ON $\chi^2$-STATISTICS WITH VARIABLE INTERVALS

BY

ANADI RANJAN ROY

TECHNICAL REPORT NO. 1

PREPARED UNDER CONTRACT Nonr-225(21)
(NR-042-993)
FOR
OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

JANUARY 30, 1956
ON $\chi^2$ -STATISTICS WITH VARIABLE INTERVALS

By

Anadi Ranjan Roy

TECHNICAL REPORT NO. 1

PREPARED UNDER CONTRACT Nonr-225(21)
(NR-042-993)
FOR
OFFICE OF NAVAL RESEARCH

REPRODUCTION IN WHOLE OR IN PART IS PERMITTED FOR
ANY PURPOSE OF THE UNITED STATES GOVERNMENT

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

JANUARY 30, 1956
ON $\chi^2$-STATISTICS WITH VARIABLE INTERVALS

by

Anadi Ranjan Roy

I. Introduction

The $\chi^2$ statistic is used as a test of goodness of fit in various situations. The usual situation met with is the one where it is required to find out if an observed sample could arise from a specified probability distribution. The structure of the $\chi^2$-test is as follows: The range of the random variable is divided into $k$ fixed mutually exclusive point sets $S_i$. If $x_1, \ldots, x_n$ ($n > k$) is the observed sample, the number $m_i$ of the $x$'s falling in the set $S_i$ is noted. If $p_i$ is the probability of a single observation falling in $S_i$ derived on the basis of the specified probability distribution (it is assumed that $p_i > 0$, $i=1,2,\ldots,k$), the sampling distribution of

$$ R_n = \sum_{i=1}^{k} \frac{(m_i-np_i)^2}{np_i} $$

(1)

tends, as $n \to \infty$, to the distribution defined by the frequency function

$$ \frac{1}{k-1} \frac{1}{\Gamma\left(\frac{k-1}{2}\right)} \frac{k-3}{2} x^{\frac{k-3}{2}} e^{-\frac{x}{2}} \left( x > 0 \right) $$

(2)

which is called the $\chi^2$ distribution with $k-1$ degrees of freedom [c.f. 1]. The above holds whenever $p_i$ is the true value of the probability of an observation falling in $S_i$.

Consider now the case where only the form of the parent distribution from
which the sample is supposed to arise is known but the values of the parameters, say \( \theta_1, \theta_2, \ldots, \theta_s \) entering in the form are not known. An estimate based on the number of observations falling in the different classes \( S_i \) will be called a limited information estimate. Let \((\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_k)\) be a best asymptotically normal limited information estimate of \((p_1, p_2, \ldots, p_k)\), such as the minimum \( \chi^2 \) estimate or the maximum likelihood estimate (based only on the limited information). Then the asymptotic sampling distribution of \( \tilde{R}_n \), where \( \tilde{R}_n \) is the same as \( R_n \) with \( p_i \) replaced by \( \tilde{p}_i \), is a \( \chi^2 \) with \( n-s-1 \) degrees of freedom [c.f. 1].

In the above formulation the sample observations have been grouped into different classes fixed arbitrarily, and in so doing the information on the values of the individual observations in the sample has been partially ignored. It is thus expected that if this information could be better utilized it would result in an improvement of the test of goodness of fit. This consideration led Lehmann and Chernoff [c.f. 3] to use the (full information) maximum likelihood estimates \( \hat{\theta}_i \) of the parameter \( \theta_i \) in the population form and replace \( p_i \) in \( R_n \) by \( \hat{p}_i \) calculated by using \( \hat{\theta}_i \)'s in place of \( \theta_i \)'s. They have the following result: The asymptotic distribution of

\[
\tilde{R}_n = \sum_{i=1}^{k} \frac{(m_i - n\hat{p}_i)^2}{n\hat{p}_i}
\]

is that of

\[
\sum_{i=1}^{k-s-1} y_i^2 + \sum_{i=k-s}^{k-1} \lambda_i y_i^2
\]

where the \( y_i \)'s are independently normally distributed with mean zero and
unit variance, and the \( \lambda_i \)'s are between 0 and 1, possibly depending on the \( s \) parameters, \( \theta_1, \ldots, \theta_s \). One disturbing feature in the distribution (4) of \( \hat{R}_n \) is that the \( \lambda \)'s may depend on the unknown parameters \( \theta_1, \theta_2, \ldots, \theta_s \) so that the distribution is not known to that extent.

II. Summary

It seems natural to utilize the information in the individual sample observations in defining the class intervals in addition to obtaining suitable estimates of \( p_i \), as done by Lehmann and Chernoff. In what follows the condition that the class intervals are fixed has been relaxed. The classes \( S_1, S_2, \ldots, S_k \) have been defined in terms of some suitably chosen estimate \( \hat{\theta} \) of \( \theta \) based on the random sample \( x_1, x_2, \ldots, x_n \), i.e., \( S_1 \) is a function of \( \hat{\theta} \) for \( i=1,2,\ldots,k \). Let \( m_i \) be the number of sample observations in \( S_1(\hat{\theta}) \). Let

\[
p_i^*(\hat{\theta}) = \Pr \{ x \in S_1(\hat{\theta}) | \theta = \hat{\theta} \} = \int_{S_1(\hat{\theta})} p(x|\theta = \hat{\theta}) \, dx
\]

where \( x \) is independent of the observations in the original sample and has the same density function \( p(x|\theta) \) as that from which the sample arises. In other words \( p_i^*(\hat{\theta}) \) is the probability of \( x \) lying in \( S_1(\hat{\theta}) \) when \( \theta \) takes the value \( \hat{\theta} \), i.e., to say \( \hat{\theta} \) is taken as a constant for the computation of \( p_i^*(\hat{\theta}) \). We construct the statistic

\[
R_n^* = \sum_{i=1}^{k} \frac{(m_i - np_i^*(\hat{\theta}))^2}{np_i^*(\hat{\theta})}
\]
It is shown that by defining the classes $S_1$ in terms of suitable functions of the maximum likelihood estimate of $\theta$ it is always possible to obtain the asymptotic distribution of $R_n^*$ in the form (4) and in case the parameters involved are either of location or scale type, the $\lambda$'s in the form of the distribution are independent of the parameters and also lie between 0 and 1. It is obvious that if the parameters are scale ($\theta_1$) and location ($\theta_2$) the asymptotic distribution of $R_n^*$ is independent of the parameters so long as one uses intervals of the type

$$S_1(\theta_1, \theta_2) = \left\{ a_{i-1} \theta_1 + \theta_2, a_i \theta_1 + \theta_2 \right\} .$$

A multi-parametric generalization has been given. The theory has been applied to some well known distributions and the actual forms of the asymptotic distribution of $R_n^*$ have been worked out in these cases.

III. Results

Section I. The following example illustrates the basic ideas that are used in the general case in the next section.

**Example 1.** Asymptotic distribution of $R_n^*$ for a normal distribution with unknown mean and known variance with two classes.

Let

$$x_1, \ldots, x_n$$

be a random sample from a normal distribution with unknown mean $\mu$ and known variance, which is taken to be 1 without any loss of generality. Let the range of $x$ viz., $(-\infty, \infty)$ be divided into the two class intervals

$$(-\infty, \bar{x}) \text{ and } (\bar{x}, \infty)$$
where

\[ \bar{x} = \frac{1}{n} (x_1 + \ldots + x_n) . \]

Let \( m \) be the frequency of \( x_\alpha \)'s \( \alpha = 1, 2, \ldots, n \) in the interval \((-\infty, \bar{x})\). We thus have a binomial frequency classification as follows:

<table>
<thead>
<tr>
<th>Class Interval</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>((-\infty, \bar{x}))</td>
<td>(m)</td>
</tr>
<tr>
<td>((\bar{x}, \infty))</td>
<td>(n-m)</td>
</tr>
<tr>
<td>Total</td>
<td>(n)</td>
</tr>
</tbody>
</table>

Let

\[ p^*(\bar{x}) = \Pr\{ x < \bar{x} \mid \mu = \bar{x} \} . \]

Obviously \( p^*(\bar{x}) = 1/2 \). The statistic we are interested in is thus

\[ R_n^* = \frac{(m-np^*(\bar{x}))^2}{np^*(\bar{x})} + \frac{(n-m-n(1-p^*(\bar{x})))^2}{n(1-p^*(\bar{x}))} \]

\[ = \frac{(m-n/2)^2}{n/4} . \]

If we let

\[ g(x_\alpha) = \begin{cases} 1 & \text{if } x_\alpha < \bar{x} \\ 0 & \text{if } x_\alpha > \bar{x} \end{cases}, \quad \alpha = 1, 2, \ldots, n \]

we have

\[ m = \sum_{\alpha=1}^{n} g(x_\alpha) . \]
Since \( x_{\alpha} - \bar{x} \) is a normal random variable with mean 0 for every \( \alpha \), we have
\[
\Pr \{ x_{\alpha} < \bar{x} \} = \frac{1}{2}, \quad \alpha = 1, 2, \ldots, n.
\]
Thus
\[
E(m) = E\left[ \sum_{\alpha=1}^{n} g(x_{\alpha}) \right]
= \sum_{\alpha=1}^{n} \Pr(x_{\alpha} < \bar{x})
= \frac{n}{2}.
\]
In the next section we shall apply the Central Limit theorem to essentially show that the distribution of
\[
y_n = \frac{m - \frac{n}{2}}{\sigma_m}
\]
where \( \sigma_m^2 \) is the variance of \( m \), tends to normal with mean 0 and variance 1. Denoting the distribution of \( y_n \) by \( \mathcal{N}(y_n) \) and the limiting distribution by \( \lim \mathcal{N}(y_n) \) we can say
\[
\lim \mathcal{N}(y_n) = N(0, 1).
\]
Thus
\[
\lim \mathcal{N}(y_n^2) = \mathcal{N}(x_1^2),
\]
a \( x^2 \) distribution with 1 d.f. But
\[
R_n^* = \frac{\sigma_m}{n/4} (y_n^2).
\]
Thus

\[ \lim \mathcal{I}(R_n^*) = \mathcal{I}(A x_1^2) , \]

if we can show that

\[ \lim_{n \to \infty} \frac{\sigma_m^2}{n/4} = A. \]

To show this we calculate \( \sigma_m^2 \). We have,

\[ E(m^2) = E\left[ \frac{n}{\alpha=1} g(x_\alpha) \right]^2 \]

\[ = E\left[ \sum_{\alpha=1}^n (g(x_\alpha))^2 \right] + E\left[ \sum_{\alpha, \beta=1}^n (g(x_\alpha)(g(x_\beta))) \right] \]

\[ = n E[(g(x_1))^2] + n(n-1) E[(g(x_1))(g(x_2))] \]

\[ = n P\{x_1 < \bar{x}\} + n(n-1) P\{x_1 < \bar{x}, x_2 < \bar{x}\} \]

\[ = \frac{n}{2} + n(n-1) \frac{1}{2\pi} \text{Cos}^{-1}\left(\frac{1}{n-1}\right) \]

since

\[ P\{x_1 < \bar{x}\} = \frac{1}{2} \]

and

\[ P\{x_1 < \bar{x}, x_2 < \bar{x}\} = \frac{1}{2\pi} \text{Cos}^{-1}\left(\frac{1}{n-1}\right) \ [c.f. 1] . \]

Thus

\[ E(m^2) = \frac{n}{2} + \frac{n(n-1)}{2\pi} \left(\frac{\pi}{2} - \text{Sin}^{-1}\left(\frac{1}{n-1}\right)\right) . \]
Thus

\[ \sigma_m^2 = E(m^2) - [E(m)]^2 \]

\[ = \frac{n(n-1)}{4} + \frac{n}{2} - \frac{n(n-1)}{2\pi} \sin^{-1} \frac{1}{n} \sin^{-1} \frac{1}{n-1} - \frac{n^2}{4} \]

\[ = n^{\frac{1}{4}} - \frac{n-1}{2\pi} \sin^{-1} \frac{1}{n-1} \, . \]

Thus

\[ \frac{\sigma_m^2}{n^{1/4}} = 1 - \frac{4(n-1)}{2\pi n} \sin^{-1} \frac{1}{n-1} \]

\[ = (1 - \frac{2}{n}) + o(1) \, . \]

Thus the asymptotic distribution of \( R_n^* \) is that of \( (1 - \frac{2}{n}) \chi_1^2 \), where \( \chi_1^2 \) is a \( \chi^2 \) distribution with 1 d.f.

