Ann. Inst. Statist. Math.
29 (1977), Part A, 359-378

THE LIKELIHOOD RATIO CRITERION AND THE ASYMPTOTIC EXPANSION OF ITS DISTRIBUTION

TAKESI HAYAKAWA

(Received June 7, 1976; revised Feb. 15, 1977)

Summary

Asymptotic expansion of the distribution of the likelihood ratio criterion (LRC) for testing a composite hypothesis is derived under null hypothesis and a correction factor ρ which makes the term of order $1/n$ in the asymptotic expansion of the distribution of it vanish is obtained. The problem is extended to the case of a general composite hypothesis and of Pitman's local alternatives. The asymptotic distribution of LRC for a simple hypothesis is studied under a fixed alternative.

1. Introduction

Let $X=[\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n]$ be an $m\times n$ observation matrix, where $\boldsymbol{x}_\alpha$'s are independently and identically distributed with probability density function $f(\boldsymbol{x}|\boldsymbol{\theta})$ depending on an unknown parameter $\boldsymbol{\theta}=(\theta_1, \cdots, \theta_p)$ with p components.

The problem considered is that of testing a composite hypothesis H_0 : $\boldsymbol{\theta}_2=\boldsymbol{\theta}_{20}$, against H: $\boldsymbol{\theta}_2\neq\boldsymbol{\theta}_{20}$, where $\boldsymbol{\theta}'=(\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')$, $\boldsymbol{\theta}_1'=(\theta_1, \cdots, \theta_q)$, $\boldsymbol{\theta}_2'=(\theta_{q+1}, \cdots, \theta_p)$ and $\boldsymbol{\theta}_{20}'=(\theta_{q+1,0}, \cdots, \theta_{p,0})$ is a specified $(p-q)$-dimensional vector. The likelihood ratio criterion (LRC) was proposed by Neyman and Pearson [8] as a method for testing a composite hypothesis. Wilks [12] showed that the limiting distribution of the log-likelihood ratio criterion $-2\log\lambda$, based on n observations, is a central chi-square distribution with $(p-q)$ degrees of freedom when the null hypothesis H_0 is true. The limiting distribution of the LRC under a particular sequence of Pitman's alternative H_n : $\boldsymbol{\theta}_2=\boldsymbol{\theta}_{20}+\boldsymbol{\varepsilon}/\sqrt{n}$, $\boldsymbol{\varepsilon}=(\varepsilon_{q+1}, \cdots, \varepsilon_p)$, has been studied by Davidson and Lever [2] and the asymptotic expansion of its distribution up to order $1/\sqrt{n}$ under H_n has been studied by Hayakawa [4]. (Peers [9] studied the case of simple hypothesis.)

The purpose of this paper is to present the asymptotic expansion of the distribution of LRC up to the order $1/n$ under the null hypothesis and a correction factor ρ which makes the term of order $1/n$ in the asymptotic distribution of $-2\log\lambda$ vanish in Section 2, and the asymptotic expansion of the distribution of LRC under a general composite hypothesis and Pitman's alternatives, respectively, in Section 3. We also study the asymptotic behavior of the distribution of $-2\log\lambda$ for a simple hypothesis under a fixed alternative H_ε: $\boldsymbol{\theta}=\boldsymbol{\theta}_0+\boldsymbol{\varepsilon}$ in Section 4.

2. Expansion of the LRC under the null hypothesis

In this section we discuss some notations needed for our consideration and the asymptotic expansion of the distribution of a LRC under the composite hypothesis.

Let the likelihood ratio criterion λ for H_0 versus H be

$$\lambda=\prod_{\alpha=1}^{n}\frac{f(\boldsymbol{x}_\alpha|\tilde{\boldsymbol{\theta}}_1,\boldsymbol{\theta}_{20})}{f(\boldsymbol{x}_\alpha|\hat{\boldsymbol{\theta}}_1,\hat{\boldsymbol{\theta}}_2)} \tag{1}$$

where $\hat{\boldsymbol{\theta}}'=(\hat{\boldsymbol{\theta}}_1',\hat{\boldsymbol{\theta}}_2')$ is the maximum likelihood estimator for $\boldsymbol{\theta}$ under H and $\tilde{\boldsymbol{\theta}}_1$ is that for $\boldsymbol{\theta}_1$ under H_0.

Defining the log-likelihood function by $L(\boldsymbol{\theta})=\sum_{\alpha=1}^{n}\log f(\boldsymbol{x}_\alpha|\boldsymbol{\theta})$ we have

$$\log\lambda=L(\tilde{\boldsymbol{\theta}}_1,\boldsymbol{\theta}_{20})-L(\hat{\boldsymbol{\theta}}_1,\hat{\boldsymbol{\theta}}_2)\ .$$

In the asymptotic expansion of the log-likelihood function the following notations and assumptions will be adopted.

(i) The function $L(\boldsymbol{\theta})$ is regular with respect to $\boldsymbol{\theta}$ derivatives up to and including those of fourth order.

(ii) Any function evaluated at the point $\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}$ will be distinguished by the addition of a circumflex.

(iii) Any function evaluated at the point $\boldsymbol{\theta}_1=\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2=\boldsymbol{\theta}_{20}$ will be distinguished by the addition of a tilde.

(iv) Let

$$v_i=\sqrt{n}(\hat{\theta}_i-\theta_i)\ ,\qquad \boldsymbol{v}'=(v_1,\cdots,v_p)\ ,$$

$$w_i=\sqrt{n}(\tilde{\theta}_i-\theta_i)\ ,\qquad \boldsymbol{w}'=(w_1,\cdots,w_q)\ ,$$

i.e., $\boldsymbol{v}=\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})$, $\boldsymbol{w}=\sqrt{n}(\tilde{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_1)$.

(v) Let

$$y_{i_1\cdots i_l}=n^{-l/2}\sum_{\alpha=1}^{n}y_{\alpha i_1\cdots i_l}\ ,\qquad y_{\alpha i_1\cdots i_l}=\frac{\partial^l\log f(\boldsymbol{x}_\alpha|\boldsymbol{\theta})}{\partial\theta_{i_1}\cdots\partial\theta_{i_l}}$$

$$i_1,\cdots,i_l=1,\cdots,p;\ l=1,2,3,4\ .$$

$$\boldsymbol{y}=(y_1,\cdots,y_p)'\ ,\qquad Y=(y_{ij})\ ,\quad Y_{\cdots}=(y_{ijk})\ ,\quad Y_{\cdots\cdot}=(y_{ijkl})$$

$$\kappa_{i,j}=\mathrm{E}(y_iy_j)=\mathrm{E}(y_{\alpha i}y_{\alpha j}), \qquad \kappa_{ij}=\mathrm{E}(y_{ij})=\mathrm{E}(y_{\alpha ij}).$$

$$\kappa_{i,j,k}=\sqrt{n}\ \mathrm{E}(y_iy_jy_k)=\mathrm{E}(y_{\alpha i}y_{\alpha j}y_{\alpha k}),$$

$$\kappa_{i,jk}=\sqrt{n}\ \mathrm{E}(y_iy_{jk})=\mathrm{E}(y_{\alpha i}y_{\alpha jk}),$$

$$\kappa_{ijk}=\sqrt{n}\ \mathrm{E}(y_{ijk})=\mathrm{E}(y_{\alpha ijk}).$$

$$\kappa_{i,j,k,l}=\mathrm{E}(y_{\alpha i}y_{\alpha j}y_{\alpha k}y_{\alpha l}), \qquad \kappa_{i,jkl}=\mathrm{E}(y_{\alpha i}y_{\alpha jkl}),$$

$$\kappa_{i,j,kl}=\mathrm{E}(y_{\alpha i}y_{\alpha j}y_{\alpha kl}), \qquad \kappa_{ij,kl}=\mathrm{E}(y_{\alpha ij}y_{\alpha kl}), \qquad \kappa_{ijkl}=\mathrm{E}(y_{\alpha ijkl}).$$

By the regularity conditions we have

$$K=(\kappa_{i,j})=\mathrm{E}(\boldsymbol{y}\boldsymbol{y}')=-\mathrm{E}(Y)=-K_{..}.$$

We assume the non-singularity of Y with probability one and the positive definiteness of K which is a Fisher's information matrix. The partitioned matrices are denoted as

$$Y=\begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22}\end{bmatrix}\begin{matrix} q \\ p-q\end{matrix} \qquad K=\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22}\end{bmatrix}\begin{matrix} q \\ p-q\end{matrix}.$$
$$\quad q \quad p-q \qquad\qquad q \quad p-q$$

(vi) For three and four suffix quantities, the following summation notations will be used: Let A be a $p\times p$ matrix, and $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ be $p\times 1$ column vectors,

