SEQUENTIAL ESTIMATION IN THE UNIFORM DENSITY

by

Peter J. Cooke

TECHNICAL REPORT NO. 144

April 21, 1969

PREPARED UNDER CONTRACT Nonr-225(52)
(NR-342-022) FOR
OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

SEQUENTIAL ESTIMATION IN THE UNIFORM DENSITY

by

Peter J. Cooke

The main problem to be solved here may be described as follows: let $X_1, X_2, \ldots$ be independent random variables, each with density $f_\theta(x) = \frac{1}{\theta}$ over $(0,\theta)$ and zero elsewhere. It is desired to estimate the unknown parameter θ by an interval of length at most d units and with confidence at least $1-\alpha$, for some specified $d > 0$ and α in $(0,1)$. An exact solution and an asymptotic theory for a sequential procedure are given in sections 1 and 2, respectively.

The procedure proposed in this paper is optimal in the sense that the expected number of observations is minimized. It is also minimax in that the maximum possible number of observations is minimized.

1. The Procedure.

By an estimation rule δ, we understand the specification of a stopping rule, which for given $X_1, X_2, \ldots$ determines the number N of observations to be made, together with a function which we also denote by δ, mapping the possible $(X_1, X_2, \ldots, X_N)$ into sets of possible values of θ. Associated with a particular δ is a function $\gamma(\theta) = P_\theta\{\theta\epsilon\delta(X_1, X_2, \ldots)\}$ which is the probability that $\delta(X_1, X_2, \ldots)$ contains θ in the sequel to be called the confidence function.

For the problem to be solved here, without loss of generality we may suppose $d = 1$, since for any other positive d, $\frac{x_i}{d}$ is uniformly distributed over $(0, \frac{\theta}{d})$. Hence we may consider the problem as one of estimating $\frac{\theta}{d}$ by an interval of at most unit length.

The sets we shall use to estimate θ are intervals of the form $\delta(X_1,X_2,\ldots,X_N) = (\hat{X}_N, g(\hat{X}_N)]$, where $\hat{X}_N$ denotes the maximum of $X_1,X_2,\ldots,X_N$. Clearly, δ maps the sample space into intervals of length ≤ 1 unit on $(0,\theta+1)$ if $\hat{X}_N \leq g(\hat{X}_N) < \hat{X}_N+1$. Our confidence requirement is $\gamma(\theta) = P_\theta\{\hat{X}_N < \theta \leq g(\hat{X}_N)\} \geq 1-\alpha$ for all θ. Since $\hat{X}_N$ is sufficient for θ in the fixed sample size case, $(N,\hat{X}_N)$ is sufficient for θ in the sequential case. (See Lehmann [3], p. 3.32.) But N is a function of $\hat{X}_N$, so $\hat{X}_N$ alone is sufficient for θ.

The sequential procedure we shall adopt is as follows:

(1) Observe $X_1,X_2,\ldots$ until for the first time $\hat{X}_N \leq a_N$,

where $a_1,a_2,\ldots$ form a non-decreasing sequence of non-negative real numbers.

(2) If this occurs at $N = n$, make the statement '$\hat{X}_n < \theta \leq g(\hat{X}_n)$'.

The confidence function associated with this statement will be at least $1-\alpha$ for every θ for an appropriately chosen sequence $\{a_j\}$.

The stopping sets for this procedure are determined by the sequence $\{a_j\}$. Clearly, the stopping set, S_n, of points $(X_1,X_2,\ldots,X_n)$ at which sampling stops at $N = n$, and the continuation set, C_n, the complement of S_n in n-dimensional Cartesian space, may be determined successively for $n = 1,2,\ldots$ by $S_1 = (0,a]$ and the recurrence relation $S_n = (C_{n-1} \cap (0,a_n]^{n-1}) \times (0,a_n]$. Thus we find

$S_2 = (a_1,a_2]x(0,a_2]$, $S_3 = (a_1,a_2]x(a_2,a_3]x(0,a_3] \cup (a_2,a_3]x(0,a_3]x(0,a_3]$, etc.

