SEQUENTIAL ESTIMATION IN THE UNIFORM DENSITY, II

BY

PETER J. COOKE

TECHNICAL REPORT NO. 175
MAY 18, 1971

THIS RESEARCH WAS SPONSORED BY THE ARMY RESEARCH OFFICE
OFFICE OF NAVAL RESEARCH, AND AIR FORCE OFFICE OF
SCIENTIFIC RESEARCH BY CONTRACT NO.
N00014-67-A-0112-0053 (NR-042-267)

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
SEQUENTIAL ESTIMATION IN THE UNIFORM DENSITY, II

by

PETER J. COOKE

TECHNICAL REPORT NO. 175
May 18, 1971

PREPARED UNDER CONTRACT N00014-67-A-0112-0053
NR-042-267
OFFICE OF NAVAL RESEARCH

Herbert Solomon, Project Director

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

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
SEQUENTIAL ESTIMATION IN THE UNIFORM DENSITY, II

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 use a two-stage sequential procedure to estimate the unknown parameter $\theta$ by an interval of length at most $\delta$ units and with confidence at least $1-\alpha$, for some specified $\delta > 0$ and $\alpha$ in $(0,1)$.

The procedure proposed in this paper satisfies an admissibility criterion which may be stated in terms of the maximum possible number of observations or, alternatively, the expected number of observations.

In sections 4 and 5 a sequential procedure is described for constructing unbiased point estimates of $\theta$ which are based on the maximum of the observations.

1. Introduction.

Suppose we observe $X_1, X_2, \ldots$ from the uniform density on $(0, \theta)$ and wish to use a two-stage sequential procedure to estimate the parameter $\theta$ by a $1-\alpha$ level confidence interval of length at most $\delta$ units, for some specified $\delta > 0$ and $\alpha$ in $(0,1)$.

As in [1], we may suppose $\delta = 1$, since for any other positive $\delta$, $\frac{X_i}{\delta}$ is uniformly distributed over $(0, \frac{\theta}{\delta})$. Thus we may consider the problem as one of estimating $\frac{\theta}{\delta}$ by an interval of at most unit length.
2. The Problem.

The two-stage sequential procedures we shall consider in this paper depend on a nondecreasing sequence of positive constants \( h_0, h_1, h_2, \ldots \) and are of the form:

1. For some fixed positive integer \( m \), observe \( \hat{X}_m \), the maximum of \( X_1, X_2, \ldots, X_m \).

2. Calculate a function \( \nu(\hat{X}_m) \), the number of additional observations to be taken, where \( \nu(x) = n \) if \( h_{n-1} < x \leq h_n \).

3. Make the statement \( \hat{X}_N < \theta \leq \hat{X}_N + 1 \), where \( N = m + \nu(\hat{X}_m) \).

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

We will show how to choose \( h_0, h_1, h_2, \ldots \) so that \( \gamma(\theta) = P(\hat{X}_N < \theta \leq \hat{X}_N + 1) \geq 1 - \alpha \) for every \( \theta \). Clearly, for any reasonable procedure \( h_n \to \infty \) as \( n \to \infty \).

The distribution of \( \hat{X}_N \) may be determined as follows:

\[
P_{\theta} (\nu(\hat{X}_m) = n) = P_{\theta} (h_{n-1} < \hat{X}_m \leq h_n)
\]

\[
= \begin{cases} 
0 & \theta \leq h_{n-1} \\
\frac{\theta^{m-h_m}}{\theta^{n-1}} & h_{n-1} < \theta \leq h_n \\
\frac{h_n^{m-h_{n-1}}}{\theta^{m}} & \theta > h_n
\end{cases} \quad (2.1)
\]


Also,
\[ P_{\theta}(\hat{X}^m(y) = \begin{cases} 1, & y \geq \theta \\ \left(\frac{y}{\theta}\right)^m, & 0 < y < \theta \end{cases} \] (2.2)

and
\[ P_{\theta}(\hat{X}_{m+n} \leq x | \hat{X}_m = y) = \begin{cases} 0, & x < y \\ \left(\frac{x}{\theta}\right)^n, & x \geq y \end{cases} \] (2.3)

