ORNL-4225

Contract No. W-7405-eng-26

MATHEMATICS DIVISION

PART 1 - THE LIMITING DISTRIBUTION OF THE LIKELIHOOD RATIO STATISTIC $-2\ln\lambda_n$ UNDER A CLASS OF LOCAL ALTERNATIVES
PART 2 - MINIMUM AVERAGE RISK DECISION PROCEDURES FOR THE NONCENTRAL CHI-SQUARE DISTRIBUTION

William Edwin Lever

Submitted as a dissertation to the Graduate School of the Florida State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

MARCH 1968

OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee
operated by
UNION CARBIDE CORPORATION
for the
U.S. ATOMIC ENERGY COMMISSION

PRECEDING PAGE BLANK NOT FILMED.

ACKNOWLEDGMENTS

This dissertation was directed by Professor Richard G. Cornell and initial suggestions of his led the author to this topic. The author gratefully acknowledges the many suggestions of Dr. Cornell and the other members of his Supervisory Committee: Drs. R. A. Bradley, H. C. Griffith, S. K. Katti, and W. A. Thompson, Jr. The author also wishes to acknowledge the valuable assistance of Drs. I. R. Savage and H. Bergström.

The portion of this work which appears as Part 1 is the product of the joint work of Dr. Roger R. Davidson and the author; it also appears as a portion of Davidson's dissertation and has been issued as a Florida State University Statistics report M126.

The author wishes to acknowledge the financial assistance he received. He was a fellow in the Biometry Training program of the Department of Statistics of the Florida State University, which was funded by The National Institute of General Medical Sciences, Training Grant Number 5 T01 GM913, and he also received financial assistance from The National Aeronautics and Space Administration Grant Number NGR-10-004-029. In addition the author wishes to acknowledge a grant from The National Science Foundation to the Florida State University Computing Center whose facilities he used extensively in this research. He also wishes to point out that the work done by Dr. Davidson on Part 1 was supported by The Office of Naval Research, Contract NONR-988(08).

The author also wishes to acknowledge the support of the Mathematics Division of the Oak Ridge National Laboratory, operated by Union Carbide Corporation for the United States Atomic Energy Commission in the final preparation of this research. He also gratefully thanks Mrs. Charlotte Montgomery and Mrs. Betty Mashburn for the many hours they spent typing this dissertation.

William Edwin Lever

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

LIST OF TABLES

N 68-27633

PART 1

I. MAXIMUM LIKELIHOOD PROPERTIES

1.1 Introduction and summary

The likelihood ratio test was proposed by Neyman and Pearson [1928] as a method for testing a composite hypothesis. Subsequently, Wilks [1938] showed that, when the null hypothesis is true, the likelihood ratio statistic $-2\ln\lambda_n$, based on a sample of size n, has a limiting central chi-square distribution with ℓ degrees of freedom, where ℓ equals the number of functionally independent parameters specified by the null hypothesis. In a rather extensive paper, Wald [1943] offered a proof that the non-null distribution of $-2\ln\lambda_n$ converges uniformly in distribution to the noncentral chi-square distribution. However, Wald made several assumptions that are difficult to verify for specific utilizations of his results, for example, the assumptions that the maximum likelihood estimates are uniformly consistent and that the likelihood ratio test is uniformly consistent. Wald stated that uniform consistency of maximum likelihood estimates would be proved in a subsequent paper but this paper did not appear.

Many authors have used Wald's result, but they appear to disregard the uniform consistency assumptions. The purpose of this study is to give a proof that $-2\ln\lambda_n$ converges in distribution (not necessarily uniformly) to a noncentral chi-square distribution under local alternatives to the null hypothesis, the assumptions made being more readily verified. Consideration is limited to maximum likelihood estimates that are solutions to the likelihood equations obtained for the maximization process. This approach was also used by Chanda [1954] and subsequently by Bradley and Gart [1962].

1.2 Notation, assumptions, and preliminary lemmas

Let $f(\underline{x},\underline{\theta})$ denote a density or probability function (continuous or discrete) where $\underline{x}$ is a p-dimensional random vector with values over a region R independent of $\underline{\theta} = (\theta_1,\ldots,\theta_k)$, a parameter vector lying in a k-dimensional parameter space Ω. Let $\{\underline{x}_\alpha\}, \alpha=1,\ldots,n$, be n independent observation vectors on $\underline{x}$. For brevity the development below is given for random vectors $\underline{x}$ which are continuous. However, the entire discussion applies to $\underline{x}$ discrete if integration is replaced throughout by summation.

For convenience, the following notation will be introduced. Let $\partial \ln f(\underline{x},\underline{\theta})/\partial\theta_r$ and $\partial f(\underline{x},\underline{\theta})/\partial\theta_r$ be denoted by $\partial \ln f/\partial\theta_r$ and $\partial f/\partial\theta_r$ respectively, with similar notation being employed for second- and third-order derivatives. In addition, $\partial \ln f/\partial\theta_r|_{\underline{\theta}'}$ will denote the value of $\partial \ln f/\partial\theta_r$ at the point $\underline{\theta}'\epsilon\Omega$ with the same convention used for other functions.

The following assumptions will be made and will be designated as Assumption A.

A1. For almost all $\underline{x}\epsilon R$ and for all $\underline{\theta}\epsilon\Omega$,

$$\frac{\partial \ln f}{\partial\theta_r}, \quad \frac{\partial^2 \ln f}{\partial\theta_r\partial\theta_s} \text{ and } \frac{\partial^3 \ln f}{\partial\theta_r\partial\theta_s\partial\theta_t}$$

exist for $r,s,t = 1,\ldots,k$.

A2. For all $f(\underline{x},\underline{\theta})$ that are densities, for almost all $\underline{x}\epsilon R$ and for every $\underline{\theta}\epsilon\Omega$,

$$\left|\frac{\partial f}{\partial\theta_r}\right| < F_r(\underline{x}) \quad \text{and} \quad \left|\frac{\partial^2 f}{\partial\theta_r\partial\theta_s}\right| < F_{rs}(\underline{x}),$$

where $F_r(\underline{x})$ and $F_{rs}(\underline{x})$ are integrable over R, $r,s = 1,\ldots,k$. These assumptions permit certain interchanges of order of differentiation and integration or summation.

A3. For every $\underline{\theta}\epsilon\Omega$, the matrix $\underline{C}(\underline{\theta}) = [C_{rs}(\underline{\theta});\ r,s = 1,\ldots,k]$ with

$$C_{rs}(\underline{\theta}) = E_{\underline{\theta}}\left[\left.\frac{\partial \ln f}{\partial\theta_r}\right|_{\underline{\theta}} \left.\frac{\partial \ln f}{\partial\theta_s}\right|_{\underline{\theta}}\right]$$

is positive definite with finite determinant.

A4. For almost all $\underline{x}\epsilon R$ and for all $\underline{\theta}\epsilon\Omega$,

$$\left|\frac{\partial^3 \ln f}{\partial\theta_r\partial\theta_s\partial\theta_t}\right| < H_{rst}(\underline{x}),$$

and

$$E_{\underline{\theta}}[H_{rst}(\underline{x})] < M < \infty,$$

and there exists a positive real number $\overline{\eta}_1$, such that

$$E_{\underline{\theta}}\left[\left|H_{rst}(\underline{x}) - E_{\underline{\theta}}\{H_{rst}(\underline{x})\}\right|^{1+\overline{\eta}_1}\right] < L < \infty$$

for $r,s,t = 1,\ldots,k$, and where M and L are positive constants. These assumptions are essentially those given by Chanda [1954] which in turn are the multiparameter extension of those given by Cramér [1946, p.500]. The one exception is the last part of Assumption A4 above. Reasons for adding this assumption will be made apparent in the discussion in Lemma 1.6. The assumptions on the first and second derivatives of $f(\underline{x},\underline{\theta})$ and $\ln f(\underline{x},\underline{\theta})$ are similar to assumptions made by Wald, in the sense that they allow differentiation and integration or summation to be interchanged as was assumed by Wald.

In addition to Assumption A, the following assumption will be made and will be designated as Assumption B.

B. There exists a positive real number ν_2 such that whenever

$$\|\underline{\theta}'' - \underline{\theta}'\| \equiv \sum_{r=1}^{k} |\theta_r'' - \theta_r'| \leq \nu_2,\ \underline{\theta}',\ \underline{\theta}''\epsilon\Omega,$$

$$E_{\underline{\theta}'}\left[\left(\frac{\partial^2 \ln f}{\partial\theta_r\partial\theta_s}\bigg|_{\underline{\theta}''}\right)^2\right] < T < \infty$$

for $r,s = 1,\dots,k$, and where T is a positive constant.

Assumption B is closely related to Assumption III(b) of Wald [1943, p. 429]. Wald assumed the existence of functions $\psi_{rs}(\underline{x},\underline{\theta}''', \delta)$ and $\varphi_{rs}(\underline{x},\underline{\theta}''', \delta)$, respectively the greatest lower bound and the least upper bound of $\partial^2 \ln f/\partial\theta_r\partial\theta_s$ for any $\underline{\theta}''$ in the region, $\|\underline{\theta}'' - \underline{\theta}'''\| \leq \delta$, about $\underline{\theta}'''$, such that $E_{\underline{\theta}'}[\{\psi_{rs}(\underline{x},\underline{\theta}''', \delta)\}^2]$ and $E_{\underline{\theta}'}[\{\varphi_{rs}(\underline{x},\underline{\theta}''',\delta)\}^2]$ are bounded whenever $\|\underline{\theta}''' - \underline{\theta}'\| \leq \frac{1}{2}\nu_2$ and $\delta < \frac{1}{2}\nu_2$.

The following assumption, designated as Assumption C, is identical with Assumption V of Wald [1943, p. 429].

C. There exists a positive real number $\overline{\eta}_2$ such that

$$E\left[\left(\left|\frac{\partial \ln f}{\partial \theta_r}\right|_{\underline{\theta}}\right)^{2+\overline{\eta}_2}\right] < K < \infty, \quad r = 1,\ldots,k,$$

for all $\underline{\theta}\epsilon\Omega$, where K is a positive constant. This assumption is needed to prove that

$$Y_r^n(\underline{\theta}^*) = \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} \frac{\partial \ln f(\underline{x}_\alpha, \underline{\theta}^*)}{\partial \theta_r}, \quad r = 1,\ldots,k,$$

has a multivariate normal distribution.

An example of a class of distributions which satisfy these assumptions is the exponential family

$$f(\underline{x}|\underline{\theta}) = \text{Exp}\ \{A(\underline{\theta})B(\underline{x}) + C(\underline{x}) + D(\underline{\theta})\}$$

such that

(A) The third partial derivatives of $A(\underline{\theta})$ and $D(\underline{\theta})$ are bounded functions of $\underline{\theta}$.

(B) $E_{\underline{\theta}}\ \{|B(\underline{x})|^{2+\eta}\} < \infty$

for some $\eta > 0$.

(C) The matrix

$$\left[\frac{\partial^2 A(\underline{\theta})}{-\partial\theta_i \partial\theta_j} E_{\underline{\theta}}\{B(\underline{x})\} + \frac{\partial^2 D(\underline{\theta})}{\partial\theta_i \partial\theta_j}\ ;\ i,j = 1,\ldots,k\right]$$

is positive definite, where $\underline{\theta} = (\theta_1,\ldots,\theta_k)$.

<u>Lemma 1.1.</u> Given Assumptions A and B, there exists a positive real number S such that

$$E_{\underline{\theta}'}\left[\left(\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}''}\right)^2\right] < S < \infty, \; r = 1,\ldots k,$$

whenever $\| \underline{\theta}'' - \underline{\theta}' \| \leq \nu_2$.

<u>Proof:</u> It follows from Assumptions A and B that

$$E_{\underline{\theta}}\left[\left(\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}}\right)^2\right] = -E_{\underline{\theta}}\left[\frac{\partial^2 \ln f}{\partial \theta_r^2}\bigg|_{\underline{\theta}}\right] < T^{1/2} < \infty$$

for all $\underline{\theta} \varepsilon \Omega$. Now consider the Taylor series expansion of $\partial \ln f/\partial \theta_r \big|_{\underline{\theta}''}$ about $\underline{\theta} = \underline{\theta}'$, namely

$$\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}''} = \frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}'} + \sum_{s=1}^{k} (\theta_s'' - \theta_s') \frac{\partial^2 \ln f}{\partial \theta_r \partial \theta_s}\bigg|_{\underline{\theta}'''},$$

where $\underline{\theta}'''$ lies between $\underline{\theta}'$ and $\underline{\theta}''$. Thus

$$E_{\underline{\theta}'}\left[\left(\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}''}\right)^2\right] = E_{\underline{\theta}'}\left[\left(\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}'} + \sum_{s=1}^{k} (\theta_s'' - \theta_s') \frac{\partial^2 \ln f}{\partial \theta_r \partial \theta_s}\bigg|_{\underline{\theta}'''}\right)^2\right]$$

$$\leq E_{\underline{\theta}'}\left[\left(\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}'}\right)^2\right] + 2 \sum_{s=1}^{k} |\theta_s'' - \theta_s'| \left| E_{\underline{\theta}'}\left[\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}'} \frac{\partial^2 \ln f}{\partial \theta_r \partial \theta_s}\bigg|_{\underline{\theta}'''}\right]\right|$$

$$+ \sum_{s=1}^{k} \sum_{t=1}^{k} |\theta_s'' - \theta_s'| \, |\theta_t'' - \theta_t'| \left| E_{\underline{\theta}'}\left[\frac{\partial^2 \ln f}{\partial \theta_r \partial \theta_s}\bigg|_{\underline{\theta}'''} \frac{\partial^2 \ln f}{\partial \theta_r \partial \theta_t}\bigg|_{\underline{\theta}'''}\right]\right|.$$

It now follows from Assumption B and the application of the Cauchy-Schwartz inequality together with the first part of the proof of this lemma that

$$\left| E_{\underline{\theta}'}\left[\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}'} \frac{\partial^2 \ln f}{\partial \theta_r\, \partial \theta_s}\bigg|_{\underline{\theta}'''} \right] \right|$$

$$\leq \left\{ E_{\underline{\theta}'}\left[\left(\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}'} \right)^2 \right] E_{\underline{\theta}'}\left[\left(\frac{\partial^2 \ln f}{\partial \theta_r\, \partial \theta_s}\bigg|_{\underline{\theta}'''} \right)^2 \right] \right\}^{1/2} < T^{3/4} \quad \text{for } r,s \;\; 1,\ldots,k,$$

and

$$\left| E_{\underline{\theta}'}\left[\frac{\partial^2 \ln f}{\partial \theta_r\, \partial \theta_s}\bigg|_{\underline{\theta}'''} \frac{\partial^2 \ln f}{\partial \theta_r\, \partial \theta_t}\bigg|_{\underline{\theta}'''} \right] \right|$$

$$\leq \left\{ E_{\underline{\theta}'}\left[\left(\frac{\partial^2 \ln f}{\partial \theta_r\, \partial \theta_s}\bigg|_{\underline{\theta}'''} \right)^2 \right] E_{\underline{\theta}'}\left[\left(\frac{\partial^2 \ln f}{\partial \theta_r\, \partial \theta_t}\bigg|_{\underline{\theta}'''} \right)^2 \right] \right\}^{1/2} < T$$

for $r,s,t = 1,\ldots,k$ whenever $\| \underline{\theta}'' - \underline{\theta}' \| \leq \nu_2$. Thus

$$E_{\underline{\theta}'}\left[\left(\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}''} \right)^2 \right] < T^{1/2} + 2\nu_2 T^{3/4} + \nu_2^2 T = S < \infty$$

whenever $\| \underline{\theta}'' - \underline{\theta}' \| = \sum_{r=1}^{k} |\theta_r'' - \theta_r'| \leq \nu_2$.

<u>Lemma 1.2.</u> Let

$$A_r(\underline{\theta}, \underline{\theta}') = E_{\underline{\theta}'}\left[\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}} \right], \quad r = 1,\ldots,k, \tag{1.2.1}$$

$$B_{rs}(\underline{\theta}, \underline{\theta}') = E_{\underline{\theta}'}\left[\frac{\partial^2 \ln f}{\partial \theta_r\, \partial \theta_s}\bigg|_{\underline{\theta}} \right], \quad r,s = 1,\ldots,k. \tag{1.2.2}$$

(Note that from Assumption A2 $B_{rs}(\underline{\theta},\underline{\theta}) = -C_{rs}(\underline{\theta})$.) Then under Assumptions A and B, for $r,s = 1,\ldots,k$,

(i) for all $\underline{\theta},\underline{\theta}',\underline{\theta}''\epsilon\Omega$ where $\|\underline{\theta}-\underline{\theta}'\| < 1/2\nu_2$, and ν_2 is given by Assumption B, $A_r(\underline{\theta},\underline{\theta}'')$ and $B_{rs}(\underline{\theta},\underline{\theta}'')$ are continuous at $\underline{\theta}'' = \underline{\theta}'$,

(ii) $C_{rs}(\underline{\theta}) = -B_{rs}(\underline{\theta},\underline{\theta})$ is continuous in $\underline{\theta}$.

<u>Proof:</u> (i) Let ϵ be an arbitrary positive real number. It must be shown that there exist positive real numbers δ_1 and δ_2 such that when $\|\underline{\theta}''-\underline{\theta}'\| \leq \delta_1$, $W_r = |A_r(\underline{\theta},\underline{\theta}'') - A_r(\underline{\theta},\underline{\theta}')| < \epsilon$, and when $\|\underline{\theta}''-\underline{\theta}'\| \leq \delta_2$, $V_{rs} = |B_{rs}(\underline{\theta},\underline{\theta}'') - B_{rs}(\underline{\theta},\underline{\theta}')| < \epsilon$. Now

$$W_r = |A_r(\underline{\theta},\underline{\theta}'') - A_r(\underline{\theta},\underline{\theta}')| = \left|\int_R \left.\frac{\partial \ln f}{\partial\theta_r}\right|_{\underline{\theta}} \{f(\underline{x},\underline{\theta}'') - f(\underline{x},\underline{\theta}')\}d\underline{x}\right|.$$

Expanding $f(\underline{x},\underline{\theta}'')$ in a Taylor series expansion about $\underline{\theta}'' = \underline{\theta}'$ one obtains

$$W_r = \left|\sum_{s=1}^{k}(\theta_s''-\theta_s')\int_R \left.\frac{\partial \ln f}{\partial\theta_r}\right|_{\underline{\theta}} \left.\frac{\partial \ln f}{\partial\theta_s}\right|_{\underline{\theta}'''} f(\underline{x},\underline{\theta}''')d\underline{x}\right|$$

for some $\underline{\theta}'''$ such that $\|\underline{\theta}'''-\underline{\theta}'\| < \|\underline{\theta}''-\underline{\theta}'\|$. Now if $\|\underline{\theta}''-\underline{\theta}'\| \leq 1/2\nu_2$, $\|\underline{\theta}'''-\underline{\theta}'\| < 1/2\nu_2$ and, since $\|\underline{\theta}-\underline{\theta}'\| < 1/2\nu_2$ by assumption, then $\|\underline{\theta}'''-\underline{\theta}\| < \nu_2$. Thus it follows from the Cauchy-Schwartz inequality and Lemma 1.1 that

$$W_r \leq \sum_{s=1}^{k}|\theta_s''-\theta_s'|\left\{\int_R \left(\left.\frac{\partial \ln f}{\partial\theta_r}\right|_{\underline{\theta}}\right)^2 f(\underline{x},\underline{\theta}''')d\underline{x}\int_R \left(\left.\frac{\partial \ln f}{\partial\theta_s}\right|_{\underline{\theta}'''}\right)^2 f(\underline{x},\underline{\theta}''')d\underline{x}\right\}^{1/2}$$

$$< S^{1/2}\,T^{1/2}\,\|\underline{\theta}''-\underline{\theta}'\|.$$

Thus if $\delta_1 = \min(\epsilon S^{-1/2}T^{-1/2}, 1/2\nu_2)$, then $W_r < \epsilon$ whenever $\|\underline{\theta}''-\underline{\theta}'\| \leq \delta_1$. Similarly,

$$V_{rs} = |B_{rs}(\underline{\theta},\underline{\theta}'') - B_{rs}(\underline{\theta},\underline{\theta}') = \left| \int_R \frac{\partial^2 \ln f}{\partial\theta_r \, \partial\theta_s}\bigg|_{\underline{\theta}} \{f(\underline{x},\underline{\theta}'')-f(\underline{x},\underline{\theta}')\}\, d\underline{x}\right|$$

$$= \left| \sum_{t=1}^{k} (\theta_t''-\theta_t') \int_R \frac{\partial^2 \ln f}{\partial\theta_r \, \partial\theta_s}\bigg|_{\underline{\theta}} \frac{\partial \ln f}{\partial\theta_t}\bigg|_{\underline{\theta}'''} f(\underline{x},\underline{\theta}''')\, d\underline{x}\right|$$

for $\underline{\theta}'''$ such that $\|\underline{\theta}'''-\underline{\theta}'\| < \|\underline{\theta}''-\underline{\theta}'\|$. Now if $\|\underline{\theta}''-\underline{\theta}'\| \leq 1/2\nu_2$ then $\|\underline{\theta}'''-\underline{\theta}\| \leq \nu_2$, so that it follows from the Cauchy-Schwartz inequality, Assumption B and Lemma 1.1, that

$$V_{rs} \leq \sum_{t=1}^{k} |\theta_t''-\theta_t'| \left\{ \int_R \left(\frac{\partial^2 \ln f}{\partial\theta_r \, \partial\theta_s}\bigg|_{\underline{\theta}}\right)^2 f(\underline{x},\underline{\theta}''')\, d\underline{x} \right.$$

$$\left. \int_R \left(\frac{\partial \ln f}{\partial\theta_t}\bigg|_{\underline{\theta}'''}\right)^2 f(\underline{x},\underline{\theta}''')\, d\underline{x}\right\}^{1/2} < T^{3/4} \|\underline{\theta}''-\underline{\theta}'\| .$$

Thus if $\delta_2 = \min(\epsilon T^{-3/4}, 1/2\nu_2)$, then $V_{rs} < \epsilon$ whenever $\|\underline{\theta}''-\underline{\theta}'\| \leq \delta_2$.

(ii) Let ϵ be an arbitrary positive real number. It must be shown that there exists a positive real number δ such that when $\|\underline{\theta}'-\underline{\theta}\| \leq \delta$, $U_{rs} = |C_{rs}(\underline{\theta}')-C_{rs}(\underline{\theta})| = |B_{rs}(\underline{\theta}',\underline{\theta}')-B_{rs}(\underline{\theta},\underline{\theta})| < \epsilon$. By the triangular inequality

$$U_{rs} \leq |B_{rs}(\underline{\theta}',\underline{\theta}')-B_{rs}(\underline{\theta},\underline{\theta}')| + |B_{rs}(\underline{\theta},\underline{\theta}')-B_{rs}(\underline{\theta},\underline{\theta})| .$$

From Part (i) of Lemma 1.2 there exists a δ_2 such that the second term on the right-hand side is less than $\epsilon/2$ whenever $\|\underline{\theta}'-\underline{\theta}\| \leq \delta_2$.

Now use of a Taylor series expansion of $\partial^2 \ln f/\partial\theta_r\,\partial\theta_s|_{\underline{\theta}'}$ about $\underline{\theta}' = \underline{\theta}$, leads to

$$|B_{rs}(\underline{\theta}',\underline{\theta}')-B_{rs}(\underline{\theta},\underline{\theta}')| = \left|\sum_{t=1}^{k}(\theta_t'-\theta_t)\int_R \frac{\partial^3 \ln f}{\partial\theta_r\,\partial\theta_s\,\partial\theta_t}\Big|_{\underline{\theta}''} f(\underline{x},\underline{\theta}')d\underline{x}\right|$$

for some $\underline{\theta}''$ such that $\|\underline{\theta}''-\underline{\theta}\| < \|\underline{\theta}'-\underline{\theta}\|$. Then by Assumption A4

$$|B_{rs}(\underline{\theta}',\underline{\theta}')-B_{rs}(\underline{\theta},\underline{\theta}')| < \sum_{t=1}^{k}|\theta_t'-\theta_t| E_{\underline{\theta}'}\left[\left|\frac{\partial^3 \ln f}{\partial\theta_r\,\partial\theta_s\,\partial\theta_t}\Big|_{\underline{\theta}''}\right|\right]$$

$$< M\|\underline{\theta}'-\underline{\theta}\| .$$

Let $\delta_3 = 1/2\ \epsilon M^{-1}$. Thus if $\delta = \min(\delta_2,\delta_3)$, then $U_{rs} < \epsilon$ whenever $\|\underline{\theta}'-\underline{\theta}\| \leq \delta$.

1.3 Asymptotic properties of null and non-null maximum likelihood estimators

Consider the test of the composite hypothesis,

$$H_o:\ \underline{\theta} = \underline{\theta}^{\omega} = ({}_1\underline{\theta}^o, {}_2\underline{\theta})$$

where $\underline{\theta}^{\omega}\epsilon\Omega$ and where ${}_1\underline{\theta}^o = (\theta_1^o,\ldots,\theta_\ell^o)$ is specified while ${}_2\underline{\theta} = (\theta_{\ell+1},\ldots,\theta_k)$ remains unspecified. Attention is restricted to the following class of local alternatives: $\{\underline{\theta}^n\}$, a sequence of "true" values of $\underline{\theta}$ such that $\underline{\theta}^n = ({}_1\underline{\theta}^n, {}_2\underline{\theta}^*)$ where $\theta_i^n = \theta_i^o + \delta_{in}/\sqrt{n}$ with $\lim_{n\to\infty} \delta_{in} = \delta_i$, $i = 1,\ldots,\ell$, and where ${}_2\underline{\theta}^*$ is the vector of true values of the nuisance parameters ${}_2\underline{\theta}$. Setting $\underline{\theta}^* = ({}_1\underline{\theta}^o, {}_2\underline{\theta}^*)$ one sees that $\lim_{n\to\infty} \underline{\theta}^n = \underline{\theta}^*$.

Henceforth the notation $\underline{a} = ({}_1\underline{a}, {}_2\underline{a})$ indicates a k-component vector whose component-vectors ${}_1\underline{a}$ and ${}_2\underline{a}$ have ℓ and $(k-\ell)$ components respectively and a matrix

$$\underline{A} = \begin{bmatrix} \underline{A}_{11} & \underline{A}_{12} \\ \underline{A}_{21} & \underline{A}_{22} \end{bmatrix}$$

indicates a $k \times k$ matrix whose principal submatrices $\underline{A}_{11}$ and $\underline{A}_{22}$ are $\ell \times \ell$ and $(k-\ell) \times (k-\ell)$ respectively. In addition, if $\underline{B} = \underline{A}^{-1}$ has similar submatrices $\underline{B}_{ii}$, write $\underline{A}_{ii}^{-1} = \overline{\underline{B}}_{ii}$ and $\underline{B}_{ii}^{-1} = \overline{\underline{A}}_{ii}$ in order to distinguish between submatrices of inverse matrices and inverses of submatrices, $i = 1,2$.

Denote the joint likelihood function by

$$\varphi(\underline{x},\underline{\theta}) = \prod_{\alpha=1}^{n} f(\underline{x}_\alpha,\underline{\theta}).$$

Then the likelihood equations under alternative and null hypotheses respectively are the following:

$$\frac{\partial \ln\varphi(\underline{x},\underline{\theta})}{\partial\theta_r} = 0, \; r = 1,\ldots,k, \tag{1.3.1}$$

$$\frac{\partial \ln\varphi(\underline{x},\underline{\theta}^{\omega})}{\partial\theta_r} = 0, \; r = \ell + 1,\ldots,k. \tag{1.3.2}$$

The following notation is introduced for use in subsequent sections.

$$L^n_r(\underline{\theta}) = \frac{1}{n}\frac{\partial \ln\varphi(\underline{x},\underline{\theta})}{\partial\theta_r} = \frac{1}{n}\sum_{\alpha=1}^{n}\frac{\partial \ln f(\underline{x}_\alpha,\underline{\theta})}{\partial\theta_r}, \qquad r=1,\ldots,k, \quad (1.3.3)$$

$$L^n_{rs}(\underline{\theta}) = \frac{1}{n}\frac{\partial^2 \ln\varphi(\underline{x},\underline{\theta})}{\partial\theta_r\,\partial\theta_s} = \frac{1}{n}\sum_{\alpha=1}^{n}\frac{\partial^2 \ln f(\underline{x}_\alpha,\underline{\theta})}{\partial\theta_r\,\partial\theta_s}, \quad r,s=1,\ldots,k, \quad (1.3.4)$$

$$L^n_{rst}(\underline{\theta}) = \frac{1}{n}\frac{\partial^3 \ln\varphi(\underline{x},\underline{\theta})}{\partial\theta_r\,\partial\theta_s\,\partial\theta_t} = \frac{1}{n}\sum_{\alpha=1}^{n}\frac{\partial^3 \ln f(\underline{x}_\alpha,\underline{\theta})}{\partial\theta_r\,\partial\theta_s\,\partial\theta_t}, r,s,t=1,\ldots,k. \quad (1.3.5)$$

Lemma 1.3. Under Assumptions A and B, for the given sequence $\{\underline{\theta}^n\}$ and any ϵ, $\eta > 0$, the following hold for all $\underline{\theta}$ such that $\|\underline{\theta}-\underline{\theta}^*\| \leq 1/2\nu_2$:

(i) $P_1 = P[|L^n_r(\underline{\theta}) - A_r(\underline{\theta},\underline{\theta}^*)|<\eta, \text{ for all } r=1,\ldots,k|\underline{\theta}^n]>1-\epsilon,$

(ii) $P_2 = P[|L^n_{rs}(\underline{\theta}) - B_{rs}(\underline{\theta},\underline{\theta}^*)|<\eta,$

for all $r,s=1,\ldots,k|\underline{\theta}^n] > 1-\epsilon$, for $n > N(\epsilon,n)$.

Proof: For the given sequence $\{\underline{\theta}^n\}$, one can find for each $\alpha > 0$ a positive integer $N(\alpha)$ such that $\|\underline{\theta}^n-\underline{\theta}^*\| \leq \alpha$ whenever $n > N(\alpha)$.

(i) By the general form of the Tchebyshev inequality, the triangular inequality, the definitions of $A_r(\underline{\theta},\underline{\theta}^n)$ and $L^n_r(\underline{\theta})$ given in equations (1.2.1) and (1.3.3) respectively, and for any $\eta > 0$

$$P_1 \geq 1 - \sum_{r=1}^{k} P[|L_r^n(\underline{\theta})-A_r(\underline{\theta},\underline{\theta}^*)| > \eta|\underline{\theta}^n]$$

$$\geq 1 - \sum_{r=1}^{k} \frac{E_{\underline{\theta}^n}[\{L_r^n(\underline{\theta})-A_r(\underline{\theta},\underline{\theta}^n)\}^2] + [A_r(\underline{\theta},\underline{\theta}^n)-A_r(\underline{\theta},\underline{\theta}^*)]^2}{\eta^2} .$$

If $\|\underline{\theta}^n-\underline{\theta}^*\| \leq 1/2\nu_2$, then $\|\underline{\theta}^n-\underline{\theta}\| \leq \nu_2$. Thus by Lemma 1.1

$$E_{\underline{\theta}^n}[\{L_r^n(\underline{\theta})-A_r(\underline{\theta},\underline{\theta}^n)\}^2] = \frac{1}{n} E_{\underline{\theta}^n}\left[\left\{\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}} - A_r(\underline{\theta},\underline{\theta}^n)\right\}^2\right]$$

$$\leq \frac{1}{n} E_{\underline{\theta}^n}\left[\left(\frac{\partial \ln f}{\partial \theta_r}\bigg|_{\underline{\theta}}\right)^2\right] < \frac{S}{n}$$

for $n > N(1/2\nu_2)$, and by Lemma 1.2 there exists a positive real number δ_1 such that for $\|\underline{\theta}^n-\underline{\theta}^*\| \leq \delta_1$, $[A_r(\underline{\theta},\underline{\theta}^n)-A_r(\underline{\theta},\underline{\theta}^*)]^2 < 1/2\ \epsilon\eta^2/k$. Thus, if $n > \max[2Sk/\epsilon\eta^2, kN(1/2\nu_2), N(\delta_1)]$, then $P[|L_r^n(\underline{\theta}) - A_r(\underline{\theta},\underline{\theta}^*)| > \eta|\underline{\theta}^n] < \epsilon$ for any small $\epsilon > 0$ and hence Part (i) of Lemma 1.3 follows.

(ii) By the general form of the Tchebyshev inequality, the triangular inequality, the definition of $L_{rs}^n(\underline{\theta})$ and $B_{rs}(\underline{\theta},\underline{\theta}^n)$ given in equations (1.3.4) and (1.2.2) and for any $\eta > 0$,

$$P_2 \geq 1 - \sum_{r-1}^{k} \sum_{s=1}^{k} P[|L_{rs}^n(\underline{\theta})-B_{rs}(\underline{\theta},\underline{\theta}^*)| > \eta|\underline{\theta}^n]$$

$$> 1 - \sum_{r=1}^{k} \sum_{s=1}^{k} \frac{E_{\underline{\theta}^n}[\{L_{rs}^n(\underline{\theta})-B_{rs}(\underline{\theta},\underline{\theta}^n)\}^2] + [B_{rs}(\underline{\theta},\underline{\theta}^n)-B_{rs}(\underline{\theta},\underline{\theta}^*)]^2}{\eta^2} .$$

Now if $\|\underline{\theta}^n-\underline{\theta}^*\| \leq 1/2\nu_2$, then $\|\underline{\theta}^n-\underline{\theta}\| \leq \nu_2$. Thus, by Assumption B,

$$E_{\underline{\theta}^n}[\{L^n_{rs}(\underline{\theta})-B_{rs}(\underline{\theta},\underline{\theta}^n)\}^2] = \frac{1}{n} E_{\underline{\theta}^n}\left[\left\{\frac{\partial^2 \ln f}{\partial\theta_r\ \partial\theta_s}\bigg|_{\underline{\theta}} - B_{rs}(\underline{\theta},\underline{\theta}^n)\right\}^2\right]$$

$$\leq \frac{1}{n} E_{\underline{\theta}^n}\left[\left(\frac{\partial^2 \ln f}{\partial\theta_r\ \partial\theta_s}\bigg|_{\underline{\theta}}\right)^2\right] < \frac{T}{n}$$

for $n > N(1/2\nu_2)$, and by Lemma 1.2 there exists a positive real number δ_2 such that for $\|\underline{\theta}^n-\underline{\theta}^*\| \leq \delta_2$, $[B_{rs}(\underline{\theta},\underline{\theta}^n)-B_{rs}(\underline{\theta},\underline{\theta}^*)]^2 < 1/2\epsilon\eta^2/k^2$. Thus if $n > \max[2Tk^2/\epsilon\eta^2, k^2N(1/2\nu_2), N(\delta_2)]$, then

$$P_2 \geq 1 - \sum_{r=1}^{k}\sum_{s=1}^{k} P[|L^n_{rs}(\underline{\theta})-B_{rs}(\underline{\theta},\underline{\theta}^*)| > \eta|\underline{\theta}^n] > 1 - \epsilon$$

for any small $\epsilon > 0$ and hence Part (ii) of Lemma 1.3 is proved.

The following result is a restatement of Lemma 2 of Aitchison and Silvey [1958, p. 819] and is thus given without proof.

Lemma 1.4. If $\underline{g}$ is a continuous function mapping R^k into itself with the property that, for every $\underline{\theta}$ such that $\|\underline{\theta}-\underline{\theta}^*\| = \delta$, $\delta > 0$, $\sum_{r=1}^{k} g_r(\underline{\theta})(\theta_r-\theta_r^*) < 0$, then there exists a point $\hat{\underline{\theta}}$ such that $\|\hat{\underline{\theta}}-\underline{\theta}^*\| < \delta$ for which $\underline{g}(\hat{\underline{\theta}}) = \underline{0}$.

