Sequential Estimation of a Restricted Mean Parameter of an Exponential Family *

by

George P. McCabe, Jr.

Purdue University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 240

August 1970

* Supported in part by NIH Fellowship 1-F01-GM42814-01 at Columbia University.
Sequential Estimation of a Restricted Mean Parameter of an Exponential Family

by

George P. McCabe, Jr.

Purdue University

1. **Introduction and summary.** The estimation of restricted parameters by fixed sample size rules has been considered by Hammersley [1]. A sequential solution to the problem of estimating the mean of a normal distribution when it is some unknown integer and a general method for solving problems of this sort are presented by Robbins in [4]. Based on the work of Robbins, a sequential procedure for estimating the parameter of a Poisson distribution when it is known to be an integer is given in [3].

The results obtained herein represent a generalization of the work dealing with the normal and Poisson cases. A class of sequential procedures is proposed and bounds on the error probabilities are obtained. The expected sample sizes are investigated and a weak form of optimality is demonstrated under certain conditions.

2. **Statement of the problem.** Let the distribution \( F(X) \) of a random variable \( X \) be a member of some exponential family, i.e.,

\[
dF(X) = f_\theta(X) d\mu(X) = \exp(\theta T(X) - c(\theta)) d\mu(X) : \theta \in \Omega, \]

where \( \mu \) is some \( \sigma \)-finite measure on the real line. It is assumed that \( \Omega \) is countable and can be ordered so that \( \theta_i < \theta_{i+1} \) for all \( i \). Let \( X_1, X_2, \ldots \) be a sequence of iid random variables distributed as \( X \).

From a finite number of observations on the sequence, it is desired to estimate the true value of \( \theta \) with a uniformly small probability of error.

It is well known that \( \theta \), the set of \( \theta \) for which \( \int \exp(\theta T(X)) d\mu(X) < \infty \) is

* Supported in part by NIH Fellowship 1-F01-GM42814-01 at Columbia University.
convex. It is assumed that $\theta_i$ is in the interior of $S$ for all $i$. For
$\theta \in \text{int}(S)$, the moment generating function exists and the following properties can be easily deduced [5]:
(a) 
$$E_{\theta}T(X)=dc(\theta)/d\theta=c'(\theta) < \infty,$$
(b) 
$$0 \leq \text{var}_{\theta}T(X)=c''(\theta) < \infty, \text{ and}$$
(c) if for some $\theta \in \text{int}(S)$, $c''(\theta) > 0$, then $c''(\theta) > 0$ for all $\theta \in \text{int}(S)$.
In view of (c), the degenerate case where $c''(\theta) = 0$ for some $\theta \in \text{int}(S)$ is excluded. Then, $c'(\theta)$ is a strictly increasing function of $\theta$ for $\theta \in \text{int}(S)$.

For convenience, the symbols $E_i$ and $P_i$ will denote expectation and probability respectively under the condition that $\theta_i$ is the true value of the parameter.

For every $\theta_i \in \Omega$, let
$$g(\theta_i)=c'(\theta_i), g_+(\theta_i)=g_+(\theta_{i-1})=(c(\theta_i)-c(\theta_{i-1}))/\theta_i-\theta_{i-1}.$$

3. A sequential solution. Let $f^n_\theta(X_1,\ldots,X_n)=\prod_{j=1}^n f_\theta(X_j)$. Let $\alpha>1$ be be fixed and define a stopping rule by
$$N=\inf\{n \geq 1: \min\left(\frac{f^n_\theta}{f^n_{\theta_{i-1}}},\frac{f^n_\theta}{f^n_{\theta_{i+1}}}\right) > \alpha \text{ for some } \theta_i \in \Omega \}.$$

It will be shown that the $\theta_i$ in the definition of $N$ is unique. Accordingly, the terminal decision rule will be to estimate that $\theta$ is $\theta_i$. The following lemmas will be needed to restate the form of the procedure and to show that $P_i(N < \infty)=1$ for all $i$.

