ON CHI-SQUARED GOODNESS OF FIT TESTS FOR SAMPLING FROM MORE THAN ONE POPULATION WITH POSSIBLY CENSORED DATA

BY

KATHLEEN LAMBORN

TECHNICAL REPORT NO. 21
JUNE 15, 1969

PREPARED UNDER THE AUSPICES OF
PUBLIC HEALTH SERVICE GRANT USPHS-2TI GM 25-11

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
ON CHI-SQUARED GOODNESS OF FIT TESTS FOR SAMPLING FROM MORE THAN ONE POPULATION WITH POSSIBLY CENSORED DATA

By
Kathleen Lamborn

TECHNICAL REPORT NO. 21
June 15, 1969

PREPARED UNDER THE AUSPICES OF
PUBLIC HEALTH SERVICE GRANT USPHS 2TI GM 25-11

Reproduction in Whole or in Part is Permitted for any Purpose of the United States Government

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
I. Introduction

For testing "goodness of fit" the classical test is the $X^2$ goodness of fit test. The distribution theory was originally developed for the multinomial distribution. In the case of a sample from a continuous distribution, the above test can be applied by choosing fixed intervals $(-\infty, z_1), \ldots, (z_{r-1}, \infty)$. A multinomial random variable is then defined with $p_i = F(z_i) - F(z_{i-1})$, the probability that the original random variable takes on a value in the $i^{th}$ interval. Suppose it is necessary to estimate a $q$ dimensional vector of parameters $\theta$. If $\hat{\theta}$ is the maximum likelihood estimate based on the sample, $p_i(\hat{\theta})$ is the probability of an observation falling in the $i^{th}$ interval given that $\hat{\theta}$ is the true value of the parameter, $n_i$ is the number observed to fall in the $i^{th}$ interval, and $N$ is the total sample size, then Chernoff and Lehmann [1] showed that the statistic

$$X^2 = \sum_{i=1}^{r} \frac{(n_i - Np_i(\hat{\theta}))^2}{Np_i(\hat{\theta})}$$

is distributed asymptotically as

$$\sum_{i=1}^{r} a_i x_i^2$$

where the $x_i's$ are independent $N(0, 1)$. The $a_i's$ are characteristic roots of a matrix. The $a_i's$ lie between zero and one. One equals zero and $r - q - 1$ of them equal one.
Another approach to testing goodness of fit for continuous distributions is to select the endpoints of the intervals, \( g_i(\hat{\theta}) \), so that

\[
p_i = \int_{g_{i-1}(\hat{\theta})}^{g_i(\hat{\theta})} p(t|\hat{\theta}) dt
\]

has a predetermined value. This situation has been considered by A. R. Roy [2] and G. S. Watson [3, 4]. It develops that the asymptotic distribution is again of the form in (2), the difference being in the matrix which determines the \( a_i \) values. However, these values again remain between zero and one inclusive, with one value equal to zero. Roy gives conditions under which more can be said about the values.

In considering an exponential survival study where the expected survival of the \( j^{th} \) individual, \( \lambda_j^{-1} \), was assumed to be related to some concomitant variable, e.g. white blood cell count, the need arose for a goodness of fit test for the case where the sample is from a finite number, \( K \), of different populations. (See Zippin and Lamborn [5]). It was also desired to allow for censored data. It will be shown below that the method of proof employed by Roy can be extended to find the asymptotic distribution of a statistic similar to that in (1) which will handle these situations. It turns out that the asymptotic distribution is again of the form (2) with the \( a_i \) values lying between zero and one inclusive. Additional conditions allow the values to be more completely specified. Further, if the data are uncensored, one value equals zero.
II. Structure of the Test

It is assumed that the sample is drawn from a finite set of populations $Q_k$, $k = 1, \ldots, K$. No specific distribution is required for the probability that a given observation will be from $Q_k$, but it is assumed that once the observation is taken the population from which it came can be determined. The random variable to be observed, $t$, is assumed to have a distribution depending on $\theta$, a q dimensional vector, and on the population from which it came. Let $p_j(t|\theta)$ be the density function associated with the $j$th observation when $\theta$ is the true value of the parameter and the population to which the $j$th observation belongs is known. Thus for fixed $\theta$, $p_j(t|\theta)$ is one of $K$ density functions depending on the population from which the $j$th sample was drawn.