$K_{...}\circ\boldsymbol{a}\circ\boldsymbol{b}\circ\boldsymbol{c}=\sum_{i,j,k}\kappa_{ijk}a_ib_jc_k$ ·····scalor

$K_{...}\circ\boldsymbol{a}\circ\boldsymbol{b}=(\sum_{j,k}\kappa_{ijk}a_jb_k)$ ········column vector with p components designated by index i

$K_{...}\circ\boldsymbol{a}=(\sum_{k}\kappa_{ijk}a_k)$ ·············$p\times p$ matrix with (i, j)th element $\sum_k \kappa_{ijk}a_k$

$K_{...}\circ A\circ\boldsymbol{b}=\sum_{i,j,k}\kappa_{ijk}a_{ij}b_k$ ·········scalor

$$K_{.,.,.}*A*B*C*K_{...}=\sum\kappa_{i,j,k}\kappa_{pqr}a_{ip}b_{jq}c_{kr},$$

$$K_{..,..}\circledast A\circledast B=\sum\kappa_{ij,kl}a_{ik}b_{jl} \quad \text{or} \quad \sum\kappa_{ij,kl}a_{il}b_{jk}.$$

We also use the following abbreviate notations:

$$K_{...}(\circ\boldsymbol{a})^3=K_{...}\circ\boldsymbol{a}\circ\boldsymbol{a}\circ\boldsymbol{a}$$

$$K_{.,.,.}(*A)^3*K_{...}=K_{.,.,.}*A*A*A*K_{...}.$$

Let the set of indices $\{1, 2,\cdots, p\}$ be partitioned into $\{1, 2,\cdots, q\}$ and $\{q+1,\cdots, p\}$ and let vectors and matrices be correspondingly partitioned

$\boldsymbol{a}'=(\boldsymbol{a}_1', \boldsymbol{a}_2')$, etc. Summation with respect to the indices in the divided set will be denoted as follows:

$$K_{111}\circ \boldsymbol{a}_1\circ \boldsymbol{b}_1\circ \boldsymbol{c}_1=\sum_1^q \kappa_{ijk}a_ib_jc_k$$

$$K_{112}\circ \boldsymbol{a}_1\circ \boldsymbol{b}_1\circ \boldsymbol{c}_2=\sum_{i,j=1}^q\sum_{\alpha=q+1}^p \kappa_{ij\alpha}a_ib_jc_\alpha$$

$$K_{212}\circ A_{21}\circ \boldsymbol{a}_2=\sum_{i=1}^q\sum_{\alpha,\beta=q+1}^p \kappa_{\alpha i\beta}a_{\alpha i}a_\beta\ .$$

(vii) Symbols o_p and O_p denote the orders of magnitude in probability sense.

Expanding $L(\tilde{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_{20})$ in a Taylor series about the point $\boldsymbol{\theta}_1=\hat{\boldsymbol{\theta}}_1$, $\boldsymbol{\theta}_2=\hat{\boldsymbol{\theta}}_2$, we have

$$\begin{aligned}2\log\lambda&=2\{L(\tilde{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_{20})-L(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2)\}\\&=(\boldsymbol{u}-\boldsymbol{v})'\hat{Y}(\boldsymbol{u}-\boldsymbol{v})+\frac{1}{3}\hat{Y}_{\dots}(\circ(\boldsymbol{u}-\boldsymbol{v}))^3\\&\quad+\frac{1}{12}\hat{Y}_{\dots.}(\circ(\boldsymbol{u}-\boldsymbol{v}))^4+o_p(1/n)\end{aligned}$$

where

$$\boldsymbol{u}'=(\boldsymbol{w}', 0')\ ,\qquad \boldsymbol{v}'=(\boldsymbol{v}_1', \boldsymbol{v}_2')\ .$$

Noting that

$$\hat{Y}=Y+Y_{\dots}\circ\boldsymbol{v}+\frac{1}{2}Y_{\dots.}(\circ\boldsymbol{v})^2+o_p(1/n)\ ,$$

$$\hat{Y}_{\dots}=Y_{\dots}+Y_{\dots.}\circ\boldsymbol{v}+o_p(1/n)\ ,$$

$$\hat{Y}_{\dots.}=Y_{\dots.}+o_p(1/n)\ ,$$

we have

$$\begin{aligned}(2)\qquad 2\log\lambda&=(\boldsymbol{u}-\boldsymbol{v})'Y(\boldsymbol{u}-\boldsymbol{v})\\&\quad+\frac{1}{3}Y_{\dots}(\circ(\boldsymbol{u}-\boldsymbol{v}))^3+Y_{\dots}\circ(\boldsymbol{u}-\boldsymbol{v})\circ(\boldsymbol{u}-\boldsymbol{v})\circ\boldsymbol{v}\\&\quad+\frac{1}{12}Y_{\dots.}(\circ(\boldsymbol{u}-\boldsymbol{v}))^4+\frac{1}{3}Y_{\dots.}(\circ(\boldsymbol{u}-\boldsymbol{v}))^3\circ\boldsymbol{v}\\&\quad+\frac{1}{2}Y_{\dots.}(\circ(\boldsymbol{u}-\boldsymbol{v}))^2(\circ\boldsymbol{v})^2+o_p(1/n)\ .\end{aligned}$$

The equation satisfied by $\boldsymbol{v}$ can be written as

$$0=\hat{\boldsymbol{y}}=\boldsymbol{y}+Y\boldsymbol{v}+\frac{1}{2}Y_{\dots}\circ\boldsymbol{v}\circ\boldsymbol{v}+\frac{1}{6}Y_{\dots.}(\circ\boldsymbol{v})^3+o_p(1/n)\ .$$

Inverting this equation with respect to $\boldsymbol{v}$, recursively, we have

$$\begin{aligned}(3)\qquad \boldsymbol{v}=&-Y^{-1}\boldsymbol{y}-\frac{1}{2}Y^{-1}[Y_{\dots}(\circ Y^{-1}\boldsymbol{y})^2]\\&-\frac{1}{2}Y^{-1}[Y_{\dots}\circ Y^{-1}\boldsymbol{y}\circ Y^{-1}(Y_{\dots}\circ Y^{-1}\boldsymbol{y}\circ Y^{-1}\boldsymbol{y})]\\&+\frac{1}{6}Y^{-1}Y_{\dots.}(\circ Y^{-1}\boldsymbol{y})^3+o_p(1/n)\ .\end{aligned}$$

Putting $\boldsymbol{\theta}_2=\boldsymbol{\theta}_{20}$ in (3) as we are handling the asymptotic theory of λ under H_0: $\boldsymbol{\theta}_2=\boldsymbol{\theta}_{20}$, we have the asymptotic expansion of $\boldsymbol{v}$ at $(\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')=(\boldsymbol{\theta}_1', \boldsymbol{\theta}_{20}')$. Similary, expanding $\tilde{\boldsymbol{y}}_1$ at $\boldsymbol{\theta}_1=\boldsymbol{\theta}_1, \boldsymbol{\theta}_2=\boldsymbol{\theta}_{20}$,

$$0=\tilde{\boldsymbol{y}}_1=\boldsymbol{y}_1+Y_{11}\circ\boldsymbol{w}+\frac{1}{2}Y_{111}\circ\boldsymbol{w}\circ\boldsymbol{w}+\frac{1}{6}Y_{1111}(\circ\boldsymbol{w})^3+o_p(1/n)$$

and therefore

$$\begin{aligned}(4)\qquad \boldsymbol{u}=&-Z_0\boldsymbol{y}-\frac{1}{2}Z_0(Y_{\dots}(\circ Z_0\boldsymbol{y})^2)\\&-\frac{1}{2}Z_0[Y_{\dots}\circ Z_0\boldsymbol{y}\circ Z_0(Y_{\dots}\circ Z_0\boldsymbol{y}\circ Z_0\boldsymbol{y})]\\&+\frac{1}{6}Z_0(Y_{\dots.}(\circ Z_0\boldsymbol{y})^3)+o_p(1/n)\ ,\end{aligned}$$

where

$$Z_0=\begin{bmatrix}Y_{11}^{-1} & O\\ O & O\end{bmatrix}.$$

Inserting these values of $\boldsymbol{u}$ and $\boldsymbol{v}$ into (2), we have the asymptotic expansion of the likelihood ratio criterion for the composite hypothesis up to order $O_p(1/n)$ as follows:

$$(5)\qquad 2\log\lambda=\boldsymbol{y}'Z\boldsymbol{y}+q_1+q_2+o_p(1/n)\ ,$$

$$q_1=\frac{1}{3}Y_{\dots}(\circ Z\boldsymbol{y})^3+Y_{\dots}\circ Z\boldsymbol{y}\circ Z_0\boldsymbol{y}\circ Y^{-1}\boldsymbol{y}\ ,$$

$$\begin{aligned}q_2=&\frac{1}{4}Y_{\dots}(\circ Y^{-1}\boldsymbol{y})^2\circ Y^{-1}(Y_{\dots}(\circ Y^{-1}\boldsymbol{y})^2)\\&-\frac{1}{4}Y_{\dots}(\circ Z_0\boldsymbol{y})^2\circ Z_0(Y_{\dots}(\circ Z_0\boldsymbol{y})^2)+\frac{1}{12}Y_{\dots.}(\circ Z\boldsymbol{y})^4\end{aligned}$$

$$+\frac{1}{6}Y_{....}(\circ Z\boldsymbol{y})^2\circ Y^{-1}\boldsymbol{y}\circ(Z+3Z_0)\boldsymbol{y}$$

$$-\frac{1}{3}Y_{....}(\circ Y^{-1}\boldsymbol{y})^3\circ Z\boldsymbol{y}\,,$$

where

$$Z=Y^{-1}-Z_0 \qquad \text{and} \qquad ZYZ=Z\,,\quad ZYZ_0=0\,.$$

It should be noted that rank $Z=p-q$.

