SOME RESTRICTED ESTIMATION AND TESTING HYPOTHESES PROBLEM WITH SINGULARITY ASSUMPTION

Naeem A.Soliman*

ABSTRACT

The limiting distribution of the likelihood ratio statistic (LR) under a class of local alternatives as developed earlier by Davidson and Lever, is extended to the case of singular information matrices.

1. INTRODUCTION

One of the important problems in statistical inference is that of deciding, on the basis of n independent observations x_1, $x_2, .., x_n$ on a p-dimensional random variable X, whether or not a finite dimensional parameter $\theta = (\theta_1, \theta_2, ..., \theta_k)$, involved in the distribution function $F(., \theta)$ of the random variable X, belongs to a proper subset ω of the set Ω of all possible parameters.

In general, the procedure is to apply a certain test criterion, to decide with a certain degree of confidence whether to accept or reject a terminal hypothesis

---

* Department of Statistics, Faculty of Sience, King Abdulaziz University, Jeddah, Saudi Arab.

concerning the problem. This problem has recived considerable attention in the literature.

In this paper the criterion of the likelihood ratio test will be considered. We will generalize the results of Davidson and Lever (1970) on the limiting distribution of the likelihood ratio statistic (LR) under a class of local alternatives to the case when the information matrix, $B(\theta)$ is singular.

The typical problem to be considered is that we have a p-dimensional random variable $\mathbf{X}$ whose distribution function depends on k parameters $\theta_1, \theta_2, \ldots \theta_k$ which are not mathematically independent but satisfy q functional relationships,

$$h_i(\underline{\theta}) = 0 , \qquad i = 1, 2, \ldots, q, \; q< k, \qquad (1.1)$$

where;

$$\theta = (\theta_1, \theta_2, \ldots , \theta_k)' \qquad (1.2)$$

We assume that the possible values of $\underline{\mathbf{X}}$ lies in a subset R of the p-dimensional space. The k x k information matrix $B(\theta)$ defined by $B(\theta) = (B_{ij}(\theta))$, $i,j = 1, 2, \ldots, k$,

$B_{ij}(\theta) = \int \frac{\partial \ln f(t,\theta)}{\partial \theta_i} \cdot \frac{\partial \ln f(t,\theta)}{\partial \theta_j} d F(t,\theta)$ is of rank k-q and hence is singular. The null hypothesis is that the true unknown parameter point θ_o belongs to a subset ω of Ω such that the elements of ω satisfy the additional r-q constraints, $q < r$;

$$h_i(\theta) = 0, \quad i=q+1, \ldots, r. \qquad (1.3)$$

The limiting central chi-square distribution of LR for testing H_o in (1.3) when $B(\theta)$ is singular had been developed by EL-HELBAWY and Soliman (1983).The recent paper is concerned with with the limiting distribution of LR under the following sequence of local alternatives:

$\{\theta^N\}$ a sequence of true values of θ such that, each θ satisfies the following conditions;

(a) $h_i\ (\theta^N) = 0, \qquad i = 1, 2, \dots , q ,$ (1.4)

these are the identifiability conditions for the k parameters.

(b) $h_i\ (\theta^N) = \delta_i^N/\sqrt{N} , \qquad i = q + 1, .., r,$ (1.5)

with;

$$\lim_{N\to\infty} \delta_i^N = \delta_i , \qquad i = q + 1 , \dots , r. \quad (1.6)$$

Note that

$$\lim_{N\to\infty} h_i(\theta^N) = h_i(\theta_o) = 0 \qquad i = 1, 2, \dots\dots, r \quad (1.7)$$

2. TRANSFORMATION TECHNIQUE AND ASSUMPTIONS

We consider the following transformation on the parameter space,

$$\mathbf{h} = \mathbf{h}\ (\theta) = (\mathbf{h}_1\ (\theta), \mathbf{h}_2\ (\theta)) \quad (2.1)$$

such that;

$$\mathbf{h}_1\ (\theta) = (h_1(\theta), h_2(\theta), \dots , h_r\ (\theta)). \quad (2.2)$$

This transformation is required to satisfy the following properties (c.f. Bradley and Gart, 1962, p. 209);

(a) There exists a vector;

$$\mathbf{h}_2 (\theta) = (h_{r+1} (\theta), \ldots, h_k (\theta)) \quad (2.3)$$

such that the inverse relationships

$$\theta(\mathbf{h}) = (\theta_1(\mathbf{h}), \theta_2(\mathbf{h}), \ldots, \theta_k (\mathbf{h})), \quad (2.4)$$

exists, where $\mathbf{h}$ is defined by (2.1).

(b) The first, second and third-order partial derivatives of $\theta_i(\mathbf{h})$, $i = 1, \ldots, k$, with repect to $\mathbf{h}$ exist, and are bounded.

(c) The greatest lower bound, with respect to $\theta \epsilon \Omega$ of the absolute value of the jacobian $\left|\frac{\partial(h_1, \ldots, h_k)}{\partial(\theta_1, \ldots, \theta_k)}\right|$ is positive.

