AD-A192 043   STOPPING RULES AND ORDERED FAMILIES OF DISTRIBUTIONS   1/1
               (U) NORTH CAROLINA UNIV AT CHAPEL HILL CENTER FOR
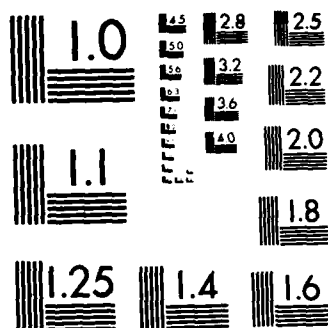   STOCHASTIC PROCESSES   J BATHER DEC 87 TR-219
UNCLASSIFIED   AFOSR-TR-88-0350 F49620-85-C-0144   F/G 12/3   NL

REPORT DOCUMENTATION PAGE

1b. RESTRICTIVE MARKINGS	

AD-A192 843

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	**3. DISTRIBUTION / AVAILABILITY OF REPORT** Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report No. 219	**5. MONITORING ORGANIZATION REPORT NUMBER(S)** AFOSR·TR· 88-0358

6a. NAME OF PERFORMING ORGANIZATION University of North Carolina	**6b. OFFICE SYMBOL** *(If applicable)*	**7a. NAME OF MONITORING ORGANIZATION** AFOSR/NM

6c. ADDRESS (City, State, and ZIP Code) Statistics Dept. 321-A Phillips Hall 039-A Chapel Hill, NC 27514	**7b. ADDRESS (City, State, and ZIP Code)** AFOSR/NM Bldg 410 Bolling AFB DC 20332-6448

8a. NAME OF FUNDING / SPONSORING ORGANIZATION AFOSR	**8b. OFFICE SYMBOL** *(If applicable)* NM	**9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER** AFOSR No. F49620 85C 0144.

8c. ADDRESS (City, State, and ZIP Code) AFOSR/NM Bldg 410 Bolling AFB DC 20332-6448	**10. SOURCE OF FUNDING NUMBERS**

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102F	2304	AS	

11. TITLE (Include Security Classification)
Stopping rules and ordered families of distributions

12. PERSONAL AUTHOR(S)
Bather, J.

13a. TYPE OF REPORT preprint	**13b. TIME COVERED** FROM 9/87 TO 8/88	**14. DATE OF REPORT (Year, Month, Day)** December 1987	**15. PAGE COUNT** 15

16. SUPPLEMENTARY NOTATION
N/A

17.	COSATI CODES		**18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)**
FIELD	GROUP	SUB-GROUP	Key Words & Phrases: confidence intervals; monotone
			likelihood ratios; stochastic ordering; sequential decisions.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
There are good reasons for using sequential methods in some statistical decision problems, but a stopping rule that is helpful for deciding whether $\theta > 0$ or $\theta < 0$ may not be so good for estimating θ. This paper considers the construction of confidence bounds on a real parameter and investigates the relation between the ordering of boundary points that are accessible under the stopping rule and the natural ordering of the parameter space.

| **20. DISTRIBUTION / AVAILABILITY OF ABSTRACT** ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	**21. ABSTRACT SECURITY CLASSIFICATION** Unclassified/unlimited
22a. NAME OF RESPONSIBLE INDIVIDUAL Maj. Brian Woodruff	**22b. TELEPHONE (Include Area Code)** (202) 767-5026	**22c. OFFICE SYMBOL** NM

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted.
All other editions are obsolete.

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina
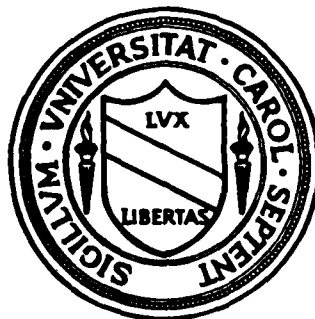
STOPPING RULES AND ORDERED FAMILIES OF DISTRIBUTIONS

by

John Bather

Technical Report No. 219

December 1987

STOPPING RULES AND ORDERED FAMILIES OF DISTRIBUTIONS

John Bather[*]

Mathematics Division, University of Sussex

and

Center for Stochastic Processes, University of North Carolina

at Chapel Hill

ABSTRACT

There are good reasons for using sequential methods in some statistical
decision problems, but a stopping rule that is helpful for deciding whether
$\theta > 0$ or $\theta < 0$ may not be so good for estimating θ. This paper considers the
construction of confidence bounds on a real parameter and investigates the
relation between the ordering of boundary points that are accessible under the
stopping rule and the natural ordering of the parameter space.