For example, in the exponential survival study mentioned earlier $p_j = \lambda_j e^{-\lambda_j t}$. $\lambda_j$ is a function of $\theta$ and $x_j$, the concomitant variable. In one case considered $\lambda_j^{-1} = a + bx_j$. The number of populations, $K$, is the number of distinct values which $x_j$ can achieve, and observations with the same $x$ value are said to be from the same population. For fixed $\theta$, if $x_j = x_k$, i.e. the two samples come from the same population, then $\lambda_j = \lambda_k$ so that $p_j(t|\theta) = p_k(t|\theta)$.

Divide the range of $t$ into $r$ intervals. The endpoints of the intervals, $g_{i,j}^j(\theta)$, vary depending on the population to which the sample belongs so that

$$\int_{g_{i,j}^j(\theta)}^{g_{i-1,j}^j(\theta)} p_j(t|\theta) dt$$
is constant over \( j \). (In the exponential survival study, for instance, the endpoints were chosen so that the probability of death during each of four intervals was .25 for each patient.)

Let \( m_{ik} \) = number of events from population \( Q_k \) observed to fall in the \( i \)th interval. Define \( m_i = \sum_{k=1}^{K} m_{ik} \). Then the statistic is

\[
R^* = \frac{r}{\tilde{\Sigma}_i} \frac{(m_i - \frac{1}{r} \tilde{\Sigma}_i \tilde{p}_{ij}^*(\tilde{\theta}))^2}{\tilde{\Sigma}_j \tilde{p}_{ij}^*(\tilde{\theta})}.
\]

\( \tilde{\theta} \) is the maximum likelihood estimate of \( \theta \) or some asymptotically equivalent estimate.

\( \tilde{p}_{ij}^*(\theta) \) = probability the \( j \)th event is observed to occur in the \( i \)th interval when \( \theta \) is the true value of the parameter. If the data are censored, this is the probability of the event occurring in the \( i \)th interval and before the time of withdrawal, which is known.

III. Asymptotic Distribution of \( R^*_r \)

Assumptions:

1) \( \frac{\partial}{\partial \theta} p_j(t|\theta) \) is uniformly bounded by an integrable function.

2) \( g_{ij}(\theta) \) is continuous in \( \theta \) and its partial derivatives exist.

3) \( p_j(t|\theta) \) satisfies the usual regularity conditions so that the maximum likelihood estimate \( \hat{\theta} \) satisfies

\[
\hat{\theta} = \theta + \frac{1}{n} \sum_j r_j + o_p(1/\sqrt{n}) \quad \text{where}
\]
\( f_j = J^{-1} \frac{\partial}{\partial \theta} \ln p_j(t|\theta) \) if \( t_j \) is observed

\[ = J^{-1} \frac{\partial}{\partial \theta} \ln \int_{T_j}^{\infty} p_j(t|\theta)dt \] if have withdrawal at \( T_j \)

and \( nJ \) is the information matrix.

iv) The sample is drawn at random from the populations \( Q_1, \ldots, Q_K \) with probability \( q_1, \ldots, q_K \) respectively.

**Theorem.** Under the assumptions above \( R^* \) is distributed asymptotically as

\[ Q^*_R = \frac{E}{\sum a_i x_i^2} \]

The \( x_i \)'s are independent \( N(0, 1) \) random variables. The \( a_i \)'s are the characteristic roots of \( \Sigma^*_R \).

In order to define \( P^* \) and \( \Sigma^* \) it is necessary to introduce some notation.

Let \( p^{*j}_{ij}(\theta) = p^{*j}_{ij} \) where \( \theta \) is the true value of the vector of parameters.

\[
P^* = \frac{1}{n} \begin{bmatrix} \Sigma^*_p & 0 & \ldots & 0 \\
J_{1j} & 0 & \ldots & 0 \\
& \vdots & \ddots & \vdots \\
& \vdots & & \ddots \\
0 & \ldots & 0 & \Sigma^*_{j_{rr}}
\end{bmatrix}
\]

\[
p^*p^* = \frac{1}{n} \begin{bmatrix} \Sigma^*_p & \Sigma^*_p & \ldots & \Sigma^*_p \\
J_{1j} & J_{1j}^2 & \ldots & J_{1j}r_{jr} \\
& \vdots & \ddots & \vdots \\
& \vdots & & \ddots \\
0 & \ldots & 0 & \Sigma^*_{j_{rr}}
\end{bmatrix}
\]
U is a $q \times r$ matrix with the elements of the $i^{th}$ column equal to

$$\sum_j \int \frac{g_{i,j}(\theta)}{g_{i-1,j}(\theta)} \frac{\partial}{\partial \theta} \ell_j(t|\theta) dt \text{ where}$$

$$
\ell_j(t|\theta) = p_j(t|\theta) \quad t < T_j \\
= 0 \quad \text{otherwise} .
$$

