

SEQUENTIAL PROCEDURES FOR A CLASS OF DISTRIBUTIONS RELATED TO THE UNIFORM

P. J. COOKE AND M. K. VAGHOLKAR

(Received Nov. 2, 1979; revised Feb. 14, 1981)

Summary

The main purpose of this paper is to examine some sequential inference procedures, mainly in the one-sided hypothesis testing situation, for the parameter of a class of distributions related to the Uniform distribution on $(0, \theta)$. The procedures are all based on the sequence of maxima of independent and identically distributed random variables and, for the case of a single (upper) boundary, procedures which are optimal in the sense of minimizing the average sample size are discussed. The impossibility of using two boundaries is demonstrated, thus leaving the best procedure with one boundary as the optimal procedure. The novelty of this procedure is that it uniformly minimizes the average sample size.

1. Introduction

Suppose X_1, X_2, \dots are independent random variables, each with cumulative distribution function $F(x)$ and density $f(x)$. Let $Y_n = \max(X_1, X_2, \dots, X_n)$, $n=1, 2, \dots$ and define a stopping rule N by

$$(1) \quad N = \text{first integer } n \geq 1 \text{ for which } Y_n \notin (b_n, a_n),$$

where $\{a_n\}$ and $\{b_n\}$ are sequences of real numbers typically non-decreasing since $\{Y_n\}$ is stochastically increasing with n . Then, for $n \geq 2$

$$\begin{aligned} P(N > n) &= P(b_1 < Y_1 < a_1, \dots, b_{n-1} < Y_{n-1} < a_{n-1}, b_n < Y_n < a_n) \\ &= P(b_i < Y_i < a_i, i=1, 2, \dots, n-1) \\ &\quad \cdot P(b_n < Y_n < a_n | b_i < Y_i < a_i, i=1, 2, \dots, n-1) \\ &= P(N > n-1) P(b_n < Y_n < a_n | b_{n-1} < Y_{n-1} < a_{n-1}) \end{aligned}$$

since the sequence $\{Y_n\}$ is Markov. Thus induction gives

$$P(N > n) = P(N > 1) \prod_{i=2}^n P(b_i < Y_i < a_i | b_{i-1} < Y_{i-1} < a_{i-1}).$$

For $b_i > a_{i-1}$, $P(b_i < Y_i < a_i | b_{i-1} < Y_{i-1} < a_{i-1}) = P(b_i < X_i < a_i) = F(a_i) - F(b_i)$, though typically $b_i \leq a_{i-1}$ for each $i > 1$. Henceforth we will make this assumption, in which case

$$P(b_{i-1} < Y_{i-1} < a_{i-1}) = F^{i-1}(a_{i-1}) - F^{i-1}(b_{i-1}),$$

$$\begin{aligned} P(b_i < Y_i < a_i, b_{i-1} < Y_{i-1} < a_{i-1}) \\ = F(a_i)[F^{i-1}(a_{i-1}) - F^{i-1}(b_{i-1})] - F(b_i)[F^{i-1}(b_i) - F^{i-1}(b_{i-1})] \end{aligned}$$

and, for $i \geq 2$,

$$\begin{aligned} P(b_i < Y_i < a_i | b_{i-1} < Y_{i-1} < a_{i-1}) \\ = F(a_i) - F(b_i) \left[\frac{F^{i-1}(b_i) - F^{i-1}(b_{i-1})}{F^{i-1}(a_{i-1}) - F^{i-1}(b_{i-1})} \right]. \end{aligned}$$

Thus, since $P(N > 1) = F(a_1) - F(b_1)$, for $n \geq 2$

$$(2) \quad P(N > n) = [F(a_1) - F(b_1)] \prod_{i=2}^n \left\{ F(a_i) - F(b_i) \left[\frac{F^{i-1}(b_i) - F^{i-1}(b_{i-1})}{F^{i-1}(a_{i-1}) - F^{i-1}(b_{i-1})} \right] \right\}.$$