1. INTRODUCTION

There are many ways of defining an ordering of probability distributions so that large values of the parameter which labels the distributions correspond to large values of the random variables themselves. For example, Lehmann (1955) gives a comparison of several definitions and an application to the sequential probability ratio test. We shall be mainly concerned with stochastic ordering, which is well-known in applied probability, and the stronger relation of ordering by monotone likelihood ratios (m.l.r.). The latter is familiar in the theory of hypothesis testing, where it leads to uniformly most powerful one-sided tests for fixed samples from a distribution in a 1-dimensional exponential family. A closely related application of m.l.r. is the construction of uniformly most accurate confidence bounds on the unknown parameter: see Lehmann's book (1959), pages 78-80. The aim of this paper is to investigate confidence bounds determined after a random stopping time. As we shall see, the above optimality property of one-sided confidence bounds is usually lost when we allow sequential sampling. However, for a large class of stopping rules, we can define an ordering of the boundary points so that the distributions of stopping points are stochastically ordered with respect to the parameter. In general, this weaker ordering relation seems to be the best that can be obtained, which underlines the need for caution when interpreting confidence bounds and intervals based on sequential data.

In a recent book, Siegmund (1985) illustrates the advantages and disadvantages of sequential methods. For hypothesis testing, we can achieve a significance level and power comparable to fixed sample procedures with a substantial reduction in expected sample size. On the other hand, sequential sampling often leads to less accurate estimation. In the design of clinical trials, estimation may be a secondary consideration: for ethical reasons, it is

important to reject an inferior treatment as soon as possible, rather than using it on a larger sample of patients to improve the estimates of its performance. Previous research has been mainly concerned with procedures for comparing different treatments and stopping rules designed to control the error probabilities of decisions about their relative merits. More recently, Siegmund and others have turned to questions of estimation from sequences of observations produced by various stopping rules. In particular, they have made substantial progress in constructing confidence intervals, in spite of the difficult probability calculations and delicate approximations often needed to deal with curved stopping boundaries.

The construction of confidence intervals after a stopping time is based on standard methods. Consider first a random variable Z whose distribution depends on a real parameter θ and write $\varphi(z,\theta) = P_\theta(Z \geq z)$. In general, a lower confidence bound for θ can be obtained by using the fact that

$$P_\theta(\varphi(Z,\theta) \geq \alpha) \geq 1-\alpha, \tag{1}$$

for any fixed α, $0 \leq \alpha \leq 1$. Let $\underline{\theta}(z) = \inf\{\theta: \varphi(z,\theta) \geq \alpha\}$, so that $\varphi(z,\theta) \geq \alpha \Rightarrow \underline{\theta}(z) \leq \theta$. It follows from this definition and (1) that

$$P_\theta(\underline{\theta}(Z) \geq \theta) \geq P_\theta(\varphi(Z,\theta) \geq \alpha) \geq 1-\alpha. \tag{2}$$

Thus, $\underline{\theta}(z)$ is a lower confidence bound for θ, given an observed value z of the random variable. Upper confidence bounds $\overline{\theta}(z)$ can be constructed similarly, for the same confidence coefficient $1-\alpha$, and we then obtain confidence intervals with coefficient $1-2\alpha$ from the inequality

$$P_\theta(\underline{\theta}(Z) \leq \theta \leq \overline{\theta}(Z)) \geq 1-P_\theta(\underline{\theta}(Z) > \theta) - P_\theta(\overline{\theta}(Z) < \theta).$$

This probability is at least $1-2\alpha$, because of (2) and a similar property of $\overline{\theta}(Z)$.

The above argument extends in a straightforward manner to more complicated sample spaces where the data consists of stopped sequences of observations.