Therefore, for \( x \leq \), we have
\[ P_{\theta}(\hat{X}_{m+n} \leq x, N = m+n) = P_{\theta}(\hat{X}_{m+n} \leq x, h_{n-1} < \hat{X}_m \leq h_n) \]
\[ = \int_{h_{n-1}}^{h_n} P_{\theta}(\hat{X}_{m+n} \leq x | \hat{X}_m = y) dP_{\theta}(\hat{X}_y) \]
\[ = \left(\frac{x}{\theta}\right)^n \int_{h_{n-1}}^{h_n} \frac{my^{m-1}}{\theta^m} dy \] (2.4)

\[ = \begin{cases} 0, & x < h_{n-1} \\ \left(\frac{x}{\theta}\right)^n \frac{\left(\frac{x}{h_{n-1}}\right)^m}{\theta^m}, & h_{n-1} < x \leq h_n \\ \left(\frac{x}{\theta}\right)^n \frac{\left(\frac{h_n}{h_{n-1}}\right)^m}{\theta^m}, & x \geq h_n \end{cases} \]

Thus, for \( x \leq \theta \),
\[ P_\theta(\hat{X}_N \leq x) = \sum_r P_\theta(\hat{X}_{m+r} \leq x, \, N = m+r) \]
\[ = \sum_{r=0}^{v(x)-1} \frac{r^m}{\theta^m} \left( \frac{x-h}{\theta^r} \right)^{r-1} + \left( \frac{x-h}{\theta^r} \right)^{v(x)-1} \]

where \( h_{-1} = 0 \).

Using (2.5) we may write the error probability, \( \alpha(\theta) \), in the following way: for \( h_{n-1} < \theta-1 < h_n \),

\[ \alpha(\theta) = P(\hat{X}_N \leq \theta-1) \]
\[ = \sum_{r=0}^{n-1} \frac{(h_{r+1}^{m}-h_r^{m})}{\theta^m} \left( \frac{\theta-1}{\theta} \right)^{r} + \frac{(\theta-1)^{m}-h_{n-1}^{m}}{\theta^m} \left( \frac{\theta-1}{\theta} \right)^{n} \]
\[ = \theta^{-(m+1)} \sum_{r=0}^{n-1} h_{r}^{m} \left( \frac{\theta-1}{\theta} \right)^{r} + \left( \frac{\theta-1}{\theta} \right)^{m+n} . \]  

(2.6)

The requirement \( \gamma(\theta) = P(\hat{X}_N < \theta < \hat{X}_N + 1) \geq 1-\alpha \) is equivalent to \( \alpha(\theta) \leq \alpha \) for all \( \theta \). For \( \theta < 1 \), \( \alpha(\theta) = 0 \) so that \( \gamma(\theta) = 1 \). The admissibility criterion we shall adopt for the two-stage procedures of this paper is similar to the one used for the procedures of [1] and is as follows: considering only two-stage procedures for some fixed \( m \geq 1 \), of all such procedures for which \( \alpha(\theta) \leq \alpha \) for all \( \theta \), a procedure is admissible if every other procedure with smaller \( v \)-function for some \( \theta \) has larger \( v \)-function for at least one \( \theta' < \theta \). As in [1], the solution to be investigated will easily be seen to also satisfy an admissibility criterion of the same form expressed in terms of the expected number of observations rather than the \( v \)-function.
By using arguments similar to those given in [1], we find that because of our optimality criterion, for each \( n \) we choose \( h_n \) as large as possible, i.e., \( h_n \) is the largest root of

\[
(x+1)^{-m-1} \sum_{r=0}^{n-1} h_r^m \left( \frac{x}{x+1} \right)^r + \left( \frac{x}{x+1} \right)^{m+n} = \alpha .
\]  