Under the transformation given by (2.1), the hypothesis H_o in (1.3) can be expressed as;

$$H_o: \quad \mathbf{h} = \mathbf{h}^{\omega} = (\mathbf{h}_1^{\omega}, \mathbf{h}_2), \ \mathbf{h}_1^{\omega} = [h_i : h_i (\theta) = 0, \ i = 1, \ldots, r]$$

and $\mathbf{h}_2$ is unspecified. (2.5)

Attention is restricted to the following sequence of local alternatives:

$$\mathbf{h}^N = (\mathbf{h}_1^N, \mathbf{h}_2^o),$$

where;

$$h_i^N = h_i(\theta) + \frac{\delta_{iN}^T}{\sqrt{N}}, \quad \text{with} \ \lim_{N\to\infty} \delta_{iN}^T = \delta_i^T, i = q + 1, \ldots, r,$$

and whereas,

$$h_i^N = h_i (\theta) = 0, \ i = 1, \ldots, q, \quad (2.6)$$

which are the identifiability conditions of the original parameters $\theta_1, \ldots, \theta_k$, where we assumed that $B(\theta)$ is of rank k-q, and $\mathbf{h}_2^\circ$ is the vector of true values of $\mathbf{h}_2$. Setting,

$$\mathbf{h}^\circ = (\mathbf{h}_1^\circ, \mathbf{h}_2^\circ), \quad (2.7)$$

then

$$\lim_{N\to\infty} h^N = h^o. \quad (2.8)$$

Let $f^T(X,\mathbf{h})$ be the transformed density function of $\mathbf{X}$ where $\mathbf{h} \in \Omega^T$, the image of Ω under the above transformation. We introduce the following assumptions, which are analogous to that given by Davidson and Lever (1970);

A1. For almost all $X \in R$ and for all $\mathbf{h} \in \Omega^T$,

$$\frac{\partial \ln f^T(X,h)}{\partial h_i}, \ \frac{\partial^2 \ln f^T(X,h)}{\partial h_i \partial h_j} \ and \ \frac{\partial^3 \ln f^T(X,h)}{\partial h_i \partial h_j \partial h_t}$$

exists for i, j, t = q+1, q+2, .. , k,

where the first q parameters, $h_1, h_2, \ldots, h_q$ equals zeors.

A2. For almost all $X \in R$ and for every $\mathbf{h} \in \Omega^T$,

$$\left|\frac{\partial f^T(X,h)}{\partial h_i}\right| < M_i(X) \ and \ \left|\frac{\partial^2 f^T(X,h)}{\partial h_i \partial h_j}\right| < M_{ij}(X)$$

where $M_i(X)$ and $M_{ij}(X)$ are integrable over R, i,j=q+1,..,k.

These Assumptions permit certain interchanges of order of differentiation and integration.

A3. For every $\mathbf{h} \in \Omega^T$, the (k-q) x (k-q) martix $B(\mathbf{h})$ = $[B_{ij}(\mathbf{h}),\ i,\ j = q+1,..,k]$ with,

$$B_{ij}(\mathbf{h}) = E_h \left[\frac{\partial \ln f^T(X,h)}{\partial h_i}\Big|_h \cdot \frac{\partial \ln f^T(X,h)}{\partial h_j}\Big|_h\right] \quad (2.9)$$

is positive definite with finite determinant. Where we have the first q paramters $h_1, .., h_q$ equals zeros, and since the original parameters $\theta_1, \theta_2, .., \theta_k$ are identifiable by introducing q constraints on it, and since the new parameters h_1, .., h_k are functions in θ this implies that h_1, .., h_k are identifiable but q of them equals zeros, hence the information matrix $B(h)$ is a function only in the remaning k-q parameters and has rank k-q, that is nonsingular.

A4. For almost all $X \in R$ and for all $h \in \Omega^T$,

$$\left|\frac{\partial^3 \ln f^T(X,h)}{\partial h_i \partial h_j \partial h_t}\right| < H_{ijt}(X)$$

where E_h $(H_{ijt}(X)) < Q < \infty$, Q is positive real number, for all $h \in \Omega^T$ and i, j, t = q + 1, .., k.

A5. There exists positive real numbers, τ, ξ such that whenever

$$\|h'' - h'\| = \sum_{i=q+1}^{k} |h''_i - h'_i| < \tau, \; h', h'' \in \Omega^T;$$

$$E_{h'}\left[\frac{\partial^2 \ln f^T(X,h)}{\partial h_i \partial h_j}\Big|_{h''}\right]^2 < \xi < \infty \; for \; i, j = q+1, \ldots, k.$$

A6. There exists positive real numbers η and U such that,

$$E_h\left[\left|\frac{\partial \ln f^T(X,h)}{\partial h_i}\right|_h^{2+\eta}\right] < U < \infty$$

for all $h \in \Omega^T$ and i = q + 1, ..., k.

This Assumption will be used to prove that;

$$\frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} \frac{\partial \ln f^T(X_\alpha, h)}{\partial h_i}, \qquad i = q+1, \ldots, k.$$

have a limitng multivariate normal distribution.

A7. There exists positive real numbers N and L such that;

$$E_h\left[\left|H_{ijt}(X) - E_h\{H_{ijt}(X)\}\right|^{1+N}\right] < L <, \infty,$$

for all $h \in \Omega^T$ and i,j,t = q+1,...,k where $H_{ijt}(x)$ is defined by A4.