Theorem 1.1. (The consistency of the two estimators $\hat{\underline{\theta}}, \hat{\hat{\underline{\theta}}}$.) Given Assumptions A and B and the given sequence $\{\underline{\theta}^n\}$ of local alternatives,

(i) there exists a sequence $\{\hat{\underline{\theta}}^n\}$ of solutions to the likelihood equations (1.3.1) which converge in probability to $\underline{\theta}^*$,

(ii) there exists a sequence $\{\hat{\hat{\underline{\theta}}}^n = ({}_1\underline{\theta}^o, {}_2\hat{\hat{\underline{\theta}}})\}$ of solutions to the likelihood equations (1.3.2) which converge in probability to $\underline{\theta}^*$.

<u>Proof:</u> (i) Consider the likelihood equations

$$L_r^n(\underline{\theta}) = 0, \ r = 1,\dots,k,$$

obtained from (1.3.1) by dividing through by n and noting the definition of $L_r^n(\underline{\theta})$ given in (1.3.3). It follows from Lemma 1.3 that given η and ϵ, arbitrary positive constants, there exists a positive integer $N(\eta,\epsilon)$ such that, for all $\underline{\theta}$ such that $\|\underline{\theta}-\underline{\theta}^*\| < 1/2\ \nu_2$,

$$P[|L_r^n(\underline{\theta})-A_r(\underline{\theta},\underline{\theta}^*)| < \eta, \text{ for all } r = 1,\dots,k|\underline{\theta}^n] > 1 - \epsilon$$

for $n > N(\eta,\epsilon)$. Thus with probability greater than $1 - \epsilon$ the following hold for $n > N(\eta,\epsilon)$ given $\underline{\theta}^n$:

$$|L_r^n(\underline{\theta})-A_r(\underline{\theta},\underline{\theta}^*)| < \eta, \quad r = 1,\dots,k,$$

$$\{L_r^n(\underline{\theta})-A_r(\underline{\theta},\underline{\theta}^*)\}(\theta_r-\theta_r^*) < \eta|\theta_r-\theta_r^*|, \ r = 1,\dots,k,$$

and

$$\sum_{r=1}^k \{L_r^n(\underline{\theta})-A_r(\underline{\theta},\underline{\theta}^*)\}(\theta_r-\theta_r^*) < \eta \sum_{r=1}^k |\theta_r-\theta_r^*| = \eta\|\underline{\theta}-\underline{\theta}^*\| .$$

Now, for each $\underline{\theta}$ such that $\|\underline{\theta}-\underline{\theta}^*\| = \delta < 1/2\ \nu_2$, one has

$$\sum_{r=1}^k L_r^n(\underline{\theta})(\theta_r-\theta_r^*) < \sum_{r=1}^k A_r(\underline{\theta},\underline{\theta}^*)(\theta_r-\theta_r^*) + \eta\delta.$$

By expanding $\partial \ln f(\underline{x},\underline{\theta})/\partial\theta_r$ in a Taylor series about $\underline{\theta} = \underline{\theta}^*$ in the expression for $A_r(\underline{\theta},\underline{\theta}^*)$, noting that $A_r(\underline{\theta}^*,\underline{\theta}^*) = 0$, multiplying by $(\theta_r-\theta_r^*)$ and summing over r, one may obtain

$$\sum_{r=1}^{k} A_r(\underline{\theta},\underline{\theta}^*)(\theta_r-\theta_r^*) = \sum_{r=1}^{k}\sum_{s=1}^{k}(\theta_r-\theta_r^*)(\theta_s-\theta_s^*)B_{rs}(\underline{\theta}^*,\underline{\theta}^*)$$

$$+\frac{1}{2}\sum_{r=1}^{k}\sum_{s=1}^{k}\sum_{t=1}^{k}(\theta_r-\theta_r^*)(\theta_s-\theta_s^*)(\theta_t-\theta_t^*)E_{\underline{\theta}^*}[L_{rst}^n(\underline{\theta}')]$$

for some $\underline{\theta}'$ such that $\| \underline{\theta}'-\underline{\theta}^*\| < \| \underline{\theta}-\underline{\theta}^*\|$. If $\| \underline{\theta}-\underline{\theta}^*\| = \delta \leq \nu_1$, then, by Assumption A4,

$$\sum_{r=1}^{k} A_r(\underline{\theta},\underline{\theta}^*)(\theta_r-\theta_r^*) < \sum_{r=1}^{k}\sum_{s=1}^{k}(\theta_r-\theta_r^*)(\theta_s-\theta_s^*)B_{rs}(\underline{\theta}^*,\underline{\theta}^*) + 1/2\delta^3 M .$$

In addition, from Assumption A2, the matrix $[B_{rs}(\underline{\theta}^*,\underline{\theta}^*)] = [-C_{rs}(\underline{\theta}^*)]$ which is negative definite by Assumption A3 so that there exists a $\beta > 0$ (namely the characteristic root of the matrix $[B_{rs}(\underline{\theta}^*,\underline{\theta}^*)]$ which is smallest in absolute value) such that

$$\sum_{r=1}^{k}\sum_{s=1}^{k}(\theta_r-\theta_r^*)(\theta_s-\theta_s^*)B_{rs}(\underline{\theta}^*,\underline{\theta}^*) < -\beta\delta^2.$$

Thus for $\| \underline{\theta}-\underline{\theta}^*\| = \delta \leq \min(\nu_1, 1/2\nu_2)$ one has

$$\sum_{r=1}^{k} L_r^n(\underline{\theta})(\theta_r-\theta_r^*) < -\beta\delta^2 + 1/2\delta^3 M + \eta\delta.$$

Thus for arbitrary $\eta \leq 1/2\delta^2 M$

$$\sum_{r=1}^{k} L_r^n(\underline{\theta})(\theta_r-\theta_r^*) < -\beta\delta^2 + \delta^3 M \leq 0$$

if $\delta \leq \min(\beta/M, \nu_1, 1/2\nu_2)$. One may now apply Lemma 1.4 to conclude that

$$P[L_r^n(\underline{\theta})] = 0, \text{ for all } r = 1,\ldots,k \text{ for some } \underline{\hat{\theta}}^n \text{ such that}$$

$$\| \underline{\hat{\theta}}^n-\underline{\theta}^*\| < \delta | \underline{\theta}^n] > 1 - \epsilon \text{ for } n > N(\eta,\epsilon).$$

Thus there exists a sequence of roots $\{\underline{\hat{\theta}}^n\}$ to the likelihood equation (1.3.1) with probability greater than $1 - \epsilon$, for n sufficiently large, in the region $\| \underline{\hat{\theta}}^n-\underline{\theta}^*\| < \delta$. Since ϵ and δ are arbitrary and may be taken small, $\{\underline{\hat{\theta}}^n\}$ converges in probability to $\underline{\theta}^*$.

It remains to show that the sequence $\{\underline{\hat{\theta}}^n\}$ of estimators are maximum likelihood estimators. By Lemma 1.3 there exists an $N(\epsilon,\eta)$ such that whenever $\| \underline{\theta}-\underline{\theta}^*\| < 1/2\nu_2$,

$$P[|L_{rs}^n(\underline{\theta})-B_{rs}(\underline{\theta},\underline{\theta}^*)| < \eta \text{ for all } r,s = 1,\ldots,k|\underline{\theta}^n] > 1-\epsilon$$

for $n > N(\epsilon,\eta)$. In addition, it follows from the proof of part (ii) of Lemma 1.2 that $B_{rs}(\underline{\theta},\underline{\theta}^*)$ is continuous at $\underline{\theta} = \underline{\theta}^*$, $r,s = 1,\ldots,k$. Since the matrix $[B_{rs}(\underline{\theta}^*,\underline{\theta}^*);\ r,s = 1,\ldots,k] = [-C_{rs}(\underline{\theta}^*);\ r,s = 1,\ldots,k]$ is negative definite by Assumption A3, one can find a ν_3 such that the matrix $[B_{rs}(\underline{\theta},\underline{\theta}^*);\ r,s = 1,\ldots,k]$ is negative definite for all $\underline{\theta}$ such that $\| \underline{\theta}-\underline{\theta}^*\| < \nu_3 \leq 1/2\nu_2$. Thus, one can find an N_ϵ such that, whenever $\|\underline{\theta}-\underline{\theta}^*\| < \nu_3$,

$$P[\text{the matrix } [L^n_{rs}(\underline{\theta});\ r,s = 1,\ldots,k] \text{ is negative definite}] > 1-\epsilon$$

for $n > N_\epsilon$, so that with arbitrarily large probability $\underline{\hat{\theta}}^n$ yields a relative maximum and thus is a maximum likelihood estimator.

(ii) A parallel argument suffices to show that there exists a sequence of roots $\{\underline{\hat{\hat{\theta}}}^n\}$ to the likelihood equations

$$L_r(\underline{\theta}^\omega) = 0,\ r = \ell,\ldots,k$$

which converge in probability to $\underline{\theta}^*$, where it is recalled that

$${}_1\underline{\hat{\hat{\theta}}}^n = {}_1\underline{\theta}^\omega = {}_1\underline{\theta}^* = {}_1\underline{\theta}^o \text{ by definition.}$$

Lemma 1.5. Under Assumptions A, B, and C, and given the sequence $\{\underline{\theta}^n\}$, the vector $\underline{Y}^n(\underline{\theta}^n) = [Y^n_r(\underline{\theta}^n);\ r = 1,\ldots,k] = [\sqrt{n}\ L^n_r(\underline{\theta}^n);\ r = 1,\ldots,k]$ has a cumulative distribution function (c.d.f.) designated by $\mathcal{F}^n(\underline{y})$ such that

$$\lim_{n\to\infty} \mathcal{F}^n(\underline{y}) = \eta(\underline{y};\underline{\theta},\underline{C})$$

where $\eta(\underline{y};\underline{\theta},\underline{C})$ is the multivariate normal c.d.f. with mean vector $\underline{0}$ and dispersion matrix $\underline{C} = [C_{rs}(\underline{\theta}^*);\ r,s = 1,\ldots,k]$.

Proof: In the proof given here it will be demonstrated that Assumptions A, B, and C suffice for the multivariate limit theorem given by Bergström [1963, pp. 320-321]. A special case of this theorem is now restated in a form which is more suitable for the present context:

Let $\{\underline{U}^{\alpha n}\}$ and $\{\underline{V}^{\alpha n}\}$, $\alpha = 1,\ldots,n$ be sequences of k-dimensional random variables having c.d.f.'s $\{F^n\}$ and $\{G^n\}$ respectively and zero expectations, $n = 1,2,\ldots$. Then the n-fold convolutions $\{F^{n^{*n}}\}$ and $\{G^{n^{*n}}\}$, the distributions of $\sum_{\alpha=1}^{n} \underline{U}^{\alpha n}$ and $\sum_{\alpha=1}^{n} \underline{V}^{\alpha n}$ respectively, converge at all points of continuity to limits and these limits are equal if and only if the following sets of limits exist and satisfy the conditions below.

(i) For each $\underline{\eta} = \{\eta_1 \ldots,\eta_k\}$ where $\eta_r > 0$, $r = 1,\ldots,k$

$$\lim_{n\to\infty} n\, P[|U_r^n| > \eta_r;\ \text{for all } r = 1,\ldots,k]$$

$$= \lim_{n\to\infty} n\, P[|V_r^n| > \eta_r;\ \text{for all } r = 1,\ldots,k]$$

(ii) $$\lim_{n\to\infty} n \int_{|u_r|\le\gamma} u_r dF^n(\underline{u}) = \lim_{n\to\infty} n \int_{|v_r|\le\gamma} v_r dG^n(\underline{v}),$$

(iii) $$\lim_{n\to\infty} n \int_{|u_r|,|u_s|\le\gamma} u_r u_s dF^n(\underline{u}) = \lim_{n\to\infty} n \int_{|v_r|,|v_s|\le\gamma} v_r v_s dG^n(\underline{v}),$$

for $r = 1,\ldots,k$ and $r,s = 1,\ldots,k$ respectively.

The following correspondences will be made in the application of this theorem. Let

$$U_r^{\alpha n} = n^{-1/2} \left.\frac{\partial \ln f}{\partial \theta_r}\right|_{\underline{\theta}^n},\quad r = 1,\ldots,k,\ \alpha = 1,\ldots,n.$$

Then $Y_r^n(\underline{\theta}^n) = \sum_{\alpha=1}^{n} U_r^{\alpha n} = \sqrt{n}\ L_r^n(\underline{\theta}^n)$, $r = 1,\ldots,k$. Suppose that $\left.\frac{\partial \ln f}{\partial \theta_r}\right|_{\underline{\theta}^n}$, $r = 1,\ldots,k$, have joint c.d.f. H^n; it is known that

$$E_{\underline{\theta}^n}\left(\left.\frac{\partial \ln f}{\partial \theta_r}\right|_{\underline{\theta}^n}\right) = 0,\ r = 1,\ldots,k,$$ and that the variance-covariance

matrix of the k random variables at $\underline{\theta} = \underline{\theta}^n$ is $\underline{C}(\underline{\theta}^n)$. The c.d.f. of $\underline{U}^{\alpha n} = [U_r^{\alpha n};\ r = 1,\ldots,k]$ is then $F^n(\underline{u}) = H^n(\sqrt{n}\,\underline{u})$. In addition, let $\underline{V}^{\alpha n}$, $\alpha = 1,\ldots,n$, have the k-element multivariate normal c.d.f. with mean vector zero and variance-covariance matrix $n^{-1}\underline{C}(\underline{\theta}^n)$, and let $\underline{W}^n = \sum_{\alpha=1}^{n} \underline{V}^{\alpha n}$.

It will now be shown that Assumptions A, B, and C are sufficient for conditions (i), (ii), and (iii). Consider first the sequence $\{F^n\}$

(i) $$nP\left[|U_r^{\alpha n}|>\eta_r;\ \text{for all } r=1,\ldots,k|\underline{\theta}^n\right] \leq n \sum_{r=1}^{k} P\left[|U_r^{\alpha n}| > \eta_r|\underline{\theta}^n\right]$$

$$= n \sum_{r=1}^{k} P\left[\left|\left.\frac{\partial \ln f}{\partial \theta_r}\right|_{\underline{\theta}^n}\right| > n^{1/2}\ \eta_r|\underline{\theta}^n\right]$$

$$\leq n \sum_{r=1}^{k} \frac{E_{\underline{\theta}^n}\left[\left|\left.\frac{\partial \ln f}{\partial \theta_r}\right|_{\underline{\theta}^n}\right|\right]^{2+\bar{\eta}_2}}{(n^{1/2}\eta_r)^{2+\bar{\eta}_2}} < \frac{K}{n^{1/2\bar{\eta}_2}} \sum_{r=1}^{k} \eta_r^{-2-\bar{\eta}_2}\ .$$

by a form of the Tchebyshev inequality and by Assumption C, where $\bar{\eta}_2$ denotes the constant of Assumption C. Therefore,

$$\lim_{n \to \infty} n\ P[|U_r^{\alpha n}| > \eta_r;\ \text{for all } r = 1,\ldots,k|\underline{\theta}^n] = 0.$$

(ii) $$n \int_{|u_r| \leq \gamma} u_r dF^n(\underline{u}) = n \int_{|u_r| \leq \gamma} u_r dH^n(\sqrt{n}\,\underline{u}) = -\,n \int_{|u_r| > \gamma} u_r dH^n(\sqrt{n}\,\underline{u}),$$

the latter form following since $E_{\underline{\theta}^n}\left(\frac{\partial \ln f}{\partial \theta_r}\Big|_{\underline{\theta}^n}\right) = 0$. Set $t_r = \sqrt{n}\,u_r$

and

$$n\left|\int_{|u_r| \leq \gamma} u_r dF^n(\underline{u})\right| \leq \sqrt{n} \int_{|t_r| > \sqrt{n}\,\gamma} |t_r| dH^n(\underline{t}) \leq \sqrt{n} \int_{|t_r| > \sqrt{n}\,\gamma} |t_r| \left(\frac{|t_r|}{\sqrt{n}\,\gamma}\right)^{1+\bar{\eta}_2} dH^n(\underline{t})$$

$$< \frac{K}{n^{1/2\bar{\eta}_2}\,\gamma^{1+\bar{\eta}_2}} \quad \text{by Assumption C.}$$

Therefore,

$$\lim_{n \to \infty} n \int_{|u_r| \leq \gamma} u_r dF^n(\underline{u}) = 0, \; r = 1,\ldots,k.$$

(iii) It will now be shown that

$$\lim_{n \to \infty} n \int_{|u_r|,|u_s| \leq \gamma} u_r u_s dF^n(\underline{u}) = C_{rs}(\underline{\theta}^*), \; r,s = 1,\ldots,k.$$

Now

$$n \int\limits_{|u_r|,|u_s| \le \gamma} u_r u_s dF^n(\underline{u}) = n \int\limits_{|u_r|,|u_s| \le \gamma} u_r u_s dH^n (\sqrt{n}\, \underline{u})$$

$$= n \int\limits_{E_k} u_r u_s dH^n (\sqrt{n}\, \underline{u}) - n \int\limits_{|u_r| \le \gamma, |u_s| > \gamma} u_r u_s dH^n (\sqrt{n}\, \underline{u})$$

$$- n \int\limits_{|u_r| > \gamma, |u_s| \le \gamma} u_r u_s dH^n (\sqrt{n}\, \underline{u}) - n \int\limits_{|u_r| > \gamma, |u_s| > \gamma} u_r u_s dH^n (\sqrt{n}\, \underline{u}),$$

where E_k denotes k-dimensional Euclidean space. Then, using (ii) of Lemma 1.2, one obtains

$$\lim_{n \to \infty} n \int\limits_{E_k} u_r u_s dF^n(\underline{u}) = \lim_{n \to \infty} \int\limits_{E_k} t_r t_s dH^n(\underline{t})$$

$$= \lim_{n \to \infty} C_{rs}(\underline{\theta}^n) = C_{rs}(\underline{\theta}^*), \quad r,s = 1,\ldots,k.$$

An examination of the second integral on the right-hand side of the expression above yields

$$\left| n \int\limits_{|u_r| \le \gamma, |u_s| > \gamma} u_r u_s dF^n(\underline{u}) \right| \le \gamma n \int\limits_{|u_s| > \gamma} |u_s| \, dF^n(\underline{u}), \quad r,s = 1,\ldots,k.$$

From the discussion in (ii) above, the right-hand side can be made arbitrarily small for n sufficiently large. Thus

$$\lim_{n \to \infty} n \int_{|u_r| \leq \gamma, |u_s| > \gamma} u_r u_s dF^n(\underline{u}) = \lim_{n \to \infty} n \int_{|u_r| \leq \gamma, |u_s| > \gamma} u_r u_s dH^n (\sqrt{n}\, \underline{u}) = 0, \quad r,s = 1,\ldots,k.$$

Similarly

$$\lim_{n \to \infty} n \int_{|u_r| > \gamma, |u_s| \leq \gamma} u_r u_s dF^n(\underline{u}) = \lim_{n \to \infty} n \int_{|u_r| > \gamma, |u_s| \leq \gamma} u_r u_s dH^n (\sqrt{n}\, \underline{u}) = 0, \quad r,s = 1,\ldots,k.$$

Finally, it may be seen with the use of the Cauchy-Schwartz inequality and Assumption C that

$$\left| n \int_{|u_r|, |u_s| > \gamma} u_r u_s dF^n(\underline{u}) \right| = \left| n \int_{|u_r|, |u_s| > \gamma} u_r u_s dH^n(\sqrt{n}\, \underline{u}) \right|$$

$$\leq \left[n \int_{|u_r|, |u_s| > \gamma} u_r^2 dH^n (\sqrt{n}\, \underline{u}) \right]^{1/2} \left[n \int_{|u_r|, |u_s| > \gamma} u_s^2 \, dH^n (\sqrt{n}\, \underline{u}) \right]^{1/2}$$

$$= \left[\int_{|t_r|, |t_s| > \sqrt{n}\, \gamma} t_r^2 \, dH^n(\underline{t}) \right]^{1/2} \left[\int_{|t_r|, |t_s| > \sqrt{n}\, \gamma} t_s^2 \, dH^n(\underline{t}) \right]^{1/2}$$

$$\leq \frac{1}{n^{1/2\bar{\eta}_2} \gamma^{\bar{\eta}_2}} \left[\int_{|t_r| > \sqrt{n}\, \gamma} |t_r|^{2+\bar{\eta}_2} dH^n(\underline{t}) \right]^{1/2} \left[\int_{|t_s| > \sqrt{n}\, \gamma} |t_s|^{2+\bar{\eta}_2} dH^n(\underline{t}) \right]^{1/2}$$

$$< \frac{K}{n^{1/2\bar{\eta}_2} \gamma^{\bar{\eta}_2}} .$$

Therefore

$$\lim_{n \to \infty} n \int_{|u_r|,|u_s|>\gamma} u_r u_s dF^n(\underline{u}) = \lim_{n \to \infty} n \int_{|u_r|,|u_s|>\gamma} u_r u_s dH^n(\sqrt{n}\,\underline{u}) = 0, \; r,s = 1,\ldots,k.$$

Combining the four integrals one obtains

$$\lim_{n \to \infty} n \int_{|u_r|,|u_s|\leq\gamma} u_r u_s dF^n(\underline{u}) = C_{rs}(\underline{\theta}^*), \; r,s = 1,\ldots,k.$$

Now consider the sequence $\{G^n\}$ of normal distributions $\{\eta[\underline{v};\, \underline{\theta}, n^{-1}\underline{C}(\underline{\theta}^n)]\}$. It is clear that the existence of third absolute moments implies that the limits corresponding to (i), (ii), and (iii) exist and equal zero, zero, and $C_{rs}(\underline{\theta}^*)$, $r,s = 1,\ldots,k$ respectively. Thus the conditions of Bergström's theorem are satisfied and hence $\{\underline{Y}^n(\underline{\theta}^n)\}$ and $\{\underline{W}^n\}$ have the same limiting distribution. But $\underline{W}^n$ has the multivariate normal c.d.f., $\eta[\underline{w};\, \underline{\theta},\, \underline{C}(\underline{\theta}^n)\}$ with limit, $\eta[\underline{w};\underline{\theta},\underline{C}(\underline{\theta}^*)]$, as $n \to \infty$, through the continuity of $\underline{C}(\underline{\theta}^n)$ as established in Lemma 1.2, (ii). Hence

$$\lim_{n \to \infty} \mathcal{F}^n(\underline{y}) = \eta[\underline{y};\, \underline{\theta},\, \underline{C}(\underline{\theta}^*)].$$

<u>Lemma 1.6.</u> Given Assumptions A1 and A4, L^n_{rst}, as defined in (1.3.5) has a limiting bound in probability; namely, for any $\epsilon > 0$

$$P\{|L^n_{rst}| < M;\text{ for all } r,s,t = 1,\ldots,k|\underline{\theta}^n\} > 1 - \epsilon$$

for n sufficiently large, where M is the positive constant introduced in Assumption A4.

<u>Proof:</u> By Assumptions A1 and A4

$$|L_{rst}| \leq \frac{1}{n} \sum_{\alpha=1}^{n} \left| \frac{\partial^3 \ln f(\underline{x}_\alpha, \underline{\theta})}{\partial\theta_r \; \partial\theta_s \; \partial\theta_t} \right| < \frac{1}{n} \sum_{\alpha=1}^{n} H_{rst}(\underline{x}_\alpha)$$

for all $\underline{\theta} \in \Omega$.

Chanda [1954] considered a sequence of independent and identically distributed random variables so that the existence of the mean of $H_{rst}(\underline{x})$ was sufficient to conclude that

$$p \lim_{n \to \infty} \left[\frac{1}{n} \sum_{\alpha=1}^{n} H_{rst}(\underline{x}_\alpha) - E\{H_{rst}(\underline{x})\} \right] = 0.$$

However in the present situation there is a triangular array of random variables which are identically distributed within rows but for which the distribution is changing from one row to the next. Thus stronger conditions of the type given in Assumption A4 appear to be required.

It will now be shown that

$$p \lim_{n \to \infty} \left[\frac{1}{n} \sum_{\alpha=1}^{n} H_{rst}(\underline{x}_\alpha) - E_{\underline{\theta}^n}\{H_{rst}(\underline{x})\}; \right.$$

$$\left. \text{for all } r,s,t = 1,\ldots,k \middle| \underline{\theta}^n \right] = 0.$$

This may be accomplished by appealing to the multivariate limit theorem of Bergström, which was restated in the proof of Lemma 1.5, to show that the distribution of the vector

$$\underline{Z}^n = \left[\frac{1}{n} \sum_{\alpha=1}^{n} H_{rst}(\underline{x}_\alpha) - E_{\underline{\theta}^n}\{H_{rst}(\underline{x})\};\ r,s,t = 1,\ldots,k \right]$$

converges at all points of continuity to a distribution having all its mass concentrated at zero, namely,

$$G(\underline{z}) = \begin{cases} 0 & \text{if any } z_{rst} < 0 \\ 1 & \text{if all } z_{rst} \geq 0 . \end{cases}$$

The application of this theorem requires a demonstration that Assumptions A1 and A4 are sufficient for conditions (i), (ii), and (iii) of Bergström's theorem. Let the following correspondences be given. Setting the vector

$$\underline{U}^{\alpha n} = \left[\frac{1}{n} (H_{rst}(\underline{x}_\alpha) - E_{\underline{\theta}^n}\{H_{rst}(\underline{x})\});\ r,s,t = 1,\ldots,k \right]$$

one obtains $\underline{Z}^n = \sum_{\alpha=1}^{n} \underline{U}^{\alpha n}$. Also let $\underline{V}^{\alpha n}$, $\alpha = 1,\ldots,n$ have the c.d.f. G of the point mass distribution given above for all n. Consider first the sequence $\{F^n\}$ of c.d.f.'s for $\underline{U}^{\alpha n}$, $\alpha = 1,\ldots,n$.

(i) $$n\, P\left[|U^{\alpha n}_{rst}| > \eta_{rst};\ \text{for all } r,s,t = 1,\ldots,k|\underline{\theta}^n\right]$$

$$\leq n \sum_{r=1}^{k}\sum_{s=1}^{k}\sum_{t=1}^{k} P\left[\,|U^{\alpha n}_{rst}| > \eta_{rst}|\underline{\theta}^n\right]$$

$$= n \sum_{r=1}^{k}\sum_{s=1}^{k}\sum_{t=1}^{k} P\left[|H_{rst}(\underline{x}_\alpha) - E_{\underline{\theta}^n}\{H_{rst}(\underline{x})\}| > n\eta_{rst}|\underline{\theta}^n\right]$$

$$\leq n \sum_{r=1}^{k}\sum_{s=1}^{k}\sum_{t=1}^{k} \frac{E_{\underline{\theta}^n}\left[|H_{rst}(\underline{x}) - E_{\underline{\theta}^n}\{H_{rst}(\underline{x})\}|\right]^{1+\overline{\eta}_1}}{n^{1+\overline{\eta}_1}\ \eta_{rst}^{1+\overline{\eta}_1}}$$

$$< \frac{L}{n^{\overline{\eta}_1}} \sum_{r=1}^{k}\sum_{s=1}^{k}\sum_{t=1}^{k} \eta_{rst}^{-1-\overline{\eta}_1}$$

by a form of the Tchebyshev inequality and by Assumption A4. Therefore

$$\lim_{n\to\infty} n\, P\left[|U^{\alpha n}_{rst}| > \eta_{rst};\ \text{for all } r,s,t = 1,\ldots,k|\underline{\theta}^n\right] = 0.$$

(ii) $$n \int_{|u_{rst}|\leq \gamma} u_{rst}\, dF^n(\underline{u}) = -\, n \int_{|u_{rst}|> \gamma} u_{rst}\, dF^n(\underline{u})$$

since the integral over the entire range is zero. Now

$$n\left|\int_{|u_{rst}|> \gamma} u_{rst}\, dF^n(\underline{u})\right| \leq n \int_{|u_{rst}|> \gamma} |u_{rst}|\left(\frac{|u_{rst}|}{\gamma}\right)^{\overline{\eta}_1} dF^n(\underline{u}) < \frac{L}{(n\gamma)^{\overline{\eta}_1}}$$

by the definition of u_{rst} and Assumption A4. Therefore

$$\lim_{n \to \infty} n \int_{|u_{rst}| \leq \gamma} u_{rst} dF^n(\underline{u}) = 0, \quad r,s,t = 1,\dots,k.$$

(iii) $$\left| n \int_{|u_{rst}|,|u_{r's't'}| \leq \gamma} u_{rst} u_{r's't'} dF^n(\underline{u}) \right|$$

$$\leq \left\{ n \int_{|u_{rst}|,|u_{r's't'}| \leq \gamma} u_{rst}^2 dF^n(\underline{u}) \right\}^{1/2} \left\{ n \int_{|u_{rst}|,|u_{r's't'}| \leq \gamma} u_{r's't'}^2 dF^n(\underline{u}) \right\}^{1/2}$$

by the Cauchy-Schwartz inequality. Now

$$n \int_{|u_{rst}|,|u_{r's't'}| \leq \gamma} u_{rst}^2 dF^n(\underline{u}) \leq n \int_{|u_{rst}| \leq \gamma} u_{rst}^2 dF^n(\underline{u}) .$$

Then, if $0 < \bar{\eta}_1 < 1$, one has $u_{rst}^2 = |u_{rst}|^{1+\bar{\eta}_1} |u_{rst}|^{1-\bar{\eta}_1}$, so that over the region of integration

$$n \int_{|u_{rst}| \leq \gamma} u_{rst}^2 dF^n(\underline{u}) \leq n\gamma^{1-\bar{\eta}_1} \int_{|u_{rst}| \leq \gamma} |u_{rst}|^{1+\bar{\eta}_1} dF^n(\underline{u}) < L\gamma^{1-\bar{\eta}_1}/n^{\bar{\eta}_1}$$

again by the definition of u_{rst} and Assumption A4. Hence

$$n \int_{|u_{rst}|,|u_{r's't'}| \leq \gamma} u_{rst}^2 dF^n(\underline{u}) \; < \; L\gamma^{1-\bar{\eta}_1}/n^{\bar{\eta}_1}$$

and thus

$$\lim_{n \to \infty} \int_{|u_{rst}|,|u_{r's't'}| \leq \gamma} u_{rst}u_{r's't'} dF^n(\underline{u}) = 0, \; r,s,t = 1,\ldots,k.$$

Now consider the sequence $\underline{V}^{\alpha n}$, $\alpha = 1,\ldots,n$, of random variables having the point mass distribution for all n. It is clear that the limits corresponding to (i), (ii), and (iii) exist and equal zero. Thus the conditions of Bergström's theorem are satisfied and hence $\{\underline{Z}^n\}$ converges in distribution to the point mass distribution, or equivalently

$$p \lim_{n \to \infty} \left[\frac{1}{n} \sum_{\alpha=1}^{n} H_{rst}(\underline{x}_\alpha) - E_{\underline{\theta}^n}\{H_{rst}(\underline{x})\}|\underline{\theta}^n\right] = 0, \; r,s,t = 1,\ldots,k.$$

Now

$$P\left[|L_{rst}^n| < M; \text{ for all } r,s,t = 1,\ldots,k|\underline{\theta}^n\right]$$

$$\geq P\left[\frac{1}{n} \sum_{\alpha=1}^{n} H_{rst}(\underline{x}_\alpha) < M; \text{ for all } r,s,t = 1,\ldots,k|\underline{\theta}^n\right]$$

$$\geq P\left[|\frac{1}{n} \sum_{\alpha=1}^{n} H_{rst}(\underline{x}_\alpha) - E_{\underline{\theta}^n}\{H_{rst}(\underline{x})\}| < M - E_{\underline{\theta}^n}\{H_{rst}(\underline{x})\};\right.$$

$$\left. \text{for all } r,s,t = 1,\ldots,k|\underline{\theta}^n\right].$$

By Assumption A4, $E_{\underline{\theta}}\{H_{rst}(\underline{x})\} < M$ for all $\underline{\theta} \in \Omega$, and by the demonstration above, the final probability can be made arbitrarily close to unity. Hence, for any $\epsilon > 0$

$$P\{|L^n_{rst}| < M; \text{ for all } r,s,t = 1,\ldots,k|\underline{\theta}^n\} > 1 - \epsilon$$

for n sufficiently large.

Lemma 1.7. Let $\{\underline{\tilde{\theta}}^n\}$ be a sequence of parameter estimates in Ω which converge in probability to $\underline{\theta}^*$ under the parameter sequence $\{\underline{\theta}^n\}$. Then, given Assumptions A and B and the sequence $\{\underline{\theta}^n\}$,

$$p \lim_{n \to \infty} [L^n_{rs}(\underline{\tilde{\theta}}^n)|\underline{\theta}^n] = -C_{rs}(\underline{\theta}^*), \; r,s = 1,\ldots,k.$$

Proof: Consider the Taylor series expansion of $L^n_{rs}(\underline{\tilde{\theta}}^n)$ about $\underline{\theta} = \underline{\theta}^*$

$$L^n_{rs}(\underline{\tilde{\theta}}^n) = L^n_{rs}(\underline{\theta}^*) + \sum_{t=1}^{k} (\tilde{\theta}^n_t - \theta^*_t) \, L^n_{rst}(\underline{\theta}')$$

for some $\underline{\theta}'$ such that $\|\underline{\theta}' - \underline{\theta}^*\| \leq \|\underline{\tilde{\theta}}^n - \underline{\theta}^*\|$. Then from Lemma 1.3

$$p \lim_{n \to \infty} \left[L^n_{rs}(\underline{\theta}^*)|\underline{\theta}^n\right] = -C_{rs}(\underline{\theta}^*), \; r,s = 1,\ldots,k$$

while from Lemma 1.6, $L^n_{rst}(\underline{\theta}')$, $r,s,t = 1,\ldots,k$, has a limiting bound in probability for all $\underline{\theta} \in \Omega$. Since $p \lim_{n \to \infty} [\{\underline{\tilde{\theta}}^n - \underline{\theta}^*\}|\underline{\theta}^n] = 0$, it then follows with the use of Slutsky's theorem (cf. Cramér [1946], p. 255) that

$$p \lim_{n \to \infty} \left[\sum_{t=1}^{k} (\tilde{\theta}^n_t - \theta^*_t) \, L^n_{rst}(\underline{\theta}')|\underline{\theta}^n\right] = 0, \; r,s = 1,\ldots,k.$$

Therefore, a second application of Slutsky's theorem gives

$$p\lim_{n\to\infty}\left[L_{rs}^{n}(\underline{\tilde{\theta}}^{n})\,|\,\underline{\theta}^{n}\right] = -C_{rs}(\underline{\theta}^{*}),\ r,s = 1,\ldots,k.$$

Theorem 1.2. (Limiting normality)

Let $\underline{\hat{\theta}}^{n}$ and $\underline{\hat{\hat{\theta}}}^{n}$ be the two sets of estimates of Theorem 1.1. Let the matrix $\underline{C} = \underline{C}(\underline{\theta}^{*}) = [C_{rs}(\underline{\theta}^{*});\ r,s = 1,\ldots,k]$, $\underline{\Sigma} = \underline{C}^{-1}$ and $\underline{\delta} = (\delta_1,\ldots,\delta_\ell,0,\ldots,0)$. Then under Assumptions A, B and C, and for the given sequence $\{\underline{\theta}^{n}\}$,

(i) $\sqrt{n}\,(\underline{\hat{\theta}}^{n} - \underline{\theta}^{*})$ has a limiting multivariate normal distribution with mean $\underline{\delta}$ and variance-covariance matrix $\underline{\Sigma}$,

(ii) $\sqrt{n}\,({}_{2}\underline{\hat{\hat{\theta}}}^{n} - {}_{2}\underline{\theta}^{*})$ has a limiting multivariate normal distribution with mean ${}_{2}\underline{\theta}$ and variance-covariance matrix $\underline{\bar{\Sigma}}_{22}$.

