Multivariate Analysis and Its Applications
IMS Lecture Notes – Monograph Series (1994) Volume 24

LOCALLY BEST INVARIANT TESTS FOR MULTIVARIATE NORMALITY IN CURVED FAMILIES WITH μ KNOWN

By Takeaki Kariya and Edward I. George

Hitotsubashi University and University of Texas at Austin

This paper is a continuation of Kariya and George (1992) and derives the LBI tests and their asymptotic null and nonnull distributions in such curved families as an arithmetic normal mixture, a geometric normal mixture, an exponential-d family, when location parameter μ is known.

1. Introduction and Summary. Generalizing the arguments in Kuwana and Kariya (1991), Kariya and George (1992) formulated a testing problem in an elliptically contoured curved family, derived a general form of the LBI (locally best invariant) test and the null and nonnull distributions of the LBI test, and proposed a measure of the local departure of an elliptically contoured curved family from normality. In this paper, we treat the special case where location parameter $\mu(\in R^p)$ is known, because the LBI test is quite different from the one when μ is unknown and because the location invariance is not available in a multivariate linear model with iid errors as will be discussed below. In an arithmetic normal mixture, a geometric normal mixture, and an exponential-d family as subfamilies, the problem is discussed in details.

In our model, a deviation from the normal family with mean μ known

$$\mathrm{N}_\mu = \{N_p(\mu, \Sigma) : \Sigma \in \mathbb{S}(p)\}. \tag{1.1}$$

is described by a real parameter θ where a specific value, say $\theta = 0$, corresponds to the normal family N_μ in (1.1). where $\mathbb{S}(p)$ denotes the set of $p \times p$ positive definite matrices. Since $N_p(\mu, \Sigma)$ is a location and scale family of (μ, Σ), it is natural to consider the location and scale family with pdf's of the form

$$p_\theta(x \mid \mu, \Sigma) = |\Sigma|^{-1/2} f_\theta((x-\mu)'\Sigma^{-1}(x-\mu)), \tag{1.2}$$

* Kariya's portion is in part supported by the Ministry of Education, Grant General C 04630014.

AMS 1980 Subject Classifications: Primary 62H15.

Key words and phrases: Curved model, locally best invariant test for multivariate normality, Mardia's test.

where $f_\theta(\cdot)$ is known for each parameter $\theta \in I$. We take p_θ to be a normal pdf when $\theta = 0$ so that $f_0(z) = (2\pi)^{-p/2}\exp[-z/2]$. Hence the family described by the pdf's in (1.2) may be regarded as a "curve" passing through the normal family (1.1). We also assume that $\frac{\partial}{\partial\theta} f_\theta(z)$ is continuous in $(z, \theta) \in [0, \infty) \times I$. Note that every pdf in (1.2) is elliptically symmetric. For general properties of such distributions, we refer the reader to Kelker (1970), Kariya and Sinha (1988) and Fang and Zhang (1990). For each curved family (1.2), we consider the problem of testing

$$H : \theta = 0 \quad \text{vs} \quad K : \theta > 0 \tag{1.3}$$

based on an iid sample $(x_1, \cdots, x_n)$, and derive the LBI test and its asymptotic null and nonnull distributions.

Then $f_\theta(\cdot)$ is specialized and the following three classes are considered in details:

(1) Arithmetic mixtures: $p_\theta(\cdot) = (1-\theta)p_0(\cdot) + \theta q(\cdot)$

(2) Geometric mixtures: $p_\theta(\cdot) = c(\theta)p_0(\cdot)^{1-\theta}q(\cdot)^\theta$

(3) Exponential-d family: $p_\theta(\cdot) = c(\theta)\exp(-\frac{1}{2}d_\theta(\cdot))$.

where $q(x|\mu, \Sigma) = |\Sigma|^{-1/2}g((x-\mu)'\Sigma^{-1}(x-\mu))$ with known g, and $d_\theta(z)$ is a linear function of powers of z such that $d_0(z) = z$. In the classes (1) and (2), the LBI tests are shown to measure the difference between the functional forms of f_0 and g. Normal-t mixtures and normal exponential power mixtures are considered as special cases. An important special case of (3) obtained by $d_\theta(z) = z^{\theta+1}$ is the family of exponential power distributions treated by Kuwana and Kariya (1991). In Kariya and George (1992), it is shown that when μ is unknown, the exponential-d family with $\frac{\partial}{\partial\theta} d_\theta(z)|_{\theta=0} = -kz^2$ yields Mardia's kurtosis test (1970) as the LBI test. We also show this for the case where μ is known.