3. THE ASYMPTOTIC PROPERTIES OF NULL AND NON-NULL MAXIMUM LIKELIHOOD ESTIMATORS

The following two lemmas are restatements of lemmas 1 and 2 of Davidson and lever (1970)

LEMMA1: Under Assumptions A1 - A5, there exists a positive real number S such that;

$$E_{h'}[(\frac{\partial \ln f^T(X,h)}{\partial h_i}|_{h''})^2] < S < \infty \quad , i = q+1, \ldots, k,$$
$$\text{whenever } \|h'' - h'\| < \tau .$$

LEMMA 2: Under Assumptions A1 - A5, and for i,j = q + 1, ..., K;

(i) C_i (**h**, **h"**) and G_i (**h**, **h"**) are continuous at **h"** = **h'** for **h**, **h'** and **h"** ϵ Ω^T where $\|\mathbf{h} - \mathbf{h}'\| \leq \frac{1}{2}\tau$

(ii) B_{ij} (**h**) = - G_{ij} (**h**, **h**) is continuous in **h**, where,

$$C_i\ (h,h') = E_{h'}\ [\frac{\partial \ln f^T(X,h)}{\partial h_i}|_h] , \qquad i = q+1, \ldots, k \quad (3.1)$$

$$G_{ij}\ (h,h') = E_{h'}\ [\frac{\partial \ln f^T(X,h)}{\partial h_i \partial h_j}|_h] , \quad i,j = q+1, \ldots, k \quad (3.2)$$

Now, the likelihood equations under H_o in (2.5) and under the alternatives in (2.6) are respectively;

$$\frac{\partial \ln L(X,h^{\omega})}{\partial h_i} = 0 , \quad i = r+1, \ldots, k \qquad (3.3)$$

and,

$$\frac{\partial \ln L(X,h)}{\partial h_i} = 0 \quad , i = q+1, \ldots, k \qquad (3.4)$$

The asymptotic properties of the maximum likelihood estimates which arise as solution to (3.3) and (3.4) will be developed by Theorm 1. We introduce the following notations for use in subsequent discussions;

$$D_i^n(h) = \frac{1}{n}\frac{\partial \ln L(X,h)}{\partial h_i} = \frac{1}{n}\sum_{\alpha=1}^{n}\frac{\partial \ln f^T(X_\alpha,h)}{\partial h_i}, \quad i = q+1, .., k \tag{3.5}$$

$$D_{ij}^n(h) = \frac{1}{n}\frac{\partial^2 \ln L(X,h)}{\partial h_i \partial h_j} \qquad i,j = q+1, \ldots, k \tag{3.6}$$

$$D_{ijt}^n(h) = \frac{1}{n}\frac{\partial^3 \ln L(X,h)}{\partial h_i \partial h_j \partial h_t} \quad i,j,t = q+1, \ldots, k \tag{3.7}$$

LEMMA 3: For the sequence $\{\mathbf{h}^N\}$ of local alternatives, and under Assumptions A 1 - A5, the following hold for all $\mathbf{h}$ and such that $\|\mathbf{h} - \mathbf{h}^\circ\| \le \frac{1}{2}\tau$:

(i) $P \lim_{n\to\infty} [D_i^n(\mathbf{h}) | \mathbf{h}^N] = C_i(\mathbf{h}, \mathbf{h}^\circ)$, $i = q+1, \ldots, k$

(ii) $P \lim_{n\to\infty} [D_{ij}^n(\mathbf{h}) | \mathbf{h}^N] = G_{ij}(\mathbf{h}, \mathbf{h}^\circ)$, $i,j = q+1, \ldots, k$.

This lemma is a restatement of lemma 3 of davidson and Lever (1970) and hence we gave it without proof. The following result is a restatement of lemma 2 of Aitchison and Silvey (1958, P-819) and will be used in the proof of the consistency of the maximum likelihood estimates which arise as solution of (3.3) and (3.4).

LEMMA 4: If J is a continuous function mapping R^K into itself with property that, for every $\mathbf{h}$ such that

$\|\mathbf{h} - \mathbf{h}^\circ\| = \delta, \delta > 0, \sum_{i=q+1}^{k} J_i(h)(h_i - h_i^\circ) < 0$, then there exists a point $\hat{\mathbf{h}}$ such that $\|\hat{\mathbf{h}} - \mathbf{h}^\circ\| < \delta$ for which $J(\hat{\mathbf{h}}) = 0$.

Theorem 1: Under Assumptions A1 - A5 and the sequence $\{\mathbf{h}^N\}$ of local alternatives;

(i) There exists a sequence $\{\hat{\mathbf{h}}_n\}$ of solutions to the likelihood equations (3.4) which converge in probability to $\mathbf{h}^\circ$.

(ii) There exists a sequence $\{\hat{\hat{\mathbf{h}}}_n = \hat{\hat{\mathbf{h}}}_2\}$ of solutions to the likelihood equations (3.3) which converge in probability to $\mathbf{h}^\circ$.