To obtain the moment generating function (MGF) of $S=-2\log\lambda$ we use the multivariate Edgeworth expansion for the joint density function of $\boldsymbol{y}$, Y, $Y_{...}$ and $Y_{....}$ up to the order $1/n$, which is stated as follows:

$$f_1=f_0\{1+A/\sqrt{n}+B/n\}+o(1/n) \tag{6}$$

where

$$f_0=(2\pi)^{-p/2}|K|^{-1/2}\exp\left(-\frac{1}{2}\boldsymbol{y}'K^{-1}\boldsymbol{y}\right)\prod\delta(y_{ij}-\kappa_{ij})$$
$$\cdot\prod\delta(y_{ijk}-\kappa_{ijk}/\sqrt{n})\prod\delta(y_{ijkl}-\kappa_{ijkl}/n)$$

$$A=\frac{1}{6}\{K_{.,.,.}(\circ K^{-1}\boldsymbol{y})^3-3K_{.,.,.}\circ K^{-1}\circ K^{-1}\boldsymbol{y}\}-K_{.,..}\circ K^{-1}\boldsymbol{y}\circ D_{..}$$

$$B=\frac{1}{2}\{K_{..,..}(\circ D_{..})^2-(\operatorname{tr}KD_{..})^2\}$$
$$+\frac{1}{2}(K_{.,.,..}\circ K^{-1}\circ D_{..}-K_{.,.,..}(\circ K^{-1}\boldsymbol{y})^2\circ D_{..})$$
$$+\frac{1}{2}(p-\boldsymbol{y}'K^{-1}\boldsymbol{y})\operatorname{tr}KD_{..}-K_{.,...}\circ K^{-1}\boldsymbol{y}\circ D_{..}$$
$$-\frac{1}{2}K_{.,..}\circ K^{-1}(K_{.,..}\circ D_{..})\circ D_{..}+\frac{1}{2}(K_{.,..}\circ K^{-1}\boldsymbol{y}\circ D_{..})^2$$
$$+\frac{1}{24}\{K_{.,.,.,.}(\circ K^{-1}\boldsymbol{y})^4-6K_{.,.,.,.}\circ K^{-1}(\circ K^{-1}\boldsymbol{y})^2$$
$$+3K_{.,.,.,.}(\circ K^{-1})^2\}$$
$$-\frac{1}{8}\{p(p+2)-2(p+2)\boldsymbol{y}'K^{-1}\boldsymbol{y}+(\boldsymbol{y}'K^{-1}\boldsymbol{y})^2\}$$
$$-\frac{1}{2}K_{.,.,.}\circ K^{-1}\circ K^{-1}(K_{.,..}\circ D_{..})$$
$$+\frac{1}{2}K_{.,.,.}(\circ K^{-1}\boldsymbol{y})^2\circ K^{-1}(K_{.,..}\circ D_{..})$$

$$
\begin{aligned}
&+\frac{1}{2}K_{.,.,.}\circ K^{-1}\circ K^{-1}\boldsymbol{y}\cdot K_{.,..}\circ K^{-1}\boldsymbol{y}\circ D_{..}\\
&-\frac{1}{2}K_{.,.,.}(\circ K^{-1}\boldsymbol{y})^{3}K_{.,..}\circ K^{-1}\boldsymbol{y}\circ D_{..}\\
&+\frac{1}{72}\{-9K_{.,.,.}\circ K^{-1}\circ K^{-1}(K_{.,.,.}\circ K^{-1})-6K_{.,.,.}(*K^{-1})^{3}*K_{.,.,.}\\
&\qquad+18K_{.,.,.}\circ K^{-1}\circ K^{-1}(K_{.,.,.}(\circ K^{-1}\boldsymbol{y})^{2})\\
&\qquad+9(K_{.,.,.}\circ K^{-1}\circ K^{-1}\boldsymbol{y})^{2}\\
&\qquad+18K_{.,.,.}(*K^{-1})^{2}*K^{-1}\boldsymbol{y}\boldsymbol{y}'K^{-1}*K_{.,.,.}\\
&\qquad+(K_{.,.,.}(\circ K^{-1}\boldsymbol{y})^{3})^{2}-9K_{.,.,.}(\circ K^{-1}\boldsymbol{y})^{2}\circ K^{-1}(K_{.,.,.}(\circ K^{-1}\boldsymbol{y})^{2})\\
&\qquad-6K_{.,.,.}\circ K^{-1}\circ K^{-1}\boldsymbol{y}\cdot K_{.,.,.}(\circ K^{-1}\boldsymbol{y})^{3}\}
\end{aligned}
$$

$$D_{..}=(d_{bc})\ , \qquad d_{bc}=\delta'(y_{bc}-\kappa_{bc})/\delta(y_{bc}-\kappa_{bc})$$

$$D_{...}=(d_{abc})\ , \qquad d_{abc}=\delta'(y_{abc}-\kappa_{abc}/\sqrt{n})/\delta(y_{abc}-\kappa_{abc}/\sqrt{n})$$

where δ's are Dirac delta functions satisfing

$$\delta(x-a)=0\ , \qquad x\neq a\ ,$$

$$\int\delta(x-a)dx=1\ ,$$

$$\int h(\cdots, x,\cdots)\delta(x-a)dx=h(\cdots, a,\cdots)\ ,$$

$$\int h(\cdots, x,\cdots)\delta^{(l)}(x-a)dx=(-1)^{l}\frac{\partial^{l}h(\cdots, x,\cdots)}{\partial x^{l}}\bigg|_{x=a}\ .$$

As this expression is represented at the any point $\boldsymbol{\theta}$, the MGF $M_1(t)$ should be evaluated at the value H_0: $\boldsymbol{\theta}_2=\boldsymbol{\theta}_{20}$, that is,

$$
\begin{aligned}
M_1(t)&=\mathrm{E}\,[\exp(tS)\,|\,H_0]\\
&=\int\exp(tS)f_1(\boldsymbol{y}, Y, Y_{...}, Y_{....}\,|\,\boldsymbol{\theta}_1, \boldsymbol{\theta}_{20})d\boldsymbol{y}\,dY\,dY_{...}dY_{....}+o(1/n)\ .
\end{aligned}
$$

By the use of some relations with respect to K, $K_{...}$, $K_{....}$, etc., which are derived from the regularity conditions for the density function,

$$
\begin{aligned}
&K_{....}\circ K^{-1}\circ K^{-1}+4K_{.,...}\circ K^{-1}\circ K^{-1}+K_{..,..}\circ K^{-1}\circ K^{-1}\\
&\qquad+2K_{..,..}\circledast K^{-1}\circledast K^{-1}+2K_{.,.,..}\circ K^{-1}\circ K^{-1}\\
&\qquad+4K_{.,.,..}\circledast K^{-1}\circledast K^{-1}+K_{.,.,.,.}\circ K^{-1}\circ K^{-1}=0
\end{aligned}
$$

$$
\begin{aligned}
&K_{...}\circ K^{-1}\circ K^{-1}(K_{...}\circ K^{-1})+K_{...}\circ K^{-1}\circ K^{-1}(K_{.,.,.}\circ K^{-1})\\
&\qquad+K_{...}\circ K^{-1}\circ K^{-1}(K_{.,..}\circ K^{-1})\\
&\qquad+2K_{...}\circ K^{-1}\circ K^{-1}(K_{..,.}\circ K^{-1})=0
\end{aligned}
$$

