The Annals of Mathematical Statistics
1972, Vol. 43, No. 3, 803-813

SEQUENTIAL ESTIMATION OF A POISSON INTEGER MEAN[1]

BY GEORGE P. MCCABE, JR.
Purdue University

Let $X_1, X_2, \cdots$ be a sequence of i.i.d. Poisson random variables with mean λ. It is assumed that true value of the parameter λ lies in the set $\{0, 1, 2, \cdots\}$. From observations on the sequence it is desired to estimate the true value of the parameter with a uniformly (for all λ) small probability of error. There is no fixed sample size rule which can accomplish this. A sequential procedure based on a likelihood ratio criterion is investigated. The procedure, which depends on a parameter $\alpha > 1$, is such that (i) $P_\lambda(\text{error}) < 2/(\alpha - 1)$ for all λ, and (ii) $E_\lambda(\text{sample size}) \sim k_\lambda \log \alpha$, as $\alpha \to \infty$, where $k_\lambda = (1 - \lambda \log(1 + 1/\lambda))^{-1}$. The procedure is asymptotically optimal as $\alpha \to \infty$.

1. Introduction and summary. One observes a sequence of random variables $X_1, X_2, \cdots$ which are identically and independently distributed Poisson variables with mean λ, i.e.

$$P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n) = \prod_{i=1}^n e^{-\lambda}\lambda^{x_i}/x_i! \qquad \text{for} \quad x_i = 0, 1, \cdots$$

It is assumed that λ is an unknown nonnegative integer which one would like to estimate with an arbitrarily small uniform (for all λ) bound on the probability of error.

The problem of estimating restricted parameters was first considered by Hammersley (1950) from a fixed sample size point of view. The present work is based on a paper by Robbins (1970) in which he proposes a general approach to this type of problem and gives procedures for estimating a normal integer mean. In contrast to the normal case, there is no fixed sample size procedure which will insure an arbitrarily small uniform bound on the error probabilities for the Poisson case.

In this work, a class of procedures is proposed, the associated error probabilities and expected sample sizes are investigated and a weak form of optimality is demonstrated.

2. Fixed sample size approach. Note that $EX_i = \lambda$ and $\text{Var}(X_i) = \lambda$. Also, for a sample of size n, $\bar{X}_n = (X_1 + X_2 + \cdots + X_n)/n$ is unbiased and sufficient for λ, and for the unrestricted parameter space $[0, \infty)$, it is a maximum likelihood estimator. In addition, for large n, the quantity $(\bar{X}_n - \lambda)/(\lambda/n)^{\frac{1}{2}}$ is approximately normal with mean zero and variance one.

A class of reasonable procedures can be characterized as follows: For $i = 0, 1 \cdots$, choose i_- such that $i - 1 \leqq i_- < i$ and set $i_+ = (i + 1)_-$. Then, given a sample of size n, estimate that $\lambda = i$ if $i_- < \bar{X}_n \leqq i_+$. A typical rule in this class

Received February 8, 1971; revised August 1971.

[1] Supported in part by NIH Fellowship 1-F01-GM-42814-01 at Columbia University.

is that with $i_+ = i + \frac{1}{2}$. The maximum likelihood estimates for this problem are discussed by Hammersley (1950). In this case, $i_+ = 1/\log((i+1)/i)$ for $i > 0$ and $i_+ = 0$ for $i = 0$. Also, $i_+ - i \to \frac{1}{2}$ as $i \to \infty$.

Let P_i^* denote the probability of error when i is the true value of the parameter λ. Now,

$$P_i^* = P_i(\bar{X}_n \leqq i_-) + P_i(\bar{X}_n > i_+) .$$

Using the results of Blackwell and Hodges (1959) for large deviation probabilities and assuming that $i_+ = i + \frac{1}{2}$ for all i, it follows that

$$\log P_i^* \sim -n((i + \tfrac{1}{2}) \log((i + \tfrac{1}{2})/i) - \tfrac{1}{2}) \qquad \text{as} \quad n \to \infty .$$

Using

$$\log((i + \tfrac{1}{2})/i) = 1/2i - 1/8i^2 + o(i^{-2}) \qquad \text{as} \quad i \to \infty ,$$

gives

$$\log P_i^* \sim -n/8i \qquad \text{as} \quad n \to \infty \quad \text{and then} \quad i \to \infty ; \tag{2.1}$$

or to be more precise,

$$\lim_{i\to\infty} (8i \lim_{n\to\infty} n^{-1} \log P_i^*) = -1 .$$

Clearly, for a preassigned value of n, it is not possible to insure a small uniform bound on the error probabilities. This is seen to be true for any fixed sample size rule by considering the standard test of the hypothesis $\lambda = i$ versus $\lambda = i + 1$ for sufficiently large i.

