THE EGYPTAIN STATISTICAL JOURNAL
ISSR. CAIRO UNIV., VOL.,41, NO.2, 1997

ON A CONSTRAINED TESTING HYPOTHESES PROBLEM

NAEEM AHMED SOLIMAN
Department of Statistics, Faculty of Science
King Abdulaziz University, Jeddah 21441 SAUDI ARABIA
E-Mail: ZCC3488@SAKAAU03.BITNET

Key words: Likelihood ratio test; Asymptotic non-central chi-square distribution; Class of local alternatives; Singular information matrix; Equality constraints; Lagrangian multiplier test; Identifiability conditions; Restricted likelihood equations; Positive definite.

ABSTRACT

This paper gives a proof that the likelihood ratio statistic, based on a sample $x=(x_1,\ldots,x_n)$ on a p-dimensional random variable X, converges in distribution to a noncentral chi-square distribution under a class of local alternatives, for a multi-dimensional parameter space. A proof of uniform convergence for this situation was given by Wald (1943) whose assumptions include the uniform consistency of the maximum likelihood estimates and of the likelihood ratio test. The assumptions utilized in this paper can be more directly verified in applications than those required by Wald. This paper is concerned with the case in which the information matrix is not of full rank. This generalizes the results of Silvey (1959), Davidson and Lever (1970) and El-Helbawy and Soliman (1983).

1. INTRODUCTION

The main object in many statistical situations is to test a null hypothesis H_0 that the true unknown parameter $\theta=(\theta_1,\theta_2,\ldots,\theta_k)$, involved in the distribution function $F(\cdot,\theta)$ of a given random variable X, satisfies certain equality constraints. Aitchison and Silvey (1958) developed the Lagrangian multiplier test, for testing H_0. They demonstrated that, under assumptions analogous to those of Cramer (1946), the Lagrangian multiplier test statistic (LMS) has an asymptotic chi-square distribution under H_0. Silvey (1959) proved that, under assumptions analogous to those of Wald (1943) each of the LMS and the likelihood ratio test statistic $(-2\ln\lambda)$ has the same asymptotic chi-square distribution, when the information matrix $B(\theta)$ is nonsingular. Davidson and Lever (1970) developed the asymptotic distribution of $-2\ln\lambda$ under a class of local alternatives when the information matrix is nonsingular.

It often happens, either for reason of symmetry or for some other reason, that the information matrix $B(\theta_0)$, where θ_0 is the true parameter value, is not positive definite. For instance a multinomial distribution describes an experiment in which an individual can fall into any of s classes, so we have s dimensional parameter space, R^s that is really $s-1$ dimensional. So that $B(\theta_0)$ will not be positive definite. Since subsequent theory concerning the asymptotic distribution of associated random variables makes considerable use of the inverse of $B(\theta_0)$, this theory no longer applies. This paper provides a method for overcoming this difficulty.

El-Helbawy and Soliman (1983) developed the asymptotic chi-square distribution of $-2\ln\lambda$ for testing equality constraints when $B(\theta)$ is singular. The object of the present paper is to generalize the results so as to cover both cases of singular and nonsingular information matrices.

Let X be a p-dimensional random vector whose distribution function $F(\cdot,\theta)$ depends on a k-dimensional parameter $\theta=(\theta_1,\theta_2,\ldots,\theta_k)$ which are not mathematically independent but satisfy q functional relationships,

$$h_i(\theta)=0 \quad , \qquad i=1,2,\ldots,q \quad , \quad q<k \; . \qquad (1.1)$$

We assume that the possible values of X lie in a subset R of the p-dimensional space. The k×k information matrix $B(\theta)$ defined by,

$$B(\theta)=(B_{ij}(\theta)) \quad , \quad i,j=1,2,\ldots,k ,$$

$$B_{ij}(\theta)=\int_R \frac{\partial \ln f(t,\theta)}{\partial\theta_i}\frac{\partial \ln f(t,\theta)}{\partial\theta_j}\,dF(t,\theta) ,$$

is of rank $k-q$ and hence is singular.

$$H_0:\theta_0 \in \omega \subset \Omega ,$$

such that the elements of ω satisfy the additional $r-q$ constraints, $q<r<k$,

$$h_i(\theta)=0 \quad , \qquad i=q+1,\ldots,r \; . \qquad (1.2)$$

This paper concerned with deriving the limiting distribution of $-2\ln\lambda$ under the following sequence of local alternatives,

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

(a) $\quad h_i(\theta^N)=0 \qquad , i=1,2,\ldots,q , \qquad (1.3)$

these are the identifiability conditions for the k parameters.

(b) $\quad h_i(\theta^N) = \delta_i^N/\sqrt{N} \qquad , i = q+1,\ldots,r ,$ (1.4)

with;

$$\lim_{N\to\infty} \delta_i^N = \delta_i , \qquad i = q+1,\ldots,r , \tag{1.5}$$

Note that

$$\lim_{N\to\infty} h_i(\theta^N) = h_i(\theta_o) = 0 , \quad i = 1,\ldots,r , \tag{1.6}$$

An approach for getting the distribution had been given by Soliman (1994). Here we establish an alternative approach.

2. ASSUMPTIONS

The assumptions needed for the derivation of the limiting distribution of the likelihood ratio test statistic $(-2\ln\lambda)$ when $B(\theta)$ is singular are given in the Appendix. Assumptions A.1-A.12 are restatements of those given by Silvey (1959). Assumptions A.13-A.15 are restatements of those given by Davidson and Lever (1970).

3. Asymptotic Properties of Maximum Likelihood Estimators Under the Sequence of Local Alternatives

Consider the class of local alternatives $\{\theta^N\}$ given by (1.3)-(1.6). H_o and H_a are restated as follows;

$$H_o : \theta_o \in \omega = \Omega \cap [\theta : h_j(\theta) = 0, \quad j = q+1,\ldots,r] ,$$

$$H_a : \theta_o \in \Omega : [\theta : h_j(\theta) = 0 \quad , \quad j = 1,\ldots,q < r] , \quad \theta \notin \omega .$$

Since $\lim_{N\to\infty} \theta^N = \theta_o$, it follows that the information matrix,

$$B(\theta^N) = [B_{ij}(\theta^N); B_{ij}(\theta^N) = E_{\theta^N}\left(\frac{\partial \ln f(t,\theta|\theta^N)}{\partial\theta_i} \cdot \frac{\partial \ln f(t,\theta|\theta^N)}{\partial\theta_j}\right), i,j = 1,\ldots,k]$$

tends to $B(\theta_o)$ as $N \to \infty$.

Under assumptions A.1-A.15 and the sequence $\{\theta^N\}$ of local alternatives, there exists a sequence $\{\hat{\hat{\theta}}_n(\cdot,\omega)\}$ of maximum likelihood estimators (MLE's) which are restricted to the r-conditions, $h_j(\theta) = 0$, $j = 1,\ldots,r$, and will emerge as solution of the following equations

$$\frac{\partial}{\partial\theta_i}\ln L_n(x,\theta)+\sum_{j=1}^{r}\Psi_j\frac{\partial h_j(\theta)}{\partial\theta_i}=0,\qquad i=1,2,\ldots,k\,, \tag{3.1}$$

$$h_j(\theta)=0\,,\qquad j=1,2,\ldots,r\,, \tag{3.2}$$

where $\Psi_1,\ldots,\Psi_r$ are Lagrangian multipliers. There exists a sequence $\{\theta_n^*(\cdot,\Omega)\}$ of the unrestricted MLE's which maximizes the likelihood function subject to the identifiability conditions (1.1), and emerge as solutions of the following equations,

$$\frac{\partial}{\partial\theta_i}\ln L_n(x,\theta)+\sum_{j=1}^{q}\tau_j\frac{\partial h_j(\theta)}{\partial\theta_i}=0,\qquad i=1,2,\ldots,k\,, \tag{3.3}$$

$$h_j(\theta)=0\,,\qquad j=1,2,\ldots,q\,, \tag{3.4}$$

where $\tau_1,\ldots,\tau_q$ are Lagrangian multipliers and $\ln L_n\, f(x,\cdot)=\sum_{i=1}^{n}\ln f(x_i,\cdot)$.