$$K_{...}(*K^{-1})^3*K_{...}+K_{...}(*K^{-1})^3*K_{.,.,.}+3K_{...}(*K^{-1})^3*K_{.,..}=0$$

$$\begin{aligned}&K_{.,..}\circ K^{-1}\circ K^{-1}(K_{...}\circ K^{-1})+K_{.,..}\circ K^{-1}\circ K^{-1}(K_{.,.,.}\circ K^{-1})\\&\quad+K_{.,..}\circ K^{-1}\circ K^{-1}(K_{.,..}\circ K^{-1})\\&\quad+2K_{.,..}\circ K^{-1}\circ K^{-1}(K_{..,.}\circ K^{-1})=0\end{aligned}$$

$$\begin{aligned}&K_{.,..}(*K^{-1})^3*K_{...}+K_{.,..}(*K^{-1})^3*K_{.,.,.}+K_{.,..}(*K^{-1})^3*K_{..,.}\\&\quad+K_{.,..}(*K^{-1})^3*K_{.,..}+K_{.,..}(*K^{-1})^3*K_{..,.}=0\end{aligned}$$

we find the MGF of $-2\log\lambda$ as given by (5), to

$$(7)\qquad M_1(t)=(1-2t)^{-(p-q)/2}\left[1+\frac{1}{24n}\{A_2d^2+A_1d\}+o(1/n)\right],$$

where

$$d=2t/(1-2t)\ ,$$

$$\begin{aligned}A_2=12\{&K_{.,..}*M*M*A*K_{.,.,.}+K_{.,..}*A*M*M*K_{...}\\&+K_{.,..}*M*M*A*K_{.,..}+2K_{.,..}*A*M*M*K_{..,.}\}\end{aligned}$$

$$A_1=l(K^{-1})-l(A)\ ,$$

$$\begin{aligned}l(K^{-1})=\;&3K_{....}(\circ K^{-1})^2+12K_{..,..}(\circledast K^{-1})^2+12K_{.,.,..}(\circledast K^{-1})^2\\&+12K_{.,...}(\circ K^{-1})^2+3K_{...}\circ K^{-1}\circ K^{-1}(K_{...}\circ K^{-1})\\&+12K_{.,..}\circ K^{-1}\circ K^{-1}(K_{...}\circ K^{-1})\\&+12K_{.,..}\circ K^{-1}\circ K^{-1}(K_{..,.}\circ K^{-1})+6K_{...}(*K^{-1})^3*K_{...}\\&+4K_{.,.,.}(*K^{-1})^3*K_{...}+24K_{.,..}(*K^{-1})^3*K_{...}\\&+12K_{.,..}(*K^{-1})^3*K_{..,.}\end{aligned}$$

$$M=K^{-1}-A\ ,\qquad A=\begin{bmatrix}K_{11}^{-1} & O\\ O & O\end{bmatrix}\begin{matrix}q\\ p-q\end{matrix}\qquad K=\begin{bmatrix}K_{11} & K_{12}\\ K_{21} & K_{22}\end{bmatrix}\begin{matrix}q\\ p-q\end{matrix}\ .$$

The inversion of (7) gives the following theorem.

THEOREM 1. *The asymptotic expansion of the distribution of the likelihood ratio criterion for a composite hypothesis under the null hypothesis is as follows*:

$$\begin{aligned}(8)\qquad &\mathrm{P}\{-2\log\lambda\leq x\,|\,H_0\}\\&=P_f+\frac{1}{24n}\{A_2P_{f+4}-(2A_2-A_1)P_{f+2}+(A_2-A_1)P_f\}+o(1/n)\end{aligned}$$

where $P_f=\mathrm{P}\{\chi_f^2\leq x\}$, χ_f^2 *is a central chi-square random variable with* f *degrees of freedom and* $f=p-q$, *and* A_1 *and* A_2 *are given in* (7), *respectively.*

Remark 1. In case of a simple hypothesis $H_0: \boldsymbol{\theta}=\boldsymbol{\theta}_0$, the vanishing of A implies $A_2=0$ in (7), and (8) becomes

$$\text{(9)} \qquad \mathrm{P}\{-2\log\lambda \leqq x \mid H_0\} = P_f + \frac{l(K^{-1})}{24n}\{P_{f+2}-P_f\} + o(1/n)$$

and $f=p$.

Differentiating $M_1(t)$ under H_0 and setting $t=0$, we have the asymptotic expectation of $-2\log\lambda$ as follows:

$$\text{(10)} \qquad \mathrm{E}[-2\log\lambda \mid H_0] = p + \frac{l(K^{-1})}{12n} + o(1/n) .$$

Combining (9) and (10), we are able to give a correction factor ρ which makes the term of order $1/n$ of these expressions vanish simultaneously:

$$\text{(11)} \qquad \rho = 1 - \frac{l(K^{-1})}{12np} + o(1/n) .$$

Wilks [12] expected that the magnitude of the second term of the asymptotic expansion of the distribution would be of order $1/\sqrt{n}$. (9) shows that the convergence of $-2\log\lambda$ to a central chi-square is more rapid than it is expected. It is also of interest to note that we can find a correction factor ρ for a simple hypothesis.

Remark 2. It is of interest to note that the term A_2 in the expression (8) vanishes for a particular distribution. Let $\boldsymbol{x}$ be a random variable with Darmois-Koopman type probability density function expressed as

$$f(\boldsymbol{x} \mid \boldsymbol{\theta}) = h(\boldsymbol{x}) \exp\left\{\sum_{i=1}^{p} \theta_i u_i(\boldsymbol{x}) + V(\boldsymbol{\theta})\right\}$$

where $V(\boldsymbol{\theta})$ is differentiable with respect to $\boldsymbol{\theta}$ up to the fourth order. We assume that $f(\boldsymbol{x} \mid \boldsymbol{\theta})$ also satisfies the regularity conditions. As the second derivatives of $\log f(\boldsymbol{x} \mid \boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ become a function of $\boldsymbol{\theta}$, not depending on a random variable $\boldsymbol{x}$, that is,

$$\frac{\partial^2 \log f}{\partial\theta_i \partial\theta_j} = \frac{\partial^2 V}{\partial\theta_i \partial\theta_j} , \qquad i, j = 1, 2, \cdots, p ,$$

we have by the regularity condition

$$\kappa_{i,jk} = \mathrm{E}\left[\frac{\partial^2 \log f}{\partial\theta_j \partial\theta_k} \frac{\partial \log f}{\partial\theta_i}\right] = \frac{\partial^2 V}{\partial\theta_j \partial\theta_k} \mathrm{E}\left[\frac{\partial \log f}{\partial\theta_i}\right] = 0 .$$

This implies that $K_{\cdot,\cdot\cdot}=0$, that is, $A_2=0$ in (8). Noting the relations

$$\kappa_{i,j,k} = -\kappa_{ijk}$$

$$\kappa_{ij,kl} = -\kappa_{i,j,kl}\ , \qquad \kappa_{i,jkl} = 0$$

$$\kappa_{ijkl} + \kappa_{i,j,k,l} = \kappa_{ij,kl} + \kappa_{ik,jl} + \kappa_{il,jk}\ ,$$

the coefficient A_1 in (7) can be simply represented as follows:

$$A_1 = l(K^{-1}) - l(A)\ ,$$

$$l(K^{-1}) = 3K_{....}(\circ K^{-1})^2 + 3K_{...} \circ K^{-1} \circ K^{-1}(K_{...} \circ K^{-1}) + 2K_{...}(* K^{-1})^3 * K_{...}\ .$$

Thus we have again a correction factor ρ as

$$\rho = 1 - \frac{A_1}{12np} + o(1/n)\ .$$

Recently the asymptotic distribution theory of the LRC's for various hypotheses about the parameters of the multivariate normal population has been studied by the method developed for the normal distribution theory (Box [1], Sugiura and Fujikoshi [10], Sugiura [11], Fujikoshi [3], etc.). Lee, Krishnaiah and Chang [6] have given the asymptotic expansion of the distribution of a statistic, which has the moments expressed by Gamma functions, up to the order $1/n^{14}$, where n is a sample size. For these expressions it is of some interest that we are able always to obtain a correction factor ρ which makes a rapid convergence of LRC to a limiting distribution.

We will show in Example 1 how to get the asymptotic expansion of the distribution of LRC for a particular hypothesis of a multivariate normal population by the use of Theorem 1. However, the derivation is laborious.