**Lemma 3.1.** Let $\theta_1, \theta_2, \theta_3 \in \Omega$ with $\theta_1 < \theta_2 < \theta_3$. Then, $\min\left(\frac{f^n_{\theta_2}}{f^n_{\theta_1}},\frac{f^n_{\theta_2}}{f^n_{\theta_3}}\right) > \alpha$
if and only if
$$g_-(\theta_2)\times(\log \alpha)/n(\theta_2-\theta_1) \leq \overline{T}_n \leq g_+(\theta_2)-(\log \alpha)/n(\theta_3-\theta_2), \text{ where}$$
$$\overline{T}_n=(T(X_1)+\ldots+T(X_n))/n.$$
Proof. Suppose that $f_{\theta_2}^n / f_{\theta_1}^n = \exp(n(\theta_2 - \theta_1)\bar{T}_n - n(c(\theta_2) - c(\theta_1))) \geq \alpha$. Taking logs and rearranging gives $\bar{T}_n \geq (c(\theta_2) - c(\theta_1)) / (\theta_2 - \theta_1) + (\log \alpha) / n(\theta_2 - \theta_1) = g_-(\theta_2) + (\log \alpha) / n(\theta_2 - \theta_1)$. Since all steps are reversible, it follows that $f_{\theta_2}^n / f_{\theta_1}^n \geq \alpha$ if and only if the left-hand inequality of (3.1) is true. Similarly, $f_{\theta_2}^n / f_{\theta_3}^n \geq \alpha$ if and only if the right-hand inequality of (3.1) is true. Combining these two facts gives the desired result. \(\square\)

From Lemma 3.1, it follows that the sequential procedure can be rewritten as follows: stop at $\bar{N} = n$ as soon as

(3.2) $g_-(\theta_i) + (\log \alpha) / n(\theta_i - \theta_{i-1}) < \bar{T}_n \leq g_+(\theta_i) - (\log \alpha) / n(\theta_{i+1} - \theta_i)$

is true for some $i$ and guess that $\theta$ is this $\theta_i$.

Lemma 3.2. Let $\theta_1, \theta_2, \theta_3 \in \Theta$ with $\theta_1 < \theta_2 < \theta_3$. Then

(3.3) $g_-(\theta_2) < g(\theta_2) < g_+(\theta_2)$.

Proof. $c(\theta_2) - c(\theta_1) = \int_{\theta_1}^{\theta_2} g(\theta) d\theta < (\theta_2 - \theta_1) g(\theta_2)$ since $g(\theta) = c'(\theta)$ is a strictly increasing function of $\theta$. Dividing by $(\theta_2 - \theta_1)$ and noting the definition of $g_-(\theta_2)$ gives the left-hand inequality. The right-hand part follows in a similar manner upon noting that $c(\theta_3) - c(\theta_2) = \int_{\theta_2}^{\theta_3} g(\theta) d\theta > (\theta_3 - \theta_2) g(\theta_2)$. \(\square\)

Lemma 3.3. On the set $\{N = n\}$, there is a unique $i$ such that (3.2) is true.

Proof. From (3.2) it follows that on the set $\{N = n\}$, $\bar{T}_n$ satisfies $g_-(\theta_i) < \bar{T}_n < g_+(\theta_i)$ for some $i$. Now, since $g_-(\theta_1) < g_+(\theta_1) = g_-(\theta_{i+1})$ it is not possible for $\bar{T}_n$ to be simultaneously in more than one interval of the
form \((g_-(\theta_i), g_+(\theta_i))\).

**Theorem 3.1.** \(P_i(N < \infty) = 1\) for all \(i\).

**Proof.** Let \(i\) be fixed and assume that \(\theta_i\) is the true value of the parameter. Now, \(E_i T(X) = g(\theta_i)\) and \(\text{var}_i T(X) = \sigma_i^2(\theta_i)\) which is finite. Hence, 

\[
\bar{T}_n \rightarrow g(\theta_i) \quad \text{almost surely as} \ n \rightarrow \infty.
\]

Now, note that 

\[
g_+(\theta_i)(\log \alpha) / n(\theta_i - \theta_{i-1}) \rightarrow g_+(\theta_i)
\]

and 

\[
g_-(\theta_i)(\log \alpha) / n(\theta_{i+1} - \theta_i) \rightarrow g_-(\theta_i) \quad \text{as} \ n \rightarrow \infty.
\]

Recalling that \(g_i(\theta_i) < g(\theta_i) < g_i(\theta_i)\) by Lemma 3.2, choose \(\varepsilon\) such that 

\[
0 < \varepsilon < \min(g_i(\theta_i) - g(\theta_i), g(\theta_i) - g_i(\theta_i)).
\]