<u>Proof</u>: Follows the technique of proof of Davidson and Lever (1970):

(i) Consider the likelihood equations $D_i^n(\mathbf{h}) = 0$, $i = q+1, \ldots, k$ as defined under the alternatives sequence $\{\mathbf{h}^N\}$ by (3.4). Given ν, ϵ; an arbitrary positive constants, it follows from lemma 3, that there exists a positive integer $N(\nu, \epsilon)$ such that, for all $\mathbf{h}$ with $\|\mathbf{h} - \mathbf{h}^\circ\| \le \frac{1}{2}\tau$,

$P\,[\,|D_i^n(\mathbf{h}) - C_i(\mathbf{h}, \mathbf{h}^\circ)| < \nu$, for all $i = q+1, \ldots, k \mid \mathbf{h}^N] > 1-\epsilon$, for $n > N(\nu, \epsilon)$. Thus, for $n > N(\nu, \epsilon)$ we have,

$$P\,[\sum_{i=q+1}^{k} (D_i^n(h) - C_i(h, h^\circ))\,(h_i - h_i^\circ) < \nu\,\|h - h^\circ\| \,|\, h^N] > 1 - \epsilon. \tag{3.8}$$

so that, for each $\mathbf{h}$ such that $\|\mathbf{h} - \mathbf{h}^\circ\| = \delta \le \frac{1}{2}\tau$;

$$P\,[\sum_{i=q+1}^{k} D_i^n(h)\,(h_i - h_i^\circ) < \sum_{i=q+1}^{k} C_i(h, h^\circ)\,(h_i - h_i^\circ) + \nu\,\delta \,|\, h^N] > 1-\epsilon. \tag{3.9}$$

By expanding $\dfrac{\partial \ln f^T(X, h)}{\partial h_i}$ in a Taylor series about $\mathbf{h} = \mathbf{h}^\circ$ in the expression for $C_i(\mathbf{h}, \mathbf{h}^\circ)$, multiplying by $(h_i - h_i^\circ)$ and summing over i, and noting that $C_i(\mathbf{h}^\circ, \mathbf{h}^\circ) = 0$, we have;

$$\sum_{i=q+1}^{k} C_i(h, h^\circ)\,(h_i - h_i^\circ) < \sum_{i=q+1}^{k} \sum_{j=q+1}^{k} (h_i - h_i^\circ)\,(h_j - h_j^\circ)\,G_{ij}(h^\circ, h^\circ) + \frac{1}{2}\delta^3 M \tag{3.10}$$

using Asumption A4 provided $\|\mathbf{h} - \mathbf{h}^{\circ}\| = \delta$.

By Assumption A3, the matrix $[G_{ij}(\mathbf{h}^{\circ}, \mathbf{h}^{\circ})] = [-B_{ij}(\mathbf{h}^{\circ})]$ is negative definite, so that there exists a $\gamma > 0$, the smallest characteristic root of the matrix $[B_{ij}(\mathbf{h}^{\circ})]$, such that;

$$\sum_{i=q+1}^{k} \sum_{j=q+1}^{k} (h_i - h_i^o)(h_j - h_j^o)\, G_{ij}(h^{\circ}, h^{\circ}) \le -\gamma\delta^2 \qquad (3.11)$$

Thus, for arbitrary $\nu \le \frac{1}{2}\delta^2 M$, and $\|\mathbf{h}-\mathbf{h}^{\circ}\| = \delta < \min(\gamma/M, \frac{1}{2}\tau)$ we get;

$$P[\sum_{i=q+1}^{k} D_i^n(h)(h_i - h_i^o) < -\gamma\delta^2 + \delta^3 M < 0 | h^N] > 1-\epsilon, \text{ for } n > N(\nu, \epsilon) \quad (3.12)$$

Now, applying Lemma 4 to conclude that for each $n > N(\nu, \epsilon)$, there exists an $\hat{\mathbf{h}}_n(., \Omega^T)$ such that $\|\hat{\mathbf{h}}_n(., \Omega^T) - \mathbf{h}^{\circ}\| < \delta$ for which,

$P[D_i^n(\hat{\mathbf{h}}_n) = 0$ for all $i = q+1, \ldots, k | \mathbf{h}^N] > 1-\epsilon$.

Thus there exists a sequence of roots $\{\hat{\mathbf{h}}_n(., \Omega^T)\}$ to the likelihood equations (3.4) in the region $\|\hat{\mathbf{h}}_n(., \Omega^T) - \mathbf{h}^{\circ}\| < \delta$ with probability greater than $1-\epsilon$. Since δ and ϵ are arbitrary and may be taken small, then the sequence $\{\hat{\mathbf{h}}_n(., \Omega^T)\}$ converges in probability to $\mathbf{h}^{\circ}$.

(ii) A parallel argument suffices to show that there exists a sequence of roots $\{\hat{\hat{\mathbf{h}}}_n(., \omega)$ to the likelihood equations.

$D_i(\mathbf{h}^{\omega}) = 0, \qquad i = r+1, \ldots, k$

which converges in probability to $\mathbf{h}^{\circ}$, where,

$$\hat{\hat{\mathbf{h}}}_{in}(., \omega) = \mathbf{h}_i^{\omega} = \mathbf{h}_i^{\circ} = 0 \qquad (3.13)$$

by definition.