Example 1 (The calculation of this example is partly due to Miss Y. Kikuchi). Let $\boldsymbol{x}$ be an m dimensional normal random vector with mean $\boldsymbol{\mu}$ and covariance matrix Σ, and $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N$ a random sample from this population. We test the simple hypothesis

$$H_0:\ \Sigma^{-1} = \Theta = \Theta_0 \text{ (specified), given } \boldsymbol{\mu} = \mathbf{0}$$

against

$$H:\ \Sigma^{-1} = \Theta \neq \Theta_0, \text{ given } \boldsymbol{\mu} = \mathbf{0}\ .$$

The probability density function of this normal population belongs to the family of Darmois and Koopman and the likelihood ratio criterion for this H_0 is given by

$$\lambda=\left(\frac{e}{N}\right)^{Nm/2}|\Theta_0 S|^{N/2}\operatorname{etr}\left(-\frac{1}{2}\Theta_0 S\right)$$

where $S=\sum_{\alpha=1}^{N} \boldsymbol{x}_\alpha \boldsymbol{x}'_\alpha$. Without loss of generality we are able to assume $\Theta_0=I_m$. Rearranging parameters as $\boldsymbol{\theta}=(\theta_{11}, \theta_{22}, \cdots, \theta_{mm}, \theta_{12}, \cdots, \theta_{m-1,m})$, we have the expectations of first and second derivatives of log-likelihood function,

$$\mathrm{E}\left[\frac{\partial^2 L}{\partial\theta_{ii}\partial\theta_{jj}}\right]=\frac{1}{2}(\theta^{ij})^2, \qquad \mathrm{E}\left[\frac{\partial L}{\partial\theta_{ii}}\frac{\partial L}{\partial\theta_{pq}}\right]=\theta^{ip}\theta^{iq}.$$

$$\mathrm{E}\left[\frac{\partial^2 L}{\partial\theta_{pq}\partial\theta_{rs}}\right]=\theta^{pr}\theta^{qs}+\theta^{ps}\theta^{qr}, \text{ respectively}.$$

Making third and fourth derivatives of L with respect to θ_{ij}'s, we have the following identities with a little algebra

$$K_{....}\circ K^{-1}\circ K^{-1}=-(2m^3+5m^2+5m)$$

$$K_{...}\circ K^{-1}\circ K^{-1}(K_{...}\circ K^{-1})=2m^3+4m^2+2m$$

$$K_{...} * K^{-1} * K^{-1} * K^{-1} * K_{...}=m^3+3m^2+4m\ ,$$

which gives

$$A_1=2m^3+3m^2-m\ .$$

Thus we have the following asymptotic expression of the distribution of λ under H_0;

$$\mathrm{P}\{-2\log\lambda\leq x\}=P_f+\frac{1}{24N}m(2m^2+3m-1)\{P_{f+2}-P_f\}+o(1/n)\ ,$$

where $P_f=\mathrm{P}\{\chi_f^2\leq x\}$, and $f=m(m+1)/2$. This result agrees with (2.10) in Sugiura [11] up to order $1/N$. Sugiura considered testing a hypothesis $H_0:\ \Sigma=\Sigma_0$, against $\Sigma\neq\Sigma_0$ by the use of a modified likelihood ratio criterion having an unbiasedness,

$$\lambda^*=\left(\frac{e}{n}\right)^{mn/2}|\Theta_0 A|^{n/2}\operatorname{etr}\left(-\frac{1}{2}\Theta_0 A\right),$$

where $n=N-1$, $\Theta=\Sigma^{-1}$, $A=\sum_{\alpha=1}^{N}(\boldsymbol{x}_\alpha-\bar{\boldsymbol{x}})(\boldsymbol{x}_\alpha-\bar{\boldsymbol{x}})'$.

3. Expansion of λ under a general composite hypothesis

In the discussion of Section 2 the attention has been restricted to the test of null hypothesis in which certain components of parameters

$\boldsymbol{\theta}$ have been specified. We generalize this result to a test of hypothesis $H_{\omega 0}$: $\boldsymbol{\theta} \in \omega$, where ω is the subspace of all vectors $\boldsymbol{\theta} \in \Theta$ for which there is a transformation $\boldsymbol{\xi}(\boldsymbol{\theta})$ such that

$$\boldsymbol{\xi}_2=(\xi_{q+1}(\boldsymbol{\theta}),\cdots,\xi_p(\boldsymbol{\theta}))=(\xi_{q+1,0},\cdots,\xi_{p,0})=\boldsymbol{\xi}_{20} ,$$

where $\xi_{q+1,0},\cdots,\xi_{p,0}$ are constants. The transformation $\boldsymbol{\xi}(\boldsymbol{\theta})$ is required to satisfy the following properties:

(a) There exists a vector $\boldsymbol{\xi}_1(\boldsymbol{\theta})=(\xi_1(\boldsymbol{\theta}),\cdots,\xi_q(\boldsymbol{\theta}))'$ such that the inverse transformation $\boldsymbol{\theta}(\boldsymbol{\xi})=(\theta_1(\boldsymbol{\xi}),\cdots,\theta_p(\boldsymbol{\xi}))'$ exists, where $\boldsymbol{\xi}'=(\boldsymbol{\xi}_1',\boldsymbol{\xi}_2')$.

(b) The partial derivatives of $\boldsymbol{\theta}(\boldsymbol{\xi})$ with respect to $\boldsymbol{\xi}$ exist up to order four, and are bounded and continuous function of $\boldsymbol{\xi}$.

(c) The transformation should be one to one, that is, the Jacobian $|\partial(\boldsymbol{\xi})/\partial(\boldsymbol{\theta})|\neq 0$.

Let $\hat{\boldsymbol{\theta}}$ be a maximum likelihood estimator of $\boldsymbol{\theta}\in\Theta$, then by the one to one correspondence between $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$, $\boldsymbol{\xi}(\hat{\boldsymbol{\theta}})=\hat{\boldsymbol{\xi}}$ becomes also a maximum likelihood estimator of $\boldsymbol{\xi}$. By this reason the likelihood ratio criterion of testing the general composite hypothesis

$$H_{\omega 0}: \boldsymbol{\xi}_2=\boldsymbol{\xi}_{20} \quad \text{against} \quad H_{\omega}: \boldsymbol{\xi}_2\neq\boldsymbol{\xi}_{20}$$

is expressed as

$$\lambda=\prod_{\alpha=1}^{n}\frac{f(\boldsymbol{x}\,|\,\tilde{\boldsymbol{\xi}}_1,\boldsymbol{\xi}_{20})}{f(\boldsymbol{x}\,|\,\hat{\boldsymbol{\xi}}_1,\hat{\boldsymbol{\xi}}_2)} , \tag{12}$$

where $\tilde{\boldsymbol{\xi}}_1'=(\tilde{\xi}_1,\cdots,\tilde{\xi}_q)$ is a m.l.e. of $\boldsymbol{\xi}_1$ under $H_{\omega 0}$. The expression (12) is completely similar to (1) except the notation $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$. This implies that the distribution theory with respect to λ under the general composite hypothesis can be handled by similar way as the case of a composite hypothesis.

THEOREM 2. *Under the required conditions for the density function and $\boldsymbol{\xi}(\boldsymbol{\theta})$, the asymptotic expansion of the distribution of the likelihood ratio criterion λ under $H_{\omega 0}$ can be expressed as follows:*

$$\begin{aligned} (13)\qquad & \mathrm{P}\{-2\log\lambda\leqq x\,|\,H_{\omega 0}\} \\ &=P_f+\frac{1}{24n}\{A_{2\omega}P_{f+4}-(2A_{2\omega}-A_{1\omega})P_{f+2}+(A_{2\omega}-A_{1\omega})P_f\}+o(1/n) \end{aligned}$$

where $A_{2\omega}$ and $A_{1\omega}$ are the values of A_2 and A_1 in (7) at $\boldsymbol{\xi}=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_{20})$ and $f=p-q$.

To express (13) in terms of $\boldsymbol{\theta}$, we have to know the values of $A_{2\omega}$ and $A_{1\omega}$ at ω. By the regularity conditions required for transformation,

$$\kappa_{ij}(\boldsymbol{\xi})=\kappa_{pq}\tau_{ip}\tau_{jq} , \qquad \kappa_{i,j}(\boldsymbol{\xi})=\kappa_{i,j}\tau_{ip}\tau_{jq}$$

(14)
$$\kappa_{ijk}(\boldsymbol{\xi})=\kappa_{pqr}\tau_{ip}\tau_{jq}\tau_{kr}+\kappa_{pq}\tau_{ij,pp}\tau_{jq}+\kappa_{pq}\tau_{jk,qq}\tau_{ip}+\kappa_{pq}\tau_{ik,pp}\tau_{jq}$$
$$\kappa_{i,jk}(\boldsymbol{\xi})=\kappa_{p,qr}\tau_{ip}\tau_{jq}\tau_{kr}+\kappa_{p,q}\tau_{ip}\tau_{jk,qq}$$
$$\kappa_{i,j,k}(\boldsymbol{\xi})=\kappa_{p,q,r}\tau_{ip}\tau_{jq}\tau_{kr} \qquad \text{etc.},$$

where

$$\tau_{ip}=\frac{\partial\theta_p}{\partial\xi_i}, \qquad \tau_{ij,pp}=\frac{\partial^2\theta_p}{\partial\xi_i\partial\xi_j}.$$

The functions of the fourth order derivatives can be obtained as a similar way.