Let $x_1, x_2, \ldots$ be independent observations from the probability density

$$f(x;\theta) = \exp\{\theta x - \psi(\theta)\}, \tag{3}$$

with respect to a σ-finite measure μ on the real line. This represents a 1-dimensional exponential family of distributions with mean $\psi'(\theta)$ and variance $\psi''(\theta) > 0$. The natural parameter space of the family is a real interval and we suppose that the true value of θ in this interval is unknown. Let $S_n = \sum_1^n x_i$ for $n \geq 1$ and note that any stopped sequence of observations $(x_1, x_2, \ldots, x_n)$ is represented by the sufficient statistic (n,s), where $S_n = s$. Suppose that the stopping time is defined by splitting the (n,s) plane into continuation and stopping regions. It is helpful to think of upper and lower stopping sets, separated by the continuation region. Since $\psi''(\theta) > 0$, the mean $\psi'(\theta)$ is increasing in θ and we can imagine that, roughly speaking, points (n,s) in the upper stopping set favour higher values of θ. However, we need a more precise ordering relation on the points of the stopping region before we can justify constructing confidence bounds on θ.

The idea is to order all the accessible boundary points (n,s) in a counter-clockwise sense around the continuation region, but this can be done in several ways. The definition given in Section 4 of this paper has been used previously by Siegmund (1978), in obtaining confidence intervals for a normal mean, and also by Jennison and Turnbull (1983), for estimating a binomial probability. The authors of the second paper considered a different ordering of boundary points based on the ratio s/n, the maximum likelihood estimate of the unknown probability, but this produced very similar results in the cases they computed. Siegmund (1985) also mentions the ratio s/n as a possible basis for ordering normal data.

In general, we must determine a function $\varphi(n,s,\theta)$ which represents the probability of stopping the sequence of observations at a point "above" (n,s).

The function can then be used to find confidence bounds on θ in the same way as $\varphi(z,\theta)$. It is important to note that the construction described earlier does not depend explicitly on any relation between the parameter θ and the ordering of the sample space. In principle, any ordering of points (n,s) in the stopping region can produce valid confidence bounds. However, it is possible to ensure that the function $\varphi(n,s,\theta)$ is increasing in θ. This general property does not seem to have been established previously, although it is obviously useful in computing confidence bounds from the formula

$$\underline{\theta}(n,s) = \inf\{\theta: \varphi(n,s,\theta) \geq \alpha\}. \tag{4}$$

The rest of this paper is organized as follows. The next section gives a simple example which demonstrates the advantages and disadvantages of sequential procedures. It also suggests an arbitrariness in the construction of confidence bounds which is not easy to eliminate entirely. However, we can distinguish between bounds that are formally valid and more sensible constructions that are also related to the unknown parameter. Section 3 is concerned with partial orderings of probability distributions. It contains a brief description of and comparison between stochastic ordering and ordering by m.l.r. There is a new result about the most likely permutation of a number of independent random variables from an ordered family of distributions. The property that the most likely permutation corresponds to the ordering of the family holds for m.l.r. but not, in general, for stochastic ordering: see Proposition 4. The ordering of boundary points for the random walk $\{S_n, n \geq 1\}$ is considered in Section 4 and it is proved that, for a large class of stopping times N, the distributions of the random point (N,S_N) are stochastically ordered with respect to θ. This means that the function $\varphi(n,s,\theta)$ is increasing in θ. The final section of the paper gives another illustration, based on a simple acceptance/rejection scheme for diffusion processes with unknown drift parameters. It shows that the intuitive counter-clockwise ordering of random

stopping points with respect to the drift can be destroyed by conditioning on rejection. The paper concludes with some tentative remarks about the design of stopping rules.

2. EXAMPLE

Consider independent Bernoulli trials with success probability p and let $x_i = \pm 1$, according as the i-th trial results in success or failure, i=1,2,... . Thus, we have an exponential family of distributions which can be expressed in the form (3) by writing $\theta = \frac{1}{2}\log(p/q)$, $\psi(\theta) = \log(e^{\theta}+e^{-\theta})$, where q=1-p. However, the usual notation for Bernoulli trials will be more convenient here. Suppose that we must decide, after observing a number of trials, whether p > ½ or not. Various stopping rules will be considered for the random walk $\{S_n\}$, $S_n = \sum_1^n x_i$. In each case, the terminal decision at a point (n,s) with $S_n = s$ will depend only on the sign of s : we conclude that p > ½ if s > 0, that p < ½ if s < 0 and we choose either decision by tossing a coin if s = 0.