In the above example the coefficient of \( \chi_1^2 \), giving the asymptotic distribution of \( R_n^* \) is \( (1 - \frac{2}{n}) \) which is a constant independent of the unknown parameter \( \mu \).

Section 2. We shall generalize the above results. We shall first give the generalization for the uniparametric case.

A. Uniparametric Case

Let \( F(x|\theta) \) be the c.d.f. of a random variable \( x \), admitting of a density function \( p(x|\theta) \), \( \theta \) taking values over an open interval. We assume \( p(x|\theta) \) to be continuous in both \( x \) and \( \theta \). Let

\[ x_1, x_2, \ldots, x_n \]
be a random sample from the population with $p(x|\theta)$ as the density function. Let $\hat{\theta}(x_1, x_2, \ldots, x_n)$ be an estimate of $\theta$ based on the above random sample with the property that there exists a function $f$ such that

$$(6) \qquad \hat{\theta} - \theta = \frac{1}{n} \sum_{i=1}^{n} f(x_i) + \varepsilon$$

where

(i) $\varepsilon = o_p\left(\frac{1}{\sqrt{n}}\right)$

(ii) $E[f(x)] = 0$ for all $\theta$

and (iii) $\text{Var}[f(x)]$ is finite for all $\theta$.

Thus, $(\hat{\theta} - \theta) = o_p\left(\frac{1}{\sqrt{n}}\right)$

Such estimates exist, e.g., if $\hat{\theta}$ is the maximum likelihood estimate of $\theta$, and if the regularity conditions [c.f. 1, page 500] hold, then $\hat{\theta} - \theta$ satisfies (6) where

$$f(x) = \frac{\frac{d}{d\theta} \log p(x|\theta)}{E\left\{\left(\frac{d}{d\theta} \log p(x|\theta)\right)^2\right\}}.$$ 

Let the range of $x$, viz. $(-\infty, \infty)$ be divided into $k$ non-overlapping intervals $S_i$, $i=1, 2, \ldots, k$ in terms of $\hat{\theta}$ as follows:

$$S_i(\hat{\theta}) = \{x: g_i^{-}(\hat{\theta}) < x \leq g_i(\hat{\theta})\}$$

where $g_i(t)$ is a function of $t$ such that $\frac{dg_i(t)}{dt}$ exists and is continuous in $t$ for all $i$, and

$$-\infty = g_0(\hat{\theta}) < g_1(\hat{\theta}) < \ldots < g_{k-1}(\hat{\theta}) < g_k(\hat{\theta}) = \infty.$$
The following notations will be used in what follows:

\[
\begin{align*}
\Pi_1^*(\theta) &= F[g_1(\theta) | \theta] - F[g_{i-1}(\theta) | \theta] \\
v_1(\theta) &= g_1(\theta) p \left\{ g_1(\theta) | \theta \right\} - g_{i-1}(\theta) p \left\{ g_{i-1}(\theta) | \theta \right\},
\end{align*}
\]

where the prime notation stands for the first differential coefficient with respect to \( \theta \).

\[
(7) \quad \begin{align*}
v_1(\theta) &= \int \frac{\partial}{\partial \theta} p(x | \theta) \, dx, \text{ where } f(x) \text{ is as in (6)} \\
u_1(\theta) &= \int \frac{\partial}{\partial \theta} p(x | \theta) \, dx \\
\sigma^2(\theta) &= \text{Var } [f(x)].
\end{align*}
\]

Let \( m_1 \) = number of \( x_\alpha \)'s \( \alpha = 1, 2, \ldots, n \) in \( S_i(\theta), i = 1, 2, \ldots, k \). We are interested in finding the asymptotic distribution of the statistic

\[
R_n^* = \sum_{i=1}^{k} \frac{(m_i - np_i^*(\hat{\theta}))^2}{np_i^*(\hat{\theta})},
\]

as defined in (5).

We shall now prove the following lemma.

**Lemma 1.** If \( \frac{d}{d\theta} p(x | \theta) \) is uniformly bounded by an integrable function, then

\[
(8) \quad \Pi_1^*(\hat{\theta}) - \Pi_1^*(\theta) = (u_1(\theta) + v(\theta))(\hat{\theta} - \theta) + o_p \left( \frac{1}{\sqrt{n}} \right).
\]
Proof: \( P_1(\hat{\theta}) - P_1(\theta) = \{F[g_1(\hat{\theta})|\theta = \hat{\theta}] - F[g_{1-1}(\hat{\theta})|\theta = \hat{\theta}]\} \)

\[-\{F[g_1(\theta)|\theta] - F[g_{1-1}(\theta)|\theta]\}\]

\[= \{F[g_1(\hat{\theta})|\theta = \hat{\theta}] - F[g_1(\theta)|\theta]\} - \{F[g_{1-1}(\hat{\theta})|\theta = \hat{\theta}] - F[g_{1-1}(\theta)|\theta]\}.\]

By Taylor's expansion, we have,

\[F[g_1(\hat{\theta})|\theta = \hat{\theta}] - F[g_1(\theta)|\theta] = (\hat{\theta} - \theta) \frac{d}{d\theta} F[g_1(\theta)|\theta] + o_p \left( \frac{1}{\sqrt{n}} \right)\]

since \( \hat{\theta} - \theta = o_p \left( \frac{1}{\sqrt{n}} \right) \) and \( \frac{d}{d\theta} F[g_1(\theta)|\theta] \) as computed below is continuous.

\[\frac{d}{d\theta} F[g_1(\theta)|\theta] = \frac{d}{d\theta} \int_{-\infty}^{\infty} g_1(\theta) p(x|\theta) \, dx\]

\[= \left[ g_1'(\theta) p(g_1(\theta)|\theta) + \int_{-\infty}^{\infty} \frac{d}{d\theta} p(x|\theta) \, dx \right].\]

Similarly,

\[F[g_{1-1}(\hat{\theta})|\theta = \hat{\theta}] - F[g_{1-1}(\theta)|\theta] = (\hat{\theta} - \theta) \left[ g_{1-1}'(\theta) p(g_{1-1}(\theta)|\theta) \right.\]

\[+ \left. \int_{-\infty}^{\infty} g_{1-1}'(\theta) \frac{d}{d\theta} p(x|\theta) \, dx \right] + o_p \left( \frac{1}{\sqrt{n}} \right).\]

Substituting the result follows. Q.E.D.
Let
\[ f_\alpha(i) = \begin{cases} 
1 & \text{if } x_\alpha \in S_i(\theta), \text{i.e., if } g_{i-1}(\theta) < x_\alpha \leq g_i(\theta) \\
0 & \text{otherwise}
\end{cases} \]
\[ i=1,2,\ldots,k \]
\[ \alpha=1,2,\ldots,n \]

\textbf{Lemma 2.} Under the condition of Lemma 1,

\[ m_i = \sum_{\alpha=1}^{n} b_i(x_\alpha) + n(\hat{\theta} - \theta) v_i(\theta) + o_p(\sqrt{n}) \tag{9} \]

where \( v_i(\theta) \) is given by (7).

\textbf{Proof:} \( m_i = \text{Number of } x_\alpha's \text{ in } (g_{i-1}(\hat{\theta}), g_i(\hat{\theta})) \)

\[ = \text{Number of } x_\alpha's \text{ in } (g_{i-1}(\theta), g_i(\theta)) \]
\[ + \text{sgn } [g_i(\hat{\theta}) - g_i(\theta)] \left\{ \text{Number of } x_\alpha's \text{ between } g_i(\hat{\theta}) \text{ and } g_i(\theta) \right\} \]
\[ - \text{sgn } [g_{i-1}(\hat{\theta}) - g_{i-1}(\theta)] \left\{ \text{Number of } x_\alpha's \text{ between } g_{i-1}(\hat{\theta}) \text{ and } g_{i-1}(\theta) \right\} \]

Let \( A = \text{Number of } x_\alpha's \text{ between } g_i(\hat{\theta}) \text{ and } g_i(\theta). \)

The desired result follows readily once we show that

\[ (i) \quad A = n ! (\hat{\theta} - \theta) g_1'(\theta) | p(g_i(\theta)|\theta) + o_p(\sqrt{n}) \]

We have,
\[ g_i(\hat{\theta}) - g_i(\theta) = (\hat{\theta} - \theta) g_1'(\theta) + o_p\left(\frac{1}{\sqrt{n}}\right) \]

where \( \hat{\theta} - \theta = o_{p\left(\frac{1}{\sqrt{n}}\right)} \). Thus (i) would follow if we can show that for a rectangular distribution
\[ F_n(u) - u = o_p\left(\frac{1}{\sqrt{n}}\right) \]
where \( u = O_p\left( \frac{1}{\sqrt{n}} \right) \) and \( F_n(z) \) is the sample c.d.f., i.e., \( F_n(z) = \frac{1}{n} \) (the number of sample observations \( \leq z \)). This is true if we can show that given \( \epsilon, \eta > 0 \) there exist an \( N \) such that

\[
P\left( \left| F_n(u) - u \right| > \frac{1}{\sqrt{n}} \right) < \epsilon \text{ for all } n > N.
\]

Since \( u = O_p\left( \frac{1}{\sqrt{n}} \right) \), there exist a \( k > 0 \) such that

\[
P\left( u > \frac{k}{\sqrt{n}} \right) \leq \frac{\epsilon}{2} \text{ for all } n.
\]

It therefore suffices to show that there exists an \( N \) such that

\[
P\left\{ \sup_{0 \leq a \leq \frac{k}{\sqrt{n}}} \left| F_n(a) - a \right| > \frac{1}{\sqrt{n}} \right\} < \epsilon \text{ for all } n > N
\]

or, in other words

\[
(ii) \quad P\left\{ \sup_{0 \leq a \leq \frac{k}{\sqrt{n}}} \left| F_n(a) - a \right| > \frac{1}{\sqrt{n}} \right\} \to 0 \text{ as } n \to \infty.
\]

By Tchebychev's inequality we have for any positive integer \( i \)

\[
P\left\{ \left| F_n\left( \frac{\ln i}{2\sqrt{n}} \right) - \frac{\ln i}{2\sqrt{n}} \right| > \frac{\eta}{2\sqrt{n}} \right\} \leq \frac{\frac{1}{n} \frac{\ln i}{2\sqrt{n}} \left( 1 - \frac{\ln i}{2\sqrt{n}} \right)}{\frac{\eta^2}{4n}} \leq \frac{\frac{1}{n} \frac{\ln i}{2\sqrt{n}}}{\frac{\eta^2}{4n}} \leq \frac{2i}{\eta \sqrt{n}}
\]
Thus

\[
\mathbb{P} \left\{ 0 \leq i \leq \frac{2k}{\eta} + 1 \mid \frac{F_n \left( \frac{i \eta}{2 \sqrt{n}} \right) - \frac{i \eta}{2 \sqrt{n}}}{2 \sqrt{n}} > \frac{\eta}{2 \sqrt{n}} \right\} \leq \sum_{i=1}^{\frac{2k}{\eta} + 1} \frac{2i}{\eta \sqrt{n}} \\
\leq \frac{2}{\eta \sqrt{n}} \frac{1}{2} \left( \frac{2k}{\eta} + 1 \right) \left( \frac{2k}{\eta} + 2 \right) \\
\rightarrow 0 \text{ as } n \rightarrow \infty.
\]