If the transformation for $\boldsymbol{\theta}$ to $\boldsymbol{\xi}$ is an Affine transformation, i.e., $\boldsymbol{\theta}=G\boldsymbol{\xi}+\boldsymbol{\eta}$, where $G=(g_{ij})$, $i, j=1, 2,\cdots, p$, is a non-singular constant matrix, the expressions given in (14) become more simpler ones:

$$\kappa_{ij}(\boldsymbol{\xi})=\kappa_{pq}(\boldsymbol{\theta})g_{pi}g_{qj}, \qquad \kappa_{i,j}(\boldsymbol{\xi})=\kappa_{p,q}(\boldsymbol{\theta})g_{pi}g_{qj}$$
$$\kappa_{ijk}(\boldsymbol{\xi})=\kappa_{pqr}(\boldsymbol{\theta})g_{pi}g_{qj}g_{rk}, \qquad \kappa_{i,jk}(\boldsymbol{\xi})=\kappa_{p,qr}(\boldsymbol{\theta})g_{pi}g_{qj}g_{rk}$$
$$\kappa_{i,j,k}(\boldsymbol{\xi})=\kappa_{p,q,r}(\boldsymbol{\theta})g_{pi}g_{qj}g_{rk}, \qquad \text{etc.}$$

It is easy to see that $l(K^{-1})$ of (7) is invariant under this Affine transformation, but $l(A)$ is not invariant.

We discuss two examples which cover the testing hypothesis of a multivariate normal population as the particular cases.

Example 2.1. We consider the following composite hypothesis

$$H_0:\ \boldsymbol{\theta}_1=\theta\boldsymbol{e}, \quad \boldsymbol{\theta}_2=\boldsymbol{\theta}_{20}, \qquad \text{against } H: H_0 \text{ is not true},$$

where $\boldsymbol{\theta}_1=(\theta_1,\cdots,\theta_q)$, $\boldsymbol{\theta}_2=(\theta_{q+1},\cdots,\theta_p)$, $\boldsymbol{e}=(1,\cdots,1)'_{q\times 1}$ $\boldsymbol{\theta}_{20}=(\theta_{q+1,0},\cdots,\theta_{p,0})$ specified vector. By considering a following Affine transformation such that

$$\begin{bmatrix}\boldsymbol{\xi}_1\\ \boldsymbol{\xi}_2\end{bmatrix}=G^{-1}\begin{bmatrix}\boldsymbol{\theta}_1\\ \boldsymbol{\theta}_2\end{bmatrix}-\begin{bmatrix}0\\ \boldsymbol{\theta}_{20}\end{bmatrix}$$

where

$$G^{-1}=\left[\begin{array}{cccccc:cccc}
1 & & & & & & & & & \\
-1 & 1 & & & & & & & & \\
-1 & & \cdot & & & & & & & \\
\cdot & & & \cdot & & & & & & \\
\cdot & & & & \cdot & & & & & \\
-1 & \cdot & \cdot & \cdot & \cdot & 1 & & & & \\
\hdashline
 & & & & & & 1 & & & \\
 & & & & & & & 1 & & \\
 & & & & & & & & \cdot & \\
 & & & & & & & & & 1
\end{array}\right]\begin{matrix} \\ \\ q \\ \\ \\ \\ \\ p-q \\ \\ \\ \end{matrix}$$

the hypothesis H_0 is reduced to a simple form,

$$H_{\omega 0}:\ \xi_2=\xi_3=\cdots=\xi_p=0\ , \qquad \text{against } H_\omega : H_{\omega 0} \text{ is not true}$$

and the likelihood ratio criterion becomes

$$\lambda=\prod_{\alpha=1}^{n}\frac{f(\boldsymbol{x}_\alpha\,|\,\tilde{\xi}_1, 0,\cdots, 0)}{f(\boldsymbol{x}_\alpha\,|\,\hat{\boldsymbol{\xi}}_1, \hat{\boldsymbol{\xi}}_2)}\ , \tag{15}$$

where $\tilde{\xi}_1$ is a maximum likelihood estimator of ξ_1 under $H_{\omega 0}$. It is easy to check that the first term of $-2\log\lambda$ becomes

$$2\log\lambda=\boldsymbol{u}'V\boldsymbol{u}+O_p(1/n)\ ,$$

where

$$\boldsymbol{u}=\frac{1}{\sqrt{n}}\frac{\partial L}{\partial\boldsymbol{\xi}}=G\frac{1}{\sqrt{n}}\frac{\partial L}{\partial\boldsymbol{\theta}}=G\boldsymbol{y}\ ,$$

$$V=U^{-1}-V_0=U^{-1}-\begin{bmatrix} a & \mathbf{0}' \\ \mathbf{0} & O\end{bmatrix}, \qquad a=u_{11}^{-1}=\left(\frac{1}{n}\frac{\partial^2 L}{\partial\xi_1^2}\right)^{-1}$$

$$U=\frac{1}{n}\frac{\partial^2 L}{\partial\boldsymbol{\xi}\partial\boldsymbol{\xi}'}=GYG'\ .$$

Noting $u_{11}=(\boldsymbol{e}', \mathbf{0}')Y(\boldsymbol{e}', \mathbf{0}')'=\boldsymbol{e}'Y_{11}\boldsymbol{e}$, we have a quadratic form with respect to $\boldsymbol{y}$ and Y as

$$\boldsymbol{u}'V\boldsymbol{u}=\boldsymbol{y}'\left\{Y^{-1}-\frac{1}{\boldsymbol{e}'Y_{11}\boldsymbol{e}}\begin{bmatrix}\boldsymbol{e}\boldsymbol{e}' & O\\ O & O\end{bmatrix}\right\}\boldsymbol{y}=\boldsymbol{y}'Z\boldsymbol{y}\ .$$

It is seen that

$$ZYZ=Z \quad\text{and}\quad \operatorname{rank} Z=\operatorname{rank} ZY=\operatorname{tr} ZY=p-1\ ,$$

which implies that $\boldsymbol{u}'V\boldsymbol{u}$ approaches to a central chi-square random variable with $p-1$ degrees of freedom in law as n tends to infinity.

This type of hypothesis can be found in testing hypothesis concerning with a multivariate normal population.

(a) Let $\boldsymbol{x}$ be an m dimensional multivariate normal random vector with mean zero and covariance matrix Σ. Testing the sphericity H_0: $\Sigma=\sigma^2 I_m$, against H: $\Sigma\neq\sigma^2 I_m$, σ^2 unspecified, is included in this case.

(b) Let $\boldsymbol{x}$ be distributed as above (a). Testing the hypothesis of an intraclass correlation model H_0: $\sigma_{ij}=\rho$, $i\neq j$ and $\sigma_{ii}=1$, $i, j=1,\cdots, m$, against $H: H_0$ is not true, is included in this case.

Example 2.2. For testing a hypothesis

$$H_0:\ \boldsymbol{\theta}_2=\theta\boldsymbol{e}\ , \qquad \text{against } H:\ \boldsymbol{\theta}_2\neq\theta\boldsymbol{e},\ \boldsymbol{e}=(1, 1,\cdots, 1)'_{(p-q)\times 1}$$

the first term of $2\log\lambda$ can be expressed by doing a similar way as

$$u'Vu=y'\left\{Y^{-1}-\begin{bmatrix} I_q & 0 \\ O & e \end{bmatrix} Y_{11}^{-1}(\xi)\begin{bmatrix} I_q & O \\ 0' & e' \end{bmatrix}\right\}y=y'Zy\ ,$$

where

$$Y_{11}(\xi)=\begin{bmatrix} I_q & O \\ 0' & e' \end{bmatrix} Y \begin{bmatrix} I_q & 0 \\ O & e \end{bmatrix}\ ,$$

and

$$ZYZ=Z\ , \qquad \text{rank}\, Z=\text{tr}\, ZY=p-q-1\ .$$

This implies that $-2\log\lambda$ approaches to a central chi-square with $p-q-1$ degrees of freedom. We are able to construct many hypothesis for parameters of multivariate normal population.