Since μ is assume to be known, without loss of generality we assume $\mu = 0$ in the sequel (by replacing x by $x - \mu$). It should be noted that for a given curved family (1.2), replacing μ by $\bar{x}$ in the LBI test when μ is known does not yield the LBI test when μ is unknown. Indeed, the LBI test under unknown μ is rather complicated (see Kariya and George (1992)). To state another reason why we consider the case where μ is known, let a multivariate linear model be

$$U = Z\gamma + E \tag{1.4}$$

where $U = (u_1, \cdots, u_n)' : n \times p, Z = (z_1, \cdots, z_n)' : n \times k$ is nonrandom, $\gamma : k \times p$ is unknown and $E = (e_1, \cdots, e_n)' : n \times p$ is an error matrix with e_i's iid and $E(e_i) = 0$. It is often assumed that e_i's are iid $N(0, \Sigma)$. To check this assumption in our setting, we need to derive the LBI test under (1.2). However, even though $\gamma' z_i$ is a location parameter in each $u_i = \gamma' z_i + e_i$, there is no group-invariance structure which gets rid of the nuisance parameter γ under the iid assumption for e_i's with (1.2) except for the normal case $\theta = 0$. It is noted that $(\widehat{\gamma}, \widehat{E}'\widehat{E})$ is <u>not</u> a sufficient statistic for (1.2) except for $\theta = 0$, where $\widehat{\gamma} = (Z'Z)^{-1}Z'U$ and $\widehat{E} = U - Z\widehat{\gamma}$. Consequently no LBI test is available for the model (1.4). In such a situation, the procedure we propose here is that assuming $\gamma' z_i$ is known, we derive the LBI test based on $e_i = u_i - \gamma' z_i$ and then substitute $\widehat{e}_i = u_i - \widehat{\gamma}' z_i$ for e_i. This problem is briefly discussed in the last section.

2. The LBI Test for Normality. In this section we derive the general form of the LBI test of $H : \theta = 0$ vs $K : \theta > 0$ under the family p_θ in (1.2) with $\mu = 0$. Let $X = (x_1, \cdots, x_n)' : n \times p$ and let $Gl(p)$ denote the group of $p \times p$ nonsingular matrices acting on X and (θ, Σ) by

$$g \cdot X = XA' \quad \text{and} \quad g \cdot (\theta, \Sigma) = (\theta, A\Sigma A') \tag{2.1}$$

respectively where $g = A \in Gl(p)$. Clearly the problem of testing (1.4) under (1.2) when $\mu = 0$ is left invariant under (2.1) and a maximal invariant is

$$W(X) \equiv \frac{1}{n} X S^{-1} X' \equiv YY' \quad \text{with}$$
$$Y \equiv (y_1, \cdots, y_n)' = \frac{1}{\sqrt{n}} X S^{-1/2} \quad \text{and} \quad S = \frac{1}{n} X'X. \tag{2.2}$$

Also a maximal invariant parameter is θ. This implies that the distribution of W, denoted by P_θ^W, depends on θ only so that we can assume $\Sigma = I$ in our invariant analysis. Furthermore, any $Gl(p)$-invariant test ϕ of $H : \theta = 0$ versus $K : \theta > 0$ will be a function only of W. The following result provides a characterization of the locally best invariant (LBI) test.

THEOREM 2.1. *Let $\pi(\phi, \theta)$ be the power function of a $Gl(p)$-invariant test ϕ of $H : \theta = 0$ vs $K : \theta > 0$ under p_θ in (1.2) when $\mu = 0$. Let $h(YA'|\theta) = \prod_{j=1}^n p_\theta(Ay_j|0, I)$. Suppose for all θ in a neighborhood of $\theta = 0$,*

$$\frac{\partial}{\partial\theta} \int_{Gl(p)} h(YA'|\,\theta)|\det A|^{n-p} dA$$
$$= \int_{Gl(p)} \frac{\partial}{\partial\theta} h(YA'|\,\theta)|\det A|^{n-p} dA. \tag{2.3}$$