We now turn to the stopping rules. Rule 1 is to take 4 observations and then reach a decision about p, according to the sign of S_4. Rule 2 is a sequential modification of it: we observe x_1 and x_2 and stop if $S_2 = \pm 2$, but if $S_2 = 0$ we take another 4 observations. This modification can be used repeatedly to produce a series of rules. Rule k is specified as follows: observe the sequence $\{S_1, S_2, \ldots\}$ and stop as soon as $S_{2n} = \pm 2$, but if $S_{2n} = 0$ for n = 1,2,...,k-1, take 4 more observations and stop at n = 2k+2. The limiting form of these rules, with k = ∞, can be regarded as a sequential probability ratio test. As we shall see, Rule ∞ is an improvement on its predecessors both with regard to expected sample size and with regard to error probability.

Let $m_k(p)$ denote the expected sample size for Rule k and let $a_k(p)$ be the probability of reaching a terminal decision that p > ½. The corresponding error probability is given by:

$$e_k(p) = a_k(p), \quad 0 \le p \le \tfrac{1}{2}, \tag{5}$$

$$e_k(p) = 1 - a_k(p), \quad \tfrac{1}{2} \le p \le 1.$$

We now prove that, for all p,

$$m_1(p) \ge m_2(p) \ge \ldots \ge m_\infty(p), \tag{6}$$

$$e_1(p) \ge e_2(p) \ge \ldots \ge e_\infty(p). \tag{7}$$

<u>Proof.</u> Clearly, $m_1(p) = 4$ and $m_2(p) = 2 + 2pq m_1(p)$. It follows that $m_2(p) \le m_1(p)$, with equality only if $p = \tfrac{1}{2}$. Then by considering S_2, $m_{k+1}(p) = 2 + 2pq m_k(p)$ and an inductive argument shows that $m_{k+1}(p) \le m_k(p)$ for all k. As $k \to \infty$, $m_k(p) \to m_\infty(p) = 2(p^2 + q^2)^{-1}$.

The proof of (7) is similar, but we need to use (5). Note that $a_1(p) = p^2(1 + 2q)$ and $a_{k+1}(p) = p^2 + 2pq a_k(p)$ for $k = 1, 2, \ldots$. It is easy to show by induction that $a_{k+1}(p) > a_k(p)$ if $0 < p < \tfrac{1}{2}$ and $a_{k+1}(p) < a_k(p)$ if $\tfrac{1}{2} < p < 1$. We also have $a_k(p) = p$, whenever $p = 0$, $\tfrac{1}{2}$ or 1, so the relations (7) follow immediately from (5). In fact, explicit formulae for $a_k(p)$ can be obtained and, in the limit, $a_\infty(p) = p^2(p^2 + q^2)^{-1}$.

The inequalities (6) and (7) show that, so far as terminal decisions are concerned, the performance of the stopping rules improves as k increases. However, their relative merits for estimation are quite different. Consider first the unbiased estimation of p. It turns out that there is just one unbiased estimator $\hat{p}_k$, based on Rule k. In particular, $\hat{p}_1 = (S_4 + 4)/8$ and its variance is $pq/4$. For $k \ge 2$, $\hat{p}_k = (S_2 + 2)/4$ is unbiased and this has variance $pq/2$. Standard methods can be used to verify that $\hat{p}_k$ is the unique unbiased estimator of p, but we shall omit the details. Thus, for Rules $2, 3, \ldots$, the minimum variance unbiased estimator of p depends only on the first two Bernoulli trials. From this point of view, Rule 1 is preferable.

Now consider confidence bounds on p. Instead of making a comparison of different rules, we shall restrict attention to Rule 2. The stopping region consists of 7 points: $(2, -2)$, $(6, -4)$, $(6, -2)$, $(6, 0)$, $(6, 2)$, $(6, 4)$, $(2, 2)$, and