Suppose we consider procedures truncated at some integer M , say, in which case $a_M = b_M$. (This possible violation of our assumption for the pair a_{M-1}, b_M does not effect any of our calculations.) We can now find expressions for any of the moments of N using (2). In particular,

$$\begin{aligned} (3) \quad E(N) &= \sum_{n=0}^{M-1} P(N > n) \\ &= 1 + F(a_1) - F(b_1) + [F(a_1) - F(b_1)] \\ &\quad \cdot \sum_{n=2}^{M-1} \prod_{i=2}^n \left\{ F(a_i) - F(b_i) \left[\frac{F^{i-1}(b_i) - F^{i-1}(b_{i-1})}{F^{i-1}(a_{i-1}) - F^{i-1}(b_{i-1})} \right] \right\}. \end{aligned}$$

Suppose further that F depends on a single unknown parameter θ (and hence we attach a subscript θ to F) and that H_0 is a hypothesis about the value of θ to be tested using a sequential procedure with stopping rule of the form (1) and with decision rule

$$(4) \quad \text{reject } H_0 \text{ if } Y_N \geq a_N, \quad \text{accept } H_0 \text{ if } Y_N \leq b_N.$$

For such a procedure, for $n \geq 2$,

$$\begin{aligned} P_\theta(N = n, \text{ reject } H_0) &= P_\theta(b_1 < Y_1 < a_1, \dots, b_{n-1} < Y_{n-1} < a_{n-1}, Y_n \geq a_n) \\ &= P_\theta(b_1 < Y_1 < a_1, \dots, b_{n-1} < Y_{n-1} < a_{n-1}, X_n \geq a_n) \\ &= P_\theta(N > n-1) P_\theta(X_n \geq a_n). \end{aligned}$$

Hence, using (2), the procedure has power function

$$\begin{aligned}
 (5) \quad \beta(\theta) &= P_\theta(\text{reject } H_0) \\
 &= \sum_{n=1}^M P_\theta(N=n, \text{reject } H_0) \\
 &= 1 - F_\theta(a_1) + [F_\theta(a_1) - F_\theta(b_1)][1 - F_\theta(a_2)] \\
 &\quad + [F_\theta(a_1) - F_\theta(b_1)] \sum_{n=3}^M [1 - F_\theta(a_n)] \\
 &\quad \cdot \prod_{i=2}^{n-1} \left\{ F_\theta(a_i) - F_\theta(b_i) \left[\frac{F_\theta^{i-1}(b_i) - F_\theta^{i-1}(b_{i-1})}{F_\theta^{i-1}(a_{i-1}) - F_\theta^{i-1}(b_{i-1})} \right] \right\}.
 \end{aligned}$$

If the decisions to reject and accept H_0 are reversed in (4), then of course the power function is one minus the expression in (5). The expected sample size, which we now denote by $E_\theta(N)$, is given by (3) with F replaced by F_θ . Equations (3) and (5) hold with $M = \infty$ for untruncated procedures.

Suppose now that the density corresponding to $F_\theta(x)$ is $f_\theta(x)$, where $f_\theta(x) = c(\theta)g(x)$ for $x \in (0, \theta)$ and $f_\theta(x) = 0$ otherwise and $g(x)$ is a known function. This includes the case in which $f_\theta(x)$ is a specified density $g(x)$ truncated at the unknown point $x = \theta$. If $G(x) = \int_0^x g(u)du$, then $c(\theta) = 1/G(\theta)$.

For the class of distributions $\{f_\theta(x), \theta > 0\}$, $Y_n = \max(X_1, X_2, \dots, X_n)$ is sufficient for θ for a fixed sample size procedure with sample size n and hence by Fay's lemma (see, for example, Lehmann [1]), Y_n and N are jointly sufficient for θ for a sequential procedure based on the stopping rule N defined in (1). Without loss of generality we will restrict the discussion henceforth to the Uniform distribution on $(0, \theta)$ since, if X has density $f_\theta(x)$, then the random variable $G(X)$ has the Uniform density on $(0, G(\theta))$.