We assumed that each of the sequences of estimators $\{\hat{\hat{\theta}}_n(\cdot,\omega)\}$ and $\{\theta_n^*(\cdot,\Omega)\}$ satisfies the set of assumptions A.1-A.12 and this will ensure the consistency of the two sets of estimates.

For brevity, we shall denote $\hat{\hat{\theta}}_n(\cdot,\omega)$ by $\hat{\hat{\theta}}_n$ and $\theta_n^*(\cdot,\Omega)$ by θ_n^*.

The Asymptotic Distribution of $\hat{\hat{\theta}}_n(\cdot,\omega)$:

Under assumptions A.1-A.12 and since $\hat{\hat{\theta}}_n\xrightarrow{p}\theta_o$ and using implications b and c of assumption A.9, then for almost all x,

$$Dn^{-1}\ln L_n(x,\hat{\hat{\theta}}_n\big|\theta^N)=Dn^{-1}\ln L_n(x,\theta_o\big|\theta^N)$$

$$+[D^2Z(\theta_o)+o(1)](\hat{\hat{\theta}}_n-\theta_o)\,, \tag{3.5}$$

where D and D^2 are as defined in the implications a and b of assumption A.9 and Z is as defined in assumption A.1.

Also, according to the continuity of the first partial derivatives of the functions h_j, for almost all x

$$H'_{\hat{\hat{\theta}}_n}=H'_{\theta_o}+o(1)\,. \tag{3.6}$$

$$h(\hat{\hat{\theta}}_n)=[H_{\theta_o}+o(1)][\hat{\hat{\theta}}_n-\theta_o]\,. \tag{3.7}$$

For almost all x, if n is sufficiently large, $\hat{\hat{\theta}}_n$, with a certain Lagrangian multiplier $\Psi_n(x)$, will satisfy the restricted likelihood equations (3.1) and (3.2). So we have

$$Dn^{-1}\ln L_n(x,\theta_0|\theta^N)+[D^2 Z(\theta_0)+o(1)][\hat{\hat{\theta}}_n-\theta_0]+H'_{\hat{\hat{\theta}}_n}\Psi_n=0\ . \quad (3.8)$$

$$[H_{\theta_0}++o(1)][\hat{\hat{\theta}}_n-\theta_0]=0\ . \quad (3.9)$$

Since, under H_0; $Z(\theta_0)$ is a maximum of Z in the set ω, then from (3.8) and (3.6) we have

$$Dn^{-1}\ln L_n(x,\theta_0|\theta^N)+[D^2 Z(\theta_0)+o(1)](\hat{\hat{\theta}}_n-\theta_0)+[H'_{\theta_0}+o(1)]\Psi_n=0. \quad (3.10)$$

Let $D^2 Z(\theta_0)=-B(\theta_0)$. Then (3.10) reduces to

$$[B(\theta_0)+o(1)](\hat{\hat{\theta}}_n-\theta_0)-[H'_{\theta_0}+o(1)]\Psi_n=Dn^{-1}\ln L_n(x,\theta_0|\theta^N). \quad (3.11)$$

Combining (3.11) and (3.9) we get

$$\begin{bmatrix} B(\theta_0)+o(1) & -H'_{\theta_0}+o(1) \\ H_{\theta_0}+o(1) & 0 \end{bmatrix}\begin{bmatrix} \hat{\hat{\theta}}_n-\theta_0 \\ \Psi_n \end{bmatrix}=\begin{bmatrix} Dn^{-1}\ln L_n(x,\theta_n|\theta^N) \\ 0 \end{bmatrix}, \quad (3.12)$$

for almost any x.

Since $\sqrt{n}H_{\theta_0}(\hat{\hat{\theta}}_n-\theta_0)\approx 0$, as shown by El-Helbawy and Soliman, (1983), it follows from (3.12) that

$$\sqrt{n}\begin{bmatrix} B(\theta_0)+H'_{1\theta_0}H_{1\theta_0} & -H'_{\theta_0} \\ H_{\theta_0} & 0 \end{bmatrix}\begin{bmatrix} \hat{\hat{\theta}}_n-\theta_0 \\ \Psi_n \end{bmatrix}=\begin{bmatrix} Dn^{-\frac{1}{2}}\ln L_n(x,\theta_0|\theta^N) \\ 0 \end{bmatrix}, (3.13)$$

Now, the results introduced by lemmas 1, 2, 3 and 5 of Davidson and Lever (1970) will be used to prove the following result.

Lemma 1

Under assumptions A.1-A.15 and the sequence $\{\theta^N\}$ of local alternatives, the sequence of vectors

$$\left\{\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta_i}\ln L_n(x,\theta_0|\theta^N),\quad i=1,\dots,k\right\}\ ,$$

converges in distribution, to the multivariate normal distribution with mean vector $B(\theta_0)\delta$ and variance matrix $B(\theta_0)$, where δ is the k×1 vector defined by $\delta=(0,\ \ ,0,\delta_{q+1},\ \ ,\delta_r,0,\ \ ,0)$.

Proof

Expanding $\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta_i}\ln L_n(x,\theta_o|\theta^N)$ by Taylor series expansion about $\theta=\theta^N$, we obtain

$$\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta_i}\ln L_n(x,\theta_o|\theta^N)=\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta_i}\ln L_n(x,\theta^N)$$

$$+\sqrt{n}\sum_{j=1}^{k}(\theta_{oi}-\theta_i^N)\left\{\frac{1}{n}\frac{\partial^2}{\partial\theta_i\,\partial\theta_j}\ln L_n(x,\tilde{\theta}^N|\theta^N)\right\}$$

$$,i=1,\ldots,k,$$

for some $\tilde{\theta}^N$ such that $\|\tilde{\theta}^N-\theta_o\|<\|\theta_o-\theta^N\|$.

Now, $\lim_{N\to\infty}\theta^N=\theta_o$ implies that $\lim_{N\to\infty}\tilde{\theta}^N=\theta_o$. It then follows from lemmas (2) and (3) of Davidson and Lever (1970) that for $i,j=1,\ldots,k$

$$p\lim_{n\to\infty}\left[\frac{1}{n}\frac{\partial^2}{\partial\theta_i\,\partial\theta_j}\ln L_n(x,\tilde{\theta}^N|\theta^N)\right]=\frac{1}{n}\frac{\partial^2}{\partial\theta_i\,\partial\theta_j}\ln L_n(x,\theta_o)$$

$$=C_{ij}(\theta,\theta_o)=-B_{ij}(\theta_o).$$

Then by Slutsky's theorem [see, for example Laha. and Rohatgi (1979)]

$$p\lim_{n\to\infty}\left[\sqrt{n}\sum_{j=1}^{k}(\theta_{oi},\theta_i^N)\{\frac{1}{n}\frac{\partial^2}{\partial\theta_i\,\partial\theta_j}\ln L_n(x,\tilde{\theta}^N|\theta^N)\}\right]=\sum_{j=1}^{k}\delta_j B_{ij}(\theta_o),$$

$$i=1,\ldots,k,$$

and since by lemma 4 of Davidson and Lever (1970), $\{\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta_i}\ln L_n(x,\theta^N), i=1,\ldots,k\}$ converges in distribution to the multivariate normal distribution $(0,B(\theta_o))$. Hence it follows that,

$$\left\{\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta_i}\ln L_n(x,\theta_o|\theta^N),\quad i=1,\ldots,k\right\},$$

converges in distribution to the multivariate normal distribution $(B(\theta_o)\delta, B(\theta_o))$.