Let \(n_0\) be such that \(n \geq n_0\) implies that 

\[
g_+(\theta_i)(\log \alpha) / n(\theta_i - \theta_{i-1}) < g(\theta_i) - \varepsilon
\]

and 

\[
g_-(\theta_i)(\log \alpha) / n(\theta_{i+1} - \theta_i) > g(\theta_i) + \varepsilon.
\]

Then for \(n > n_0\), 

\[
P_i(g(\theta_i) - \varepsilon < \bar{T}_n < g(\theta_i) + \varepsilon) \leq P_i(N < n) \text{ by Lemma 3.1 and the definition of } N.
\]

Since \(\varepsilon > 0\) is fixed and 

\[
\bar{T}_n \overset{a.s.}{\rightarrow} g(\theta_i),
\]

taking limits as \(n \rightarrow \infty\) gives 

\[
1 = \lim_{n \rightarrow \infty} P_i(g(\theta_i) - \varepsilon < \bar{T}_n < g(\theta_i) + \varepsilon) \leq \lim_{n \rightarrow \infty} P_i(N < n) = P(N < \infty).
\]

For each possible parameter value \(\theta_i\), there are two quantities of interest concerning the sample size. The first, which will be designated \(m_i\), is the minimum sample size for which a guess of \(\theta_i\) is possible, and the second, which will be designated \(n_i\), is the minimum sample size such that the stopping interval (3.2) for \(\bar{T}_n\) includes the point \(g(\theta_i)\).

Thus, from (3.2),

\[
m_i = \inf\{n \geq 1 : g_-(\theta_i)(\log \alpha) / n(\theta_i - \theta_{i-1}) \leq g_+(\theta_i)(\log \alpha) / n(\theta_{i+1} - \theta_i)\}
\]

and
\[ n_i = \inf\{n \geq 1 : g_-(\theta_i) + (\log n) / n(\theta_i - \theta_{i-1}) \leq g(\theta_i) \leq g_+(\theta_i) - (\log n) / n(\theta_{i+1} - \theta_i)\}. \]

For convenience, \( n_i(m_i) \) will be identified with any real number less than or equal to \( n_i(m_i) \) and greater than \( n_i-1(m_i-1) \).

**Lemma 3.4.**

\[ m_i = (\log n) / (\theta_{i+1} - \theta_i) \leq (g_+(\theta_i) - g_-(\theta_i)) / (\theta_i - \theta_{i-1})(\theta_{i+1} - \theta_i) \]

\[ \geq 4(\log n) / (g_+(\theta_i) - g_-(\theta_i))(\theta_{i+1} - \theta_i) \]

and

\[ n_i = (\log n) / \min((g(\theta_i) - g_-(\theta_i))(\theta_i - \theta_{i-1}), (g_+(\theta_i) - g(\theta_i))(\theta_{i+1} - \theta_i)). \]

**Proof.** (3.4): By definition, \( m_i \) is the minimum \( n \) such that \( g_-(\theta_i) + (\log n) / n(\theta_i - \theta_{i-1}) \leq g_+(\theta_i) - (\log n) / n(\theta_{i+1} - \theta_i) \). Solving for \( n \) gives the first part of (3.4). Now, let \( \theta_i + \theta_{i-1} = a \) and \( \theta_i - \theta_{i-1} = b \). Then,

\[ \theta_{i+1} - \theta_i = a - b \text{ and } (\theta_{i+1} - \theta_{i-1}) / (\theta_i + \theta_{i-1}) = a / (ab - b^2). \]

Note that \( a, b \) and \( a - b \) are all positive. Considering \( a \) to be a fixed positive number and setting the derivative of the above expression equal to zero yields the root \( b = a / 2 \). A check of the second derivative shows that the expression evaluated at this point is minimized. Thus, \( a / (ab - b^2) \geq 4 / a = 4 / (\theta_{i+1} - \theta_{i-1}) \). Using this inequality and the first part of (3.4) gives the second part.