We will now determine the distribution function of $\hat{X}_N$. On the basis of our sampling procedure, for $\theta \geq a_n$ we have

$$P_\theta(N=n) = P_\theta\{(X_1,X_2,\ldots,X_n)\epsilon S_n\}$$

$$= \frac{1}{\theta^n} \int_{S_n}\!\!\cdots\!\int dx_1 dx_2,\ldots,dx_n$$

$$= \frac{b_n}{\theta^n}\,, \quad \text{for some } b_n\,.$$

For $a_{n-1} < \theta \leq a_n$, $P_\theta(N=n) = 1-P_\theta(N \leq n-1) = 1 - \sum_{r=0}^{n-1} \frac{b_r}{\theta^r}$. Define $b_o = 0$. Thus,

$$P_\theta(N=n) = \begin{cases} 0 \quad , & \theta \leq a_{n-1} \\ 1 - \sum_{r=1}^{n-1} \frac{b_r}{\theta^r} \quad , & a_{n-1} < \theta \leq a_n \\ \frac{b_n}{\theta^n} \quad , & \theta \geq a_n \end{cases} \tag{1.1}$$

For $\theta = a_n$, $P_\theta(N \leq n) = \sum_{r=1}^{n} \frac{b_r}{\theta^r} = 1$. Hence

$$b_n = a_n^n - b_1 a_n^{n-1} - b_2 a_n^{n-2} \cdots\cdots b_{n-1} a_n, \quad n = 1,2,\ldots \tag{1.2}$$

We will denote by $B_n(x)$ the polynomial obtained by replacing a_n

in the right hand side of (1.2) by x; i.e.,

$$B_n(x) = x^n - b_1 x^{n-1} - b_2 x^{n-2} \cdots\cdots b_{n-1} x \,. \tag{1.3}$$

Thus, we infer, for $a_{n-1} < x \le a_n$ and $x \le \theta$, $P_\theta(\hat{X}_n \le x, N=n) = \dfrac{B_n(x)}{\theta^n}$.

Now $P_\theta(\hat{X}_n \le x | N=n) = 1$, for $x \ge a_n$. Hence

$$P_\theta(\hat{X}_N \le x) = \sum_{r=1}^{\infty} P_\theta(\hat{X}_r \le x,\ N=r)$$

$$= \sum_{r=1}^{\nu(x)-1} P_\theta(N=r) + P_\theta\{\hat{X}_{\nu(x)} \le x, N = \nu(x)\} \quad ,$$

where

$$\nu(x) = j \text{ if } a_{j-1} < x \le a_j \,.$$

It follows that for all x in $(a_{n-1}, a_n]$ and $x \le \theta$,

$$P_\theta(\hat{X}_N \le x) = \sum_{r=1}^{-1} \frac{b_r}{\theta^r} + \frac{B_n(x)}{\theta^r} \tag{1.4}$$

For $x > \theta$ each side of this equation equals 1.

Clearly, $\nu(\theta)$ is the maximum number of observations which could be required; i.e., $\nu(\theta) = \max_\theta N$, the largest n for which $P_\theta(N=n) > 0$.

Consider now procedures with terminal statement '$\hat{X}_N < \theta \le \hat{X}_N+1$.' Using (1.4), the probability that this statement is untrue is given by

$$\alpha(\theta) \equiv P_\theta(\hat{X}_N \le \theta-1) = \sum_{r=1}^{n-1} \frac{b_r}{\theta^r} + \frac{B_n(\theta-1)}{\theta^n} \,, \quad a_{n-1} < \theta-1 \le a_n \tag{1.5}$$

Hence the requirement $\gamma(\theta) \ge 1-\alpha$ is equivalent to $\alpha(\theta) \le \alpha$ for all θ.

The optimality criterion we shall adopt is as follows: of all procedures for which $\alpha(\theta) \leq \alpha$ for all θ, a procedure is optimal if every other procedure with smaller ν-function for some θ has larger ν-function for at least one $\theta' < \theta$. The solution to be investigated will easily be seen to also satisfy an optimality criterion of the same form expressed in terms of the expected number of observations rather than the ν-function.

