REMARKS ON A STOPPING TIME

BY NORMAN STARR AND MICHAEL B. WOODROOFE

UNIVERSITY OF MICHIGAN, ANN ARBOR

Communicated by Paul A. Smith, October 7, 1968

1. *Some Inequalities.*—Throughout this section $x_1, x_2, \ldots$ will be independent, identically distributed (i.i.d.) random variables with finite expectation Ex_1. Put $\bar{x}_n = (\sum_1^n x_i)/n$, $n \geq 1$, and let $\{c_n\}$ be any sequence of constants and m any positive integer. A stopping time of the sequence $x_1, x_2, \ldots$ is defined by

$$\begin{aligned} N &= \text{smallest } n \geq m \text{ such that } \bar{x}_n \leq c_n \\ &= \infty \text{ if no such } n \text{ exists, i.e., if } \bar{x}_n > c_n \text{ for every } m \leq n < \infty. \end{aligned}$$

We assume that $P(N < \infty) = 1$, so that $\bar{x}_N$ is well defined.

THEOREM 1. *If $E\bar{x}_N$ exists, then $E\bar{x}_N \leq Ex_1$.*

Proof: We may and shall assume that $Ex_1 = 0$. For any $n \geq m$ and any $i = 1, 2, \ldots, n$, we therefore have

$$\int_{(N>n)} x_i\, dP = \int \ldots \int_A \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \int_{\alpha}^{\infty} x_i dF(x_i) dF(x_n) \ldots dF(x_{i+1}) dF(x_{i-1}) \ldots dF(x_1) \geq 0,$$

where A denotes the set of values of $x_1, \ldots, x_{i-1}$ for which $N > i - 1$ and

$$\alpha = \max\left\{kc_k - \sum_{j=1, j\neq i}^{k} x_j; \quad i \leq k \leq n\right\}.$$

It follows that

$$\int_{(N>n)} \bar{x}_n dP \geq 0 \quad (n > m),$$

and since x_n is independent of the event $N > n - 1$,

$$\int_{(N>n-1)} \bar{x}_n dP = \frac{(n-1)}{n} \int_{(N>n-1)} \bar{x}_{n-1}\, dP \leq \int_{(N>n-1)} \bar{x}_{n-1}\, dP \quad (n > m).$$

Hence for every $n > m$

$$\begin{aligned} \int_{(N\leq n)} \bar{x}_N dP &= \sum_{i=m}^{n-1} \int_{(N=1)} \bar{x}_i dP + \int_{(N>n-1)} \bar{x}_n dP - \int_{(N>n)} \bar{x}_n dP \\ &\leq \sum_{i=m}^{n-1} \int_{(N=i)} \bar{x}_i dP + \int_{(N>n-1)} \bar{x}_n dP \\ &\leq \sum_{i=m}^{n-1} \int_{(N=i)} \bar{x}_i dP + \int_{(N>n-1)} \bar{x}_{n-1}\, dP \\ &= \sum_{i=m}^{n-2} \int_{(N=i)} \bar{x}_i dP + \int_{(N>n-2)} \bar{x}_{n-1} dP \\ &\leq \ldots \leq \int_{(N=m)} \bar{x}_m dP + \int_{(N>m)} \bar{x}_{m+1}\, dP \leq \int \bar{x}_m dP = E\bar{x}_m = 0. \end{aligned}$$

The result follows on letting $n \to \infty$.

Theorem 1 and Corollary 2 (below) were communicated to us by H. Robbins.

Remarks: (*a*) The theorem remains true if the x_i are merely assumed to be independent with $Ex_i = Ex_1$ for all $i \geq 1$. However, if the c_i are allowed to depend on $x_1, \ldots, x_{i-1}$, we can have $E\bar{x}_N > Ex_1$.