Now let \( a \) be such that for a positive integer \( i \)

\[
\frac{(i-1)\eta}{2 \sqrt{n}} \leq a \leq \frac{i \eta}{2 \sqrt{n}}.
\]

Then

\[
0 \leq a \leq \frac{k}{\sqrt{n}} \implies 0 \leq i \leq \frac{2k}{\eta} + 1.
\]

Then we have,

\[
F_n \left( \frac{(i-1)\eta}{2 \sqrt{n}} \right) - \frac{i \eta}{2 \sqrt{n}} \leq F_n (a) - a \leq F_n \left( \frac{i \eta}{2 \sqrt{n}} \right) - \frac{(i-1)\eta}{2 \sqrt{n}}
\]
or,

\[
F_n \left( \frac{(i-1)\eta}{2 \sqrt{n}} \right) - \frac{(i-1)\eta}{2 \sqrt{n}} - \frac{\eta}{2 \sqrt{n}} \leq F_n (a) - a \leq F_n \left( \frac{i \eta}{2 \sqrt{n}} \right) - \frac{i \eta}{2 \sqrt{n}} + \frac{\eta}{2 \sqrt{n}}
\]
Thus
\[
0 \leq i \leq \frac{2k}{\eta} + 1 \left| F_n \left( \frac{i\eta}{2\sqrt{n}} \right) - \frac{i\eta}{2\sqrt{n}} \right| \leq \frac{\eta}{2\sqrt{n}}
\]
implies
\[
0 \leq a \leq \frac{k}{\sqrt{n}} \left| F_n(a) - a \right| \leq \frac{\eta}{\sqrt{n}}.
\]
Therefore,
\[
P \left\{ \left. \left| F_n \left( \frac{i\eta}{2\sqrt{n}} \right) - \frac{i\eta}{2\sqrt{n}} \right| \leq \frac{\eta}{2\sqrt{n}} \right| 0 \leq i \leq \frac{2k}{\eta} + 1 \right\}
\leq P \left\{ \left. \left| F_n(a) - a \right| \leq \frac{\eta}{\sqrt{n}} \right| 0 \leq a \leq \frac{k}{\sqrt{n}} \right\}.
\]
But by (iii) the probability on the left hand side of the above inequality tends to 1 as \( n \to \infty \) and thus the same is true of the probability on the right hand side. In other words
\[
P \left\{ \left. \left| F_n(a) - a \right| > \frac{\eta}{\sqrt{n}} \right| 0 \leq a \leq \frac{k}{\sqrt{n}} \right\} \to 0 \text{ as } n \to \infty
\]
Hence the lemma follows. Q.E.D.

By substituting for \( p_1^*(\hat{\theta}) - p_1^*(\hat{\theta}) \) and \( m_1 \) from (8) and (9) and using (6) for the expression \( \hat{\theta} - \theta \), we have,
\[ \frac{1}{\sqrt{n}} (m_i - np_i^*(\hat{\theta})) = \frac{1}{\sqrt{n}} (m_i - np_i^*(\theta)) - \sqrt{n} (p_i^*(\hat{\theta}) - p_i^*(\theta)) \]

\[ = \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} [b_i(x_{\alpha}) - p_i^*(\theta)] - \sqrt{n} (\hat{\theta} - \theta) u_i(\theta) + \eta \]

where \( \eta = o_p(1) \)

\[ = \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} [b_i(x_{\alpha}) - p_i^*(\theta) - u_i(\theta) f(x_{\alpha})] + \eta \]

\[ = \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} h_i(x_{\alpha}) + \eta \]

where

\[ h_i(x_{\alpha}) = [b_i(x_{\alpha}) - p_i^*(\theta) - u_i(\theta) f(x_{\alpha})] \quad i=1,2,\ldots,k. \]

Thus by the Mann and Wald result [c.f. 4] the limiting distribution of

\[ \left[ \frac{m_1 - np_1^*(\hat{\theta})}{\sqrt{n}}, \frac{m_2 - np_2^*(\hat{\theta})}{\sqrt{n}}, \ldots, \frac{m_k - np_k^*(\hat{\theta})}{\sqrt{n}} \right] \]

is the same as the limiting distribution of

\[ \left[ \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} h_1(x_{\alpha}), \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} h_2(x_{\alpha}), \ldots, \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} h_k(x_{\alpha}) \right] \]

\[ = \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} [h_1(x_{\alpha}), h_2(x_{\alpha}), \ldots, h_k(x_{\alpha})]. \]

To find the latter limiting distribution we notice that
\[ E[h_i(x_{\alpha})] = \frac{1}{\sqrt{n}} E[h_i(x_{\alpha}) - p_i(\theta) - u_i f(x_{\alpha})] \]

\[ = 0 \text{ for all } i \text{ and } \alpha. \]

Also for \( \alpha \neq \beta \),

\[ E[h_i(x_{\alpha}) h_j(x_{\beta})] = E[h_i(x_{\alpha})] E[h_j(x_{\beta})] = 0 \]

since \( x_{\alpha} \) and \( x_{\beta} \) are independent. Thus the vectors \([h_1(x_{\alpha}), h_2(x_{\alpha}), \ldots, h_k(x_{\alpha})]\) and \([h_1(x_{\beta}), h_2(x_{\beta}), \ldots, h_k(x_{\beta})]\) are mutually uncorrelated for \( \alpha \neq \beta \) and thus the Central Limit theorem applies to

\[ \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} [h_1(x_{\alpha}), h_2(x_{\alpha}), \ldots, h_k(x_{\alpha})] \]

which has thus the limiting distribution \( N(0, \Sigma^*) \) where \( \Sigma^* \) is a \( k \times k \) matrix whose \((i,j)\)th term is given by

\[ \sigma_{ij}^*(\theta) = E[h_i(x_{\alpha}) h_j(x_{\alpha})] . \]

We have

\[ \sigma_{11}^* = E\{h_1(x_{\alpha})\}^2 \]

\[ = E[b_1(x_{\alpha}) - p_1^*(\theta) - u_1(\theta) f(x_{\alpha})]^2 \]

\[ = E[b_1(x_{\alpha}) - p_1^*(\theta)]^2 - 2u_1(\theta) E[b_1(x_{\alpha}) - p_1^*(\theta)] f(x_{\alpha}) + u_1^2(\theta) E[f(x_{\alpha})]^2 \]

\[ = E[b_1(x_{\alpha}) - p_1^*(\theta)]^2 - 2u_1(\theta) E[b_1(x_{\alpha}) f(x_{\alpha})] + u_1^2(\theta) E[f(x_{\alpha})]^2 \]

since \( E[f(x_{\alpha})] = 0 \).
or, \( \sigma_{11}^* = p_1^*(\theta) - p_2^*(\theta) - 2u_1(\theta) \int_{g_{1-1}(\theta)}^{g_1(\theta)} r(x) \mu(x; \theta) \, dx + u_1^2(\theta) \text{Var}[r(x)] \)

since by definition \( b_i(x_\alpha) = \begin{cases} 1 & \text{if } g_{i-1}(\theta) < x_\alpha \leq g_i(\theta) \\ 0 & \text{otherwise} \end{cases} \)

(10) \[ \sigma_{11}^* = p_1^*(\theta) - p_2^*(\theta) - 2u_1(\theta) w_1(\theta) + u_1^2(\theta) \sigma^2(\theta) \]

by using the notation (7). Similarly it can be shown that

(11) \[ \sigma_{ij}^*(\theta) = E[g_i(x_\alpha) g_j(x_\alpha)] \]

\[ = -p_i^*(\theta) p_j^*(\theta) - u_i(\theta) w_j(\theta) - u_j(\theta) w_i(\theta) + \sigma^2(\theta) u_i(\theta) u_j(\theta), \]

for \( i \neq j \).

Henceforth, if an expression \( q(\theta) \) depends on \( \theta \) we shall suppress the dependence and simply write \( q \) for \( q(\theta) \). This will not apply to functions of \( \hat{\theta} \). Thus

(12) \[ \sum^* = P^* - p^* \cdot \cdot \cdot \cdot u \cdot w' - wu' + \sigma^2 uu' \]

where

\[ \begin{pmatrix} p_1^* & 0 & \ldots & 0 \\ 0 & p_2^* & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & p_k^* \end{pmatrix}, \quad p_i^* = p_i^*(\theta) \]
\[ p^* = (p_1^*, p_2^*, \ldots, p_k^*)', \quad p_1^* = p_1(\theta) \]
\[ u = (u_1, u_2, \ldots, u_k)', \quad u_i = u_i(\theta) \]
\[ w = (w_1, w_2, \ldots, w_k)', \quad w_i = w_i(\theta) \]

where \( p^* \)'s, \( u \)'s and \( w \)'s are given by (7) and primes mean transpose. Then

\[
\lim \sqrt{n} \left( \frac{m_1 - np_1^* (\hat{\theta})}{\sqrt{n}}, \frac{m_2 - np_2^* (\hat{\theta})}{\sqrt{n}}, \ldots, \frac{m_k - np_k^* (\hat{\theta})}{\sqrt{n}} \right) = N(0, \Sigma^*)
\]

where \( \Sigma^* \) is given by (12). Since by (8) \( p_1^* (\hat{\theta}) = p_1^* (\theta) + \sigma \left( \frac{1}{\sqrt{n}} \right) \) by \( \hat{\theta} - \theta = o_p \left( \frac{1}{\sqrt{n}} \right) \) we have by the Mann and Wald result [c.f. 4]

\[
\lim \sqrt{n} \left( \frac{m_1 - np_1^* (\hat{\theta})}{\sqrt{np_1^* (\hat{\theta})}}, \frac{m_2 - np_2^* (\hat{\theta})}{\sqrt{np_2^* (\hat{\theta})}}, \ldots, \frac{m_k - np_k^* (\hat{\theta})}{\sqrt{np_k^* (\hat{\theta})}} \right) = N(0, \Sigma^* - 1/2 \Sigma^* - 1/2).
\]

By Lemma 1 in the appendix the asymptotic distribution of \( R_n^* \) is that of

\[ \frac{\sum_{i=1}^{k} \lambda_i z_i^2}{\sum_{i=1}^{k} \lambda_i} \] where \( \sum(z_i) = N(0,1), i=1,2,\ldots,k \) and \( \lambda_1, \lambda_2, \ldots, \lambda_k \) are the characteristic roots of \( \Sigma^* - 1/2 \Sigma^* - 1/2 \). To find these characteristic roots we prove the following lemma.

**Lemma 3.** Let

(i) \( \hat{\theta} \) be an estimate of \( \theta \) satisfying (6)

(ii) \( p^*, p^*, u \) and \( w \) be defined as in (12)

(iii) \( \Sigma^* = p^* - p^* p^* - u w' - w u' + \sigma^2 u u' \)
(iv) the rank of the matrix

\[ C = p \star p' + uu' + uu - \sigma^2 uu' \]

be \( r \).

Then the matrix \( P^{-1} \sum^* \) has at least one characteristic root equal to 0, exactly \( k-r \) characteristic roots equal to 1 and the remaining \( r-1 \) characteristic roots are non-negative and different from 1.