Lemma 2

Subject to assumptions A.1-A.15 and under the sequence $\{\theta^N\}$ of local alternatives, the random vector

$$\sqrt{n}\begin{bmatrix}\hat{\hat{\theta}}_n - \theta_o \\ \Psi_n\end{bmatrix}$$

is asymptotically jointly distributed multivariate normal with mean vector $\begin{bmatrix}P_{\theta_o} B(\theta_o)\delta \\ 0\end{bmatrix}$ and variance matrix $\begin{bmatrix}P_{\theta_o} & 0 \\ 0 & S\end{bmatrix}$, where $S = -R_{\theta_o} - \begin{bmatrix}I_q & 0 \\ 0 & 0\end{bmatrix}$.

Proof

From (3.13) we have

$$\sqrt{n}\begin{bmatrix}\hat{\hat{\theta}}_n - \theta_o \\ \Psi_n\end{bmatrix} \approx \begin{bmatrix}P_{\theta_o} & Q'_{\theta_o} \\ Q_{\theta_o} & R_{\theta_o}\end{bmatrix}\begin{bmatrix}J \\ 0\end{bmatrix}, \qquad (3.14)$$

where, $J = (n^{-\frac{1}{2}} \frac{\partial}{\partial \theta_i} \ln L_n(x, \theta_o | \theta^N), i = 1, \ldots, k$ is a k-dimensional random vector.

From lemma 1, J has a limiting multivariate normal distribution with mean vector $B(\theta_o)\delta$ and variance matrix $B(\theta_o)$.

Then the $(k+r)$ dimensional random vector $\begin{bmatrix}J \\ 0\end{bmatrix}$ is asymptotically distributed according to the normal distribution with mean vector $\begin{bmatrix}B(\theta_o)\delta \\ 0\end{bmatrix}$ and variance matrix $\begin{bmatrix}B(\theta_o) & 0 \\ 0 & 0\end{bmatrix}$.

Then from (3.14),

$$\sqrt{n}\begin{bmatrix}\hat{\hat{\theta}}_n - \theta_o \\ \Psi_n\end{bmatrix}$$

is asymptotically jointly distributed multivariate normal with mean vector

$$\begin{bmatrix}P_{\theta_o} B(\theta_o)\delta \\ 0\end{bmatrix}$$

and variance-covariance matrix

$$\begin{bmatrix} P_{\theta_0} & Q'_{\theta_0} \\ Q_{\theta_0} & R_{\theta_0} \end{bmatrix} \begin{bmatrix} B_{\theta_0} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P_{\theta_0} & Q'_{\theta_0} \\ Q_{\theta_0} & R_{\theta_0} \end{bmatrix} = \begin{bmatrix} P'_{\theta_0} B_{\theta_0} P_{\theta_0} & P'_{\theta_0} B_{\theta_0} Q'_{\theta_0} \\ Q_{\theta_0} B_{\theta_0} P_{\theta_0} & Q_{\theta_0} B_{\theta_0} Q'_{\theta_0} \end{bmatrix}$$

$$= \begin{bmatrix} P_{\theta_0} & 0 \\ 0 & S \end{bmatrix}$$

where. $P'_{\theta_0} B_{\theta_0} P_{\theta_0} = P_{\theta_0}$, $P'_{\theta_0} B_{\theta_0} Q'_{\theta_0} = Q_{\theta_0} B_{\theta_0} P_{\theta_0} = 0$,

and $S = -R_{\theta_0} - \begin{bmatrix} I_q & 0 \\ 0 & 0 \end{bmatrix}$.

Corollary 1

From (3.12) and (3.14) we have

$$\sqrt{n}(\hat{\hat{\theta}}_n - \theta_0) \approx n^{-\frac{1}{2}} P_{\theta_0} D \ln L_n(x, \theta_0 | \theta^N) , \quad (3.15)$$

$$\sqrt{n}\,\psi_n \approx n^{-\frac{1}{2}} Q_{\theta_0} D \ln L_n(x, \theta_0 | \theta^N) . \quad (3.16)$$

The Asymptotic Distribution of $\theta_n^*(\cdot, \Omega)$:

In a similar manner, as used above to derive the asymptotic distribution of $\hat{\hat{\theta}}(\cdot, \omega)$, we can get

$$D n^{-1} \ln L_n(x, \theta_n^* | \theta^N) = D n^{-1} \ln L_n(x, \theta_0 | \theta^N)$$

$$+ [D^2 Z(\theta_0) + o(1)][\theta_n^* - \theta_0], \quad (3.17)$$

$$H'_{1\theta_n^*} = H'_{1\theta_0} + o(1) \quad . \quad (3.18)$$

where. $H_{1\theta_0}$ is the $q \times k$ matrix $(h_{ij}(\theta_n^*))$, $h_{ij}(\theta_n^*) = \dfrac{\partial h_i(\theta_n^*)}{\partial \theta_j}$, $i = 1, \ldots, q$, $j = 1, \ldots, k$

$$h(\theta_n^*) = [H_{1\theta_0} + o(1)][\theta_n^* - \theta_0]. \quad (3.19)$$

Hence

$$D n^{-1} \ln L_n(x, \theta_0 | \theta^N) + [D^2 Z(\theta_0) + o(1)][\theta_n^* - \theta_0] + H'_{1\theta_n^*} \tau = 0 , \quad (3.20)$$

$$[H_{1\theta_0} + o(1)][\theta_n^* - \theta_0] = 0 \quad (3.21)$$

This follows from equations (3.3), (3.4), (3.18) and (3.19). But $DZ(\theta_0)=0$ where $Z(\theta_0)$ is a maximum in Ω of $Z(\theta)$ and $D^2 Z(\theta_0) = -B(\theta_0)$. Then (3.20) reduces to

$$[B(\theta_0)+o(1)][\theta_n^* - \theta_0] - [H'_{1\theta_0} + o(1)]\tau = Dn^{-1} \ln L_n(x,\theta_0|\theta^N) + o(1). \quad (3.22)$$

Combining (3.21) and (3.22) we get

$$\begin{bmatrix} B(\theta_0)+o(1) & -H'_{1\theta_0}+o(1) \\ H_{1\theta_n^*}+o(1) & 0 \end{bmatrix} \begin{bmatrix} \theta_n^* - \theta_0 \\ \tau \end{bmatrix} = \begin{bmatrix} Dn^{-1} \ln L_n(x,\theta_0|\theta^N)+o(1) \\ 0 \end{bmatrix}. \quad (3.23)$$

Define,

$$\begin{bmatrix} U_{\theta_0} & L'_{1\theta_0} \\ L_{\theta_0} & W_{\theta_0} \end{bmatrix} = \begin{bmatrix} B(\theta_0)+H'_{1\theta_0} H_{1\theta_0} & -H'_{1\theta_0} \\ H_{1\theta_n} & 0 \end{bmatrix}^{-1}. \quad (3.24)$$