Consider the case $0 < \theta - 1 \leq a_1$. Using (1.5) and (1.3) we have $\alpha(\theta) = \frac{\theta-1}{\theta}$, which we require to be less than or equal to α. Hence a_1 must be such that $\frac{a_1}{a_1+1} \leq \alpha$. Suppose $\frac{a_1^*}{a_1^*+1} = \alpha$. Let P and P* be procedures associated with a_1 and a_1^* and with ν-functions $\nu(\theta)$ and $\nu(\theta)^*$, respectively. If we put $a_1 = a_1^*$ we have $\nu(\theta) = 1$ for θ in $(0, a_1^*]$. Suppose we choose $a_1 < a_1^*$. Then the procedure P cannot be optimal, since for any procedure P' associated with a_1' in $(a_1, a_1^*]$, $\nu(\theta)' < \nu(\theta)$ for θ in $(a_1, a_1']$, but $\nu(\theta)'$ is not greater than $\nu(\theta)$ for any $\theta < a_1'$. Since this is true for any $a_1 < a_1^*$ and a_1' in $(a_1, a_1^*]$, P cannot be optimal for $a_1 < a_1^*$. Hence we put $a_1 = a_1^*$; i.e., we choose a_1 as large as possible. From (1.2), $b_1 = a_1$. In general, because of our optimality criterion, for each n we choose a_n as large as possible; i.e., a_n is the largest x for which

$$\sum_{r=1}^{n-1} \frac{b_r}{(x+1)^r} + \frac{x^n - b_1 x^{n-1} - b_2 x^{n-2} - \cdots - b_{n-1} x}{(x+1)^n} = \alpha \qquad (1.6)$$

and b_n is determined using (1.2). It should be noticed that $a_2 = a_1$; i.e., N cannot take the value 2 and so the procedure must be started by taking one observation and if the observed $X_1 > a_1$, an additional two observations. The second decision whether or not to continue sampling is then based on the observed value of $\hat{X}_3$.

Values (correct to 4 significant figures) of the first 20 members of the sequence $\{a_j\}$ are given in table 1 for $\alpha = 0.05$ and 0.01. Members of the sequence $\{b_j\}$ are required for evaluating the expected number of observations, hereafter to be denoted by $E_\theta(N)$. However, in section 2 we will derive asymptotic expansions for this function which are independent of the higher members of the sequence $\{b_j\}$. Thus (in table 2) we only tabulate (correct to 4 significant figures) the first 10 members of this sequence for the above values of α.

From (1.6) we have $\sum_{r=1}^{n-1} \frac{b_r}{(x+1)^r} + \frac{B_n(x)}{(x+1)^n} < \alpha$ for x in (a_{n-1}, a_n). Hence $\alpha(\theta) < \alpha$ for θ in the intervals $(a_{n-1}+1, a_n+1)$, $n=1,2,\ldots,$ where a_o is defined to be zero. However, for these values of θ, the confidence will equal $1-\alpha$ if we slightly modify the procedure and estimate θ by an interval of the form $(\hat{X}_N, g(\hat{X}_N)]$, where $g(\hat{X}_N)-\hat{X}_N < 1$. Clearly, we require $P_\theta\{g(\hat{X}_N) < \theta\} = P_\theta\{\hat{X}_N < g^{-1}(\theta)\} = \alpha$.

There will exist a value $x(\alpha,\theta)$ of x, depending on α and θ, for which $P_\theta(\hat{X}_N < x) = \alpha$. Hence $g^{-1}(\theta) = x(\alpha,\theta)$. For the optimal procedure $\nu(\theta)$ cannot take the value 2. Suppose $\nu(\theta) = n$, $n \neq 2$. Then $P_\theta(\hat{X}_N \leq x) = \sum_{r=1}^{n-1} \frac{b_r}{\theta^r} + \frac{B_n(x)}{\theta^n}$. Thus the largest root of the polynomial $\alpha\theta^n - b_1\theta^{n-1} - b_2\theta^{n-2} - \cdots - b_{n-1}\theta - B_n(x) = 0$ when x is replaced by the observed value of $\hat{X}_n$ is the required value of $g(\hat{X}_n)$, for a specified α. Clearly, $g(a_n) = a_n + 1$ since $\alpha(a_n+1) = \alpha$ for $n=1,2,\ldots$.