let us label these 1,2,...,7 in counter-clockwise order. Their labels here are related to p in the following sense. Let $\varphi(j,p)$ be the probability that the random walk stops at a point whose label is at least j. Then $\varphi(j,p)$ is non-decreasing in p for j=1,2,...,7. This is a consequence of the general result which will be proved in Section 4. However, there are several possible orderings of the 7 points with this property. For example, it is not difficult to verify that it also holds if the labels (1,2,3,4,5,6,7) are replaced by (1,3,2,4,6,5,7), respectively. Another ordering of the stopping region that is plausible from a different point of view is obtained on replacing the original labels by (2,1,3,4,5,7,6). It is arguable that the point (6,4), representing 5 successes in 6 trials, indicates higher values of p than the point (2,2). To fix ideas, suppose that we observe a sequence of Bernoulli trials which terminates at the point (6,4). In order to construct a lower confidence bound on p, we need to specify the set A of boundary points above the data. There are $2^6 = 64$ possible definitions of A. It is easy to see that 32 of these would lead to the trivial claim that $0 \leq p \leq 1$, but the others produce confidence bounds that make more sense. For example, the 3 possible orderings mentioned above would lead to different statements of the form: $\underline{p} \leq p \leq 1$, for the same confidence coefficient.

3. ORDERING OF RANDOM VARIABLES

We now consider two different partial orderings of probability distributions on the real line. A brief outline of their properties is given below. Let Y and Z be random variables with distribution functions G and H, respectively. Note that, if we define $G^{-1}(u) = \inf\{y: G(y) \geq u\}$, for $0 < u < 1$, then the distribution of Y can be described by writing $Y = G^{-1}(U)$, where U is uniformly distributed on $[0,1]$.

<u>Definition 1.</u> We say that Y is stochastically less than Z and write $Y \leq_{st} Z$ if

$E(v(Y)) \leq E(v(Z))$ for every bounded increasing function v on $\mathbb{R}$.

__Proposition 1.__ $Y \leq_{st} Z \iff G(t) \geq H(t)$, $t \in \mathbb{R}$. Hence, if $Y \leq_{st} Z$, we can write $Y = G^{-1}(U)$, $Z = H^{-1}(U)$, where U is uniformly distributed on [0,1]. For this representation of the joint distribution, the inequality $Y \leq Z$ always holds.

Suppose further that Y and Z have probability densities g and h, with respect to a common σ-finite measure μ on $\mathbb{R}$.

__Definition 2.__ We say that Y is less than or equal to Z in the sense of monotone likelihood ratio and write $Y \leq_r Z$ if $h(t)/g(t)$ is non-decreasing in $t \in \mathbb{R}$ (excluding t such that $g(t) = h(t) = 0$).

__Proposition 2.__ $Y \leq_r Z \Rightarrow Y \leq_{st} Z$.

Proofs of the above results can be found in Lehmann's paper and there is a clear exposition, for discrete random variables, in the paper by Whitt (1979). This also introduces the notion of uniform conditional stochastic order (u.c.s.o.), which is investigated more generally in Whitt (1980). For our purposes, it will be enough to note one of the results from the last paper. For distributions with probability densities on the real line, u.c.s.o. is equivalent to m.l.r. in the following sense. Let $B \subset \mathbb{R}$ be a Borel set and consider the probability distributions of Y_B and Z_B obtained by conditioning Y and Z, respectively, on the event B.

__Proposition 3.__ $Y_B \leq_{st} Z_B$ for every event B $\iff$ $Y \leq_r Z$.

There is another way of seeing that ordering by m.l.r. is stronger than stochastic ordering. Suppose that we have an ordered sequence of random variables $Y_1, Y_2, \ldots, Y_k$ and that they have probability densities $g_1, g_2, \ldots, g_k$

with respect to μ. Suppose further that they are independent of one another and consider whether the most likely ordering of their observed values is

$$Y_1 \leq Y_2 \leq \ldots \leq Y_k.$$

__Proposition 4.__ (i) Let $Y_1, Y_2, \ldots, Y_k$ be independent random variables and suppose that $Y_1 \leq_r Y_2 \leq_r \ldots \leq_r Y_k$. Let $\pi = (\pi_1, \pi_2, \ldots, \pi_k)$ be any permutation of the integers $1, 2, \ldots, k$. Then

$$P(Y_1 \leq Y_2 \leq \ldots \leq Y_k) \geq P(Y_{\pi_1} \leq Y_{\pi_2} \leq \ldots \leq Y_{\pi_k}).$$

(ii) This property does not hold, in general, for $k \geq 3$ independent random variables such that $Y_1 \leq_{st} Y_2 \leq_{st} \ldots \leq_{st} Y_k$.