Since $\sqrt{n}\, H'_{1\theta_0}(\theta_n^* - \theta_0) \approx 0$ it follows by using (3.23) that

$$\sqrt{n} \begin{bmatrix} B(\theta_0)+H'_{1\theta_0} H_{1\theta_0} & -H'_{1\theta_0}+o(1) \\ H_{1\theta_n}+o(1) & 0 \end{bmatrix} \begin{bmatrix} \theta_n^* - \theta_0 \\ \tau \end{bmatrix} \approx \begin{bmatrix} Dn^{-\frac{1}{2}} \ln L_n(x,\theta_0|\theta^N) \\ 0 \end{bmatrix}. \quad (3.25)$$

Lemma 3

Under assumptions A.1-A.12 and the sequence $\{\theta^N\}$ of local alternatives, the random vector,

$$\sqrt{n} \begin{bmatrix} \theta_n^* - \theta_0 \\ \tau \end{bmatrix},$$

is asymptotically normally distributed with mean vector $\begin{bmatrix} U_{\theta_0} B(\theta_0)\delta \\ 0 \end{bmatrix}$ and variance matrix $\begin{bmatrix} U_{\theta_0} & 0 \\ 0 & W_{\theta_0} \end{bmatrix}$.

The proof is analogous to that given in lemma 2 and follows from (3.24), (3.25) and the result stated in lemma 1, that the vector $(Dn^{-\frac{1}{2}} \ln L_n(x,\theta_0|\theta^N))$ has a limiting normal distribution with mean vector $B(\theta_0)\delta$ and variance matrix $B(\theta_0)$.

Corollary 2

$$\sqrt{n}(\theta_n^* - \theta_o) \approx n^{-\frac{1}{2}} U_{\theta_o} D \ln L_n(x, \theta_o|\theta^N) + o_p(1) , \qquad (3.26)$$

$$\sqrt{n}\,\tau \approx n^{-\frac{1}{2}} L'_{\theta_o} D \ln L_n(x, \theta_o|\theta^N) + o_p(1) . \qquad (3.27)$$

Lemma 4

Subject to assumptions A.1-A.12 and the sequence $\{\theta^N\}$ of local alternatives

$$-2\ln\lambda = n(\hat{\hat{\theta}}_n - \theta_n^*)'[B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}](\hat{\hat{\theta}}_n - \theta_n^*) + o_p(1) .$$

Proof

From corollaries (1) and (2), $\left\|\hat{\hat{\theta}}_n - \theta_n^*\right\| = o_p(n^{-\frac{1}{2}})$.

Expanding $\ln L_n(\cdot, \hat{\hat{\theta}}_n)$ about $\theta = \theta_n^*$ by Taylor's theorem, we have,

$$\ln L_n(\cdot, \hat{\hat{\theta}}_n) = \ln L_n(\cdot, \theta_n^*) + \tfrac{1}{2}(\hat{\hat{\theta}}_n - \theta_n^*)'[D^2 \ln L_n(\cdot, \theta_n^*)](\hat{\hat{\theta}}_n - \theta_n^*) + o_p(1).$$

Also by Taylor's theorem,

$$n^{-1} D^2 \ln L_n(\cdot, \theta_n^*) = n^{-1} D^2 \ln L_n(\cdot, \theta_o|\theta^N) + o_p(1)$$

$$= -B(\theta_o) + o_p(1) ,$$

where $D^2 Z(\theta_o) = -B(\theta_o)$.

Since $nH_{\theta_o}(\hat{\hat{\theta}}_n - \theta_o) \approx 0$, and $nH_{\theta_o}(\theta_n^* - \theta_o) \approx 0$, then, $nH_{\theta_o}(\hat{\hat{\theta}}_n - \theta_n^*) \approx 0$.

Hence

$$\ln\lambda = \ln L_n(\cdot, \hat{\hat{\theta}}_n|\theta^N) - \ln L_n(\cdot, \theta_n^*|\theta^N)$$

$$= -\tfrac{1}{2} n(\hat{\hat{\theta}}_n - \theta_n^*) - [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o} + o_p(1)](\hat{\hat{\theta}}_n - \theta_n^*) + o_p(1) ,$$

But, $\left\|\hat{\hat{\theta}}_n - \theta_n^*\right\| = o_p(n^{-\frac{1}{2}})$. Then

$$-2\ln\lambda = n(\hat{\hat{\theta}}_n - \theta_n^*)'[B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}](\hat{\hat{\theta}}_n - \theta_n^*) + o_p(1) ,$$

under the sequence of local alternatives.

Lemma 5

Subject to assumptions A.1-A.12 and under the sequence $\{\theta^N\}$ of local alternatives, the likelihood ratio statistic $-2\ln\lambda$ satisfies the relation

$$-2\ln\lambda = -n\Psi_n' R_{\theta_0}^{-1}\Psi_n + n\tau' W_{\theta_0}^{-1}\tau \quad ,$$

where Ψ, τ, R and W are as defined in (3.1), (3.3), (3.14), (3.24) and lemma 4 of Davidson and Lever (1970).

Proof

We have from lemma 4,

$$-2\ln\lambda \approx n(\hat{\hat{\theta}}_n - \theta_n^*)'[B(\theta_0) + H_{1\theta_0}' H_{1\theta_0}](\hat{\hat{\theta}}_n - \theta_n^*) \quad .$$

From corollaries (1) and (2), we have

$$\sqrt{n}(\hat{\hat{\theta}}_n - \theta_n^*) \approx n^{-\frac{1}{2}}[P_{\theta_0} - U_{\theta_0}]D\ln L_n(x,\theta_0|\theta^N) \quad ,$$

where P_{θ_0} and U_{θ_0} are as defined before.

In lemma A.1 of El-Helbawy and Soliman (1983), we proved that

$$(P_{\theta_0} - U_{\theta_0})'[B(\theta_0) + H_{1\theta_0}' H_{1\theta_0}](P_{\theta_0} - U_{\theta_0}) = L_{\theta_0}' W_{\theta_0}^{-1} L_{\theta_0} - Q_{\theta_0}' R_{\theta_0}^{-1} Q_{\theta_0} \, .$$

Then under the sequence $\{\theta^N\}$ of local alternatives.

$$-2\ln\lambda \approx n^{-1}[D\ln L_n(x,\theta_0|\theta^N)]' L_{\theta_0}' W_{\theta_0}^{-1} L_{\theta_0}[D\ln L_n(x,\theta_0|\theta^N)]$$

$$- n^{-1}[D\ln L_n(x,\theta_0|\theta^N)]' Q_{\theta_0}' R_{\theta_0}^{-1} Q_{\theta_0}[D\ln L_n(x,\theta_0|\theta^N)] .$$

Since from corollaries 1 and 2

$$\sqrt{n}\Psi_n \approx n^{-\frac{1}{2}} Q_{\theta_0} D\ln L_n(x,\theta_0|\theta^N) \ ,$$

and

$$\sqrt{n}\tau \approx n^{-\frac{1}{2}} L_{\theta_0} D\ln L_n(x,\theta_0|\theta^N) \quad .$$

It follows that

$$-2\ln\lambda \approx n(H_{\theta_0}'\Psi_n)[B(\theta_0) + H_{1\theta_0}' H_{1\theta_0}](H_{\theta_0}'\Psi_n)'$$

$$-n(H_{1\theta_0}'\tau_n)[B(\theta_0) + H_{1\theta_0}' H_{1\theta_0}](H_{\theta_0}'\tau_n)'$$

where $R_{\theta_0}^{-1} = -H_{\theta_0}[B(\theta_0) + H_{1\theta_0}' H_{1\theta_0}]H_{\theta_0}'$,

and $W_{\theta_0}^{-1} = -H_{1\theta_0}[B(\theta_0) + H_{1\theta_0}' H_{1\theta_0}]H_{1\theta_0}'$.