**Proof:** We have identically

\[ \sum_{i=1}^{k} p_i^* = 1 \]

\[ \sum_{i=1}^{k} w_i = \int_{-\infty}^{\infty} f(x) p(x|\theta) \, dx = 0 \quad \text{by (6)} \]

\[ \sum_{i=1}^{k} u_i = \int_{-\infty}^{\infty} \frac{d}{d\sigma} p(x|\theta) \, dx = \frac{d}{d\sigma} \int_{-\infty}^{\infty} p(x|\theta) \, dx = 0. \]

Thus all the rows and columns of \( \sum^* \) add up to 0. It is easy to see that the rank of \( \sum^* \) and hence the rank of \( P^{-1} \sum^* \) is between \( k-1 \) and \( k-r \). Thus \( P^{-1} \sum^* \) has at least one characteristic root equal to 0. Moreover, since \( C \) is of rank \( r \), so is \( P^{-1} C \), \( P^* \) being non-singular and since

\[ P^{-1} \sum^* = I - P^{-1} C \]

\( P^{-1} \sum^* \) has exactly \( k-r \) characteristic roots equal to 1. The remaining \( r-1 \) characteristic roots of \( P^{-1} \sum^* \) are non-negative, since \( P^{-1/2} \sum^* P^{-1/2} \) is
non-negative definite. Q.E.D.

**Corollary 1.** If in Lemma 3 u and w are linearly dependent, say \( w = au \), then \( P^{*-1\sum_*} \) has 1 characteristic root equal to 0, k-2 characteristic roots equal to 1 and the remaining non-negative characteristic root is given by

\[
\lambda = 1 - (2a - \sigma^2) \sum_{i=1}^{k} \frac{u_i^2}{p_i}.
\]

If \( 2a < \sigma^2, 0 \leq \lambda < 1 \).

**Proof:** We have,

\[
C = p^* \bar{p}^* + 2a u u' - \sigma^2 u u' = p^* \bar{p}^* + (2a - \sigma^2) u u'
\]

\( C \) depends only on the two vectors \( \bar{p}^* \) and \( u \) which are linearly independent, since \( \sum_{i=1}^{k} p_i^* = 1 \) and \( \sum_{i=1}^{k} u_i = 0 \). Thus the rank of \( C \) is 2 and the first part of the Corollary follows from Lemma 3 by putting \( r = 2 \).

To find the value of the remaining non-negative characteristic root \( \lambda \) of \( P^{*-1\sum_*} \) we use the fact that the sum of the characteristic roots of a matrix is equal to the trace of the matrix. Thus

\[
k - 2 + \lambda = \text{trace} \, P^{*-1\sum_*} = \text{trace} \left[ I - P^{*-1}(p^* \bar{p}^* + (2a - \sigma^2) u u') \right]
\]

\[
= k - (2a - \sigma^2) \sum_{i=1}^{k} \frac{u_i^2}{p_i}.
\]
Thus

$$\lambda = 1 - (2a - \sigma^2) \sum_{i=1}^{k} \frac{u_i^2}{p_i^*}.$$ 

Evidently $0 \leq \lambda < 1$ since by hypothesis $2a > \sigma^2$, and $\lambda$ is non-negative and $p_i^*$ is positive. Q.E.D.

Theorem 1. Let

(i) the density function $p(x|\theta)$ satisfy the usual regularity conditions for maximum likelihood estimates [c.f. Cramer 1, p. 500]

(ii) $\hat{\theta}(x_1, x_2, ..., x_n)$ be a maximum likelihood estimate of $\theta$ based on a random sample $x_1, x_2, ..., x_n$ from $p(x|\theta)$

(iii) the range of $x$ viz. $-\infty$ to $\infty$ be divided into $k$ intervals $S_i(\hat{\theta})$ $i=1,2,...,k$ as follows:

$$S_i(\hat{\theta}) = \{x: g_{i-1}(\hat{\theta}) < x \leq g_i(\hat{\theta})\}$$

where $g_i(t)$ is a function of $t$ such that $\frac{dg_i(t)}{dt}$ exists and is continuous for all $i=0,1,...,k$, and

$$-\infty = g_0(\hat{\theta}) < g_1(\hat{\theta}) < ... < g_k(\hat{\theta}) = \infty$$

(iv) $m_i$ = Number of $x_{\alpha}$'s in $S_i(\hat{\theta})$, $\alpha=1,2,...,n$

(v) $p_i^*(\hat{\theta}) = F[g_i(\hat{\theta})|\theta = \hat{\theta}] - F[g_{i-1}(\hat{\theta})|\theta = \hat{\theta}]$

where $F$ is the c.d.f. corresponding to the d.f. $p(x|\theta)$.

Then the asymptotic distribution of the statistic
\[
R_n^* = \sum_{i=1}^{k} \frac{(m_i - np_i^*(\theta))^2}{np_i^*(\theta)}
\]

is given by

(15)
\[\chi^2_{k-2} + \lambda z^2,\]

where \(\chi^2_{k-2}\) is a \(\chi^2\) distribution with \(k-2\) d.f., \(z\) is an independent normal variate with mean 0 and variance 1 and

(16)
\[\lambda = 1 - \frac{1}{b^2(\theta)} \sum_{i=1}^{k} \frac{u_i^2(\theta)}{p_i^*(\theta)}\]

where \(b^2(\theta) = E[\frac{d}{d\theta} \log p(x|\theta)|\theta]^2\) and \(u_i(\theta), p_i^*(\theta)\) are as defined in (7).

**Proof:** If \(\hat{\theta}\) is a maximum likelihood estimate of \(\theta\) it is well known that under the regularity conditions in (1) we have

\[\hat{\theta} - \theta = \frac{1}{n} \sum_{\alpha=1}^{n} f(x_{\alpha}) + \eta \text{ where } \eta = o_p\left(\frac{1}{\sqrt{n}}\right)\]

where

\[f(x) = \frac{d}{d\theta} \log p(x|\theta) \quad \frac{1}{b^2(\theta)}\]

Thus

\[E[f(x)] = \frac{1}{b^2(\theta)} E\left[\frac{d}{d\theta} \log p(x|\theta)\right]\]

\[= \frac{1}{b^2(\theta)} \int_{-\infty}^{\infty} \frac{d}{d\theta} p(x|\theta) \, dx\]
thus

\[ E[f(x)] = \frac{1}{b^2(\theta)} \frac{d}{d\theta} \int_{-\infty}^{\infty} p(x|\theta) \, dx \]

\[ = 0 \]

\[ \text{Var}[f(x)] = E[f(x)]^2 = \frac{1}{b^2(\theta)} . \]

Thus by (13) the asymptotic distribution of

\[ \frac{m_1 - np_1^*(\hat{\theta})}{\sqrt{np_1^*(\hat{\theta})}}, \frac{m_2 - np_2^*(\hat{\theta})}{\sqrt{np_2^*(\hat{\theta})}}, \ldots, \frac{m_k - np_k^*(\hat{\theta})}{\sqrt{np_k^*(\hat{\theta})}} \]

is \( \mathcal{N}(0, p^*-1/2 \sum p^*-1/2) \) where by (12)

\[ \sum p^* - p^* p^* - u w - w u' + u u' \]

since \( \sigma^2 = \text{Var}[f(x)] = \frac{1}{b^2} \). But

\[ w_1(\theta) = \int g_1(\theta) f(x) p(x|\theta) \, dx \]

\[ = \frac{1}{b^2(\theta)} \int g_1(\theta) \left[ \frac{d}{d\theta} \log p(x|\theta) \right] p(x|\theta) \, dx \]

\[ = \frac{1}{b^2(\theta)} \int g_1(\theta) \frac{d}{d\theta} p(x|\theta) \, dx \]
or

\[ v_1(\theta) = \frac{1}{b^2(\theta)} u_1(\theta). \]

Thus

\[ w = \frac{1}{b^2} u. \]

Substituting in \[ \sum^* \] we have

\[ \sum^* = p^* - p^* p^* - \frac{u u'}{b^2}. \]

Therefore, the rank of \[ C = p^* p^* + \frac{u u'}{b^2} \] is 2 and by Lemma 3, \[ p^* - \sum^* \] has 1 characteristic root equal to 0, \( k-2 \) characteristic roots equal to 1 and 1 characteristic root \( \lambda \) between 0 and 1 [Corollary to Lemma 3] given by

\[ \lambda = 1 - \frac{1}{b^2} \sum_{i=1}^{k} \frac{u_i^2}{p_i}. \]

By Lemma 1 in the appendix the asymptotic distribution of \( R_n^* \) is given by

\[ \chi^2_{k-2} + \lambda z^2 \] where \( \chi^2_{k-2} \) is a \( \chi^2 \) distribution with \( k-2 \) d.f. and \( z \) is an independent normal variate with mean 0 and variance 1. Q.E.D.

In the above, \( \lambda \) given by (16) normally depends on \( \theta \) through \( u_i, p_i^* \) and \( b \). There is, however, one important case where \( \lambda \) can be freed from its dependence on \( \theta \). This is the so-called translation parameter case. We formulate the result in the form of the following theorem.
Theorem 2. Let

(i) $\hat{\theta}$ be a translation parameter in $p(x|\theta)$, i.e., $p(x|\theta)$ if of the form $p(x - \theta)$

(ii) $\hat{\theta}$ be the maximum likelihood estimate of $\theta$

(iii) $g_i(\hat{\theta}) = \hat{\theta} + d_i$, $i=0,1,2,...,k$, where $d_i$'s are constants [of course, $-\infty = d_0 < d_1 < ... < d_k = \infty$ analogous to the relation $-\infty = g_0(\hat{\theta}) < g_1(\hat{\theta}) < ... < g_k(\hat{\theta}) = \infty$].

Then the asymptotic distribution of $R^*_n$ is of the form (15), where $\lambda$ is independent of $\theta$.

Proof: All that we need show is that $\lambda$ is independent of $\theta$. If $g_i(\hat{\theta}) = \hat{\theta} + d_i$ we have

$$p_i^*(\theta) = \int_{\theta + d_{i-1}}^{\theta + d_i} p(x - \theta) \, dx$$

$$= \int_{d_{i-1}}^{d_i} p(y) \, dy \text{ by change of variable } y = x - \theta.$$

Thus $p_i^*(\theta)$ is independent of $\theta$. Again

$$b^2(\theta) = E_\theta \left( \frac{d}{d\theta} \log p(x - \theta) \right)^2$$

$$= \int_{-\infty}^{\infty} \left( \frac{d}{dx} p(x - \theta) \right)^2 \, dx \text{ since } \frac{dp}{dx} = - \frac{dp}{d\theta}$$
or \[ b^2(\theta) = \int_{-\infty}^{\infty} \frac{d}{dy} \frac{p(y)^2}{p(y)} \, dy. \]

Thus \( b^2 \) is independent of \( \theta \). We have

\[
\begin{align*}
u_i(\theta) &= \int_{\theta+d_{i-1}}^{\theta+d_i} \frac{d}{d\theta} p(x-\theta) \, dx \\
&= \int_{\theta+d_{i-1}}^{\theta+d_i} -\frac{d}{dx} p(x-\theta) \, dx \\
&= -p^*(d_i) + p^*(d_{i-1}).
\end{align*}
\]

Thus \( u_i \) is also independent of \( \theta \). Therefore, \( \lambda \) which depends only on \( p^*(\theta), u_i(\theta) \) and \( b(\theta) \) is independent of \( \theta \). Q.E.D.

**Corollary 1.** If \( \theta \) is a scale parameter in \( p(x|\theta) \), \( \hat{\theta} \) is a maximum likelihood estimate of \( \theta \), then by taking \( g_i(\theta) = d_i \hat{\theta} \), where \( d_i \) is a constant, the asymptotic distribution of \( R_n^* \) is of the form (15), where \( \lambda \) is independent of \( \theta \).