2. Optimal procedures

We will restrict the discussion to follow to tests of a hypothesis of the form $H_0: \theta \leq \theta_0$ against the alternative $H_1: \theta > \theta_0$, where θ_0 is a specified constant. It is well known (see, for example, Lehmann [2]) that for random samples of fixed size n from the Uniform distribution on $(0, \theta)$, every size α test based on Y_n is uniformly most powerful for testing H_0 against H_1 . The procedure which minimizes the power function uniformly for $\theta \leq \theta_0$ is the one with critical region $Y_n \geq \theta_0(1-\alpha)^{1/n}$. This procedure has power function $\beta(\theta) = 1 - (1-\alpha)(\theta_0/\theta)^n$ for $\theta \geq \theta_0(1-\alpha)^{1/n}$ and $\beta(\theta) = 0$ if $\theta \leq \theta_0(1-\alpha)^{1/n}$. The sample size required for $\beta(\theta_1) = 1 - \beta$ for some specified $\theta_1 > \theta_0$ and $\beta \in (0, 1)$ is

$$n(\alpha, \beta) = \log \left(\frac{1-\alpha}{\beta} \right) / \log \left(\frac{\theta_1}{\theta_0} \right).$$

To avoid some algebraic complications in the discussion to follow we will ignore the fact that $n(\alpha, \beta)$ is not usually an integer.

The question of whether or not we can improve on the best fixed sample size procedure through a sequential sampling plan now arises. Thus we seek a sequential test with size α and power function equal to $1-\beta$ at $\theta=\theta_1$ and with expected sample size not larger than $n(\alpha, \beta)$ for any θ and smaller than $n(\alpha, \beta)$ for some θ . The first procedure to come to mind is the Sequential Probability Ratio Test (SPRT) of $H_0: \theta=\theta_0$ versus $H_1: \theta=\theta_1$ with error probabilities (α, β) . However, the likelihood ratio is constant for $x<\theta_0$, so that the SPRT procedure reduces to: reject H_0 as soon as you observe an X_i greater than or equal to θ_0 , otherwise accept H_0 after $\log(1/\beta)/\log(\theta_1/\theta_0)$ observations. Thus the SPRT necessarily has size $\alpha=0$. Its optimality property is preserved; that is, among all tests with size $\alpha=0$, the expected sample size is minimized at both $\theta=\theta_0$ and $\theta=\theta_1$. Clearly, the expected sample size equals $n(0, \beta)$ for $\theta\leq\theta_0$, but is smaller than $n(0, \beta)$ for $\theta>\theta_0$. Tests with size zero are of limited interest and what we seek are sequential tests, with size $\alpha>0$, which are better than the corresponding fixed sample size procedure and which are in some sense optimal.

Samuel-Cahn [3] has considered truncated sequential procedures based on a single sequence $\{a_n\}$ with $a_1\leq a_2\leq\cdots\leq a_M\leq\theta_0$. Thus the stopping and decision rules are as in (1) and (4), respectively, except that $b_1=b_2=\cdots=b_{M-1}=0$ and $b_M=a_M$. The procedure is therefore to take no more than M observations and accept H_0 if and only if the boundary is not reached. From (3) and (5), expressions for the expected sample size and power function of such a procedure reduce to the following:

$$E_\theta(N) = 1 + \sum_{n=1}^{M-1} \prod_{i=1}^n \left(\frac{a_i}{\theta} \right), \quad \theta \geq \theta_0,$$

$$\beta(\theta) = 1 - \prod_{i=1}^M \left(\frac{a_i}{\theta} \right), \quad \theta \geq \theta_0.$$

It follows that if $\beta(\theta_0)=\alpha$ and $\beta(\theta_1)=1-\beta$, then $\prod_{i=1}^M a_i=(1-\alpha)\theta_0^M$, $M=n(\alpha, \beta)=\log((1-\alpha)/\beta)/\log(\theta_1/\theta_0)$ and $\beta(\theta)=1-(1-\alpha)(\theta_0/\theta)^M$ for $\theta\geq\theta_0$. Thus the truncation point equals the number of observations required by the best fixed sample size procedure and the power functions coincide for $\theta\geq\theta_0$.