(i)*The slope of $\pi(\phi, \theta)$ at $\theta = 0$ is given by*

$$\frac{\partial}{\partial\theta}\pi(\phi, \theta)|_{\theta=0} = nE_0[\phi(W)K(W)] \tag{2.4}$$

where

$$K(W) = \frac{1}{n}\sum_{j=1}^n E_b \Psi(b\|y_j\|^2) \tag{2.5}$$

with

$$\Psi(z) = \frac{\partial}{\partial \theta} \log f_\theta(z)|_{\theta=0} = f'_0(z)/f_0(z) \propto f'_0(z) \exp[z/2] \qquad (2.6)$$

and $b \sim \chi^2_n, \chi^2$*-distribution with d.f. (degrees of freedom)* n.

(ii) $E_0 K(W) = 0$.

(iii) *The test which rejects* H *for large values of* $K(W)$ *is LBI for* H *vs* K.

PROOF. The proof is similar to the one given in Kuwana and Kariya (1991), and hence it is outlined. By Wijsman's Theorem (1967) or Andersson (1978), the power function of ϕ is given by

$$\pi(\theta, \phi) = \int \phi(W) q(W|\theta) dP_0^w, \qquad (2.7)$$

where $q(W|\theta) = H(Y|\theta)/H(Y|0)$, with

$$H(Y \mid \theta) = \int_{Gl(p)} h(YA'|\theta)| \det A|^n \nu(dA) \qquad (2.8)$$

and $\nu(dA) = |\det A|^{-p} dA$. Here using the continuous differentiability of $p_\theta(\cdot|0, I)$ and the compactness of the space $I\!D$ of $Y, H(Y|\theta)$ is bounded on $I\!D \times [0, \varepsilon]$ for some $\varepsilon > 0$. Thus the differentiation of π at $\theta = 0$ can be performed beneath the integral sign and making use of (2.3), the derivative of q at $\theta = 0$ leads to the derivative inside the integral sign;

$$\begin{aligned} \frac{\partial}{\partial \theta} h(YA'|\theta)|_{\theta=0} &= \sum_{j=1}^{n} \frac{p'_0(Ay_j)}{p_0(Ay_j)} h(YA'|0) \\ &= \sum_{j=1}^{n} \Psi(y'_j A' A y_j) h(YA'|0). \end{aligned} \qquad (2.9)$$

Since $h(YA'|0) = (2\pi)^{-np/2} \exp[-\frac{1}{2} \mathrm{tr} A'A]$, we obtain

$$\frac{\partial}{\partial \theta} q(W|\theta)|_{\theta=0} = \sum_{j=1}^{n} E_A[\Psi(y'_j A' A y_j)] \qquad (2.10)$$

where $E_A(\,\cdot\,)$ is the expectation of $\cdot$ under the pdf $f(A) \propto \exp[-\frac{1}{2} \mathrm{tr}\, A'A]|\det A|^{n-p}$ where $\int_{Gl(p)} f(A) dA = 1$. Arguing in the same way as in Kuwana and Kariya (1992), we obtain

$$E_A[\Psi(y'_j A' A y_j)] = E_b[\Psi(b \|y_j\|^2)] \qquad (2.11)$$

where $b \sim \chi^2_n$. Substituting (2.11) into (2.10) yields (2.4).

The special case $\phi \equiv \alpha$ inserted into (2.4) shows $E_0[K(W)] = 0$. Finally, maximizing $E_0[\phi(W)K(W)]$ since the alternative is $\theta > 0$, the LBI result follows from the generalized Neyman-Pearson Lemma.

The function $\Psi(z) = \frac{\partial}{\partial\theta} \log f_\theta(z)|_{\theta=0}$ in (2.6) is the efficient score at $\theta = 0$ for testing $H : \theta = 0$ vs $K : \theta > 0$ for the one dimensional family $z \sim f_\theta(z)$; see Cox and Hinkley (1974). Thus $K(W)$ in (2.5) may be interpreted as an averaged efficient score statistic in the direction of the alternative. Note that as long as $f'_0(\cdot)$ is available, $K(W)$ can be computed by Monte Carlo approximation under $b \sim \chi^2_n$. The following alternative expression for $K(W)$ is obtained by transforming b into $c = b(1 - \|y_j\|^2)$.

