)}{\Gamma(\kappa + 1)\pi^{d/2}|\Delta|^{1/2}}\left(1 - (x - \mu)'\Delta^{-1}(x - \mu)\right)^{\kappa}$$
$$\times I\{(x - \mu)'\Delta^{-1}(x - \mu) \leq 1\},$$

for some vector $\mu \in \mathbb{R}^d$ and some symmetric positive definite matrix Δ, $\kappa \in \mathbb{R}$, $\kappa > -1$ [see Fang, Kotz and Ng (1989), Section 3.4]. We have $E[X] = \mu$ and

$$E[(X - \mu)(X - \mu)'] = \frac{1}{d + 2\kappa + 2}\Delta.$$

Since $(X - \mu)'\Delta^{-1}(X - \mu)$ has the Beta distribution with density

$$\frac{\Gamma(d/2 + \kappa + 1)}{\Gamma(d/2)\Gamma(\kappa + 1)}u^{d/2 - 1}(1 - u)^{\kappa}, \qquad 0 \leq u \leq 1,$$

it is easily seen that for $X \sim \mathrm{MPII}_d(\kappa, \mu, \Delta)$

$$m_4 = d(d + 2)\frac{d + 2 + 2\kappa}{d + 4 + 2\kappa},$$

$$m_6 = d(d + 2)(d + 4)\frac{(d + 2 + 2\kappa)^2}{(d + 4 + 2\kappa)(d + 6 + 2_{\kappa})}$$

and thus

$$\alpha_1 = 6 - \frac{6[(d + 6 + 2\kappa)d + 16(\kappa + 1)]}{(d + 4 + 2\kappa)(d + 6 + 2\kappa)},$$

$$\alpha_2 = 6 - \frac{12(3d + 6\kappa + 10)}{(d + 4 + 2\kappa)(d + 6 + 2\kappa)}.$$

EXAMPLE 2.5 (Symmetric Kotz type distributions). The random vector X is said to have a *symmetric Kotz type distribution* [denoted by $X \sim \mathrm{MK}_d(a, r, s, \mu, \Delta)$] if the density of X is of the form

$$f(x) = c_d|\Delta|^{-1/2}\left[(x - \mu)'\Delta^{-1}(x - \mu)\right]^{a-1}\exp\left(-r\left[(x - \mu)'\Delta^{-1}(x - \mu)\right]^{s}\right),$$

$r, s > 0$, $2a + d > 2$, for some vector $\mu \in \mathbb{R}^d$ and some symmetric positive definite matrix Δ [see Fang, Kotz and Ng (1989), Section 3.2]. The normalizing constant c_d is

$$c_d = \frac{s\Gamma(d/2)}{\pi^{d/2}\Gamma((2a + d - 2)/(2s))}r^{(2a + d - 2)/(2s)}.$$

We have $E[X] = \mu$ and

$$E[(X - \mu)(X - \mu)'] = \frac{\Gamma((2a + d)/(2s))}{dr^{1/s}\Gamma((2a + d - 2)/(2s))}\Delta.$$

Since the density of $(X-\mu)'\Delta^{-1}(X-\mu)$ is given by

$$\frac{\pi^{d/2}}{\Gamma(d/2)} c_d u^{d/2+a-2} \exp(-ru^s), \qquad u > 0,$$

straightforward calculations yield for $X \sim \mathrm{MK}_d(a, r, s, \mu, \Delta)$

$$m_4 = d^2 \frac{\Gamma((2a+d+2)/(2s))\Gamma((2a+d-2)/(2s))}{\Gamma^2((2a+d)/(2s))},$$

$$m_6 = d^3 \frac{\Gamma((2a+d+4)/(2s))\Gamma^2((2a+d-2)/(2s))}{\Gamma^3((2a+d)/(2s))}.$$