To close this section we show the asymptotic behavior of $-2\log\lambda$ under a sequence of Pitman's alternative $H\omega_n$ such that

$$H\omega_n:\ \xi_2=\xi_{20}+\varepsilon/\sqrt{n}$$

where n is a sample size. The asymptotic expansion of the distribution of LRC under H_n: $\theta=\theta_{20}+\varepsilon/\sqrt{n}$ has been obtained by Theorem 1 in Hayakawa [4]. By using a representation (12) of LRC and Theorem 1 of Hayakawa, the asymptotic expansion of the distribution of $-2\log\lambda$ under H_{ω_n} is given as follows:

$$\text{(16)} \qquad \mathrm{P}\{-2\log\lambda\leq x\}=P_f(\delta^2)+\frac{1}{\sqrt{n}}\sum_{k=0}^{3} a_k P_{f+2k}(\delta^2)+O(1/n)$$

where

$$f=p-q\ , \qquad \delta^2=\frac{1}{2}\varepsilon' K_{22\cdot 1}\varepsilon\ ,$$

$$a_2=-\frac{1}{6}K_{.,.,.}(\circ\,\varepsilon^*)^3$$

$$a_1=-\frac{1}{6}\{K_{...}(\circ\,\varepsilon^*)^3-2K_{.,.,.}(\circ\,\varepsilon^*)^3+3K_{...}\circ A\circ\varepsilon^* +6K_{.,..}\circ A\circ\varepsilon^*+3K_{2..}\circ\varepsilon\circ\varepsilon^*\circ\varepsilon^*+3K_{2,..}\circ\varepsilon\circ\varepsilon^*\circ\varepsilon^*\}$$

$$a_0=-\frac{1}{6}\{K_{.,.,.}(\circ\,\varepsilon^*)^3-K_{...}(\circ\,\varepsilon^*)^3-3K_{...}\circ A\circ\varepsilon^* -6K_{.,..}\circ A\circ\varepsilon^*-3K_{2..}\circ\varepsilon\circ\varepsilon^*\circ\varepsilon^*-3K_{2,..}\circ\varepsilon\circ\varepsilon^*\circ\varepsilon^*\}$$

$$\varepsilon^*=\begin{bmatrix} K_{11}^{-1}K_{12} \\ -I_{p-q} \end{bmatrix}\varepsilon\ .$$

$K_{.,.,.}$, $K_{.,..}$, $K_{...}$, A and $\boldsymbol{\varepsilon}^*$ are the values of these at $(\boldsymbol{\xi}_1, \boldsymbol{\xi}_{20})$.

4. Expansion of λ under a fixed alternative

In this section we shall only consider the asymptotic distribution of the normalized likelihood ratio criterion under a simple fixed alternative H_ε: $\boldsymbol{\theta}=\boldsymbol{\theta}_0+\boldsymbol{\varepsilon}$, $\boldsymbol{\varepsilon}$ is a fixed vector.

Expanding $L(\boldsymbol{\theta}_0)$ at $\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}$ and noting $\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)=\boldsymbol{v}+\sqrt{n}\,\boldsymbol{\varepsilon}$,

$$
\begin{aligned}
(17)\qquad & 2\{L(\boldsymbol{\theta}_0)-L(\hat{\boldsymbol{\theta}})\} \\
&=(\boldsymbol{v}+\sqrt{n}\,\boldsymbol{\varepsilon})'\hat{Y}(\boldsymbol{v}+\sqrt{n}\,\boldsymbol{\varepsilon})-\frac{1}{3}\hat{Y}_{...}(\circ(\boldsymbol{v}+\sqrt{n}\,\boldsymbol{\varepsilon}))^3 \\
&\quad+\frac{1}{12}\hat{Y}_{....}(\circ(\boldsymbol{v}+\sqrt{n}\,\boldsymbol{\varepsilon}))^4-\cdots .
\end{aligned}
$$

Again expanding $\hat{y}_{ij}, \hat{y}_{ijk}, \cdots$ at $\boldsymbol{\theta}$ and rearranging (17), we have the following asymptotic expansion;

$$
\begin{aligned}
&2n\{K(\boldsymbol{\theta}_0)-K(\boldsymbol{\theta})\}+2\sqrt{n}\{u_0-u+\boldsymbol{\varepsilon}'\boldsymbol{y}-\boldsymbol{\varepsilon}'K_{..}\boldsymbol{v}+\boldsymbol{\varepsilon}'W_{..}\boldsymbol{v}\} \\
&\quad+\boldsymbol{v}'K_{..}\boldsymbol{v}+K_{...}\circ\boldsymbol{v}\circ\boldsymbol{v}\circ\boldsymbol{\varepsilon}+\boldsymbol{v}'W_{..}\boldsymbol{v}+W_{...}\circ\boldsymbol{v}\circ\boldsymbol{v}\circ\boldsymbol{\varepsilon}+o_p(1) ,
\end{aligned}
$$

where

$$K(\boldsymbol{\theta})=\mathrm{E}\,[\log f(\boldsymbol{x}\,|\,\boldsymbol{\theta})]=\int\{\log f(\boldsymbol{x}\,|\,\boldsymbol{\theta})\}f(\boldsymbol{x}\,|\,\boldsymbol{\theta})d\boldsymbol{x} ,$$

$$K(\boldsymbol{\theta}_0)=\mathrm{E}\,[\log f(\boldsymbol{x}\,|\,\boldsymbol{\theta}_0)]$$

$$u_0=\frac{1}{\sqrt{n}}\{L(\boldsymbol{\theta}_0)-nK(\boldsymbol{\theta}_0)\} ,\qquad u=\frac{1}{\sqrt{n}}\{L(\boldsymbol{\theta})-nK(\boldsymbol{\theta})\} ,$$

$$W_{..}=(w_{ij})=Y-K_{..} ,\qquad w_{ij}=\frac{1}{n}\{L_{ij}(\boldsymbol{\theta})-n\kappa_{ij}(\boldsymbol{\theta})\} ,$$

$$W_{...}=(w_{ijk})=\sqrt{n}(Y_{...}-K_{...}) ,\qquad w_{ijk}=\frac{1}{n}\{L_{ijk}(\boldsymbol{\theta})-n\kappa_{ijk}(\boldsymbol{\theta})\} .$$

where

$$L_{ij}(\boldsymbol{\theta})=\frac{\partial^2 L}{\partial\theta_i\partial\theta_j}\quad\text{and}\quad L_{ijk}(\boldsymbol{\theta})=\frac{\partial^3 L}{\partial\theta_i\partial\theta_j\partial\theta_k} .$$

Replacing $\boldsymbol{v}$ by (3), we have the normalized asymptotic expansion of $-2\log\lambda$.

$$
\begin{aligned}
S_H&=[-2\log\lambda+2n\{K(\boldsymbol{\theta}_0)-K(\boldsymbol{\theta})\}]/\sqrt{n} \\
&=2(u-u_0)-\boldsymbol{y}'Y^{-1}\boldsymbol{y}/\sqrt{n}+o_p(1/\sqrt{n}) .
\end{aligned}
$$

To find the MGF of S_H we need the Edgeworth expansion of $(u_0, u, \boldsymbol{y}, Y)$, which is given as follows.

$$f_2=f_0\left[1+\frac{1}{\sqrt{n}}\left\{\frac{1}{6}\{H_{\cdot,\cdot,\cdot}(\circ M^{-1}\boldsymbol{z})^3-3H_{\cdot,\cdot,\cdot}\circ M^{-1}\circ M^{-1}\boldsymbol{z}\}\right.\right.$$
$$\left.\left.-H_{\cdot,\cdot\cdot}\circ M^{-1}\boldsymbol{z}\circ D_{\cdot\cdot}\right\}+o(1/n)\right],$$

where

$$f_0=(2\pi)^{-(p+2)/2}|M|^{-1/2}\exp\left\{-\frac{1}{2}\boldsymbol{z}'M^{-1}\boldsymbol{z}\right\}\prod_{i,j}\delta(y_{ij}-\kappa_{ij})$$

$$\boldsymbol{z}'=(u_0, u, \boldsymbol{y}')$$

$$M=\mathrm{E}(\boldsymbol{z}\boldsymbol{z}')=\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22}\end{bmatrix}=\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & K\end{bmatrix}\begin{matrix}2\\p\end{matrix}$$
$$\qquad\qquad\qquad\qquad\qquad\quad 2 \qquad p$$

$$H_{\cdot,\cdot,\cdot}=(h_{i,j,k}), \qquad h_{i,j,k}=\mathrm{E}(z_iz_jz_k)$$

$$H_{\cdot,\cdot\cdot}=(h_{i,jk}), \qquad h_{i,jk}=\mathrm{E}(z_iy_{jk})$$

$$D_{\cdot\cdot}=(d_{ij}), \qquad d_{ij}=\delta'(y_{ij}-\kappa_{ij})/\delta(y_{ij}-\kappa_{ij}).$$

We find after some lengthy algebra with respect to the normal distribution that the MGF of S_H is given by

$$M_2(t)=\exp\{2t^2\boldsymbol{e}'M_{11}\boldsymbol{e}\}\left[1+\frac{q_1}{\sqrt{n}}+o(1/n)\right],$$

where

$$q_1=pt+t^3\left\{4\boldsymbol{e}'M_{12}M_{22}^{-1}M_{21}\boldsymbol{e}+\frac{4}{3}H_{1,1,1}(\circ\boldsymbol{e})^3\right\},$$

$$\boldsymbol{e}'=(-1, 1),$$

$$H_{1,1,1}(\circ\boldsymbol{e})^3=h_{0,0,1}+h_{0,1,0}+h_{1,0,0}+h_{1,1,1}-\{h_{0,0,0}+h_{0,1,1}+h_{1,0,1}+h_{1,1,0}\}.$$

Putting $\tau^2=4\boldsymbol{e}'M_{11}\boldsymbol{e}$, we have the following theorem.