COROLLARY 2.1. *The LBI test statistic $K(W)$ in (2.5) may be expressed as*

$$K(W) \equiv (2\pi)^{p/2}\frac{1}{n}\sum_{j=1}^{n}(1 - \|y_j\|^2)^{-n/2}E_b\left[f_0\left(\frac{b\|y_j\|^2}{1-\|y_j\|^2}\right)\right] \tag{2.12}$$

with $b \sim \chi^2_n$.

Next we obtain asymptotic distributions under p_θ in (1.2) of statistics of the form

$$T(W) = \frac{1}{2}\sum_{j=1}^{n} E_b H(b\|y_j\|^2) \tag{2.13}$$

with $b \sim \chi^2_n$, and $\|y_j\|^2$ as in (2.2). These results will then be applied to the test statistic $K(W)$ in (2.5) which is a special case of T when $H = \Psi$.

THEOREM 2.2. *For $T(w)$ in (2.13), suppose $H : R \to R$ satisfies*

(a) *$H(z)$ is continuously twice differentiable,*

(b) *$E_0[H(\|x_1\|^2)]^2, E_\theta[\|x_1\|^2 H(\|x_1\|^2)], E_\theta[\|x_1\|^2 H'(\|x_1\|^2)] < \infty$ and*

(c) *there exists a function of the form $G(u) = \sum_{k=-N_1}^{N_2} \alpha_k u^k$ for some N_1, N_2 such that $|H''(u)| \le G(u)$ and $E_\theta[\|x_1\|^4 G(\|x_1\|^2)] < \infty$.*

Then as $n \to \infty$,

$$\sqrt{n}[T(W) - \mu(\theta)] \to N(0, v(\theta)) \tag{2.14}$$

where

$$\mu(\theta) = E_\theta[H(\|x_1\|^2)] \quad \text{and}$$
$$v(\theta) = \mathrm{Var}_\theta\left\{H(\|x_1\|^2) - \frac{1}{p}\|x_1\|^2 E_\theta[\|x_1\|^2 H'(\|x_1\|^2)]\right\}. \tag{2.14a}$$

PROOF. The proof is completely analogous to Kuwana and Kariya (1991), and so omitted here.

The asymptotic null distribution of T is directly obtained from Theorem 2.2 with $\theta = 0$, while the asymptotic nonnull distribution under contiguous alternatives is given by

COROLLARY 2.2. *Suppose $T(W)$ satisfies (a)–(c) of Theorem 2.2. Then the asymptotic nonnull distribution of $\sqrt{n}[T(W) - \mu(0)]$ under the contiguous alternatives $\theta_n = \omega/\sqrt{n}$ with $\omega > 0$ is $N(\omega\mu'(0), v(0))$ where $\mu'(0) = \frac{\partial}{\partial\theta}\mu(\theta)|_{\theta=0}$.*

PROOF. Write $\sqrt{n}[T(W)-\mu(0)] = \sqrt{n}[T(W)-\mu(\omega/\sqrt{n})]+\sqrt{n}[\mu(\omega/\sqrt{n}) -\mu(0)]$. The result follows from $\lim_{n\to\infty}\sqrt{n}[\mu(\omega/\sqrt{n}) - \mu(0)] = \omega\mu'(0)$ and $\lim_{n\to\infty} v(\omega/\sqrt{n}) = v(0)$.

3. Specific LBI Tests.

(I) The LBI Test Against Arithmetic Mixtures

The following result yields the LBI test when $\mu = 0$, against the family of arithmetic mixture alternatives;

$$p_\theta(x|0,\Sigma) = (1-\theta)|\Sigma|^{-1/2} f_0(x'\Sigma^{-1}x) + \theta|\Sigma|^{-1/2} g(x'\Sigma^{-1}x) \tag{3.1}$$

where $g(x'x)$ is a fixed pdf on R^p.

THEOREM 3.1. *Suppose*

$$\int_{Gl(p)} \prod_{j=1}^{n} g(y'_j A'Ay_j)|\det A|^{n-p} dA < \infty. \tag{3.2}$$

Then the LBI test of $H:\theta = 0$ vs $K:\theta > 0$ in (3.1) rejects H for large values of

$$\begin{aligned} T(W) =& \frac{1}{n}\sum_{j=1}^{n} E_b\left[\frac{g(b\|y_j\|^2)}{f_0(b\|y_j\|^2)}\right] \\ =&(2\pi)^{p/2}\frac{1}{n}\sum_{j=1}^{n}(1+\eta_j)^{n/2} E_b[g(b\eta_j)] \end{aligned} \tag{3.3}$$

where $b \sim \chi^2_n$ and $\eta_j = \|y_j\|^2/(1-\|y_j\|^2)$.