Using (1.1), for θ in the interval $(a_{n-1}, a_n]$, the expected sample size is given by

$$E_\theta(N) = \sum_{r=1}^{n-1} \frac{r \cdot b_r}{\theta^r} + n\left(1 - \sum_{r=1}^{n-1} \frac{b_r}{\theta^r}\right) = n - \sum_{r=1}^{n-1} (n-r) \cdot \frac{b_r}{\theta^r} . \tag{1.7}$$

As a variation of the procedure already described, we may consider taking observations in groups of m, $m > 1$. Let $a_n^{(m)}$ denote the n-th member of the sequence which determines the optimal procedure (i.e., the sequence in which members are chosen successively as large as possible). Thus we have $0 = a_1^{(m)} = a_2^{(m)} = \cdots = a_{m-1}^{(m)} < a_m^{(m)} = a_{m+1}^{(m)} = \cdots = a_{2m-1}^{(m)} < a_{2m}^{(m)} = \cdots$, $b_j^{(m)} = 0$ for j not a multiple of m and $0 < b_m^{(m)} < b_{2m}^{(m)} < \cdots$, where $\{b_j^{(m)}\}$ is the sequence determined by $\{a_j^{(m)}\}$ using (1.2). Also, using (1.7), for n a multiple of m, say $n = zm$, and θ in $(a_{n-m}^{(m)}, a_n^{(m)}]$ we have

$$E_\theta(N) = n - \sum_{r=1}^{z-1} \frac{(n-rm) b_{rm}^{(m)}}{\theta^{rm}} = n - \sum_{r=1}^{z-1} \frac{rm b_{n-rm}^{(m)}}{\theta^{n-rm}} \tag{1.8}$$

2. Asymptotic Theory.

Before embarking on the asymptotic theory we will consider some sequences of numbers which will prove to be useful. From the binomial theorem, for m a non-negative integer we have $\frac{1}{m!} \sum_{r=0}^{\infty} r^{(m)} x^{r-m} = (1-x)^{-(m+1)}$, where $r^{(m)} = \frac{r!}{m!}$. It follows that

$$\sum_{r=0}^{\infty} r^{(m)} x^{r} = \frac{m! x^{m}}{(1-x)^{m+1}} . \tag{2.1}$$

For positive integral powers we have

$$r^{m} = \sum_{s=0}^{m} S(m,s) r^{(s)} \tag{2.2}$$

where the numbers $\{S(m,s)\}$, $m \geq s$, are the Stirling numbers of the second kind. These numbers are tabulated in Table XXII of Fisher and Yates [1] under the title 'Initial Differences of Powers of Natural Numbers.'

We now define a sequence of numbers $\{p_j\}$, where

$$p_m = \sum_{r=0}^{\infty} r^{m} e^{-r} , \quad m=0,1,2,\ldots \tag{2.3}$$

Thus, using (2.2) and (2.1), with $x = \frac{1}{e}$, we have

$$p_m = \sum_{s=0}^{m} S(m,s) \sum_{r=0}^{\infty} r^{(s)} e^{-r} = \sum_{s=0}^{m} \frac{S(m,s) s! e}{(e-1)^{s+1}} .$$

Let $q_m = \frac{p_m}{p_o}$. Thus

$$q_m = \sum_{s=0}^{m} \frac{S(m,s) s!}{(e-1)^{s}} , \quad m=1,2,\ldots \tag{2.4}$$

In particular, $q_1 = \frac{1}{e-1}$ and $q_2 = \frac{1}{e-1} + \frac{2!}{(e-1)^2}$. Later members of sequence $\{q_j\}$ may readily be evaluated using the Fisher and Yates table.

Let us assume that a_n and b_n have power series expansions of the following forms:

$$a_n = \frac{n}{k}\left(1 + \frac{d_1}{n} + \frac{d_2}{n^2} + \dots\right) \qquad (2.5)$$

$$b_n = c\left(\frac{n}{k}\right)^n\left(1 + \frac{e_1}{n} + \frac{e_2}{n^2} + \dots\right) \qquad (2.6)$$

where c, k and the members of the sequences of coefficients $\{d_j\}$ and $\{e_j\}$ are constants.