(2.7)

Values (correct to 4 significant figures) of \( h_n \) for \( n = 0 \) to 15 and \( \alpha = .05 \) and .01 in the cases \( m = 5 \) and 10 are given in table 1.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \alpha = .05 )</th>
<th>( \alpha = .01 )</th>
<th>( \alpha = .05 )</th>
<th>( \alpha = .01 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.219</td>
<td>.6614</td>
<td>2.862</td>
<td>1.710</td>
</tr>
<tr>
<td></td>
<td>1.219</td>
<td>.6614</td>
<td>2.863</td>
<td>1.710</td>
</tr>
<tr>
<td></td>
<td>1.685</td>
<td>.9542</td>
<td>3.363</td>
<td>2.032</td>
</tr>
<tr>
<td></td>
<td>1.804</td>
<td>1.020</td>
<td>3.547</td>
<td>2.148</td>
</tr>
<tr>
<td></td>
<td>2.183</td>
<td>1.265</td>
<td>3.925</td>
<td>2.393</td>
</tr>
<tr>
<td></td>
<td>2.532</td>
<td>1.358</td>
<td>4.172</td>
<td>2.549</td>
</tr>
<tr>
<td></td>
<td>2.694</td>
<td>1.581</td>
<td>4.508</td>
<td>2.767</td>
</tr>
<tr>
<td></td>
<td>2.889</td>
<td>1.689</td>
<td>4.781</td>
<td>2.940</td>
</tr>
<tr>
<td></td>
<td>3.210</td>
<td>1.901</td>
<td>5.099</td>
<td>3.146</td>
</tr>
<tr>
<td></td>
<td>3.421</td>
<td>2.018</td>
<td>5.384</td>
<td>3.328</td>
</tr>
<tr>
<td></td>
<td>3.729</td>
<td>2.222</td>
<td>5.694</td>
<td>3.528</td>
</tr>
<tr>
<td></td>
<td>3.951</td>
<td>2.346</td>
<td>5.985</td>
<td>3.714</td>
</tr>
<tr>
<td></td>
<td>4.250</td>
<td>2.545</td>
<td>6.290</td>
<td>3.911</td>
</tr>
<tr>
<td></td>
<td>4.479</td>
<td>2.673</td>
<td>6.584</td>
<td>4.099</td>
</tr>
<tr>
<td></td>
<td>4.772</td>
<td>2.868</td>
<td>6.887</td>
<td>4.295</td>
</tr>
<tr>
<td>15</td>
<td>5.007</td>
<td>2.999</td>
<td>7.183</td>
<td>4.484</td>
</tr>
</tbody>
</table>

Table 1: Values of \( h_n \).
3. The Expected Number of Observations.

For the two-stage sequential procedure described in section 2, the expected number of additional observations taken is given by

\[ E_\theta(N-m) = \sum_r rP_\theta(N-m=r) \]

\[ = \sum_r rP_\theta(h_{r-1} < x_m \leq h_r) \]

\[ = \frac{(h_m^m - h_0^m)}{\theta^m} + \frac{2(h_2^m - h_1^m)}{\theta^m} + \ldots \]

\[ + \frac{(n-1)(h_n^m - h_{n-2}^m)}{\theta^m} + n(1 - \frac{h_{n-1}^m}{\theta^m}) \text{ for } h_{n-1} < \theta \leq h_n. \]

Thus, (3.1) gives

\[ E(N) = m+n - \frac{1}{\theta^m} \sum_{r=0}^{n-1} h_r^m \text{ for } h_{n-1} < \theta \leq h_n. \]


Suppose we sequentially sample \( x_1, x_2, \ldots \) from the uniform density on \((0, \theta)\) and wish to find an unbiased point estimate of \( \theta \). In [1], sequential sampling procedures are described for estimating \( \theta \) by an interval of fixed length. These procedures depend on a nondecreasing sequence of positive constants \( a_1, a_2, \ldots \) and are of the form:

1. Observe \( x_1, x_2, \ldots \) until for the first time \( \hat{x}_N \leq a_N \), where \( \hat{x}_N \) is the maximum of \( x_1, x_2, \ldots, x_N \).
In [1] it is shown how to choose the sequence $a_1, a_2, \ldots$ to provide a procedure which is admissible in a sense there defined when the terminal statement is $\hat{X}_N < \theta \leq \hat{X}_N + 1$.