__Proof.__ We shall establish Part (i) by showing that

$$g_1(y_1)g_2(y_2)\ldots g_k(y_k) \geq g_{\pi_1}(y_1)g_{\pi_2}(y_2)\ldots g_{\pi_k}(y_k) \tag{8}$$

holds at every point of $C_k = \{(y_1, y_2, \ldots, y_k): y_1 \leq y_2 \leq \ldots \leq y_k\}$. The required result will then follow by integrating over the set C_k.

Note first that $g_1(y_1)g_2(y_2) \geq g_2(y_1)g_1(y_2)$ if $y_1 \leq y_2$, so (8) holds for $k = 2$. Now let $k \geq 3$ and assume that

$$g_1(y_1)g_2(y_2)\ldots g_{k-1}(y_{k-1}) \geq g_{\sigma_1}(y_1)g_{\sigma_2}(y_2)\ldots g_{\sigma_{k-1}}(y_{k-1}) \tag{9}$$

holds in C_{k-1}, for any permutation σ of $1, 2, \ldots, k-1$. If $\pi_k = k$, then (8) is a trivial consequence of (9), so we may assume that, for some $j < k$, $\pi_j = k$ and $\pi_k < k$. We define σ in (9) by $\sigma_i = \pi_i$ if $i \neq j$, $i \leq k-1$, and $\sigma_j = \pi_k$. It is now a straightforward matter to deduce (8), by using the fact that, since $y_j \leq y_k$,

$$g_k(y_k) \geq g_k(y_j)g_{\pi_k}(y_k)/g_{\pi_k}(y_j) = g_{\pi_j}(y_j)g_{\pi_k}(y_k)/g_{\sigma_j}(y_j).$$

The proof of Part (ii) is based on a counter-example. Let Y_1', Y_2', Y_3' be independent and uniformly distributed on $[-1, 1]$. Define $Y_1 = \min(Y_1', 0)$, $Y_2 = Y_2'$, $Y_3 = \max(Y_3', 0)$. Proposition 1 can be used to show that $Y_1 \leq_{st} Y_2 \leq_{st} Y_3$. On the other hand,

$$P(Y_1 \leq Y_2 \leq Y_3) = \frac{1}{4}, \text{ but } P(Y_2 \leq Y_1 \leq Y_3) = P(Y_1 \leq Y_3 \leq Y_2) = \frac{3}{8}.$$

It is easily shown that the result of Part (i) remains true for stochastic ordering in the case k=2. However, it does not hold for $k \geq 4$ and this can be demonstrated by extending the above example.

4. ORDERING OF STOPPED SEQUENCES

We now return to the random walk model described earlier, with $S_n = \sum_1^n x_i$, $n \geq 1$, where the steps x_i are generated by independent observations from the distribution defined by (3). We restrict attention to stopping times specified by two sequences of numbers. Let $-\infty \leq a_n \leq b_n \leq \infty$ for n=1,2,... and let

$$N = \min\{n \geq 1 : S_n \notin (a_n, b_n)\}. \tag{10}$$

Clearly, $N \leq m = \min\{n \geq 1 : a_n = b_n\}$, if this is finite. If $a_n < b_n$ for all n, we set $m = \infty$ and the stopping time N may be infinite. However, it is assumed in such cases that N is finite with probability 1, for any value of the parameter θ. This means that the stopping point (N, S_N) always has a proper distribution.

The stopping region associated with N consists of points (n,s) such that $1 \leq n \leq m$ and either $s \leq a_n$ or $s \geq b_n$. This can be regarded as a totally ordered set.

Definition 3. Let (n,s) and (n',s') be points of the stopping region. We say that (n',s') is above (n,s) and write $(n',s') \rangle (n,s)$ if one of the following conditions holds:

      (i) $n' = n$ and $s' \geq s$,

      (ii) $n' < n$ and $s' \geq b_{n'}$,

      (iii) $n' > n$ and $s \leq a_n$.

For example, it is a straightforward matter to check that either

$(n',s') \rangle (n,s)$ or $(n,s) \rangle (n',s')$, for every pair of stopping points, and that the relation $\rangle$ is transitive.

We are now in a position to prove the main result. Let

$$\varphi(n,s,\theta) = P_\theta((N,S_N) \rangle (n,s)). \tag{11}$$

Theorem. Under the above conditions, the function φ is non-decreasing in θ:

$$\varphi(n,s,\theta') \geq \varphi(n,s,\theta)$$

whenever $\theta' > \theta$, for any point (n,s) in the stopping region.