Samuel-Cahn has proved that the procedure with boundary $a_1=(1-\alpha)\theta_0$, $a_2=a_3=\cdots=a_M=\theta_0$ is optimal in the sense that $E_\theta(N)$ is minimized uniformly for $\theta\geq\theta_0$. Thus a_1 takes care of the requirement that the size is $\alpha>0$ and after that the procedure is of the same form as the SPRT. Indeed, when $\alpha=0$ they are the same. However, despite

the proven optimality of the above procedure, the heavy emphasis on X_1 might make a potential user of the scheme somewhat wary of it. Thus in Section 3 we consider a broader class of stopping rules. Before doing so, suppose we attempt to formulate the fixed width confidence interval problem in terms of a stopping rule of the form used by Samuel-Cahn for the hypothesis testing problem. We see immediately that no matter how the a_i 's are chosen, $P_\theta(N=\infty) > 0$ for some values of θ . Truncation seems the obvious next step, but with no knowledge of θ it is impossible to determine a truncation value. The only case in which this problem can be solved with a Samuel-Cahn type stopping rule is when θ is known to be smaller than some constant θ_0 , say, in which case we can determine a truncation value M (depending on θ_0 of course). For purposes of illustration, and indeed without loss of generality, suppose we consider the unit length confidence interval (Y_N, Y_N+1) . We find

$$\begin{aligned}
 (6) \quad P_\theta(Y_N < \theta - 1) &= \left(1 - \frac{1}{\theta}\right)^M, \quad 1 \leq \theta \leq a_1 + 1 \\
 &= 1 - \frac{1}{\theta} - \frac{1}{\theta} \sum_{n=2}^k \prod_{i=1}^{n-1} \frac{a_i}{\theta} - \left[1 - \left(1 - \frac{1}{\theta}\right)^{M-k}\right] \prod_{i=1}^k \frac{a_i}{\theta}, \\
 &\quad a_k + 1 < \theta \leq a_{k+1} + 1, \quad k = 1, 2, \dots, M-1,
 \end{aligned}$$

where the sum is to be taken as zero for $k=1$.

From (6), given that we will not consider any boundary points larger than $\theta_0 - 1$, we find that $P_\theta(Y_N < \theta - 1) \leq \alpha$ for all $\theta \leq \theta_0$ if and only if $a_1 = a_2 = \dots = a_M = \theta_0 - 1$ and $M = \log \alpha / \log(1 - (1/\theta_0))$, the sample size required by the best fixed sample size procedure. Also,

$$E_\theta(N) = \begin{cases} M, & \theta \leq \theta_0 - 1 \\ \theta(1 - \alpha) - 1, & \theta_0 - 1 < \theta \leq \theta_0. \end{cases}$$

Thus the above formulation leads to a procedure which is a slim improvement over the best fixed sample size procedure since all it does is take care of the obvious defect of that procedure in that if you observe an $X_i > \theta_0 - 1$ you might as well stop sampling since you now have an interval which contains θ with probability one.

3. The two boundary approach

The question now is whether or not we can improve on the above optimal solution, in the sense of reducing the average sample size, by using a stopping rule of the form (1) and a decision rule of the form (4) with at least one of b_1, b_2, \dots, b_{M-1} nonzero. If we require the power function to satisfy $\beta(\theta_0) = \alpha$ and $\beta(\theta_1) = 1 - \beta$, the truncation point must

again be $M=n(\alpha, \beta)$ since, if the procedure is truncated at some M and if the acceptance region for H_0 is A , then A is a subset of the M -dimensional cube with sides of length θ_0 and $P_{\theta_0}(Y_N \in A) = \text{Vol}(A)/\theta_0^M = 1-\alpha$. Also, for $\theta \geq \theta_0$, $P_{\theta}(Y_N \in A) = \text{Vol}(A)/\theta^M = (1-\alpha)(\theta_0/\theta)^M$, so that $\beta(\theta) = 1 - (1-\alpha)(\theta_0/\theta)^M$ for $\theta \geq \theta_0$. Hence $\beta(\theta_1) = 1-\beta$ implies $M=n(\alpha, \beta)$.

