Metrika (1989) 36:149–159

Monotonicity of the Power Function and Unbiasedness of Some Likelihood Ratio Tests

Lanxiang Chen[1]

A natural generalization of the p-dimensional normal distribution is provided by elliptically contoured distributions. In the case of the normal distribution the likelihood ratio tests (LRT) of null-hypothesis of the form

i) $\Sigma = I$,

ii) $\Sigma = I$ and $\mu = 0$,

have well known properties. This paper contains an investigation of the question of how far these properties are conserved when this more general family of distributions is considered. It is shown that the unbiasedness of the tests and the monotonicity of rheir power functions can still be proved for a large subfamily of these distributions.

1 Introduction and Notation

A p-dimensional random vector X has an elliptically contoured distribution if X has a density of the form

$$f(x) = |\Sigma|^{-\frac{1}{2}} g((x-\mu)^\tau \Sigma^{-1}(x-\mu)), \quad x \in \mathbb{R}^p \tag{1.1}$$

where $g(\cdot)$ is a positive measurable function, Σ a positive definite $p \times p$ matrix and $\mu \in \mathbb{R}^p$. We will then write $X \sim EC_p(\mu, \Sigma, g)$.

[1] Lanxiang Chen, Department of Mathematics, Tongji University, Si Ping 1239, Shanghai, V.R. China.

0026–1335/89/3–4/149–159 $2.50

The theory of elliptically contoured distributions was developed by Schoenberg (1938), Kelker (1970), Cambanis, Huang and Simons (1981) and generalized to random matrices by Anderson and Fang (1982a, 1982b), Good and Jensen (1981), Jensen (1984), Kariya and Sinha (1985).

A random $N \times p$ matrix X has an elliptically contoured distribution if for an $N \times p$ matrix M, a positive definite $p \times p$ matrix Σ and a non-negative measurable function g the density of X is given by

$$f(x) = |\Sigma|^{-N/2} g(\operatorname{tr} \Sigma^{-1}(x - M)^\tau (x - M)), \quad x \in \mathbb{R}^{N \times p}. \tag{1.2}$$

We will use the notation $X \sim LEC_{N \times p}(M, \Sigma, g)$.

Anderson and Fang (1982a and 1982b) gave likelihood ratio tests (LRT) for various hypotheses about the forms of M and Σ. It turns out that the test statistics have the same forms as the corresponding ones for the normally distributed case.

Optimality properties of these tests were proved by Nagao (1967), Sugiura and Nagao (1968) and by Anderson and Das Gupta (1964) for the normally distributed case. In this paper these results are generalized to elliptically contoured distributions.

We shall need the following results in this paper (see Anderson and Fang 1982b).

a) If $X \sim LEC_{N \times p}(M, \Sigma, g)$ then the density of the random matrix $W = (x - M)^\tau (x - M)$ is given by

$$c \cdot |\Sigma|^{-\frac{N}{2}} |w|^{\frac{N-p-1}{2}} g(\operatorname{tr} \Sigma^{-1} w), \quad w \in \mathbb{R}^{p \times p} \tag{1.3}$$

We call these distributions generalized Wishart distributions and use the notation $W \sim GW_p(\Sigma, N, g)$.

b) If $X \sim LEC_{N \times p}(e_N \mu^\tau, \Sigma, g)$ with $N > p$ and $e_N = (1, 1, \ldots, 1)^\tau$ and if

(*) $g(\cdot)$ is monotone decreasing and differentiable

then

1.) for known μ the ML-estimator of Σ is given by $\hat{\Sigma} = \lambda_0 W$ where $W = (X - e_N \mu^\tau)^\tau (X - e_N \mu^\tau)$ and λ_0 an appropriate constant;

2.) if μ is unknown the ML-estimators of μ and Σ are given by

$$\hat{\mu} = \bar{X} \quad \text{and} \quad \hat{\Sigma} = \lambda_0 V$$

with

$$V = (X - e_N \bar{X}^\tau)^\tau (X - e_N \bar{X}^\tau) \quad \text{and} \quad \bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_{(i)} .$$

Where μ is known we shall assume in the following without loss of generality that $\mu = 0$.