Now, by using a Theorem of Cramèr Wold (cf. Cramèr, 1962, P. 105-6), and the result of lemma 2, namely, $\lim_{N\to\infty} B(h^N) = B$ and using a parallel argument of the proof of lemma 5 of Davidson and Lever (1970) to demonstrate that Assumptions A1 - A 6 suffice for the Levy-Feller Theorem (cf. Loeve, 1963, P. 295), the following Lemma can be stated as a restatement of lemma 5 of Davidson and Lever (1970).

<u>LEMMA 5</u>: Under Assumptions A1 - A 6 and the sequence $\{h^N\}$ of local alternatives, the sequence of vectors,

$$D^n (h^N) = [\sqrt{n}\, D_i^n (h^N), \quad i = q + 1, \ldots, k].$$

converges in distribution to the multivariate normal distibution,

$N(0, B, (h^\circ))$ with mean vector O and variance matrix $B(h^\circ) = [B_{ij} (h^\circ), i,j = q + 1, \ldots, k]$.

Lemmas 6 and 7 below, is a restatment of lammas 6 and 7 of Davidson and Lever (1970).

<u>LEMMA 6</u>: Under Assumptions A1 - A4 and A7, $D^n_{ijt} (h)$ as defined by (3.7) has a limiting bound in probability namely for any $\epsilon > 0$,

$P\,[|D^n_{ijt} (h)| > M | h^N] < \epsilon$, $i,j,t = q + 1, \ldots, k$, for n sufficiently large, where M is the positive constant introduced in Assumption A4.

<u>LEMMA 7:</u> Let $\{\hat{h}_n (., \Omega^T)\}$ be a sequence of parameter estimates in Ω^T which converge in probability to h° under the parametric sequence $\{h^N\}$. Then under the parametric

sequence $\{\mathbf{h}^N\}$,

$$P \lim_{n\to\infty} [D_{ij}^n(\hat{h}_n(.,\Omega^T)|h^N)] = -B_{ij}(h^o), \qquad i,j = q+1,\ldots,k.$$

In the following Theorem 2, we shall state the limiting normality of $\hat{\mathbf{h}}_n$ $(.,\ \Omega^T)$ and $\hat{\hat{\mathbf{h}}}_n$ $(.,\ \omega)$ the maximum likelihood estimates under the alternatives and the null hypotheses respectively.

Theorem 2: Let $\hat{\mathbf{h}}_n$ $(.,\ \Omega^T)$ and $\hat{\hat{\mathbf{h}}}_n$ $(.,\ \omega)$ be the two sets of estimators of Theorem 1. Let the matrices $\mathrm{B} = \mathrm{B}\ (\mathbf{h}^\circ)$ and $\Sigma = \mathrm{B}^{-1}$, and let $\delta = (\delta_{q+1},\ \delta_{q+2},\ \ldots,\ \delta_r,\ 0,\ \ldots,\ 0)$. Then under Assumptions A1-A7 and for the sequence $\{\mathbf{h}^N\}$ of local alternatives;

(i) $\sqrt{n}\ (\hat{\mathbf{h}}_n\ (.,\ \Omega^T) - \mathbf{h}^\circ)$ has a limiting multivariate normal distriburtion with mean vector δ and variance covariance matrix Σ.

(ii) $\sqrt{n}\ (\hat{\hat{\mathbf{h}}}_{2n}\ (.,\ \omega) - \mathbf{h}_2^\circ)$ has a limiting multivariate normal distribution with mean vector O_2 and variance convariance matrix $\bar{\Sigma}_{22}$ where O_2 is null vector of k-r components and $\bar{\Sigma}_{22} = [\mathrm{B}_{22}\ (\mathbf{h}^\circ)]^{-1}$.

Proof: Follows the technique of Davidson and Lever (1970); expanding $\sqrt{n}\ \mathrm{D}_i^n\ (\mathbf{h}^\circ)$ about $\mathbf{h} = \mathbf{h}^N$ by Taylor's series expansion we get.

$$\sqrt{n}D_i^n(h^o) = \sqrt{n}D_i^n(h^N) + \sqrt{n}\sum_{s=q+1}^{r} (h_s^o - h_s^N)D_{is}^n\ (\tilde{h}^N), \quad i=q+1,\ldots,k. \tag{3.14}$$

where

$$h_i^N = h_i^\circ, \qquad i = r+1,\ \ldots,\ k,$$

and for some $\tilde{\mathbf{h}}^N$ such that $\|\tilde{\mathbf{h}}^N - \mathbf{h}^\circ\| < \|\mathbf{h}^\circ - \mathbf{h}^N\|$.

Now, $\lim_{N\to\infty} \mathbf{h}^N = \mathbf{h}^\circ$ implies that, $\lim_{N\to\infty} \tilde{\mathbf{h}}^N = \mathbf{h}^\circ$, so that, it follows from lemma 7 and slutsky's Theorem that for i= q+1, ...,k.

$$P\lim_{N\to\infty} [\sqrt{n} \sum_{s=q+1}^{r} (h_s^\circ - h_s^N)\ D_{is}^n\ (\tilde{h}^N|h^N] = \sum_{s=q+1}^{r} \delta_s B_{is}(h^\circ). \quad (3.15)$$

From lemma 5, D^n $(\mathbf{h}^N)$ converges in distibution to the multivariate normal distribution (MVN), (0, B $(\mathbf{h}^\circ)$), it follows that, [$\sqrt{n}$ D_i^n $(\mathbf{h}^\circ)$, i = q + 1 , ..., k] has as its limiting MVN (δ B $(\mathbf{h}^\circ)$, B $(\mathbf{h}^\circ)$).