Hence, with the aim of devising a decision procedure that will insure a small uniform bound on the error probabilities, one is led to consider sequential procedures.

3. Sequential approach. For $\lambda > 0$, let $S_n = X_1 + \cdots + X_n$, and define

$$f_\lambda^n = f_\lambda(X_1, \cdots, X_n) = e^{-n\lambda}\lambda^{S_n}/(X_1! \cdots X_n!) .$$

For $\lambda = 0$, let $f_0^n = \lim_{\lambda\to 0} f_\lambda^n$. Thus, f_0^n equals 1 or 0 according as S_n is zero or positive. Now for i and j positive,

$$f_i^n/f_j^n = e^{-n(i-j)}(i/j)^{S_n} ,$$

and f_1^n/f_0^n equals f_i^n or ∞ according as S_n is zero or positive. This is consistent with the above if the convention is adopted that $(i/j)^{S_n} = 1$ when $j = 0$ and $S_n = 0$.

LEMMA 3.1. *For $\lambda > 0$, $f_{\lambda-1}^n/f_\lambda^n$ is a monotone increasing function of λ.*

Now, let

$$\begin{aligned} L_i^n &= \min(f_i^n/f_{i+1}^n, f_i^n/f_{i-1}^n) && \text{for} \quad i > 0 , \\ &= f_0^n/f_1^n && \text{for} \quad i = 0 . \end{aligned}$$

Consider the following rule:

Fix $\alpha > 1$. Let N be the smallest $n \geqq 1$ such that $L_i^n \geqq \alpha$ for some i. Stop at stage $N = n$ and estimate that λ is the i for which $L_i^n \geqq \alpha$. First, note that there is no ambiguity in the estimate since $L_i^n \geqq \alpha$ for some i implies that $L_j^n < 1$ for all $j \neq i$.

The form of the rule can be considerably simplified as follows:

Suppose $i > 0$. Then $L_i^n \geqq \alpha$ implies that

$$f_i^n/f_{i+1}^n = e^n(i/(i+1))^{S_n} \geqq \alpha$$

or

$$\bar{X}_n \leqq i_+ - i_+(\log \alpha)/n ,$$

where $\bar{X}_n = S_n/n$ and $i_+ = 1/\log((i+1)/i)$.

Similarly, for $i > 1$, $L_i^n \geqq \alpha$ implies

$$\bar{X}_n \geqq i_- + i_-(\log \alpha)/n ,$$

where $i_- = (i-1)_+ = 1/\log(i/(i-1))$.

For completeness let $i_- = 0$ for $i = 1$. Also, $f_1^n/f_0^n = e^{-n}$ or ∞ according as S_n is zero or positive. Thus, $L_0^n \geqq \alpha$ implies $n \geqq \log \alpha$ and $S_n = 0$; and $L_1^n \geqq \alpha$ implies $f_1^n/f_0^n \geqq \alpha$ or simply $S_n > 0$. Thus, the rule can be rewritten as follows:

(3.1) Stop at $N = n$ as soon as one of the following is true:

(a) $S_n > 0$ and for some $i > 0$, $i_- + i_-(\log \alpha)/n \leqq \bar{X}_n \leqq i_+ - i_+(\log \alpha)/n$; guess that $\lambda = i$;

(b) $S_n = 0$ and $n \geqq \log \alpha$; guess $\lambda = 0$.

Note that as $n \to \infty$, $i_-(\log \alpha)/n \to 0$, $i_+(\log \alpha)/n \to 0$ and $\bar{X}_n$ converges almost surely to λ, an integer. Thus if $i_- < i < i_+$, the procedure will terminate with probability one. This is seen to be true from the inequality

$$(n+1)^{-1} < \log((n+1)/n) < n^{-1} .$$

It is interesting to note that as $i \to \infty$, $i_+ - i \to \frac{1}{2}$.