(*b*) Consider a gambler who plays a succession of fair games and stops with x_N, where the $\{c_n\}$ sequence is such that $\lim_{n\to\infty}\sup c_n > 0$. Then $P(N < \infty) = 1$, and hence if $E\bar{x}_N$ exists, it is ≤ 0; in fact, $E\bar{x}_N < 0$, unless N is constant with probability 1. On the other hand, if the x_i are i.i.d., then by Wald's lemma on cumulative sums, $E(N\bar{x}_N) = 0$, provided $EN < \infty$.

The sequence $\{\bar{x}_n, \mathfrak{F}_n;\ n = \ldots, m+1, m\}$ is a martingale with a last element $\bar{x}_m$, where $\mathfrak{F}_n$ denotes the σ-algebra generated by $\bar{x}_n, \bar{x}_{n+1}, \ldots$. From a theorem of Doob (ref. 2, p. 317), we obtain

LEMMA 1. *Let $f(\cdot)$ be a nonnegative continuous convex function on an interval I for which $P(x_1 \in I) = 1$. Then*

$$E(f^\alpha(\bar{x}_N)) \leq \left(\frac{\alpha}{\alpha - 1}\right)^\alpha E(f^\alpha(\bar{x}_m)) \qquad (\alpha > 1).$$

In particular, if $E(|\bar{x}_1|^\alpha) < \infty$ for some $\alpha > 1$, then $E|\bar{x}_N| < \infty$.

THEOREM 2. *Let $c_n = f(n)$, where $f(\cdot)$ is increasing and convex on $[0, \infty)$. If the x_i are nonnegative, then*

$$EN \leq 1 + f^{-1}\{Ex_1 + f(m-1)P(x_n \leq c_n \text{ for all } n \geq m)\}.$$

Proof: Since $c_n \to \infty$ as $n \to \infty$, $P(N < \infty) = 1$. Consider the random variable

$$\begin{aligned} M &= \text{largest } n \geq m \text{ such that } \bar{x}_n > c_n \\ &= m - 1 \text{ if no such } n \text{ exists; i.e., if } \bar{x}_n \leq c_n \text{ for every } m \leq n < \infty. \end{aligned}$$

Then $N \leq M + 1$, and $E\bar{x}_M = Ex_1$, since the event $M = k$ belongs to $\mathfrak{F}_k$ ($k \geq m - 1$). Hence

$$\begin{aligned} f(E(N-1)) &\leq f(EM) \leq E(f(M)) = \int_{(M\geq m)} f(M)dP + \int_{(M=m-1)} f(m-1)dP \\ &\leq \int_{(M\geq m)} \bar{x}_M dP + f(m-1)P(M = m-1) \leq \\ &\qquad E\bar{x}_M + f(m-1)P(M = m-1) \\ &= Ex_1 + f(m-1)P(x_n \leq c_n \text{ for all } n \geq m), \end{aligned}$$

and the theorem follows.

COROLLARY 1. *If $c_n = c(n + 1)$, $c > 0$, and the x_i are nonnegative, then*

$$EN \leq \frac{Ex_1}{c} + mP(\bar{x}_n \leq c_n \text{ for all } n \geq m).$$

Remarks: (*c*) Theorem 2 is an extension of results given in references 3 and 6.

(*d*) A slightly weaker form of Corollary 1 follows directly from Theorem 1. In fact, since $\bar{x}_N \geq c(N - 1)$ on $(N > m)$,

$$E(c(N-1)) \leq \int_{(N>m)} \bar{x}_N dP + \int_{(N=m)} c(m-1) \leq E\bar{x}_N + c(m-1)P\ (N = m),$$

so by Theorem 1,

$$EN \leq 1 + \frac{Ex_1}{c} + (m - 1)P(\bar{x}_m \leq c_m).$$

2. *Applications to Sequential Confidence Intervals.*—Let $y_1, y_2, \ldots$ be i.i.d. $N(\mu, \sigma^2)$ with μ, σ unknown. We seek a random interval I of fixed width $2d > 0$ such that $P(\mu \in I) \geq \gamma$, $0 < \gamma < 1$, for all values of μ, σ. If σ were known, then $I_n = (\bar{y}_n - d, \bar{y}_n + d)$ would be such an interval if $n \geq a^2\sigma^2/d^2$, with a defined by the equation