(i) Expanding D_i^n $(\mathbf{h}^\circ)$ about $\mathbf{h} = \hat{\mathbf{h}}_n$ (., Ω^T), so that,

$$D_i^n\ (h^\circ) = -\sum_{j=q+1}^{k} (\hat{h}_{jn} - h_j^\circ)\ D_{ij}^n\ (\tilde{h}_n), \quad i= q + 1 , \ldots ,k \quad (3.16)$$

for some $\tilde{\mathbf{h}}_n$ such that $\|\tilde{\mathbf{h}}_n - \mathbf{h}^\circ\| < \|\hat{\mathbf{h}}_n\ (., \Omega^T) - \mathbf{h}^\circ\|$.

Now by Theorem 1,

$$P\lim_{n\to\infty} [\hat{h}_n(.,\Omega^T)\,|h^N] = h^\circ. \quad (3.17)$$

so that,

$$P\lim_{n\to\infty} [\tilde{h}_n\ |h^N] = h^\circ \quad (3.18)$$

Then using lemma 7 we get;

$$\sqrt{n}D_i^n(h^\circ) = \sum_{j=q+1}^{k} \sqrt{n}(\hat{h}_{jn}-h_j^\circ)\ [B_{ij}(h^\circ) + o_p(1)],\ i=q+1,\ldots,K. \quad (3.19)$$

where o_p (1) denotes a quantity which converges to zero in probability under $\{\mathbf{h}^N\}$. The matrix [B_{ij} $(\mathbf{h}^\circ)$ + o_p (1), i,j = q+1,..k] approaches $B(\mathbf{h}^\circ)$ in probability for large n and thus may be inverted to give the matrix [σ_{ij} $(\mathbf{h}^\circ)$ + o_p (1), i,j = q+1,..,k] where,

$\Sigma = [\sigma_{ij}\ (\mathbf{h}^{\circ}),\ i,j = q+1, \ldots, k]$. Thus it follows that,

$$\sqrt{n}\ (\hat{h}_{in} - h_i^{o}) = \sqrt{n} \sum_{j=q+1}^{k} D_i^{n}\ (h^{o})\{\sigma_{ji}(h^{o}) + o_p(1)\},\ i=q+1, \ldots, k. \qquad (3.20)$$

Then from lemma 1 of chiang (1956, p. 338) it follows that, $\sqrt{n}\ (\hat{\mathbf{h}}_n\ (.,\ \Omega^T) - \mathbf{h}^{\circ})$ has a limiting multivariate normal distribution with mean vector $\delta\ B\ B^{-1} = \delta$ and variance covariance matrix, $B^{-1}\ B\ B^{-1} = B^{-1} = \Sigma$.

(ii) Expanding $D_i^{n}\ (\mathbf{h}^{\circ})$ about $\mathbf{h} = \hat{\hat{\mathbf{h}}}_n\ (.,\ \omega),\ i = r+1, .., k$ to obtain, since $\hat{\hat{h}}_{in} = h_i^{\circ}$, $\quad i = q+1, \ldots, r$;

$$D_i^{n}\ (h^{o}) = -\sum_{j=r+1}^{k} (\hat{\hat{h}}_{jn} - h_j^{o})\ D_{ij}^{n}\ (\tilde{h}_n)\ ,\ i = r+1, \ldots, k.$$

where $\|\tilde{h}_n - h^{o}\| < \|\hat{h}_n(.,\omega) - h^{o}\|$.

By Theorem 1; $P \lim_{n\to\infty} [\hat{\hat{\mathbf{h}}}_n\ (.,\ \omega)\,|\,\mathbf{h}^{N}] = \mathbf{h}^{\circ}$ and as in part (i) one can obtain from lemma7,

$$\sqrt{n}(\hat{\hat{h}}_{in} - h_i^{o}) = \sum_{j=r+1}^{k} \sqrt{n} D_j^{n}(h^{o})\{\bar{\sigma}_{ji}(h^{o}) + o_p(1)\},\quad i=r+1, \ldots, k. \qquad (3.21)$$

where $\bar{\Sigma}_{22} = [\bar{\sigma}_{ij}\ (\mathbf{h}^{\circ}),\ i,j = r+1, ., k]$.

Now $_2D^{n}\ (\mathbf{h}^{\circ}) = [\sqrt{n}\ D_j^{n}\ (\mathbf{h}^{\circ}),\ i = r+1, \ldots, k]$ has a limiting normal distribution $N\ (_2O,\ B_{22}\ (\mathbf{h}^{\circ}))$ as it is the marginal distribution of the last k-r components of $D^{n}\ (\mathbf{h}^{\circ})$.

Again, using lemma 1 of chiang (1956, p 338) we see that, $\sqrt{n}\ (\hat{\hat{\mathbf{h}}}_{2n}\ (.,\ \omega) - \mathbf{h}^{\circ}_2)$ has a limiting multivariate normal distribution with mean vector $_2O$ and variance-covariance matrix $\bar{\Sigma}_{22}\ B_{22}\ \bar{\Sigma}_{22} = \bar{\Sigma}_{22} = [B_{22}\ (\mathbf{h}^{\circ})]^{-1}$.