4. Sample size. Recall that a guess of $\lambda = 0$ implies that $n \geqq \log \alpha$. Also, note that for large α and small n,

$$i_- + i_-(\log \alpha)/n > i_+ - i_+(\log \alpha)/n .$$

But, $i_- + i_-(\log \alpha)/n$ decreases to i_- and $i_+ - i_+(\log \alpha)/n$ increases to i_+ as $n \to \infty$. Thus, for each i, there is a minimum sample size, call it m_i, which is the smallest sample size which will admit a guess of $\lambda = i$. For conciseness, m_i will be identified with any number less than m_i and greater than $m_i - 1$. To find m_i, set

$$i_- + i_-(\log \alpha)/n = i_+ - i_+(\log \alpha)/n .$$

Solving for n gives

$$\begin{aligned} (4.1) \qquad m_i &= \log \alpha && \text{for } i = 0, 1 ; \\ &= (\log \alpha)(\log((i+1)/(i-1)))/\log(i^2/(i^2-1)) && \text{for } i > 1 . \end{aligned}$$

Note that $n \geqq m_i$ does not imply

$$i_- + i_-(\log \alpha)/n \leqq i \leqq i_+ - i_+(\log \alpha)/n .$$

It will be necessary to use the minimum value of n, which will be denoted by n_i, such that this expression is valid. Clearly $n_0 = \log \alpha$. For $i > 1$, $i_- + i_-(\log \alpha)/n = i$ implies

$$n = (\log \alpha)/(i \log (i/(i-1)) - 1) ,$$

and for $i = 1$, the inequality $i_- + i_-(\log \alpha)/n < i$ is valid for all $n \geqq 1$. Similarly, for $i \geqq 1$,

$$i = i_+ \dot{-} i_+(\log \alpha)/n$$

implies

$$n = (\log \alpha)/(1 - i \log ((i+1)/i)) .$$

Hence,

$$\begin{aligned} n_i &= (\log \alpha)/\min (i \log (i/(i-1)) - 1,\ 1 - i \log ((i+1)/i)) && \text{for} \quad i > 1 , \\ &= (\log \alpha)/(1 - \log 2) && \text{for} \quad i = 1 , \end{aligned}$$

where nonintegral values are interpreted in the obvious manner.

Integrating the Taylor expansion for x^{-1} about the point $(x+1)/2$ gives $\log x = 2(z + z^3/3 + \cdots)$ where $z = (x-1)/(x+1)$, $x > 0$. It follows that for $x > 1$, $\log x > 2(x-1)/(x+1)$. (This inequality could have been obtained in a more statistical manner by an application of Jensen's inequality.) Setting $x = (i+1)/(i-1)$ yields $\log ((i+1)/(i-1)) > 2/i$, which upon rearrangement gives

$$1 - i \log ((i+1)/i) < i \log (i/(i-1)) - 1 .$$

Hence,

$$\begin{aligned} n_i &= (\log \alpha)/(1 - i \log ((i+1)/i)) && \text{for} \quad i \geqq 1 , \\ &= \log \alpha && \text{for} \quad i = 0 . \end{aligned} \tag{4.2}$$

Note that as $i \to \infty$,

$$n_i \sim 2i \log \alpha . \tag{4.3}$$

5. Error probabilities. Let P_i denote the probability of error when $\lambda = i$. Then,

$$P_i = \sum_{j \neq i} \sum_{n \geqq m_j} \sum_{A_{n,j}} f_i^n ,$$

where $A_{n,j}$ is the set where $X_1, \cdots, X_n$ are such that $N = n$ and the estimated value of λ is j. Note $P_0 = 0$. Now, for $i > 0$, let

$$a_i = \sum_{j < i} \sum_{n \geqq m_j} \sum_{A_{n,j}} f_i^n ,$$

and

$$b_i = \sum_{j > i} \sum_{n \geqq m_j} \sum_{A_{n,j}} f_i^n .$$

Thus, $P_i = a_i + b_i$. Now,

$$a_i = \sum_{j<i} \sum_{n \geqq m_j} \sum_{A_{n,j}} f_j^n \left(\frac{f_{j+1}^n}{f_j^n}\right) \cdots \left(\frac{f_i^n}{f_{i-1}^n}\right).$$