**Proof:** \( p(x|\theta) \) is of the following form

\[
\frac{1}{\theta} p\left(\frac{x}{\theta}\right) \, dx.
\]

If we make the transformation

\[ y = \log x \]
and put

\[ \phi = \log \theta \]

the distribution of \( y \) in terms of \( \phi \) is

\[ p(e^y - \phi) e^y - \phi \ dy \]

whose \( \phi \) appears as a translation parameter and then by Theorem 2 the corollary follows.

B. Applications

**Example 1.** Normal distribution with unknown mean \( \theta \) and variance 1

\[ p(x|\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} (x-\theta)^2} \cdot \]

The maximum likelihood estimate of \( \theta \) is

\[ \hat{\theta}(x_1, x_2, \ldots, x_n) = \bar{x} = \frac{1}{n} (x_1 + x_2 + \ldots + x_n) \cdot \]

We have,

\[ b^2 = E\left[ \frac{d}{d\theta} \log p(x|\theta) \right]^2 \]

\[ = E[(x-\theta)]^2 \]

\[ = 1 \cdot \]

We set

\[ g_i(\hat{\theta}) = \bar{x} + a_i \]

where \(-\infty = a_0 < a_1 < \ldots < a_k = \infty \).
We have,

\[ p^*_1 = \frac{1}{\sqrt{2\pi}} \int_{\theta + a_{1-1}}^{\theta + a_1} e^{-\frac{1}{2} (x - \theta)^2} dx \]

\[ = \frac{1}{\sqrt{2\pi}} \int_{a_{1-1}}^{a_1} e^{-\frac{y^2}{2}} dy \]

\[ = N(a_1) - N(a_{1-1}) , \]

where \( N \) represents the c.d.f. of the unit normal probability distribution.

\[ u_1 = \int_{\theta + a_{1-1}}^{\theta + a_1} \frac{d}{d\theta} p(x|\theta) dx \]

\[ = \frac{1}{\sqrt{2\pi}} \begin{pmatrix} \frac{a_1^2}{2} & -\frac{a_{1-1}^2}{2} \\ e^{-\frac{a_1^2}{2}} & -e^{-\frac{a_{1-1}^2}{2}} \end{pmatrix} . \]

By Theorem 2, the asymptotic distribution of \( R^*_n \) is that of \( \chi^2_{k-2} + \lambda z^2 \)
where \( \chi^2_{k-2} \) is a \( \chi^2 \) distribution with \( k-2 \) d.f., \( z \) is a normal variable with 0 mean and variance 1, and

\[ \lambda = 1 - \sum_{i=1}^{k} \frac{u_{1i}^2}{p_i^*} \]

where \( u_{1i} \) and \( p_i^* \) are as computed above. It is to be noted that both \( u_{1i} \) and \( p_i^* \) are independent of \( \theta \) and thus the asymptotic distribution of \( R^*_n \) is parameter free.
Example 2. Normal distribution with 0 mean and variance $\sigma^2$

$$p(x, \theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad \theta > 0.$$ 

The maximum likelihood estimate of $\sigma^2$ is given by

$$\sigma^2 = s^2 = \frac{1}{n} \sum_{1}^{n} x_i^2.$$ 

We have,

$$\frac{d}{d\theta} \log p(x, \theta) = -\frac{1}{\theta} + \frac{x^2}{\theta^3} = \frac{1}{\theta} \left( \frac{x^2}{\theta^2} - 1 \right).$$

Consequently,

$$b^2 = \frac{1}{\theta^2} E_\theta \left\{ \left( \frac{x^2}{\theta^2} - 1 \right)^2 \right\}$$

$$= \frac{1}{\theta^6} E_\theta \left\{ x^2 - \theta^2 \right\}^2 = \frac{1}{\theta^6} E_\theta (x^4 - 2\sigma^2 x^2 + \sigma^4)$$

$$= \frac{2}{\sigma^2}.$$ 

If we now take

$$g_i(\theta) = s_i$$

where $-\infty = a_0 < a_1 < \ldots < a_k = \infty$. We have,

$$p^*_i = \frac{1}{\sqrt{2\pi}\sigma} \int_{a_{i-1}\sigma}^{a_i\sigma} e^{-\frac{x^2}{2\sigma^2}} dx$$
or \[ p_i^* = \frac{1}{\sqrt{2\pi}} \int_{a_{i-1}}^{a_i} e^{-\frac{y^2}{2}} \, dy \]

\[ = N(a_i) - N(a_{i-1}) \]

Also,

\[ u_i = \frac{1}{\sqrt{2\pi \theta^2}} \int_{a_{i-1} \theta}^{a_i \theta} \left( \frac{x^2}{\theta^2} - 1 \right) e^{-\frac{x^2}{2\theta^2}} \, dx \]

\[ = \frac{1}{\theta} \frac{1}{\sqrt{2\pi}} \left( \begin{array}{cc} \frac{a_i}{\theta} & \frac{a_{i-1}}{\theta} \\ \frac{a_i}{\theta} & \frac{a_{i-1}}{\theta} \end{array} \right) \left( \begin{array}{c} a_i e^{-\frac{a_i^2}{2\theta^2}} - a_{i-1} e^{-\frac{a_{i-1}^2}{2\theta^2}} \\ a_i e^{-\frac{a_i^2}{2\theta^2}} - a_{i-1} e^{-\frac{a_{i-1}^2}{2\theta^2}} \end{array} \right) \]

\[ = t_i \theta \]

where

\[ t_i = \frac{1}{2\sqrt{\pi}} \left( \begin{array}{cc} \frac{a_i}{\theta} & \frac{a_{i-1}}{\theta} \\ \frac{a_i}{\theta} & \frac{a_{i-1}}{\theta} \end{array} \right) \left( \begin{array}{c} a_i e^{-\frac{a_i^2}{2\theta^2}} - a_{i-1} e^{-\frac{a_{i-1}^2}{2\theta^2}} \\ a_i e^{-\frac{a_i^2}{2\theta^2}} - a_{i-1} e^{-\frac{a_{i-1}^2}{2\theta^2}} \end{array} \right) \]

By Theorem 2, the asymptotic distribution of \( R_n^* \) is that of

\[ \chi^2_{k-2} + \lambda z^2 \]

where \( \chi^2_{k-2} \) has a \( \chi^2 \) distribution with \( k-2 \) d.f. \( z \) has a normal distribution with mean 0 and variance 1, and

\[ \lambda = 1 - \frac{1}{b^2} \sum_{i=1}^{k} \frac{u_i^2}{p_i^*} \]
or \[ \lambda = 1 - \theta^2 \sum_{i=1}^{k} \frac{t_{i}^2}{\theta^2 x} \]

\[ = 1 - \sum_{i=1}^{k} \frac{t_{i}}{p_{i}} \]

Note that the distribution is independent of \( \theta \), since \( t_{i} \) and \( p_{i} \) are so.

**Example 3. Gamma distribution with scale parameter**

\[
p(x|\theta) = \frac{1}{\theta^q x^{q-1} e^{-\frac{x}{\theta}}} \quad x > 0, \infty > \theta > 0.
\]

The maximum likelihood estimate of \( \theta \) is given by

\[ \hat{\theta}(x_1, x_2, \ldots, x_n) = \frac{x}{q} \]

where \( \bar{x} \) is the arithmetic mean of \( x_1, x_2, \ldots, x_n \). We have

\[
\frac{d}{d\theta} \log p(x|\theta) = \frac{1}{\theta^2} (x - \theta q).
\]

Thus,

\[ b^2(\theta) = \frac{1}{\theta^4} E_\theta (x - \theta q)^2 \]

\[ = \frac{q}{\theta^2} \]

Set

\[ z_1(\hat{\theta}) = a_1 \frac{x}{q} \]
where \( 0 = a_0 < a_1 < \ldots < a_k = \infty \). We have,

\[
p^*_1(\theta) = \frac{1}{\theta^q r(q)} \int_{a_{i-1}}^{a_i} x^{q-1} e^{-\frac{x}{\theta}} dx
\]

\[
= \frac{1}{r(q)} \int_{a_{i-1}}^{a_i} y^{q-1} e^{-\frac{y}{\theta}} dy .
\]

Thus \( p^*_1(\theta) \) is independent of \( \theta \). According to the notation used before we shall refer to \( p^*_1(\theta) \) as \( p^*_1 \). Also,

\[
u_1 = \int_{a_{i-1}}^{a_i} \frac{d}{d\theta} p(x|\theta) dx
\]

\[
= \frac{1}{\theta^{q+2} r(q)} \int_{a_{i-1}}^{a_i} (x - q\theta) x^{q-1} e^{-\frac{x}{\theta}} dx
\]

\[
= \frac{1}{\theta r(q)} \int_{a_{i-1}}^{a_i} (y - q) y^{q-1} e^{-\frac{y}{\theta}} dy
\]

\[
= \frac{1}{\theta} v_1
\]

where

\[
v_1 = \frac{1}{r(q)} \int_{a_{i-1}}^{a_i} (y - q) y^{q-1} e^{-\frac{y}{\theta}} dy .
\]
By Theorem 2, the asymptotic distribution of $R_n^*$ is that of

$$\chi^2_{k-2} + \lambda z^2$$

where $\chi^2_{k-2}$ has a $\chi^2$ distribution with $k-2$ d.f., $z$ has a normal distribution with mean 0 and variance 1, and

$$\lambda = 1 - \frac{1}{b^2} \sum_{i=1}^{k} \frac{u_i^2}{p_i}$$

$$= 1 - \frac{\theta^2}{q} \sum_{i=1}^{k} \frac{v_i^2}{\theta^2 p_i^*}$$

$$= 1 - \frac{1}{q} \sum_{i=1}^{k} \frac{v_i^2}{p_i^*}$$

The distribution is thus independent of $\theta$, since $v_i$ and $p_i^*$ are so.

C. Multiparametric Generalization

Let $F(x|\theta)$ be the c.d.f. of a random variable $x$ admitting of a density function $p(x|\theta)$ where $\theta$ is vector valued with $q$ components, i.e.,

$$\theta = \begin{pmatrix}
\theta_1 \\
\theta_2 \\
\vdots \\
\theta_q
\end{pmatrix}$$

As in the uniparametric case we assume $p(x|\theta)$ to be continuous in both $x$ and $\theta$. We shall use the following notation
and a prime will indicate transpose.

Let \( \hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_q) \) be an estimate of \( \theta \) based on a random sample \( x_1, x_2, \ldots, x_n \) of size \( n \) from \( F(x|\theta) \) with the property that there exist functions \( f_i(x) \) \( i=1,2,\ldots,q \) such that

\[
\begin{align*}
\text{(i) } & \quad \frac{\hat{\theta}_i - \theta_i}{\sigma_i} = \frac{1}{n} \sum_{\alpha=1}^{n} f_i(x_\alpha) + \epsilon_i, \quad i=1,2,\ldots,q \\
\text{where } & \quad \epsilon_i = o_p\left(\frac{1}{\sqrt{n}}\right) \\
\text{(ii)} & \quad E[f_i(x)] = 0 \\
\text{(iii)} & \quad \text{Var}[f_i(x)] \text{ is finite.}
\end{align*}
\]

(17)

We shall represent (17) in the vector form

\[
\theta - \hat{\theta} = \frac{1}{n} \sum_{\alpha=1}^{n} f(x_\alpha) + \epsilon
\]

where

\[
f = (f_1, f_2, \ldots, f_q)'
\]

\[
\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_q)'.
\]
Then \( \hat{\theta} - \theta = O_p \left( \frac{1}{\sqrt{n}} \right) \). Such estimates exist, e.g., if \( \hat{\theta} \) is a maximum likelihood estimate of \( \theta \), and if the usual regularity conditions hold,

\[
f(x) = J^{-1} \frac{\partial}{\partial \theta} \log p(x|\theta)
\]

where

\[
J = E \left\{ \frac{\partial^2}{\partial \theta \partial \theta} \log p(x|\theta) \right\}
\]

is the information matrix.