Lemma 1: Let $X \sim LEC_{N \times p}(e_N \mu^\tau, \Sigma, g)$. Then $V \sim GW_p(\Sigma, N-1, g_1)$ with

$$g_1(t) = c \int_0^\infty r^{\frac{p}{2}-1} g(t+r)dr.$$

Proof: We use an orthogonal transformation $Z = QX$ where the last row of Q is $(\sqrt{N})^{-1} \cdot e_N^\tau$.

Then

$$Z = \begin{pmatrix} Z_1 \\ Z_{(N)}^\tau \end{pmatrix} \sim LEC_{N \times p}\left(\begin{pmatrix} 0 \\ \sqrt{N} \cdot \mu^\tau \end{pmatrix}, \Sigma, g \right)$$

where $Z_{(N)} = \sqrt{N} \cdot \bar{X}$ and the density of Z_1 is

$$\int_{\mathbb{R}^p} |\Sigma|^{-N/2} g(\text{tr}\, \Sigma^{-1} z_1^\tau z_1 + (z_{(N)} - \sqrt{N}\mu)^\tau \Sigma^{-1} (z_{(N)} - \sqrt{N}\mu)) dz_{(N)},$$

$$z_1 \in \mathbb{R}^{(N-1) \times p}.$$

$$Z_1 \sim LEC_{(N-1) \times p}(0, \Sigma, g_1) \quad \text{with} \quad g_1(t) = \int_{\mathbb{R}^p} g(t + y^\tau y) dy.$$

Since $\int_{u^\tau u = t} du = \frac{\pi^{p/2}}{\Gamma(p/2)} t^{\frac{p}{2}-1}$ (see Srivastava and Khatri 1979), and $V = Z_1^\tau Z_1$ it follows that $V \sim GW_p(\Sigma, N-1, g_1)$.

2 Monotonicity of the Modified LRT for the Hypothesis $\Sigma = I$

Let $X \sim LEC_{N\times p}(e_N\mu^\tau, \Sigma, g)$. In the following we investigate LRT for the hypothesis $\Sigma = I$ against $\Sigma \neq I$ for known and for unknown μ.

I. $\mu = 0$

The critical region K (of the LRT) for $\Sigma = I$ against $\Sigma \neq I$ is of the form

$$K = \{w : |w|^{\frac{N}{2}} g(\mathrm{tr}\, w) < c\} \tag{2.1}$$

where $w = x^\tau x$ and c is uniquely determined by the test level α (see Anderson and Fang 1982b). First we discuss the above test in the case $p = 1$. Let $X \sim EC_N(0, \sigma^2, g)$, $x \in \mathbb{R}^N$. The critical region of the LRT for $\sigma^2 = 1$ against $\sigma^2 \neq 1$ is

$$K = \{u : u^{\frac{N}{2}} g(u) < c\}$$

where $U = X^\tau X$ has density

$$f_\sigma(u) = \pi^{N/2}\Gamma^{-1}\left(\frac{N}{2}\right)\sigma^{-N}u^{N/2-1}g\left(\frac{u}{\sigma^2}\right).$$

Theorem 1: Let $X \sim EC_N(0, \sigma^2, g)$ and let $g(\cdot)$ have property (*). Let

a) $g(ct)/g(t)$ be monotone increasing for $0 < c < 1$.

b) $h'(t) = 0$ have a unique solution, where $h(t) = t^{N/2}g(t)$.

Then the power function $\beta(\sigma^2)$ of the LRT for $\sigma^2 = 1$ against $\sigma^2 \neq 1$ is a monotone increasing function of $|\sigma^2 - 1|$.

Remarks:

1) Condition a) means that the family of distributions for U has the monotone likelihood ratio property.

2) If $m(t)$ is positive and nondecreasing, then $g(t) = \exp\left(-\int_0^t m(x)dx\right)$ satisfies (*) and a).