We are assuming that a_n asymptotically approaches the linear function $\frac{1}{k}(n+d_1)$ of n. If a_n increases approximately as $\frac{n}{k}$ for n large, $b_n = a_n^n - b_1 a_n^{n-1} - b_2 a_n^{n-2} - \dots - b_{n-1} a_n \approx$ $\approx \left(\frac{n}{k}\right)^n\left(1 - \frac{b_1 k}{n} - \frac{b_2 k^2}{n^2} - \dots - \frac{b_{n-1} k^{n-1}}{n^{n-1}}\right)$. Thus, under the assumption that $\frac{d_1}{k}$ is small compared with $\frac{n}{k}$, it seems reasonable to assume the expansion (2.6) for b_n.

We will now find expressions for the constants c, k, d_1, d_2 and e_1.

The equations (1.2) and (1.6) which determine a_n and b_n may be written

$$\sum_{r=0}^{n-1} \frac{b_{n-r}}{a_n^{n-r}} = 1 \qquad (2.7)$$

$$\sum_{r=0}^{n-1} \frac{b_{n-r}}{(a_n+1)^{n-r}} = \alpha \qquad (2.8)$$

Formally substituting the power series expansions for a_n and b_n into (2.7) and (2.8) and equating coefficients of powers of n leads to equations which may be solved in succession. Coefficients of terms of higher order than n^{-3} involve series $\sum_{r=0}^{n-1} r^m e^{-r}$, $m = 0$ to 4, which, when n is large may be replaced, with exponentially small error, by $\sum_{r=0}^{\infty} r^m e^{-r}$. Solving the equations and simplifying leads finally to

$$k = -\log_e \alpha$$

$$d_1 = -\frac{k}{2} - q_1 - b_1\left(\frac{1}{\alpha} - 1\right)$$

$$e = \frac{e^{d_1}}{p_o}$$

$$d_2 = -b_1\left(\frac{1}{\alpha} - 1\right)\left(\frac{q_2}{2} + \frac{d_1^2}{2} + d_1 q_1 + kb_1 - d_1\right) + \frac{kb_1}{\alpha}$$
$$- kb_2\left(\frac{1}{\alpha} - 1\right) - \frac{q_3}{2} - \frac{1}{k}\left[(d_1+k)^2 - d_1^2\right]\left(\frac{3q_2}{4} - \frac{q_1}{2}\right)$$
$$- \frac{1}{k}\left[(d_1+k)^3 - d_1^3\right]\left(\frac{q_1}{2} - \frac{1}{3}\right) - \frac{1}{8k}\left[(d_1+k)^4 - d_1^4\right]$$

$$e_1 = d_2 - \frac{q_2}{2} - \frac{d_1^2}{2} - d_1 q_1 - kb_1$$

Table 2.1 gives the numerical values of d_1 and d_2 correct to 4 and 3 significant figures, respectively, for $\alpha = 0.05$ and 0.01 for the standard procedure P and procedures in which an initial set of $m \geq 2$ observations are taken.

	P		m = 2		m > 2	
α	d_1	d_2	d_1	d_2	d_1	d_2
0.05	-3.080	0.67	-2.080	-4.97	-2.080	-0.25
0.01	-3.885	1.69	-2.885	-4.86	-2.885	0.77

Table 2.1

We will now derive asymptotic expansions for $E_\theta(N)$. Equation (1.7) may be written as follows:

$$E_\theta(N) = n - \sum_{r=0}^{n-1} \frac{rb_{n-r}}{\theta^{n-r}} , \ a_{n-1} < \theta \leq a_n .$$

Thus, if $\theta = a_n - x$, $0 \leq x < a_n - a_{n-1}$

$$E_\theta(N) = n - \sum_{r=0}^{n-1} \frac{rb_{n-r}}{(a_n - x)^{n-r}} . \tag{2.9}$$

From (2.5)

$$a_n - x = \frac{n}{k}\left(1 + \frac{d_1 - kx}{n} + \frac{d_2}{n^2} + \dots\right) . \tag{2.10}$$