Let the range of \( x \), viz. \((-\infty , \infty)\) be partitioned into \( x(k > q) \) intervals \( S_i, i=1,2,\ldots,k \) in terms of the estimate \( \hat{\theta} \) as follows:

\[
S_i(\hat{\theta}) = \{ x: g_{i-1}(\hat{\theta}) < x \leq g_i(\hat{\theta}) \}
\]

where \( g_i(\theta) \) is a continuous function of \( \theta \) with continuous derivative \( \frac{\partial}{\partial \theta} g_i(\theta) \) for \( i=0,\ldots,k \) and \(-\infty = g_0(\hat{\theta}) < g_1(\hat{\theta}) < \ldots < g_k(\hat{\theta}) = \infty\).

We shall use the following notations in what follows:

\[
\begin{align*}
& p_i^\dagger(\theta) = F[g_i(\theta)|\theta] - F[g_{i-1}(\theta)|\theta] \\
& v_i(\theta) = p[g_i(\theta)|\theta] \frac{\partial}{\partial \theta} g_i(\theta) + p[g_{i-1}(\theta)|\theta] \frac{\partial}{\partial \theta} g_{i-1}(\theta) \\
\end{align*}
\]

\[
(19)
\begin{align*}
& w_i(\theta) = \int_{g_{i-1}(\theta)}^{g_i(\theta)} f(x) p(x|\theta) \, dx \\
& u_i(\theta) = \int_{g_{i-1}(\theta)}^{g_i(\theta)} \frac{\partial}{\partial \theta} p(x|\theta) \, dx \\
& \gamma = \text{Cov} [f(x)] .
\end{align*}
\]
It is to be noted that $v_1, w_1, u_1$ are $q \times 1$ column vectors and $\sum$ is a $q \times q$ matrix. Let

$$m_1 = \text{Number of } x_\alpha's \text{ for } \alpha = 1, \ldots, n \text{ falling in } S_1(\hat{\theta})$$

and

$$p_i^*(\hat{\theta}) = \mathbb{P}[g_i(\hat{\theta}) | \theta = \hat{\theta}] - \mathbb{P}[g_{i-1}(\hat{\theta}) | \theta = \hat{\theta}].$$

Let

$$b_i(x_\alpha) = \begin{cases} 1 & \text{if } x_\alpha \in S(\theta), \text{ i.e., if } g_{i-1}(\theta) < x_\alpha \leq g_i(\theta) \\ 0 & \text{otherwise} \end{cases} \quad \alpha = 1, 2, \ldots, n.$$ 

We are interested in finding the asymptotic distribution of

$$R_n^* = \sum_{i=1}^{k} \frac{(m_i - np_i^*(\hat{\theta}))^2}{np_i^*(\hat{\theta})}.$$ 

The following results hold.

**Lemma 4.** If $\frac{\partial}{\partial \theta_i}$ is uniformly bounded by an integrable function $h_i(x)$ for $i = 1, 2, \ldots, q$, then

(a) \[ p_i^*(\hat{\theta}) - p_i^*(\theta) = (\hat{\theta} - \theta) \left[ v_i(\theta) + u_i(\theta) \right] + o_p\left( \frac{1}{\sqrt{n}} \right). \]

(b) \[ m_1 = \sum_{\alpha=1}^{n} b_i(x_\alpha) + n(\hat{\theta} - \theta)' v_i(\theta) + o_p(\sqrt{n}). \]

**Proof:** For (a), as in Lemma 1. For (b), as in Lemma 2. Q.E.D.

It is easy to derive, just as in the uniparametric case that
\[
\lim \left( \frac{m_1 - np_1^*(\hat{\theta})}{\sqrt{n}}, \frac{m_2 - np_2^*(\hat{\theta})}{\sqrt{n}}, \ldots, \frac{m_k - np_k^*(\hat{\theta})}{\sqrt{n}} \right) = N(0, \Sigma^*)
\]

where

\[
\Sigma^* = (\sigma^*_{ij})
\]

\[
\sigma^*_{ii} = p_1^*(\theta) - p_1^*(\hat{\theta}) - u_i'(\theta) w_i'(\theta) - w_i'(\theta) u_i'(\theta) + u_i'(\theta) \Sigma u_i'(\theta)
\]

\[
\sigma^*_{ij} = p_i^*(\theta) p_j^*(\theta) - u_i'(\theta) w_j'(\theta) - w_i'(\theta) u_j'(\theta) + u_i'(\theta) \Sigma u_j'(\theta)
\]

Henceforth, we shall follow the convention that if a function \( q(\theta) \) depends on \( \theta \) we shall suppress the dependence by writing \( q \) for \( q(\theta) \). This will not apply to function of \( \hat{\theta} \).

Thus \( \Sigma^* \) is given by,

\[
(20) \quad \Sigma^* = \mathbf{P}^* - \mathbf{P} \mathbf{P}^* - \mathbf{U}' \mathbf{W} - \mathbf{W}' \mathbf{U} + \mathbf{U}' \mathbf{U}
\]

where

\[
\mathbf{P}^* = \begin{pmatrix} \mathbf{P}_1 & 0 & \ldots & 0 \\ 0 & \mathbf{P}_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathbf{P}_k \end{pmatrix}
\]

\[
\mathbf{P}^* = (p_1^*, p_2^*, \ldots, p_k^*)'
\]

\( \mathbf{U} \) is the matrix formed by \( u_1, \ldots, u_k \) as columns, and

\( \mathbf{W} \) is the matrix formed by \( w_1, \ldots, w_k \) as columns.

Both \( \mathbf{U} \) and \( \mathbf{W} \) are \( q \times k \) matrices.
It follows by the Mann and Wald result [c.f. 4]

$$\lim \frac{\sigma_{l} \text{np}_{l}^*(\hat{\theta}) \sigma_{k} \text{np}_{k}^*(\hat{\theta})}{\sqrt{\sigma_{l}^2 \text{np}_{l}^*(\hat{\theta}) \sigma_{k}^2 \text{np}_{k}^*(\hat{\theta})}} = N(0, \sigma_{-1}^{2} \sum \sigma_{-1}^{2})$$

By Lemma 1 in the appendix the asymptotic distribution of $R_n^*$ is that of

$$\sum_{i=1}^{k} \lambda_i z_i^2$$

where, $z_1, z_2, \ldots, z_k$ are mutually independent normal variates with mean 0 and variance 1, and $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the characteristic roots of $P_{-1}^{*} \sum^{*}$. To find these characteristic roots we prove the following lemma

**Lemma 5.** Let

(i) $\hat{\theta}$ be an estimate of $\theta$ satisfying (28)

(ii) $P^*, p^*, U$ and $W$ are as defined in (30)

(iii) the rank of the matrix

$$C = p^* p'^* + U' W + W' U - U' \sum U$$

is $r$

(iv) $\sum^{*} = P^* - C$.

Then (a) the matrix $P_{-1}^{*} \sum^{*}$ has 1 characteristic root equal to 0, $k-r$ characteristic roots equal to 1 and $r-1$ non-negative characteristic roots.

(b) Moreover, if

(i) rank of $W$ is $q$, and

(ii) there exists a symmetric positive definite $q \times q$ matrix $D$ such that $U = D W$ and $2D - D \sum D$ is positive definite.

Then the rank of $C$ is $1+q$, and the $q$ non-negative characteristic roots
of $p_1^* \ldots \sum \ldots \sum \ldots p_k^*$ lie between 0 and 1.

Proof: The proof of part (a) of the lemma is similar to that of Lemma 3. To prove part (b) we have,

$$C = p^* p^* + U' W + W' U - U \sum U$$

$$= p^* p^* + W' (2D - D \sum D) W,$$

since $U = DW$ by hypothesis. Since by hypothesis $W$ is of rank $q$ and $2D - D \sum D$ is positive definite, the rank of $W' (2D - D \sum D) W$ is $q$ and thus $C$ is of rank $q+1$, $p^*$ being linearly independent of $w_i$'s, $i=1, \ldots, k$. Thus by part (a) of the lemma, $P_{\sum \sum}^*$ has $q$ positive characteristic roots different from 0 and 1. To show that these roots lie between 0 and 1, we have

$$P_{\sum \sum}^* = I - P_{\sum \sum}^* C$$

$$= I - P_{\sum \sum}^* [p^* p^* + W' (2D - D \sum D) W].$$

Now by hypothesis $2D - D \sum D$ is a positive definite $q \times q$ matrix. Thus $W' (2D - D \sum D) W$ is non-negative definite. Moreover, $p^* p^*$ is non-negative definite and $P_{\sum \sum}^*$ is positive definite, diagonal. Thus we can express $P_{\sum \sum}^*$ as follows

$$P_{\sum \sum}^* = I - F$$

where $F$ is a non-negative definite matrix. On the other hand $P_{\sum \sum}^* P_{\sum \sum}^*$ is also non-negative definite, $P^*$ being positive definite and diagonal and $\sum^*$ being a covariance matrix. Thus the characteristic roots of $P_{\sum \sum}^*$ lie between 0 and 1 and the lemma follows. Q.E.D.
Theorem 3. If

(i) the density function \( p(x|\theta) \), where \( \theta \) is a column q-vector, satisfies the usual regularity conditions for maximum likelihood estimates

(ii) \( \hat{\theta}(x_1, \ldots, x_n) \) is the maximum likelihood estimate of \( \theta \) based on a random sample \( x_1, \ldots, x_n \) from \( p(x|\theta) \)

(iii) the range of \( x \) viz. \((-\infty, \infty)\) is divided into \( k \) intervals \( S_i \), \( i=1,2,\ldots,k \) as follows:

\[
S_i(\theta) = \{ x : g_{i-1}(\hat{\theta}) < x \leq g_i(\hat{\theta}) \}
\]

where \( g_i(t_1, \ldots, t_q) \) is a function of \( t = (t_1, t_2, \ldots, t_q)' \) such that \( \frac{\partial}{\partial t} g_i(t) \) exists and is continuous for all \( i=1,2,\ldots,k \) and

\[-\infty = g_0(\hat{\theta}) < g_1(\hat{\theta}) < \ldots < g_k(\hat{\theta}) = \infty\]

(iv) \( m_i = \text{Number of } x_\alpha \text{'s, } \alpha=1,2,\ldots,n \text{ in } S_i(\hat{\theta}) \)

(v) \( p_i^*(\hat{\theta}) = F[g_i(\hat{\theta})|\theta = \hat{\theta}) - F(g_{i-1}(\hat{\theta})|\theta = \hat{\theta}) \)

Then the matrix

\[
P_{**}^{-1} = I - P_{**}^{-1}[p^* p^*'] + W' J W,
\]

where \( \sum^* \), \( P^* \) and \( W \) are as defined in (20) and \( J = E_{\theta} \left[ \frac{\partial^2}{\partial \theta \partial \theta} \log p(x|\theta) \right] \)

has 1 characteristic root equal to 0, exactly \( k-q-1 \) characteristic roots equal to 1 and \( q \) characteristic roots \( \lambda_1, \lambda_2, \ldots, \lambda_q \) lying between 0 and 1. The asymptotic distribution of

\[
R_n^* = \sum_{i=1}^{k} \frac{(m_i - np_i(\hat{\theta}))^2}{np_i(\hat{\theta})}
\]
is that of

\[ \chi^2_{k-q-1} + \lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_q z_q^2 \]

where \( \chi^2_{k-q-1} \) is a \( \chi^2 \) distribution with \( k-q-1 \) d.f. and \( z_1, z_2, \ldots, z_q \) are mutually independent normal variates with mean 0 and variance 1.