Then

$$\Sigma^* = P^* - P^*P^* - \frac{1}{n^2} U'J^{-1}U .$$

The major part of the proof of this result consists of rewriting

$$(m_1 - \sum_j p_j^* (\hat{\theta}), \ldots, m_r - \sum_j p_j^* (\hat{\theta}))$$

as the sum of a sequence of $r$ dimensional random variables plus terms which when divided by $\sqrt{n}$ go to zero as $n \to \infty$.

By extending the procedure in [2] one can prove:

**Lemma.** Let

$$b_i(t_j) = 1 \text{ if } t_j \in (g_{i-1,j}(\theta), g_{i,j}(\theta)) \text{ and is observed}$$

$$= 0 \text{ otherwise} ,$$

then

$$m_i - \sum_j b_i(t_j) = \sum_j \left( L_j(g_{i,j}(\hat{\theta})|\theta) - L_j(g_{i,j}(\theta)|\theta) \right)$$

$$- \sum_j [L_j(g_{i-1,j}(\hat{\theta})|\theta) - L_j(g_{i-1,j}(\theta)|\theta)] + o_p(\sqrt{n})$$

where

$$L_j(t|\theta) = \int_0^t \ell_j(x|\theta) dx .$$
Proof. Throughout the proof $L_j(t|\theta)$ will be written as $L_j(t)$ for simplicity. Observation $j$ makes a nonzero contribution to $m_j - \sum_{j} b_i(t_j)$ only if it is observed to occur between $g_{ij}(\hat{\theta})$ and $g_{ij}(\theta)$ or $g_{i-1,j}(\hat{\theta})$ and $g_{i-1,j}(\theta)$. An event falling between $g_{ij}(\hat{\theta})$ and $g_{ij}(\theta)$ makes a positive contribution if $g_{ij}(\hat{\theta}) > g_{ij}(\theta)$ and a negative one if the inequality is reversed. Exactly the opposite is true for an event between $g_{i-1,j}(\hat{\theta})$ and $g_{i-1,j}(\theta)$. The direction of the inequalities will be the same for all observations taken from a given population. Write $j \in Q_k$ if the $j^{th}$ observation comes from population $Q_k$. Then it is sufficient to prove that the number of events from population $Q_k$ observed to occur between $g_{ij}(\hat{\theta})$ and $g_{ij}(\theta)$ is

$$ |\sum_{j \in Q_k} [L_j(g_{ij}(\hat{\theta})) - L_j(g_{ij}(\theta))]| + o_p(\sqrt{n}) .$$

(3)

Since the sample is taken at random from the total population, $F(n_k \to \infty \text{ as } n \to \infty) = 1$, where $n_k$ is the number of observations taken from population $Q_k$. Therefore the proof follows if (3) can be proved with $o_p(\sqrt{n})$ replaced by $o_p(\sqrt{n_k})$.

Let

$$Y_{ij}(t) = L_j(t) - L_j(g_{ij}(\theta)) \quad t < T_j$$

$$= 0 \quad \text{otherwise}$$

$Y_{ij}(t)$ has density $f_Y = 1$ on $(-L_j[g_{ij}(\theta)], L_j(T_j) - L_j[g_{ij}(\theta)])$ with mass $1 - L_j(T_j)$ at 0.
\[ t_j < T_j \] and occurring between \( g_{ij}(\hat{\theta}) \) and \( g_{ij}(\theta) \) is equivalent to \( Y_{ij}(t_j) \in (0, L_j(g_{ij}(\hat{\theta})) - L_j(g_{ij}(\theta))) \). Define

\[
Z_{ij}(a) = \begin{cases} 1 & Y_{ij}(t_j) \in (0, a) \\ 0 & \text{otherwise}, \end{cases}
\]

where \( a \) may be either positive or negative.

\[
|L_j(g_{ij}(\hat{\theta})) - L_j(g_{ij}(\theta))| \text{ is bounded above by } |F_j(g_{ij}(\hat{\theta})) - F_j(g_{ij}(\theta))| = o_p(1/\sqrt{n}) \text{ where } F_j \text{ is the distribution of the } j^\text{th} \text{ observation given from which population it comes}. \]