Proof. We shall couple together two realisations of the random walk, corresponding to the parameter values θ and θ'. It follows from (3) that the likelihood ratio for a single observation x is $\exp\{(\theta'-\theta)x-(\psi(\theta')-\psi(\theta))\}$ and this is increasing in x if $\theta' > \theta$. Hence, we can associate random variables X and X' with θ and θ', respectively, such that $X \leq_r X'$. Let $F(x;\theta)$ be the distribution function determined by (3). Then, according to Propositions 1 and 2, we can describe the two distributions by writing $X = F^{-1}(U;\theta)$ and $X' = F^{-1}(U;\theta')$, where U is uniformly distributed on $[0,1]$. Now let $u_1,u_2,\ldots$ be independent observations from the uniform distribution and consider the realisations generated by setting $S_n = \sum_1^n x_i$, $S_n' = \sum_1^n x_i'$, where $x_i = F^{-1}(u_i;\theta)$, $x_i' = F^{-1}(u_i;\theta')$, and hence $x_i \leq x_i'$ always holds. We must compare the stopping points associated with θ and θ'. Given the sequence $(u_1,u_2,\ldots)$, we can apply (10) to determine points (N,S_N) and $(N',S_N'.)$, say. Then it follows from Definition 3 and the fact that $S_n' \geq S_n$, for all $n \geq 1$, that $(N',S_N'.) \rangle (N,S_N)$. Thus, we have generated the stopping points from independent uniform random variables in such a way that the event $[(N,S_N) \rangle (n,s)]$ is contained in the event $[(N',S_N'.) \rangle (n,s)]$. The theorem follows immediately.

Remark. The above argument can also be used to show that, for any bounded increasing function v defined on the stopping region, $E(v(N,S_N)) \leq E(v(N',S_{N'}))$ if $\theta < \theta'$. Hence, the random point (N,S_N) is stochastically increasing in θ, in the sense of Definition 1.

5. ILLUSTRATION

The monotonicity of the function $\varphi(n,s,\theta)$ established in the theorem is useful in constructing confidence bounds, but it does not mean that such bounds are optimal. The uniformly most accurate confidence bounds mentioned earlier are obtained only if the random stopping point (N,S_N) is increasing with respect to θ in the stronger sense of m.l.r. This is exceptional: roughly speaking, fixed samples lead to uniformly most accurate one-sided confidence intervals, but random stopping times do not.

We can easily see why, by examining likelihood ratios for the exponential model (3). After n observations, suppose we find that $S_n = s$. The likelihood ratio for parameter values $\theta < \theta'$ is $\exp\{(\theta'-\theta)s-(\psi(\theta')-\psi(\theta))n\}$. This is increasing in s, so if the sample size is fixed in advance at n, the m.l.r. property holds. Now suppose that (n,s) and (n',s') are points of the stopping region with $n' \neq n$. The second point yields a higher likelihood ratio if and only if

$$s'-s \geq \frac{(\psi(\theta')-\psi(\theta))}{(\theta'-\theta)} (n'-n).$$

We could extend this comparison of points to produce an ordering relation on the stopping region but, in general, the relation would depend on our choice of parameter values, since $\psi''(\theta) > 0$ and the coefficient $(\psi(\theta')- \psi(\theta))/(\theta'-\theta)$ is not constant. In the case of independent Bernoulli trials, it is not difficult to devise random stopping times in such a way that the m.l.r. property holds, but it is worth noting that in the example discussed in Section 2, only Rule 1 with a fixed sample size produces a stopping region that has the m.l.r.

property.

Finally, consider a simple acceptance/rejection scheme based on a process in continuous time. Let $S(t) = \theta t + W(t)$ for $t \geq 0$, where $S(0) = 0$ and $W(t)$ is a standard Wiener process. Suppose that $S(t)$ is a summary of responses in $[0, t]$ to a new medical treatment and that positive values of the unknown parameter θ represent a higher risk of serious adverse effects. Let b and m be fixed positive numbers and let the decision procedure be specified as follows: stop and reject the treatment as soon as $S(t) = b$ if this occurs for some $t < m$; accept the new treatment if $S(t) < b$ for $0 \leq t \leq m$. A detailed evaluation of this procedure is given in Siegmund's book (1985): see Chapter 3. Here, the aim is to illustrate some consequences of Propositions 3 and 4.