3) If $h(t)$ or $\log h(t)$ is concave, then $h(t)$ satisfies b)

Proof: Since $\infty > c \int_0^\infty u^{\frac{N}{2}-1} g(u)du$, it follows that

$$2^{-\frac{N}{2}} h(2t) = t^{\frac{N}{2}} g(2t) \leqslant t^{\frac{N}{2}-1} \int_t^{2t} g(u)du \leqslant \int_t^{2t} u^{\frac{N}{2}-1} g(u)du \to 0$$

for $t \to \infty$, that $h(0) = 0$ and $h(t) \geqslant 0$.

Therefore $t^{N/2} g(t) = c$ has two solutions $c_1 < c_2$ for $0 < c < h(t_0)$ with $h'(t_0) = 0$ from b).

The critical region K of the LRT is thus the union of the intervals $[0, c_1)$ and (c_2, ∞). It follows that

$$\beta(\sigma^2) = \int_K f_\sigma(u)du = c\left[\int_0^{\sigma^{-2}c_1} u^{\frac{N}{2}-1} g(u)du + \int_{\sigma^{-2}c_2}^{\infty} u^{\frac{N}{2}-1} g(u)du\right]$$

and

$$\frac{d\beta}{d\sigma^2} = c \cdot (\sigma^2)^{-\frac{N}{2}+1} c_1^{\frac{N}{2}} g(\sigma^{-2}c_2)\left[\frac{g(c_1)}{g(c_2)} - \frac{g(\sigma^{-2}c_1)}{g(\sigma^{-2}c_2)}\right] \gtreqless 0, \quad \text{for } \sigma^2 \gtreqless 1$$

from a).

In the case $p > 1$, X has density (1.2) with $M = 0$ and critical region K as in (2.1) and we have $W = X^\tau X \sim GW_p(\Sigma, N, g)$. Since the test problem and K are invariant under the transformation group $\{X \to X\Gamma : \Gamma \text{ orthogonal } p \times p \text{ matrix}\}$ we can assume that $\Sigma = \Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_p)$ where $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_p$ are the eigenvalues $ch_i(\Sigma)$ of Σ.

Lemma 2: Let $W \sim GW_p(\Lambda, N, g)$ and $W_0 = \operatorname{diag}(W_{11}, \ldots, W_{pp})$ where W_{ii} is the i-th diagonal element of W, $i = 1, \ldots, p$. Then $R = W_0^{-1/2} \cdot W \cdot W_0^{-1/2}$ and $(W_{11}, \ldots, W_{pp})$ are mutually independent. The density of R is

$$c\,|r|^{\frac{N-p-1}{2}} \tag{2.2}$$

and the density of $(W_{11}, \ldots, W_{pp})$ is

$$c \prod_{j=1}^{p} \lambda_j^{-\frac{N}{2}} w_{jj}^{\frac{N}{2}-1} g\left(\sum_{j=1}^{p} \frac{w_{jj}}{\lambda_j} \right). \tag{2.3}$$

The proof is similar to the proof in the case of normal distributions.

Theorem 2. Let $X \sim LEC_{N \times p}(0, \Sigma, g)$ and let $g(\cdot)$ have property (*). Suppose that

a) $g(a + ct)/g(a + t)$ is monotone increasing for $0 < c < 1$ and $a \geqslant 0$,

b) $h'(t) = 0$ has a unique solution, where $h(t) = t^{N/2} g(a + t)$ and $a \geqslant 0$.

Then the power function of the LRT for $\Lambda = I$ against $\Lambda \neq I$ is a monotone increasing function of $|\lambda_i - 1|, i = 1, \ldots, p$.