Substituting this expansion and the expansion (2.6) for b_n into (2.9), taking the sum from $r = 0$ to ∞ and simplifying leads to

$$E_\theta(N) = n-(kb_1+q_1e^{kx}) + \frac{1}{n}\{[c_1q_1 + \frac{q_3}{2} - d_2q_1 + \frac{q_1}{2}(d_1-kx)^2$$

$$+ q_2(d_1-kx)]e^{kx} + kb_1(d_1-kx)-k^2b_2\} + O(\frac{1}{n^2}) \text{ for } \theta = a_n-x,$$

$$0 \leq x < a_n-a_{n-1} . \quad (2.11)$$

We may also expand $E_\theta(N)$ in powers of $\frac{1}{\theta}$ rather than $\frac{1}{n}$ if θ and n are both large. In this case, if $\theta = a_n$ then $n = k\theta-d_1 + O(\frac{1}{\theta})$ and hence, using (2.11) with $x = 0$ we have

$$E_\theta(N) = k(\theta + \frac{1}{2}) + b_1\{(\frac{1}{\alpha} - 1)-k\} + O(\frac{1}{\theta}) \quad (2.12)$$

For assessing the efficiency of the sequential procedures discussed in this paper, a convenient standard is provided by the optimal fixed sample size procedure; for $\theta > 1$, $\mu(\theta) = \dfrac{\log \alpha}{\log(1 - \frac{1}{\theta})}$ is the least n for which $P_\theta(\hat{X}_n < \theta \leq \hat{X}_n+1) \geq 1-\alpha$. Since $k = \log 1/\alpha$, $\mu(\theta) = k(\theta - \frac{1}{2}) + O(\frac{1}{\theta})$. Thus, using (2.12), $\lim\limits_{\theta\to\infty} \dfrac{\mu(\theta)}{E_\theta(N)} = 1$ so that the sequential procedures may be said to be asymptotically 100 percent efficient, relative to the optimal fixed sample size procedure.

We will now derive asymptotic expansions for $E_\theta(N)$ when the observations are taken in groups of m. It is sufficient to consider only the case in which n is a multiple of m, since if it is not, $a_n^{(m)} = a_{xm}^{(m)}$ where x is the largest integer smaller then $\frac{n}{m}$. Hence, let n be a multiple of m, say $n = zm$. Then, using (1.8)

$$E_\theta(N) = n - \sum_{r=0}^{z-1} \frac{rmb_{n-rm}^{(m)}}{(a_n^{(m)}-x)^{n-rm}} \quad \text{for} \quad \theta = a_n^{(m)}-x,\ 0 \le x < a_n^{(m)}-a_{n-m}^{(m)} \tag{2.13}$$

Expanding the right hand side of (2.13) in powers of n leads to terms of higher order than n^{-2} involving the series $\sum_{r=0}^{z-1} (rm)^j e^{-rm}$, $j = 1,2,3$. However, for fixed m and large z, these series may be replaced, with exponentially small error, by $\sum_{r=0}^{\infty} (rm)^j e^{-rm}$.

We now define a sequence $\{q_j^{(m)}\}$, where $q_j^{(m)} = (\frac{e-1}{e}) \sum_{r=0}^{\infty} (rm)^j e^{-rm}$, $j = 0,1,2,\ldots$. Thus, for example, $q_o^{(m)} = \frac{(e-1)e^{m-1}}{(e^m-1)}$ and $q_1^{(m)} = \frac{m(e-1)e^{m-1}}{(e^m-1)^2}$. Hence we are led to

$$E_\theta(N) = n-e^{kx}\{q_1^{(m)} + \frac{1}{n}[e_1q_1^{(m)} + \frac{q_3^{(m)}}{2} - d_2q_1^{(m)} + \frac{1}{2}(d_1-kx)^2q_1^{(m)}$$

$$+ (d_1-kx)q_2^{(m)}] + O(\frac{1}{n^2})\} \quad \text{for} \quad m > 2,\ \theta = a_n^{(m)}-x, \quad 0 \le x < a_n^{(m)}-a_{n-m}^{(m)}$$

and for $m = 2$

$$E_\theta(N) = 2z-e^{kx}\{q_1^{(2)} + \frac{1}{2z}[e_1q_1^{(2)} + \frac{q_3^{(2)}}{2} - d_2q_1^{(2)} + \frac{1}{2}(d_1-kx)^2q_1^{(2)} + (d_1-kx)q_2^{(2)}$$

$$+ k^2b_2^{(2)}e^{-kx}] + O(\frac{1}{z^2}) \quad \text{for} \quad \theta = a_{2z}^{(2)}-x,\ 0 \le x < a_{2z}^{(2)}-a_{2(z-1)}^{(2)} .$$