PROOF. It is easy to see that (2.3) of Theorem 2.1 will hold for (3.1) as long as (3.2) is satisfied. The two expressions are immediately obtained from Theorem 2.1 and Corollary 2.1 using $K(W) = 1 + T(W)$.

The first expression of $T(W)$ in (3.3) shows that the LBI test compares g and f_0 by a weighted average of likelihood ratios. The second expression for

$T(W)$ can be further specified when $g(\cdot)$ is analytic on $[0,\infty)$ and so expanded as

$$g(z)=\sum_{k=0}^{\infty}\alpha_k z^k/k! \quad \text{with} \quad \alpha_k=\partial^k g(z)/\partial z^k|_{z=0} \tag{3.4}$$

COROLLARY 3.1. *When $g(z)$ satisfying (3.2) is of the form (3.4), the LBI test for H vs K in (3.1) is that which rejects H for large values of*

$$T(W)=\frac{1}{n}\sum_{j=1}^{n}(1+\eta_j)^{n/2}\Big[\sum_{k=0}^{\infty}r_k(\eta_j)\beta_k\Big] \tag{3.5}$$

where $\eta_j=\|y_j\|^2/(1-\|y_j\|^2)$, $r_k(\eta_j)=\frac{\Gamma(\frac{n}{2}+k)}{k!\Gamma(\frac{n}{2})}\eta_j^k$ *and* $\beta_k=(2\pi)^{p/2}2^k\alpha_k$.

PROOF. The proof is straightforward.

The LBI test statistic in (3.5) has the following interesting interpretation.

(1) When $g(z)$ is analytic on $[0,\infty)$, the entire form of $g(z)$ is determined by the derivative coefficients α_k's at $z=0$. In th case of a normal pdf, $\alpha_{kN}=(2\pi)^{-p/2}(-2)^{-k}$ and hence β_k in (3.5) measures the difference between the fixed pdf g for which $\beta_k=(-1)^k\alpha_k/\alpha_{kN}$ and the normal pdf f_0 for which $\beta_k=(-1)^k$.

(2) The value $w_{jj}=\|y_j\|^2=\frac{1}{n}x'_jS^{-1}x_j(\leq 1)$ is regarded as the distance of the j-th observation x_j from the origin relative to the sample covariance matrix S. If θ is close to 1 in (3.1) and if $g(z)$ has a heavier (or thinner) tail relative to the normal case, them the $\|y_j\|^2$'s and hence the η_j's tend to be larger (or smaller) as a whole. Thus the $\|y_j\|^2$'s and η_j's reflect the mixture parameter θ of f_θ in (3.1) and the form of g. And if θ is close to 1, the $\|y_j\|^2$'s and η_j's will reflect more of the form of g, whereas if θ is close to 0, they will tend to exhibit more of the features of the normal case f_0.

(3) Hence the deviation of g from normality is detected by the weights $r_k(\eta_j)$ on β_k in (3.5). When $\|y_j\|^2$ and hence η_j is large, $r_k(\eta_j)$ will put more weight on the β_k for large k. Thus, when some $\|y_j\|^2$ is large, the ratios $\beta_k=\alpha_k/\alpha_{kN}$ of higher derivatives are more likely to be tested against the deviation form normality when n is fixed. Conversely, when all the $\|y_j\|^2$'s are small, the lower derivatives get more weights for testing normality because the $\|y_j\|^2$'s do not contradict with the null hypothesis in higher order derivatives. Also, as n becomes larger, the increased weight due to a large $\|y_j\|^2$ becomes more pronounced through the term $(1+\eta_j)^{n/2}$. Hence as n gets large, the higher derivatives get more weight for testing normality.