(3.5): The condition \( g_-(\theta_i) + (\log n) / n(\theta_i - \theta_{i-1}) \leq g(\theta_i) \) implies that \( n \geq (\log n) / (g(\theta_i) - g_-(\theta_i))(\theta_i - \theta_{i-1}) \), while \( g(\theta_i) \leq g_+(\theta_i) - (\log n) / n(\theta_{i+1} - \theta_i) \) implies that \( n \geq (\log n) / (g_+(\theta_i) - g(\theta_i))(\theta_{i+1} - \theta_i) \). Combining yields (3.5). \( \square \)

Let \( A_{n,k} = \{N = n, \text{ estimate } \theta = \theta_k\} \) and
Lemma 3.5 On the set $\Lambda_{n,k}$, 
$$f^n_{\theta_j}/f^n_{\theta_k} \leq \exp(-\log \alpha) b(j,k)$$ 
for all $n \geq m_k$ and $j \neq k$.

Proof. Suppose that $j > k$. Now, 
$$\log(f^n_{\theta_j}/f^n_{\theta_k}) = n(c(\theta_k)-c(\theta_j)+\theta_j-\theta_k)T_n$$
and on the set $\Lambda_{n,k}$, 
$$T_n \leq g_-(\theta_{k+1})-(\log \alpha)/n(\theta_{k+1}-\theta_k)$$
by (3.2). Thus, since 
$$\theta_j > \theta_k,$$
(3.6) 
$$\log(f^n_{\theta_j}/f^n_{\theta_k}) \leq n(c(\theta_k)-c(\theta_j)+\theta_j-\theta_k)g_-(\theta_{k+1})$$
$$-\log(\alpha)(\theta_j-\theta_k)/(\theta_{k+1}-\theta_k))$$

Also, from the definition of $g_-(\theta_i)$, it follows that 
$$c(\theta_j)-c(\theta_k) = \sum_{i=k+1}^j (c(\theta_i)-c(\theta_{i-1})) = \sum_{i=k+1}^j (\theta_i-\theta_{i-1})g_-(\theta_i).$$

Also, $(\theta_j-\theta_k)g_-(\theta_{k+1})$ can be rewritten as 
$$\sum_{i=k+1}^j (\theta_i-\theta_{i-1})g_-(\theta_{k+1}).$$
Combining the above with (3.6) gives 
(3.7) 
$$\log(f^n_{\theta_j}/f^n_{\theta_k}) \leq -(\log \alpha)(\theta_j-\theta_k)/(\theta_{k+1}-\theta_k)$$
$$-n \sum_{i=k+2}^j (\theta_i-\theta_{i-1})(g_-(\theta_i)-g_-(\theta_{k+1})).$$

Now from (3.4), $n \geq m_k$ implies that $n \geq 4(\log \alpha)/(g_+(\theta_k)-g_-(\theta_k))(\theta_{k+1}-\theta_{k-1})$.

Thus, since all the terms in the summation of (3.7) are positive,
\[
\log \left( \frac{f^n_{\theta_j}}{f^n_{\theta_k}} \right) \leq -(\log \alpha) \left[ \frac{(\theta_j-\theta_k)}{(\theta_{k+1}-\theta_k)} \right] - 4(\log \alpha) \sum_{i=k+2}^{j} \frac{(\theta_i-\theta_{i-1})}{(\theta_{i+1}-\theta_i)} \left( g_-(\theta_i) - g_-(\theta_{i+1}) \right)/(g_+(\theta_k) - g_-(\theta_k))(\theta_{k+1} - \theta_{k-1}).
\]

Noting that \( g_-(\theta_{k+1}) = g_+(\theta_k) \) gives the desired result. The case where \( j < k \) is treated in a completely similar manner using the fact that on \( A_{n,k} \),

\[
\overline{T}_n \geq g_-(\theta_k) + (\log \alpha)/n(\theta_k - \theta_{k-1}).
\]

Let \( P_j \) denote the probability that an incorrect estimate is given when \( \theta = \theta_j \).