Proof: Using the Lemma 2 we obtain the critical region

$$K^* = \left\{ (r, w_{11}, \ldots, w_{pp}) : |r|^{N/2} \prod_{j=1}^{p} (w_{jj})^{N/2} g\left(\sum_{j=1}^{p} w_{jj} \right) \leqslant c \right\}.$$

Let

$$K_r^* = \left\{ (w_{11}, \ldots, w_{pp}) : \prod_{j=1}^{p} (w_{jj})^{N/2} g\left(\sum_{j=1}^{p} w_{jj} \right) \leqslant c\,|r|^{-N/2} \right\}$$

and

$$\beta(\lambda_1, \ldots, \lambda_p | R = r) = P((W_{11}, \ldots, W_{pp}) \in K_r^* | R = r)$$

then

$$\beta(\lambda_1, \ldots, \lambda_p) = E\beta(\lambda_1, \ldots, \lambda_p) | R).$$

If w_{jj} $(j \neq i)$ are fixed it can be established, analogously to the proof of Theorem 1, that $\beta(\lambda_1, \ldots, \lambda_p | R = r)$ is a monotone increasing function of $|\lambda_i - 1|$. Therefore $\beta(\lambda_1, \ldots, \lambda_p)$ has the same property.

II. $\mu \neq 0$

T.e. $X \sim LEC_{N \times p}(e_N \mu^\tau, \Sigma, g)$. The critical region of the modified LRT for the hypothesis $\Sigma = I$ against $\Sigma \neq I$ is

$$K_1 = \{v : |v|^{\frac{N-1}{2}} g(\text{tr } v) < c\} \qquad \text{with} \qquad V = (X - e_N \bar{X}^\tau)^\tau (X - e_N \bar{X}^\tau).$$

Theorem 3: Let $X \sim LEC_{N \times p}(e_N \mu^\tau, \Lambda, g)$ and let $g(\cdot)$ have property (*). Suppose that

a) $g_1(a + ct)/g_1(a + t)$ is monotone increasing for $0 < c < 1$ and $a \geqslant 0$ where $g_1(t) = c \int_0^\infty r^{\frac{p}{2}-1} g(t + r) dr$

b) $h'(t) = 0$ has a unique solution, where $h(t) = t^{\frac{N-1}{2}} g(a + t), a \geqslant 0$.

Then the power function of the modified LRT for $\Lambda = I$ against $\Lambda \neq I$ is a monotone increasing function of $|\lambda_i - 1|, i = 1, \ldots, p$.

Proof: From the proofs of Lemma 1 and Lemma 2 we know that $V \sim GW_p(\Lambda, N - 1, g)$ and that $(V_{11}, \ldots, V_{pp})$ has density (2.3) with v_{ii} in place of w_{ii} and $N - 1$ in place of N. Thus for fixed v_{jj} $(j \neq i)$ and $\Lambda = I$

$$\infty > \int_0^\infty v_{ii}^{\frac{N-1}{2}-1} g_1(v_{ii} + \sum_{j \neq i} v_{jj}) dv_{ii}$$

$$= \int_{\mathbb{R}} \left(\int_0^\infty v_{ii}^{\frac{N-1}{2}-1} g(v_{ii} + \sum_{j \neq i} v_{jj} + y^\tau y) dv_{ii} \right) dy$$

and

$$\infty > \int_0^\infty v_{ii}^{\frac{N-1}{2}-1} g(v_{ii} + a) dv_{ii}.$$

The remainder can be proved analogously to the proofs of Theorem 1 and Theorem 2.

Remark: Sufficient conditions for g and g_1 to satisfy the hypothesis in theorem 2 and 3 can be given analogously to the remarks following theorem 1.

3 Unbiasedness of the LRT for the Hypothesis $\Sigma = I$ and $\mu = 0$

Let $X \sim LEC_{N\times p}(e_N\mu^\tau, \Sigma, g)$. The critical region of the LRT for the hypothesis $\Sigma = I$ and $\mu = 0$ against $\Sigma \neq I$ or $\mu \neq 0$ is given by (see Anderson and Fang 1982b)

$$K = \{(v, X) : |v|^{\frac{N}{2}} g(\mathrm{tr}\,(v + N\bar{X}\bar{X}^\tau)) < c\}.$$

To prove the unbiasedness of this test we only have to show the following inequalities

(1) $P(K|\mu = 0, \Sigma = I) \leqslant p(K|\mu = 0, \Sigma)$ for $\Sigma > 0$,

(2) $P(K|\mu = 0, \Sigma) \leqslant p(K|\mu, \Sigma)$ for $\Sigma > 0$.