Finally, the asymptotic null and nonnull distributions of $T(W)$ in (3.3) with $H(z)=g(z)/f_0(z)$ under assumptions (a)-(c) are

$$\begin{aligned}&\sqrt{n}[T(W)-1]\to N(0,v(0)) \quad \text{for} \quad \theta=0\\ &\sqrt{n}[T(W)-1]\to N(\omega\mu'(0),v(0)) \quad \text{for} \quad \theta_n=\omega/\sqrt{n}\end{aligned} \tag{3.6}$$

where

$$v(\theta) = \mathrm{Var}_\theta \left[\frac{g(||x_1||^2)}{f_0(||x_1||^2)} - \frac{1}{p}||x_1||^2 \delta(\theta) \right]$$

and

$$\delta(\theta) = E_\theta \left[||x_1||^2 \frac{g'(||x_1||^2) f_0(||x_1||^2) - g(||x_1||^2) f'_0(||x_1||^2)}{f_0(||x_1||^2)^2} \right]$$

EXAMPLE 3.1. Normal-t arithmetic mixture. Suppose the mixing distribution for g in (3.1) is the multivariate t distribution with m degrees of freedom

$$|\Sigma|^{-1/2} \frac{\Gamma(m+p)/2}{(\pi m)^{p/2}\Gamma(m/2)} \left[1 + \frac{1}{m} x'\Sigma^{-1}x \right]^{-(m+p)/2} \tag{3.7}$$

To insure that condition (3.2) of Theorem 3.1 is satisfied we shall require $m \geq n-p$. From (3.3), the LBI test is that which rejects H for large values of

$$T(W) = \frac{1}{n}\sum_{j=1}^{n} E_b H(b||y_j||^2) \quad \text{with}$$

$$H(z) = \frac{\Gamma(m+p)/2}{(m/2)^{p/2}\Gamma(m/2)} \left[1 + \frac{1}{m} z \right]^{-(m+p)/2} \exp[z/2] \tag{3.8}$$

where $b \sim \chi^2_n$. this expression for $T(W)$ can be evaluated by Monte Carlo approximation.

Unfortunately here $E_\theta[H(||x_j||^2)]^2 = \infty$ for all θ, so that condition (b) of Theorem 2.2 is not satisfied. Thus, these results cannot be used to obtain asymptotic distributions for $T(W)$.

EXAMPLE 3.2. Normal-exponential power arithmetic mixture. Suppose the mixing distribution for g in (3.1) is a member of the exponential power family

$$|\Sigma|^{-1/2} \frac{\alpha\Gamma(p/2)}{(2^{1/\alpha}\pi)^{p/2}\Gamma(p/2\alpha)} \exp\left\{ -\frac{1}{2}[(x'\Sigma^{-1}x)]^{\alpha} \right\}. \tag{3.9}$$

Note that condition (3.2) of Theorem 3.1 is satisfied by (3.9) as long as $\alpha > 0$. From (3.3), the LBI test is that which rejects H for large values of

$$T(W) = \frac{1}{n}\sum_{j=1}^{n} E_b H(b||y_j||^2),$$

$$H(z) = \frac{2^{p(\alpha-1)/2\alpha}\alpha\Gamma(p/2)}{\Gamma(p/2\alpha)} \exp\{(z - z^{\alpha})/2\} \tag{3.10}$$

where $b \sim \chi^2_n$. This expression for $T(W)$ can be evaluated by Monte Carlo approximation.

Finally, when $\alpha > 1$, the asymptotic null and nonnull distributions of $T(W)$ here may be obtained from Theorem 2.2 and Corollary 2.2 (or (3.6)) since in this case assumptions (a)-(c) are satisfied. These assumptions are not satisfied when $\alpha < 1$.

(2) The LBI Test Against Geometric Mixtures

The following result yields the LBI test when $\mu = 0$, against the family of geometric mixture alternatives

$$p_\theta(x|0, \Sigma) = |\Sigma|^{-1/2} c(\theta)[f_0(x'\Sigma^{-1}x)]^{1-\theta}[g(x'\Sigma^{-1}x)]^\theta \tag{3.11}$$

where $g(x'x)$ is a fixed pdf on R^p.