**Theorem 3.2.** \[ P_j \leq \sum_{k+j} \alpha^{-b(j,k)}. \]

**Proof.**

\[
P_j = \sum_{k+j} \sum_{n \geq m_k} \int_{A_{n,k}} \frac{f^n_{\theta_j}}{f^n_{\theta_k}} f_{\theta_k},
\]

where the differential term is omitted. Now,

\[
P_i \leq \sum_{k+j} \alpha^{-b(j,k)} \sum_{n \geq m_k} \int_{A_{n,k}} \frac{f^n_{\theta_j}}{f^n_{\theta_k}} f_{\theta_k}
\]

so,

\[
P_i \leq \sum_{k+j} \alpha^{-b(j,k)} \sum_{n \geq m_k} \int_{A_{n,k}} \frac{f^n_{\theta_j}}{f^n_{\theta_k}} f_{\theta_k}
\]

by the previous lemma. Hence,

\[
P_i \leq \sum_{k+j} \alpha^{-b(j,k)} (1 - P_k) \leq \sum_{k+j} \alpha^{-b(j,k)}.
\]

Let

\[
a(j,k) = \begin{cases} 
(\theta_j - \theta_k)/(\theta_{k+1} - \theta_k) & \text{for } j > k \\
(\theta_{k-1} - \theta_j)/(\theta_k - \theta_{k-1}) & \text{for } j < k 
\end{cases}
\]

since \( a(j,k) \leq b(j,k) \), the following is evident:

**Corollary.** \[ P_j \leq \sum_{k+j} \alpha^{-a(j,k)}. \]
Example 1. \( \theta_k = ka + c \) where \( a \neq 0 \) and \( c \) are arbitrary real numbers.

Without loss of generality, \((X_i\) could be replaced by \(-X_i\)) it is assumed that \( a > 0 \) so that \( \theta_k \) is an increasing function of \( k \) as hypothesized. Now, 
\[
(\theta_j - \theta_k) = (k-j)a, \ \text{so} \ a(j,k) = |j-k|.
\]

Thus,
\[
P_j \leq \sum_{k \neq j} a^{-|j-k|} = 2/(\alpha-1).
\]

Note that this example includes the case of normal variables with mean \( \mu_k = k \) and known variance \( \sigma^2 \).

Example 2. \( \theta_k = a/k + c \) for \( k \geq 1 \), where \( a \neq 0 \) and \( c \) are arbitrary real numbers. As in Example 1, it can be assumed that \( a > 0 \). Now,
\[
(\theta_j - \theta_k) = (k-j)a/kj. \ \text{For} \ k > j,
\]
\[
a(j,k) = (k-j)(k-1)/j \geq k-j. \ \text{Thus,}
\]
\[
(3.8) \sum_{k > j} a^{-a(j,k)} < \sum_{k > j} a^{-k-j} = 1/(\alpha-1).
\]

Similarly, for \( k < j \), \( a(j,k) = (j-k)(k+1)/j \). Now, letting \( j^* = \) the greatest integer less than or equal to \( j/2 \), it follows that
\[
\sum_{k < j} a^{-a(j,k)} < \sum_{k=1}^{j^*} a^{-(k+1)/2} + \sum_{k=j^*+1}^{j-1} a^{-(j-k)/2}
\]
\[
< \sum_{i > 1} (\alpha^{-1/2})^i + \sum_{j > 0} (\alpha^{-1/2})^i = (1+\alpha^{-1/2})/(\alpha^{1/2} - 1)
\]

Combining with (3.8) gives
\[
P_j < (2 + \alpha^{1/2} + \alpha^{-1/2})/(\alpha - 1) \sim \alpha^{-1/2} \text{ as } \alpha \to \infty.
\]
Although this bound is perhaps a bit crude, it is nonetheless, a uniform bound on the error probability which goes to zero as $\alpha \to \infty$. One might expect that a better bound could be obtained which would be asymptotic to $2 \alpha^{-1}$ as $\alpha \to \infty$. On the other hand, it can be shown that $P_j \leq 2 \alpha^{-1}$ as $\alpha \to \infty$ for each $j$. Clearly, from (3.9), the sum of the terms for $k > j$ is asymptotically less than or equal to $\alpha^{-1}$. Also,

$$
\alpha \sum_{k>j} \alpha^{-a(j,k)} = \sum_{k<j} \alpha^{-((j-k)(k+1)-j)/j} \sim 1 \text{ as } \alpha \to \infty,
$$

since there are only a finite number of terms and the exponent $((j-k)(k+1)-j)/j$ equals zero for $k=j-1$ and is positive for $k < j-1$. Thus $P_j \leq 2\alpha^{-1}$ as $\alpha \to \infty$ for every $j$.