Lemma 3: Let $X \sim LEC_{N\times p}(e_N\mu^\tau, \Sigma, g)$ then $W = (X - e_N\mu^\tau)^\tau(X - e_N\mu^\tau)$ and $T = N(\bar{X} - \mu)^\tau W^{-1}(\bar{X} - \mu)$ are mutually independent, indeed $W \sim GW_p(\Sigma, \mathrm{N}, \mathrm{g})$ and $T \sim B\left(\frac{p}{2}, \frac{N-p}{2}\right)$.

The proof is similar to the proof in the case of normal distributions.

Lemma 4: Let $X \sim LEC_{N\times p}(M(k_1, \ldots, k_N), \Sigma, g)$ with $M(k_1, \ldots, k_N) = (k_1\mu_{(1)}, \ldots, k_N\mu_{(N)})^\tau$ for $0 \leqslant k_i \leqslant 1, i = 1, \ldots, N$, and let $g(\cdot)$ have property (*).

Let $\bar{K}$ be a subset of $\mathbb{R}^{N\times p}$ with the following property:

For $i = 1, \ldots, N$ and fixed $x_{(j)} \in \mathbb{R}^N, j \neq i$, the set

$$\bar{K}_i = \{y_{(i)} | y \in \bar{K}, y_{(j)} = x_{(j)}, \; j \neq i\} \subset \mathbb{R}^p$$

is convex and symmetric about the origin.

Then $P(X \in \bar{K})$ is a monotone decreasing function of k_i, $i = 1, \ldots, N$.

Proof: We shall only prove the lemma for $i = N$.

Let $0 \leqslant \tilde{k}_N < k_N \leqslant 1$,

$$X = \begin{pmatrix} X_1 \\ X_{(N)}^\tau \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} M_1 \\ k_N \mu_{(N)}^\tau \end{pmatrix}.$$

Then $X_1 \sim LEC_{(N-1)\times p}(M_1, \Sigma, g_1)$ with $g_1(t) = c \int_0^\infty r^{\frac{p}{2}-1} g(t+r)dr$.

If $h(M_1, \Sigma)$ is the density of X_1, the conditional density of $X_{(N)}$ given $X_1 = x_1$ is

$$q(x_{(N)} - k_N \mu_{(N)})$$

with $q(y) = |\Sigma|^{-N/2} g[y^\tau \Sigma^{-1} y + \text{tr}\, \Sigma^{-1}(x_1 - M_1)^\tau (x_1 - M_1)]/h(M_1, \Sigma)$.

The function $q(y)$ satisfies

(i) $q(y) = q(-y), \quad y \in \mathbb{R}^n$,

(ii) $\int_E q(y)dy < \infty$,

For any $X_1 \in \mathbb{R}^{(N-1)\times p}$ and $u \geqslant 0$ the set

$$Ku = \{y \,|\, q(y) \geqslant u\} \text{ is a convex set for } 0 \leqslant u < \infty.$$

From Theorem 10.2.1 (Srivastava and Khatri 1979, p. 299), it follows that

$$\int_{\bar{K}_N} q(x_{(N)} - k_N \mu_{(N)})dx_{(N)} \leqslant \int_{\bar{K}_N} q(x_{(N)} - \tilde{k}_N \mu_{(N)})dx_{(N)}$$

from which the assertion follows.

Clearly $P(X \in \bar{K})$ assumes its maximum at the point $M = 0$.

Theorem 4: Let $X \sim LEC_{N\times p}(e_N \mu^\tau, \Sigma, g)$ and let $g(\cdot)$ satisfy the assumptions of Theorem 2. Then the LRT for $\Sigma = I$ and $\mu = 0$ against $\mu \neq 0$ or $\Sigma \neq I$ is unbiased.