Since

$$(f_k^n / f_{k-1}^n) \leqq (f_m^n / f_{m-1}^n) \qquad \text{for} \quad m \leqq k$$

by Lemma 3.1 and $f_j^n / f_{j+1}^n \geqq \alpha$ on $A_{n,j}$, it follows that

$$\max((f_{j+1}^n / f_j^n), \cdots, (f_i^n / f_{i-1}^n)) = (f_{j+1}^n / f_j^n) \leqq \alpha^{-1}$$

on $A_{n,j}$.
Therefore,

$$\begin{aligned} a_i &\leqq \sum_{j<i} \sum_{n \geqq m_j} \sum_{A_{n,j}} f_j^n \alpha^{-(i-j)} \\ &= \sum_{j<i} (\alpha^{-(i-j)} \sum_{n \geqq m_j} \sum_{A_{n,j}} f_j^n). \end{aligned}$$

Since

$$\sum_{n \geqq m_j} \sum_{A_{n,j}} f_j^n = 1 - P_j \leqq 1,$$

it follows that

$$a_i \leqq \sum_{j=0}^{i-1} \alpha^{-(i-j)}$$

or $a_i \leqq (1 - \alpha^{-i})/(\alpha - 1)$.
In an entirely analogous fashion, it can be shown that $b_i \leqq 1/(\alpha - 1)$. Combining gives

$$\begin{aligned} P_i &\leqq (2 - \alpha^{-i})/(\alpha - 1) \\ &< 2/(a - 1) \qquad \text{for all } i. \end{aligned} \tag{5.1}$$

Hence, by appropriate choice of α, one can obtain an arbitrarily small uniform bound on the error probabilities.

6. Expected sample size. As in Section 4, when considering sample sizes, no distinction will be made between n, an integer, and any real number less than n but greater than $n - 1$. Recall that $S_n = 0$ for every n when $\lambda = 0$; so $N = \log \alpha$ and $E_0 N = \log \alpha$.

Now, for $i \geqq 1$, let

$$k_i = 1/(1 - i \log((i+1)/i)),$$

and let $k_0 = 1$. Then $n_i = k_i \log \alpha$.
Recall that n_i is the smallest sample size such that

$$i_- + i_-(\log \alpha)/n \leqq i \leqq i_+ - i_+(\log \alpha)/n.$$

Let $i \geqq 1$ and $k > k_i$ be fixed. Let $n = k \log \alpha$. Thus, $n > n_i$ and

$$\begin{aligned} P_i(N > n) &\leqq P_i(i_- + i_-(\log \alpha)/n > \bar{X}_n) \\ &\qquad + P_i(i_+ - i_+(\log \alpha)/n < \bar{X}_n) \\ &= P_i(a > z_n) + P_i(b < z_n), \end{aligned}$$

where $a = i_- - i + i_-/k$, $b = i_+ - i - i_+/k$, and $z_n = \bar{X}_n - i$.
Since $k > k_i$ it follows that $a < 0$ and $b > 0$. Also, $b < -a$ by the same argument used to find n_i in Section 4. Therefore,

$$P_i(N > n) \leqq P_i(|z_n| > b) .$$

The latter probability can be bounded using the Markov inequality with $r = 4$ (see, for example, Loève (1963)). This gives

$$P_i(|z_n| > b) \leqq b^{-4} E z_n^4 .$$

Now,

$$E z_n^4 = n^{-4} E(X - ni)^4 ,$$

where X is a Poisson random variable with mean ni. Hence

$$E z_n^4 = n^{-4}(ni + 3n^2 i^2) .$$

Therefore,

$$P_i(N > n) \leqq K n^{-2} = K(k \log \alpha)^{-2} , \tag{6.1}$$

where $K = b^{-4}(i/m_i + 3i^2)$.
By letting $\alpha \to \infty$ in the above expression, it is seen that $P_i(N > n) \to 0$ as $\alpha \to \infty$. Since k was arbitrary, subject only to the condition $k > k_i$, it follows that N is asymptotically less than or equal to n_i as $\alpha \to \infty$; or to be more precise, letting $k = k_i(1 + \varepsilon)$,

$$\lim_{\alpha \to \infty} P_i(N \leqq n_i(1 + \varepsilon)) = 1 \qquad \text{for any } \varepsilon > 0 \text{ and all } i. \tag{6.2}$$

The study of the behavior of $E_i N$ as α gets large will be aided by the following:

LEMMA 6.1. *Let $i \geqq 1$ and $k' \geqq k > k_i$. Then there exists a positive number K which may depend on i and k but not on k' or α such that*

$$P_i(N > k' \log \alpha) \leqq K(k' \log \alpha)^{-2} . \tag{6.3}$$

PROOF. Let $n = k \log \alpha$ and $n' = k' \log \alpha$. By (6.1), $P_i(N > n') \leqq K'(n')^{-2}$, where

$$K' = (i_+ - i - i_+/k')^{-4}(i/m_i + 3i^2)^4 .$$

But,

$$K' \leqq K = (i_+ - i - i_+/k)^{-4}(i/m_i + 3i^2)^4$$

since $k \leqq k'$. Therefore,

$$P_i(N > n') \leqq K(n')^{-2} .$$ □

THEOREM 6.1. *For $i \geqq 0$,*

$$\limsup_{\alpha \to \infty} (n_i^{-1} E_i N) \leqq 1 . \tag{6.4}$$

PROOF. The case where $i = 0$ has already been considered. Let $i \geqq 1$ and $k > k_i$ be fixed. Set $n = k \log \alpha$. For convenience, it will be assumed that $k \log \alpha$ is an integer.
Now,

$$\begin{aligned} E_i N &= \textstyle\sum_{j\geq 1} j P_i(N=j) \\ &\leq n + \textstyle\sum_{j>n} j P_i(N=j) \\ &= n + (n+1)P_i(N>n) + \textstyle\sum_{j>n} P_i(N>j) . \end{aligned}$$

By the previous lemma,

$$(n+1)P_i(N>n) \leq K(k \log \alpha + 1)(k \log \alpha)^{-2} .$$

Clearly, as $\alpha \to \infty$, this term goes to zero. Applying the previous lemma to each term of the last summation above gives

$$\textstyle\sum_{j>n} P_i(N>j) \leq K \sum_{j>k \log \alpha} j^{-2} .$$

This series is clearly convergent. Thus, as $\alpha \to \infty$, this term also approaches zero. Hence,

$$n^{-1} E_i N \leq 1 + o((\log \alpha)^{-1}) .$$

Since k was arbitrary subject only to $k > k_i$, (6.3) follows. □

Note that $(2i)^{-1}k_i \to 1$ as $i \to \infty$. Therefore,

$$\limsup_{i\to\infty} ((2i)^{-1} \limsup_{\alpha\to\infty} (\log \alpha)^{-1} E_i N) \leq 1 . \tag{6.5}$$

7. Optimality. The following two lemmas will be useful in proving the main result of this section. Let F_n be the σ-algebra generated by $(X_i, \cdots, X_n)$.

LEMMA 7.1. *Let N be any stopping rule with $P_i(N < \infty) = 1$ and let A be any set such that $A \cap \{N=n\}$ is in F_n for all n. If $P_i(A) > 0$ and $P_{i+1}(A) > 0$, then*

$$E_i(\log (f_i^N / f_{i+1}^N) \mid A) \geq \log (P_i(A)/P_{i+1}(A)) . \tag{7.1}$$

PROOF.

$$\begin{aligned} E_i(\log (f_i^N/f_{i+1}^N) \mid A) &= -E_i(\log (f_{i+1}^N/f_i^N) \mid A) \\ &\geq -\log E_i((f_{i+1}^N/f_i^N) \mid A) \end{aligned}$$

by Jensen's inequality. Let $A_n = \{N = n\} \cap A$. Then,

$$\begin{aligned} E_i((f_{i+1}^N/f_i^N) \mid A) &= (P_i(A))^{-1} \textstyle\sum_n \sum_{A_n} (f_{i+1}^n/f_i^n) f_i^n \\ &= (P_i(A))^{-1} \textstyle\sum_n \sum_{A_n} f_{i+1}^n \\ &= P_{i+1}(A)/P_i(A) . \end{aligned}$$

Substituting this expression into the above inequality gives the desired result. □

LEMMA 7.2. *For any $\alpha > 1$, let N be any stopping rule such that $P_i(N < \infty) = 1$ for all i and let there be an associated terminal decision rule with the property that $P_i(error) \leq 2/(\alpha - 1)$ for all i. Then, for every i,*

$$\liminf_{\alpha\to\infty} (\log \alpha)^{-1} E_i(\log (f_i^N/f_{i+1}^N)) \geq 1 . \tag{7.2}$$