Substituting again for $R_{\theta_0}^{-1}$ and $W_{\theta_0}^{-1}$ in the expression of $-2\ln\lambda$, we get

$$-2\ln\lambda = -n\Psi_n' R_{\theta_0}^{-1}\Psi_n + n\tau' W_{\theta_0}^{-1}\tau \quad .$$

The proof is completed.

The following theorem states the asymptotic equivalence between the likelihood ratio statistic, $-2\ln\lambda$ and the Lagrangian multipliers statistic, $n\Psi' R_{\theta_0}^{-1}\Psi$, under the sequence $\{\theta^N\}$ of local alternatives.

Theorem 1

Subject to assumptions A.1-A.12 and under the sequence $\{\theta^N\}$ of local alternatives.

$$-2\ln\lambda \approx -n\Psi' R_{\theta_0}^{-1}\Psi \quad .$$

Proof

From lemma 5 we have

$$-2\ln\lambda \approx -n\Psi' R_{\theta_0}^{-1}\Psi + n\tau' W_{\theta_0}^{-1}\tau \quad ,$$

We shall prove that $n\tau' W_{\theta_0}^{-1}\tau$ vanishes in the limit.

From (3.25), we have

$$\sqrt{n}\begin{bmatrix} B(\theta_0)+H_{1\theta_0}' H_{1\theta_0} & -H_{1\theta_0}' \\ H_{1\theta_0} & 0 \end{bmatrix}\begin{bmatrix} \theta_n^* - \theta_0 \\ \tau \end{bmatrix} \approx \begin{bmatrix} E \\ 0 \end{bmatrix} \quad , \qquad (3.28)$$

where E is a k-dimensional random vector which is asymptotically distributed, under the sequence $\{\theta^N\}$, according to a multivariate normal distribution with mean vector $B(\theta_0)\delta$ and variance matrix $B(\theta_0)$ (lemma 1).

Consider the transformation $W = VE$ where V is a non-singular matrix such that;

$$V(B(\theta_0)+H_{1\theta_0}' H_{1\theta_0})V' = I_k \qquad (3.29)$$

$$VB(\theta_0)V' = \begin{bmatrix} I_{k-q} & 0 \\ 0 & 0 \end{bmatrix} \qquad (3.30)$$

$$VH_{1\theta_0}' H_{1\theta_0} V' = \begin{bmatrix} 0 & 0 \\ 0 & I_q \end{bmatrix} \qquad (3.31)$$

where from lemma 1, W is a k-dimensional random vector, $W' = (w_1, \ldots, w_k)$ which is normally distributed with mean vector $VB(\theta_0)\delta = V'^{-1}\delta$ and variance matrix $VB(\theta_0)V' = \begin{bmatrix} I_{k-q} & 0 \\ 0 & 0 \end{bmatrix}$.

This implies that $w_1, w_2, \ldots, w_{k-q}$ are independently normally distributed $\left(\sum_{j=1}^{k-q} V_{ij}^{-1}\,\delta_j ;1\right)$ and $w_{k-q+1}, w_{k-q+2}, \ldots, w_k = 0$.

From (3.28) and the above transformation we have

$$\sqrt{n}\begin{bmatrix}(V'V)^{-1} & -H'_{1\theta_o} \\ H_{1\theta_o} & 0\end{bmatrix}\begin{bmatrix}\theta_n^* - \theta_o \\ \tau\end{bmatrix} \approx \begin{bmatrix}V^{-1}W \\ 0\end{bmatrix}. \qquad (3.32)$$

Pre-multiplying both sides of (3.32) by the matrix $\begin{bmatrix}V & 0 \\ 0 & I_q\end{bmatrix}$ we get

$$\sqrt{n}\begin{bmatrix}V'^{-1} & -VH'_{1\theta_o} \\ H_{1\theta_o} & 0\end{bmatrix}\begin{bmatrix}\theta_n^* - \theta_o \\ \tau\end{bmatrix} \approx \begin{bmatrix}W \\ 0\end{bmatrix}, \qquad (3.33)$$

and

$$W'W = n\begin{bmatrix}\theta_n^* - \theta_o \\ \tau\end{bmatrix}'\begin{bmatrix}(V'V)^{-1} & -H'_{1\theta_o} \\ H_{1\theta_o} & H_{1\theta_o}V'VH'_{1\theta_o}\end{bmatrix}\begin{bmatrix}\theta_n^* - \theta_o \\ \tau\end{bmatrix}. \qquad (3.34)$$

Since, $\sqrt{n}\,H_{1\theta_o}(\theta_n^* - \theta_o) \approx 0$ we have

$$W'W \approx n(\theta_n^* - \theta_o)'B(\theta_o)(\theta_n^* - \theta_o) + n\tau' H_{1\theta_o}V'VH'_{1\theta_o}\tau, \qquad (3.35)$$

and

$$W \approx \sqrt{n}\,V'^{-1}(\theta_n^* - \theta_o) - \sqrt{n}\,VH'_{1\theta_o}\tau\ . \qquad (3.36)$$

Pre-multiplying both sides of the last equation by $-H_{1\theta_o}V'$ we get

$$-H_{\theta_o}V'W \approx \sqrt{n}\,H_{1\theta_o}V'VH'_{\theta_o}\tau\ .$$

But,

$$W'VH'_{1\theta_o}H_{1\theta_o}V'W = W'\begin{bmatrix}0 & 0 \\ 0 & I_q\end{bmatrix}W = 0\ ,$$

where $VH'_{1\theta_o}H_{1\theta_o}V' = \begin{bmatrix}0 & 0 \\ 0 & I_q\end{bmatrix}$ and $w_{k-q+1} = w_{k-q+2} = \ldots = w_k = 0$.

Hence $H_{1\theta_o}V'W = 0$, which implies that, $-\sqrt{n}\,H_{1\theta_o}V'VH'_{1\theta_o}\tau \approx 0$, where $-H_{1\theta_o}V'W \approx \sqrt{n}\,H_{1\theta_o}V'VH'_{1\theta_o}\tau$.

Then $-n\tau' H_{1\theta_o}V'VH_{1\theta_o}\tau \approx 0$.

Since $-n\tau' H_{1\theta_0} V' V H'_{1\theta_0} \tau = n\tau' W_{1\theta_0}^{-1} \tau \approx 0$ hence, $-2\ln\lambda \approx -n\Psi' R_{\theta_0}^{-1} \Psi$. This completes the proof.

Theorem 2

Under assumptions A.1-A.12 and the sequence $\{\theta^N\}$ of local alternatives, the random variable $-n\Psi' R_{\hat{\theta}_0}^{-1} \Psi$ is asymptotically distributed as χ^2 distribution with $r-q$ degrees of freedom and non-centrality parameter

$$\alpha = \delta' H_{\hat{\theta}_0} [H_\theta \{B(\theta_0) + H'_{1\theta_0} H_{1\theta_0}\}^{-1} H'_\theta]^{-1} \dot{H}_{\hat{\theta}} \delta \ .$$

Proof

Since the matrix $B(\theta_0) + H'_{1\theta_0} H_{1\theta_0}$ is positive definite and the rank of $B(\theta_0)$ is $k-q$, there exists a non-singular matrix V satisfies (3.29)-(3.31).