5. The Estimator.

We will use the stopping rule (1), but our terminal decision will be '$\theta = p(\hat{X}_N)$', where $p(x)$ is a nondecreasing function of $x$. In what follows we will adopt the notation of [1].

The density function of $\hat{X}_N$ (see [1], (2.4)) is as follows:

for $x \leq \theta$,

$$f_{\theta}(x) = \frac{B_n'(x)}{\theta^n}, \quad a_{n-1} < x \leq a_n$$

(5.1)

where

$$B_n'(x) = nx^{n-1} - (n-1)b_n x^{n-2} - (n-2)b_{n-1} x^{n-3} - \ldots - b_{n-1}$$

(5.2)

For $p(\hat{X}_N)$ to be an unbiased estimate of $\theta$, its expected value must equal $\theta$. Thus, if $\nu(\theta) = n$, we must have

$$E_{\theta} p(\hat{X}_N) = \frac{1}{\theta} \int_0^{a_1} B_1'(x)p(x)dx + \frac{1}{\theta^2} \int_{a_1}^{a_2} B_2'(x)p(x)dx + \ldots$$

$$+ \frac{1}{\theta^n} \int_{a_{n-1}}^{\theta} B_n'(x)p(x)dx = \theta.$$  

(5.3)
Suppose we define a sequence of functions \( \{\rho_n(\theta)\} \) by

\[
\rho_n(\theta) = \int_{a_{n-1}}^{\theta} B_n'(x)p(x)dx , \quad n=1,2,\ldots .
\] (5.4)

Then, using (5.3), we have

\[
\rho_n(\theta) = \theta^{n+1} - \rho_1(a_1)\theta^n - \rho_2(a_2)\theta^{n-1} - \ldots - \rho_{n-1}(a_{n-1})\theta .
\] (5.5)

Differentiating both sides of (5.4) with respect to \( \theta \) gives

\[
p(x) = \frac{\rho_n'(x)}{B_n'(x)} \quad \text{for} \quad a_{n-1} < x \leq a_n
\] (5.6)

where \( B_n'(x) \) is given by (5.2) and

\[
\rho_n'(x) = (n+1)x^n - n\rho_1(a_1)x^{n-1} - (n-1)\rho_2(a_2)x^{n-2} - \ldots - \rho_{n-1}(a_{n-1}) .
\] (5.7)

Thus, the estimator \( p(\hat{x}_N) \) is increasing, continuous and a rational algebraic function for every \( \hat{x}_N \) in the intervals \( (a_{n-1}, a_n] \), \( n=1,2,\ldots \) where \( a_0 \equiv 0 \). The functions \( \rho_n(\theta) \) are easily determined by \( \rho_1(\theta) = 2\theta \) and the recurrence relation (5.5).

REFERENCES

A two-stage sequential procedure is proposed for finding bounded length confidence intervals for the parameter of the uniform distribution on \((0,\theta)\). The procedure satisfies admissibility criteria in terms of the maximum number of observations required and the expected number of observations.

Unbiased point estimation of the parameter \(\theta\) using sequential methods is also discussed.
### UNCLASSIFIED

**Security Classification**

<table>
<thead>
<tr>
<th>KEY WORDS</th>
<th>LINK A</th>
<th>LINK B</th>
<th>LINK C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-stage</td>
<td>ROLE</td>
<td>WT</td>
<td>ROLE</td>
</tr>
<tr>
<td>Sequential</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Admissibility</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unbiased</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**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 number, 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 (T5), (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