PROOF. Let $C_i = \{$estimate that λ is $i\}$. Then, $P_i(\text{error}) = P_i(C_i^c)$. Without loss of generality, assume that $P_i(C_j) > 0$ for all i and j. Any decision rule can be modified on a set of arbitrarily small probability to meet this condition. Now, let i be fixed. Clearly,

$$E_i \log (f_i^N/f_{i+1}^N) = P_i(C_i)E_i(\log (f_i^N/f_{i+1}^N) \mid C_i) + P_i(C_i^c)E_i(\log (f_i^N/f_{i+1}^N) \mid C_i^c) .$$

Applying the previous lemma, first with $A = C_i$ and then with $A = C_i^c$ yields

$$E_i \log (f_i^N/f_{i+1}^N) \geqq P_i(C_i) \log (P_i(C_i)/P_{i+1}(C_i)) + P_i(C_i^c) \log (P_i(C_i^c)/P_{i+1}(C_i^c)) . \tag{7.3}$$

Now, $P_i(C_i^c) \leqq 2/(\alpha - 1)$, so, $P_i(C_i) \geqq (\alpha - 3)/(\alpha - 1)$ and

$$P_{i+1}(C_i) < P_{i+1}(C_{i+1}^c) \leqq 2/(\alpha - 1) .$$

Thus,

$$P_i(C_i) \log (P_i(C_i)/P_{i+1}(C_i)) \geqq ((\alpha - 3)/(\alpha - 1)) \log ((\alpha - 3)/2) ,$$

which is asymptotic to $\log \alpha$ as $\alpha \to \infty$.
Also,

$$P_i(C_i^c) \log (P_i(C_i^c)/P_{i+1}(C_i^c)) > P_i(C_i^c) \log P_i(C_i^c)$$

which approaches zero as $\alpha \to \infty$ since $P_i(C_i^c) \leqq 2/(\alpha - 1)$. Combining the above with (7.3) gives the desired result. □

THEOREM 7.1. *For any $\alpha > 1$, let (N^*, d^*) be the stopping rule and terminal decision function described in Section 3, and let (N, d) be any stopping rule and associated terminal decision function such that $E_i N < \infty$ and $P_i(error) \leqq 2/(\alpha - 1)$ for all i. Then, for every i,*

$$\limsup_{\alpha \to \infty} (E_i N^*/E_i N) \leqq 1 . \tag{7.4}$$

PROOF. Since the Poisson variables $X_1, X_2, \cdots$ are identically and independently distributed, and $E_i N < \infty$ for all i, the following well-known equality is valid for all i:

$$E_i(\log (f_i^N/f_{i+1}^N)) = (E_i N)(E_i(\log (f_i(x)/f_{i+1}(x))) .$$

Recall that

$$f_i(x)/f_{i+1}(x) = e(i/(i + 1))^x \qquad \text{for all } i \geqq 1 ,$$

so

$$\begin{aligned} E_i(\log (f_i(X)/f_{i+1}(X))) &= E_i(1 + X \log (i/(i + 1))) \\ &= 1 - i \log ((i + 1)/i) \\ &= k_i^{-1} . \end{aligned}$$

This is also valid for $i = 0$. Hence,

$$E_i N = k_i E_i(\log (f_i^N/f_{i+1}^N)) \qquad \text{for all } i .$$

Now, by Lemma 7.2,

$$\liminf_{\alpha \to \infty} (\log \alpha)^{-1} E_i(\log (f_i^N/f_{i+1}^N)) \geqq 1$$

so,

$$\liminf_{\alpha\to\infty} ((k_i \log \alpha)^{-1} E_i N) \geqq 1 .$$

But by Theorem (6.1),

$$\limsup_{\alpha\to\infty} (n_i^{-1} E_i N^*) \leqq 1$$

for all i. Therefore,

$$\limsup_{\alpha\to\infty} (E_i N^* / E_i N) \leqq 1 . \qquad \square$$

8. Comparison of fixed and sequential procedures. The sequential procedure is obviously far superior to any fixed sample size procedure since it is only with a sequential approach that one can obtain an arbitrarily small uniform bound on the error probabilities for the whole parameter space.

Let i be fixed and suppose that one could somehow (perhaps by a two-stage sampling procedure) pick a sample size which would give a reasonable bound on the error probability for the true parameter, i.e. using (2.1), pick n such that $\log P_i^* = -n/8i$. Suppose further that i and n are large enough for this expression to validly approximate the fixed sample size error probability and for the expression $E_i N \leqq 2i \log \alpha$ to be roughly true by virtue of (6.5). Of course, a knowledge of i is being assumed and a great deal of approximation is involved, but these facts will be temporarily neglected. Now let

$$\log (2/(\alpha - 1)) = -n/8i .$$

Then,

$$\log (\alpha - 1) = (n/8i) + \log 2 .$$

Multiplying by $2i$, one sees that as $\alpha \to \infty$, roughly speaking, $E_i N$ will be asymptotically less than or equal to $n/4$. Thus, even if it were possible to preselect the appropriate sample size for a fixed procedure, the sequential approach requires only about $\frac{1}{4}$ as many observations to attain the same error probability for α and i large.