<u>The Asymptotic Distribution of LR</u>:

Now, in a way similar to that given by Davidson and Lever (1970) we shall drive the Asymptotic distribution of the likelihood ratio statistic (LR) under the sequence $\{h^N\}$ of local alternatives, when the information matrix B (θ) is singular, but under the transformation $h = h(\theta)$ we introduced a new parameters $(h_1, h_2, .., h_k)$ where $h_1 = h_2 = ... = h_q = 0$, the necessary conditions for identifiability of the original parameters $\theta_1, ..., \theta_k$, and in this case the information matrix of the new parameters $h_{q+1}, ..., h_k$; $B(h)$ is non-singular.

Expanding $\ln L(X, h^o)$ about $\hat{h}_n(., \Omega^T)$ and $\hat{\hat{h}}_n(., \omega)$ respectively, we have,

$$\ln L(X,h^o) = \ln L(X,\hat{h}_n) + \frac{1}{2} \sum_{i=q+1}^{k} \sum_{j=q+1}^{k} n(\hat{h}_{in}-h_i^o)(\hat{h}_{jn}-h_j^o)\, D_{ij}^n(\tilde{h}_n),$$

$$\ln L(X,h^o) = \ln L(X,\hat{\hat{h}}_n) + \frac{1}{2} \sum_{i=r+1}^{k} \sum_{j=r+1}^{k} n(\hat{\hat{h}}_{in}-h_i^o)(\hat{\hat{h}}_{jn}-h_j^o)\, D_{ij}^n(\tilde{\tilde{h}}_n),$$

for some $\tilde{h}_n$ such that $\|\tilde{h}_n - h^o\| < \|\hat{h}_n - h^o\|$ and some $\tilde{\tilde{h}}_n = (h_1^-, \tilde{\tilde{h}}_{2n})$, such that $\|\tilde{\tilde{h}}_n - h^o\| < \|\hat{\hat{h}}_n - h^o\|$. Since by Theorem 1,

$$P\lim_{n\to\infty} [\hat{h}_n | h^N] = P\lim_{n\to\infty} [\hat{\hat{h}}_n | h^N] = h^o,$$

then, $P\lim_{n\to\infty} [\tilde{h}_n | h^N] = P\lim_{n\to\infty} [\tilde{\tilde{h}}_n | h^N] = h^o$.

Thus using lemma 7, we have;

$$LR = 2\,[\ln L(X, \hat{h}_n) - \ln L(X, \hat{\hat{h}}_n)].$$

$$= n(\hat{h}_n - h^o)\, B(h^o)\, (\hat{h}_n - h^o)' - n(\hat{\hat{h}}_{2n} - h_2^o)\, B_{22}(h^o)\, (\hat{\hat{h}}_{2n} - h_2^o)' + o_p(1) \qquad (3.22)$$

where, $\hat{\hat{h}}_{2n} = (\hat{\hat{h}}_{r+1,n}, \hat{\hat{h}}_{r+2,n}, \ldots, \hat{\hat{h}}_k)$,

$$B_{22}(h^{\circ}) = \{E_h[\frac{\partial \ln f^T(X,h)}{\partial h_i}|_h \frac{\partial \ln f^T(X,h)}{\partial h_j}|_h], \quad i,j = r+1,\ldots,k\}$$

Consider the second term of the right-handside of LR in (3.22), it follows from (3.21) that,

$$n(\hat{\hat{h}}_{2n}-h_2^{\circ})B_{22}(h^{\circ})(\hat{\hat{h}}_{2n}-h_2^{\circ})' = [{}_2D^n(h^{\circ})\bar{\Sigma}_{22}+o_p(1)]\ B_{22}(h^{\circ})\ [{}_2D^n(h^{\circ})\bar{\Sigma}_{22}+o_p(1)]'$$

$$= {}_2D^n\ (h^{\circ})\bar{\Sigma}_{22}\{{}_2D^n(h^{\circ})\}'+o_p(1) \qquad (3.23).$$

where ${}_2D^n\ (h^{\circ})\ [\sqrt{n}\ D_i^n\ (h^{\circ}), \quad i = r+1, \ldots, k]$.

Now By Theorem 2, $\sqrt{n}\ (\hat{h} - h^{\circ})$ has a limiting distribution and by (3.19) we have,

$$D^n\ (h^{\circ}) = \sqrt{n}\ (\hat{h}_n - h^{\circ})\ B(h^{\circ}) + o_p\ (1)$$

Then,

$${}_2D^n(h^{\circ})\bar{\Sigma}_{22}[{}_2D^n(h^{\circ})]' = n(\hat{h}_n-h^{\circ})\ A(\hat{h}_n-h^{\circ})'+o_p(1), \qquad (3.24)$$

where,

$$A = \begin{bmatrix} B_{12}\ (h^{\circ}) \\ B_{22}\ (h^{\circ}) \end{bmatrix} [B_{22}\ (h^{\circ})]^{-1}\ [B_{21}(h^{\circ}),\ B_{22}\ (h^{\circ})]$$

$$= \begin{bmatrix} B_{12}(h^{\circ})[B_{22}\ (h^{\circ})]^{-1}\ B_{21}(h^{\circ}) & B_{12}\ (h^{\circ}) \\ B_{21}\ (h^{\circ}) & B_{22}\ (h^{\circ}) \end{bmatrix}$$