Now, define a k-dimensional random vector $T = (t_1, \ldots, t_k)$ by $T = VJ$ where, $J = (n^{-\frac{1}{2}} D \ln L_n(x, \theta_0 | \theta^N))$ has a limiting normal distribution with mean vector $B(\theta_0)\delta$ and variance matrix $B(\theta_0)$. Then T is asymptotically normally distributed with mean vector distributed with mean vector $VB(\theta_0)\delta = V^{-1}\delta$ and variance matrix $VB(\theta_0)V' = \begin{bmatrix} I_{k-q} & 0 \\ 0 & 0 \end{bmatrix}$ which implies that $t_1, t_2, \ldots, t_{k-q}$ are independently normally distributed $\left(\sum_{j=1}^{k-q} V_{ij}^{-1} \delta_j ; 1\right)$ and $t_{k-q+1} = t_{k-q+2} = \ldots = t_k = 0$. From (3.13) we have

$$\sqrt{n} \begin{bmatrix} B(\theta_0) + H'_{1\theta_0} H_{1\theta_0} & -H'_{\theta_0} \\ H_\theta & 0 \end{bmatrix} \begin{bmatrix} \hat{\hat{\theta}}_n - \theta_0 \\ \Psi \end{bmatrix} \approx \begin{bmatrix} J \\ 0 \end{bmatrix} , \qquad (3.37)$$

which is equivalent to

$$\sqrt{n} \begin{bmatrix} (V'V)^{-1} & -H'_{\theta_0} \\ H_\theta & 0 \end{bmatrix} \begin{bmatrix} \hat{\hat{\theta}}_n - \theta_0 \\ \Psi \end{bmatrix} \approx \begin{bmatrix} J \\ 0 \end{bmatrix} = \begin{bmatrix} V^{-1}T \\ 0 \end{bmatrix} . \qquad (3.38)$$

From (3.38), we get

$$T'T \approx n(\hat{\hat{\theta}}_n - \theta_0)' B(\theta_0)(\hat{\hat{\theta}}_n - \theta_0) + n\Psi' H_{\theta_0} V' V H'_{\theta_0} \Psi \ ,$$

where $T'T$ is distributed as non-central χ^2 with $k-q$ degrees of freedom and non-centrality parameter,

$$\gamma = (V^{-1}\delta)' V' B(\theta_o) V' V^{-1}\delta = \delta' B(\theta_o)\delta \quad .$$

Since the rank of $n(\hat{\hat{\theta}}_n - \theta_o)' B(\theta_o)(\hat{\hat{\theta}}_n - \theta_o)$, when expressed in terms of $T'T$ is at most $k-r$ and the rank of $n\Psi' H_{\theta_o} V' V H'_{\theta_o} \Psi$ when expressed in terms of $T'T$ is at most $r-q$, then, by applying Fisher-Cochran's theorem (Rao, 1962), $n(\hat{\hat{\theta}}_n - \theta_o)' B(\theta_o)(\hat{\hat{\theta}}_n - \theta_o)$ and $n\Psi' H_{\theta_o} V' V H'_{\theta_o} \Psi$ are asymptotically independently distributed as non-central χ^2 with $k-r$ and $r-q$ degrees of freedom and non-centrality parameters ξ_1, ξ_2 respectively, which we will determine below.

To determine ξ_1, ξ_2, we formulate $n(\hat{\hat{\theta}}_n - \theta_o)' B(\theta_o)(\hat{\hat{\theta}}_n - \theta_o)$ and $n\Psi' H_{\theta_o} V' V H'_{\theta_o} \Psi$ in terms of $t_1, t_2, \ldots, t_{k-q}$ as quadratic forms as follows.

From (3.14) and (3.37) we have

$$\sqrt{n}(\hat{\hat{\theta}}_n - \theta_o) \approx P_{\theta_o} V^{-1} T \quad .$$

Then

$$\begin{aligned} n(\hat{\hat{\theta}}_n - \theta_o)' B(\theta_o)(\hat{\hat{\theta}}_n - \theta_o) &\approx (P_{\theta_o} V^{-1} T)' B(\theta_o)(P_{\theta_o} V^{-1} T) \\ &= T'(V')^{-1} P'_{\theta_o} B(\theta_o) P_{\theta_o} V^{-1} T \\ &= T'(V')^{-1} P_{\theta_o} V^{-1} T \,, \end{aligned} \tag{3.39}$$

where $P'_{\theta_o} B(\theta_o) P_{\theta_o} = P_{\theta_o}$ and $\sqrt{n}\Psi = Q_{\theta_o} J = Q_{\theta_o} V^{-1} T$.
Then

$$\sqrt{n}\Psi' = T'(V')^{-1} Q'_{\theta_o} \quad .$$

Hence

$$n\Psi' H_{\theta_o} V' V H'_{\theta_o} \Psi = T'(V')^{-1} Q'_{\theta_o} H_{\theta_o} V' V H'_{\theta_o} Q_{\theta_o} V^{-1} T \quad .$$

Since

$$\begin{aligned} Q'_{\theta_o} H_{\theta_o} &= [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}] H'_{\theta_o} R_{\theta_o} H_{\theta_o} \\ &= V' V H'_{\theta_o} R_{\theta_o} H_{\theta_o} \,, \end{aligned}$$

and

$$H'_{\theta_o} Q_{\theta_o} = [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}] P'_{\theta_o} - I_k = (V'V)^{-1} P'_{\theta_o} - I_k \quad .$$

then

$$n\Psi' H_{\theta_o} V' V H'_{\theta_o} \Psi = T'(V')^{-1} V' V H'_{\theta_o} R_{\theta_o} H_{\theta_o} V' V[(V'V)^{-1} P'_{\theta_o} - I_k] V^{-1} T$$

$$= T'(V')^{-1} V' V H'_{\theta_o} R_{\theta_o} H_{\theta_o} V' V (V' V)^{-1} P'_{\theta_o} V^{-1} T$$

$$- T' V H'_{\theta_o} R_{\theta_o} H_{\theta_o} V' V V^{-1} T$$

$$= T' V H'_{\theta_o} R_{\theta_o} H_{\theta_o} P'_{\theta_o} V^{-1} T - T' V H'_{\theta_o} R_{\theta_o} H_{\theta_o} V' T$$

$$= -T' V H'_{\theta_o} R_{\theta_o} H_{\theta_o} V' T \quad , \qquad (3.40)$$

where

$$H_{\theta_o} P'_{\theta_o} = 0 \ .$$

Let

$$(V')^{-1} P_{\theta_o} V^{-1} = A_1 \ ,$$

and

$$-V H'_{\theta_o} R_{\theta_o} H_{\theta_o} V' = A_2 \ .$$

Applying Fisher-Cochran's theorem, we obtain

$$A_1 A_2 = (V')^{-1} P_{\theta_o} H'_{\theta_o} R_{\theta_o} H_{\theta_o} V = 0 \ ,$$

which implies the independence of the terms

$$n(\hat{\hat{\theta}}_n - \theta_o)' B(\theta_o)(\hat{\hat{\theta}}_n - \theta_o) \text{ and } n\Psi' H_{\theta_o} V' V H'_{\theta_o} \Psi .$$