There are only a finite number of populations. Therefore there exists a \( k \) such that

\[
P(\sup_j |L_j(g_{ij}(\hat{\theta})) - L_j(g_{ij}(\theta))| > k/\sqrt{n}) < \epsilon/2 \text{ for all } n. \tag{4}
\]

Let \( a_j = L_j(g_{ij}(\hat{\theta})) - L_j(g_{ij}(\theta)) \). If it can be shown that

\[
\left| \sum_{j \in Q_k^n} Z_{ij}(a_j) - |a_j| \right| = o_p(\sqrt{n_k}) \text{ uniformly for } |a_j| < k/\sqrt{n} \tag{5}
\]

then using (4) the lemma will follow.

For simplicity, consider now a sample from a single population where \( n \) refers to the size of that sample. Choose a sequence of integers \( i_1, \ldots, i_n \) all with the same sign as \( g_{ij}(\hat{\theta}) - g_{ij}(\theta) \). (Assume they are positive. The proof follows similarly if they are negative.) Fix \( \eta > 0 \), and \( \delta \) a positive integer. Let \( i_j \) be the greatest integer less than or equal to \( \delta \) such that \( i_j \eta/2\sqrt{n} \) is within the range of \( Y_{ij} \).

Using Chebychev's inequality
\[
P\left( \frac{1}{n} \left| \sum_{j} (Z_{i,j} \eta/2\sqrt{n}) - i_{j} \eta/2\sqrt{n} \right| > \eta/2\sqrt{n} \right) \leq \frac{1}{n^2} \sum_{j} \left( \frac{2k/\eta + 1}{\eta^2/4n} \right) \leq 2b/\eta \sqrt{n}
\]

\[
P\left( \sup_{0 \leq \delta \leq 2k/\eta + 1} \frac{1}{n} \left| \sum_{j=1}^{n} (Z_{i,j} \eta/2\sqrt{n}) - i_{j} \eta/2\sqrt{n} \right| > \eta/2\sqrt{n} \right) \leq \sum_{\delta=1}^{[2k/\eta + 1]} 2b/\eta \sqrt{n} \leq (2/\eta \sqrt{n}) (2k/\eta + 1)(2k/\eta + 2) \to 0 \text{ as } n \to \infty.
\]

Let \( c_{j} \) be such that
\[
i_{j} \eta/2\sqrt{n} \leq c_{j} \leq (i_{j} + 1) \eta/2\sqrt{n}
\]
then
\[
\sum_{j} (Z_{i,j} \eta/2\sqrt{n}) - (i_{j} + 1) \eta/2\sqrt{n} \leq \sum_{j} (Z_{j}(c_{j}) - c_{j}) \leq \sum_{j} (Z_{j}[(i_{j} + 1) \eta/2\sqrt{n}] - i_{j} \eta/2\sqrt{n})
\]
from which it follows that
\[
\sup_{0 \leq \delta \leq 2k/\eta + 1} \left| \frac{1}{n} \sum_{j} (Z_{j}(i_{j} \eta/2\sqrt{n}) - i_{j} \eta/2\sqrt{n}) \right| \leq \eta/2\sqrt{n} \text{ implies } \\
\sup_{0 \leq c_{j} \leq k/\sqrt{n}} \left| \frac{1}{n} \sum_{j} (Z(c_{j}) - c_{j}) \right| \leq \eta/\sqrt{n}.
\]
This follows since \( c_{j} \leq k/\sqrt{n} \) implies \( i_{j} \leq 2k/\eta \).

\[
P\left( \sup_{0 \leq \delta \leq 2k/\eta + 1} \left| \frac{1}{n} \sum_{j} (Z_{j}(i_{j} \eta/2\sqrt{n}) - i_{j} \eta/2\sqrt{n}) \right| \leq \eta/2\sqrt{n} \right) \leq P\left( \sup_{0 \leq c_{j} \leq k/\sqrt{n}} \left| \frac{1}{n} \sum_{j} (Z(c_{j}) - c_{j}) \right| \leq \eta/\sqrt{n} \right).
\]

Therefore \( P\left( \sup_{0 \leq c_{j} \leq k/\sqrt{n}} \left| \frac{1}{n} \sum_{j} (Z(c_{j}) - c_{j}) \right| > \eta/\sqrt{n} \right) \to 0 \text{ as } n \to \infty \) as long as the \( c_{j} \)'s satisfy (6) for some sequence of \( i_{j} \)'s.