**Theorem 3.3.** If $f_{\theta_{i-1}}(X)/f_{\theta_i}(X) \leq f_{\theta_{k-1}}(X)/f_{\theta_k}(X)$ for all $i \leq k$ and all $X$, then

$$(3.9) \quad P_j \leq 2/(\alpha-1) \text{ for all } j.$$

**Proof.** It follows immediately from the hypothesis that

$$
f^n_{\theta_{i-1}}/f^n_{\theta_i} \leq f^n_{\theta_{k-1}}/f^n_{\theta_k} \text{ for all } k \leq i \text{ and all } (X_1, \ldots, X_n).$$

Now,

$$P_j = \sum P_j \text{ (guess } \theta=\theta_k).$$

Suppose $k > j$. Then,

$$P_j \text{ (guess } \theta=\theta_k) = \sum_{n>m_k} \int_{A_{n,k}} f^n_{\theta_j}$$

$$= \sum_{n>m_k} \int_{A_{n,k}} (f^n_{\theta_j}/f^n_{\theta_j+1}) \cdots (f^n_{\theta_{k-1}}/f^n_{\theta_k}) f^n_{\theta_k}.$$
Since \( f_{\theta_{k-1}}^{n} / f_{\theta_{k}}^{n} \leq \alpha^{-1} \) on \( A_{n,k} \) and \( f_{\theta_{i-1}}^{n} / f_{\theta_{i}}^{n} \leq f_{\theta_{k-1}}^{n} / f_{\theta_{k}}^{n} \), it follows that

\[
P_j (\text{guess } \theta = \theta_k) \leq \alpha^{-(k-j)} \sum_{n > m_k} \int_{A_{n,k}} f_{\theta_k}^{n} = \alpha^{-(k-j)} (1 - p_k) \leq \alpha^{-(k-j)}.
\]

In an entirely analogous fashion, it can be shown that

\[
P_j (\text{guess } \theta = \theta_k) \leq \alpha^{-(j-k)} \text{ for } k < j. \quad \text{Therefore,}
\]

\[
P_j \leq \sum_{k \neq j} \alpha^{-|j-k|} = 2/(\alpha - 1).
\]

**Example 3.** The \( X_i \) are Poisson with mean \( \lambda_j = j \) for \( j \geq 1 \). In this case, \( \theta_j = \log(j) \) and the hypothesis of Theorem 3.3 is satisfied. Hence,

\[
P_j \leq 2/(\alpha - 1) \text{ for all } j \geq 1.
\]

**4. Asymptotic sample size.** Recall that the quantity \( n_i \) given by (3.5) is the minimum sample size such that the stopping interval (3.2) for \( \overline{T}_n \) includes the point \( g(\theta_i) \). For every \( i \), let \( k_i \) be such that \( n_i = k_i \log \alpha \).

**Theorem 4.1.** When \( \theta_i \) is the true value of the parameter \( \theta \),

\[
(4.1) \quad N \leq n_i \text{ as } \alpha \to \infty.
\]

**Proof.** Let \( i \) and \( k > k_i \) be fixed. Let \( n = k \log \alpha \).

\[
P_i (N > n) \leq P_i (g_i (\theta_i) + (\log \alpha)/n(\theta_i - \theta_{i-1}) > \overline{T}_n) +
\]

\[
P_i (g_i (\theta_i) - (\log \alpha)/n(\theta_{i+1} - \theta_i) < \overline{T}_n).
\]
Letting
\[ a(k) = g_-(\theta_i) - g(\theta_i) + 1/k(\theta_i - \theta_{i-1}), \]
\[ b(k) = g_+(\theta_i) - g(\theta_i) - 1/k(\theta_{i+1} - \theta_i), \]
and
\[ z_n = \bar{T}_n - g(\theta_i), \] it follows that \( a(k) < 0, \ b(k) > 0 \) since \( k > k_i \), and
\[ P_i(N > n) \leq P_i(z_n < a(k)) + P_i(z_n > b(k)). \]

Now, let \( d(k) = \min(b(k), -a(k)) \) Then,
\[ P_i(N > n) \leq P_i(|z_n| > d(k)). \]

Applying the Markov inequality [2] for \( r=3 \) gives
\[ P_i(N > n) \leq E_i|z_n|^3/(d(k))^3. \]

Now,
\[ E_i|z_n|^3 \leq n^{-2} E_i|x - \theta_i|^3 \]
where the random variable \( X \) has the same distribution as each of the iid random variables when \( \theta = \theta_i \). Since all moments exist,
\[ E_i|x - \theta_i|^3 < \infty. \]

Letting \( K(k) = (d(k))^{-3} E_i|x - \theta_i|^3 \) gives
\[ P_i(N > n) \leq K(k)n^{-2} = K(k)(k \log \alpha)^{-2}. \]

It follows that \( P_i(N > k \log \alpha) \to 0 \) as \( \alpha \to \infty \) for \( k \) fixed.