Proof: First consider a Taylor series expansion of $Y_r^{n}(\underline{\theta}^{*})$ about $\underline{\theta} = \underline{\theta}^{n}$. Since $\theta_r^{n} = \theta_r^{*}$, $r = \ell + 1,\ldots,k$ this expansion is given by

$$Y_r^{n}(\underline{\theta}^{*}) = Y_r^{n}(\underline{\theta}^{n}) + \sqrt{n}\sum_{s=1}^{\ell}(\theta_s^{*} - \theta_s^{n})L_{rs}^{n}(\underline{\tilde{\theta}}^{n}),\ r = 1,\ldots,k,$$

for some $\underline{\tilde{\theta}}^{n}$ such that $\|\underline{\tilde{\theta}}^{n} - \underline{\theta}^{*}\| \le \|\underline{\theta}^{*} - \underline{\theta}^{n}\|$. Now $\lim_{n\to\infty}\underline{\theta}^{n} = \underline{\theta}^{*}$ implies $\lim_{n\to\infty}\underline{\tilde{\theta}}^{n} = \underline{\theta}^{*}$ so that it follows from Lemma 1.7 that

$$p\lim_{n\to\infty}[L_{rs}^{n}(\underline{\tilde{\theta}}^{n})\,|\,\underline{\theta}^{n}] = -\,C_{rs}(\underline{\theta}^{*}),\ r,s = 1,\ldots,k.$$

Thus, it follows from Slutsky's theorem that

$$\underset{n\to\infty}{\mathrm{p\ lim}}\left[\sqrt{n}\sum_{s=1}^{\ell}(\theta_s^* - \theta_s^n)L_{rs}^n(\underline{\tilde{\theta}}^n)\,|\,\underline{\theta}^n\right] = \sum_{s=1}^{\ell}\delta_s C_{rs}(\underline{\theta}^*) = \sum_{s=1}^{k}\delta_s C_{rs}(\underline{\theta}^*),$$

$$r,s = 1,\ldots,k.$$

By Lemma 1.5, $\underline{Y}^n(\underline{\theta}^n)$ has a limiting multivariate normal distribution with mean vector $\underline{\theta}$ and variance-covariance matrix $\underline{C}$. Hence it follows from Lemma 1 of Chiang [1956, p. 338] that $[Y_r^n(\underline{\theta}^*);\ r=1,\ldots,k]$ has a limiting multivariate normal distribution with mean vector $\underline{\delta}\ \underline{C}$ and variance-covariance matrix $\underline{C}$.

(i) Consider now the Taylor series expansion of $L_r^n(\underline{\theta}^*)$ about $\underline{\theta} = \underline{\hat{\theta}}^n$ so that

$$L_r^n(\underline{\theta}^*) = -\sum_{s=1}^{k}(\hat{\theta}_s^n - \theta_s^*)L_{rs}^n(\underline{\tilde{\theta}}^n),\ r = 1,\ldots,k,$$

for some $\underline{\tilde{\theta}}^n$ such that $\|\underline{\tilde{\theta}}^n - \underline{\theta}^*\| \le \|\underline{\hat{\theta}}^n - \underline{\theta}^*\|$. Now by Theorem 1.1, $\underset{n\to\infty}{\mathrm{p\ lim}}\ [\underline{\hat{\theta}}^n|\underline{\theta}^n] = \underline{\theta}^*$ so that $\underset{n\to\infty}{\mathrm{p\ lim}}\ [\underline{\tilde{\theta}}^n|\underline{\theta}^n] = \underline{\theta}^*$. Then by Lemma 1.7

$$\underset{n\to\infty}{\mathrm{p\ lim}}\ [L_{rs}^n(\underline{\tilde{\theta}}^n)|\underline{\theta}^n] = -\,C_{rs}(\underline{\theta}^*),\ r,s = 1,\ldots,k.$$

Therefore

$$L_r^n(\underline{\theta}^*) = \sum_{s=1}^{k}(\hat{\theta}_s^n - \theta_s^*)\{C_{rs}(\underline{\theta}^*) + o_p(1)\},\ r=1,\ldots,k, \qquad (1.3.6)$$

where $o_p(1)$ denotes a quantity which converges to zero in probability under $\{\underline{\theta}^n\}$. For large n, the matrix $[C_{rs}(\underline{\theta}^*) + o_p(1);\ r,s=1,\ldots,k]$

approaches $\underline{C}$ in probability and thus may be inverted to give the matrix $[\sigma_{rs}(\underline{\theta}^*) + o_p(1);\ r,s = 1,\ldots,k]$ where $\underline{\Sigma} = [\sigma_{rs}(\underline{\theta}^*);\ r,s = 1,\ldots,k]$. Thus one has

$$(\hat{\theta}_r^n - \theta_r^*) = \sum_{s=1}^{k} L_s^n(\underline{\theta}^*)\ \{\sigma_{sr}(\underline{\theta}^*) + o_p(1)\},\ r = 1,\ldots,k.$$

After multiplication by $\sqrt{n}$ and use of the definition in Lemma 1.5 that $Y_s^n(\underline{\theta}) = \sqrt{n}\ L_s^n(\underline{\theta})$, $s = 1,\ldots,k$, it follows that

$$\sqrt{n}\ (\hat{\theta}_r^n - \theta_r^*) = \sum_{s=1}^{k} Y_s^n(\underline{\theta}^*)\ \{\sigma_{sr}(\underline{\theta})^* + o_p(1)\},\ r = 1,\ldots,k. \quad (1.3.7)$$

It then follows from Lemma 1 of Chiang [1956, p. 338] that $\sqrt{n}(\hat{\underline{\theta}}^n - \underline{\theta}^*)$ has a limiting multivariate normal distribution with mean vector $\underline{\delta}\ \underline{C}\ \underline{C}^{-1} = \underline{\delta}$ and variance-covariance matrix $\underline{C}^{-1}\underline{C}\ \underline{C}^{-1} = \underline{C}^{-1} = \underline{\Sigma}$.

(ii) By an analogous argument one can expand $L_r^n(\underline{\theta}^*)$ about $\underline{\theta} = \hat{\hat{\underline{\theta}}}^n$, $r = \ell+1,\ldots,k$, to obtain, since $\hat{\hat{\theta}}_r^n = \theta_r^* = \theta_r^o$, $r = 1,\ldots,\ell$,

$$L_r^n(\underline{\theta}^*) = -\sum_{s=\ell+1}^{k} (\hat{\hat{\theta}}_s^n - \theta_s^*) L_{rs}^n(\tilde{\hat{\underline{\theta}}}^n),\ r = \ell+1,\ldots,k$$

for some $\tilde{\hat{\underline{\theta}}}^n$ such that $\|\tilde{\hat{\underline{\theta}}}^n - \underline{\theta}^*\| \leq \|\hat{\hat{\underline{\theta}}}^n - \underline{\theta}^*\|$. By Theorem 1.1, $p\lim_{n\to\infty} [\hat{\hat{\underline{\theta}}}^n | \underline{\theta}^n] = \underline{\theta}^*$ and as above one can obtain from Lemma 1.7

$$p\lim_{n\to\infty} [L_{rs}^n(\tilde{\hat{\underline{\theta}}}^n) | \underline{\theta}^n] = -C_{rs}(\underline{\theta}^*),\ r,s = \ell+1,\ldots,k.$$

Therefore

$$L_r^n(\underline{\theta}^*) = \sum_{s=\ell+1}^{k} (\hat{\hat{\theta}}_s^n - \theta_s^*)\{C_{rs}(\underline{\theta}^*) + o_p(1)\}, \; r = \ell+1,\ldots,k. \quad (1.3.8)$$

Now for large n, the matrix $[C_{rs}(\underline{\theta}^*) + o_p(1);\ r,s = \ell+1,\ldots,k]$ approaches the matrix $[C_{rs}(\underline{\theta}^*);\ r,s = \ell+1,\ldots,k]$ and may be inverted to give the matrix $[\bar{\sigma}_{rs}(\underline{\theta}^*) + o_p(1);\ r,s = \ell+1,\ldots,k]$ where $\underline{\bar{\Sigma}}_{22} = [\bar{\sigma}_{rs}(\underline{\theta}^*);\ r,s = \ell+1,\ldots,k] = \underline{C}_{22}^{-1}$. Thus one has

$$(\hat{\hat{\theta}}_r^n - \theta_r^*) = \sum_{s=\ell+1}^{k} L_s^n(\underline{\theta}^*) + o_p(1)\}, \; r = \ell+1,\ldots,k,$$

or, after multiplying through by $\sqrt{n}$,

$$\sqrt{n}\,(\hat{\hat{\theta}}_r^n - \theta_r^*) = \sum_{s=\ell+1}^{k} Y_s^n(\underline{\theta}^*)\,\{\bar{\sigma}_{sr}(\underline{\theta}^*) + o_p(1)\}, \; r = \ell+1,\ldots,k. \quad (1.3.9)$$

Now ${}_2\underline{Y}^n(\underline{\theta}^*)$ has a limiting multivariate normal distribution with mean vector ${}_2\underline{\theta}$ and variance-covariance matrix $\underline{C}_{22}$ as it is the marginal distribution of the last $(k-\ell)$ components of $\underline{Y}^n(\underline{\theta}^*)$ whose limiting distribution was obtained at the beginning of the proof of this theorem. Thus again using Lemma 1 of Chiang [1956, p. 338], one sees that $\sqrt{n}\,({}_2\hat{\hat{\underline{\theta}}}^n - {}_2\underline{\theta}^*)$ has a limiting multivariate normal distribution with mean vector ${}_2\underline{\theta}$ and variance-covariance matrix $\underline{\bar{\Sigma}}_{22}\underline{C}_{22}\underline{\bar{\Sigma}}_{22} = \underline{\bar{\Sigma}}_{22}$.

II. ASYMPTOTIC DISTRIBUTIONS

2.1 Asymptotic distribution of $-2\ln\lambda_n$

It is now possible to prove the following theorem which is the main result of interest in this study.

Theorem 2.1. Let $\{\underline{\theta}^n\}$ be the sequence of local alternatives with $\theta_i^n = \theta_i^o + \delta_{in}/\sqrt{n}$, $\lim_{n\to\infty} \delta_{in} = \delta_i$, $i = 1,\ldots,\ell$, and $\theta_i^n = \theta_i^*$, $i=\ell+1,\ldots,k$ where ${}_2\underline{\theta}^*$ is the vector of true values of the nuisance parameters. Let $F_\ell(\bar{\lambda}^2,t)$ be the distribution function of the non-central chi-square distribution with ℓ degrees of freedom and non-centrality parameter $\bar{\lambda}^2$. If λ_n is the likelihood ratio statistic for testing the hypotheses H_o: $\underline{\theta} = \underline{\theta}^\omega$ where $\underline{\theta}^\omega = ({}_1\underline{\theta}^o, {}_2\underline{\theta}) \in \Omega$, then under Assumptions A, B, and C,

$$\lim_{n\to\infty} P[-2\ln\lambda_n \le t|\underline{\theta}^n] = F_\ell(\bar{\lambda}^2,t) ,$$

where $\bar{\lambda}^2 = {}_1\underline{\delta}\, \underline{\Sigma}_{11}^{-1}\, {}_1\underline{\delta}' = {}_1\underline{\delta}\, \underline{\bar{C}}_{11}\, {}_1\underline{\delta}'$ if the hypothesis to be tested is false, and $\bar{\lambda}^2 = 0$ if the hypothesis to be tested is true.

Proof: Consider the Taylor-series expansions of $\ln\phi(\underline{x},\underline{\theta}^*)$ about $\hat{\underline{\theta}}^n$ and $\hat{\hat{\underline{\theta}}}^n$ respectively:

$$\ln\phi(\underline{x},\underline{\theta}^*) = \ln\phi(\underline{x},\hat{\underline{\theta}}^n) + \frac{1}{2} \sum_{r=1}^{k} \sum_{s=1}^{k} n(\hat{\theta}_r^n - \theta_r^*)(\hat{\theta}_s^n - \theta_s^*) L_{rs}^n(\tilde{\underline{\theta}}^n),$$

$$\ln\phi(\underline{x},\underline{\theta}^*) = \ln\phi(\underline{x},\hat{\hat{\underline{\theta}}}^n) + \frac{1}{2} \sum_{r=\ell+1}^{k} \sum_{s=\ell+1}^{k} n(\hat{\hat{\theta}}_r^n - \theta_r^*)(\hat{\hat{\theta}}_s^n - \theta_s^*) L_{rs}^n(\approx{\underline{\theta}}^n),$$

for some $\tilde{\underline{\theta}}^n$ such that $\|\tilde{\underline{\theta}}^n - \underline{\theta}^*\| \le \|\hat{\underline{\theta}}^n - \underline{\theta}^*\|$ and some $\overset{\approx}{\underline{\theta}}^n$ such that

$\|\underline{\tilde{\tilde{\theta}}}^n - \underline{\theta}^*\| \le \|\underline{\hat{\hat{\theta}}}^n - \underline{\theta}^*\|$ where $\underline{\tilde{\tilde{\theta}}}^n = ({}_1\underline{\theta}^o, {}_2\underline{\tilde{\tilde{\theta}}}^n)$. Now

$\operatorname*{p\,lim}_{n\to\infty} [\underline{\hat{\theta}}^n | \underline{\theta}^n] = \operatorname*{p\,lim}_{n\to\infty} [\underline{\hat{\hat{\theta}}}^n | \underline{\theta}^n] = \underline{\theta}^*$ by Theorem 1.1 so that

$\operatorname*{p\,lim}_{n\to\infty} [\underline{\tilde{\theta}}^n | \underline{\theta}^n] = \operatorname*{p\,lim}_{n\to\infty} [\underline{\tilde{\tilde{\theta}}}^n | \underline{\theta}^n] = \underline{\theta}^*$. By Lemma 1.7

$$\operatorname*{p\,lim}_{n\to\infty} [L^n_{rs}(\underline{\tilde{\theta}}^n) | \underline{\theta}^n] = \operatorname*{p\,lim}_{n\to\infty} [L^n_{rs}(\underline{\tilde{\tilde{\theta}}}^n) | \underline{\theta}^n] = - C_{rs}(\underline{\theta}^*),$$

$$r, s = 1, \ldots, k.$$

Thus

$$\begin{aligned} - 2\ln\lambda_n &= 2[\ln\phi(\underline{x}, \underline{\hat{\theta}}^n) - \ln\phi(\underline{x}, \underline{\hat{\hat{\theta}}}^n)] \\ &= \sum_{r=1}^{k} \sum_{s=1}^{k} n(\hat{\theta}^n_r - \theta^*_r)(\hat{\theta}^n_s - \theta^*_s) C_{rs}(\underline{\theta}^*) \\ &\quad - \sum_{r=\ell+1}^{k} \sum_{s=\ell+1}^{k} n(\hat{\hat{\theta}}^n_r - \theta^*_r)(\hat{\hat{\theta}}^n_s - \theta^*_s) C_{rs}(\underline{\theta}^*) + o_p(1). \end{aligned}$$

In matrix form this becomes

$$- 2\ln\lambda_n = n(\underline{\hat{\theta}}^n - \underline{\theta}^*)\underline{C}(\underline{\hat{\theta}}^n - \underline{\theta}^*)' - n({}_2\underline{\hat{\hat{\theta}}}^n - {}_2\underline{\theta}^*)\underline{C}_{22}({}_2\underline{\hat{\hat{\theta}}}^n - {}_2\underline{\theta}^*)' + o_p(1).$$

It has been shown by Doob [1935, p. 164] that, if one has two sequences of random variables the first of which converges in distribution while the second converges in probability to zero, then the product sequence converges in probability to zero. Hence, it follows from Equation (1.3.9) that

$$\sqrt{n}({}_2\underline{\hat{\hat{\theta}}}^n - {}_2\underline{\theta}^*) = {}_2\underline{Y}^n(\underline{\theta}^*)\ \underline{\Sigma}_{22} + o_p(1),$$

and that

$$n({}_2\hat{\underline{\theta}}^n - {}_2\underline{\theta}^*)\ \underline{C}_{22}({}_2\hat{\underline{\theta}}^n - {}_2\underline{\theta}^*)'$$

$$= [{}_2\underline{Y}^n(\underline{\theta}^*)\ \underline{\Sigma}_{22} + o_p(1)]\ \underline{C}_{22}[{}_2\underline{Y}^n(\underline{\theta}^*)\ \underline{\Sigma}_{22} + o_p(1)]'$$

$$= {}_2\underline{Y}^n(\underline{\theta}^*)\ \underline{\Sigma}_{22}\ \underline{C}_{22}\ \underline{\Sigma}_{22}\ \{{}_2\underline{Y}^n(\underline{\theta}^*)\}' + o_p(1)$$

$$= {}_2\underline{Y}^n(\underline{\theta}^*)\ \underline{\Sigma}_{22}\ \{{}_2\underline{Y}^n(\underline{\theta}^*)\}' + o_p(1).$$

Now multiplying by $\sqrt{n}$ in Equation (1.3.6), noting that $\sqrt{n}\ (\hat{\underline{\theta}}^n - \underline{\theta}^*)$ has a limiting distribution by Theorem 1.2, and again using Doob's result, one obtains

$$\underline{Y}^n(\underline{\theta}^*) = \sqrt{n}\ (\hat{\underline{\theta}}^n - \underline{\theta}^*)\ \underline{C} + o_p(1)$$

so that

$${}_2\underline{Y}^n(\underline{\theta}^*) = \sqrt{n}\ (\hat{\underline{\theta}}^n - \underline{\theta}^*) \begin{bmatrix} \underline{C}_{12} \\ \underline{C}_{22} \end{bmatrix} + o_p(1).$$

Thus

$$n({}_2\hat{\underline{\theta}}^n - {}_2\underline{\theta}^*)\ \underline{C}_{22}({}_2\hat{\underline{\theta}}^n - {}_2\underline{\theta}^*)'$$

$$= [\sqrt{n}\ (\hat{\underline{\theta}}^n - \underline{\theta}^*) \begin{bmatrix} \underline{C}_{12} \\ \underline{C}_{22} \end{bmatrix} + o_p(1)]\ \underline{C}_{22}^{-1}\ [\sqrt{n}\ (\hat{\underline{\theta}}^n - \underline{\theta}^*) \begin{bmatrix} \underline{C}_{12} \\ \underline{C}_{22} \end{bmatrix} + o_p(1)]' + o_p(1)$$

$$= n(\hat{\underline{\theta}}^n - \underline{\theta}^*)\ \underline{D}\ (\hat{\underline{\theta}}^n - \underline{\theta}^*)' + o_p(1),$$

where

$$\underline{D} = \begin{bmatrix} \underline{C}_{12} \\ \underline{C}_{22} \end{bmatrix} \underline{C}_{22}^{-1}\ [\underline{C}_{21}, \underline{C}_{22}] = \begin{bmatrix} \underline{C}_{12}\ \underline{C}_{22}^{-1}\ \underline{C}_{21} & , & \underline{C}_{12} \\ \underline{C}_{21} & & \underline{C}_{22} \end{bmatrix}.$$

Therefore,

$$- 2\ln\lambda_n = n(\underline{\hat{\theta}}^n - \underline{\theta}^*)[\underline{C} - \underline{D}](\underline{\hat{\theta}}^n - \underline{\theta}^*)' + o_p(1)$$

$$= n({}_1\underline{\hat{\theta}}^n - {}_1\underline{\theta}^o)[\underline{C}_{11} - \underline{C}_{12}\,\underline{C}_{22}^{-1}\,\underline{C}_{21}]({}_1\underline{\hat{\theta}}^n - {}_1\underline{\theta}^o)' + o_p(1)$$

$$= n({}_1\underline{\hat{\theta}}^n - {}_1\underline{\theta}^o)\ \underline{\Sigma}_{11}^{-1}({}_1\hat{\theta}^n - {}_1\underline{\theta}^o)' + o_p(1)$$

$$= n({}_1\underline{\hat{\theta}}^n - {}_1\underline{\theta}^o)\ \underline{\overline{C}}_{11}({}_1\underline{\hat{\theta}}^n - {}_1\underline{\theta}^o)' + o_p(1).$$

Thus

$$- 2\ln\lambda_n = \sum_{r=1}^{\ell}\ \sum_{s=1}^{\ell} n(\hat{\theta}_r^n - \theta_r^o)(\hat{\theta}_s^n - \theta_s^o)\ \overline{C}_{rs}(\underline{\theta}^*) + o_p(1).$$

Now, since $\sqrt{n}\ ({}_1\underline{\hat{\theta}}^n - {}_1\underline{\theta}^o)$ has a limiting normal distribution with mean vector ${}_1\underline{\delta}$ and variance-covariance matrix $\underline{\Sigma}_{11}$ by Theorem 1.2, it is clear that

$$Q_n = \sum_{r=1}^{\ell}\ \sum_{s=1}^{\ell}\ n(\hat{\theta}_r^n - \theta_r^o)(\hat{\theta}_s^n - \theta_s^o)\ \overline{C}_{rs}(\underline{\theta}^*)$$

has a limiting noncentral chi-square distribution with ℓ degrees of freedom and with noncentrality parameter

$$\overline{\lambda}^2 = \lim_{n\to\infty}\ \sum_{r=1}^{\ell}\ \sum_{s=1}^{\ell} n(\theta_r^n - \theta_r^o)(\theta_s^n - \theta_s^o)\overline{C}_{rs}(\underline{\theta}^*) = {}_1\underline{\delta}\ \underline{\overline{C}}_{11}\ {}_1\underline{\delta}'.$$

Thus, since $- 2\ln\lambda_n$ and Q_n have the same limiting distribution in that they differ only by a quantity which is converging to zero in probability, one has

$$\lim_{n\to\infty} P[-2\ln\lambda_n \leq t|\underline{\theta}^n] = F_\ell(\overline{\lambda}^2,\ t).$$

2.2 Sampling from associated populations

Situations often arise in which observations are made on several distinct populations which are related in the sense that their underlying probability distributions depend on common parameters. Such populations have been referred to as "associated" by Bradley and Gart [1962] who have extended the results of Chanda [1954] on the large-sample properties of maximum likelihood estimators to such situations. In Section 2.4 of their study, Bradley and Gart have commented that an analagous extension of the results given in earlier sections of the present study should be straightforward. That this is in fact the case will now be indicated.

Let $f_h(\underline{x}_h,\underline{\theta})$, $h = 1,\ldots,m$, denote m density or probability functions (continuous or discrete) where $\underline{x}_h$ is a random vector with values over a region R_h independent of $\underline{\theta} = (\theta_1,\ldots,\theta_k)$, a parameter vector lying in a k-dimensional parameter space Ω. Each $f_h(\underline{x}_h,\underline{\theta})$ need not depend on all $\theta_i, i=1,\ldots,k$. Let $\{\underline{x}_{h\alpha}\}, \alpha=1,\ldots,n_h$, be n_h independent observations on $\underline{x}_h$, $h=1,\ldots,m$. Then in the case where the random vectors $\underline{x}_h$ have continuous c.d.f.s, the assumptions of Section 1.2 are generalized as follows:

A'1. For almost all $\underline{x}_h \in R_h$ and for all $\underline{\theta} \in \Omega$

$$\frac{\partial \ln f_h}{\partial \theta_r}, \quad \frac{\partial^2 \ln f_h}{\partial \theta_r \partial \theta_s} \quad \text{and} \quad \frac{\partial^3 \ln f_h}{\partial \theta_r \partial \theta_s \partial \theta_t}$$

exist for $r, s, t = 1,\ldots,k$; $h = 1,\ldots,m$.

A'2. For all $f_h(\underline{x}_h,\underline{\theta})$ that are densities, for almost all $\underline{x}_h \in R_h$ and for every $\underline{\theta} \in \Omega$,

$$\left| \frac{\partial f_h}{\partial \theta_r} \right| < F_{hr}(\underline{x}_h) \text{ and } \left| \frac{\partial^2 f_h}{\partial \theta_r \partial \theta_s} \right| < F_{hrs}(\underline{x}_h),$$

where $F_{hr}(\underline{x}_h)$ and $F_{hrs}(\underline{x}_h)$ are integrable over R_h, $r,s = 1,\ldots,k$; $h = 1,\ldots,m$. These assumptions will permit certain interchanges of differentiation and integration.

A'3. For every $\underline{\theta} \in \Omega$, the matrix $\underline{C}(\underline{\theta}) = [C_{rs}(\underline{\theta});\ r,s = 1,\ldots,k]$ with

$$C_{rs}(\underline{\theta}) = \sum_{h=1}^{m} \mu_h \, E_{\underline{\theta}} \left[\frac{\partial \ln f_h}{\partial \theta_r} \bigg|_{\underline{\theta}} \; \frac{\partial \ln f_h}{\partial \theta_s} \bigg|_{\underline{\theta}} \right]$$

is positive definite with finite determinant. One defines the constant $\mu_h = n_h/N$ where $N = \sum_{h=1}^{m} n_h$, $h=1,\ldots,m$.

A'4. For almost all $\underline{x}_h \in R_h$ and for all $\underline{\theta} \in \Omega$

$$\left| \frac{\partial^3 \ln f_h}{\partial \theta_r \partial \theta_s \partial \theta_t} \right| < H_{hrst}(\underline{x}_h),$$

and

$$E_{\underline{\theta}} \left[H_{hrst}(\underline{x}_h) \right] < M_h < \infty,$$

and there exists a positive real number η_1, such that

$$E_{\underline{\theta}} \left[\left| H_{hrst}(\underline{x}_h) - E_{\underline{\theta}}\{H_{hrst}(\underline{x}_h)\} \right|^{1+\eta_1} \right] < L_h < \infty$$

for $r,s,t = 1,\ldots,k$; $h = 1,\ldots,m$, and where M_h and L_h are positive constants, $h = 1,\ldots,m$.

B' . There exists a positive real number ν_2 such that whenever

$$\|\underline{\theta}'' - \underline{\theta}'\| \leq \nu_2, \ \underline{\theta}', \underline{\theta}'' \in \Omega,$$

$$E_{\underline{\theta}'}\left[\left(\frac{\partial^2 \ln f_h}{\partial\theta_r\partial\theta_s}\Big|_{\underline{\theta}''}\right)^2\right] < T_h < \infty$$

for r, s = 1,...,k, and where T_h is a positive constant, h = 1,...,m.

C' . There exists a positive real number η_2 such that

$$E_{\underline{\theta}}\left[\left|\frac{\partial \ln f_h}{\partial\theta_r}\Big|_{\underline{\theta}}\right|^{2+\eta_2}\right] < K_h < \infty,$$

r = 1,...,k, for all $\underline{\theta} \in \Omega$, where K_h is a positive constant, h = 1,...,m.

Subject to Assumptions A' and B', Lemmas 1.1 and 1.2 of Section 1.2 clearly hold for each h, h = 1,...,m.

Suppose now that one wishes to consider the following hypotheses:

$$H_o: \underline{\theta} = \underline{\theta}^{\omega} = ({}_1\underline{\theta}^o, {}_2\underline{\theta}) \text{ and } H_1: \underline{\theta} = ({}_1\underline{\theta}^N, {}_2\underline{\theta}^*)$$

with $\theta_i^N = \theta_i^o + \delta_{iN}/\sqrt{N}$, where as $n_h = \mu_h N$, h = 1,...,m, and N become large, one has $\lim_{N\to\infty} \delta_{iN} = \delta_i$, i = 1,...,$\ell$.

Denote the joint likelihood function by

$$\phi(\underline{x}, \underline{\theta}) = \prod_{h=1}^{m} \prod_{\alpha=1}^{n_h} f_h(\underline{x}_\alpha, \underline{\theta}).$$

Then the likelihood equations under alternative and null hypotheses respectively are the following:

$$\frac{\partial \ln\phi(\underline{x},\underline{\theta})}{\partial\theta_r} = 0 \ , \ r = 1,\ldots,k, \tag{2.2.1}$$

$$\frac{\partial \ln \phi(\underline{x}, \underline{\theta}^{\omega})}{\partial \theta_r} = 0, \quad r=\ell+1, \ldots, k. \tag{2.2.2}$$

Now suppose that

$$L_r^N(\underline{\theta}) = \frac{1}{N} \frac{\partial \ln \phi(\underline{x}, \underline{\theta})}{\partial \theta_r} = \frac{1}{N} \sum_{h=1}^{m} \sum_{\alpha=1}^{n_h} \frac{\partial \ln f_h(\underline{x}_{h\alpha}, \underline{\theta})}{\partial \theta_r}, \quad r=1, \ldots, k,$$

$$L_{rs}^N(\underline{\theta}) = \frac{1}{N} \frac{\partial^2 \ln \phi(\underline{x}, \underline{\theta})}{\partial \theta_r \partial \theta_s} = \frac{1}{N} \sum_{h=1}^{m} \sum_{\alpha=1}^{n_h} \frac{\partial^2 \ln f_h(\underline{x}_{h\alpha}, \underline{\theta})}{\partial \theta_r}, \quad r, s=1, \ldots, k,$$

$$L_{rst}^N(\underline{\theta}) = \frac{1}{N} \frac{\partial^3 \ln \phi(\underline{x}, \underline{\theta})}{\partial \theta_r \partial \theta_s \partial \theta_t} = \frac{1}{N} \sum_{h=1}^{m} \sum_{\alpha=1}^{n_h} \frac{\partial^3 \ln f_h(\underline{x}_{h\alpha}, \underline{\theta})}{\partial \theta_r \partial \theta_s \partial \theta_t}, \quad r, s, t=1, \ldots, k.$$

Then one may obtain the following generalization of Lemma 1.3 by simply noting that a finite linear combination of continuous functions is itself continuous.

Lemma 1.3'. Under Assumptions A' and B' and for the given sequence $\{\underline{\theta}^N\}$, the following hold for all $\underline{\theta}$ such that $\|\underline{\theta} - \underline{\theta}^*\| \leq 1/2\, \nu_2$:

(i) $$\underset{N \to \infty}{\text{p lim}} [L_r^N(\underline{\theta}) | \underline{\theta}^N] = \sum_{h=1}^{m} \mu_h E_{\underline{\theta}^*} \left[\frac{\partial \ln f_h}{\partial \theta_r} \Big|_{\underline{\theta}} \right], \quad r=1, \ldots, k,$$

(ii) $$\underset{N \to \infty}{\text{p lim}} [L_{rs}^N(\underline{\theta}) | \underline{\theta}^N] = \sum_{h=1}^{m} \mu_h E_{\underline{\theta}^*} \left[\frac{\partial^2 \ln f_h}{\partial \theta_r \partial \theta_s} \Big|_{\underline{\theta}} \right], \quad r, s=1, \ldots, k.$$

The remaining results generalize directly from those given in previous sections when $C_{rs}(\underline{\theta})$ as defined in Assumption A3 is replaced by the definition given in A'3, namely

$$C_{rs}(\underline{\theta}) = \sum_{h=1}^{m} \mu_h E_{\underline{\theta}} \left[\frac{\partial \ln f_h}{\partial \theta_r} \bigg|_{\underline{\theta}} \frac{\partial \ln f_h}{\partial \theta_s} \bigg|_{\underline{\theta}}\right] = -\sum_{h=1}^{m} \mu_h E_{\underline{\theta}} \left[\frac{\partial^2 \ln f_h}{\partial \theta_r \partial \theta_s} \bigg|_{\underline{\theta}}\right].$$

These extensions hinge on the fact that one is dealing with finite linear combinations of quantities which satisfy the assumptions which were sufficient for the results for a single population. In particular one has the following generalization of Theorem 2.1.

Theorem 2.1'. Let $\{\underline{\theta}^N\}$ be the sequence of local alternatives with $\theta_i^N = \theta_i^o + \delta_{iN}/\sqrt{N}$, where as $n_h = \mu_h N$, $h = 1,\ldots,m$ and $N \to \infty$, $\lim_{N \to \infty} \delta_{iN} = \delta_i$, $i = 1,\ldots,\ell$, and with $\theta_i^N = \theta_i^*$, $i = \ell + 1,\ldots k$, where ${}_2\underline{\theta}^*$ is the vector of true values of the nuisance parameters. Let $F_\ell(\overline{\lambda}^2, t)$ be the distribution function of the noncentral chi-square distribution with ℓ degrees of freedom and noncentrality parameter $\overline{\lambda}^2$. If λ_N is the likelihood ratio statistic for testing the hypothesis H_o: $\underline{\theta} = \underline{\theta}^\omega$ where $\underline{\theta}^\omega = ({}_1\underline{\theta}^o, {}_2\underline{\theta}) \in \Omega$, then under Assumptions A', B', and C',

$$\lim_{N \to \infty} P[-2\ln\lambda_N \leq t \,|\, \underline{\theta}^N] = F_\ell(\overline{\lambda}^2, t),$$

where $N = \sum_{h=1}^{m} n_h$, $\overline{\lambda}^2 = {}_1\underline{\delta}\, \underline{\Sigma}_{11}^{-1}\, {}_1\underline{\delta}' = {}_1\underline{\delta}\, \underline{C}_{11}\, {}_1\underline{\delta}'$ if the hypothesis to be tested is false, and $\overline{\lambda}^2 = 0$ if the hypothesis to be tested is true.

2.3 Transformations on the parameter space

In the discussion thus far attention has been restricted to tests of a null hypothesis in which certain components of a parameter vector $\underline{\theta}$ have been specified. In this section the asymptotic theory will be generalized to tests of a more general hypothesis of the form

$$H_o: \underline{\theta}\epsilon\omega$$

where ω is the space of all $\underline{\theta}\epsilon\Omega$ for which there is a transformation $\underline{\xi}(\underline{\theta})$ such that

$$_1\underline{\xi}(\underline{\theta}) = (\xi_1(\underline{\theta}),\ldots,\xi_\ell(\underline{\theta})) = {}_1\underline{\xi}^o = (\xi_1^o,\ldots,\xi_\ell^o)$$

where $\xi_1^o,\ldots,\xi_\ell^o$, $\ell \leq k$, are constants. In addition the transformation $\underline{\xi}(\underline{\theta})$ will be required to satisfy the following peoperties:

(a) There exists a vector

$$_2\underline{\xi}(\underline{\theta}) = (\xi_{\ell+1}(\underline{\theta}),\ldots,\xi_k(\underline{\theta}))$$

such that the inverse relationships

$$\underline{\theta}(\underline{\xi}) = (\theta_1(\underline{\xi}),\ldots,\theta_k(\underline{\xi}))$$

exist, where $\underline{\xi} = \underline{\xi}(\underline{\theta}) = ({}_1\underline{\xi}(\underline{\theta}),{}_2\underline{\xi}(\underline{\theta}))$.

(b) The first-, second- and third-order partial derivatives of $\theta_r(\underline{\xi})$ with respect to $\underline{\xi}$ exist, and there exist positive real numbers R_{ri}, R_{rij} and R_{rijh} such that

$$\left|\frac{\partial\theta_r(\underline{\xi})}{\partial\xi_i}\right| < R_{ri},$$

$$\left|\frac{\partial^2\theta_r(\underline{\xi})}{\partial\xi_i\partial\xi_j}\right| < R_{rij},$$

$$\left|\frac{\partial^3\theta_r(\underline{\xi})}{\partial\xi_i\,\partial\xi_j\,\partial\xi_h}\right| < R_{rijh},$$

those of Section 1.2 or assumptions which imply the same results, then Theorem 2.1, the main result of interest in this study, may be generalized to the situation presently being considered as follows:

Theorem 2.1^T. Let λ_n^T be the likelihood ratio statistic for testing the hypothesis H_o: $\underline{\xi} = \underline{\xi}^\omega$ against the local alternatives $\{\underline{\xi}^n\}$, where $\underline{\xi}^\omega = ({}_1\underline{\xi}^o, {}_2\underline{\xi})$ and $\underline{\xi}^n = ({}_1\underline{\xi}^n, {}_2\underline{\xi}^*) \epsilon\ \Omega^T$, and where the transformation $\underline{\xi} = \underline{\xi}(\underline{\theta})$ satisfies properties (a), (b), and (c). Let $F_\ell(\bar{\lambda}^2, t)$ be the distribution function of the noncentral chi-square distribution with ℓ degrees of freedom and noncentrality parameter $\bar{\lambda}^2$. Then under Assumptions A, B, and C

$$\lim_{n \to \infty} P[-2\ln\lambda_n^T \leq t|\underline{\xi}^n] = F_\ell(\bar{\lambda}^2, t)$$

where $\bar{\lambda}^2 = {}_1\underline{\delta}^T\, \underline{\Sigma}_{11}^{T^{-1}}\, {}_1\underline{\delta}^{T'}$ with ${}_1\underline{\delta}^T = (\delta_1^T, \ldots, \delta_\ell^T)$, $\underline{\Sigma}^T = \underline{C}^{T^{-1}}$ and

$$\underline{C}^T = \left[- E_{\underline{\xi}^*} \left\{ \frac{\partial^2 \ln f^T}{\partial \xi_i\, \partial \xi_j}\bigg|_{\underline{\xi}^*} \right\} ;\ i,j = 1, \ldots, k \right] .$$

The properties of $f^T(\underline{x}, \underline{\xi})$ will now be discussed in general and then given a detailed development. Let A^T, B^T, and C^T denote the assumptions on $f^T(\underline{x}, \underline{\xi})$ which correspond to Assumptions A, B, and C on $f(\underline{x}, \underline{\theta})$. It may be established, using Assumptions A, B, and C and the properties (a), (b), and (c) of the transformation $\underline{\xi}(\underline{\theta})$, that Assumptions $A1^T$, $A2^T$, $A3^T$, B^T, and C^T are satisfied. However, it cannot be shown that Assumption $A4^T$ is satisfied in that the existence of a function $H_{ijh}^T(\underline{x})$ corresponding to $H_{rst}(\underline{x})$ cannot be guaranteed. In the first four sections of this study seven lemmas and three theorems are given and of these only Lemmas 1.2 and 1.6 and for all $\underline{\theta} \epsilon \Omega$ and r, i, j, h = 1,...,k.