In the special case $s = 1$ which includes the normal distribution for $a = 1$, we have for $X \sim \mathrm{MK}_d(a, r, 1, \mu, \Delta)$

$$\alpha_1 = 6\frac{d^2(d+2a) + 8(a-1)^2}{(d+2)(d+2a-2)^2},$$

$$\alpha_2 = 6\frac{d^2(2a+d)(2a+d+2)}{(d+2)(d+4)(d+2a-2)^2}.$$

EXAMPLE 2.6 (Symmetric multivariate Pearson Type VII distribution). The random vector X is said to have a *symmetric multivariate Pearson Type VII distribution* [denoted by $X \sim \mathrm{MPVII}_d(\kappa, a, \mu, \Delta)$] if X has the density

$$f(x) = \frac{\Gamma(a)}{\Gamma(a-d/2)(\pi\kappa)^{d/2}|\Delta|^{1/2}}\left(1 + \frac{1}{\kappa}(x-\mu)'\Delta^{-1}(x-\mu)\right)^{-a}$$

for some vector $\mu \in \mathbb{R}^d$ and some symmetric positive definite matrix Δ, $a > d/2$, $\kappa > 0$ [see Fang, Kotz and Ng (1989), Section 3.3]. We have $E[X] = \mu$ and

$$E[(X-\mu)(X-\mu)'] = \frac{\kappa}{2a-d-2}\Delta, \qquad a > d/2 + 1.$$

This class includes the class of *multivariate t-distributions* for $a = (d+\kappa)/2$ and $\kappa \in \mathbb{N}$. Since $(X-\mu)'\Delta^{-1}(X-\mu)$ has the density

$$\frac{1}{B(d/2, a-d/2)} \kappa^{-d/2} u^{d/2-1}\left(1 + \frac{u}{\kappa}\right)^{-a}, \qquad u > 0,$$

some calculations give

$$m_4 = d(d+2)\frac{2a-d-2}{2a-d-4}$$

$$m_6 = d(d+2)(d+4)\frac{(2a-d-2)^2}{(2a-d-4)(2a-d-6)}$$

and thus

$$\alpha_1 = 6\frac{(d+2)(2a-d+2)+(2a-d-2)^2}{(2a-d-4)(2a-d-6)},$$

$$\alpha_2 = 6\frac{(2a-d-2)^2}{(2a-d-4)(2a-d-6)}.$$

Note that m_4 and m_6 are finite if and only if $2a > d+4$ and $2a > d+6$, respectively.

3. The limit distribution of $b_{1,d}$ in the nondegenerate case. As before, assume without loss of generality that $E[X] = 0$ and $E[XX'] = I_d$. Let

$$h_1(x) = E\left[(x'X)^3\right], \qquad x \in \mathbb{R}^d. \tag{3.1}$$

We now consider the case that the kernel h figuring in (2.4) is nondegenerate, that is,

$$0 < \mathrm{Var}(h_1). \tag{3.2}$$

Since

$$h_1(x) = \sum_{i=1}^{d} x_i^3 E\left[W_i^3\right] + 3\sum_{i\neq j} x_i^2 x_j E\left[W_i^2 W_j\right] + 6\sum_{i<j<k} x_i x_j x_k E\left[W_i W_j W_k\right],$$

where $x = (x_1, \dots, x_d)'$, $X = (W_1, \dots, W_d)'$, we see that the weak assumption that the support of P^X has positive d-dimensional Lebesgue measure implies that (3.2) is equivalent to the condition $\beta_{1,d} > 0$ [see, e.g., Okamoto (1973)]. That is, we are virtually dealing with the case of positive multivariate skewness in what follows. Let

$$T_n = \binom{n}{2}^{-1} \sum_{1\le i<j\le n} h(X_i, X_j)$$

be the U-statistic associated with h, and let V_n be as in (2.3). From (2.2) we deduce that

$$V_n - T_n = o_P(n^{-1/2}), \tag{3.3}$$

which in view of the well-known result on the asymptotic normality of U-statistics with nondegenerate kernel [see, e.g., Serfling (1980), page 192], implies that $n^{1/2}(V_n - \beta_{1,d})$ has a nondegenerate normal limit distribution.