Proof: Firstly we prove the above inequality (1).

Because of the invariance property of this test we can assume that $X \sim LEC_{N \times p}(0, \Lambda, g)$ with $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_p)$. From Lemma 3 we see that the distribution of

$$\frac{|V|}{|W|} = 1 - N\bar{X}^T W^{-1} \bar{X} = 1 - T$$

does not depend upon W and Σ.

Then the critical region is

$$K = \{(w, t) : (1 - t)^{N/2} |w|^{N/2} g(\text{tr } w) < c\}.$$

Let

$$K_t = \{w : |w|^{N/2} g(\text{tr } w) < K(1 - t)^{-N/2}\} \quad \text{and}$$

$$\beta(\lambda_1, \ldots, \lambda_p | T = t) = P(W \in K_t | T = t).$$

From Theorem 2 we obtain that $\beta(\lambda_1, \ldots, \lambda_p | T = t)$ is a monotone increasing function of $|\lambda_i - 1|$. Since $P(K) = E\beta(\lambda_1, \ldots, \lambda_p | T = t)$, $P(K)$ reaches its minimum at $\Lambda = I$. That is inequality (1) holds.

In order to prove inequality (2) we proceed as in Lemma 1.

$$K = \{(z_1, z_{(N)}) : |z_1^T z_1|^{N/2} g(\text{tr } z_1^T z_1 + z_{(N)}^T z_{(N)}) < c\}.$$

For $z_1 \in \mathbb{R}^{(N-1) \times p}$ the set $\{z_{(N)} : z_{(N)}^T z_{(N)} \leqslant g^{-1}(c|z_1^T z_1|^{-N/2}) - \text{tr } z_1^T z_1\}$ is convex and symmetric about the origin.

It follows from Lemma 4 that $P(X \in \bar{K} | \mu \neq 0, \Sigma) \leqslant P(X \in \bar{K} | \mu = 0, \psi)$, i.e. that inequality (2) holds.

Acknowledgement: The author would like to thank Prof. Fang for giving the idea for this paper and to thank Prof. Lehn and Prof. Wegmann for valuable hints during its preparation. Thanks also to the referees for useful references.

References

Anderson TW (1955) The integral of a symmetric unimodal function over an symmetric convex set and some probability inequalities. Proc Am Math Soc 6:170–176

Anderson TW, Das Gupta S (1964) Monotonicity of the power function of some tests of independence between two sets of variates. Ann Math Statist 35:206–208

Anderson TW, Fank KT (1982a) On the theory of multivariate elliptically contoured distributions and application. Technical Report No 54 ONR, Contract No 0014-75-C-0442, Stanford, USA

Anderson TW, Fang KT (1982b) Maximum likelihood estimator and likelihood ratio criteria for multivariate elliptically contoured distribution. Technical Report No 1, Aro Contract DAAG 29-82-K-0156 Stanford, USA

Combanis S, Huang S, Simons G (1981) On the theory of elliptically contoured distributions. J Multivariate Analysis 11:368–385

Good IJ, Jensen DR (1981) Invariant distributions associated with matrix laws under structural symmetry. J Royal Statist Soc B 43:327–332

Jensen DR (1984) The structure of ellipsoidal distributions. Biometrical J 26:779–787

Kariya T, Sinha BK (1985) Nonull and optimality robustness of some tests. Ann Statist 13:1182–1197

Nagao H (1967) Monotonicity of the modified likelihood ratio test for a covariance matrix. J Sci Hiroshima University A1, 31:147–150

Schoenberg IJ (1938) Metric spaces and completely monotone functions. Ann Math 39:811–841

Srivastava MS, Khatri CG (1979) An introduction to multivariate statistics. North-Holland, New York

Sugiura N, Nagao H (1968) Unbiasedness of some test criteria for the equality of two covariance matrices. Ann Math Statist 39:1686–1692

Received 18. 8. 1987
(Revised version 30. 8. 1988)