THEOREM 3. *Under the fixed alternative, the normalized likelihood ratio criterion* $S_H=[-2\log\lambda+2n\{K(\boldsymbol{\theta}_0)-K(\boldsymbol{\theta})\}]/\sqrt{n}$ *has the following asymptotic expansion of the distribution.*

$$\mathrm{P}\{S_H/\tau\leq x\}=\Phi(x)-\frac{1}{\sqrt{n}}\left[p\Phi^{(1)}(x)/\tau+\left\{4\boldsymbol{e}'M_{12}M_{22}^{-1}M_{21}\boldsymbol{e}\right.\right.$$
$$\left.\left.+\frac{4}{3}H_{1,1,1}(\circ\boldsymbol{e})^3\right\}\Phi^{(3)}(x)/\tau^3\right]+o(1/n), \tag{18}$$

where $\tau^2=4e'M_{11}e$ *and* $\Phi^{(r)}(x)$ *means the r-th derivative of the standard normal distribution function* $\Phi(x)$.

Under the null hypothesis H_0: $\boldsymbol{\theta}=\boldsymbol{\theta}_0$, we have $u=u_0$. This means $\tau^2=0$, by which the asymptotic non-null distribution $-2\log\lambda$ has a singularity at the null hypothesis, so that the formula given by Theorem 3 does not give the good approximation near the null hypothesis.

Example 3. We shall examine the same hypothesis testing as Example 1. The parameter $\boldsymbol{\theta}$ is arranged as

$$\boldsymbol{\theta}'=(\sigma_{11},\cdots,\sigma_{mm},\sigma_{12},\cdots,\sigma_{m-1,m}) .$$

To find the covariance matrix M of $\boldsymbol{z}$, without loss of generality, we take the expectations of $\boldsymbol{z}$ at $\Sigma=I_m$. By doing this, $\boldsymbol{z}$ should be read as

$$u_0=-\frac{1}{2}\log|2A|-\frac{1}{2}\operatorname{tr}A\boldsymbol{x}\boldsymbol{x}' ,$$

$$u=-\frac{1}{2}\log|2I_m|-\frac{1}{2}\operatorname{tr}\boldsymbol{x}\boldsymbol{x}' ,$$

$$\boldsymbol{y}'=\frac{1}{\sqrt{n}}\left(\frac{\partial L}{\partial\sigma_{11}}\bigg|_{\Sigma=I},\cdots,\frac{\partial L}{\partial\sigma_{mm}}\bigg|_{\Sigma=I},\frac{\partial L}{\partial\sigma_{12}}\bigg|_{\Sigma=I},\cdots,\frac{\partial L}{\partial\sigma_{m-1,m}}\bigg|_{\Sigma=I}\right) ,$$

where

$$\boldsymbol{x}\sim N(0,I_m) \quad\text{and}\quad A=\Sigma_0^{-1/2}\Sigma\Sigma_0^{-1/2} .$$

Then,

$$M=\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22}\end{bmatrix} ,$$

$$M_{11}=\frac{1}{2}\begin{bmatrix} \operatorname{tr}A^2 & \operatorname{tr}A \\ \operatorname{tr}A & m\end{bmatrix} ,$$

$$M_{22}=\begin{bmatrix} \frac{1}{2}I_m & \\ & I_{m(m-1)/2}\end{bmatrix}$$

$$M_{12}=-\left[\begin{array}{c|c} \frac{1}{2}a_{11},\cdots,\frac{1}{2}a_{mm} & a_{12},\cdots,a_{m-1,m} \\ \hline \frac{1}{2},\cdots,\frac{1}{2} & 0,\cdots,0 \end{array}\right] .$$

We have

$$\tau^2=4e'M_{11}e=2\,\mathrm{tr}\,(A-I)^2\,,$$

$$4e'M_{12}M_{22}^{-1}M_{21}e=2\,\mathrm{tr}\,(A-I)^2\,,$$

$$H_{1,1,1}(\circ\, e)^3=h_{1,1,1}+3h_{1,0,0}-\{h_{0,0,0}+3h_{1,1,0}\}\,,$$

$$h_{0,0,0}=\mathrm{E}\,(u_0^3)=-\mathrm{tr}\,A^3\,,$$

$$h_{1,0,0}=h_{0,1,0}=h_{0,0,1}=\mathrm{E}\,(u_0^2u)=-\mathrm{tr}\,A^2\,,$$

$$h_{1,1,0}=h_{1,0,1}=h_{0,1,1}=\mathrm{E}\,(u_0u^2)=-\mathrm{tr}\,A\,,$$

$$h_{1,1,1}=\mathrm{E}\,(u^3)=-m\,.$$

This gives

$$H_{1,1,1}(\circ\, e)^3=\mathrm{tr}\,(A-I)^3\,.$$

Combining these results, we have the following expression.

$$\begin{aligned}\mathrm{P}\,\{[-2\log\lambda-n\{\mathrm{tr}\,(A-I)-\log|A|\}]/\sqrt{n}\,\tau\leqq x\}\\ =\Phi(x)-\frac{1}{6\sqrt{n}}\{3m(m+1)\Phi^{(1)}(x)/\tau+4(m+2\,\mathrm{tr}\,A^3\\ -3\,\mathrm{tr}\,A^2)\Phi^{(3)}(x)/\tau^3\}+o(1/\sqrt{n})\,.\end{aligned}$$

This agrees with (2.16) in Sugiura [11].

Acknowledgment

The author wishes to express his gratitude to the referee for his helpful comments in revising a paper.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria, *Biometrika*, **36**, 317-346.
[2] Davidson, R. R. and Lever, W. E. (1970). The limiting distribution of the likelihood ratio statistic under a class of local alternatives, *Sankhyā*, **32**, 209-224.
[3] Fujikoshi, Y. (1970). Asymptotic expansions of the distributions of test statistics in multivariate analysis, *Jour. Sci. Hiroshima Univ.* Ser. A-I Math., **34**, 73-144.
[4] Hayakawa, T. (1975). The likelihood ratio criterion for a composite hypothesis under a local alternative, *Biometrika*, **62**, 451-460.
[5] Lawley, D. N. (1956). A general method for approximating to the distribution of the likelihood ratio criteria, *Biometrika*, **43**, 295-303.
[6] Lee, J. C., Krishnaiah, P. R. and Chang, T. C. (1976). On the distribution of the likelihood ratio test statistic for compound symmetry, *S. Afric. Statist. J.*, **10**, 49-62.
[7] Mann, H. B. and Wald, A. (1943). On the stochastic limit and order relationships, *Ann. Math. Statist.*, **14**, 217-226.

[8] 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.

[9] Peers, H. W. (1971). Likelihood ratio and associate test criteria, *Biometrika*, **58**, 577-587.

[10] Sugiura, N. and Fujikoshi, Y. (1969). Asymptotic expansions of the non-null distributions of the likelihood ratio criteria for multivariate linear hypothesis and independence, *Ann. Math. Statist.*, **40**, 942-952.

[11] Sugiura, N. (1969). Asymptotic expansions of the distributions of the likelihood ratio criteria for covariance matrix, *Ann. Math. Statist.*, **40**, 2051-2063.

[12] Wilks, S. S. (1938). The large sample distribution of the likelihood ratio for testing composite hypothesis, *Ann. Math. Statist.*, **9**, 60-62.

Ann. Inst. Statist. Math.
39 (1987), Part A, 681

CORRECTION TO
"THE LIKELIHOOD RATIO CRITERION AND THE ASYMPTOTIC EXPANSION OF ITS DISTRIBUTION"

TAKESI HAYAKAWA

(This Annals Vol. 29, No. 3 (1977), pp. 359-378)

The expressions for the constants A_2 in (7) and $A_{2\omega}$ in (13) are incorrect. The correct expressions are $A_2=A_{2\omega}=0$.

This fact is also noted in Cordeiro [1] and Harris [2].

REFERENCES

[1] Cordeiro, G. M. (1987). On the corrections to the likelihood ratio statistics, *Biometrika*, **74**, 265-274.

[2] Harris, P. (1986). A note on Bartlett adjustments to likelihood ratio tests, *Biometrika*, **73**, 735-737.