LEMMA 3.1. *Under the conditions* (2.2) *and* (3.2), *we have*

$$b_{1,d}(X_1, \dots, X_n) = V_n - \frac{3}{\sqrt{n}} tr(\mathbf{A}_n B) - 6a'\overline{X}_n + o_P(n^{-1/2}),$$

where $\mathbf{A}_n$ *is given in* (2.9),

$$a = E\left[|X|^2 X\right] \tag{3.4}$$

and

$$B = E\left[X_1(X_2'X_1)^2X_2'\right]. \tag{3.5}$$

PROOF. In view of the asymptotic normality of $\sqrt{n}(V_n - \beta_{1,d})$, we have to expand $b_{1,d}(X_1, \dots, X_n)$ neglecting terms which are of order $o_P(n^{-1/2})$. Using (2.10), it follows that

$$\begin{aligned}\frac{1}{n^2}\sum_{i,j}\left(X_i'S^{-1}X_j\right)^3 &= V_n - \frac{3}{\sqrt{n}}tr\left(\mathbf{A}_n\frac{1}{n^2}\sum_{i,j}X_j\left(X_i'X_j\right)^2X_i'\right) + o_P(n^{-1/2})\\ &= V_n - \frac{3}{\sqrt{n}}tr(\mathbf{A}_nB) + o_P(n^{-1/2}).\end{aligned}$$

Furthermore,

$$\frac{1}{n^2}\sum_{i,j}\left(X_i'S^{-1}X_j\right)^2X_j'S^{-1}\overline{X} = E\left[|X|^2X\right]'\overline{X} + o_P(n^{-1/2}).$$

Counting the number of symmetric cases yields the assertion. □

To state the main result of this section, let $B = (b_{ij})_{i,j=1,\dots,d}$, with B given in (3.5). Furthermore, let

$$\begin{aligned}u &= (2, -3b_{11}, -3b_{12}, \dots, -3b_{1d}, -3b_{21}, \dots, -3b_{dd}, -6a')'\\ &\in \mathbb{R}^{1+d^2+d},\end{aligned} \tag{3.6}$$

with a given in (3.4). Define Z_i to be the $(1 + d^2 + d)$-dimensional random vector

$$\begin{aligned}Z_i = \big(h_1(X_i) - \beta_{1,d}, X_{i1}^2 - 1, X_{i1}X_{i2}, \dots, X_{i1}X_{id},\\ X_{i2}X_{i1}, \dots, X_{id}^2 - 1, X_i'\big)',\end{aligned} \tag{3.7}$$

where h_1 is given in (3.1) and $X_i = (X_{i1}, \dots, X_{id})'$.

THEOREM 3.2. *Under the conditions* $E[X] = 0$, $E[XX'] = I_d$, $E[|X|^6] < \infty$ *and* $\operatorname{Var}(h_1(X)) > 0$, *we have*

$$\sqrt{n}\left(b_{1,d}(X_1, \dots, X_n) - \beta_{1,d}\right) \to_{\mathcal{D}} \mathcal{N}(0, \sigma^2),$$

where

$$\sigma^2 = u'E[Z_1Z_1']u.$$

PROOF. Let

$$\hat{T}_n = \sum_{i=1}^{n} E[T_n|X_i] - (n-1)\beta_{1,d}$$

be the Hajek projection of T_n [Serfling (1980), page 188]. Since

$$\hat{T}_n - \beta_{1,d} = \frac{2}{n}\sum_{i=1}^{n}\left[h_1(X_i) - \beta_{1,d}\right]$$

and $E[(\hat{T}_n - T_n)^2] = O(n^{-2})$, we deduce from (3.3) that $V_n = \hat{T}_n + o_P(n^{-1/2})$. Invoking Lemma 3.1 we obtain

$$\sqrt{n}\,(b_{1,d}(X_1,\dots,X_n) - \beta_{1,d}) = u'\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i + o_P(1),$$

with u given in (3.6) and Z_i in (3.7). Since $Z_1,\dots,Z_n$ are i.i.d. with $E[Z_i] = 0$, the assertion of Theorem 3.2 follows from the multivariate central limit theorem and the continuous mapping theorem. □

REMARK. Observe that in view of affine invariance the assertion of Theorem 3.2 is more general than stated.