**Proof:** Since \( \hat{\theta} \) is the maximum likelihood estimate of \( \theta \), we have in the representation of (18) of \( \hat{\theta} - \theta \),

\[ f(x) = J^{-1} \frac{\partial}{\partial \theta} \log p(x|\theta) \]

\[ \Sigma = \text{Cov} [f(x)] \]

\[ = E \left\{ [f(x)][f(x)]' \right\} \]

\[ = E \left\{ [J^{-1} \left( \frac{\partial}{\partial \theta} \log p(x|\theta) \right)] \left[ \frac{\partial}{\partial \theta} \log p(x|\theta) \right] J^{-1} \right\} \]

\[ = J^{-1} E \left[ \frac{\partial}{\partial \theta} \log p(x|\theta) \frac{\partial}{\partial \theta} \log p(x|\theta) \right] J^{-1} \]

\[ = J^{-1} . \]

Now, using the notation in (19), we have

\[ w_1 = \int_{g_{1-1}(\theta)} g_1(\theta) f(x) p(x|\theta) \, dx \int_{g_{1-1}(\theta)} dx \]

\[ = J^{-1} \int_{g_{1-1}(\theta)} g_1(\theta) \left\{ \frac{\partial}{\partial \theta} \log p(x|\theta) \right\} p(x|\theta) \, dx \]
or \[ w_i = J^{-1} \int_{g_i(\theta)} \frac{\partial}{\partial \theta} p(x|\theta) \, dx \]
\[ = J^{-1} u_i. \]

Therefore,

\[ U = J W \]

where \( U \) and \( W \) are as defined in (20). It is easy to verify that

\[ P^{* -1} \Sigma^{*} = I - P^{* -1} (p^{*} p^{*'} + W' J W) = I - P^{* -1} (p^{*} p^{*'} + U' J^{-1} U). \]

Also in Lemma 5,

\[ D = J \]

and \[ \Sigma = J^{-1}. \]

Thus \( 2D - D \Sigma D = J \) which is a positive definite matrix. The rank of \( W \) is also \( q \). Thus all the conditions of Lemma 5, part (b) hold. Thus \( P^{* -1} \Sigma^{*} \) has 1 characteristic root equal to 0, \( k-q-1 \) characteristic roots equal to 1 and \( q \) characteristic roots \( \lambda_1, \lambda_2, \ldots, \lambda_q \) lying between 0 and 1. The asymptotic distribution of \( R_n^* \) now follows from Lemma 1 in the appendix.

Q.E.D.

Theorem 2 and its corollary can also be generalized to the multi-parametric case. As a matter of fact if a density function involves scale and translation parameters alone, the same principle can be applied to free the \( \lambda \)'s in the asymptotic distribution of \( R_n^* \) from their dependence on the parameters. Following is an example to elucidate it.
D. Applications

1. Normal with unknown mean and variance

Let $x$ be a normal random variable with the density function

$$p(x|\theta) = \frac{1}{\sqrt{2\pi}\theta_2} e^{-\frac{1}{2\theta_2^2}(x - \theta_1)^2}$$

where

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}.$$ 

We have,

$$E\left(\frac{\partial}{\partial \theta_1} \log p(x|\theta)\right)^2 = \frac{1}{\theta_2^2}$$

$$E\left(\frac{\partial}{\partial \theta_2} \log p(x|\theta)\right)^2 = \frac{2}{\theta_2^2}$$

and,

$$E\left(\frac{\partial}{\partial \theta_1} \log p(x|\theta) \frac{\partial}{\partial \theta_2} \log p(x|\theta)\right) = 0.$$

Thus,

$$J = \begin{pmatrix} \frac{1}{\theta_2^2} & 0 \\ 0 & \frac{2}{\theta_2^2} \end{pmatrix}.$$ 

Let

$$x_1, x_2, \ldots, x_n$$
be a random sample of elements whose density is \( p(x|\theta) \)

\[
(\hat{\theta}_1, \hat{\theta}_2) = (\bar{x}, s)
\]

is the maximum likelihood estimate of \( \theta \), where

\[
\hat{\theta}_1 = \bar{x} = \frac{1}{n} \sum_{\alpha=1}^{n} x_\alpha
\]

\[
\hat{\theta}_2 = s = \frac{1}{n} \sum_{\alpha=1}^{n} (x_\alpha - \bar{x})^2
\]

Let the range of \( x \) viz. \((-\infty, \infty)\) be divided into subsets as follows:

\[
S_1(\bar{x}, s) = \left\{ x: \bar{x} + a_1 \cdot s < x < \bar{x} + a_1 \cdot s \right\}
\]

where \(-\infty = a_0 < a_1 < \ldots < a_{k-1} < a_k = \infty\), i.e., we take

\[
g_1(\theta) = \hat{\theta}_1 + a_1 \hat{\theta}_2.
\]

We have

\[
p^*_1 = \int_{g_{1-1}(\theta)} \int g_1(\theta) p(x|\theta) \, dx
\]

\[
= \frac{1}{\sqrt{2\pi} \theta_2} \int_{\frac{\theta_1 + a_1 \theta_2}{\theta_1 + a_1 \theta_2}}^{\frac{\theta_1 + a_1 \theta_2}{\theta_1 + a_1 \theta_2}} e^{-\frac{1}{2\theta_2^2} (x - \theta_1)^2} \, dx
\]

\[
= \frac{1}{\sqrt{2\pi}} \int_{a_1-l}^{a_1} e^{-\frac{y^2}{2}} \, dy
\]

\[
= N(a_1) - N(a_{1-1})
\]
\[ u_1 = \left( \begin{array}{c}
\int_{\theta_1 - \sigma_1}^{\theta_1 + \sigma_1} \frac{g_1(\theta)}{\theta_1} p(x|\theta) \, dx \\
\int_{\theta_1 - \sigma_1}^{\theta_1 + \sigma_1} \frac{\partial}{\partial \theta_1} p(x|\theta) \, dx \\
\int_{\theta_1 - \sigma_1}^{\theta_1 + \sigma_1} \frac{g_1(\theta)}{\theta_2} p(x|\theta) \, dx \\
\int_{\theta_1 - \sigma_1}^{\theta_1 + \sigma_1} \frac{\partial}{\partial \theta_2} p(x|\theta) \, dx
\end{array} \right) \]

\[ = \frac{1}{\sqrt{2\pi}} \left( \int_{\theta_1 - \sigma_1}^{\theta_1 + \sigma_1} \frac{(x - \theta_1)}{\sigma_2} e^{-\frac{1}{2}(x - \theta_1)^2} \, dx \right) \]

\[ = \frac{1}{\sqrt{2\pi}} \left( \int_{\theta_1 - \sigma_1}^{\theta_1 + \sigma_1} \frac{1}{\sigma_2} e^{-\frac{1}{2}(x - \theta_1)^2} \left[ -\frac{1}{2\sigma_2} (x - \theta_1)^2 + \frac{1}{\sigma_1^2} \frac{(x - \theta_1)^2}{\sigma_2^2} - \frac{1}{2\sigma_2} (x - \theta_1)^2 \right] \, dx \right) \]

\[ = \frac{1}{\sqrt{2\pi}} \left( \int_{\theta_1 - \sigma_1}^{\theta_1 + \sigma_1} \frac{1}{\theta_2} e^{-\frac{1}{2}(y^2 - 1)} \, dy \right) \]

\[ = \frac{1}{\sqrt{2\pi}} \left( \int_{\theta_1 - \sigma_1}^{\theta_1 + \sigma_1} y e^{-\frac{1}{2}(y^2 - 1)} \, dy \right) \]

\[ = \frac{1}{\sqrt{2\pi}} \left( \int_{\theta_1 - \sigma_1}^{\theta_1 + \sigma_1} \left( y^2 - 1 \right) e^{-\frac{1}{2}y^2} \, dy \right) \]
Thus

\[
u_1 = \frac{1}{\sqrt{2\pi} \theta_2} \begin{pmatrix}
-\frac{s_{i-1}^2}{2} & -\frac{s_i^2}{2} \\
\theta_2 e^{-\frac{s_{i-1}^2}{2}} & \theta_2 e^{-\frac{s_i^2}{2}} \\
\end{pmatrix}
\]

\[
\begin{pmatrix}
s_i \\
\sqrt{2} t_i \\
\theta_2 \\
\end{pmatrix}
\]

where

\[
s_1 = \frac{1}{\sqrt{2\pi}} \begin{pmatrix}
-\frac{s_{i-1}^2}{2} & -\frac{s_i^2}{2} \\
\theta_2 e^{-\frac{s_{i-1}^2}{2}} & \theta_2 e^{-\frac{s_i^2}{2}} \\
\end{pmatrix}
\]

and

\[
t_i = \frac{1}{2 \sqrt{\pi}} \begin{pmatrix}
-\frac{s_{i-1}^2}{2} & -\frac{s_i^2}{2} \\
\theta_2 e^{-\frac{s_{i-1}^2}{2}} & \theta_2 e^{-\frac{s_i^2}{2}} \\
\end{pmatrix}
\]

Therefore,

\[
U^{-1} \nu U = \begin{pmatrix}
s_1 \\
\theta_2 \\
\sqrt{2} t_1 \\
\theta_2 \\
\vdots \\
\sqrt{2} t_k \\
\theta_2 \\
\end{pmatrix} \cdot \begin{pmatrix}
\theta_2 & 0 \\
0 & \theta_2 \\
\sqrt{2} t_1 & \sqrt{2} t_2 \\
\theta_2 & \theta_2 \\
\vdots \\
\sqrt{2} t_k & \theta_2 \\
\theta_2 & \theta_2 \\
\end{pmatrix} \cdot \begin{pmatrix}
s_1 \\
\theta_2 \\
\sqrt{2} t_1 \\
\theta_2 \\
\vdots \\
\sqrt{2} t_k \\
\theta_2 \\
\end{pmatrix}
\]
\[
U'J^{-1}U = \begin{pmatrix}
  s_1^2 t_1 & s_1 s_2 t_1 t_2 & \cdots & s_1 s_k t_1 t_k \\
  s_1 s_2 t_1 t_2 & s_2^2 t_2 & \cdots & s_2 s_k t_2 t_k \\
  \vdots & \vdots & \ddots & \vdots \\
  s_1 s_k t_1 t_k & s_2 s_k t_2 t_k & \cdots & s_k^2 t_k \\
\end{pmatrix}
\]

\[
= s s' + t t'
\]

where

\[
s = (s_1, s_2, \ldots, s_k)
\]

and \[
t = (t_1, t_2, \ldots, t_k).
\]

Thus we have,

\[
\sum^* = P^* - p^* p'^* - s s' - t t'.
\]

Therefore,

\[
P^{*1/2} \sum^* P^{*1/2} = I - (P^{*1/2} p^*) (P^{*1/2} p^*)' - (P^{*1/2} s) (P^{*1/2} s)' - (P^{*1/2} t) (P^{*1/2} t)'
\]

\[
= I - A A'
\]

where \( A \) is the \( k \times 3 \) matrix

\[
(p^{*1/2} p^* \quad p^{*1/2} s \quad p^{*1/2} t).
\]

By Theorem 3, \( P^{*1/2} \sum^* P^{*1/2} \) (or \( P^{*1/2} \sum^* \)) has 1 characteristic root equal to 0, exactly \( k - 3 \) characteristic roots equal to 1 and 2.
characteristic roots $\lambda_1$ and $\lambda_2$, $0 \leq \lambda_1, \lambda_2 < 1$. Then the asymptotic distribution of $R_n^*$ is that of

$$\chi^2_{k-3} + \lambda_1 z_1^2 + \lambda_2 z_2^2$$

where $z_1$ and $z_2$ are normal variates with 0 mean and variance 1.