In the tables to follow, when a number is given in brackets, it is the exponent to the base 10, of the number immediately preceding it; e.g. 2.101(4) = 21,010.

$n \backslash \alpha$	0.05	0.01
1	2.632(-2)	1.010(-2)
	2.632(-2)	1.010(-2)
	2.880(-1)	1.045(-1)
	5.003(-1)	2.403(-1)
	7.791(-1)	3.978(-1)
	1.065	5.753(-1)
	1.372	7.642(-1)
	1.686	9.617(-1)
	2.007	1.165
10	2.332	1.372
	2.660	1.582
	2.990	1.794
	3.321	2.007
	3.654	2.221
	3.986	2.436
	4.320	2.652
	4.653	2.867
	4.987	3.083
	5.321	3.300
20	5.654	3.516

Table 1: Values of a_n

$n \backslash \alpha$	0.05	0.01
1	2.632(-2)	1.010(-2)
	zero	zero
	1.952(-2)	1.031(-3)
	4.631(-2)	2.944(-3)
	2.197(-1)	8.369(-3)
	1.080	2.964(-2)
	6.714	1.210(-1)
	4.706(1)	5.697(-1)
	3.773(2)	3.016
	3.361(3)	1.776(1)
10	3.307(4)	1.150(2)

Table 2: Values of b_n

3. Comments.

Graybill and Connell [2] have proposed a two-stage sequential procedure for the problem considered in this paper, in which the information of the first sample is ignored once the size of the second sample is determined. However, in practice this approach is not likely to be acceptable. A solution making use of information from the entire sample would seem preferable. The author has worked out details of a two-stage procedure based on the largest observation. This procedure has optimality properties similar to those of the procedure of section 1. This work, together with details of unbiased point estimation of θ based on the procedure of section 1 will appear in a later publication.

The author is indebted to E. W. Bowen for proposing these problems and for many helpful discussions.

REFERENCES

[1] Fisher, R. A. and Yates, F. (1963). Statistical Tables for Biological, Agricultural and Medical Research, (6^{th} ed.), Oliver and Boyd, London.

[2] Graybill, F. A. and Connell, T. L. (1964). Sample Size Required to Estimate the Parameter in the Uniform Density Within d Units of the True Value, J. Am. Stat. Assoc., 59, 550-556.

[3] Lehmann, E. L. (1949). Notes on the Theory of Estimation, lecture notes recorded by Colin Blyth, Univ. of Calif, Berkeley, Calif.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Statistics Department Stanford University Stanford, California	2a. REPORT SECURITY CLASSIFICATION 2b. GROUP

3. REPORT TITLE

SEQUENTIAL ESTIMATION IN THE UNIFORM DENSITY

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (Last name, first name, initial)

COOKE, Peter J.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
April 21, 1969	15	3

8a. CONTRACT OR GRANT NO. Nonr-225(52) b. PROJECT NO. NR-342-022 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 144 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

A solution is given to the problem of sequentially estimating (to within $d > 0$ units of the true value) the parameter θ of a uniform distribution on $(0,\theta)$.

The solution is optimal in the sense that the expected number of observations is minimized. It is also minimax in that the maximum possible number of observations is minimized.

Some tables are appended.

DD FORM 1 JAN 64 1473

Unclassified

Security Classification

Unclassified

Security Classification

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
sequential						
parameter						
uniform						
confidence function						
asymptotic						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (corporate author) issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year, or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (*either by the originator or by the sponsor*), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (*paying for*) the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS), (S), (C), or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.

DD FORM 1 JAN 64 1473 (BACK)

Unclassified

Security Classification