Therefore, the second term of the right-hand side of LR,

$$n(\hat{\hat{h}}_{2n} - h^{\circ}_2)B_{22}(h^{\circ})(\hat{\hat{h}}_{2n} - h^{\circ}_2)' = n(\hat{h}_n - h^{\circ})\ A(\hat{h}_n - h^{\circ})' + o_p\ (1).$$

Then;

$$LR = n(\hat{h}_{1n}-h^{\circ}_1)\ [B_{11}(h^{\circ}) - B_{12}\ (h^{\circ})\ [B_{22}(h^{\circ})]^{-1}\ B_{21}\ (h^{\circ})].$$

$$(\hat{h}_{1n} - h^{\circ}_1)' + o_p(1)$$

$$= n(\hat{h}_{1n} - h_1^{\circ})\ \bar{B}_{11}(h^{\circ})(\hat{h}_{1n} - h^{\circ}_1)' + o_p\ (1) \qquad (3.25)$$

, where $\bar{B}_{11}\ (h^{\circ}) = \Sigma^{-1}_{11}$,

$$\hat{\mathbf{h}}_{1n} = (\hat{h}_{q+1,n}, \hat{h}_{q+2,n}, \ldots \hat{h}_{rn}),$$

$$B_{11}(\mathbf{h}^\circ) = [E_h[\frac{\partial \ln f^T(X,h^\circ)}{\partial h_i}|_h \frac{\partial \ln f^T(X,h^\circ)}{\partial h_j}|_h], i,j=q+1,..,r],$$

$$B_{12}(\mathbf{h}^\circ) =$$

$$[E_h[\frac{\partial \ln f^T(X,h^\circ)}{\partial h_i}|_h \frac{\partial \ln f^T(X,h^\circ)}{\partial h_j}|_h], i=q+1,..,r; j=r+1,..,q.$$

Now, since $\sqrt{n}$ $(\hat{\mathbf{h}}_{1n} - \mathbf{h}_1^\circ)$ has a limiting normal distribution with mean vector $\delta_1 = \delta_{q+1}, \delta_{q+2}, \ldots, \delta_r)$ and variance-convariance matrix Σ_{11} then by Theorem 2, it follows that, $g_n = n$ $(\hat{\mathbf{h}}_{1n} - \mathbf{h}_1^\circ)$ $\bar{B}_{11}(\mathbf{h}^\circ)$ $(\hat{\mathbf{h}}_{1n} - \mathbf{h}_1^\circ)'$ has a limiting non central chi-square distribution with non-centrality parameter.

$$\alpha = \lim_{n\to\infty} \sum_{i=q+1}^{r} \sum_{j=q+1}^{r} n(h_{in}-h_i^\circ)(h_{jn}-h_j^\circ)\bar{B}_{ij}(h^\circ) = \delta_1 \bar{B}_{11}(h^\circ)\delta_1' \quad (3.26)$$

Since g_n and LR have the same limiting distribution in that they differ only by a quantity which converges to zero in probability, we can establish the following Theorem;

<u>Theorem 3</u>:

Under the sequence $[\mathbf{h}^N]$ of local alternatives and Assumptions A1 - A7, the likelihood ratio statistic LR for testing the hypothesis $H_o = \mathbf{h} = \mathbf{h}^\omega$, where $\mathbf{h}^\omega = (\mathbf{h}_1^\omega, \mathbf{h}_2) \in \Omega^T$, the image of Ω under the transformation $\mathbf{h} = \mathbf{h}(\theta)$ where θ is a k-dimensional parameter $\in \Omega$; has a limiting non-central chi-square distribution with r- q degrees of freedom and non-centrality parameter $\alpha = \delta_1 \bar{\mathbf{B}}_{11} (\mathbf{h}^\circ) \delta_1'$.

REFERENCES

Aitchinson, J. and Silvey, S.D. (1958): Maximum-Likelihood estimation of parameters subject ot restrants. Ann. Math. Stat., 29, 813-828.

Bradley, R.A. and Gart, J.J. (1962): The asymptotic properties of ML estimators when sampling from associated populations. Biometrik 49: 205-214.

Chiang, C.L. (1956): On best regular asymptotically normal estimates. Ann. Math. Stat., 27, 336-351.

Cramer, M. (1962): Random Variables and Probability Distribution Cambridge Tracts in Mathematics and Mathematical Physics, 36.

Davidson, R.R. and Lever, W.E. (1970): The limiting distribution of the likelihood ratio test statistic under a class of local alternatives. Sankhya, Series A, Vol. 32, Part 2, 209-224.

El-Helbawy, A.T. and Soliman, N.A. (1983): A note on the limiting distribution of the likelihood rato statistic for testing equality constraints. The Egyptian Statistical Journal, Vol. 27, No. 2.P. 62-71.

Loeve, M, (1963): Probability Theory, 3 rd. ed., Van Nostrand, Princeton New Jersey.