It is not difficult to see that, as in the fixed sample size problem, not only do all size α tests have the same power function for $\theta \geq \theta_0$, but the power function is uniformly minimized by choosing the critical region as far as possible from the origin. By the theorem which follows this paragraph, this amounts to choosing a_1, a_2, \dots, a_M such that $\prod_{i=1}^M a_i = (1-\alpha)\theta_0^M$, in which case b_1, b_2, \dots, b_{M-1} are all necessarily zero. It is not our main concern here to discuss in detail the procedure which uniformly minimizes $\beta(\theta)$ for $\theta < \theta_0$, though we easily find that the procedure is the one with $b_1=b_2=\dots=b_{M-1}=0$ and $a_1=a_2=\dots=a_M=\theta_0(1-\alpha)^{1/M}$ since the power function for every procedure is zero for $\theta \leq a_1$ and, in view of the theorem and since $a_1 \leq a_2 \leq \dots \leq a_M$, the largest possible value of a_1 is $\theta_0(1-\alpha)^{1/M}$.

THEOREM. *For every sequential test of $H_0: \theta \leq \theta_0$ against $H_1: \theta > \theta_0$ based on a stopping rule of the form (1) and a decision rule of the form (4) and with power function $\beta(\theta)$ satisfying both $\beta(\theta_0) = \alpha$ and $\beta(\theta_1) = 1-\beta$ for some $\alpha \in (0, 1)$ and $\beta \in (0, 1)$,*

(i)
$$\prod_{i=1}^M a_i \leq (1-\alpha)\theta_0^M,$$

(ii)
$$\prod_{i=1}^M a_i = (1-\alpha)\theta_0^M \text{ if and only if } b_1=b_2=\dots=b_{M-1}=0.$$

PROOF. $\{X_1 < a_1, X_2 < a_2, \dots, X_M < a_M\} \iff \{Y_1 < a_1, Y_2 < a_2, \dots, Y_M < a_M\} \implies \{\text{accept } H_0\}$. Thus $P_{\theta}(X_1 < a_1, X_2 < a_2, \dots, X_M < a_M) = \prod_{i=1}^M \left(\frac{a_i}{\theta}\right) \leq 1-\beta(\theta)$

and (i) follows by putting $\theta = \theta_0$.

From (5) with $F_{\theta}(x) = x/\theta$ we have

$$\beta(\theta) = 1 - \frac{a_1}{\theta} + \frac{(a_1-b_1)(\theta-a_2)}{\theta^2} + \frac{(a_1-b_1)}{\theta} \sum_{n=3}^M \frac{(\theta-a_n)}{\theta^{n-1}} \prod_{i=2}^{n-1} \left\{ a_i - b_i \left[\frac{b_i^{i-1} - b_{i-1}^{i-1}}{a_{i-1}^{i-1} - b_{i-1}^{i-1}} \right] \right\}$$

and, if we define $a_0=1, b_0=0$ and $b_0^0=0$ and arrange the terms in increasing powers of $1/\theta$ we have

$$(7) \quad \beta(\theta) = 1 - \frac{b_1}{\theta} - \sum_{n=2}^{M-1} \frac{b_n}{\theta^n} \left[\frac{b_n^{n-1} - b_{n-1}^{n-1}}{a_{n-1}^{n-1} - b_{n-1}^{n-1}} \right] \prod_{i=1}^{n-1} \left\{ a_i - b_i \left[\frac{b_i^{i-1} - b_{i-1}^{i-1}}{a_{i-1}^{i-1} - b_{i-1}^{i-1}} \right] \right\} - \frac{a_M}{\theta^M} \prod_{i=1}^{M-1} \left\{ a_i - b_i \left[\frac{b_i^{i-1} - b_{i-1}^{i-1}}{a_{i-1}^{i-1} - b_{i-1}^{i-1}} \right] \right\}.$$

It is now easy to see that if $b_1=b_2=\dots=b_{M-1}=0$, then $\prod_{i=1}^M a_i = (1-\alpha)\theta_0^M$

since $\beta(\theta_0)=\alpha$.

On the other hand, suppose at least one of b_1, b_2, \dots, b_{M-1} is strictly positive. Then, for $\theta \geq \theta_0$, the power function of the procedure is smaller than the power function of a procedure with the same upper boundary, but with each point in the lower boundary equal to zero. That is, from (7)

$$\beta(\theta) < 1 - \frac{1}{\theta^M} \prod_{i=1}^M a_i \quad \text{for } \theta \geq \theta_0.$$

Thus, if $\prod_{i=1}^M a_i = (1-\alpha)\theta_0^M$, then $\beta(\theta_0) < \alpha$. This contradicts the assumption that $\beta(\theta_0) = \alpha$ and hence, from (i), $\prod_{i=1}^M a_i < (1-\alpha)\theta_0^M$. The proof is now complete.