(c) The greatest lower bound, with respect to $\underline{\theta} \epsilon \Omega$, of the absolute value of the Jacobian

$$\partial(\xi_1,\ldots,\xi_k)/\partial(\theta_1,\ldots,\theta_k)$$

is positive.

The preceding assumptions on the transformation $\underline{\xi}(\underline{\theta})$ are identical with those given by Bradley and Gart [1962, p. 209] and Wald [1943, p. 463], with the exception of the assumption on the third partial derivatives of $\underline{\theta}(\underline{\xi})$, which have been added.

Under the transformation $\underline{\xi} = \underline{\xi}(\underline{\theta})$, the null hypothesis H_o can be expressed as

$$H_o: \underline{\xi} = \underline{\xi}^{\omega} = ({}_1\underline{\xi}^o, {}_2\underline{\xi})$$

where ${}_2\underline{\xi}$ is unspecified. As in earlier sections attention is restricted to the following class of local alternatives: $\{\underline{\xi}^n\}$, a sequence of "true" values of $\underline{\xi}$ such that $\underline{\xi}^n = ({}_1\underline{\xi}^n, {}_2\underline{\xi}^*)$ where $\xi_i^n = \xi_i^o + \delta_{in}^T/\sqrt{n}$ with $\lim\limits_{n \to \infty} \delta_{in}^T = \delta_i^T$, $i = 1,\ldots,\ell$, and where ${}_2\underline{\xi}^*$ is the vector of true values of ${}_2\underline{\xi}$. Setting $\xi^* = ({}_1\underline{\xi}^o, {}_2\underline{\xi}^*)$, it is seen that $\lim\limits_{n \to \infty} \underline{\xi}^n = \underline{\xi}^*$. Now the hypothesis $H_o: \underline{\xi} = \underline{\xi}^{\omega}$ and the sequence of local alternatives $\{\underline{\xi}^n\}$ are in the same form as those given in Section 1.3. Again let $f(\underline{x},\underline{\theta})$, the density or probability function of the random vector $\underline{x}$, satisfy Assumptions A, B and C of Section 1.2, and let $f^T(\underline{x},\underline{\xi})$, $\underline{\xi} \epsilon \Omega^T$, be the transformed density or probability function where Ω^T, the parameter space of $\underline{\xi}$, is the image of Ω under the transformation $\underline{\xi} = \underline{\xi}(\underline{\theta})$. If $f^T(\underline{x},\underline{\xi})$ can be shown to satisfy assumptions similar to

Theorem 1.1 require the use of Assumption A4. In fact, the proofs of Lemma 1.2 and Theorem 1.1 only require the property that

$$E_{\underline{\theta}'}\left[\left|\frac{\partial^3 \ln f}{\partial\theta_r\,\partial\theta_s\,\partial\theta_t}\right|_{\underline{\theta}''}\right] < M \tag{2.3.1}$$

for all $\underline{\theta}'$, $\underline{\theta}'' \in \Omega$ such that $\|\underline{\theta}' - \underline{\theta}''\| < \nu_2$, where ν_2 is defined in Assumption B. This condition may be shown to hold in the transformed case for $f^T(\underline{x},\underline{\xi})$, namely

$$E_{\underline{\xi}'}\left[\left|\frac{\partial^3 \ln f^T}{\partial\xi_i\,\partial\xi_j\,\partial\xi_h}\right|_{\underline{\xi}''}\right] < M^T \tag{2.3.2}$$

for all $\underline{\xi}'$, $\underline{\xi}'' \in \Omega^T$ such that $\|\underline{\xi}' - \underline{\xi}''\| < \nu_2^T$, where ν_2^T is such that $\|\underline{\theta}(\underline{\xi}') - \underline{\theta}(\underline{\xi}'')\| < \nu_2$ whenever $\|\underline{\xi}' - \underline{\xi}''\| < \nu_2^T$. Assumption A4 is required in its entirety in the proof of Lemma 1.6 but it is noted that Lemma 1.6 is used only in the proof of Lemma 1.7. However, if Lemma 1.7^T, the result which corresponds to Lemma 1.7, can be established independently by use of Assumptions A, B, and C and the properties (a), (b), and (c) of the transformation $\underline{\xi}(\underline{\theta})$, then Assumption $A4^T$ will not be required. That this is in fact the case will be demonstrated.

It will first be shown that $f^T(\underline{x},\underline{\xi})$ satisfies Assumptions $A1^T$, $A2^T$, $A3^T$, B^T, and C^T, and the condition given by Equation (2.3.2). Consider the following relationships:

$$\frac{\partial \ln f^T}{\partial \xi_i} = \sum_{r=1}^{k} \frac{\partial\theta_r}{\partial\xi_i}\frac{\partial \ln f}{\partial\theta_r}\ , \tag{2.3.3}$$

$$\frac{\partial^2 \ln f^T}{\partial\xi_i \partial\xi_j} = \sum_{r=1}^{k} \frac{\partial^2\theta_r}{\partial\xi_i\partial\xi_j} \frac{\partial \ln f}{\partial\theta_r} + \sum_{r=1}^{k}\sum_{s=1}^{k} \frac{\partial\theta_r}{\partial\xi_i}\frac{\partial\theta_s}{\partial\xi_j}\frac{\partial^2 \ln f}{\partial\theta_r\partial\theta_s}, \tag{2.3.4}$$

and

$$\frac{\partial^3 \ln f^T}{\partial\xi_i\partial\xi_j\partial\xi_h}$$

$$= \sum_{r=1}^{k} \frac{\partial^3\theta_r}{\partial\xi_i\partial\xi_j\partial\xi_h} \frac{\partial \ln f}{\partial\theta_r} + \sum_{r=1}^{k}\sum_{s=1}^{k}\sum_{t=1}^{k} \frac{\partial\theta_r}{\partial\xi_i}\frac{\partial\theta_s}{\partial\xi_j}\frac{\partial\theta_t}{\partial\xi_h}\frac{\partial^3 \ln f}{\partial\theta_r\partial\theta_s\partial\theta_t} \tag{2.3.5}$$

$$+ \sum_{r=1}^{k}\sum_{s=1}^{k}\left[\frac{\partial^2\theta_r}{\partial\xi_i\partial\xi_j}\frac{\partial\theta_s}{\partial\xi_h} + \frac{\partial^2\theta_r}{\partial\xi_i\partial\xi_h}\frac{\partial\theta_s}{\partial\xi_j} + \frac{\partial^2\theta_r}{\partial\xi_j\partial\xi_h}\frac{\partial\theta_s}{\partial\xi_i}\right]\frac{\partial^2 \ln f}{\partial\theta_r\partial\theta_s} .$$

By Assumption A1 on $f(\underline{x},\underline{\theta})$ and property (b) of the transformation $\underline{\xi} = \underline{\xi}(\underline{\theta})$, the above are linear combinations of partial derivatives which exist. Thus $f^T(\underline{x},\underline{\xi})$ satisfies Assumption A1^T.

To satisfy Assumption A2^T, $f^T(\underline{x},\underline{\xi})$ must be such that

$$\left|\frac{\partial f^T}{\partial\xi_i}\right| < F_i^T(\underline{x}) \quad\text{and}\quad \left|\frac{\partial^2 f^T}{\partial\xi_i\partial\xi_j}\right| < F_{ij}^T(\underline{x})$$

for almost all $\underline{x}\epsilon R$ and every $\underline{\xi}\epsilon\Omega^T$, where $F_i^T(\underline{x})$ and $F_{ij}^T(\underline{x})$ are integrable over R, i, j = 1,...,k. Now by Assumption A2 on $f(\underline{x},\underline{\theta})$ and property (b) on $\underline{\xi} = \underline{\xi}(\underline{\theta})$, one obtains, by use of expressions analogous to (1.6.3) and (1.6.4), that

$$\left|\frac{\partial f^T}{\partial\xi_i}\right| \le \sum_{r=1}^{k} \left|\frac{\partial\theta_r}{\partial\xi_i}\right|\left|\frac{\partial f}{\partial\theta_r}\right| < \sum_{r=1}^{k} R_{ri}F_r(\underline{x}) = F_i^T(\underline{x}),$$

and

$$\left|\frac{\partial^2 f^T}{\partial\xi_i\partial\xi_j}\right| \leq \sum_{r=1}^{k} \left|\frac{\partial^2\theta_r}{\partial\xi_i\partial\xi_j}\right| \left|\frac{\partial f}{\partial\theta_r}\right| + \sum_{r=1}^{k}\sum_{s=1}^{k} \left|\frac{\partial\theta_r}{\partial\xi_i}\right| \left|\frac{\partial\theta_s}{\partial\xi_j}\right| \left|\frac{\partial^2 f}{\partial\theta_r\partial\theta_s}\right|$$

$$< \sum_{r=1}^{k} R_{rij}F_r(\underline{x}) + \sum_{r=1}^{k}\sum_{s=1}^{k} R_{ri}R_{sj}F_{rs}(\underline{x}) = F_{ij}^T(\underline{x}).$$

The functions $F_i^T(\underline{x})$ and $F_{ij}^T(\underline{x})$ are integrable over R, since they are linear combinations of functions which are integrable over R.

To satisfy Assumption A3^T, the matrix $\underline{C}^T(\underline{\xi}) = [C_{ij}^T(\underline{\xi});\ i,j = 1,\ldots,k]$ with

$$C_{ij}^T(\underline{\xi}) = E_{\underline{\xi}}\left[\left.\frac{\partial \ln f^T}{\partial\xi_i}\right|_{\underline{\xi}} \left.\frac{\partial \ln f^T}{\partial\xi_j}\right|_{\underline{\xi}}\right]$$

must be positive definite with finite determinant. Now

$$C_{ij}^T(\underline{\xi}) = \sum_{r=1}^{k}\sum_{s=1}^{k} \left.\frac{\partial\theta_r}{\partial\xi_i}\right|_{\underline{\theta}} \left.\frac{\partial\theta_s}{\partial\xi_j}\right|_{\underline{\theta}} C_{rs}(\underline{\theta}). \qquad (2.3.6)$$

Then,

$$\underline{C}^T(\underline{\xi}) = \underline{D}'\underline{C}(\underline{\theta})\underline{D}$$

where

$$\underline{D} = \left[\left.\frac{\partial\theta_r}{\partial\xi_i}\right|_{\underline{\theta}};\ r,\ i = 1,\ldots,k\right].$$

Since $\underline{C}(\underline{\theta})$ is positive definite with finite determinant by Assumption A3 and $\underline{D}$ is nonsingular with finite determinant by properties (b) and (c), thus $\underline{C}^T(\underline{\xi})$ is positive definite with finite determinant.

For $f^T(\underline{x},\underline{\xi})$ to satisfy Assumption B^T there must exist a positive real number ν_2^T such that whenever $\|\underline{\xi}' - \underline{\xi}''\| < \nu_2^T$, $\underline{\xi}', \underline{\xi}'' \in \Omega^T$,

$$E_{\underline{\xi}'}\left[\left(\frac{\partial^2 \ln f^T}{\partial\xi_i \partial\xi_j}\bigg|_{\underline{\xi}''}\right)^2\right] < U < \infty$$

for i, j = 1,...,k. Now, with use of (2.3.4) one obtains

$$E_{\underline{\xi}'}\left[\left(\frac{\partial^2 \ln f^T}{\partial\xi_i \partial\xi_j}\bigg|_{\underline{\xi}''}\right)^2\right]$$

$$= E_{\underline{\theta}'}\left[\left(\sum_{r=1}^{k} \frac{\partial^2\theta_r}{\partial\xi_i\partial\xi_j}\bigg|_{\underline{\theta}''} \frac{\partial \ln f}{\partial\theta_r}\bigg|_{\underline{\theta}''} + \sum_{r=1}^{k}\sum_{s=1}^{k} \frac{\partial\theta_r}{\partial\xi_i}\bigg|_{\underline{\theta}''} \frac{\partial\theta_s}{\partial\xi_j}\bigg|_{\underline{\theta}''} \frac{\partial^2 \ln f}{\partial\theta_r\partial\theta_s}\bigg|_{\underline{\theta}''}\right)^2\right]$$

Due to the continuity of $\underline{\theta}(\underline{\xi})$, there exists a positive real number ν_2^T such that $\|\underline{\theta}' - \underline{\theta}''\| < \nu_2$ whenever $\|\underline{\xi}' - \underline{\xi}''\| < \nu_2^T$, where ν_2 is defined in Assumption B. By multiple use of the inequality $(\sum_{i=1}^{m} a_i)^2 \leq m \sum_{i=1}^{m} a_i^2$, first with m = 2 and then with m = k and $m = k^2$ respectively, one obtains

$$E_{\underline{\xi}'}\left[\left(\frac{\partial^2 \ln f^T}{\partial\xi_i\partial\xi_j}\bigg|_{\underline{\xi}''}\right)^2\right]$$

$$\leq 2E_{\underline{\theta}'}\left[\left(\sum_{r=1}^{k}\frac{\partial^2\theta_r}{\partial\xi_i\partial\xi_j}\bigg|_{\underline{\theta}''}\frac{\partial \ln f}{\partial\theta_r}\bigg|_{\underline{\theta}''}\right)^2\right]$$

$$+\ 2E_{\underline{\theta}'}\left[\left(\sum_{r=1}^{k}\sum_{s=1}^{k}\frac{\partial\theta_r}{\partial\xi_i}\bigg|_{\underline{\theta}''}\frac{\partial\theta_s}{\partial\xi_j}\bigg|_{\underline{\theta}''}\frac{\partial^2 \ln f}{\partial\theta_r\partial\theta_s}\bigg|_{\underline{\theta}''}\right)^2\right]$$

$$\leq 2k\sum_{r=1}^{k}\left(\frac{\partial^2\theta_r}{\partial\xi_i\partial\xi_j}\bigg|_{\underline{\theta}''}\right)^2 E_{\underline{\theta}'}\left[\left(\frac{\partial \ln f}{\partial\theta_r}\bigg|_{\underline{\theta}''}\right)^2\right]$$

$$+\ 2k^2\sum_{r=1}^{k}\sum_{s=1}^{k}\left(\frac{\partial\theta_r}{\partial\xi_i}\bigg|_{\underline{\theta}''}\frac{\partial\theta_s}{\partial\xi_j}\bigg|_{\underline{\theta}''}\right)^2 E_{\underline{\theta}'}\left[\left(\frac{\partial^2 \ln f}{\partial\theta_r\partial\theta_s}\bigg|_{\underline{\theta}''}\right)^2\right]$$

$$< 2k\sum_{r=1}^{k}R^2_{rij}\ S + 2k^2\sum_{r=1}^{k}\sum_{s=1}^{k}R^2_{ri.}R^2_{sj.}T = U < \infty,$$

where the last inequality follows from Assumption B, Lemma 1.1, and property (b), for $\|\underline{\xi}' - \underline{\xi}''\| < \nu_2^T$. Thus $f^T(\underline{x},\underline{\xi})$ satisfies Assumption B^T.

For Assumption C^T to hold there must exist a positive real number $\bar{\eta}_2^T$ such that

$$E_{\underline{\xi}}\left[\left|\frac{\partial \ln f^T}{\partial\xi_i}\bigg|_{\underline{\xi}}\right|^{2+\bar{\eta}_2^T}\right] < K^T < \infty,$$

$i = 1,\ldots,k$, for all $\underline{\xi}\epsilon\Omega^T$. Let

$$\rho(\underline{x},\underline{\theta}) = \max_r \left\{ \left| \frac{\partial \ln f}{\partial \theta_r} \right| \right\} .$$

Then, if $\bar{\eta}_2^T = \bar{\eta}_2$, one obtains by use of (2.3.5)

$$E_{\underline{\xi}} \left[\left| \frac{\partial \ln f^T}{\partial \xi_i} \Big|_{\underline{\xi}} \right|^{2+\bar{\eta}_2} \right] \leq \left(\sum_{r=1}^{k} \left| \frac{\partial \theta_r}{\partial \xi_i} \right| \right)^{2+\bar{\eta}_2} E_{\underline{\theta}} \left[\{\rho(\underline{x},\underline{\theta})\}^{2+\bar{\eta}_2} \right]$$

$$< \left(\sum_{r=1}^{k} R_{ri} \right)^{2+\bar{\eta}_2} K = K^T < \infty,$$

where the last inequality follows from Assumption C and property (b). Thus $f^T(\underline{x},\underline{\xi})$ satisfies Assumption C^T.

As was indicated earlier, $f^T(\underline{x},\underline{\xi})$ does not necessarily satisfy Assumption $A4^T$ but can be shown to satisfy Equation (2.3.2). Note that by use of (2.3.5) one obtains

$$E_{\underline{\xi}'} \left[\left| \frac{\partial^3 \ln f^T}{\partial \xi_i \partial \xi_j \partial \xi_h} \Big|_{\underline{\xi}''} \right| \right]$$

$$\leq \sum_{r=1}^{k} \left| \frac{\partial^3 \theta_r}{\partial \xi_i \partial \xi_j \partial \xi_h} \Big|_{\underline{\theta}''} \right| E_{\underline{\theta}'} \left[\left| \frac{\partial \ln f}{\partial \theta_r} \Big|_{\underline{\theta}''} \right| \right]$$

$$+ \sum_{r=1}^{k} \sum_{s=1}^{k} \sum_{t=1}^{k} \left| \frac{\partial \theta_r}{\partial \xi_i} \Big|_{\underline{\theta}''} \frac{\partial \theta_s}{\partial \xi_j} \Big|_{\underline{\theta}''} \frac{\partial \theta_t}{\partial \xi_h} \Big|_{\underline{\theta}''} \right| E_{\underline{\theta}'} \left[\left| \frac{\partial^3 \ln f}{\partial \theta_r \partial \theta_s \partial \theta_t} \Big|_{\underline{\theta}''} \right| \right]$$

$$+ \sum_{r=1}^{k} \sum_{s=1}^{k} \left[\left| \frac{\partial^2 \theta_r}{\partial \xi_i \partial \xi_j} \Big|_{\underline{\theta}''} \frac{\partial \theta_s}{\partial \xi_h} \Big|_{\underline{\theta}''} \right| + \left| \frac{\partial^2 \theta_r}{\partial \xi_i \partial \xi_h} \Big|_{\underline{\theta}''} \frac{\partial \theta_s}{\partial \xi_j} \Big|_{\underline{\theta}''} \right| + \left| \frac{\partial^2 \theta_r}{\partial \xi_j \partial \xi_h} \Big|_{\underline{\theta}''} \frac{\partial \theta_s}{\partial \xi_i} \Big|_{\underline{\theta}''} \right| \right]$$

$$\cdot E_{\underline{\theta}'} \left[\left| \frac{\partial^2 \ln f}{\partial \theta_r \partial \theta_s} \Big|_{\underline{\theta}''} \right| \right]$$

where $\underline{\theta}' = \underline{\theta}(\underline{\xi}')$ and $\underline{\theta}'' = \underline{\theta}(\underline{\xi}'')$. For all $\underline{\xi}'$, $\underline{\xi}'' \epsilon \Omega^T$ such that $\|\underline{\xi}' - \underline{\xi}''\| < \nu_2^T$, where by the continuity of $\underline{\theta}(\underline{\xi})$, ν_2^T is such that $\|\underline{\xi}' - \underline{\xi}''\| < \nu_2^T$ implies $\|\underline{\theta}' - \underline{\theta}''\| < \nu_2$, one has

$$E_{\underline{\xi}'}\left[\left|\frac{\partial^3 \ln f^T}{\partial\xi_i \partial\xi_j \partial\xi_h}\right|_{\underline{\xi}''}\right] < \sum_{r=1}^{k} R_{rijh}\, S^{1/2} + \sum_{r=1}^{k}\sum_{s=1}^{k}\sum_{t=1}^{k} R_{ri}R_{sj}R_{th}M$$

$$+ \sum_{r=1}^{k}\sum_{s=1}^{k} [R_{rij}R_{sh} + R_{rih}R_{sj} + R_{rjh}R_{si}]T^{1/2}$$

$$= M^T,$$

by Assumption A4, B, Lemma 1.1 and property (b). Thus $f^T(\underline{x},\underline{\xi})$ satisfies Equation (2.3.2).

Finally, consider Lemma 1.7^T which is to be proved independently by use of Assumptions A, B, and C and the properties (a), (b), and (c) of the transformation $\underline{\xi}(\underline{\theta})$.

Lemma 1.7^T. Let $\tilde{\underline{\xi}}^n$ be a sequence of parameter values in Ω^T which converge in probability to $\underline{\xi}^*$ under the sequence of local alternatives $\{\underline{\xi}^n\}$. Then given Assumptions A and B on $f(\underline{x},\underline{\theta})$, and properties (a), (b), and (c) of the transformation $\underline{\xi} = \underline{\xi}(\underline{\theta})$,

$$p \lim_{n\to\infty} [L_{ij}^{n^T}(\tilde{\underline{\xi}}^n)|\underline{\xi}^n] = - C_{ij}^T(\underline{\xi}^*), \quad i,j = 1,\ldots,k$$

where

$$L_{ij}^{n^T}(\tilde{\underline{\xi}}^n) = \frac{1}{n}\sum_{\alpha=1}^{n} \left.\frac{\partial^2 \ln f^T(\underline{x}_\alpha,\underline{\xi})}{\partial\xi_i \partial\xi_j}\right|_{\tilde{\underline{\xi}}^n} .$$

Proof: Using (1.3.3), (1.3.4), (1.3.5), and (2.3.4), one can write

$$L_{ij}^{n^T}(\underline{\tilde{\xi}}^n) = \sum_{r=1}^{k} \frac{\partial^2\theta_r}{\partial\xi_i\partial\xi_j}\Big|_{\underline{\tilde{\theta}}^n} L_r^n(\underline{\tilde{\theta}}^n) + \sum_{r=1}^{k}\sum_{s=1}^{k} \frac{\partial\theta_r}{\partial\xi_i}\Big|_{\underline{\tilde{\theta}}^n} \frac{\partial\theta_s}{\partial\xi_j}\Big|_{\underline{\tilde{\theta}}^n} L_{rs}^n(\underline{\tilde{\theta}}^n)$$

where $\underline{\tilde{\theta}}^n = \underline{\theta}(\underline{\tilde{\xi}}^n)$. Let $\underline{\theta}^* = \underline{\theta}(\underline{\xi}^*)$ and $\underline{\theta}^n = \underline{\theta}(\underline{\xi}^n)$. Then by the continuity of $\underline{\theta}(\underline{\xi})$, $\lim_{n\to\infty} \underline{\theta}^n = \underline{\theta}^*$ since $\lim_{n\to\infty} \underline{\xi}^n = \underline{\xi}^*$, and $p\lim_{n\to\infty} \{\underline{\tilde{\theta}}^n|\underline{\theta}^n\} = \underline{\theta}^*$ since $p\lim_{n\to\infty} \{\underline{\tilde{\xi}}^n|\underline{\xi}^n\} = \underline{\xi}^*$.

Expanding $L_r^n(\underline{\tilde{\theta}}^n)$ in Taylor series about $\underline{\theta} = \underline{\theta}^*$ yields

$$L_r^n(\underline{\tilde{\theta}}^n) = L_r^n(\underline{\theta}^*) + \sum_{s=1}^{k} (\tilde{\theta}_s^n - \theta_s^*) L_{rs}^n(\underline{\tilde{\tilde{\theta}}}^n), \quad r = 1,\ldots,k$$

for $\underline{\tilde{\tilde{\theta}}}^n$ such that $\|\underline{\tilde{\tilde{\theta}}}^n - \underline{\theta}^*\| \le \|\underline{\tilde{\theta}}^n - \underline{\theta}^*\|$. Now $p\lim_{n\to\infty} \{\underline{\tilde{\theta}}^n|\underline{\theta}^n\} = \underline{\theta}^*$ implies $p\lim_{n\to\infty} \{\underline{\tilde{\tilde{\theta}}}^n|\underline{\theta}^n\} = \underline{\theta}^*$, and thus

$$p\lim_{n\to\infty} \{L_{rs}^n(\underline{\tilde{\tilde{\theta}}}^n) \mid \underline{\theta}^n\} = -C_{rs}(\underline{\theta}^*)$$

by Lemma 1.7. Also, by Lemma 1.3, one has

$$p\lim_{n\to\infty} \{L_r^n(\underline{\theta}^*)\,|\,\underline{\theta}^n\} = A_r(\underline{\theta}^*,\underline{\theta}^*) = 0.$$

Thus by Slutsky's theorem

$$p\lim_{n\to\infty} \{L_r^n(\underline{\tilde{\theta}}^n)\,|\,\underline{\theta}^n\} = 0, \; r = 1,\ldots,k.$$

In addition

$$p\lim_{n\to\infty} \{L_{rs}^n(\underline{\tilde{\theta}}^n)\,|\,\underline{\theta}^n\} = -C_{rs}(\underline{\theta}^*), \quad r,s = 1,\ldots,k$$

by Lemma 1.7. Using (1.6.6) and Slutsky's theorem, one can conclude that

$$\underset{n\to\infty}{\text{p lim}} \{L_{ij}^{n^T}(\tilde{\underline{\xi}}^n)\,|\,\underline{\xi}^n\} = -\, C_{ij}^T(\underline{\xi}^*), \quad i,j = 1,\ldots,k,$$

thus completing the proof of Lemma 1.7^T

This completes the discussion of the conditions necessary for Theorem 2.1^T. The discussion of this section could also be extended to the case of associated populations and would lead to the appropriate generalization of Theorem 2.1' of Section 2.2.

N 68-27634

PART 2

III. INTRODUCTION

The purpose of this part of this dissertation is to give possible methods of making decisions about the unknown noncentrality parameters of noncentral chi-square distributions. In each of the decision procedures considered, the noncentrality parameters are assumed to be random variables with continuous distributions.

The notation $\bar{\lambda}^2$ used for the noncentrality parameter in Part 1 will be simplified to λ in Part 2. This simplification is introduced since the noncentrality parameter is referred to many times in Part 2.

The decision procedures are developed in Chapter IV. In Section 4.2, a two-decision procedure is considered on the magnitude of the noncentrality parameter λ, where the two decision regions are the two sets $\{\lambda | 0 \leq \lambda < \tau\}$ and $\{\lambda | \tau \leq \lambda < \infty\}$, and τ is a specified constant. In Section 4.3, two extensions of the work in Section 4.2 are considered. One extension is the combination of p two-decision procedures on the magnitudes of the noncentrality parameters $\lambda_1, \ldots, \lambda_p$. If $\omega_{0i} = \{(\lambda_1, \ldots, \lambda_p) | 0 \leq \lambda_i < \tau_i\}$ and $\omega_{1i} = \{(\lambda_1, \ldots, \lambda_p) | \tau_i \leq \lambda_i < \infty\}$, $i = 1, .., p$, the combination of decision procedures yields 2^p possible decisions of the form $\bigcap_{i=1}^{p} \omega_{j_i i}$, $j_i = 0, 1$, $i = 1, \ldots, p$. This combined decision procedure is developed through the use of a paper by Lehmann [1957]. The second extension is a q-decision procedure based on one noncentrality parameter, giving the q decision regions $\{\lambda | \tau_i^* \leq \lambda < \tau_{i+1}^*\}$, $i = 0, \ldots, q - 1$, where the τ_i^*, $i = 0, \ldots, q$ are specified constants, such that $\tau_i^* < \tau_{i+1}^*$, $\tau_0^* = 0$ and $\tau_q^* = +\infty$. This decision procedure is developed through the use of a paper by Karlin and Rubin [1956].

The decision procedure given in Section 4.4 is a comparison of the magnitudes of two noncentrality parameters λ_1 and λ_2, through the three decision regions: (1) $\{(\lambda_1,\lambda_2)|\lambda_1 > \lambda_2 + \tau\}$, (2) $\{(\lambda_1,\lambda_2)||\lambda_1 - \lambda_2| < \tau\}$ and (3) $\{(\lambda_1,\lambda_2)|\lambda_2 > \lambda_1 + \tau\}$, where τ is a finite positive constant. This decision procedure is obtained through the combination of two two-decision procedures which have decision regions $\{(\lambda_1,\lambda_2)|\lambda_1 < \lambda_2 + \tau\}$, $\{(\lambda_1,\lambda_2)|\lambda_1 \geq \lambda_2 + \tau\}$, $\{(\lambda_1,\lambda_2)|\lambda_2 < \lambda_1 + \tau\}$ and $\{(\lambda_1,\lambda_2)|\lambda_2 \geq \lambda_1 + \tau\}$. In all these decision procedures, the loss functions are assumed to be additive, continuous and monotone in λ for each of the decision procedures. Also when more than one noncentrality parameter is considered, their distributions are considered to be independent.

Chapter V is concerned with the presentation of aids which will help the experimenter in his use of the decision procedures developed in Chapter IV, when a gamma prior distribution is assumed for the noncentrality parameter. Tables to facilitate the use of the decision procedures are described in Section 5.3 and are displayed in the Appendix. These tables give the critical values which are needed to conduct the decision procedure developed in Chapter IV and they also give a set of values from which the new approximate parameters can be obtained, when it is approximated by a gamma distribution.

In order to use these tables, it is necessary to assume that the loss functions are linear in the noncentrality parameter and zero when the decision which is made is true. The latter restriction means the loss functions are actually regret functions. An approximate method for finding the optimal sample size for each of the decision procedures developed in Chapter IV is discussed in Section 5.4.

Chapter VI is concerned with the situation when the distribution of the statistic of concern is only known to have a noncentral chi-square distribution asymptotically. It is shown there the risk functions obtained by using the decision procedure based on the exact noncentral chi-square distribution converge to the minimum risk functions of the exact noncentral chi-square distribution.

Chapter VII contains a discussion of two examples of the decision procedures developed in Chapter IV one which deals with an exact noncentral chi-square distribution and the other is concerned with a distribution which is asymptotically a noncentral chi-square distribution. Chapter VIII contains a summary and evaluation of the investigation reported in this part of the dissertation.

IV. DECISION PROCEDURES

4.1 Introduction

This chapter will be concerned with four types of decision procedures dealing with one or more noncentrality parameters of noncentral chi-square distributions, three of them will be concerned with the magnitude of noncentrality parameters in relation to specified tolerances and the fourth will be concerned with the relative magnitude of two noncentrality parameters.

Throughout this chapter, whenever more than one noncentral chi-square random variable is considered, these random variables will be assumed to be independent. When U is a noncentral chi-square random variable whose distribution has noncentrality parameter λ and S degrees of freedom, $g(U|\lambda,S)$ will denote its probability density function.

4.2 The two-decision procedure

This section will deal with the development of a two-decision procedure for the magnitude of the noncentrality parameter. This two-decision procedure will be used to help develop a multiple decision procedure for several noncentrality parameters by use of a method given by Lehmann [1957].

The noncentrality parameter λ given in Section 4.1 will be assumed to be a continuous random variable with cumulative distribution function $F(\lambda) = \int_0^\infty f(t)dt$, where $f(t) > 0$ for $t \in (0,\infty)$. The two-decision procedure $\delta_2(U)$, where U is the noncentral chi-square random variable to be considered, will have the two decisions $d_0: \lambda \epsilon \{\lambda | 0 \leq \lambda < \tau\}$ and $d_1: \lambda \epsilon \{\lambda | \tau \leq \lambda \leq \infty\}$, where τ is a tolerance assigned by the experimenter. The loss function corresponding to $\delta_2(U)$ will be considered to be

$$L(\delta_2(U),\lambda) = \Phi_0(U) L_0(\lambda,\tau) + \Phi_1(U) L_1(\lambda,\tau), \tag{4.2.1}$$

where

(I) $L_0(\lambda,\tau) - L_1(\lambda,\tau) = \begin{cases} < 0, & \lambda < \tau \\ = 0, & \lambda = \tau \\ > 0, & \lambda > \tau; \end{cases}$

(II) $L_0(\lambda,\tau)$ is a continuous non-decreasing function of $\lambda \epsilon [0,\infty)$ and $L_1(\lambda,\tau)$ is a continuous non-increasing function of $\lambda \epsilon [0,\infty)$;

(III) The integral $\int_0^\infty L_i(\lambda,\tau) g(U|\lambda,S) f(\lambda) d\lambda$ exists for $i = 0,1$;

(IV) $\Phi_0(U) + \Phi_1(U) = 1$ for all U;

(V) $\Phi_0(U)$ has a countable number of discontinuities;

(VI) $\Phi_0(U)$ equals one on the open interior of all sets U for which $\delta_2(U) = d_0$ and zero otherwise.

Condition (VI) will not change the form of the optimal decision, since at most a countable number of points would be excluded and these would have measure zero. Conditions (IV), (V) and (VI) imply that $\delta_2(U) = d_0$ when $\Phi_0(U) = 1$ and $\delta_2(U) = d_1$ when $\Phi_1(U) = 1$, except on a set of measure zero. The functions $L_0(\lambda,\tau)$ and $L_1(\lambda,\tau)$ defined in conditions (I), (II) and (III) are the losses incurred when decisions d_0 and d_1 are made, respectively. The restrictions made on $L_0(\lambda,\tau)$ and $L_1(\lambda,\tau)$ are reasonable, since it is desirable that $L_0(\lambda,\tau)$ be less than $L_1(\lambda,\tau)$ when $\lambda < \tau$, that is, when d_0 is the correct decision, and greater than

$L_1(\lambda,\tau)$ when $\lambda > \tau$, so that d_1 is the correct decision. Similarly, the assumptions of monotonicity are reasonable since the difference between $L_0(\lambda,\tau)$ and $L_1(\lambda,\tau)$ would tend to be less when the difference between λ and τ is small than when that difference is large.

The average risk associated with $\delta_2(U)$ is given by

$$R(\delta_2,S,\tau,f) = \int_0^\infty \left[\int_0^\infty \left\{L(\delta_2(U),\lambda)g(U|\lambda,S)f(\lambda)\right\}dU\right]d\lambda.$$

The minimum average risk procedure $\delta_2'(U)$ is determined by the form of $\underline{\Phi}'(U) = (\Phi_0'(U),\Phi_1'(U))$ of $\underline{\Phi}(U) = (\Phi_0(U),\Phi_1(U))$ which minimizes $R(\delta_2,S,\tau,f)$.