EXAMPLE 3.3 (The case $d = 1$). In the case $d = 1$ we have $h_1(x) = m_3 x^3$, where $m_k = E[X^k]$, $k \geq 1$. It follows that the matrix $E[Z_1 Z_1']$ takes the form

$$E[Z_1Z_1'] = \begin{bmatrix} \operatorname{Var} h_1 & \operatorname{Cov}(h_1, X^2 - 1) & \operatorname{Cov}(h_1, X) \\ \operatorname{Cov}(h_1, X^2 - 1) & \operatorname{Var}(X^2 - 1) & \operatorname{Cov}(X^2 - 1, X) \\ \operatorname{Cov}(h_1, X) & \operatorname{Cov}(X^2 - 1, X) & \operatorname{Var}(X) \end{bmatrix}$$

$$= \begin{bmatrix} m_3^2(m_6 - m_3^2) & m_3 m_5 - m_3^2 & m_3 m_4 \\ m_3 m_5 - m_3^2 & m_4 - 1 & m_3 \\ m_3 m_4 & m_3 & 1 \end{bmatrix}.$$

Since

$$u = (2, -3m_3^2, -6m_3)',$$

a simple calculation gives

$$u'E[Z_1Z_1']u = 4m_3^2\left[m_6 - 6m_4 + 11m_3^2 - 3m_3m_5 + \tfrac{9}{4}m_3^2(m_4 - 1) + 9\right].$$

Letting

$$\sqrt{b}_{1,1} = \frac{(1/n)\sum_{i=1}^{n}(X_i - \bar{X})^3}{\left((1/n)\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^{3/2}},$$

it follows that in the case $m_3 \neq 0$

$$\begin{aligned} &\sqrt{n}\left(\sqrt{b}_{1,1} - m_3\right) \\ &\quad = \sqrt{n}\,(b_{1,1} - \beta_{1,1})\frac{1}{\sqrt{b_{1,1}} + m_3} \\ &\quad \to_{\mathscr{D}} \mathscr{N}\left(0, m_6 - 6m_4 + 11m_3^2 - 3m_3m_5 + \tfrac{9}{4}m_3^2(m_4 - 1) + 9\right), \end{aligned}$$

which is the result of Gastwirth and Owens (1977).

4. Conclusions.

REMARK 4.1. We have obtained the limit behavior of Mardia's (sample) measure $b_{1,d}$ of multivariate skewness. Under the weak condition that the support of the underlying distribution has positive d-dimensional Lebesgue measure, two possible types of limit distributions occur according as $\beta_{1,d} = 0$ or $\beta_{1,d} > 0$.

If $\beta_{1,d} = 0$, Lemma 2.1 holds true even if the distribution of X is not elliptical. In fact, $\beta_{1,d} = 0$ implies that $E[|X|^2 X] = 0$, so the proof of Lemma 2.1 carries over. Consequently, the limit law of $nb_{1,d}$ is a weighted sum of (possibly more than two) independent χ^2-variates. However, it seems to be difficult to obtain closed-form expressions for the weights in case of nonelliptical distributions satisfying $\beta_{1,d} = 0$.

REMARK 4.2. Obviously the test for multivariate normality rejecting the hypothesis H_0 for large values of $nb_{1,d}$ is consistent exactly against those alternative distributions satisfying $\beta_{1,d} > 0$. Theorem 2.2 clarifies some curiosities in the power performance of the test for multivariate normality based on multivariate skewness for the class of elliptical distributions. In Table 1 we present the result of a Monte Carlo experiment regarding the power of the test for multivariate normality based on $b_{1,d}$. For the case $d = 5$, 1000 pseudorandom samples of size $n = 20$ were taken from several elliptical distributions considered in Section 2. The entries represent the number of significant samples at the level $\alpha = 0.05$. These results, which at first sight might be striking, are to be expected in view of Theorem 2.2. In the case of the symmetric multivariate Pearson Type II distribution $\mathrm{MPII}_d(\kappa, \mu, \Delta)$, the weights α_1 and α_2 figuring in the limiting distribution of $nb_{1,d}$ increase with κ, but are both always less than (the "normal" value) 6 (Example 2.4). Consequently, the limit law of $nb_{1,d}$ for the Pearson Type II family is