9. Monte Carlo results. To investigate the procedure described in Section 3 for various values of α and λ, a Fortran program for an IBM-360-90 was written. Sequences of Poisson variables with a given mean were generated, the stopping rules and terminal decision rules were applied, and the results were tabulated. For each value of α and λ, 1000 sequences were generated.

For convenience, an arbitrary upper bound of 1000 was set on the length of the sequences. At the point of truncation, the decision function was taken to be the maximum likelihood estimate of λ. For the date presented in Table 1, this truncation point was reached for only one sequence.

For each value of the pair (α, λ), $(\alpha = 3, 5, 21, 41, 81;\ \lambda = 1, 3, 5, 10, 20)$ the following quantities are tabulated:

(a) mean $= i =$ the true value of λ;
(b) P(err) $=$ average number of incorrect decisions;
(c) TP(err) $=$ theoretical bound on the error probability $= 2/(\alpha - 1)$;

TABLE 1
Results of Monte Carlo Experiment

Mean	P(err)	TP(err)	Av-N	Tav-N	Fix-N
1	0.214	1.0	4.11	3.58	1
3	0.305	1.0	11.28	8.02	1
5	0.293	1.0	17.87	12.43	1
10	0.318	1.0	34.34	23.42	1
20	0.343	1.0	63.82	45.38	1
1	0.192	0.5	5.83	5.24	3
3	0.192	0.5	16.92	11.75	6
5	0.173	0.5	26.63	18.21	10
10	0.219	0.5	51.36	34.31	19
20	0.215	0.5	99.08	66.47	37
1	0.035	0.1	11.04	9.92	11
3	0.042	0.1	31.62	22.23	33
5	0.049	0.1	51.05	34.44	55
10	0.055	0.1	100.66	64.91	109
20	0.051	0.1	191.76	125.75	217
1	0.029	0.05	13.44	12.10	16
3	0.023	0.05	37.15	27.12	47
5	0.024	0.05	61.64	42.01	77
10	0.025	0.05	121.83	79.17	154
20	0.018	0.05	235.85	153.38	308
1	0.007	0.025	16.05	14.32	21
3	0.006	0.025	42.68	32.09	61
5	0.013	0.025	71.25	49.71	101
10	0.013	0.025	140.10	93.69	202
20*	0.010	0.025	270.85	181.50	403

* One sequence in this group was truncated at 1000.

(d) Av-N = the average sample size;

(e) Tav-N $= k_i \log \alpha =$ theoretical asymptotic bound for the expected sample size; and

(f) Fix-N = the sample size which would be required to distinguish the hypothesis $\lambda = i$ from $\lambda = i + 1$ or $i - 1$ with an error probability less than or equal to $(2 - \alpha^{-i})/(\alpha - 1)$. (When $\lambda = i$, the sequential procedure has error probability less than or equal to this quantity by (5.1)). The normal approximation was used to calculate Fix-N.

These results point out that in many cases, the true error probability may be somewhat less than the theoretical bound. This is due mostly to the inequalities introduced in the derivation of (5.1). It is not surprising that the calculated average sample size is greater than Tav-N since the latter quantity is an asymptotic bound. The average sample sizes obtained do, however, compare favorably with the corresponding fixed sample size values for the moderate values of α used.

Acknowledgments. 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

[1] BLACKWELL, D. and HODGES, J. L. (1959). The probability in the tail of a convolution. *Ann. Math. Statist.* **30** 1113–1120.

[2] HAMMERSLEY, J. M. (1950). On estimating restricted parameters. *J. Roy. Statist. Soc. Ser. B* **12** 192–240.

[3] LOÈVE, M. (1963). *Probability Theory*. Van Nostrand, Princeton.

[4] ROBBINS, H. (1970). Sequential estimation of an integer mean, in *Scientists at Work, Festschrift in honour of Herman Wold*. Almqvist and Wiksells Boktryckeri Ab, Uppsala.