Now using the definition of $L(\delta_2(U),\lambda)$ it can be seen that

$$R(\delta_2,S,\tau,f) = \int_0^\infty \int_0^\infty \Phi_0(U)[L_0(\lambda,\tau) - L_1(\lambda,\tau)]g(U|\lambda,S)f(\lambda)dUd\lambda + \text{constant} \qquad (4.2.2)$$

To find $\underline{\Phi}'(U)$, it will be necessary to change the order of integration in (4.2.2). This can be done by use of Tonelli's Theorem, Royden [1963, p. 234], since the conditions (II), (V) and (VI) guarantee that the function $\Phi_0(U)[L_0(\lambda,\tau) - L_1(\lambda,\tau)]$ will be measurable. Thus (4.2.2) can be expressed as

$$R(\delta_2,S,\tau,f) = \int_0^\infty [\Phi_0(U)H(U,\tau)]P(U)dU + \text{constant}, \qquad (4.2.3)$$

where

$$H(U,\tau) = \int_0^\infty [L_0(\lambda,\tau) - L_1(\lambda,\tau)]\pi(\lambda|U)d\lambda, \qquad (4.2.4)$$

$$P(U) = \int_0^\infty g(U|\lambda,S)f(\lambda)d\lambda$$

and

$$\pi(\lambda|U) = g(U|\lambda,S)f(\lambda)/P(U). \qquad (4.2.5)$$

Since $P(U) \geq 0$ for all U, the form of $\Phi_0(U)$ which minimizes (4.2.3) is the form of $\Phi_0(U)$ which minimizes

$$\Phi_0(U) H(U,\tau) \ , \tag{4.2.6}$$

for each value of U. Now it is easily seen that the form of $\Phi_0(U)$ which minimizes (4.2.6) is given by:

$$\Phi_0'(U) = \begin{cases} 1, & H(U,\tau) < 0 \\ 0, & H(U,\tau) \geq 0. \end{cases} \tag{4.2.7}$$

The assignment of the equality in (4.2.7) is arbitrary, since (4.2.6) will be zero when $H(U,\tau) = 0$, so the equality is assigned as to make $\Phi_0'(U) = 1$ only on open sets. This procedure says to make decision d_0 if $H(U,\tau) < 0$ and d_1 if $H(U,\tau) \geq 0$.

There are three possible situations which may occur in (4.2.7): (1) $H(U,\tau) \geq 0$ for all U, (2) $H(U,\tau) < 0$ for all U and (3) $H(U,\tau)$ changes sign at least once in the interval $(0,\infty)$. Situation (1) is when decision d_1 is always made, regardless of the value of the observed U and situation (2) is the case where decision d_0 is always made, regardless of the value of the observed U. In these two situations no experiments need be performed, since the decision of whether to accept d_0 or d_1 can be made without further evidence. In the third situation, the decision will depend on the value of U observed. So from now on, assume that situation (3) prevails; that is, $H(U,\tau)$ changes sign at least once for $U \epsilon (0,\infty)$.

Now for the decision procedure to be practical, a value $U_0 = C$ must exist such that $H(U,\tau) < 0$ for $U < C$ and $H(U,\tau) \geq 0$ for $U \geq C$, so that (4.2.7) can be expressed as

$$\Phi_0'(U) = \begin{cases} 1, & U < C \\ 0, & U \geq C, \end{cases} \tag{4.2.8}$$

where C is such that $H(C,\tau) = 0$. The remainder of this section up to the statement of Theorem 4.1 deals with the use of a paper by Karlin and Rubin [1956] to show that Equation (4.2.8) holds and may be skipped without loss of generality by the reader who is not interested in the details of this demonstration.

Now to verify Equation (4.2.8), use can be made of a paper by Karlin and Rubin [1956] on decision procedures for distributions with monotone likelihood ratios. The definition they give for a distribution to have a monotone likelihood ratio is as follows:

Definition: Let the cumulative distribution of a random variable x, when the true state of nature is described by a parameter β, have the form

$$P(x|\beta) = \int_{-\infty}^{x} p(t|\beta)\,dt, \tag{4.2.9}$$

where if $x_1 > x_2$ and $\beta_1 > \beta_2$, then the pdf satisfies

$$p(x_1|\beta_1)p(x_2|\beta_2) - p(x_1|\beta_2)p(x_2|\beta_1) \geq 0. \tag{4.2.10}$$

Any distribution of the form (4.2.9) which satisfies (4.2.10) will be said to have a monotone likelihood ratio.

The fact that the noncentral chi-square distribution has a monotone likelihood ratio is known to others, as indicated by Karlin [1956], but the author has not been able to locate a proof of this result in print, so consider the proof of the following lemma:

Lemma 4.1: The noncentral chi-square distribution has a monotone likelihood ratio.

Proof: Consider the following form of $g(U|\lambda,S)$, the pdf of the noncentral chi-square distribution:

$$g(U|\lambda,S) = \frac{e^{-1/2(U+\lambda)}}{2^{1/2S}} \sum_{j=0}^{\infty} \frac{U^{1/2S + j-1}\lambda^j}{\Gamma(1/2S+j)j!2^{2j}} . \qquad (4.2.11)$$

Now let $U_1 > U_2$ and $\lambda_1 > \lambda_2$ and also let

$$I(U_1,U_2,\lambda_1,\lambda_2) = e^{-1/2(U_1+U_2+\lambda_1+\lambda_2)}(U_1U_2)^{1/2S - 1}/2^S \qquad (4.2.12)$$

and

$$J(j,k) = \Gamma(1/2S + j)\,\Gamma(1/2S + k)j!k!\ 2^{2(j+k)} .$$

Note that $J(j,k) = J(k,j)$. Then putting $p(x|\beta) = g(U|\lambda,S)$, (4.2.10) becomes

$$g(U_1|\lambda_1,S)g(U_2|\lambda_2,S) - g(U_1|\lambda_2,S)g(U_2|\lambda_1,S)$$

$$= I(U_1,U_2,\lambda_1,\lambda_2)\left\{\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{U_1^{\,j}U_2^{\,k}}{J(j,k)}\left[\lambda_1^{\,j}\lambda_2^{\,k} - \lambda_1^{\,k}\lambda_2^{\,j}\right]\right\}. \qquad (4.2.13)$$

The double sum in (4.2.13) can be broken into two parts along the values of $j = k$. Then by renaming the subscripts and letting $\alpha = k-j$ for $k \geq j$ and using the fact that $J(j,k) = J(k,j)$, (4.2.13) can be reduced to the following:

$$I(U_1,U_2,\lambda_1,\lambda_2)\sum_{j=0}^{\infty}\sum_{\alpha=0}^{\infty}\frac{(U_1U_2\lambda_1\lambda_2)^j}{J(j,j+\alpha)}\left\{(\lambda_1^{\alpha} - \lambda_2^{\alpha})(U_1^{\alpha} - U_2^{\alpha})\right\} > 0 . \qquad (4.2.14)$$

The strict inequality holds since $U_1 > U_2$ and $\lambda_1 > \lambda_2$. Thus the noncentral chi-square distribution has a monotone likelihood ratio.

Since the noncentral chi-square distribution has a monotone likelihood ratio and satisfies (4.2.10) with a strict inequality and since $h(\lambda,\tau) = [L_0(\lambda,\tau) - L_1(\lambda,\tau)]$ satisfies the conditions of Lemma 1 as well as Corollaries 1 and 2 of the paper by Karlin and Rubin [1956, pp. 276,277], Equation (4.2.8) is the correct form of $\Phi_0'(U)$. Thus the form of $\underline{\Phi}(U) = (\Phi_0(U),\Phi_1(U))$ which minimizes (4.2.2) is given by:

$$\underline{\Phi}'(U) = \begin{cases} (1,0), & U < C \\ (0,1), & U \geq C \end{cases}, \tag{4.2.15}$$

where $H(C,\tau) = 0$.

The above results can now be summarized in the following theorem:

<u>Theorem 4.1</u>: Let U be a random variable coming from a noncentral chi-square distribution with S degrees of freedom and noncentrality parameter λ. Further assume that λ is a continuous random variable with cumulative distribution function $F(\lambda) = \int_0^\lambda f(t)dt$, where $f(t) > 0$ for $t \epsilon (0,\infty)$. Consider a two-decision procedure $\delta_2(U)$, whose decisions are d_0: $\lambda \epsilon \{\lambda | 0 \leq \lambda < \tau\}$ and d_1: $\lambda \epsilon \{\lambda | \tau \leq \lambda \leq \infty\}$, where τ is a positive tolerance, and whose loss function $L(\delta_2(U),\lambda)$ is defined by (4.2.1) and satisfies conditions (I) through (IV). The form of $\delta_2(U)$ which minimizes the average risk function $R(\delta_2,S,\tau,f)$, defined by (4.2.3), is given by:

$$\delta_2'(U) = \begin{cases} d_0, & U < C \\ d_1, & U \geq C \end{cases}, \tag{4.2.16}$$

where $H(C,\tau) = 0$, or by $\delta_2'(U) = d_0$ if $H(U,\tau) < 0$ for all U, or by $\delta_2'(U) = d_1$ if $H(U,\tau) \geq 0$ for all U. The equation $H(U,\tau)$ is defined by (4.2.4).

Some values of a function of C have been tabulated for several choices of S, τ, f and $L(\delta_2(U),\lambda)$. These tables are described in Section 5.2 and are displayed in the Appendix. Then the decision of whether to choose decisions d_0 or d_1 can be made by comparing the observed value of U to C.

4.3 Multiple decision procedures

In this section the development of two multiple decision procedures will be outlined. The first will be the combination of p two-decision procedures arising from p independent noncentral chi-square distributions. The second procedure will be a q-decision procedure based on one noncentral chi-square distribution.

In the first case, a method of combining decision procedures with additive losses will be used to combine p independent two-decision procedures of the type given in Section 4.2. This combining method is given by Lehmann [1957]. In this case p independent noncentral chi-square random variables $U_1,\ldots,U_p$, whose distributions have noncentrality parameters $\lambda_1,\ldots,\lambda_p$ and degrees of freedom $S_1,\ldots,S_p$, respectively, will be used. The noncentrality parameters $\lambda_1,\ldots,\lambda_p$ will be assumed to be independent random variables with prior distributions $F_i(\lambda_i) = \int_0^{\lambda_i} f_i(t)dt$, $f_i(t) > 0$ for $t \epsilon (0,\infty)$, $i = 1,\ldots,p$. The decision procedure of concern in this first case will be $\delta_p(\underline{U}_p)$, where $\underline{U}_p = (U_1,\ldots,U_p)$. The number of possible decisions is 2^p. These decisions are given by the Cartesian product of the p sets of decisions $\{d_{0,j}, d_{1,j}\}$, $j = 1,\ldots,p$, where $d_{0,j}$: $\underline{\lambda}_p \in \{\underline{\lambda}_p | 0 \leq \lambda_j < \tau_j\}$, $d_{1,j}$: $\underline{\lambda}_p$: $\{\underline{\lambda}_p | \tau_j \leq \lambda_j \leq \infty\}$, $j = 1,\ldots,p$, and $\underline{\lambda}_p = (\lambda_1,\ldots,\lambda_p)$. The tolerances $\tau_1,\ldots,\tau_p$ will be assigned by the experimenter.

Under the assumption of additive loss functions and the assumption that each of the individual loss functions for the decision procedures $\delta_{2,j}(\underline{U}_p)$ with decisions $d_{0,j}$ and $d_{1,j}$ obeys conditions (I) through (VI) given in Section 4.2, with U replaced by $\underline{U}_p$, the form of $\delta_p(\underline{U}_p)$ which will minimize the average risk is given by

$$\delta_p'(\underline{U}_p) = (\delta_{2,1}'(\underline{U}_p),\ldots,\delta_{2,p}'(\underline{U}_p)), \tag{4.3.1}$$

where $\delta_{2,j}'(\underline{U}_p) = \delta_{2,j}'(U_j)$ as defined by (4.2.16) with the proper subscript added. The next portion of this section up to the statement of Theorem 4.2 is just a verification of (4.3.1) and may be omitted without loss of generality by the reader who is not interested in the details of this demonstration.

Now to develop (4.3.1), let the loss incurred in making decision $d_{i_j,j}$ be $L_{i_j,j}(\lambda_j,\tau_j)$, $i_j = 0,1$, $j = 1,\ldots,p$. Then under the assumption of additive losses, the loss function for $\delta_p(\underline{U}_p)$ is

$$L(\delta_p(\underline{U}_p),\underline{\lambda}_p) = \sum_{j=1}^{p} \sum_{i_j=0}^{1} \left\{\Phi_{i_1,\ldots,i_p}(\underline{U}_p) \sum_{s=1}^{p} L_{j_s,s}(\lambda_s,\tau_s)\right\}$$

$$= \sum_{j=1}^{p} \sum_{i_j=0}^{1} \Phi_{i_j,j}(\underline{U}_p) L_{i_j,j}(\lambda_j,\tau_j)$$

$$= \sum_{j=1}^{p} L_j(\delta_{2,j}(\underline{U}_p),\lambda_j) \;, \tag{4.3.2}$$

where

$$\Phi_{i_j,j}(\underline{U}_p) = \sum_{\substack{s=1 \\ s\neq j}}^{p} \sum_{i_s=0}^{1} \Phi_{i_1,\ldots,i_p}(\underline{U}_p) ,$$

$i_j = 0,1$ and $\delta_{2,j}(\underline{U}_p)$ is the two-decision procedure with decisions $d_{0,j}$ and $d_{1,j}$ and loss function $L_j(\delta_{2,j}(\underline{U}_p),\lambda_j)$. In the above equation take

$$\Phi_{i_1,\ldots,i_p}(\underline{U}_p) = \begin{cases} 1, & \{\underline{U}_p \mid \delta_p(\underline{U}_p) = (d_{i_1,1},\ldots,d_{i_p,p}) \\ 0, & \text{otherwise} \end{cases} .$$

and

$$\sum_{j=1}^{p} \sum_{i_j=0}^{1} \Phi_{i_1,\ldots,i_p}(\underline{U}_p) = 1 \tag{4.3.3}$$

for all $\underline{U}_p$. Thus (4.3.3) gives that

$$\Phi_{0,j}(\underline{U}_p) + \Phi_{1,j}(\underline{U}_p) = 1$$

for all $\underline{U}_p$, $j = 1,\ldots,p$.

Then the average risk function associated with $\delta_p(\underline{U}_p)$ is given by

$$R(\delta_p,\underline{S}_p,\underline{\tau}_p,\underline{f}_p) = \int_0^\infty \cdots \int_0^\infty L(\delta_p(\underline{U}_p),\underline{\lambda}_p) \prod_{j=1}^{p} g(U_j|\lambda_j,S_j)f_j(\lambda_j) \prod_{j=1}^{p} dU_j \prod_{j=1}^{p} d\lambda_j$$

$$= \sum_{j=1}^{p} \int_0^\infty \cdots \int_0^\infty L_j(\delta_{2,j}(\underline{U}_p),\lambda_j) \prod_{j=1}^{p} g(U_j|\lambda_j,S_j)f_j(\lambda_j) \prod_{j=1}^{p} dU_j \prod_{j=1}^{p} d\lambda_j$$

$$= \sum_{j=1}^{p} R_j(\delta_{2,j},\underline{S}_p,\tau_j,\underline{f}_p), \tag{4.3.4}$$

where $\underline{S}_p$, $\underline{\tau}_p$ and $\underline{f}_p$ are defined as are $\underline{U}_p$ and $\underline{\lambda}_p$. Now assuming $L_j(\delta_{2,j}(\underline{U}_p),\lambda_j)$ obeys conditions (I) through (VI) of Section 4.2 for $\underline{U}_p$ instead of $U, j = 1,\ldots,p$, then the form of $\delta_p(\underline{U}_p)$ which minimizes

$R(\delta_p, \underline{S}_p \tau_p, \underline{f}_p)$ can be determined by use of the paper by Lehmann [1957] on multiple decision problems to be

$$\delta_p'(\underline{U}_p) = (\delta_{2,1}'(\underline{U}_p), \ldots, \delta_{2,p}'(\underline{U}_p)), \tag{4.3.5}$$

where $\delta_{2,j}'(\underline{U}_p)$ minimizes $R_j(\delta_{2,j}, \underline{S}_p, \tau_j, \underline{f}_p)$, $j = 1, \ldots, p$. It is easily seen by a method similar to the one used in Section 4.2 that

$$\delta_{2,j}'(\underline{U}_p) = \delta_{2,j}'(U_j) = \begin{cases} d_{0,j}, & \Phi_{0,j}'(\underline{U}_p) = \Phi_{0,j}'(U_j) = 1 \\ d_{1,j}, & \Phi_{1,j}'(\underline{U}_p) = \Phi_{1,j}'(U_j) = 1, \end{cases} \tag{4.3.6}$$

where $\Phi_{0,j}'(U_j) + \Phi_{1,j}'(U_j) = 1$ for all U_j and

$$\Phi_{0,j}'(U_j) = \begin{cases} 1, & U_j < C_j \\ 0, & U_j \geq C_j, \end{cases} \tag{4.3.7}$$

and where C_j is defined as C is in Section 4.2.

Thus the form of $\delta_p'(\underline{U}_p)$ is given by

$$\delta_p'(\underline{U}_p) = \left\{(d_{i_1,1}, \ldots, d_{i_p,p}), \text{ for } \prod_{j=1}^{p} \Phi_{i_j,j}'(U_j) = 1\right\}. \tag{4.3.8}$$

This result can now be summarized in the following theorem:

Theorem 4.2: Let $U_1, \ldots, U_p$ be p independent noncentral chi-square random variables, whose distributions have noncentrality parameters $\lambda_1, \ldots, \lambda_p$ and degrees of freedom $S_1, \ldots, S_p$, respectively. Also assume that $\lambda_1, \ldots, \lambda_p$ are independent random variables with respective prior distributions $F_j(\lambda_j) = \int_0^{\lambda_j} f_j(t)dt$, $j = 1, \ldots, p$, such that $f_j(t) > 0$ for $t \in (0, \infty)$, $j = 1, \ldots, p$. Let the decision procedure $\delta_p(\underline{U}_p)$ have the 2^p possible decisions $(d_{i_1,1}, \ldots, d_{i_p,p})$, $i_j = 0, 1$, $j=1, \ldots, p$, where $d_{0,j}$: $\underline{\lambda}_p \epsilon \{\underline{\lambda}_p | 0 \leq \lambda_j < \tau_j\}$ and $d_{1,j}$: $\underline{\lambda}_p \epsilon \{\underline{\lambda}_p | \tau_j \leq \lambda_j \leq \infty\}$. Then the form of $\delta_p(\underline{U}_p)$ which minimizes the average risk $R(\delta_p, \underline{S}_p, \underline{\tau}_p, \underline{f}_p)$ associated with it is given by:

$$\delta_p'(\underline{U}_p) = \left\{(d_{i_1,1}, \ldots, d_{i_p,p}),\ \text{when } \prod_{j=1}^{p} \Phi'_{i_j,j}(\underline{U}_p P) = 1\right\},$$

where $\Phi'_{i_j,j}(\underline{U}_p)$, $j = 1, \ldots, p$, is defined by (4.3.6) and (4.3.7).

Now consider the second case, where a q-decision procedure is to be considered on one noncentral chi-square distribution. This procedure will decide in which of several mutually exclusive intervals the magnitude of the noncentrality parameter lies. The decision procedure will be based on critical values like those used in Theorem 4.1.

To consider the form of the decision procedure let U be the noncentral chi-square random variable of concern, whose distribution has noncentrality parameter λ and degrees of freedom S. Assume also that λ is a continuous random variable with prior distribution function $F(\lambda) = \int_0^{\lambda} f(t)dt$, where $f(t) > 0$ for $t \epsilon (0,\infty)$.

Let the q-decision prodecure of concern be $\delta_q^*(U)$, where it has the following possible decisions: d_j^*: $\lambda \epsilon \{\lambda | \tau_j^* \leq \lambda < \tau_{j+1}^*\}$, $j = 0, \ldots, q - 1$, where $\tau_0^* = 0$, $\tau_q^* = +\infty$ and $\tau_j^* < \tau_{j+1}^*$, $j = 0, \ldots, q - 1$. Let the loss incurred in making decision d_j^* be

$$L_j(\lambda, \underline{\tau}_{q+1}^*) = \sum_{i=1}^{j} L_1(\lambda, \tau_i^*) + \sum_{i=j+1}^{q-1} L_0(\lambda, \tau_i^*),$$

$j = 0, \ldots, q-1$, where $\underline{\tau}_{q+1}^* = (\tau_0^*, \ldots, \tau_q^*)$ and $\sum_{i=1}^{0} L_1(\lambda, \tau_i^*) = \sum_{i=q}^{q-1} L_0(\lambda, \tau_i^*) = 0$.

Then the loss function associated with $\delta_q^*(U)$ will be

$$L(\delta_q^*(U),\lambda) = \sum_{j=0}^{q-1} \Phi_j(U) L_j(\lambda,\underline{\tau}_{q+1}^*),$$

where

$$\Phi_j(U) = \begin{cases} 1, & \{U \mid \delta_q^*(U) = d_j^*\} \\ 0, & \text{otherwise} \end{cases}, \tag{4.3.9}$$

$j = 0,\ldots, q-1$, such that $\sum_{j=0}^{q-1} \Phi_j(U) = 1$ for all $U \epsilon (0,\infty)$. So the average risk function associated with $\delta_q^*(U)$ is given by

$$R(\delta_q^*,S,\underline{\tau}_{q+1}^*,f) = \int_0^\infty \int_0^\infty L(\delta_q^*(U),\lambda) g(U|\lambda,S) f(\lambda) dU d\lambda \tag{4.3.10}$$

$$= \int_0^\infty \int_0^\infty \sum_{j=1}^{q-1} \Big[\big(\sum_{i=0}^{j-1} \Phi_i(U)\big) L_0(\lambda_j,\tau_j^*) + \big(\sum_{i=j}^{q-1} \Phi_i(U)\big) L_1(\lambda,\tau_j^*) \Big] g(U|\lambda,S) f(\lambda) dU d\lambda .$$

Now assume that

$$\Phi_{0,j}^*(U) = \sum_{i=0}^{j-1} \Phi_i(U),$$

$$\Phi_{1,j}^*(U) = \sum_{i=j}^{q-1} \Phi_i(U), \tag{4.3.11}$$

and $L_0(\lambda,\tau_j^*)$ and $L_1(\lambda,\tau_j^*)$ obey conditions (I) through (VI) of Section 4.2, $j = 1,\ldots,q$. These conditions allow the interchange of the order of integration in (4.3.10) and force $L_0(\lambda,\tau_j^*)$ and $L_1(\lambda,\tau_j^*)$, $j=1,\ldots,q-1$, to obey the assumptions of Theorem 3 of Karlin and Rubin [1956, p. 281]. Thus, since the noncentral chi-square distribution has a monotone likelihood ratio and $F(\lambda)$ is a continuous prior distribution, all the assumptions of Theorem 3 of Karlin and Rubin are satisfied. This theorem then gives the form of δ_q^* (U) which minimizes (4.3.10) to be

$$\delta_q^{*'}(U) = \{d_j^*, \text{ for } C_j \leq U < C_{j+1}\}, \tag{4.3.12}$$

where $C_0 = 0$, $C_q = +\infty$,

$$\int_0^\infty \left[L_0(\lambda,\tau_j^*) - L_1(\lambda,\tau_j^*)\right] g(C_j|\lambda,S) f(\lambda) d\lambda = 0,$$

and $C_j \leq C_{j+1}$, $j = 0,\ldots,q-1$. For $j = 1,\ldots,q-1$, these C_j values are the same C values which were obtained in Section 4.2, where τ is replaced by τ_j^*.

This result can now be summarized in the following theorem:

<u>Theorem 4.3</u>: Let U be a noncentral chi-square random variable whose distribution has noncentrality parameter λ and degrees of freedom S. Assume further that λ is a continuous random variable with prior distribution $F(\lambda) = \int_0^\lambda f(t)dt$, such that $f(t) > 0$ for $t \epsilon (0,\infty)$. Let the decision procedure $\delta_q^*(U)$ have the q decisions d_j^*: $\lambda \epsilon \{\lambda | \tau_j^* \leq \lambda < \tau_{j+1}^*\}$, where $\tau_0^* = 0$, $\tau_q^* = +\infty$ and $\tau_j^* < \tau_{j+1}^*$, $j = 0,\ldots,q-1$. Also let the loss function associated with $\delta_q^*(U)$ be given by

$$L(\delta_q^*(U),\lambda) = \sum_{j=0}^{q-1} \Phi_j(U)\{\sum_{i=1}^{j} L_1(\lambda,\tau_i^*) + \sum_{i=j+1}^{q-1} L_0(\lambda,\tau_i^*)\},$$

where $\Phi_j(U)$ is defined by (4.3.9) and where $\sum_{i=0}^{j-1} \Phi_i(U)$, $\sum_{i=j}^{q-1} \Phi_i(U)$, $L_0(\lambda,\tau_j^*)$ and $L_1(\lambda,\tau_j^*)$ obey conditions (I) through (VI) of Section 4.2, $j = 1,\ldots, q-1$. Then the form of $\delta_q^*(U)$ which minimizes the average risk $R(\delta_q^*, S, \tau_{q+1}^*, f)$ is given by $\delta_q^{*'}(U)$, where $\delta_q^{*'}(U)$ is defined by (4.3.12).

The critical C values for $\delta_q^{*'}(U)$ are obtained from the same tables as those used for $\delta_2'(U)$ and $\delta_p'(\underline{U}_p)$.

4.4 A procedure for comparing the magnitude of noncentrality parameters

In Sections 4.2 and 4.3, the decision procedures developed were concerned with the location of the noncentrality parameters in the parameter spaces, but none of these methods directly compared the magnitudes of the noncentrality parameters. So in this section, a method of comparing the magnitudes of noncentrality parameters will be presented. This method will give a decision procedure which will make decisions as to whether two noncentrality parameters are within a certain tolerance of each other and if not, in what direction the difference lies. To obtain this three-decision procedure, two two-decision procedures will be developed and then they will be combined through the use of the paper by Lehmann [1957] on multiple decision procedures. Under the assumption of linear additive loss functions, the resulting minimum risk decision procedure will be based on the comparison of the posterior expectations of the noncentrality parameters. That is, if the posterior expected values of the two noncentrality parameters are within a certain tolerance of each other, it will be decided that the two noncentrality parameters are within the same tolerance, but if the two posterior expected values are not within the tolerance, the two noncentrality parameters will be said to differ in the same direction as the posterior expected values differ.

Now to develop the decision procedure, let U_1 and U_2 be the two independent noncentral chi-square random variables of interest, whose distributions have noncentrality parameters λ_1 and λ_2 and degrees of freedom S_1 and S_2, respectively, where λ_1 and λ_2 are continuous independent random variables with prior distributions $F_i(\lambda_i) = \int_0^{\lambda_i} f_i(t)dt$, $f_i(t) > 0$ for $t\epsilon(0,\infty)$, $i = 1,2$. Let the two two-decision procedures of concern be $\overline{\delta}_i(\underline{U}_2)$, $i = 1,2$, where $\overline{\delta}_i(\underline{U}_2)$ has decisions

$$\overline{d}_{0,i}: \quad \underline{\lambda}_2 \epsilon \{\underline{\lambda}_2 | \lambda_i < \lambda_j + \tau\}$$

and (4.4.1)

$$\overline{d}_{1,i}: \quad \underline{\lambda}_2 \epsilon \{\underline{\lambda}_2 | \lambda_i \geq \lambda_j + \tau\}$$

such that $i = 1$, $j = 2$ or $i = 2$, $j = 1$, $i = 1,2$. The tolerance τ will be assigned by the experimenter. Let the loss function associated with $\overline{\delta}_i(\underline{U}_2)$ be given by

$$\overline{L}_i(\overline{\delta}_i(\underline{U}_2),\underline{\lambda}_2) = \overline{\Phi}_{0,i}(\underline{U}_2)\overline{L}_{0,i}(\underline{\lambda}_2,\tau) + \overline{\Phi}_{1,i}(\underline{U}_2)\overline{L}_{1,i}(\underline{\lambda}_2,\tau), \quad (4.4.2)$$

where

$$\overline{\Phi}_{0,i}(\underline{U}_2) = \begin{cases} 1, & \{\underline{U}_2 | \overline{\delta}_i(\underline{U}_2) = \overline{d}_{0,i}\} \\ 0, & \text{otherwise} \end{cases}, \quad (4.4.3)$$

such that $\overline{\Phi}_{0,i}(\underline{U}_2) + \overline{\Phi}_{1,i}(\underline{U}_2) = 1$ for all $\underline{U}_2$, and where $\overline{\Phi}_{0,i}(\underline{U}_2) = 1$ only on open sets. Also let

$$\overline{L}_{0,i}(\underline{\lambda}_2,\tau) = \begin{cases} 0 & , \ \lambda_i < \lambda_j + \tau \\ (\lambda_i - \lambda_j - \tau), & \lambda_i \geq \lambda_j + \tau \end{cases}$$

and (4.4.4)

$$\overline{L}_{1,i}(\underline{\lambda}_2,\tau) = \begin{cases} (\lambda_j - \lambda_i + \tau), & \lambda_i < \lambda_j + \tau \\ 0 & , \ \lambda_i \geq \lambda_j + \tau \ , \end{cases}$$

such that $i = 1$, $j = 2$ or $i = 2$, $j = 1$, $i = 1,2$, so that $\overline{L}_{0,i}(\underline{\lambda}_2,\tau)$ is the loss incurred when decision $\overline{d}_{0,i}$ is made and $\overline{L}_{1,i}(\underline{\lambda}_2,\tau)$ is the loss incurred when decision $\overline{d}_{1,i}$ is made. Note that $\overline{L}_i(\overline{\delta}_i(\underline{U}_2),\underline{\lambda}_2)$ satisfies conditions (I) through (VI) of Section 4.2 when $\overline{\Phi}_{o,i}(\underline{U}_2) = 1$ only on open sets, $i = 1,2$, for $\underline{U}_2$ and $\underline{\lambda}_2$ replacing U and λ, respectively.

The average risk function associated with $\overline{\delta}_i(\underline{U}_2)$ is given by:

$$\overline{R}_i(\overline{\delta}_i,\underline{S}_2,\tau,\underline{f}_2) = \int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty \overline{L}_i(\overline{\delta}_i(\underline{U}_2),\underline{\lambda}_2) \prod_{i=1}^{2} g(U_i|\lambda_i,S_i)f_i(\lambda_i) \prod_{i=1}^{2} dU_i \prod_{i=1}^{2} d\lambda_i, \tag{4.4.5}$$

$i = 1,2$. Since $\overline{L}_i(\overline{\delta}_i(\underline{U}_2),\underline{\lambda}_2)$ satisfies conditions (I) through (VI) of Section 4.2, the order of integration in (4.4.5) can be interchanged to give

$$\overline{R}_i(\overline{\delta}_i,\underline{S}_2,\tau,\underline{f}_2) = \int_0^\infty\int_0^\infty \overline{\Phi}_{0,i}(\underline{U}_2)$$

$$\left[\int_0^\infty\int_0^\infty \left\{\overline{L}_{0,i}(\underline{\lambda}_2,\tau) - \overline{L}_{1,i}(\underline{\lambda}_2,\tau)\right\} \prod_{i=1}^{2} \pi_i(\lambda_i|U_i) \prod_{i=1}^{2} d\lambda_i\right] \cdot \prod_{i=1}^{2} P_i(U_i) \prod_{i=1}^{2} dU_i$$

$$+ \text{ constant}, \tag{4.4.6}$$

$i = 1,2$, where $\pi_i(\lambda_i|U_i)$ and $P_i(U_i)$, $i = 1,2$, are defined as in Section 4.2. Then in a manner similar to the method used in Section 4.2, the form of $\overline{\Phi}_{0,i}(\underline{U}_2)$ which minimizes (4.4.6) can be found to be

$$\overline{\Phi}'_{0,i}(\underline{U}_2) = \begin{cases} 1, & H_i(\underline{U}_2,\tau) < 0 \\ 0, & H_i(\underline{U}_2,\tau) \geq 0, \end{cases} \tag{4.4.7}$$

where

$$H_i(\underline{U}_2,\tau) = \int_0^\infty \int_0^\infty \left[\bar{L}_{0,i}(\underline{\lambda}_2,\tau) - \bar{L}_{1,i}(\underline{\lambda}_2,\tau)\right] \prod_{i=1}^{2} \pi_i(\lambda_i | U_i) \prod_{i=1}^{2} d\lambda_i$$

$$= E(\lambda_i | U_i) - E(\lambda_j | U_j) - \tau$$

by (4.4.4), where $i = 1$, $j = 2$ or $i = 2$, $j = 1$, $i = 1,2$.

Now consider the decision procedure $\bar{\delta}(\underline{U}_2)$, where the set of possible decisions for $\bar{\delta}(\underline{U}_2)$ is the Cartesian product of the sets of possible decisions of $\bar{\delta}_i(\underline{U}_2)$, $i = 1,2$. That is, $\bar{\delta}(\underline{U}_2)$ will have the decisions $(\bar{d}_{0,1},\bar{d}_{0,2})$: $\underline{\lambda}_2 \epsilon \{\underline{\lambda}_2 | |\lambda_1 - \lambda_2| < \tau\}$, $(\bar{d}_{0,1},\bar{d}_{1,2})$: $\underline{\lambda}_2 \epsilon \{\underline{\lambda}_2 | \lambda_1 + \tau < \lambda_2\}$, $(\bar{d}_{1,1},\bar{d}_{0,2})$: $\underline{\lambda}_2 \epsilon \{\underline{\lambda}_2 | \lambda_2 + \tau < \lambda_1\}$ and $(\bar{d}_{1,1},\bar{d}_{1,2})$: $\underline{\lambda}_2$ belongs to the null set, where the loss for making the decision $(\bar{d}_{s,1},\bar{d}_{r,2})$ is the sum of the respective losses for $\bar{d}_{s,1}$ and $\bar{d}_{r,2}$, $r,s = 0,1$. The the unique form of $\bar{\delta}(\underline{U}_2)$ which minimizes its corresponding average risk function, under the assumption of additive losses, is given by

$$\bar{\delta}'(\underline{U}_2) = (\bar{\delta}_1'(\underline{U}_2),\ \bar{\delta}_2'(\underline{U}_2)) \tag{4.4.8}$$

by the use of the paper by Lehmann [1957] on multiple decision procedures, if the inconsistent decision $\bar{\delta}'(\underline{U}_2) = (\bar{d}_{1,1},\bar{d}_{1,2})$ can not be made. But $\bar{\delta}'(\underline{U}_2) = (\bar{d}_{1,1},\bar{d}_{1,2})$ on the set of $\underline{U}_2$ values given by $\{\underline{U}_2 | E(\lambda_1 | U_1) \geq E(\lambda_2 | U_2) + \tau\} \cap \{\underline{U}_2 | E(\lambda_2 | U_2) \geq E(\lambda_1 | U_1) + \tau\}$ which is a null set, so the decision $\bar{\delta}'(\underline{U}_2) = (\bar{d}_{1,1},\bar{d}_{1,2})$ can not be made. Thus $\bar{\delta}'(\underline{U}_2)$ is the minimum risk decision procedure. This result can now be summarized in the following theorem:

<u>Theorem 4.4</u>: Let U_1 and U_2 be independent noncentral chi-square random variables, whose distributions have noncentrality

parameters λ_1 and λ_2 and degrees of freedom S_1 and S_2, respectively. Also let λ_1 and λ_2 be independent random variables with prior distribution functions $F_i(\lambda_i) = \int_0^{\lambda_i} f_i(t)dt$, $f_i(t) > 0$ for $t \in (0,\infty)$, $i = 1,2$, respectively. Let the two two-decision procedures $\bar{\delta}_i(\underline{U}_2)$, $i = 1,2$, have decisions $(\bar{d}_{0,i}, \bar{d}_{1,i})$, $i = 1,2$, as defined by (4.4.1) and let the loss function associated with $\bar{\delta}_i(\underline{U}_2)$ be given by $\bar{L}_i(\bar{\delta}_i(\underline{U}_2),\underline{\lambda}_2)$, as defined by (4.4.2), where $\bar{L}_i(\bar{\delta}_i(\underline{U}_2),\underline{\lambda}_2)$ satisfies conditions (I) through (VI) of Section 4.2, $i = 1,2$. The form of $\bar{\delta}_i(\underline{U}_2)$, $i = 1,2$, which minimizes its respective average risk function is given by $\bar{\delta}_i'(\underline{U}_2)$, $i = 1,2$, respectively, where $\bar{\delta}_i'(\underline{U}_2)$ is defined by (4.4.7). Let $\bar{\delta}(\underline{U}_2)$ be the decision procedure whose set of possible decisions is given by the Cartesian product of the sets of possible decisions of $\bar{\delta}_i(\underline{U}_2)$, $i = 1,2$. Then under the assumption of additive losses, the form of $\bar{\delta}(\underline{U}_2)$ which minimizes its respective average risk function is given by $\bar{\delta}'(\underline{U}_2) = (\bar{\delta}_1'(\underline{U}_2), \bar{\delta}_2'(\underline{U}_2))$.