THEOREM 3.2. *Suppose*

$$\int_{Gl(p)} \prod_{j=1}^{n} g(y'_j A' A y_j)| \det A|^{n-p} dA < \infty. \tag{3.12}$$

Then the LBI test of $H : \theta = 0$ vs $K : \theta > 0$ in (3.11) rejects H for large values of

$$\begin{aligned} T(W) &= \frac{1}{n} \sum_{j=1}^{n} E_b \log \left[\frac{g(b\|y_j\|^2)}{f_0(b\|y_j\|^2)} \right] \\ &= \frac{1}{n} \sum_{j=1}^{n} E_b \left[\log g(b\|y_j\|^2) \right] + C \end{aligned} \tag{3.13}$$

where $b \sim \chi^2_n$ and C is a constant.

PROOF. It is straightforward to check that (2.3) of Theorem 2.1 will hold for (3.11) as long as (3.12) is satisfied. Calculation of Ψ in (2.6) yields $\Psi(z) = \log[g(z)/f_0(z)] + C'$ for some constant C', as $\Sigma_{j=1}^n E_b \log[f_0(b\|y_j\|^2)] = n \log(2\pi)^{p/2} - \frac{np}{2}$.

As opposed to $T(W)$ in (3.3), the first expression in (3.13) shows that here the LBI test compares g and f_0 by a weighted average of log likelihood ratios. Asymptotic null and nonnull distributions of $T(W)$ in (3.13) can be obtained from Theorem 2.2 and Corollary 2.2 with $H(z) = \log[g(z)/f_0(z)]$ when assumptions (a)-(c) are satisfied.

EXAMPLE 3.3. Normal-t geometric mixture. Suppose the mixing distribution in (3.11) is the multivariate t distribution with m degrees of freedom in (3.7). From Theorem 3.2, the LBI test rejects H for small values of

$$T(W) = \frac{m+p}{n} \sum_{j=1}^{n} E_b \log \left[1 + \frac{b}{m} \|y_j\|^2 \right] \tag{3.14}$$

where $b \sim \chi^2_n$. The asymptotic distributions of $T(W)$ can be obtained from Theorem 2.2 and Corollary 2.2. Assumptions (a)–(c) are easily verified using $H(z) = (m+p)\log(1+z/m), H'(z) = (m+p)/(m+z)$, and $H''(z) = -(m+p)/(m+z)^2$. Note that $\mu(0)$ and $v(0)$ may be computed using (2.14a).

EXAMPLE 3.4. Normal-exponential power geometric mixture. Suppose the mixing distribution in (3.11) is a member of the exponential power family in (3.9). From Theorem 3.2, the LBI test rejects H for small values of

$$T(W) = \left[2^\alpha \Gamma\left(\frac{n}{2}+\alpha\right)/\Gamma\left(\frac{n}{2}\right)\right] \frac{1}{n}\sum_{j=1}^{n} \|y_j\|^{2\alpha}. \tag{3.15}$$

The asymptotic distributions of $T(W)$ can be obtained from Theorem 2.2 and Corollary 2.2. Assumptions (a)-(c) are easily verified using $H(z) = z^\alpha, H'(z) = \alpha z^{\alpha-1}$, and $H''(z) = \alpha(\alpha-1)z^{\alpha-2}$. For example, under $H: \theta = 0, \sqrt{n}\,[T(W) - \mu(0)] \to N(0, v(0))$ where $\mu(0) = [2^\alpha\Gamma(\frac{p}{2}+\alpha)/\Gamma(\frac{p}{2})]$ and $v(0) = \text{Var}_0[\|x_1\|^{2\alpha} - \frac{\alpha}{p}\|x_1\|^2\mu(0)]$.

(3) The Exponential-d Family and Mardia's Test

When $\mu = 0$, the exponential-d family is given by

$$p_\theta(x|0,\Sigma) = |\Sigma|^{-1/2} c(\theta) \exp\left(-\frac{1}{2} d_\theta(x'\Sigma^{-1}x)\right), \tag{3.16}$$

where $d_\theta(z)$ is a liner function of powers of z such that $d_0(z) = z$. Note that when this family satisfies (2.3) of Theorem 2.1, the LBI test against (3.16) rejects H for large values of

$$T(W) = \sum_{j=1}^{n} E_b\left[-\frac{\partial}{\partial\theta} d_\theta(b\|y_j\|^2)\Big|_{\theta=0}\right]. \tag{3.17}$$

The asymptotic null and nonnull distributions of (3.17) may then be obtained from Theorem 2.2 and Corollary 2.2 when assumption (a)-(c) are satisfied.