Let μ denotes the mean vector of T, then

$$\xi_1 = \mu' A_1 \mu = (V^{-1}\delta)' V^{-1} P_{\theta_o} (V')^{-1} V^{-1} \delta$$

$$= \delta' (V' V)^{-1} P_{\theta_o} (V' V)^{-1} \delta$$

$$= \delta' [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}]' P_{\theta_o} [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}] \delta \ ,$$

$$\xi_2 = \mu' A_2 \mu = -(V^{-1}\delta)' V H'_{\theta_o} R_{\theta_o} H_{\theta_o} V' V^{-1} \delta$$

$$= -\delta' V^{-1} V H'_{\theta_o} R_{\theta_o} H_{\theta_o} V' V^{-1} \delta$$

$$= -\delta' H'_{\theta_o} R_{\theta_o} H_{\theta_o} \delta \ .$$

Now, from Fisher-Cochran's theorem, the above is true if the following quadratic forms are satisfied

$$Q_i = T' A_i T, \ i = 1,2 \ \text{ and } \ \sum_{j=1}^{2} \mu_i^2 = \sum_{i=1}^{2} \xi_i, \ i = 1,2 \ ,$$

where

$$Q_1 = n(\hat{\hat{\theta}}_n - \theta_o)' B(\theta_o)(\hat{\hat{\theta}}_n - \theta_o) \text{ and } Q_2 = n\Psi' H_{\theta_o} V' V H'_{\theta_o} \Psi .$$

From (3.38)

$$Q_1 = n(\hat{\hat{\theta}}_n - \theta_o)' B(\theta_o)(\hat{\hat{\theta}}_n - \theta_o) = T'(V')^{-1} P_{\theta_o} V^{-1} T = T' A_1 T,$$

and from (3.40)

$$Q_2 = n\Psi' H_{\theta_o} V' V H'_{\theta_o} \Psi = -T' V H'_{\theta_o} R_{\theta_o} H_{\theta_o} T = T' A_2 T .$$

Since

$$\sum_{i=1}^{k} \mu_i^2 = \mu' \mu = (V^{-1}\delta)'(V^{-1}\delta) = \delta'(V^{-1})' V^{-1} \delta = \delta' [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}]\delta,$$

$$\sum_{i=1}^{2} \xi_i = \xi_1 + \xi_2 = \delta' [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}]' P_{\theta_o} [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}]\delta$$

$$- \delta' H'_{\theta_o} R_{\theta_o} H_{\theta_o} \delta$$

$$= \delta' [\{B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}\}' P_{\theta_o} \{B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}\}$$

$$- H'_{\theta_o} R_{\theta_o} H_{\theta_o}]\delta ,$$

but,

$$[B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}] Q'_{\theta_o} - H'_{\theta_o} R_{\theta_o} = 0 ,$$

then

$$[B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}] Q'_{\theta_o} H_{\theta_o} - H'_{\theta_o} R_{\theta_o} H_{\theta_o} = 0 ,$$

and

$$Q'_{\theta_o} H_{\theta_o} = P_{\theta_o} [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}] - I_k .$$

Then

$$H'_{\theta_o} R_{\theta_o} H_{\theta_o} = [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}]' P_{\theta_o} [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}]$$

$$-[B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}] .$$

Hence

$$\xi_1 + \xi_2 = \delta' [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}]\delta = \mu' \mu = \sum_i \mu_i^2 .$$

Therefore, asymptotically,

$$n(\hat{\hat{\theta}}_n - \theta_o)' B(\theta_o)(\hat{\hat{\theta}}_n - \theta_o) \text{ and } n\Psi' H_{\theta_o} V' V H'_{\theta_o} \Psi ,$$

are independently distributed under the sequence $\{\theta^N\}$ of local alternatives as χ^2 with $k-r$ and $r-q$ degrees of freedom and non-centrality parameters

$$\xi_1 = \delta' [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}]' P_{\theta_o} [B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}]\delta \,, \quad (3.41)$$

$$\xi_2 = -\delta' H'_{\hat{\theta}} [H_{\theta_o} \{B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}\}^{-1} H'_{\theta_o}]^{-1} H_{\hat{\theta}} \delta \,. \quad (3.42)$$

The proof could be completed by noting that in the expression of ξ_2 if we substitute $R_{\theta_o}^{-1} \sim R_{\hat{\theta}}^{-1}$ we then have

$$n\Psi' H_{\theta_o} V' V H'_{\theta_o} \Psi = -n\Psi' R_{\theta_o}^{-1} \Psi \approx -n\Psi' R_{\hat{\theta}}^{-1} \Psi,$$

under the sequence $\{\theta^N\}$ of local alternatives is asymptotically distributed as χ^2 with $r-q$ degrees of freedom and non-centrality parameter

$$\alpha = \delta' H'_{\hat{\theta}_o} [H_\theta \{B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}\}^{-1} H'_\theta]^{-1} H_{\hat{\theta}} \delta \quad ,$$

with $\delta_1 = \cdots = \delta_q = \delta_{r+1} = \cdots = \delta_k = 0$.

We conclude by the following theorem which is the main result in this article,

<u>**Theorem 3**</u>

Under the assumptions A.1-A.15 and considering the sequence $\{\theta^N\}$ of local alternatives where $\theta_i^N = \theta_{io} + \frac{\delta_{iN}}{\sqrt{N}}$ with $\lim_{N\to\infty} \delta_{iN} = \delta_i$. The likelihood ratio test statistic $-2\ln\lambda$ for testing the null hypothesis

$$H_o : \theta \in \omega = \Omega \cap [\theta : h_j(\theta) = 0 \,, \quad j = q+1,\ldots,r] \ ,$$

where $\Omega = [\theta : h_j(\theta) = 0, j = 1,\ldots,q, q < r]$, is asymptotically distributed as non-control chi-square distribution with $r-q$ degrees of freedom and non-centrality parameter

$$\alpha = \delta' H'_{\hat{\theta}_o} [H_\theta \{B(\theta_o) + H'_{1\theta_o} H_{1\theta_o}\}^{-1} H'_\theta]^{-1} H_{\hat{\theta}} \delta \quad ,$$

where $\delta = (\delta_1, \ldots, \delta_k)$ with $\delta_1 = \cdots = \delta_q = \delta_{r+1} = \cdots = \delta_k = 0$.

<u>**Proof**</u>

The proof follows from theorems 1 and 2.

APPENDIX

A.1. For every $\theta \in \Omega$

$$Z(\theta) = \int_R \ln f(t,\theta)\, dF(t,\theta_0) \quad ,$$

exists, where $f(\cdot,\theta)$ is the density function corresponding to each θ in the parameter space and θ_0 is the true value of the parameter vector.

A.2. Ω is a convex compact subset of R^k, ω is a subset of Ω defined by,

$$\omega = \Omega \cap [\theta : h_i(\theta) = 0, \qquad i = q+1,\dots,r] \ .$$

A.3. For almost all $t \in R$, $\ln f(t,\cdot)$ is continuous on Ω.

A.4. For almost all $t \in R$, and for every $\theta \in \Omega$,

$$\frac{\partial \ln f(t,\theta)}{\partial \theta_i}, \qquad i = 1,2,\dots,k \ ,$$

exists and,

$$\left| \frac{\partial \ln f(t,\theta)}{\partial \theta_i} \right| < g_l(t) \ , \qquad i = 1,2,\dots,k \ ,$$

where, $\int_R g_l(t)\, dF(t,\theta_0)$, is finite.

A.5. The functions $h_i(\theta), i = 1,2,\dots,r$ are continuous on Ω and possesses first and second order partial derivatives which are continuous on Ω.

A.6. $\theta_0 \in \omega$ and for any other point θ of ω, $F(t,\theta) \neq F(t,\theta_0)$ for at least one t.