Since $\bar{\delta}'(\underline{U}_2)$ involves comparing the posterior expected values of noncentrality parameters, tables are given in the Appendix from which these expected values can be obtained for several choices of S and f, where $F(\lambda)$ is taken to be the gamma distribution. These tables are discussed in Section 5.2 of Chapter V.

The loss functions used in this section are not very general due to the great difficulty in computing which would arise if a more general loss function were used.

V. METHODS TO AID IN THE USE OF THE DECISION PROCEDURES

5.1 Introduction

This chapter will present ways to aid the experimenter in the use of the decision procedures developed in Chapter IV. In all the aids to be presented, the prior distribution is assumed to be the gamma distribution

$$F(\lambda|p,\theta) = \int_0^\lambda \left[\frac{t^{p-1}e^{-t/2\theta}}{\Gamma(p)\ (2\theta)^p}\right] dt, \qquad (5.1.1)$$

p, $\theta > 0$. The points that will be covered are: (1) the choice of the parameters p and θ of the prior distribution; (2) the use of the various tables given in the Appendix; and (3) the choice of the sample size.

5.2 The prior and approximate posterior distribution

As has been indicated in Section 4.1, any noncentrality parameter on which a decision is to be made is assumed to be a continuous random variable varying between zero and plus infinity. The random variable will be considered to have a gamma distribution of the form given in (5.1.1) initially. This form has been chosen for the prior distribution because it allows the noncentrality parameter to range from zero to plus infinity, it is a flexible unimodal distribution and it combines well with the noncentral chi-square distribution.

Once the gamma distribution has been decided on for the general form of the prior distribution of the noncentrality parameter, a procedure must be considered for selecting the values of the parameters of the distribution, that is, the values of p and θ, for the first experiment. If there is sufficient information available on the behavior of the noncentrality parameter, the mean and variance of this information can be estimated. Then the estimated mean and variance can be equated to the mean $p(2\theta)$ and the variance $p(2\theta)^2$, respectively, of the gamma distribution. From these equations estimates of p and θ can be computed.

Three other ways are proposed here for the selection of the initial prior distribution if sufficient information is not available to specify the mean and variance of the initial prior distribution. (1) The first method is to base the prior distribution on the sort of information which would be likely to be available if an experiment had been performed, namely, an observed value of the chi-square random variable that is associated with the noncentrality parameter. Assume this chi-square random variable has S degrees of freedom. This hypothetical observed value can be thought of as an estimate of N times the noncentrality parameter, where N is the size of the sample from which the hypothetical observed value would have been computed. Also this hypothetical observed value would have come from a distribution which has a mean of at least S and a variance of at least 2S with equality holding in both instances when the noncentrality parameter is zero. So to approximate the parameters p and θ of the initial gamma prior distribution of the noncentrality parameter, the experimenter can set the mean and variance of N times the noncentrality parameter equal to S and 2S, respectively, that is, $Np(2\theta) = S$ and $N^2p(2\theta)^2 = 2S$, and solve for p and θ. This method gives $p = 1/2S$ and $\theta = 1/N$. The experimenter could take N small, say 1, if he is fairly certain that $\lambda > 0$ and could take N larger if he thought it likely that $\lambda \doteq 0$. If the experimenter feels that he cannot make the assumption that the approximate prior distribution is of the form given above, that is, that λ is likely to be small, then two other possible courses of action are as follows: (2) The experimenter can assign a uniform weight function for the first experiment, which will give equal weight to all possible values of the noncentrality parameter λ; (3) In the case when the experimenter is choosing between decisions of the form $0 \leq \lambda < \tau$ and $\tau \leq \lambda \leq \infty$, he can assign a prior distribution which will have its median at τ and thus assign equal weight to both decisions. In (2), a uniform weight function is the same as a gamma weight function $f^*(\lambda) = \lambda^{p-1}e^{-\lambda/2\theta}$, with $p = 1$ and $\theta = +\infty$. When the decisions are of the form $0 \leq \lambda < \tau$ and $\tau \leq \lambda \leq \infty$, this weight function will assign unequal weights to these two decisions, even though it seems to be impartial. So it may be more appropriate in this case to follow (3) and use a gamma prior distribution which has its median at τ, for example, a gamma prior distribution defined by $p = 1$ and $\theta = \tau/(2\ln 2)$.

Once the prior distribution of the noncentrality parameter has been selected and the actual chi-square value has been observed, the problem of computing or approximating the posterior distribution must be faced. The posterior distribution would hopefully always have a simple form and be easily computed, but in this case that is not true. The actual form of the pdf for the posterior distribution is very cumbersome even after just one experiment has been performed and becomes even more so after each new experiment is performed. For example, if the random variable U under consideration comes from a noncentral chi-square distribution with S degrees of freedom and noncentrality parameter λ, then the posterior pdf for λ after a value of U has been observed, when the prior distribution has pdf

$$f(\lambda) = \lambda^{p-1} e^{-\lambda/2\theta} / (\Gamma(p) (2\theta)^p) \tag{5.2.1}$$

is as follows:

$$\pi(\lambda|U) = \frac{e^{-\lambda/2} \sum_{j=0}^{\infty} \frac{(U\lambda)^j}{\Gamma(1/2S+j)j!2^{2j}} \cdot \frac{\lambda^{p-1} e^{-\lambda/2\theta}}{\Gamma(p) (2\theta)^p}}{\int_0^{\infty} \left[e^{-\lambda/2} \sum_{j=0}^{\infty} \frac{(U\lambda)^j}{\Gamma(1/2S+j)j!2^{2j}} \cdot \frac{\lambda^{p-1} e^{-\lambda/2\theta}}{\Gamma(p) (2\theta)^p} \right] d\lambda} \tag{5.2.2}$$

So far this function has defied significant simplication. Thus if these decision procedures are to be used in successive experiments, some approximation must be applied to $\pi(\lambda|U)$ before it can be used as the prior distribution for the next experiment. For convenience this approximation should be of the same form as the previous prior distribution, so that only one set of critical values need be computed. It must be simple enough so that the experimenter can determine the parameter of his new approximate prior just by doing a few calculations and/or consulting a table. This can be done when the previous prior was a gamma by approximating $\pi(\lambda|U)$ by the gamma distribution of the form

$$F(\lambda|q,\varphi) = \int_0^{\lambda} \left[t^{q-1} e^{-\lambda/2\varphi} / ((2\varphi)^q \Gamma(q)) \right] dt,$$

where the parameters q and φ can be found by setting the mean $q(2\varphi)$ and the variance $q(2\varphi)^2$ of the gamma equal to the mean and variance, respectively, of the posterior distribution and then solving for q and φ. This will be an approximation the experimenter can handle, since the mean and the variance of the posterior can be determined from the tables given in the Appendix by a few simple operations. Those operations will be given in Section 5.3. While this method may not be the best possible approximation, it is easy to apply and any other approximation would be difficult for the experimenter to use.

To give an indication of the appropriateness of this approximation, the actual posterior density $\pi(\lambda|U)$ and the approximate posterior density $f(\lambda)$ were computed for values of λ from .2 in steps of .2 until $\pi(\lambda|U)$ and $f(\lambda)$ were less than 0.0000000001. Then $X = \sum_{\lambda} \frac{\{f(\lambda)-\pi(\lambda|U)\}^2}{\pi(\lambda|U)}$ was computed in each case. The value of X was computed for the following combinations of parameter values: (1) All combinations of the following parameters; one through nine in steps of two for S, one through five for p, five through twenty-five for U, and one-half through two in steps of one-half for θ and (2) For S equal to thirty-five, all combinations of five through nine for p, ten through forty in steps of ten for U and one-half through two in steps of one-half for θ. For each set of parameter values considered the value of X was found to be less than 0.125.

5.3 The use of the tables given in the Appendix

The tables which are given in the Appendix were calculated on the IBM 709 and the CDC 6400 computers at the Florida State University. These tables give values which are necessary to conduct the decision procedures given in Chapter IV under the assumption that the noncentrality parameter has a gamma prior distribution. Tables are also given to determine the parameters of the gamma distribution used to approximate the posterior distribution of the noncentrality parameter.

The tables are divided into two sets: (1) Tables 1.1.1 through 1.5.5 give a function of the critical value of the noncentral chi-square random variable which is needed to conduct the decision procedures given in Sections 4.2 and 4.3; (2) Tables 2.1 through 2.5 give values from which the critical expected values needed for the decision procedures given in Section 4.4 can be obtained and they also give values from which the parameters of the approximate gamma posterior distribution can be calculated.

For the decision procedures given in Sections 4.2 and 4.3, tables of C values are needed, where C is that value of U which satisfies (4.2.4), that is, for which

$$H(C,\tau) = \int_0^\infty [L_0(\lambda,\tau) - L_1(\lambda,\tau)]\ \pi(\lambda|C)d\lambda = 0. \tag{5.3.1}$$

Tables 1.1.1 through 1.5.5 give a function of C values to two decimal places for the following type of loss function:

$$L_0(\lambda,\tau) = \begin{cases} 0\ , & \lambda \le \tau \\ k_0(\lambda-\tau), & \lambda > \tau \end{cases}$$

$$L_1(\lambda,\tau) = \begin{cases} k_1(\tau-\lambda), & \lambda \le \tau \\ 0\ , & \lambda > \tau\ , \end{cases} \tag{5.3.2}$$

k_0, $k_1 > 0$, which is a special case of the loss function defined in Conditions (I), (II) and (III) of Section 4.2. The tables are given for only one type of loss function due to the rather long time necessary to determine each C. As was indicated in Section 5.2 only a gamma distribution or a gamma weight function with parameters p and θ are used, so that the probability density $\pi(\lambda|U)$ defined by (4.2.5) is given by

$$\pi(\lambda|U) = e^{-\lambda/[2\theta/(\theta+1)]} \sum_{j=0}^{\infty} \left[U^j \lambda^{p+j-1}/(\Gamma(1/2S+j)j!2^{2j})\right] \cdot \left[\frac{U^{1/2S} e^{-1/2U}}{2^{1/2S} p(U)}\right], \tag{5.3.3}$$

where S is the degrees of freedom of U and p(U) is such that

$$\int_0^\infty \pi(\lambda|U)d\lambda = 1.$$

Then putting (5.3.2) and (5.3.3) into (5.3.1), making the transformation $y = \lambda/[\frac{2\theta}{\theta+1}]$, performing the integration and solving for the ratio

$$R = (k_1-k_0)/k_0,$$

(5.3.1) is seen to be equivalent to

$$R = H(C^*,\tau^*,p,S) = \frac{\sum_{j=0}^{\infty} [C^{*j}(\Gamma(p+j+1)-\tau^*\Gamma(p+j))/(\Gamma(1/2S+j)j!)]}{\tau^*(p+1)\sum_{j=0}^{\infty} \frac{(C^*\tau^*)^j}{\Gamma(1/2S+j)j!} \sum_{i=0}^{\infty} \frac{(-\tau^*)^i}{i!(p+i+j)(p+i+j+1)}}, \tag{5.3.4}$$

where

$$C^{*} = \theta C/(2(\theta+1))$$

and (5.3.5)

$$\tau^{*} = (\theta+1)\tau/(2\theta) \ .$$

So the tables 1.1.1 through 1.5.5 are given in terms of C^{*} values, for selected values of S, p, τ^{*} and R. The degrees of freedom S takes on the values one through five while p takes on values which depend on 1/2S, decreasing from 1/2S in steps of 0.2 and p also takes on five values greater than 1/2S, increasing from 1/2S in steps of 0.2. The tolerance τ^{*} takes on the five integer values one through five. The ratio R_i indicates that the experimenter feels that a wrong decision in the region $\tau \leq \lambda \leq \infty$ is 10+10i times as bad as a wrong decision in the region $0 \leq \lambda < \tau$. These choices of the R's can be compared with the ratio of the Type II error (α) to the Type I error (β) minus one in the conventional hypothesis testing situation. That is, if the experimenter wishes to have α one-tenth of β or α five-hundredths of β, he would choose R to be 9 or 19, respectively.

The appropriate C^{*} values were obtained by using the Newton-Raphson interative approach on (5.3.4) for the given values of S, p and R. If a zero entry appears in Tables 1.1.1 through 1.5.5, this means that (5.3.1) does not have a zero for $C^{*} > 0$ and that the decision $\lambda > \tau$ is always taken for the particular choice of S, p and R which correspond to the zero entry.

To find an accurate approximation to the infinite sums involved, two well known lemmas on infinite series were used from Olmsted [1959].

Consider first the two single infinite sums given in (5.3.4). The method of approximation used in these two cases is given by the following lemma:

Lemma 5.1: (Theorem I, Section 720, Olmsted [1959, p. 235].)
If $a_j > 0$, $S_n = a_0+\ldots+a_n$, $S = \sum_{j=0}^{\infty} a_j$, $r_j = \frac{a_{j+1}}{a_j}$, $r_j\downarrow$ and $\lim_{j\to\infty} r_j = \rho<1$, then for any n for which $r_n < 1$,

$$S_n + a_{n+1}/(1-\rho) \leq S \leq S_n + a_{n+1}/(1 - r_{n+1}). \tag{5.3.6}$$

Let

$$S = \sum_{j=0}^{\infty} \left[\frac{C^{*j}\Gamma(p+1+j)}{\Gamma(1/2S+j)j!}\right], \tag{5.3.7}$$

$$a_j = \frac{C^{*j}\ \Gamma(p+1+j)}{\Gamma(1/2S+j)\ j!} \tag{5.3.8}$$

and

$$r_j = C^*(p+1+j)/((1/2S+j)(j+1)). \tag{5.3.9}$$

Now $r_j > r_{j+1}$ for all j, so that $r_j\downarrow$ and $\lim_{j\to\infty} r_j = 0$. Thus the lemma can be applied. Let

$$S_n = \sum_{j=0}^{n} a_j \tag{5.3.10}$$

where a_j is defined in (5.3.8). In the computations the number of terms n in the partial sum S_n was taken large enough so that both a_{n+1} and $a_{n+1}/(1-r_{n+1})$ were less than $\epsilon = 0.0000001$ for $r_{n+1} < 1$. Thus S_n

is within ϵ of S as given by (5.3.7). A similar procedure was used for the second single sum in (5.3.4).

Now for the double sum in (5.3.4), the Lemma 4.1, plus the following well known lemma were used:

<u>Lemma 5.2:</u> (Theorem given Section 715, Olmsted [1959, p. 225].) An alternating series $\sum_{i=0}^{\infty} (-1)^i C_i$, whose terms satisfy the two conditions

(i) $C_{n+1} < C_n$ for every n

(ii) $C_n \to 0$ as $n \to \infty$

converges. If S and S_n denote the sum and the partial sum of the first n terms, respectively, of the series, then

$$|S_n - S| < C_{n+1}.$$

So in the double sum in (5.3.4) consider first the alternating internal sums which are all of the form

$$S_j = \sum_{i=0}^{\infty} \frac{(-\tau^*)^i}{i!\,(p+i+j)(p+i+j+1)} . \tag{5.3.11}$$

Then let

$$C_{ij} = (\tau^*)^i/(i!\ (p+i+j)(p+i+j+1)) .$$

Now for $i > \tau^*$, $C_{i+1j} < C_{ij}$. Also,

$$\lim_{i \to \infty} C_{ij} = 0.$$

Thus for that portion of (5.3.11) for which $i > \tau^*$, the alternating series lemma can be applied. So if

$$S_{nj} = \sum_{i=0}^{n} (-1)^i C_{ij}, \tag{5.3.12}$$

then

$$|S_j - S_{nj}| < C_{n+1j}$$

for $n > \tau^*$. In the calculation of partial sums of the form (5.3.12) $M > \tau^*$ was taken large enough, so that

$$C_{n+1j} < (\epsilon)/(4k) \tag{5.3.13}$$

for $n > M$, where $k > \tau^{*(p+1)} \sum_{j=0}^{\infty} \frac{(C^*\tau^*)^j}{\Gamma(1/2S+j)j!}$ and $\epsilon = 0.0000001$. The reason for taking this kind of tolerance will become apparent later.

Now let

$$W_{NM} = (\tau^*)^{p+1} \sum_{j=0}^{N} \frac{(C^*\tau^*)^j}{\Gamma(1/2S+j)j!} S_{Mj},$$

(5.3.14)

$$W_N = (\tau^*)^{p+1} \sum_{j=0}^{N} \frac{(C^*\tau^*)^j}{\Gamma(1/2S+j)j!} S_j,$$

and

$$W = \tau^{*(p+1)} \sum_{j=0}^{\infty} \frac{(C^*\tau^*)^j S_j}{\Gamma(1/2S+j)j!} . \tag{5.3.15}$$

For $n > M > \tau^*$ satisfying (5.3.13),

$$W_{NM} - \epsilon/4 \le W_N \le W_{NM} + \epsilon/4. \tag{5.3.16}$$

Now consider

$$r_j = \frac{(C^*\tau^*)^{j+1}\ S_{j+1}}{\Gamma(1/2S+j+1)\ (j+1)!} \Big/ \frac{(C^*\tau^*)^j\ S_j}{\Gamma(1/2S+j)\ j!}$$

$$= \frac{(C^*\tau^*)}{(1/2S+j)\ (j+1)} \cdot \frac{S_{j+1}}{S_j}$$

which is such that

$$\lim_{j \to \infty} r_j = 0$$

since S_{j+1}/S_j approaches one. Also as j becomes large, r_j will be dominated by the term

$$(C^*\tau^*) \ / \ ((1/2S+j)(j+1)) \ .$$

Thus for large j, r_j is decreasing. So by Lemma 5.1 given above

$$W_N + \frac{\tau^{*(p+1)}(C^*\tau^*)^{N+1}}{\Gamma(1/2S+N+1)\ (N+1)!}\ S_{N+1} \le W \le W_N + \frac{\tau^{*(p+1}\ (C^*\tau^*)^{N+1}\ S_{N+1}}{\Gamma(1/2S+N+1)(N+1)!(1-r_{N+1})} \qquad (5.3.17)$$

for N large enough so that r_j is decreasing for $j \ge N$ and so that $r_{N+1} < 1$. Then also by (5.3.13), (5.3.16) and (5.3.17),

$$W_{N+1,M} - \epsilon/4 \leq W_{N+1} \leq W \leq W_N + \frac{(\tau^*)^{p+1} \; (C^*\tau^*)^{N+1} \; S_{N+1}}{\Gamma(1/2S+N+1)(N+1)!(1-r_{N+1})}$$

$$\leq W_{N+1,M} + \epsilon/4 + \frac{\tau^{*(p+1)}(C^*\tau^*)^{N+1}(S_{M,N+1}+\epsilon/4)}{\Gamma(1/2S+N+1)(N+1)!(1-\hat{r}_{N+1})} \tag{5.3.18}$$

where

$$1 > \hat{r}_{N+1} = \frac{(C^*\tau^*)}{(1/2S+N+1)(N+2)} \cdot \frac{(S_{M,N+2}+ \epsilon/4)}{(S_{M,N+1}- \epsilon/4)} \geq r_{N+1}, \tag{5.3.19}$$

for $M > \tau$ satisfying (5.3.16). In the computations, N was taken such that

$$\frac{\tau^{*(p+1)}(C^*\tau^*)^{N+1} \; (S_{M,N+1} + \epsilon/4)}{\Gamma(1/2S+N+1) \; (N+1)!(1-\hat{r}_{N+1})} < \epsilon/4,$$

so the double sum in (5.3.4) was computed within ϵ of its value.

The methods indicated above for calculating the sums which make up R, defined in (5.3.4), allow R to be computed correctly to more than four decimal places, that is, the values of C^* which are tabulated, are values which force the ratio (5.3.4) to be within four decimal places of the specified R value.

To find C^* values for p, τ^* and R values not listed in the tables, but which are within the range of the tables, a five point Lagrange interpolation formula for equal intervals, as given by Kunz [1957, p. 91], can be used on each p, τ^* and R to give two decimal place accuracy.

For the decision procedures given in Section 3.4, tables must be provided from which the function $E(\lambda|U,p,\theta,S)$ can be obtained with a minimum of effort, where

$$E(\lambda|U,p,\theta,S) = \frac{\int_0^\infty \left[\lambda e^{-\lambda/[\frac{2\theta}{\theta+1}]} \sum_{j=0}^{\infty} \frac{(U/4)^j \lambda^{p+j-1}}{\Gamma(1/2S+j)j!}\right] d\lambda}{\int_0^\infty \left[e^{-\lambda/[\frac{2\theta}{\theta+1}]} \sum_{j=0}^{\infty} \frac{(U/4)^j \lambda^{p+j-1}}{\Gamma(1/2S+j)j!}\right] d\lambda} . \quad (5.3.20)$$

The tables which contain the necessary values are Tables 2.1 through 2.5. To see what is tabled and how to use it, consider (5.3.20) after the integration has been carried out, so that

$$E(\lambda|U,p,\theta,S) = [\frac{2\theta}{\theta+1}] \frac{\sum_{j=0}^{\infty} (U^*)^j \frac{\Gamma(p+j+1)}{\Gamma(1/2S+j)j!}}{\sum_{j=0}^{\infty} (U^*)^j \frac{\Gamma(p+j)}{\Gamma(1/2S+j)j!}} \quad (5.3.21)$$

$$= [\frac{2\theta}{\theta+1}] [p + U^* W(p,U^*,S)] \quad (5.3.22)$$

where

$$W(p,U^*,S) = \frac{\sum_{j=0}^{\infty} (U^*)^j \frac{\Gamma(p+j+1)}{\Gamma(1/2S+j+1)j!}}{\sum_{j=0}^{\infty} (U^*)^j \frac{\Gamma(p+j)}{\Gamma(1/2S+j+1)j!}} , \quad (5.3.23)$$

and U^* is defined like C^* in (5.3.5). Tables 2.1 through 2.5 contain values of $W(p,U^*,S)$ for selected values of S, p and U^*. The values of

S and p which were used in the first set of tables were also used in this set, but p will also be allowed to take on five additional values. The reason for having these extra values will become apparent later. The values of U^* which are used are 0.5 through 15.0 in steps of 0.5 and 20.0 through 50.0 in steps of 5.0 along with 75.0 and 100.0.

The infinite sums in (5.3.23) were approximated in exactly the same manner as the single infinite sums were approximated in (5.3.4), with the same tolerance (0.0000001). This method will give an approximation of $W(p,U^*,S)$ which is accurate to more than two decimal places but the table is given to only two decimal places. To arrive at values of p and U^* which are not tabled, but are within the range of the table, linear interpolation can be used with two decimal place accuracy resulting. If a value of U^* is observed which is greater than 100.0, use the value of $W(p,100,S)$ for U^* near 100.0 and use $W(p,U^*,S) = 1.$ for U^* much greater than 100.0.

In summary, to use the table, first observe the values of p, θ, U and S. Then compute $[2\theta/(\theta+1)]$ and U^* and find the value of $W(p,U^*,S)$, interpolating if necessary, corresponding to p, U^* and S in the tables. Then compute

$$E(\lambda|U,p,\theta,S) = [2\theta/(\theta+1)]\ [p + U^*W(p,U^*,S)] \tag{5.3.24}$$

which gives the desired result. For a quick approximation to (5.3.24), one can compute

$$E(\lambda|U,p,\theta,S) \doteq [2\theta/(\theta+1)]\ [p + U^*]\ .$$

This approximation will be good if p is close to S/2 and/or if U^* is greater than 25.

To determine the parameters of the approximate gamma posterior distribution discussed in Section 5.2, the following two equations must be solved:

$$2q\varphi = E(\lambda | U,p,\theta,S)$$

and

$$q(2\varphi)^2 = E(\lambda^2 | U,p,\theta,S) - [E(\lambda | U,p,\theta,S)]^2 .$$

The solutions are

$$2\varphi = E(\lambda^2 | U,p,\theta,S) \;/\; E(\lambda | U,p,\theta,S) - E(\lambda | U,p,\theta,S)$$

and

$$q = E(\lambda | U,p,\theta,S)/(2\varphi).$$

Now the equation for 2φ can be written as follows:

$$2\varphi = (2\theta/(\theta+1)) \left[\frac{\sum_{j=0}^{\infty} (U^*)^j \frac{\Gamma(p+j+2)}{\Gamma(S/2+j)j!}}{\sum_{j=0}^{\infty} (U^*)^j \frac{\Gamma(p+j+1)}{\Gamma(S/2+j)j!}} - \frac{\sum_{j=0}^{\infty} (U^*)^j \frac{\Gamma(p+j+1)}{\Gamma(S/2+j)j!}}{\sum_{j=0}^{\infty} (U^*)^j \frac{\Gamma(p+j)}{\Gamma(S/2+j)j!}} \right]$$

$$= [2\theta/(\theta+1)]\; [1 + U^*\{W(p+1,U^*,S) - W(p,U^*,S)\}]$$

by (5.3.23). From (5.3.22)

$$q = [2\theta/(\theta+1)]\; [p + U^*W(p,U^*,S)] \;/\; (2\varphi), \tag{5.3.25}$$

where $W(p,U^*,S)$ is given in Table 2.1 through 2.5. Thus separate tables need not be prepared for q and φ and to find them, it is necessary only to look up two values in Tables 2.1 through 2.5, one under p and one under p+1.

In many cases the noncentrality parameter is of the form $N\lambda$ instead of just λ. If this is the case, the procedure for using the tables would be the same, but $(2\theta/(\theta+1))$ would be changed to $(2\theta/(N\theta+1))$, $(\theta U/(2\theta+1))$ would be changed to $(N\theta U/(2(N\theta+1)))$ and $((\theta+1)\tau/(2\theta))$ would be changed to $((\theta+N)\tau/(2N\theta))$. The decisions of interest in this case would remain as decisions on λ and would not change to decisions on $N\lambda$.

Also in some cases the noncentrality parameter would not vary between zero and infinity, but would be bounded above by some constant. This situation can be approximately handled by considering $\gamma = N\lambda$ instead of λ. In this case, the procedure for using the tables would be the same, except that $(2\theta/(\theta+1))$ would be changed to $(2N\theta/(\theta+N))$, $(\theta U/(2(\theta+1)))$ would be changed to $(\theta U/(2(\theta+N)))$ and $((\theta+1)\tau/(2\theta))$ would be changed to $((\theta+N)\tau/(2\theta))$. The decisions of interest in this case would be approximately equivalent to decisions on the magnitude of λ.

In all the tables given in the Appendix only two decimal places are given, because the author feels that in most applications at most two decimal places would be retained.

5.4 An approximate method for computing sample size

This section is concerned with an approximate method for determining the optimum sample size, that is, the sample size which will approximately minimize a cost function. In this discussion, the cost per observation and the initial set up cost are assumed to be standardized, so that they are in the same units as the risk function.

Let the parameters of the gamma prior distribution on the noncentrality parameter λ be p and θ and let the parameters of the approximate gamma posterior distribution which will be observed after the experiment is performed be q and φ. Also let the same size of this experiment be M.

Now in Part 1, to show that minus two times the natural logarithm of the likelihood ratio statistic converge in distribution to a noncentral chi-square random variable U, an assumption of local alternatives had to be made. This assumption causes the noncentrality parameter $M\gamma_M$ to converge to constant λ, as $M \to +\infty$.

In Chapter VI, it is shown that when the distribution of concern is only an asymptotic noncentral chi-square distribution and when the decision procedures developed in Chapter III are used, the average risk function converges to the minimum average risk function for the exact noncentral chi-square distribution. So to determine the approximate sample size in the asymptotic case, it is appropriate to use the properties of the exact case.

For an approximate method for finding the sample size in the asymptotic case, assume that $\gamma_M = \lambda$. Then the noncentrality parameter for the approximate noncentral chi-square distribution is $M\lambda$. This technique of setting $\gamma_M = \lambda$ to approximate the sample size has also been indicated by other authors, for example, Kendall and Stuart [1961, p. 436], Cochran [1954, p. 419] and Patnaik [1949, p. 216]. If the random variable U has an exact noncentral chi-square distribution, the noncentrality will either be in the form $M\gamma$ or a function of M.

In most cases when the noncentrality parameter is a function of M and is not of the form $M\gamma$, M appears as the upper limit of a summation of constants which are specified by the experiment and they do not contribute to the randomness of the noncentrality parameter. Thus if the noncentrality parameter, in this case is expressed in terms of the averages of these constants, it can be expressed as $M\gamma_M$. This type of noncentrality parameter is common in the analysis of variance theory. So for the sample size determination in this case consider $\gamma_M = \gamma$. Then the results given here can be applied to all three situations considered.

So now assume that the approximate or exact noncentral chi-square distribution in question has S degrees of freedom and non-centrality parameter $M\gamma$, so its density is

$$g(U|M\gamma,S) = \frac{e^{-(U+M\gamma)/2}}{2^{1/2S}} \sum_{j=0}^{\infty} \frac{U^{1/2S+j-1}(M\gamma)^j}{\Gamma(1/2S+j)j!2^{2j}} \tag{5.4.1}$$

and the density of the gamma prior distribution of γ is

$$f(\gamma) = (\gamma^{p-1}\, e^{-\gamma/2\theta}) \;/\; (\Gamma(p)\ (2\theta)^p), \tag{5.4.2}$$

or the prior uniform weight function on γ is

$$f(\gamma) = 1. \tag{5.4.3}$$

To find the approximate sample size M, a relation between the parameters of the approximate gamma posterior and the sample size must be determined. To do this, let U_M be the value of the approximate or exact chi-square random variable U, based on M observations, which would be observed when the experiment would be completed. Then the density of the approximate gamma posterior distribution given U_M would be:

$$\pi(\gamma|U_M) = (\gamma^{q-1} e^{-\gamma/2\varphi}) / (\Gamma(q)\ (2\varphi)^q), \tag{5.4.4}$$

such that

$$2\varphi = \theta_M\{1 + X_M[W(p+1,X_M,S) - W(p,X_M,S)]\}$$

and (5.4.5)

$$q = \theta_M\{p + X_M\ W(p,X_M,S)\}/2\varphi,$$

where

$$X_M = MU_M\theta \ /\ (2(M\theta+1))$$

and (5.4.6)

$$\theta_M = 2\theta \ /\ (M\theta+1)$$

for a gamma prior distribution for which $\theta < +\infty$, while

$$X_M = U_M/2$$

and (5.4.7)

$$\theta_M = 2/M$$

for the uniform weight function with $\theta = +\infty$ and $p = 1$.

Now $W(p,X_M,S)$ is tabled in the Appendix for various values of p, X_M, and S and it can be seen there that $W(p,X_M,S)$ approaches one more rapidly than X_M approaches infinity and it is close to one for all values of X_M. So for the approximation of q and φ for use in the sample size determination, let U_M equal its expected value $(1/2S + M\gamma)$, and let $W(p,X_M,S)$ and $W(p+1,X_M,S)$ both be one. Thus for $U_M = 1/2S + M\gamma$, (5.4.5) can be approximated to give

$$2\varphi = \theta_M$$

and (5.4.8)

$$q = p + X_M ,$$

where

$$X_M \sim M\gamma/2 ,$$

and (5.4.9)

$$\theta_M \sim 2/M$$

for all gamma type priors. So let the approximate parameters of the posterior gamma distribution to be used for the sample size determination be

$$\varphi = 1/M$$

and (5.4.10)

$$q = 2\theta\ p\ M.$$

The form of q is found by replacing γ by its prior expected value $2\theta p$. For an initial uniform prior weight function let

$$\varphi = 1/M$$

and (5.4.11)

$$q = M.$$

The choice of $q = M$ is somewhat arbitrary, since the expected value of γ does not exist in this case.

Now that the parameters of the posterior distribution have been approximated in relation to M, the cost functions which yield the approximate optimum sample sizes for each of the decision procedures given in Chapter IV can be given.

For the decision procedure $\delta_2'(U)$ given in Section 4.2, let the cost function be

$$F^*(M) = C_1M + C_2 + \int_0^\infty [L_0(\gamma,\tau) + L_1(\gamma,\tau)]\pi^*(\gamma)d\gamma, \qquad (5.4.12)$$

where C_1 is the standardized cost per observation, C_2 is the standardized set up cost, and both C_1 and C_2 are in the same units as the posterior risk. The losses $L_0(\gamma,\tau)$ and $L_1(\gamma,\tau)$ are given by (5.3.2). The same type of linear losses were used in computing the tables given in the Appendix. The density $\pi^*(\gamma)$ given in (5.4.12) is the density of the approximate gamma posterior distribution for sample size determination, whose parameters q and φ are given by (5.4.10) and (5.4.11).

Now substituting (5.3.2) in (5.4.12)

$$F^*(M) = C_1M + C_2 + k_0\int_\tau^\infty (\gamma-\tau)\pi^*(\gamma)d\gamma$$
$$+ k_1\int_0^\tau (\tau-\gamma)\pi^*(\gamma)d\gamma, \qquad (5.4.13)$$

where

$$\pi^*(\gamma) = [\gamma^{(2\theta pM)} e^{-\gamma/(2M)}] / [\Gamma(2\theta pM) (2/M)^{2\theta pM}]$$

where both θ and p are one if a uniform prior weight function was used. Then letting $k = \max(k_0, k_1)$, (5.4.13) can be bounded by

$$F^*(M) \leq C_1 M + C_2 + k \int_0^\infty |\gamma - 2\theta p| \pi^*(\gamma) d\gamma + k|2\theta p - \tau| .$$

Applying the Cauchy-Schwartz inequality to the above integral,

$$F^*(M) \leq C_1 M + C_2 + k[V(\gamma|(2\theta p)M, 1/M)]^{1/2} + k|2\theta p - \tau|$$

$$\leq C_1 M + C_2 + k[4\theta p/M]^{1/2} + k|2\theta p - \tau| = F_1(M).$$

Now considering M to be continuous, the minimum of $F_1(M)$ can be found, which gives an approximate minimum of $F^*(M)$, by differentiating $F_1(M)$ with respect to M. Now

$$\frac{\partial F_1(M)}{\partial M} = C_1 - k[\theta p]^{1/2} M^{-3/2} .$$

Then setting $\left.\frac{\partial F_1(M)}{\partial M}\right|_{M_0} = 0$ and solving for the approximate optimum sample size M_0, M_0 is found to be

$$M_0 = \{(\theta p)^{1/2} (k/C_1)\}^{2/3} . \tag{5.4.14}$$

The second derivative $\partial^2 F_1(M)/\partial M^2$ at the point M_0 is positive, so that M_0 yields a minimum as desired.

If instead of just one two-decision procedure on one non-centrality parameter, there are ℓ independent two-decision procedures

$\delta'_{2,i}(U_i)$, $i = 1,\ldots,\ell$ on $\gamma_1,\ldots,\gamma_\ell$ which are to be combined through the use of the Restricted Product method to give the decision procedure $\delta'_\ell(\underline{U}_\ell)$ developed in Section 4.3, the sample sizes are found separately for each of the decision procedures $\delta'_{2,i}(U_i)$, $i = 1,\ldots,\ell$, since the combined posterior risk is the sum of the posterior risks of the individual two-decision procedures.