We now show that Mardia's kurtosis test is LBI (when $\mu = 0$) against the subfamily of (3.16) obtained by restricting to $d_\theta(z)$ satisfying

$$\frac{\partial}{\partial\theta} d_\theta(z)|_{\theta=0} = -kz^2 \tag{3.18}$$

for some constant $k > 0$. For example, $d_\theta(z) = z(1-\theta z)^2$ is such a function.

THEOREM 3.3. *Consider an exponential-d family (3.16) satisfying (3.18) and (2.3) of Theorem 2.1. The LBI test for $H: \theta = 0$ vs $K: \theta > 0$ rejects H for large values of*

$$T(W) = \frac{1}{n}\sum_{j=1}^{n} [x'_j S^{-1} x_j]^2 \quad \text{with} \quad S = \frac{1}{n}\sum_{j=1}^{n} x_j x'_j \tag{3.19}$$

PROOF. The proof is straightforward.

4. Approximate LBI Test in a Linear Model. As is discussed in Section 1, when the errors e_i's in a multivariate linear model (1.4) are iid with pdf (1.2), no group-invariance structure is available to delete the unknown parameter γ. In this section, we propose the following procedure:

(i) Regard $\widehat{X} \equiv \widehat{E} = U - Z\widehat{\gamma}$ with $\widehat{\gamma} = (Z'Z)^{-1}Z'U$ as X and $\widehat{S} \equiv \frac{1}{n}\widehat{E}'\widehat{E}$ as S in Section 2.

(ii) Then substituting $\widehat{X}$ and $\widehat{S}$ for X and S in one of the LBI tests in Section 3, we obtain an approximate LBI test.

In particular, Mardia type test in this case is given by

$$\widehat{T} = \frac{1}{n}\sum_{j=1}^{n}[e'_j\widehat{S}^{-1}e_j]^2 \quad \text{with} \quad e_j = u_j - \widehat{\gamma}'z_j,$$

which is regarded as an approximate LBI test when e_i's follow the exponential-d family (3.16) with (3.18). While in the case of exponential power distribution with $d_\theta(z) = z^{1+\theta}$ in (3.16), the approximate LBI test is

$$\widehat{T} = \frac{1}{n}\sum_{j=1}^{n} e'_j\widehat{S}^{-1}e_j \log e'_j\widehat{S}^{-1}e_j$$

as in Kuwana and Kariya (1991). It is remarked that the LBI test when μ is unknown is very complicated in this distribution and that when $\theta = -\frac{1}{2}$, the distribution becomes multivariate double exponential distribution with thicker tails.

REFERENCES

ANDERSSON, S. A. (1978). Invariant measures. Technical Report No.129, Stanford University, Department of Statistics.

COX, D. R. and D. V. HINKLEY (1974). *Theoretical Statistics.* Chapman and Hall, London.

FANG K. T. and Y. T. ZHANG (1990). *Generalized Multivariate Analysis.* Science Press Beijing, Hong Kong.

HENZE, N. (1990). The asymptotic distribution of Mardia's mesure of multivariate kurtosis. Technical Report, Institute fur Mathematische Stochastikastik, Hannover University, West Germany.

HENZE, N. and B. ZIRKLER (1990). A class of invariant consistent test for nultivarite normality. Technical Report, Institut fur Mathematische Stochastikastik, Hannover University, West Germany.

KARIYA T. and B. K. SINHA (1988). *The Robustness of Statistical Tests.* Academic Press.

KARIYA T. and E. GEORGE (1992). LBI tests for multivariate normality in curved families and Mardia's test. To appear from Sankhya.

KELKER, D. (1970). Distribution theory of spherical distributions and a locationscale parameter generalization. *Sankhya* A**43**, 419–430.

KUWANA, Y. and T. KARIYA (1991). LBI tests for multivariate normality in exponential power distributions, *The Journal of Multivariate Analysis* **39**, 117–134.

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

WIJSMAN, R. A. (1967). Cross-sections of orbits and their applications to densities of maximal invariants. *Fifth Berk. Symp. Math Statist. Prob* 1, University of California, Berkeley, 389–400.

THE INSTITUTE OF ECONOMIC RESEARCH
HITOTSUBASHI UNIVERSITY
KUNITACHI
TOKYO 186, JAPAN

DEPARTMENT OF MSIS
CBA 5.202
UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TX 78712-1175, USA