Let \( c_{j} = a_{j} = L_{j}(g_{i,j}(\theta)) - L_{j}(g_{i,j}(\theta)). \) If \( 0 \leq a_{j} \leq k/\sqrt{n} \), then \( a_{j} \) satisfies the above restrictions. Following a similar argument if \( a_{j} < 0 \), then, completes the proof of (5) and the lemma follows.
Proof of Theorem.

Using Taylor's expansion and the assumption that $\frac{\partial}{\partial \theta} p_j(t|\theta)$ is uniformly bounded by an integrable function, write

$$L_j(g_{i,j}(\hat{\theta})|\theta) - L_j(g_{i,j}(\theta)|\theta) = [L_j(g_{i-1,j}(\hat{\theta})|\theta) - L_j(g_{i-1,j}(\theta)|\theta)] = \left(\hat{\theta} - \theta\right)' v_{ij}(\theta) + o_P(1/\sqrt{n})$$

where $v_{ij}(\theta) = L_j(g_{i,j}(\theta)|\theta) \frac{\partial}{\partial \theta} g_{i,j}(\theta) - L_j(g_{i-1,j}(\theta)|\theta) \frac{\partial}{\partial \theta} g_{i-1,j}(\theta)$.

($\hat{g}_{ij}(\theta) = \frac{\partial}{\partial \theta} g_{ij}(\theta)$.) Thus

$$m_1 - \sum_j b_j(t_j) = \left(\hat{\theta} - \theta\right)' \sum_j v_{ij}(\theta) + o_P(1/\sqrt{n}) .$$

Similarly

$$p_{ij}(\hat{\theta}) - p_{ij}(\theta) = L_j(g_{i,j}(\hat{\theta})|\theta) - L_j(g_{i,j}(\theta)|\theta) - [L_j(g_{i-1,j}(\hat{\theta})|\theta) - L_j(g_{i-1,j}(\theta)|\theta)]$$

$$= \left(\hat{\theta} - \theta\right)' (u_{ij}(\theta) + v_{ij}(\theta)) + o_P(1/\sqrt{n})$$

where $u_{ij}(\theta) = \int g_{i-1,j}(\theta) \frac{\partial}{\partial \theta} g_{ij}(\theta) dx$. Substituting these values gives

$$m_1 - \sum_j p_{ij}(\hat{\theta}) = \sum_j (b_j(t_j) - p_{ij}(\theta)) - \left(\hat{\theta} - \theta\right)' \sum_j u_{ij}(\theta) + o_P(1/\sqrt{n}) .$$

Substituting for $\hat{\theta} - \theta$ and using the Mann-Wald result:

$$\frac{1}{\sqrt{n}}(m_1 - \sum_j p_{ij}(\hat{\theta}), \ldots, m_r - \sum_j p_{ij}(\hat{\theta}))$$

is distributed in the limit as

$$\frac{1}{\sqrt{n}} \sum_j (h_{ij}, \ldots, h_{ij})$$

where $h_{ij} = b_i(t_j) - p_{ij}(\theta) - \frac{1}{n} \sum_j u_{ij}(\theta)$. 

10
\[ \frac{1}{\sqrt{n}} \sum_{j} \Sigma[h_{1j}, \ldots, h_{r_{j}}] \] has mean 0 and covariance matrix

\[ \Sigma^* = P^* - p^* p'^* - \frac{1}{n^2} U' J^{-1} U \]

\( \Sigma^* \) depends on the proportion of the sample from each population and the times of withdrawal. However, since \( T_j > 0 \), and the observations are randomly chosen from a finite number of populations, all the components of \( \Sigma^* \) go to constants in the limit, so that \( \Sigma^* \) is a constant in the limiting case.

Then asymptotically

\[ \frac{1}{\sqrt{n}} \sum_{j} \Sigma[h_{1j}, \ldots, h_{r_{j}}] \sim N(0, \Sigma^*) \]

\[ \left( \frac{m_1 \Sigma p^*_{1j} (\hat{\theta})}{\sqrt{\Sigma p^*_{1j} (\hat{\theta})}}, \ldots, \frac{m_r \Sigma p^*_{r_{j}} (\hat{\theta})}{\sqrt{\Sigma p^*_{r_{j}} (\hat{\theta})}} \right) \sim N(0, P^* - \frac{1}{n^2} \Sigma^* P^* - \frac{1}{n^2}) \].