We can now prove that it is not possible to have a nonzero lower boundary for the problem considered here. If we let

$$c_n = b_n \left[\frac{b_n^{n-1} - b_{n-1}^{n-1}}{a_n^{n-1} - b_{n-1}^{n-1}} \right], \quad n = 1, 2, \dots, M-1,$$

$d_1 = b_1$ and $d_n = c_n \prod_{i=1}^{n-1} (a_i - c_i)$, $n = 2, 3, \dots, M-1$, using (7) we can write

$$(8) \quad \beta(\theta) = 1 - \prod_{i=1}^M \left(\frac{a_i}{\theta} \right) \left[\prod_{i=1}^M a_i^{-1} \sum_{n=1}^{M-1} d_n \theta^{M-n} + \prod_{i=1}^{M-1} \left(1 - \frac{c_i}{a_i} \right) \right].$$

If $\prod_{i=1}^M a_i = (1-\alpha)\theta_0^M$, the theorem implies that $b_1 = b_2 = \dots = b_{M-1} = 0$. Also from the theorem, if $\prod_{i=1}^M a_i \neq (1-\alpha)\theta_0^M$, then $\prod_{i=1}^M a_i < (1-\alpha)\theta_0^M$ in which case $\prod_{i=1}^M a_i = \lambda(1-\alpha)\theta_0^M$ for some $\lambda \in (0, 1)$. But then $\prod_{i=1}^M \left(\frac{a_i}{\theta_1} \right) = \lambda(1-\alpha) \left(\frac{\theta_0}{\theta_1} \right)^M = \lambda\beta$ since $M = \log \left(\frac{1-\alpha}{\beta} \right) / \log \left(\frac{\theta_1}{\theta_0} \right)$. Thus, from (9), since $\beta(\theta_0) = \alpha$ and $\beta(\theta_1) = 1 - \beta$, we must have

$$\frac{1}{\lambda} = \prod_{i=1}^M a_i^{-1} \sum_{n=1}^{M-1} d_n \theta_0^{M-n} + \prod_{i=1}^{M-1} \left(1 - \frac{c_i}{a_i} \right) = \prod_{i=1}^M a_i^{-1} \sum_{n=1}^{M-1} d_n \theta_1^{M-n} + \prod_{i=1}^{M-1} \left(1 - \frac{c_i}{a_i} \right)$$

or $\sum_{n=1}^{M-1} d_n \theta_0^{M-n} = \sum_{n=1}^{M-1} d_n \theta_1^{M-n}$. But θ_0 and θ_1 are distinct and $d_n \geq 0$ for each n , where $1 \leq n \leq M-1$ and hence $d_n = 0$ for $n = 1, 2, \dots, M-1$ which implies that $b_1 = b_2 = \dots = b_{M-1} = 0$.

It follows that we cannot improve on the optimal single boundary procedure; that is, stating the result for the class of distributions with densities $f_\theta(x) = c(\theta)g(x)$, $0 < x < \theta$, the stopping rule of the form (1) which

uniformly minimizes the average sample size for $\theta \geq \theta_0$ is the one for which $a_1 = F_{\theta_0}^{-1}(1 - \alpha)$, $a_2 = a_3 = \dots = a_M = \theta_0$ and $b_1 = b_2 = \dots = b_{M-1} = 0$.

UNIVERSITY OF NEW SOUTH WALES

REFERENCES

- [1] Lehmann, E. L. (1949). Notes on the theory of estimation, lecture notes recorded by Colin Blyth, University of California, Berkeley.
- [2] Lehmann, E. L. (1959). *Testing Statistical Hypotheses*, John Wiley and Sons, New York.
- [3] Samuel-Cahn, E. (1974). Two kinds of repeated significance tests, and their application for the uniform distribution, *Commun. Statist.*, 3(5), 419-431.