Since the boundary prevents any overshoot, the stopping region consists of two lines in the (t,s) plane. Strictly speaking, the theorem of Section 4 does not cover processes in continuous time, but it is easy to verify that, for a counter-clockwise ordering of the boundary, we have stochastic ordering of the distributions of the terminal point with respect to θ. For two values θ and θ' of the drift parameter, the likelihood ratio at any boundary point (t,s) is

$$\exp\{(\theta'-\theta)s - \tfrac{1}{2}(\theta'^2-\theta^2)t\}. \tag{12}$$

In cases of acceptance, $t = m$ and this is increasing in s provided that $\theta' > \theta$. Rejected cases occur on the line $s = b$, for $t < m$, and there the likelihood ratio is increasing in the counter-clockwise direction (i.e. decreasing in t) if and only if $|\theta'| > |\theta|$.

Consider the results of applying the scheme independently to k different treatments and suppose the corresponding drift parameters are in the order: $\theta_1 > \theta_2 > \ldots > \theta_k$. Intuitively, it might seem that the most likely arrangement of the corresponding terminal points is $(t_1, s_1) \rangle (t_2, s_2) \rangle \ldots \rangle (t_k, s_k)$, using the obvious extension of Definition 3. However, this may not be true. Proposition 4 applies if $|\theta_1| \geq |\theta_2| \geq \ldots \geq |\theta_k|$, but let us assume that this last condition

does not hold. We can argue conditionally, if all the treatments are accepted. Because of the m.l.r. property on the line $t = m$, given acceptance, the most likely arrangement is $b \geq s_1 \geq s_2 \geq \ldots \geq s_k$. However, after rejection, the observed arrangement of the final times could be quite misleading. Conditional on rejection, we have the m.l.r. property based on $|\theta|$, rather than θ, and the possibility of false rejections (i.e. cases with $\theta_i < 0$) makes the situation more complicated. The most likely arrangement, given rejection, need not be the one with $t_1 \leq t_2 \leq \ldots \leq t_k < m$.

More generally, suppose we have a stopping region determined by two smooth boundary curves: $s = a(t)$ and $s = b(t)$. The process $\{S(t)\}$ is allowed to continue so long as $a(t) < S(t) < b(t)$. We can see by using (12) that on the upper boundary curve, the likelihood ratio is

$$\exp\{(\theta'-\theta)(b(t) - \tfrac{1}{2}(\theta+\theta')t)\}$$

and this is increasing in the counter-clockwise direction if $\theta' > \theta$ and if the derivative $b'(t) < \tfrac{1}{2}(\theta+\theta')$. In the special case where $b'(t) = 0$, we noted that negative values of the drift led to complications in relating the order of parameter values to the order of stopping points on the boundary. Here we can say roughly that the idea of a counter-clockwise ordering of boundary points remains valid, conditional on stopping near the point $(t, b(t))$, provided that we are concerned with values of the drift leading towards the boundary (i.e. $\theta > b'(t)$). Similar remarks apply to the lower boundary for values of the drift $\theta < a'(t)$. It seems that we should try to design stopping rules so that there is always a high probability that the random process will reach a stopping point where the expected increments lead towards the boundary, rather than away from it.

BIBLIOGRAPHY

Jennison, C. and Turnbull, B.W. (1983). Confidence intervals for a binomial parameter following a multistage test. Technometrics, 24, 49-58.

Lehmann, E.L. (1955). Ordered families of distributions, Ann. Math. Statist., 26, 399-419.

Lehmann, E.L. (1959). Testing statistical hypothesis, New York: Wiley.

Siegmund, D. (1978). Estimation following sequential tests, Biometrika, 65, 341-349.

Siegmund, D. (1985). Sequential analysis, New York: Springer-Verlag.

Whitt, W. (1979). A note on the influence of the sample on the posterior distribution, Journal of the American Statistical Association, 74, 424-426.

Whitt, W. (1980). Uniform conditional stochastic order, J. Appl. Prob. 17, 112-123.

END

DATE

FILMED

6-1988

DTIC