Assumption A.2-A.6 achieved that, for each θ, the sequence $(n^{-1} \ln L_n(x,\theta))$ converges for almost all x, to $Z(\theta)$ and for large n and most x, $n^{-1} \ln L_n(x,\cdot)$ will be uniformly near Z and will attain its supremum in the set ω near the point where Z attains it supremum in ω, where

$$n^{-1} \ln L_n(x,\cdot) = \sum_{i=1}^{n} \ln f(x_i,\cdot) \ .$$

A.7. θ_0 is an interior in ω.

This assumption implies that for large n and most x, the sequence of points $\hat{\theta}_n(x,\omega)$ which are the maximum likelihood estimates of θ_0 under H_0, will be an interior point of ω and will emerge as a solution of the restricted likelihood equations when $h_i(\theta)$ are differentiable.

A.8. For almost all $t \in R$, the function $\ln f(t,\cdot)$ possesses continuous second order partial derivatives in a neighborhood of θ_o and if θ belong to this neighborhood, then;

$$\left|\frac{\partial^2 \ln f(t,\theta)}{\partial\theta_i \partial\theta_j}\right| < g_2(t), \qquad i,j = 1,2,\dots,k ,$$

where, $\int_R g_2(t)\, dF(t,\theta_o)$, is finite.

A.9. For almost all $t \in R$, the function $\ln f(t,\cdot)$ possesses third order partial derivatives in a neighborhood of θ_o, if θ belong to this neighborhood of θ_o, then;

$$\left|\frac{\partial^3 \ln f(t,\theta)}{\partial\theta_i \partial\theta_j \partial\theta_s}\right| < g_3(t), \qquad i,j,s = 1,2,\dots,k ,$$

where, $\int_R g_3(t)\, dF(t,\theta_o)$, is finite.

Assumptions A.4, A.8 and A.9 imply that;

(a) The vector $DZ(\theta)$, exists for every $\theta \in \Omega$ and the sequence of vectors, $\{Dn^{-1} \ln L_n(x,\theta)\}$ converges for almost all x to $DZ(\theta)$, where the i-th element of the $k\times 1$ vector $DZ(\theta)$ is;

$$\frac{\partial Z(\theta)}{\partial\theta_i}, \qquad i = 1,2,\dots,k ,$$

and similarly $Dn^{-1} \ln L_n(x,\theta)$ is defined.

(b) The matrix $D^2 Z(\theta_o)$ exists and the sequence $\{D^2 n^{-1} \ln L_n(x,\theta_o)\}$ of matrices converges for almost all x to $D^2 Z(\theta_o)$, where the (i, j)-th element of the $k\times k$ matrix $D^2 Z(\theta_o)$ is,

$$D_{ij} Z(\theta_o) = \frac{\partial^2 Z(\theta_o)}{\partial\theta_i \partial\theta_j}, \qquad i,j = 1,2,\dots,k ,$$

and similarly $D^2 n^{-1} \ln L_n(x,\theta_o)$ is defined.

(c) For almost all x and $i,j,s = 1,2,\dots,k$, the sequence $\left\{n^{-1} \frac{\partial^3 \ln L_n(x,\theta_o)}{\partial\theta_i \partial\theta_j \partial\theta_s}\right\}$ is uniformly bounded with respect to θ in a neighborhood of θ_o.

Each of the statements in a, b and c is almost a direct consequence of the strong law of large numbers.

A.10. The r×k matrix $H(\theta_0)=h_{ij}(\theta_0)$ is of rank r. The $h_{ij}(\theta)=\partial h_i(\theta_0)/\partial\theta_j$, i=1,...,r, j=1,...,k. The k×k matrix $B(\theta_0)$ is of rank $k-q$ where $q<r$ and there exists a q×k submatrix $H_1(\theta_0)$ of $H(\theta_0)$ such that $B(\theta)+H_1'(\theta_0)H_1(\theta_0)$ is positive definite. We assume that $H_1(\theta_0)$ is composed of the first q rows of $H(\theta_0)$.

A11. The matrix $B(\theta)$ exists in a neighborhood of θ_0 and its elements are continuous functions of θ there. Hence the matrix $B(\theta)+H_1'(\theta_0)H_1(\theta_0)$ will be positive definite in a neighborhood of θ_0. Similarly $H(\theta)$ is of rank r in a neighborhood of θ_0 and so the matrix $R(\theta)$ is defined by;

$$\begin{bmatrix} P(\theta) & Q'(\theta) \\ Q(\theta) & R(\theta) \end{bmatrix} = \begin{bmatrix} B(\theta)+H_1'(\theta)H_1(\theta) & -H'(\theta) \\ H(\theta) & 0 \end{bmatrix}^{-1},$$

exists and its elements are continuous functions of θ in a neighborhood of θ_0.

From the strong convergence of $\hat{\hat{\theta}}_n(x,\omega)$ to θ_0, it follows that;

$$R^{-1}(\hat{\hat{\theta}}) = R^{-1}(\theta)+o_p(1) \ ,$$

where $o_p(1)$ is a quantity which tends to zero with probability 1 in the limit.

A.12. $\int_R \frac{\partial^2 f(t,\theta_0)}{\partial\theta_i\,\partial\theta_j}dt = 0, \qquad i,j = 1,2,\ldots,k$

Assumptions A.1-A.12 imply that, $\hat{\hat{\theta}}_n(x,\omega)$ exists and almost converges to θ_0.

A.13. There exist positive real numbers ξ_1 and ξ_2 such that,

$$E_\theta\left\{\left|g_3(x)-E_\theta(g_3(x))\right|^{1+\xi_1}\right\} < \xi_2 < \infty \ .$$

A.14. There exist positive real numbers ν and ρ such that whenever,

$$\|\theta''-\theta'\| = \sum_{i=1}^{k}\left|\theta_i''-\theta_i'\right| < \nu, \quad \theta',\theta''\in\Omega \ ,$$

$$E_\theta\left[\left(\left.\frac{\partial^2 \ln f(t,\theta_0)}{\partial\theta_i\,\partial\theta_j}\right|_{\theta''}\right)^2\right] < \rho < \infty \qquad i,j = 1,\ldots,k \ .$$

A.15. There exist positive real numbers ξ_3 and ξ_4 such that;

$$E_\theta\left[\left|\frac{\partial \ln f(t,\theta)}{\partial \theta_i}\bigg|_\theta\right|^{2+\xi_3}\right]^3 < \xi_4 < \infty ,$$

for all $\theta \in \Omega$, $i = 1,2,\ldots,k$

Bibliography

Aitchinson, J. and Silvey, S. D. (1958). Maximum-Likelihood Estimation of Parameters Subject to Restraints, Ann. Math. Statist., 29, 813-828.

Cramer, H. (1946). Mathematical Methods of Statistics, Princeton University Press, Princton, New Jersey.

----------------- (1962). Random Variables and Probability Distributions, 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 Ratio Statistic for Testing Equality Constraints, The Egyptian Statistical Journal, Vol. 27, No. 2, Dec. 1983, 62-71.

Laha, R. G. and Rohatgi, V. K. (1979) Probability Theory, John Wiley, NewYork.

Rao, C. R. (1962). Linear Statistical Inference and its Applications, John Wiley, NewYork.

Silvey, S. D. (1959). The Lagrangian Multiplier test, Ann. Math. Statist., 34, 389-407.

Soliman, N. A. (1994). Some Restricted Estimation and Testing Hypothesis Problem with Singularity Assumption, The Egyptian Statistical Journal, Vol. 38, No. 2, 231-248.

Wald, A. (1943). Tests of Statistical Hypotheses Concerning Several Parameters when the Number of Observations is Large, Trans. Amer. Math. Soc., 54, 426-482.