TABLE 1
Number of 1000 *Monte Carlo samples declared significant by the test for multinormality based on* $b_{1,d}$ ($\alpha = 0.05$, $n = 20$, $d = 5$)

Distribution	Number of significant samples
$\mathrm{MPII}_5(0, \mu, \Delta)$	0
$\mathrm{MPII}_5(1, \mu, \Delta)$	0
$\mathrm{MPII}_5(4, \mu, \Delta)$	12
$\mathrm{MK}_5(-0.25, r, 1, \mu, \Delta)$	222
$\mathrm{MK}_5(0, r, 1, \mu, \Delta)$	133
$\mathrm{MK}_5(1, r, 1, \mu, \Delta)$	47
$\mathrm{MK}_5(2, r, 1, \mu, \Delta)$	13
$\mathrm{MPVII}_5(\kappa, 3, \mu, \Delta)$	997
$\mathrm{MPVII}_5(\kappa, 5, \mu, \Delta)$	557
$\mathrm{MPVII}_5(\kappa, 10, \mu, \Delta)$	166

stochastically bounded from above by $6\chi^2_{d(d+1)(d+2)/6}$. This explains the very poor power behavior in this case.

For the symmetric Kotz type distribution $\mathrm{MK}_d(a, r, 1, \mu, \Delta)$, α_1 and α_2 are decreasing functions of a (Example 2.5), which implies that the limit distribution of $nb_{1,d}$ is stochastically decreasing with a. For $a > 1$ $(a < 1)$ the asymptotic power of Mardia's test for multivariate normality is less than (greater than) the nominal level.

For the symmetric multivariate Pearson Type VII distribution $\mathrm{MPVII}_d(\kappa, a, \mu, \Delta)$, both α_i are decreasing functions of a (Example 2.6) and approach the ("normal") value 6 as $a \to \infty$. For small values of a the test for multinormality based on $b_{1,d}$ will have very high power due to the fact that m_4 and m_6 are infinite (see Example 2.6).

REFERENCES

EATON, M. L. and PERLMAN, M. D. (1973). The nonsingularity of generalized sample covariance matrices. *Ann. Statist.* **1** 710–717.

ERDÉLYI, A., MAGNUS, W., OBERTHETTINGER, F. and TRICOMI, F. G. (1953). *Higher Transcendental Functions* **2**. McGraw-Hill, New York.

FANG, K. T., KOTZ, S. and NG, K. W. (1989). Symmetric Multivariate and Related Distributions. Chapman and Hall, London.

GASTWIRTH, J. L. and OWENS, M. E. B. (1977). On classical tests of normality. *Biometrika* **64** 135–139.

GREGORY, G. G. (1977). Large sample theory for *U*-statistics and tests of fit. *Ann. Statist.* **5** 110–123.

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 in testing normality and robustness studies. *Sankhyā Ser. B* **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. and ZEMROCH, P. J. (1975). Measures of multivariate skewness and kurtosis. *J. Roy. Statist. Soc. Ser. C* **24** 262–265.

OKAMOTO, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *Ann. Statist.* **1** 763–765.

SCHWAGER, S. J. (1985). Multivariate skewness and kurtosis. *Encyclopedia of Statistical Sciences* **6** 122–125. Wiley, New York.

SERFLING, R. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

STEIN, E. and WEISS, G. (1971). *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton Univ. Press.

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
UNIVERSITÄT HANNOVER
WELFENGARTEN 1, D-3000 HANNOVER 1
GERMANY

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
UNIVERSITÄT KARLSRUHE
ENGLERSTR. Z, D-7500 KARLSRUHE
GERMANY