To evaluate $\lambda_1$ and $\lambda_2$ we observe that the characteristic roots of $P^{-1/2} \sum P^{-1/2}$ are 1 minus the characteristic roots of $A A'$. We now use the fact that the non-zero characteristic roots of $A A'$ are the same as the non-zero characteristic roots of $A A'$. We have

$$A A' = \begin{pmatrix} P \cdot P^{-1/2} \\ s \cdot P^{-1/2} \\ t \cdot P^{-1/2} \end{pmatrix} \cdot \begin{pmatrix} P^{-1/2} & P^{-1/2} & P^{-1/2} \\ s & P^{-1/2} & s \\ t & P^{-1/2} & t \end{pmatrix}$$

$$= \begin{pmatrix} P \cdot P^{-1/2} & P \cdot P^{-1/2} & P \cdot P^{-1/2} \\ s & s & s \\ t & t & t \end{pmatrix} \cdot \begin{pmatrix} P \cdot P^{-1/2} & P \cdot P^{-1/2} & P \cdot P^{-1/2} \\ s & s & s \\ t & t & t \end{pmatrix}$$

Now,

$$p \cdot P^{-1} p = \sum_{i=1}^{k} p_i = 1$$

$$p \cdot P^{-1} s = s \cdot P^{-1} p = \sum_{i=1}^{k} s_i = 0$$

$$p \cdot P^{-1} t = t \cdot P^{-1} p = \sum_{i=1}^{k} t_i = 0$$
\[ s', p^* - 1 \, t = t', p^* - 1 \, s = \frac{k}{i = 1} \frac{s_i t_i}{p_i} \]

\[ s', p^* - 1 \, s = \sum_{i = 1}^{k} \frac{s_i^2}{p_i} \]

\[ t', p^* - 1 \, t = \sum_{i = 1}^{k} \frac{t_i^2}{p_i} \]

\[
\begin{pmatrix}
1 & 0 & 0 \\
0 & \frac{k s_i}{l p_i} & \frac{k s_{i t_i}}{l p_i} \\
0 & \frac{k s_{i t_i}}{l p_i} & \frac{k t_i}{l p_i}
\end{pmatrix}
\]

or,

\[
A^t \ A = \begin{pmatrix}
1 & 0 & 0 \\
0 & \frac{k s_i}{l p_i} & \frac{k s_{i t_i}}{l p_i} \\
0 & \frac{k s_{i t_i}}{l p_i} & \frac{k t_i}{l p_i}
\end{pmatrix}
\]

Thus \( \lambda_1 \) and \( \lambda_2 \) are the characteristic roots of the 2 x 2 matrix

\[
\begin{pmatrix}
1 - \frac{k s_i}{l p_i} & \frac{k s_{i t_i}}{l p_i} \\
\frac{k s_{i t_i}}{l p_i} & 1 - \frac{k t_i}{l p_i}
\end{pmatrix}
\]
Lemma 1. Let

\[ \lim \mathcal{N} \left( \frac{\sum_{i=1}^{m_1-nP_1}{}, \frac{m_2-nP_2}{{n_2}^*}, \ldots, \frac{m_k-nP_k}{{n_k}^*}}{np_i^*} \right) = N(0, P_{*}^{-1/2} \sum{P_{*}^{-1/2}}) \]

(ii) \[ R_n^* = \sum_{i=1}^{k} \left( \frac{(m_i-nP_i^*)^2}{np_i^*} \right). \]

Then \( \lim \mathcal{N}(R_n^*) = \mathcal{N} \left( \sum_{i=1}^{k} \lambda_i z_i^2 \right), \) where \( \lambda_1, \lambda_2, \ldots, \lambda_k \) are the characteristic roots of \( P_{*}^{-1/2} \sum{P_{*}^{-1/2}} \) and \( z_1, z_2, \ldots, z_k \) are mutually independent normal random variates with mean 0 and variance 1.

Proof: The lemma is essentially a consequence of the following known result [c.f. 3]. If \( \mathcal{N}(y) = N(0, U) \) and \( \lambda_1, \lambda_2, \ldots, \lambda_k \) are the characteristic roots of \( U \) then \( \mathcal{N}(y'y) = \mathcal{N} \left( \sum_{i=1}^{k} \lambda_i z_i^2 \right), \) where \( \mathcal{N}(z_i) = N(0, 1) \) for \( i=1,2,\ldots,k. \) This implies that the asymptotic distribution of \( R_n^* \) is that of \( \sum_{i=1}^{k} \lambda_i z_i^2 \) where \( \lambda_1, \lambda_2, \ldots, \lambda_k \) are the characteristic roots of \( P_{*}^{-1/2} \sum{P_{*}^{-1/2}} \) and \( \mathcal{N}(z_i) = N(0, 1), \) \( i=1,2,\ldots,k. \) But the characteristic roots of \( P_{*}^{-1/2} \sum{P_{*}^{-1/2}} \) are the same as the characteristic roots of \( P_{*}^{-1} \sum{P_{*}^{-1}}, \) since \( P_{*} \) being non-singular, the equation

\[ \left| P_{*}^{-1/2} \sum{P_{*}^{-1/2}} - \lambda I \right| = 0 \]

yields the same solutions of \( \lambda \) as the equation

\[ \left| \sum{P_{*}^{-1}} - \lambda P_{*} \right| = 0 \]

or as the equation

\[ \left| P_{*}^{-1} \sum{P_{*}^{-1}} - \lambda I \right| = 0. \] Q.E.D.
BIBLIOGRAPHY


ACKNOWLEDGEMENT

I am grateful to Professor Herman Chernoff for suggesting the problem and to Professors Herman Rubin and Herman Chernoff for helpful suggestions and encouragement. My thanks go to the Office of Naval Research for financial support.
<table>
<thead>
<tr>
<th>Address</th>
<th>Code</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chief of Naval Research&lt;br&gt;Office of Naval Research&lt;br&gt;Washington 25, D.C.&lt;br&gt;Attn: Physical Sciences Div., Code 420</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Chief, Bureau of Ships&lt;br&gt;Department of the Navy&lt;br&gt;Washington 25, D.C.&lt;br&gt;Attn: H. Weingarten Code 223</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Commanding Officer&lt;br&gt;U.S. Naval Powder Factory&lt;br&gt;Indianhead, Md.&lt;br&gt;Attn: F. Freshman, R. &amp; D.</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Director, Naval Research Lab.&lt;br&gt;Washington 25, D.C.&lt;br&gt;Attn: Technical Information Officer</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Document Service Center&lt;br&gt;Armed Services Technical Information Agency&lt;br&gt;Knott Building&lt;br&gt;Dayton 2, Ohio&lt;br&gt;Attn: DSC-S</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Dept. of Mathematical Statistics&lt;br&gt;University of North Carolina&lt;br&gt;Chapel Hill, N.C.</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Office of Naval Research&lt;br&gt;Department of the Navy&lt;br&gt;17th St. &amp; Constitution Ave.&lt;br&gt;Washington 25, D.C.&lt;br&gt;Attn: Code 433</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Office of the Asst. Naval Attaché for Research&lt;br&gt;Naval Attaché&lt;br&gt;American Embassy&lt;br&gt;Navy No. 100&lt;br&gt;Fleet Post Office&lt;br&gt;New York, N.Y.</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Office of Technical Services&lt;br&gt;Department of Commerce&lt;br&gt;Washington 25, D.C.</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Scientific Section&lt;br&gt;Office of Naval Research&lt;br&gt;Department of the Navy&lt;br&gt;1000 Geary St.&lt;br&gt;San Francisco 9, Calif.</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Statistical Engineering Lab.&lt;br&gt;National Bureau of Standards&lt;br&gt;Washington 25, D.C.</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Statistical Laboratory&lt;br&gt;Department of Mathematics&lt;br&gt;University of California&lt;br&gt;Berkeley 4, Calif.</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Statistical Laboratory&lt;br&gt;University of Washington&lt;br&gt;Seattle 5, Washington</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Professor O. P. Aggarwal&lt;br&gt;Statistical Laboratory&lt;br&gt;Purdue University&lt;br&gt;Lafayette, Indiana</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dr. Stephen G. Allen&lt;br&gt;244 Laurel&lt;br&gt;San Francisco, Calif.</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Professor T. W. Anderson&lt;br&gt;Dept. of Mathematical Statistics&lt;br&gt;Columbia University&lt;br&gt;New York 27, N.Y.</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Professor Fred C. Andrew&lt;br&gt;Mathematics Department&lt;br&gt;University of Nebraska&lt;br&gt;Lincoln 8, Nebraska</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Professor David Blackwell&lt;br&gt;Statistics Department&lt;br&gt;University of California&lt;br&gt;Berkeley 4, California</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Professor Russell Bradt  
Mathematics Department  
University of Kansas  
Lawrence, Kansas

Dr. Charles Boll  
Hughes Aircraft Co.  
Bldg. 12 Room 2343  
Culver City, Calif.

Professor W. G. Cochran  
Biostatistics Dept.  
The Johns Hopkins University  
Baltimore 5, Md.

Professor Lee Cronbach  
Bureau of Education Research  
1007 S. Wright  
Champaign, Illinois

Dr. Joseph Daly  
Bureau of the Census  
Washington 25, D. C.

Dr. Francis Dresch  
Stanford Research Inst.  
1915 University Ave.  
Palo Alto, Calif.

Mr. Murray A. Geisler  
Logistics Section  
The RAND Corp.  
1700 Main Street  
Santa Monica, Calif.

Professor E. J. Gumbel  
Industrial Engineering Dept.  
409 Engineering Bldg.  
Columbia University  
New York 27, N. Y.

Professor Stanley Isaacson  
716 Eastmont Dr., Apt. 2C  
Greensburg, Pa.

Professor S. Karlin  
Dept. of Mathematics  
Calif. Inst. of Technology  
Pasadena, Calif.

Professor Leo Katz  
Mathematics Department  
Michigan State College  
East Lansing, Michigan

Dr. Eugene Lukacs  
Dept. of Mathematics  
Catholic University  
Washington 17, D. C.

Mr. Monroe Norden  
Research Division  
College of Engineering  
New York University  
New York 53, N. Y.

Professor Stanley Reiter  
Dept. of Economics  
Purdue University  
Lafayette, Indiana

Professor Herman Rubin  
Dept. of Mathematics  
University of Oregon  
Eugene, Oregon

Professor Seymour Sherman  
Moore School of Electrical Engineering  
University of Pennsylvania  
Philadelphia 4, Pa.

Dr. Milton Sobel  
Bell Telephone Laboratories  
555 Union Blvd.  
Allentown, Pa.

Professor Herbert Solomon  
Teachers' College  
Columbia University  
New York 27, N. Y.

Dr. Dan Teichroew  
Numerical Analysis Research  
405 Hilgard Ave.  
Los Angeles 24, Calif.

Professor Donald Truax  
Mathematics Dept.  
California Institute of Technology  
Pasadena, Calif.
Professor W. Allen Wallis  
Committee on Statistics  
University of Chicago  
Chicago 37, Illinois  
1

Professor Irving Weiss  
Department of Mathematics  
Lehigh University  
Bethlehem, Pa.  
1

Dr. John D. Wilkes  
Office of Naval Research  
Code 200  
Washington 25, D. C.  
1

Professor S. S. Wilks  
Room 110, Fine Hall  
Box 708  
Princeton, N. J.  
1

Professor J. Wolfowitz  
Mathematics Department  
Cornell University  
Ithaca, N. Y.  
1

Professor M. A. Woodbury  
Logistics Project  
George Washington University  
Washington, D. C.  
1

Additional copies for project  
leader and assistants and  
reserve for future requirements  50