In the case when the ℓ-decision procedure $\delta^{*\prime}_\ell(U)$, which was also developed in Section 4.3, is to be used, the cost function can be expressed as the sum of the posterior risks of $(\ell-1)$ two-decision procedures $\delta'_{2,i}(U)$, with decisions $d_{0,i}$ and $d_{1,i}$, $i = 1,\ldots,(\ell-1)$, where $\tau^*_1 < \ldots < \tau^*_{(\ell-1)}$, plus M times C_1, the standardized cost per observation, and C_2, the standardized setup cost. Thus a reasonable cost function is

$$F^*_\ell(M) = C_1 M + C_2 + \sum_{i=1}^{\ell-1} \int_0^\infty \{L_0(\gamma,\tau^*_i) + L_1(\gamma,\tau^*_i)\}\ \pi^*(\gamma)d\gamma, \tag{5.4.15}$$

where $L_0(\gamma,\tau^*_i)$ and $L_1(\gamma,\tau^*_i)$ are defined as in (5.3.2), $i = 1,\ldots,(\ell-1)$, and $\pi^*(\gamma)$ is defined as before in (5.4.13). Now (5.4.15) can be approximated in the same manner as (5.4.13), so that

$$F^*_\ell(M) \leq C_1 M + C_2 + \sum_{i=1}^{\ell-1} \{k[4\theta p/M]^{1/2}\} + \sum_{i=1}^{\ell-1} [k|2\theta p - \tau^*_i|] = F_\ell(M)\ .$$

The approximate optimum sample size which minimizes $F_\ell(M)$ can now be found in the same manner as (5.4.14) was determined. Thus if $M_{\ell 0}$ is the approximate optimum sample size,

$$M_{\ell 0} = \{(\theta p)^{1/2} \ ((\ell-1)k/C_1\}^{2/3} \ . \tag{5.4.16}$$

If the magnitude of the noncentrality parameters are to be compared in the manner indicated in Section 4.4, a reasonable cost function for comparing γ_1 and γ_2 is as follows:

$$F^*(M_1,M_2) = C_1M_1 + C_2M_2 + C + \int_0^\infty \int_0^{\gamma_2+\tau} k(\gamma_1-\gamma_2-\tau)\ \pi_1^*(\gamma_1)\pi_2^*(\gamma_2)d\gamma_1 d\gamma_2$$

$$+ \int_0^\infty \int_{\gamma_2+\tau}^\infty k(\gamma_2+\tau-\gamma_1)\pi_1^*(\gamma_1)\pi_2^*(\gamma_2)d\gamma_1 d\gamma_2, \tag{5.4.17}$$

where C_i is the standardized cost of taking an observation from population i, i = 1,2, C is the standardized setup cost, and

$$\pi_i^*(\gamma_i) = [\gamma_i^{(2\theta_i p_i M_i)}\ e^{-\gamma_i/(2/M_i)}] \ / \ [\Gamma(2\theta p M_i)(2/M)^{(2\theta_i p_i M_i)}],$$

i = 1,2. Now from (5.4.17), it can be seen that

$$F^*(M_1,M_2) \leq C_1M_1 + C_2M_2 + C + \sum_{i=1}^{2} \int_0^\infty \int_0^\infty k|\gamma_i-2\theta_i p_i| \ \cdot$$

$$\pi_1^*(\gamma_1)\ \pi_2^*(\gamma_2)d\gamma_1 d\gamma_2 + k|2\theta_1 p_1-2\theta_2 p_2-\ \tau\ | \ . \tag{5.4.18}$$

Then applying the Cauchy-Schwartz inequality to the above integrals, (5.4.18) becomes

$$F^{*}(M_1,M_2) \le C_1M_1 + C_2M_2 + C + \sum_{i=1}^{2} k\left[\frac{4\theta_i p_i}{M_i}\right]^{1/2}$$

$$+ k\left|2\theta_1 p_1 - 2\theta_2 p_2 - \tau\right| = F_1(M_1,M_2).$$

Now considering M_1 and M_2 to be continuous and setting

$$\left.\frac{\partial F_1(M_1,M_2)}{\partial M_i}\right|_{(M_{10},M_{20})} = 0,$$

$i = 1,2$, M_{10} and M_{20} are found to be

$$M_{i0} = [(\theta_i p_i)^{1/2} (k/C_i)]^{2/3} , \tag{5.4.19}$$

$i = 1,2$. The values of M_{10} and M_{20} minimize $F_1(M_1,M_2)$ since the matrix of second partial derivatives of $F_1(M_1,M_2)$ is positive definite at the point (M_{10},M_{20}).

If ℓ noncentrality parameters are to be compared in the manner given in Section 4.4, the approximate optimum sample sizes can be found to be

$$M_{i0} = [(\theta_i p_i)^{1/2}((\ell-1)k/C_i)]^{2/3} ,$$

$i = 1,\ldots,\ell$, in a manner similar to the way (5.4.19) was determined.

It must be remembered that in arriving at values of φ and q for purposes of sample size determination, it was assumed that the value of M selected would not be small. Therefore, the procedures would be applicable only when the cost per unit sampled (C_i) is small in comparison to k, θ_i, p_i and ℓ.

VI. ASYMPTOTIC CONSIDERATIONS

This chapter will be concerned with sequences of random variables which are only known to have asymptotic noncentral chi-square distributions. That is, it will be assumed that the statistic U of concern has a cumulative distribution function $G_n(U|\lambda) = \int_0^U dG_n(t|\lambda)$, which is measurable with respect to λ, with $\lim_{n \to \infty} G_n(U|\lambda) = \int_0^U g(t|\lambda,S)\,dt$, where $g(t|\lambda,S)$ is the pdf of the noncentral chi-square distribution with noncentrality parameter λ and S degrees of freedom. To apply the decision procedures given in Chapter IV to random variables which only have asymptotic noncentral chi-square distributions, the sequence of risk functions, based on $G_n(U|\lambda)$ and the optimal decision procedure for the exact noncentral chi-square distribution, must converge to the minimum risk function given in Chapter IV.

The decision procedures developed in Section 4.2 and 4.3 are all based on a decision function of the form

$$\Phi_0'(U) = \begin{cases} 1, & U < C \\ 0, & U \geq C \ . \end{cases} \tag{6.1}$$

So for these cases, it is sufficient to show that the risk function based on $G_n(U|\lambda)$ and $\Phi_0'(U)$, which is

$$R_n(\delta_2,S,\tau,f) = \int_0^\infty \left[\int_0^\infty \Phi_0'(U)h(\lambda,\tau)\,dG_n(U|\lambda)\right]f(\lambda)\,d\lambda + \text{Constant} \tag{6.2}$$

converges to the risk function

$$R(\delta_2,S,\tau,f) = \int_0^\infty \Big[\int_0^\infty \Phi_0'(U)h(\lambda,\tau)g(U|\lambda,S) \Big] f(\lambda)\,d\lambda$$

$$+ \text{ Constant}, \tag{6.3}$$

where

$$h(\lambda,\tau) = L_0(\lambda,\tau) - L_1(\lambda,\tau)$$

has the properties given in Conditions I, II, and III of Section 4.2. A demonstration that $R_n(\delta_2,S,\tau,f)$ approaches $R(\delta_2,S,\tau,f)$ in the limit as n becomes large is given in the proof of the following theorem:

Theorem 6.1: Let $G_n(U|\lambda) = \int_0^U dG_n(U|\lambda)$ be a distribution function, which is measurable with respect to λ and which converges to the noncentral chi-square distribution $G(U|\lambda,S) = \int_0^U g(t|\lambda,S)dt$ and let λ be a random variable with distribution function $F(\lambda) = \int_0^\lambda f(t)dt$, where $f(t)$ is a continuous function of $t\epsilon[0,\infty)$. Also let $\Phi_0'(U)$, as defined by (6.1) be the decision function which minimizes the risk function $R(\delta_2,S,\tau,f)$, as defined by (6.3). Then the approximate risk function $R_n(\delta_2,S,\tau,f)$, as defined by (6.2), approaches $R(\delta_2,S,\tau,f)$ in the limit as n becomes large.

Proof: Putting the form of $\Phi_0'(U)$ given in (6.1) into $R(\delta_2,S,\tau,f)$ and $R_n(\delta_2,S,\tau,f)$, it is found that

$$R_n(\delta_2,S,\tau,f) = \int_0^\infty h(\lambda,\tau)\Big[\int_0^C dG_n(U|\lambda) \Big] f(\lambda)\,d\lambda$$

$$+ \text{ Constant}$$

and

$$R(\delta_2,S,\tau,f) = \int_0^\infty h(\lambda,\tau)\Big[\int_0^C g(U|\lambda,S)\,dU\Big]f(\lambda)\,d\lambda$$

$$+ \text{ Constant.}$$

Now by the properties given in Conditions I, II, and III of Section 4.2 for $h(\lambda,\tau)$, $\int_0^\infty h(\lambda,\tau)f(\lambda)d\lambda$ is finite. Then since it is assumed that $\lim_{n\to\infty} G_n(U|\lambda) = G(U|\lambda,S)$ and since $0 \leq G_n(U|\lambda) \leq 1$ for all n,

$$\lim_{n\to\infty} R_n(\delta_2,S,\tau,f) = \lim_{n\to\infty} \int_0^\infty h(\lambda,\tau)\Big[\int_0^C dG_n(U|\lambda)\Big]f(\lambda)\,d\lambda$$

$$+ \text{ Constant}$$

$$= \int_0^\infty h(\lambda,\tau)\Big[\lim_{n\to\infty}\int_0^C dG_n(U|\lambda)\Big]f(\lambda)\,d\lambda$$

$$+ \text{ Constant}$$

$$= \int_0^\infty h(\lambda,\tau)\Big[\int_0^C g(U|\lambda,S)\,dU\Big]f(\lambda)\,d\lambda + \text{ Constant}$$

$$= R(\delta_2,S,\tau,f),$$

by the Lebesgue dominated convergence theorem.

Now consider the decision procedure developed in Section 4.4. This decision procedure is based on a decision function of the form

$$\bar{\Phi}'_{0,1}(U_1,U_2) = \begin{cases} 1, & E(\lambda_1|U_1) < E(\lambda_2|U_2) + \tau \\ 0, & E(\lambda_1|U_1) \geq E(\lambda_2|U_2) + \tau \end{cases} \tag{6.4}$$

which is in the same form of Equation (4.4.7). So for this case, it is only necessary to show that the risk function based on $G_{n_1}(U_1|\lambda_1)$, $G_{n_2}(U_2|\lambda_2)$, $\bar{\Phi}'_{0,1}(U_1,U_2)$, the losses defined by Equation (4.4.4) and prior distributions $F_i(\lambda_i) = \int_0^{\lambda_i} f_i(t)dt$, $i = 1,2$, which is given by

$$\bar{R}_{1(n_1,n_2)}(\bar{\delta}_1,\underline{S}_2,\tau,\underline{f}_2)$$

$$= \int_0^\infty \int_0^\infty \Big[\int_0^\infty \int_0^\infty \bar{\Phi}'_{0,1}(U_1,U_2)(\lambda_1 - \lambda_2 - \tau)dG_{n_1}(U_1|\lambda_1)dG_{n_2}(U_2|\lambda_2)\Big]$$

$$f_1(\lambda_1)f_2(\lambda_2)d\lambda_1 d\lambda_2 + \text{Constant} \tag{6.5}$$

converges to the minimum risk function defined in Equation (4.4.5), which is

$$\bar{R}_1(\delta_1,\underline{S}_2,\tau,\underline{f}_2)$$

$$= \int_0^\infty \int_0^\infty \Big[\int_0^\infty \int_0^\infty \bar{\Phi}'_{0,1}(U_1,U_2)(\lambda_1 - \lambda_2 - \tau)g(U_1|\lambda_1,S_1)g(U_2|\lambda_2,S_2)dU_1dU_2\Big]$$

$$f_1(\lambda_1)f_2(\lambda_2)d\lambda_1 d\lambda_2 + \text{Constant}. \tag{6.6}$$

To see that if $\bar{R}_{1(n_1,n_2)}(\bar{\delta}_1,\underline{S}_2,\tau,\underline{f}_2)$ converges, it converges to $\bar{R}_1(\bar{\delta}_1,\underline{S}_2,\tau,\underline{f}_2)$, it is sufficient to prove the following theorem:

<u>Theorem 6.2</u>: Let $G_{n_1}(U_1|\lambda_1)$ and $G_{n_2}(U_2|\lambda_2)$ be distribution functions which are measurable with respect to λ_1 and λ_2 and which converge to the noncentral chi-square distributions $G(U_i|\lambda_i,S_i) = \int_0^{U_i} g(t|\lambda_i,S_i)dt$, $i = 1,2$, respectively, and let λ_i, $i = 1,2$, be random variables with $F_i(\lambda_i) = \int_0^{\lambda_i} f_i(t)dt$, such that $\int_0^\infty tf_i(t)dt$ exists and $f_i(t)$ is continuous, $i = 1,2$,. Also let $\bar{\Phi}'_{0,1}(U_1,U_2)$, as defined by

(6.4), be the decision function which minimizes the risk function $\overline{R}_1(\overline{\delta}_1,\underline{S}_2,\tau,\underline{f}_2)$, as defined by (6.6). Then at least one subsequence of the risk function $\overline{R}_{1(n_1,n_2)}(\overline{\delta}_1,\underline{S}_2,\tau,\underline{f}_2)$, as defined by (6.5), converges to $\overline{R}_1(\overline{\delta}_1,\underline{S}_2,\tau,\underline{f}_2)$ as n_1 and n_2 become large.

Proof: In the decision function $\overline{\Phi}'_{0,1}(U_1,U_2)$, defined in (6.4), let $\tau(U_2) = E(\lambda_2|U_2) + \tau$. Then (6.4) is in the same form as the decision function, defined by (4.2.7), for the two decision procedure discussed in Section 4.2 when the linear losses defined by Equation (5.3.2) for $k_0 = k_1$ and $\tau = \tau(U_2)$ are used. So that for each U_2, $\overline{\Phi}'_{0,1}(U_1,U_2)$ can be expressed in the same way as the decision function given in Equation (4.2.8) with $C = C_1(U_2)$. That is,

$$\overline{\Phi}'_{0,1}(U_1,U_2) = \begin{cases} 1, & U_1 < C_1(U_2) \\ 0, & U_1 \geq C_2(U_2) \ . \end{cases} \tag{6.7}$$

Similarly, setting $\tau(U_1) = E(\lambda_1|U_1) - \tau$, (6.4) can also be expressed in the form

$$\overline{\Phi}'_{0,1}(U_1,U_2) = \begin{cases} 1, & U_2 > C_2(U_1) \\ 0, & U_2 \leq C_2(U_1) \ . \end{cases} \tag{6.8}$$

Then putting (6.7) and (6.8) into (6.5) yields

$$\overline{R}_{1(n_1,n_2)}(\overline{\delta}_1,\underline{S}_2,\tau,\underline{f}_2)$$

$$= \int_0^\infty \int_0^\infty (\lambda_1 - \lambda_2 - \tau) H(n_1,n_2|\lambda_1,\lambda_2) f_2(\lambda_2) f_1(\lambda_1) d\lambda_1 d\lambda_2 + \text{Constant}$$

where

$$H(n_1,n_2|\lambda_1,\lambda_2) = \int_0^{\infty} \left[\int_0^{C_1(U_2)} dG_{n_1}(U_1|\lambda_1) \right] dG_{n_2}(U_2|\lambda_2)$$

$$= 1 - \int_0^{\infty} \left[\int_0^{C_2(U_1)} dG_{n_2}(U_2|\lambda_2) \right] dG_{n_1}(U_1|\lambda_1) \ .$$

Since the $G_{n_i}(U_i|\lambda_i)$ converge to noncentral chi-square distributions, use of the Lebesgue dominated convergence theorem and the Generalized Second Theorem of Helly, (Gnedenko [1963, p. 267]) leads to

$$\lim_{n_2\to\infty} \lim_{n_1\to\infty} H(n_1,n_2|\lambda_1,\lambda_2) = \lim_{n_2\to\infty} \lim_{n_1\to\infty} \int_0^{\infty} \left[\int_0^{C_1(U_2)} dG_{n_1}(U_1|\lambda_1) \right] dG_{n_2}(U_2|\lambda_2)$$

$$= \lim_{n_2\to\infty} \int_0^{\infty} \left[\int_0^{C_1(U_2)} g(U_1|\lambda_1,S_1)\,dU_1 \right] dG_{n_2}(U_2|\lambda_2)$$

$$= \int_0^{\infty} \left[\int_0^{C_1(U_2)} g(U_1|\lambda_1,S_1)\,dU_1 \right] g(U_2|\lambda_2,S_2)\,dU_2$$

$$= \int_0^{\infty} \int_0^{\infty} \bar{\Phi}'_{0,1}(U_1,U_2)\, g(U_1|\lambda_1,S_1)\, g(U_2|\lambda_2,S_2)\,dU_1 dU_2$$

$$\equiv H(\lambda_1,\lambda_2), \tag{6.9}$$

and similarly

$$\lim_{n_1\to\infty} \lim_{n_2\to\infty} H(n_1,n_2|\lambda_1,\lambda_2) = 1 - \int_0^{\infty} \left[\int_0^{C_2(U_1)} g(U_2|\lambda_2,S_2)\,dU_2 \right] g(U_1|\lambda_1,S_1)\,dU_1$$

$$\equiv H(\lambda_1,\lambda_2) \ . \tag{6.10}$$

Next consider the sequence $H(n_1,n_2|\lambda_1,\lambda_2)$ jointly in (n_1,n_2). Since $0 \le H(n_1,n_2|\lambda_1,\lambda_2) \le 1$ for all (n_1,n_2), the First Theorem of Helly, (Gendenko [1963, p.264]), indicates there exists at least one subsequence (n_{1_k},n_{2_k}) such that $\lim_{(n_{1_k},n_{2_k}) \to (\infty,\infty)} H(n_{1_k},n_{2_k}|\lambda_1,\lambda_2)$ converges. (It should be noted here, that in the First Theorem of Helly, as stated in Gnedenko, it is specified that $H(n_1,n_2,|\lambda_1,\lambda_2)$ be an increasing function of (λ_1,λ_2), but this is required only so that the limiting function will also be increasing in (λ_1,λ_2), which is not needed in this context.) Since both the iterated limits of $H(n_1,n_2|\lambda_1,\lambda_2)$ converge to $H(\lambda_1,\lambda_2)$, the joint limit of the subsequence $H(n_{1_k},n_{2_k}|\lambda_1,\lambda_2)$ also converges to $H(\lambda_1,\lambda_2)$. Thus

$$\lim_{(n_{1_k},n_{2_k}) \to (\infty,\infty)} \overline{R}_{1(n_{1_k},n_{2_k})}(\overline{\delta}_1,\underline{S}_2,\tau,\underline{f}_2)$$

$$= \lim_{n_1\to\infty} \lim_{n_2\to\infty} \overline{R}_{1(n_1,n_2)}(\overline{\delta}_1,\underline{S}_2,\tau,\underline{f}_2)$$

$$= \overline{R}_1(\overline{\delta}_1,\underline{S}_2,\tau,\underline{f}_2) \quad . \tag{6.11}$$

If either of the iterated limits given in (6.9) and (6.10) is uniform in the other index, the limit in (6.11) holds for (n_1,n_2) also.

So when the decision procedure developed in Section 4.4 is used as the decision procedure for the distributions $G_{n_1}(U_1|\lambda_1)$ and $G_{n_2}(U_2|\lambda_2)$ which only have asymptotic noncentral chi-square distributions, the approximate minimum risk function based on them and $\overline{\Phi}_{0,1}(U_1,U_2)$, when it converges, approaches the minimum risk function for the exact noncentral chi-square distribution.

VII. APPLICATIONS OF THE DECISION PROCEDURES

Two applications of the decision procedures discussed in Chapter III will be presented in this chapter. The first one is a genetic study for which it is reasonable to make assumptions which lead to the noncentral chi-square distribution for any sample size. The second application is a goodness-of-fit test which involves the use of $-2\ln\lambda_n$, where λ_n represents the likelihood ratio statistic, which is asymptotically distributed as a noncentral chi-square under local alternatives as shown in Part 1. In addition to meeting the distributional assumptions required for the use of the decision procedures developed herein, these examples are appropriate because in each of them the noncentrality parameter is a measure of the characteristic about which a decision is desired.

In the first example, human chromosomes must be described as normal or abnormal. At present only the twenty-two nonsex chromosomes will be included so that both males and females can be studied jointly. Each of these chromosomes has two long arms and two short arms. The main aim of the research is to correctly classify the chromosomes based on the average long and short arm measurements instead of requiring individual examination by a technician. So far 1440 sets of chromosomes have been observed from one hundred people and more subjects will be sampled before any chromosomes are examined for abnormality.

Let $\underline{\eta}$ be the vector of average chromosome arm lengths for an individual. Then $\underline{\eta}$ will have forty-four elements, that is, a long and short arm measurement for each of the twenty-two nonsex chromosomes. Several assumptions will be made about the probability distribution of $\underline{\eta}$. The assumptions are as follows: (1) If $\underline{\eta}$ is a vector of measurements of normal chromosomes, then $\underline{\eta}$ is a random vector from a multivariate normal distribution with mean vector $\underline{\mu}$ and variance-covariance matrix $\underline{\Sigma}$; (2) If $\underline{\eta}$ is a vector of measurements of abnormal chromosomes, then $\underline{\eta}$ is a random vector from a multivariate normal distribution with mean vector $\underline{\mu}^*$ and variance-covariance matrix $\underline{\Sigma}$; (3) The vector $\underline{\eta}$ is made up of five subvectors $\underline{\eta}_1$, $\underline{\eta}_2$, $\underline{\eta}_3$, $\underline{\eta}_4$, and $\underline{\eta}_5$ which are statistically independent. The values of $\underline{\mu}^*$ will be different for different kinds of abnormalities. The breakdown of the chromosomes into the five groups was done by visual inspection by biologists before measurements were taken on any chromosomes.

Let N be the number of sets of chromosomes observed and let $\hat{\underline{\eta}}$ be the maximum likelihood estimate of $\underline{\eta}$. Then because of the assumptions (1), (2), and (3), decisions on the normality or abnormality of chromosomes or a subset of the chromosomes are based on the noncentral chi-square random variable

$$U = N(\hat{\underline{\eta}} - \underline{\mu})' \, \underline{\Sigma}^{-1} \, (\hat{\underline{\eta}} - \underline{\mu}) \tag{7.1}$$

whose distribution has noncentrality parameter

$$\begin{aligned} \lambda &= N\gamma \\ &= N(\underline{\mu}^* - \underline{\mu})' \, \underline{\Sigma}^{-1} \, (\underline{\mu}^* - \underline{\mu}) \end{aligned} \tag{7.2}$$

if the chromosomes observed are abnormal or

$$\lambda = 0$$

if the chromosomes observed are normal. The function γ is a measure of the distance between the vectors $\underline{\mu}^*$ and $\underline{\mu}$ and thus can be used as a basis of decisions on the normality of the chromosomes. So for the remainder of the discussion on this area of application, γ will be considered to be in the form defined in (7.2), with the exception that $\underline{\eta}$ may be replaced by $\underline{\eta}_i$, $i = 1,\ldots,5$, with corresponding changes $\underline{\mu}^*$, $\underline{\mu}$, and $\underline{\Sigma}$.

Before specifying the type of decisions to be made about γ, the selection of the initial prior distribution of γ will be considered. The basic form of the initial prior will be assumed to be that of a gamma, so the tables given in the Appendix can be used. This is not an unreasonable assumption, for even if a group of individuals do not have abnormal chromosomes, the mean lengths of their arms can be expected to vary about $\underline{\mu}$ in a manner in accord with a normal probability distribution. Then γ would be a central chi-square random variable, so that its distribution would be a particular type of gamma distribution. Two possible approaches to determine the initial prior are as follows: (1) Assume that γ has a central chi-square distribution; and (2) Using the data on the 1440 sets of chromosomes observed initially, estimate the mean and variance of γ and then set these equal to the mean $p(2\theta)$ and the variance $p(2\theta)^2$, respectively, of the gamma and solve for p and θ.

Now the types of decisions will be considered in two parts: (1) Decisions of the type developed in Sections 4.2 and 4.3; and (2) Decisions of the type developed in Section 4.4. The first group of decisions to be considered will be decisions regarding the normality or extent of abnormality of a particular group of individual's chromosomes or a subset of them. These decisions will be based on N sets of chromosomes which will be observed from a particular group of individuals. Now let d_0 be the decision that the chromosomes are normal and d_1 be the decision that they are abnormal. Then d_0 will be the decision that $\gamma \in \{\gamma | 0 \leq \gamma < \tau\}$ and d_1 will be the decision that $\gamma \in \{\gamma | \tau \leq \gamma \leq \infty\}$. The value of τ would be specified by the experimenter in the following manner. He will specify a tail probability for the central chi-square distribution and τ will be the value that cuts off this probability. This method is used because if there are no abnormalities present in the chromosomes, then the γ will act like a central chi-square random variable. Later when the experimenter has more experience, including experience with abnormal chromosomes, he may wish to redefine the τ values. It is also possible that the experimenter will be able to define degrees of abnormality in the chromosomes. That is, the interval $[\tau,\infty)$ could be broken up into several subintervals, with each subinterval indicating a different degree of abnormality. So let d_0 be the decision that the chromosomes are normal and let d_i be the decision that the chromosomes have abnormality of degree i, $i = 1,\ldots,k$. Then d_0 will be the decision that $\gamma \in \{\gamma | 0 \leq \gamma < \tau_1\}$ and d_i will be the decision that $\gamma \in \{\gamma | \tau_i \leq \gamma < \tau_{i+1}\}$, $i = 1,\ldots,k$, where $\tau_{k+1} = +\infty$. In this case the τ values will be defined by the experimenter based on his past experience.

The second set of decisions will be used to compare apparent abnormalities in the chromosomes associated with different types of physical or mental abnormalities. The methods of Section 4.4 will be used to select the appropriate decisions in each comparison. For this decision procedure let γ_1 be the distance measure associated with one set of chromosomes and γ_2 be the distance measure associated with the other set. Then there are three decisions which can be considered; (1) d_0, the decision that the chromosomes associated with the first physical or mental abnormality, exhibit less abnormality than the chromosomes associated with the second physical or mental abnormality; (2) d_1, the decision that the abnormalities exhibited in both groups of chromosomes are about equal; and (3) d_2, the decision that the first group of chromosomes exhibit more abnormalities than the second group of chromosomes. Then d_0 is the decision that $(\gamma_1,\gamma_2) \in \{(\gamma_1,\gamma_2)|\gamma_2>\gamma_1+\tau\}$, d_1 is the decision that $(\gamma_1,\gamma_2) \in \{(\gamma_1,\gamma_2)||\gamma_1-\gamma_2| < \tau\}$ and d_2 is the decision that $(\gamma_1,\gamma_2) \in \{(\gamma_1,\gamma_2)|\gamma_1 > \gamma_2 + \tau\}$. The value of τ would be specified by the experimenter, since at the time that he would be ready to make these types of decisions, he would have sufficient experience to select τ.

The purpose of decision procedures of this type is to aid the biologists in linking apparent physical or mental abnormalities with abnormalities in the chromosomes so that if abnormalities are indicated in one of the subgroups of the chromosomes, future research can be concentrated on these subgroups.

In the second example a two-decision procedure will be considered. The two decisions are as follows: (1) d_0, an unknown distribution G has approximately the specified form F_0; and (2) d_1, the distribution G does not have the form F_0. An appropriate method for considering these decisions is to divide the domain S of F_0 into k subsets $S_1,\ldots,S_k$ and consider decisions based on whether $G(S_i) = F_0(S_i)$, $i = 1,\ldots,k-1$. This can be done by using the theory developed in Part 1, that is, using the asymptotic distribution of the likelihood ratio statistic

$$-2\ln\lambda_n = -2\ln\left\{\frac{\prod_{i=1}^{k}\{F_0(S_i)\}^{n_i}}{\prod_{i=1}^{k}\{\frac{n_i}{n}\}^{n_i}}\right\} \tag{7.3}$$

under the local alternatives

$$\sqrt{n}\,(G(S_i) - F_0(S_i)) \to \delta_i \tag{7.4}$$

as $n \to \infty$, $i = 1,\ldots,k-1$.

The form of (7.3) satisfies all the conditions specified in Part I, so that $-2\ln\lambda_n$, as specified by (7.3), under the local alternatives (7.4) has an asymptotic noncentral chi-square distribution with k-1 degrees of freedom and noncentrality parameter

$$\lambda = \sum_{i=1}^{k}\left[\delta_i^2 / F_0(S_i)\right]. \tag{7.5}$$

Then the decision procedure can be specified according to the magnitude of λ in the following manner: (1) Let d_0 be the decision that $\lambda \in \{\lambda | 0 \leq \lambda < \tau\}$ and (2) d_1 be the decision that $\lambda \in \{\lambda | \tau \leq \lambda \leq \infty\}$, where τ is a positive tolerance. The value of λ can be considered as a measure of the distance between G and F_0 and thus an appropriate measure of how close G is to F_0. The value of τ will depend on how close the experimenter wants G to be to F_0 before he will say that G is approximately equal to F_0. One method of determining possible τ values would be to approximate δ_i by $G(S_i) - F_0(S_i)$ and then have the experimenter specify what proportion r_i of $F_0(S_i)$ he would allow $G(S_i)$ to differ from $F_0(S_i)$ in each of the k subsets $S_1,\ldots,S_k$. That is, he would allow

$$F_0(S_i) - r_i F_0(S_i) \leq G(S_i) \leq F_0(S_i) + r_i F_0(S_i) \tag{7.6}$$

and will still be able to say that G is approximately equal to F_0. Then τ would be taken as

$$\tau = \sum_{i=1}^{k} (r_i F_0(S_i))^2 \,/\, F_0(S_i)$$

$$= \sum_{i=1}^{k} r_i^2 F_0(S_i) \;.$$

For an initial experiment very little would be known about the possible distribution of λ, so to be able to use the tables given in the Appendix it could be assumed that λ has a gamma prior distribution with parameters p and θ and the parameters p and θ could be specified so that τ would be the median of the gamma prior distribution, as was discussed in Section 5.2. This procedure would assign equal prior probability to the two decisions and thus would favor neither of them.

VIII. SUMMARY

Critical values are given for two minimum average risk decision procedures with the following respective decision regions: (1) $\{\lambda | 0 \leq \lambda < \tau\}$ or $\{\lambda | \tau \leq \lambda < \infty\}$ and (2) $\{(\lambda_1,\lambda_2) | E(\lambda_1 | U_1) < E(\lambda_2 | U_2) + \tau\}$ or $\{(\lambda_1,\lambda_2) | E(\lambda_1 | U_1) \geq E(\lambda_2 | U_2) + \tau\}$, where U, U_1, and U_2 are random variables with noncentral chi-square distributions with S, S_1 and S_2 degrees of freedom and noncentrality parameters λ, λ_1, and λ_2, respectively, where $0 \leq \lambda$, λ_1, $\lambda_2 \leq \infty$. These critical values are determined for the case when the loss function is a linear regret function and the noncentrality parameter has a gamma distribution. These decision procedures were extended by the use of the Restricted Product method of Lehmann [1957] to give several multiple decision procedures involving several independent noncentral chi-square random variables.

When the noncentrality parameter has a gamma distribution and the posterior distribution of the noncentrality parameter is to be used for the prior distribution of a future experiment, a way of approximating the posterior distribution by a gamma distribution was considered. This approximation was made so that the same set of critical values can be used for each experiment.

To aid the experimenter in determining his sample size, an approximate method for determining an optimum sample size, that is, the sample size which approximately minimizes a specified cost function, has been provided. The cost functions were formulated under the assumption that the loss function is a linear regret function and that the prior distribution of the noncentrality parameter is a gamma distribution.

To provide a larger area of possible application, it was shown that under certain conditions, the decision procedures could be applied to a random variable which had only an asymptotic noncentral chi-square distribution.

The decision procedures discussed in this paper are mainly decision procedures on a distance function which will usually measure the distance between the parameters or a subset of the parameters of a distribution under question and a standard distribution. Two such applications were given, one involving a noncentral chi-square distribution exactly, the other only asymptotically.