(The assumption is made that \( \Sigma p^*_{1j} (\hat{\theta}) > 0 \) for all \( i \).) From a standard large sample theory result it now follows that \( R_p^* \) has the asymptotic distribution stated in the theorem.

IV. Values of the Characteristic Roots

Lemma. i) All roots of \( P^* - \frac{1}{n^2} \Sigma^* P^* - \frac{1}{n^2} \) lie between zero and one inclusive.

ii) Let \( C = p^* p'^* + \frac{1}{n^2} U' J^{-1} U \). If the rank of \( C \) is \( s \), then there are exactly \( r - s \) characteristic roots equal to one.

iii) In the case of uncensored data, at least one root is equal to 0.
Proof. \( P^* \Sigma^* P^* \) has the same characteristic roots as \( P^* \Sigma^* P^* \).

\( P^* \Sigma^* = I - P^* \Sigma^* \). ii) follows immediately from this. i) follows by noting that both \( P^* \Sigma^* P^* \) and \( P^* \Sigma^* P^* C \) are positive semi-definite.

In the case of uncensored data the rows and columns of \( P^* \Sigma^* \) add to 1, so that at least one characteristic value is 0, proving iii.

In general, the values of the \( a_i \)'s will depend on the parameter values. However, in some fairly general cases where the \( \theta_i \)'s are part of scale or translation parameters only, the asymptotic distributions will be independent of \( \theta \) in the uncensored case.

As an example, suppose \( p_j(t|\theta) = l/b \frac{t+ax_j}{b} \), where \( p \) satisfies the initial regularity conditions, and \( g_{ij}(\theta) = bk_i - ax_j \).

\[
p_{ij}(\theta) = \int_{bk_{i-1} - ax_j}^{bk_i - ax_j} \frac{1}{l/b} p\left(\frac{t+ax_j}{b}\right) dt
\]

Let \( u = \frac{t+ax_j}{b} \). Then \( p_{ij}(\theta) = \int_{k_{i-1}}^{k_i} p(u) du \) and the integral is independent of \( \theta = (a, b) \).

It remains to show that \( U'J^{-1}U \) is independent of \( \theta \).

\[
U_{ij} = \left( \int_{bk_{i-1} - ax_j}^{bk_i - ax_j} \frac{1}{l/b} \frac{\partial p}{\partial a} \left(\frac{t+ax_j}{b}\right) dt \right)
\]

\[
\left( \int_{bk_{i-1} - ax_j}^{bk_i - ax_j} \frac{1}{l/b^2} \left( \frac{t+ax_j}{b} \right) + \frac{1}{b} \frac{\partial p}{\partial b} \right) dt \right)
\]

12
\[
\left( \frac{x_j}{b} \int_{k_{i-1}}^{k_i} \frac{dp}{du} (u) du \right) \\
\left( -\frac{1}{b} \int_{k_{i-1}}^{k_i} \left[ p(u) + u \frac{dp}{du} \right] du \right)
\]

\[
J = \frac{1}{n} \sum_j \mathbb{E} \left( \frac{\partial}{\partial \theta} \frac{\partial}{\partial \theta} \ln p_j(t|\theta) \right)
\]

\[
\mathbb{E} \left( \frac{\partial}{\partial a^2} \ln p_j(t|\theta) \right) = - \int \frac{1}{b} \left( \frac{\partial p}{\partial a} \right)^2 \frac{1}{p(t|\theta)} \, dt = - \frac{x_j^2}{b^2} \int \frac{dp}{du} \frac{1}{p(u)} \, du
\]

\[
\mathbb{E} \left( \frac{\partial}{\partial b^2} \ln p_j(t|\theta) \right) = - \frac{1}{b^2} \left( 1 + \int \left[ 2u \frac{dp}{du} + u^2 \left( \frac{dp}{du} \right)^2 \frac{1}{p(u)} \right] du \right)
\]

\[
\mathbb{E} \left( \frac{\partial}{\partial a \partial b} \ln p_j(t|\theta) \right) = \frac{x_j}{b^2} \int \left( \frac{dp}{du} \right)^2 u \cdot \frac{1}{p(u)} \, du
\]

Therefore \( U'J^{-1}U \) is independent of \( a \) and \( b \).

There are other instances when the asymptotic distribution is independent of the parameters. Because there are so few requirements on \( p_j(t|\theta) \) in the original theorem, a statement of the exact conditions needed cannot be made.

The author wishes to thank Professor Rupert Miller, Jr. for his many helpful suggestions during the preparation of this paper.
REFERENCES