Since \( k \) was arbitrary, subject only to \( k > k_i \), it follows that \( N \leq k_i \log \alpha \) as \( \alpha \to \infty. \)

The following lemma will be used to determine the behavior of \( E_iN \) as \( \alpha \to \infty. \)

**Lemma 4.1.** For any \( i \) and \( k' > k > k_i \), there exists a positive constant \( B \), which may depend on \( k \) and \( i \) but not on \( k' \) or \( \alpha \), such that

\[ P_i(N > k' \log \alpha) \leq B(k' \log \alpha)^{-2}. \]

**Proof.** This lemma follows immediately from the proof of the previous theorem by letting \( B = K(k) \) and noting that \( K(k') \leq K(k) \) whenever \( k' > k. \)
Theorem 4.2. For every $i$,

$$E_i N \leq n_i = \kappa_i \log \alpha \text{ as } \alpha \to \infty.$$  

**Proof.** Given the previous lemma, the proof of this theorem is identical with that given for the theorem of section 3.5 in [3] and hence will be omitted.

The following lemma follows by a slight modification of Lemma 2 of section 3.6 in [3]:

**Lemma 4.2.** For each $\alpha > 1$, let $N$ by any stopping rule such that $P_i(N < \infty) = 1$ for all $i$, and let there by a family of associated terminal decision rules with the property $P_i < 2/\alpha$ as $\alpha \to \infty$ for all $i$.

Then,

$$\log \alpha \leq \min(E_i \log(f_{\theta_i}^N / f_{\theta_i + 1}^N), E_i \log(f_{\theta_i}^N / f_{\theta_i - 1}^N)) \text{ as } \alpha \to \infty \text{ for every } i. \tag{4.4}$$

**Theorem 4.3.** For each $\alpha > 1$ let $(N, d)$ be any stopping rule and terminal decision rule satisfying the hypothesis of Lemma 4.2 and let $(N^*, d^*)$ be the corresponding rules proposed in section 3. Then,

$$E_i N^* \leq E_i N \text{ as } \alpha \to \infty \text{ for all } i. \tag{4.5}$$

**Proof.** From the previous lemma and the equality

$$E_i N = (E_i \log(f_{\theta_i}^N / f_{\theta_j}^N) / (E_i \log(f_{\theta_i} / f_{\theta_j}^N)), \text{ valid for any } j \neq i, \text{ it follows that }$$

$$E_i N \leq (\log \alpha) / \min(E_i \log(f_{\theta_i} / f_{\theta_{i+1}}^N), E_i \log(f_{\theta_i} / f_{\theta_{i-1}}^N)).$$

Now, $E_i \log(f_{\theta_i} / f_{\theta_{i+1}}^N) = (\theta_{i+1} - \theta_i)(g_+(\theta_i) - g(\theta_i))$ and $E_i \log(f_{\theta_i} / f_{\theta_{i-1}}^N) = (\theta_{i-1} - \theta_i)(g(\theta_i) - g_-(\theta_i))$. Using the above and the fact that $n_i^{-1} \log \alpha = \min((\theta_{i+1} - \theta_i)(g_+(\theta_i) - g(\theta_i)), (\theta_{i-1} - \theta_i)(g(\theta_i) - g_-(\theta_i)))$, it follows that

$$E_i N \leq n_i \text{ as } \alpha \to \infty \text{ for all } i. \text{ Therefore, since } E_i N^* \leq n_i \text{ by Theorem 4.2, }$$

$$E_i N^* \leq E_i N \text{ as } \alpha \to \infty \text{ for all } i. \Box$$
ACKNOWLEDGEMENTS

I am grateful to Professor Herbert Robbins for direction and guidance in the completion of this work which is based on part of my Ph.D. dissertation (Columbia University, 1970). I would also like to thank Professor Robert Berk for several helpful suggestions.
REFERENCES