APPENDIX

1 TABLES OF C* FOR WHICH H(C*, Δ*, P, S) = R

Table 1.1.1. - C* for Δ* = 1.0 and S = 1.0

R \ P	0.10	0.30	0.50	0.70	0.90	1.10	1.30	1.50
19.0	6.55	3.23	2.52	1.97	1.50	1.11	0.79	0.54
39.0	7.78	3.87	3.14	2.55	2.04	1.60	1.22	0.90
59.0	7.70	4.27	3.52	2.91	2.38	1.91	1.50	1.15
79.0	8.01	4.55	3.80	3.18	2.63	2.14	1.71	1.34
99.0	8.24	4.77	4.02	3.38	2.83	2.33	1.89	1.49

Table 1.1.2. - C* for Δ* = 2.0 and S = 1.0

R \ P	0.10	0.30	0.50	0.70	0.90	1.10	1.30	1.50
19.0	8.33	4.90	4.29	3.75	3.26	2.82	2.41	2.04
39.0	9.20	5.69	5.06	4.51	4.00	3.53	3.09	2.68
59.0	9.70	6.16	5.53	4.97	4.45	3.96	3.50	3.08
79.0	10.06	6.50	5.87	5.30	4.77	4.27	3.80	3.37
99.0	10.34	6.77	6.14	5.56	5.02	4.52	4.04	3.60

Table 1.1.3. - C* for Δ* = 3.0 and S = 1.0

R \ P	0.10	0.30	0.50	0.70	0.90	1.10	1.30	1.50
19.0	10.05	6.40	5.84	5.33	4.85	4.40	3.97	3.57
39.0	11.03	7.30	6.74	6.22	5.72	5.25	4.79	4.37
59.0	11.61	7.84	7.28	6.74	6.24	5.75	5.29	4.85
79.0	12.01	8.23	7.66	7.12	6.61	6.12	5.65	5.20
99.0	12.33	8.53	7.96	7.42	6.90	6.40	5.93	5.47

Table 1.1.4. - C* for Δ* = 4.0 and S = 1.0

R \ p	0.10	0.30	0.50	0.70	0.90	1.10	1.30	1.50
19.0	11.72	7.82	7.30	6.81	6.34	5.89	5.46	5.04
39.0	12.80	8.83	8.30	7.79	7.31	6.84	6.39	5.95
59.0	13.43	9.42	8.89	8.38	7.88	7.40	6.94	6.50
79.0	13.87	9.85	9.31	8.79	8.29	7.81	7.34	6.89
99.0	14.21	10.18	9.64	9.12	8.62	8.13	7.65	7.20

Table 1.1.5. - C* for Δ* = 5.0 and S = 1.0

R \ p	0.10	0.30	0.50	0.70	0.90	1.10	1.30	1.50
19.0	13.34	9.21	8.71	8.23	7.76	7.31	6.88	6.45
39.0	14.50	10.29	9.79	9.29	8.81	8.35	7.90	7.46
59.0	15.17	10.94	10.42	9.93	9.44	8.97	8.51	8.06
79.0	15.65	11.40	10.88	10.38	9.89	9.41	8.94	8.49
99.0	19.21	11.76	11.24	10.73	10.23	9.75	9.28	8.82

Table 1.2.1. - C* for Δ* = 1.0 and S = 2.0

R \ p	0.20	0.40	0.60	0.80	1.00	1.20	1.40	1.60	1.80	2.00
19.0	4.69	3.69	2.95	2.35	1.83	1.39	1.01	0.70	0.44	0.23
39.0	5.39	4.38	3.62	2.98	2.42	1.93	1.51	1.14	0.82	0.55
59.0	5.80	4.79	4.02	3.36	2.78	2.27	1.82	1.42	1.07	0.77
79.0	6.10	5.09	4.31	3.64	3.05	2.52	2.05	1.63	1.27	0.94
99.0	6.33	5.32	4.53	3.86	3.26	2.72	2.24	1.81	1.42	1.09

Table 1.2.2. - C* for Δ* = 2.0 and S = 2.0

R \ p	0.20	0.40	0.60	0.80	1.00	1.20	1.40	1.60	1.80	2.00
19.0	6.19	5.38	4.74	4.17	3.65	3.18	2.74	2.34	1.98	1.65
39.0	6.98	6.18	5.53	4.94	4.41	3.91	3.44	3.01	2.61	2.24
59.0	7.45	6.66	6.01	5.41	4.86	4.35	3.87	3.42	3.00	2.61
79.0	7.79	7.01	6.35	5.75	5.19	4.67	4.18	3.72	3.29	2.89
99.0	8.06	7.28	6.62	6.01	5.45	4.92	4.43	3.96	3.52	3.10

Table 1.2.3. - C* for Δ* = 3.0 and S = 2.0

R \ p	0.20	0.40	0.60	0.80	1.00	1.20	1.40	1.60	1.80	2.00
19.0	7.52	6.85	6.27	5.73	5.23	4.76	4.32	3.89	3.50	3.12
39.0	8.42	7.76	7.17	6.63	6.11	5.62	5.15	4.71	4.28	3.88
59.0	8.95	8.30	7.71	7.16	6.63	6.13	5.66	5.20	4.76	4.34
79.0	9.34	8.69	8.10	7.54	7.01	6.50	6.02	5.55	5.11	4.68
99.0	9.64	8.99	8.40	7.84	7.31	6.79	6.30	5.83	5.38	4.94

Table 1.2.4. - C* for Δ* = 4.0 and S = 2.0

R \ p	0.20	0.40	0.60	0.80	1.00	1.20	1.40	1.60	1.80	2.00
19.0	8.83	8.24	7.71	7.20	6.71	6.24	5.80	5.37	4.96	4.56
39.0	9.83	9.25	8.71	8.19	7.69	7.20	6.74	6.29	5.86	5.44
59.0	10.43	9.85	9.30	8.77	8.27	7.78	7.30	6.84	6.40	5.97
79.0	10.86	10.28	9.73	9.19	8.68	8.19	7.71	7.24	6.79	6.36
99.0	11.19	10.61	10.06	9.52	9.01	8.51	8.02	7.55	7.10	6.66

Table 1.2.5. - C* for Δ* = 5.0 and S = 2.0

R \ p	0.20	0.40	0.60	0.80	1.00	1.20	1.40	1.60	1.80	2.00
19.0	10.15	9.60	9.09	8.60	8.12	7.66	7.21	6.78	6.36	5.96
39.0	11.24	10.70	10.18	9.67	9.18	8.71	8.25	7.80	7.36	6.94
59.0	11.89	11.34	10.82	10.31	9.81	9.33	8.86	8.41	7.96	7.53
79.0	12.35	11.81	11.28	10.76	10.26	9.78	9.30	8.84	8.39	7.95
99.0	12.71	12.17	11.64	11.12	10.62	10.12	9.64	9.18	8.72	8.28

Table 1.3.1. - C* for Δ* = 1.0 and S = 3.0

R \ p	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30	2.50
19.0	3.24	2.59	2.04	1.55	1.13	0.77	0.46	0.20	0.00	0.00
39.0	3.95	3.27	2.68	2.15	1.69	1.27	0.91	0.60	0.34	0.12
59.0	4.37	3.69	3.07	2.52	2.03	1.59	1.20	0.86	0.57	0.32
79.0	4.68	3.98	3.36	2.80	2.29	1.83	1.42	1.06	0.74	0.47
99.0	4.92	4.21	3.58	3.01	2.49	2.02	1.60	1.22	0.89	0.60

Table 1.3.2. - C* for Δ^* = 2.0 and S = 3.0

R \ p	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30	2.50
19.0	5.11	4.51	3.97	3.47	3.01	2.59	2.20	1.84	1.51	1.22
39.0	5.92	5.31	4.75	4.23	3.74	3.29	2.86	2.47	2.11	1.77
59.0	6.41	5.79	5.22	4.69	4.19	3.72	3.27	2.86	2.48	2.12
79.0	6.76	6.14	5.56	5.02	4.51	4.03	3.57	3.15	2.75	2.38
99.0	7.03	6.41	5.83	5.28	4.76	4.27	3.81	3.38	2.97	2.59

Table 1.3.3. - C* for Δ^* = 3.0 and S = 3.0

R \ p	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30	2.50
19.0	6.65	6.09	5.57	5.08	4.62	4.18	3.77	3.37	3.00	2.66
39.0	7.56	7.00	6.47	5.96	5.48	5.02	4.58	4.16	3.76	3.38
59.0	8.11	7.54	7.00	6.48	5.99	5.52	5.07	4.63	4.22	3.83
79.0	8.50	7.93	7.38	6.86	6.36	5.88	5.42	4.98	4.56	4.15
99.0	8.81	8.23	7.68	7.16	6.65	6.16	5.70	5.25	4.82	4.41

Table 1.3.4. - C* for Δ^* = 4.0 and S = 3.0

R \ p	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30	2.50
19.0	8.08	7.55	7.05	6.57	6.11	5.67	5.24	4.84	4.45	4.08
39.0	9.09	8.55	8.04	7.55	7.07	6.61	6.17	5.74	5.33	4.93
59.0	9.69	9.15	8.63	8.13	7.64	7.17	6.72	6.28	5.86	5.45
79.0	10.12	9.57	9.05	8.54	8.05	7.57	7.11	6.67	6.24	5.82
99.0	10.45	9.91	9.38	8.87	8.37	7.89	7.42	6.97	6.53	6.11

Table 1.3.5. - C* for Δ^* = 5.0 and S = 3.0

R \ p	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30	2.50
19.0	9.46	8.95	8.46	7.99	7.53	7.09	6.66	6.25	5.85	5.46
39.0	10.55	10.04	9.54	9.05	8.58	8.12	7.68	7.25	6.83	6.42
59.0	11.20	10.68	10.17	9.68	9.20	8.74	8.28	7.84	7.41	7.00
79.0	11.66	11.14	10.63	10.13	9.65	9.17	8.72	8.27	7.83	7.41
99.0	12.02	11.49	10.98	10.48	9.99	9.52	9.05	8.60	8.16	7.73

Table 1.4.1. - C* for Δ^* = 1.0 and S = 4.0

R \ p	1.20	1.40	1.60	1.80	2.00	2.20	2.40	2.60	2.80	3.00
19.0	2.16	1.63	1.17	0.77	0.42	0.12	0.00	0.00	0.00	0.00
39.0	2.86	2.30	1.79	1.34	0.94	0.59	0.29	0.03	0.00	0.00
59.0	3.28	2.70	2.17	1.70	1.27	0.89	0.56	0.27	0.02	0.00
79.0	3.58	2.99	2.45	1.96	1.51	1.11	0.76	0.45	0.19	0.00
99.0	3.82	3.22	2.67	2.16	1.71	1.29	0.93	0.60	0.32	0.07

Table 1.4.2. - C* for Δ^* = 2.0 and S = 4.0

R \ p	1.20	1.40	1.60	1.80	2.00	2.20	2.40	2.60	2.80	3.00
19.0	4.23	3.71	3.23	2.78	2.37	1.98	1.63	1.31	1.02	0.76
39.0	5.05	4.50	3.99	3.52	3.07	2.66	2.27	1.91	1.58	1.28
59.0	5.53	4.98	4.46	3.96	3.50	3.07	2.66	2.29	1.94	1.61
79.0	5.88	5.32	4.79	4.29	3.81	3.37	2.95	2.56	2.20	1.86
99.0	6.16	5.59	5.05	4.54	4.06	3.61	3.18	2.78	2.40	2.06

Table 1.4.3. - C* for Δ^* = 3.0 and S = 4.0

R \ p	1.20	1.40	1.60	1.80	2.00	2.20	2.40	2.60	2.80	3.00
19.0	5.88	5.37	4.89	4.44	4.00	3.60	3.21	2.84	2.50	2.18
39.0	6.79	6.27	5.77	5.29	4.84	4.41	3.99	3.60	3.23	2.88
59.0	7.34	6.81	6.30	5.81	5.34	4.90	4.47	4.06	3.67	3.31
79.0	7.73	7.19	6.67	6.18	5.70	5.25	4.81	4.40	4.00	3.62
99.0	8.03	7.49	6.97	6.47	5.99	5.53	5.08	4.66	4.25	3.87

Table 1.4.4. - C* for Δ^* = 4.0 and S = 4.0

R \ p	1.20	1.40	1.60	1.80	2.00	2.20	2.40	2.60	2.80	3.00
19.0	7.37	6.88	6.40	5.95	5.51	5.09	4.69	4.30	3.94	3.58
39.0	8.38	7.87	7.38	6.91	6.45	6.01	5.59	5.18	4.79	4.41
59.0	8.97	8.46	7.96	7.48	7.01	6.56	6.13	5.71	5.31	4.92
79.0	9.40	8.88	8.38	7.89	7.42	6.96	6.52	6.09	5.68	5.28
99.0	9.73	9.21	8.70	8.21	7.73	7.27	6.82	6.39	5.97	5.56

Table 1.4.5. - C* for Δ^* = 5.0 and S = 4.0

R \ p	1.20	1.40	1.60	1.80	2.00	2.20	2.40	2.60	2.80	3.00
19.0	8.79	8.30	7.84	7.38	6.94	6.52	6.11	5.71	5.33	4.96
39.0	9.87	9.38	8.90	8.43	7.97	7.53	7.11	6.69	6.29	5.90
59.0	10.52	10.01	9.53	9.05	8.59	8.14	7.70	7.27	6.86	6.46
79.0	10.97	10.47	9.98	9.49	9.03	8.57	8.13	7.69	7.27	6.86
99.0	11.33	10.82	10.32	9.84	9.37	8.91	8.46	8.02	7.59	7.18

Table 1.5.1. - C* for Δ* = 1.0 and S = 5.0

R \ p	1.70	1.90	2.10	2.30	2.50	2.70	2.90	3.10	3.30	3.50
19.0	1.16	0.71	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00
39.0	1.84	1.35	0.92	0.53	0.19	0.00	0.00	0.00	0.00	0.00
59.0	2.25	1.74	1.28	0.87	0.50	0.18	0.00	0.00	0.00	0.00
79.0	2.55	2.02	1.55	1.12	0.73	0.38	0.08	0.00	0.00	0.00
99.0	2.78	2.25	1.76	1.31	0.91	0.55	0.24	0.00	0.00	0.00

Table 1.5.2. - C* for Δ* = 2.0 and S = 5.0

R \ p	1.70	1.90	2.10	2.30	2.50	2.70	2.90	3.10	3.30	3.50
19.0	3.40	2.93	2.49	2.09	1.71	1.37	1.06	0.77	0.52	0.29
39.0	4.20	3.71	3.24	2.80	2.39	2.02	1.66	1.34	1.04	0.77
59.0	4.68	4.17	3.69	3.24	2.81	2.41	2.04	1.70	1.38	1.09
79.0	5.03	4.51	4.02	3.55	3.12	2.70	2.32	1.96	1.63	1.32
99.0	5.30	4.77	4.27	3.80	3.35	2.93	2.54	2.17	1.82	1.50

Table 1.5.3. - C* for Δ* = 3.0 and S = 5.0

R \ p	1.70	1.90	2.10	2.30	2.50	2.70	2.90	3.10	3.30	3.50
19.0	5.13	4.66	4.21	3.79	3.38	3.00	2.64	2.31	1.99	1.69
39.0	6.04	5.54	5.07	4.63	4.20	3.79	3.40	3.04	2.69	2.36
59.0	6.57	6.07	5.59	5.13	4.69	4.27	3.87	3.48	3.12	2.77
79.0	6.96	6.45	5.96	5.49	5.04	4.61	4.20	3.81	3.43	3.08
99.0	7.26	6.75	6.25	5.78	5.32	4.88	4.46	4.06	3.68	3.31

Table 1.5.4. - C* for Δ^* = 4.0 and S = 5.0

R \ p	1.70	1.90	2.10	2.30	2.50	2.70	2.90	3.10	3.30	3.50
19.0	6.67	6.21	5.76	5.32	4.91	4.51	4.13	3.77	3.42	3.09
39.0	7.67	7.18	6.72	6.27	5.83	5.41	5.01	4.62	4.25	3.89
59.0	8.26	7.77	7.29	6.83	6.38	5.95	5.54	5.14	4.75	4.38
79.0	8.68	8.18	7.70	7.23	6.78	6.34	5.92	5.51	5.11	4.73
99.0	9.01	8.51	8.02	7.55	7.09	6.64	6.21	5.80	5.40	5.01

Table 1.5.5. - C* for Δ^* = 5.0 and S = 5.0

R \ p	1.70	1.90	2.10	2.30	2.50	2.70	2.90	3.10	3.30	3.50
19.0	8.12	7.66	7.21	6.77	6.35	5.95	5.55	5.17	4.81	4.46
39.0	9.20	8.72	8.26	7.81	7.37	6.94	6.53	6.13	5.74	5.37
59.0	9.83	9.35	8.88	8.42	7.97	7.54	7.11	6.70	6.31	5.92
79.0	10.29	9.80	9.32	8.86	8.40	7.96	7.53	7.12	6.71	6.32
99.0	10.64	10.15	9.67	9.20	8.74	8.29	7.86	7.44	7.03	6.63

2 TABLES OF W(p, θ*, S)

Table 2.1. - W(p, θ*, S) for S = 1

θ* \ p	0.10	0.30	0.50	0.70	0.90	1.10	1.30
1.0	0.33	0.74	1.00	1.19	1.34	1.46	1.57
2.0	0.50	0.84	1.00	1.11	1.20	1.27	1.34
3.0	0.66	0.90	1.00	1.07	1.14	1.19	1.24
4.0	0.79	0.93	1.00	1.05	1.10	1.15	1.19
5.0	0.86	0.95	1.00	1.04	1.08	1.12	1.15
6.0	0.90	0.96	1.00	1.04	1.07	1.10	1.13
7.0	0.92	0.97	1.00	1.03	1.06	1.08	1.11
8.0	0.94	0.97	1.00	1.03	1.05	1.07	1.10
9.0	0.95	0.98	1.00	1.02	1.04	1.07	1.09
10.0	0.95	0.98	1.00	1.02	1.04	1.06	1.08
11.0	0.96	0.98	1.00	1.02	1.04	1.05	1.07
12.0	0.96	0.98	1.00	1.02	1.03	1.05	1.07
13.0	0.97	0.98	1.00	1.02	1.03	1.05	1.06
14.0	0.97	0.98	1.00	1.01	1.03	1.04	1.06
15.0	0.97	0.99	1.00	1.01	1.03	1.04	1.05
20.0	0.98	0.99	1.00	1.01	1.02	1.03	1.04
25.0	0.98	0.99	1.00	1.01	1.02	1.02	1.03
30.0	0.99	0.99	1.00	1.01	1.01	1.02	1.03
35.0	0.99	0.99	1.00	1.01	1.01	1.02	1.02
40.0	0.99	0.99	1.00	1.01	1.01	1.01	1.02
45.0	0.99	1.00	1.00	1.00	1.01	1.01	1.02
50.0	0.99	1.00	1.00	1.00	1.01	1.01	1.02
75.0	0.99	1.00	1.00	1.00	1.01	1.01	1.01
100.0	1.00	1.00	1.00	1.00	1.00	1.00	1.01

Table 2.1. - Continued

θ* \ p	1.50	1.70	1.90	2.10	2.30	2.50
1.0	1.67	1.75	1.84	1.91	1.98	2.05
2.0	1.40	1.46	1.51	1.56	1.61	1.65
3.0	1.29	1.33	1.37	1.41	1.44	1.48
4.0	1.22	1.26	1.29	1.32	1.35	1.38
5.0	1.18	1.21	1.24	1.27	1.29	1.32
6.0	1.15	1.18	1.20	1.23	1.25	1.27
7.0	1.13	1.16	1.18	1.20	1.22	1.24
8.0	1.12	1.14	1.16	1.18	1.20	1.21
9.0	1.11	1.12	1.14	1.16	1.18	1.19
10.0	1.10	1.11	1.13	1.14	1.16	1.18
11.0	1.09	1.10	1.12	1.13	1.15	1.16
12.0	1.08	1.09	1.11	1.12	1.14	1.15
13.0	1.07	1.09	1.10	1.11	1.13	1.14
14.0	1.07	1.08	1.09	1.11	1.12	1.13
15.0	1.06	1.08	1.09	1.10	1.11	1.12
20.0	1.05	1.06	1.07	1.08	1.08	1.09
25.0	1.04	1.05	1.05	1.06	1.07	1.08
30.0	1.03	1.04	1.05	1.05	1.06	1.06
35.0	1.03	1.03	1.04	1.04	1.05	1.05
40.0	1.02	1.03	1.03	1.04	1.04	1.05
45.0	1.02	1.03	1.03	1.03	1.04	1.04
50.0	1.02	1.02	1.03	1.03	1.04	1.04
75.0	1.01	1.02	1.02	1.02	1.02	1.03
100.0	1.01	1.01	1.01	1.02	1.02	1.02

Table 2.2. - $W(p, \theta^*, S)$ for $S = 2$

θ^* \ p	0.20	0.40	0.60	0.80	1.00	1.20	1.40	1.60
1.0	0.30	0.52	0.71	0.87	1.00	1.12	1.23	1.33
2.0	0.42	0.64	0.79	0.91	1.00	1.08	1.15	1.22
3.0	0.55	0.74	0.85	0.93	1.00	1.06	1.11	1.16
4.0	0.67	0.81	0.89	0.95	1.00	1.05	1.09	1.13
5.0	0.76	0.85	0.91	0.96	1.00	1.04	1.07	1.11
6.0	0.82	0.88	0.93	0.97	1.00	1.03	1.06	1.09
7.0	0.85	0.90	0.94	0.97	1.00	1.03	1.05	1.08
8.0	0.88	0.92	0.95	0.97	1.00	1.02	1.05	1.07
9.0	0.90	0.93	0.95	0.98	1.00	1.02	1.04	1.06
10.0	0.91	0.94	0.96	0.98	1.00	1.02	1.04	1.06
11.0	0.92	0.94	0.96	0.98	1.00	1.02	1.04	1.05
12.0	0.93	0.95	0.97	0.98	1.00	1.02	1.03	1.05
13.0	0.93	0.95	0.97	0.98	1.00	1.02	1.03	1.04
14.0	0.94	0.95	0.97	0.99	1.00	1.01	1.03	1.04
15.0	0.94	0.96	0.97	0.99	1.00	1.01	1.03	1.04
20.0	0.96	0.97	0.98	0.99	1.00	1.01	1.02	1.03
25.0	0.97	0.98	0.98	0.99	1.00	1.01	1.02	1.02
30.0	0.97	0.98	0.99	0.99	1.00	1.01	1.01	1.02
35.0	0.98	0.98	0.99	0.99	1.00	1.01	1.01	1.02
40.0	0.98	0.98	0.99	0.99	1.00	1.00	1.01	1.01
45.0	0.98	0.99	0.99	1.00	1.00	1.00	1.01	1.01
50.0	0.98	0.99	0.99	1.00	1.00	1.00	1.01	1.01
75.0	0.99	0.99	0.99	1.00	1.00	1.00	1.01	1.01
100.0	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.01

Table 2.2. - Continued

θ* \ p	1.80	2.00	2.20	2.40	2.60	2.80	3.00
1.0	1.42	1.50	1.58	1.65	1.72	1.79	1.86
2.0	1.28	1.33	1.39	1.44	1.48	1.53	1.57
3.0	1.21	1.25	1.29	1.33	1.37	1.40	1.43
4.0	1.17	1.20	1.23	1.26	1.30	1.32	1.35
5.0	1.14	1.17	1.19	1.22	1.25	1.27	1.30
6.0	1.12	1.14	1.17	1.19	1.21	1.24	1.26
7.0	1.10	1.13	1.15	1.17	1.19	1.21	1.23
8.0	1.09	1.11	1.13	1.15	1.17	1.19	1.20
9.0	1.08	1.10	1.12	1.14	1.15	1.17	1.18
10.0	1.07	1.09	1.11	1.12	1.14	1.15	1.17
11.0	1.07	1.08	1.10	1.11	1.13	1.14	1.16
12.0	1.06	1.08	1.09	1.10	1.12	1.13	1.14
13.0	1.06	1.07	1.08	1.10	1.11	1.12	1.13
14.0	1.05	1.07	1.08	1.09	1.10	1.11	1.13
15.0	1.05	1.06	1.07	1.09	1.10	1.11	1.12
20.0	1.04	1.05	1.06	1.07	1.07	1.08	1.09
25.0	1.03	1.04	1.05	1.05	1.06	1.07	1.07
30.0	1.03	1.03	1.04	1.04	1.05	1.06	1.06
35.0	1.02	1.03	1.03	1.04	1.04	1.05	1.05
40.0	1.02	1.02	1.03	1.03	1.04	1.04	1.05
45.0	1.02	1.02	1.03	1.03	1.03	1.04	1.04
50.0	1.02	1.02	1.02	1.03	1.03	1.03	1.04
75.0	1.01	1.01	1.02	1.02	1.02	1.02	1.03
100.0	1.01	1.01	1.01	1.01	1.02	1.02	1.02

Table 2.3. - W(p, θ*, S) for S = 3

θ* \ p	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90
1.0	0.43	0.57	0.69	0.80	0.90	1.00	1.09	1.17
2.0	0.53	0.66	0.76	0.85	0.93	1.00	1.07	1.13
3.0	0.63	0.73	0.81	0.88	0.94	1.00	1.05	1.10
4.0	0.70	0.79	0.85	0.91	0.96	1.00	1.04	1.08
5.0	0.76	0.83	0.88	0.92	0.96	1.00	1.03	1.07
6.0	0.81	0.86	0.90	0.94	0.97	1.00	1.03	1.06
7.0	0.84	0.88	0.91	0.94	0.97	1.00	1.03	1.05
8.0	0.86	0.89	0.92	0.95	0.98	1.00	1.02	1.04
9.0	0.88	0.91	0.93	0.96	0.98	1.00	1.02	1.04
10.0	0.89	0.92	0.94	0.96	0.98	1.00	1.02	1.04
11.0	0.90	0.92	0.94	0.96	0.98	1.00	1.02	1.03
12.0	0.91	0.93	0.95	0.97	0.98	1.00	1.02	1.03
13.0	0.92	0.94	0.95	0.97	0.98	1.00	1.01	1.03
14.0	0.93	0.94	0.96	0.97	0.99	1.00	1.01	1.03
15.0	0.93	0.95	0.96	0.97	0.99	1.00	1.01	1.03
20.0	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.02
25.0	0.96	0.97	0.98	0.98	0.99	1.00	1.01	1.02
30.0	0.97	0.97	0.98	0.99	0.99	1.00	1.01	1.01
35.0	0.97	0.98	0.98	0.99	0.99	1.00	1.01	1.01
40.0	0.97	0.98	0.98	0.99	1.00	1.00	1.00	1.01
45.0	0.98	0.98	0.99	0.99	1.00	1.00	1.00	1.01
50.0	0.98	0.98	0.99	0.99	1.00	1.00	1.00	1.01
75.0	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.01
100.0	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00

Table 2.3. - Continued

θ^* \ p	2.10	2.30	2.50	2.70	2.90	3.10	3.30	3.50
1.0	1.25	1.33	1.40	1.47	1.53	1.60	1.66	1.72
2.0	1.18	1.24	1.29	1.33	1.38	1.42	1.47	1.51
3.0	1.14	1.18	1.22	1.26	1.30	1.33	1.36	1.40
4.0	1.12	1.15	1.18	1.21	1.24	1.27	1.30	1.33
5.0	1.10	1.13	1.15	1.18	1.21	1.23	1.26	1.28
6.0	1.08	1.11	1.13	1.16	1.18	1.20	1.22	1.24
7.0	1.07	1.10	1.12	1.14	1.16	1.18	1.20	1.22
8.0	1.07	1.09	1.11	1.12	1.14	1.16	1.18	1.19
9.0	1.06	1.08	1.10	1.11	1.13	1.15	1.16	1.18
10.0	1.05	1.07	1.09	1.10	1.12	1.13	1.15	1.16
11.0	1.05	1.06	1.08	1.09	1.11	1.12	1.14	1.15
12.0	1.05	1.06	1.07	1.09	1.10	1.11	1.13	1.14
13.0	1.04	1.06	1.07	1.08	1.09	1.11	1.12	1.13
14.0	1.04	1.05	1.06	1.08	1.09	1.10	1.11	1.12
15.0	1.04	1.05	1.06	1.07	1.08	1.09	1.10	1.12
20.0	1.03	1.04	1.05	1.06	1.06	1.07	1.08	1.09
25.0	1.02	1.03	1.04	1.04	1.05	1.06	1.07	1.07
30.0	1.02	1.03	1.03	1.04	1.04	1.05	1.06	1.06
35.0	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05
40.0	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05
45.0	1.01	1.02	1.02	1.03	1.03	1.03	1.04	1.04
50.0	1.01	1.02	1.02	1.02	1.03	1.03	1.03	1.04
75.0	1.01	1.01	1.01	1.02	1.02	1.02	1.02	1.03
100.0	1.01	1.01	1.01	1.01	1.01	1.02	1.02	1.02

Table 2.4. - $W(p, \theta^*, S)$ for $S = 4$

θ^* \ p	1.00	1.20	1.40	1.60	1.80	2.00	2.20	2.40
1.0	0.58	0.68	0.76	0.85	0.92	1.00	1.07	1.14
2.0	0.66	0.74	0.81	0.88	0.94	1.00	1.06	1.11
3.0	0.72	0.79	0.85	0.90	0.95	1.00	1.04	1.09
4.0	0.77	0.82	0.87	0.92	0.96	1.00	1.04	1.07
5.0	0.81	0.85	0.89	0.93	0.97	1.00	1.03	1.06
6.0	0.84	0.87	0.91	0.94	0.97	1.00	1.03	1.05
7.0	0.86	0.89	0.92	0.95	0.97	1.00	1.02	1.05
8.0	0.88	0.90	0.93	0.95	0.98	1.00	1.02	1.04
9.0	0.89	0.91	0.94	0.96	0.98	1.00	1.02	1.04
10.0	0.90	0.92	0.94	0.96	0.98	1.00	1.02	1.03
11.0	0.91	0.93	0.95	0.97	0.98	1.00	1.02	1.03
12.0	0.92	0.93	0.95	0.97	0.98	1.00	1.02	1.03
13.0	0.92	0.94	0.96	0.97	0.99	1.00	1.01	1.03
14.0	0.93	0.94	0.96	0.97	0.99	1.00	1.01	1.03
15.0	0.93	0.95	0.96	0.97	0.99	1.00	1.01	1.02
20.0	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.02
25.0	0.96	0.97	0.98	0.98	0.99	1.00	1.01	1.02
30.0	0.97	0.97	0.98	0.99	0.99	1.00	1.01	1.01
35.0	0.97	0.98	0.98	0.99	0.99	1.00	1.01	1.01
40.0	0.97	0.98	0.99	0.99	1.00	1.00	1.00	1.01
45.0	0.98	0.98	0.99	0.99	1.00	1.00	1.00	1.01
50.0	0.98	0.98	0.99	0.99	1.00	1.00	1.00	1.01
75.0	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.01
100.0	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00

Table 2.4. - Continued

θ* \ p	2.60	2.80	3.00	3.20	3.40	3.60	3.80	4.00
1.0	1.21	1.27	1.33	1.39	1.45	1.51	1.56	1.62
2.0	1.16	1.20	1.25	1.29	1.34	1.38	1.42	1.45
3.0	1.13	1.16	1.20	1.23	1.27	1.30	1.33	1.36
4.0	1.10	1.14	1.17	1.20	1.22	1.25	1.28	1.30
5.0	1.09	1.12	1.14	1.17	1.19	1.22	1.24	1.26
6.0	1.08	1.10	1.13	1.15	1.17	1.19	1.21	1.23
7.0	1.07	1.09	1.11	1.13	1.15	1.17	1.19	1.21
8.0	1.06	1.08	1.10	1.12	1.14	1.15	1.17	1.19
9.0	1.06	1.07	1.09	1.11	1.12	1.14	1.16	1.17
10.0	1.05	1.07	1.08	1.10	1.11	1.13	1.14	1.16
11.0	1.05	1.06	1.08	1.09	1.11	1.12	1.13	1.15
12.0	1.04	1.06	1.07	1.08	1.10	1.11	1.12	1.14
13.0	1.04	1.05	1.07	1.08	1.09	1.10	1.11	1.13
14.0	1.04	1.05	1.06	1.07	1.09	1.10	1.11	1.12
15.0	1.04	1.04	1.06	1.07	1.08	1.09	1.10	1.11
20.0	1.03	1.04	1.05	1.05	1.06	1.07	1.08	1.09
25.0	1.02	1.03	1.04	1.04	1.05	1.06	1.06	1.07
30.0	1.02	1.03	1.03	1.04	1.04	1.05	1.05	1.06
35.0	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05
40.0	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05
45.0	1.01	1.02	1.02	1.03	1.03	1.03	1.04	1.04
50.0	1.01	1.02	1.02	1.02	1.03	1.03	1.03	1.04
75.0	1.01	1.01	1.01	1.02	1.02	1.02	1.02	1.03
100.0	1.01	1.01	1.01	1.01	1.01	1.02	1.02	1.02

Table 2.5. - W(p, θ*, S) for S = 5

θ* \ p	1.50	1.70	1.90	2.10	2.30	2.50	2.70	2.90
1.0	0.66	0.74	0.81	0.87	0.94	1.00	1.06	1.12
2.0	0.72	0.78	0.84	0.90	0.95	1.00	1.05	1.09
3.0	0.77	0.82	0.87	0.92	0.96	1.00	1.04	1.08
4.0	0.80	0.85	0.89	0.93	0.97	1.00	1.03	1.07
5.0	0.83	0.87	0.90	0.94	0.97	1.00	1.03	1.06
6.0	0.85	0.89	0.92	0.95	0.97	1.00	1.03	1.05
7.0	0.87	0.90	0.93	0.95	0.98	1.00	1.02	1.04
8.0	0.89	0.91	0.93	0.96	0.98	1.00	1.02	1.04
9.0	0.90	0.92	0.94	0.96	0.98	1.00	1.02	1.04
10.0	0.91	0.93	0.95	0.96	0.98	1.00	1.02	1.03
11.0	0.91	0.93	0.95	0.97	0.98	1.00	1.02	1.03
12.0	0.92	0.94	0.95	0.97	0.99	1.00	1.01	1.03
13.0	0.93	0.94	0.96	0.97	0.99	1.00	1.01	1.03
14.0	0.93	0.95	0.96	0.97	0.99	1.00	1.01	1.03
15.0	0.94	0.95	0.96	0.98	0.99	1.00	1.01	1.02
20.0	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.02
25.0	0.96	0.97	0.98	0.98	0.99	1.00	1.01	1.01
30.0	0.97	0.97	0.98	0.99	0.99	1.00	1.01	1.01
35.0	0.97	0.98	0.98	0.99	0.99	1.00	1.01	1.01
40.0	0.98	0.98	0.99	0.99	1.00	1.00	1.00	1.01
45.0	0.98	0.98	0.99	0.99	1.00	1.00	1.00	1.01
50.0	0.98	0.98	0.99	0.99	1.00	1.00	1.00	1.01
75.0	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.01
100.0	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00

Table 2.5. - Continued

θ* \ p	3.10	3.30	3.50	3.70	3.90	4.10	4.30	4.50
1.0	1.18	1.23	1.29	1.34	1.39	1.44	1.49	1.54
2.0	1.14	1.18	1.22	1.26	1.30	1.34	1.38	1.41
3.0	1.11	1.15	1.18	1.21	1.25	1.28	1.31	1.34
4.0	1.10	1.13	1.15	1.18	1.21	1.23	1.26	1.28
5.0	1.08	1.11	1.13	1.16	1.18	1.20	1.23	1.25
6.0	1.07	1.10	1.12	1.14	1.16	1.18	1.20	1.22
7.0	1.07	1.09	1.11	1.12	1.14	1.16	1.18	1.20
8.0	1.06	1.08	1.10	1.11	1.13	1.15	1.16	1.18
9.0	1.05	1.07	1.09	1.10	1.12	1.13	1.15	1.16
10.0	1.05	1.06	1.08	1.09	1.11	1.12	1.14	1.15
11.0	1.05	1.06	1.07	1.09	1.10	1.11	1.13	1.14
12.0	1.04	1.06	1.07	1.08	1.09	1.11	1.12	1.13
13.0	1.04	1.05	1.06	1.08	1.09	1.10	1.11	1.12
14.0	1.04	1.05	1.06	1.07	1.08	1.09	1.10	1.12
15.0	1.03	1.05	1.06	1.07	1.08	1.09	1.10	1.11
20.0	1.03	1.04	1.04	1.05	1.06	1.07	1.08	1.09
25.0	1.02	1.03	1.04	1.04	1.05	1.06	1.06	1.07
30.0	1.02	1.02	1.03	1.04	1.04	1.05	1.05	1.06
35.0	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05
40.0	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05
45.0	1.01	1.02	1.02	1.03	1.03	1.03	1.04	1.04
50.0	1.01	1.02	1.02	1.02	1.03	1.03	1.03	1.04
75.0	1.01	1.01	1.01	1.02	1.02	1.02	1.02	1.03
100.0	1.01	1.01	1.01	1.01	1.01	1.02	1.02	1.02

REFERENCES

Part 1

Aitchison, J. and Silvey, S. D. [1958]. Maximum-likelihood estimation of parameters subject to restraints. Ann. Math. Statist. 29, 813-828.

Bergström, Harald. [1963]. Limit theorems for convolutions. John Wiley and Sons, New York.

Bradley, R. A. and Gart, J. J. [1962]. The asymptotic properties of ML estimators when sampling from associated populations. Biometrika 49, 205-214.

Chanda, K. C. [1954]. A note on the consistency and maxima of the roots of likelihood equations. Biometrika 41, 56-61.

Chiang, C. L. [1956]. On best regular asymptotically normal estimates. Ann. Math. Statist. 27, 336-351.

Cramér, H. [1946]. Mathematical methods of statistics. Princeton University Press, Princeton, New Jersey.

Doob, J. L. [1935]. The limiting distributions of certain statistics. Ann. Math. Statist. 6, 160-169.

Neyman, J. and Pearson, E. S. [1928]. On the use and interpretation of certain test criteria for purposes of statistical inference. Biometrika 20 A, 175-240, 263-294.

Wald, A. [1943]. Tests of statistical hypotheses concerning several parameters when the number of observations is large. Trans. Amer. Math. Soc. 54, 426-482.

Wilks, S. S. [1938]. The large-sample distribution of the likelihood ratio for testing composite hypotheses. Ann. Math. Statist. 9, 60-62.

Part 2

Cochran, W. C. [1954]. Some methods for strengthening the common χ^2 test. Biometrics 10, 417-451.

Gnedenko, B. V. [1963]. The theory of probability. Chelsea Publishing Company, New York.

Karlin, S. [1956]. Decision theory for Pólya type distributions. Case of two actions, I. Proceedings the Fourth Berkeley Symposium on Mathematical Statistics and Probability I, 115-128.

Karlin, S. and Rubin, H. [1956]. The theory of decision procedures with monotone likelihood ratio. Ann. Math. Statist. 27, 272-299.

Kendall, M. G. and Stuart, A. [1961]. The advanced theory of statistics. Vol. 2. Hafner, New York.

Kunz, K. S. [1957]. Numerial analysis. McGraw-Hill, New York.

Lehmann, E. L. [1957]. A theory of some multiple decision problems I. Ann. Math. Statist. 28, 1-26.

Olmsted, S. M. H. [1959]. Real variables. Appleton-Century-Crofts, New York.

Patnaik, P. B. [1949]. The non-central χ^2- and F-distributions and their applications. Biometrika 36, 202-232.

Royden, H. L. [1963]. Real analysis. Macmillan, New York.

ORNL-4225
UC-32 — Mathematics and Computers

INTERNAL DISTRIBUTION

1. Biology Library
2-4. Central Research Library
5-6. ORNL — Y-12 Technical Library
Document Reference Section
7-26. Laboratory Records Department
27. Laboratory Records, ORNL R.C.
28. A. A. Brooks
29. D. A. Gardiner
30. D. Gosslee
31. T. L. Hebble
32. G. K. Haeuslein
33. A. S. Householder
34. N. C. Hull
35. C. K. Johnson
36. W. H. Jordan
37. M. Kastenbaum
38. C. E. Larson
39. Elmon Leach
40-64. W. E. Lever (Y-12)
65. H. G. MacPherson
66. F. L. Miller, Jr.
67. S. E. Moore
68. M. J. Skinner
69. D. A. Sundberg
70. D. R. Vondy
71. A. M. Weinberg

EXTERNAL DISTRIBUTION

72-73. R. G. Cornell, Department of Statistics, Florida State University, Tallahassee, Florida 32306
74. R. R. Davidson, Department of Mathematics, University of Victoria, Victoria, British Columbia, Canada
75. J. H. Barrett, Mathemathics Department, University of Tennessee, Knoxville, Tennessee
76. Eva Sobocinski, Librarian, Documents Room, Computing Center, University of Notre Dame, Notre Dame, Indiana 46556
77. Laboratory and University Division, AEC, ORO
78-368. Given distribution as shown in TID-4500 under Mathematics and Computers category (25 copies — CFSTI)