$$2\Phi(a) - 1 = \gamma,$$

$\Phi(\cdot)$ denoting the $N(0,1)$ distribution function. When σ is unknown, let $s_n^2 = \sum_{i=1}^n (y_i - \bar{y}_n)^2/(n - 1)$, $n \geq 2$, and for some fixed integer $m_0 \geq 2$ define as in reference 4

$$\begin{aligned} N_0 &= \text{smallest } n \geq m_0 \text{ such that } s_n^2 \leq d^2 n/a^2 \\ &= \infty \text{ if no such } n \text{ exists.} \end{aligned}$$

Then $P(N < \infty)$, and it is known[1, 5] that

$$\lim_{d \to 0} P(\mu \in I_{N_0}) = \alpha, \qquad \lim_{d \to 0} \frac{E(N_0)}{(a^2\sigma^2/d^2)} = 1.$$

To apply the results of section 1, we use the well-known fact that for $n \geq 2$, s_n^2 may be written as $s_n^2 = \sigma^2 \bar{x}_{n-1}$, where $x_1, x_2, \ldots$ are i.i.d. with a chi-squared distribution with 1 degree of freedom. Thus with $c_n = c(n + 1)$, $c = d^2/a^2\sigma^2$, $m = m_0 - 1$, we have $N_0 = N + 1$ where

$$N = \text{smallest } n \geq m \text{ such that } \bar{x}_n \leq c_n,$$

and $s_{N_0}^2 = \sigma^2 \bar{x}_N$. From Theorem 1 follows

COROLLARY 2. $E(s_{N_0}^2) < \sigma^2$.

Let $G_m(\cdot)$ denote the chi-squared distribution function with m degrees of freedom. Then from Corollary 1 we have

COROLLARY 3. $EN_0 < 1 + (a^2\sigma^2/d^2) + mG_m\,[m(m + 1)d^2/a^2\sigma^2]$, *and hence*

$$\limsup_{d \to 0} \left(EN_0 - \frac{a^2\sigma^2}{d^2} \right) \leq 1.$$

We now consider $P(\mu \not\in I_{N_0})$, which, because $\bar{y}_n$ is independent of $s_2^2, \ldots, s_n^2$, may be written as

$$\begin{aligned} P(\mu \not\in I_{N_0}) &= \sum_{n=m}^{\infty} P(N_0 = n) P(|\bar{y}_n - \mu| \geq d | N_0 = n) \\ &= 2 \sum_{n=m}^{\infty} P(N_0 = n) \left[1 - \Phi\left(\frac{d\sqrt{n}}{\sigma} \right) \right] = 2E\left(1 - \Phi\left(\frac{d\sqrt{N_0}}{\sigma} \right) \right). \end{aligned}$$

The function $(1 - \Phi(\beta\sqrt{\cdot}))^\alpha$ is convex on $[0, \infty)$ for any $\alpha > 0$ and $\beta > 0$, and hence from Corollary 3 and Jensen's inequality

$$P(\mu \not\in I_{N_0}) \geq 2\left(1 - \Phi\left(\frac{d}{\sigma}\sqrt{EN_0}\right)\right)$$

$$\geq 2\left\{1 - \Phi\left[\frac{d}{\sigma}\sqrt{1 + \frac{a^2\sigma^2}{d^2} + mG_m\left(\frac{m(m+1)d^2}{a^2\sigma^2}\right)}\right]\right\}$$

$$\geq 2\left\{1 - \Phi\left[a + \frac{d}{\sigma}\sqrt{1 + mG_m\left(\frac{m(m+1)d^2}{a^2\sigma^2}\right)}\right]\right\},$$

$$P(\mu \in I_{N_0}) \leq 2\Phi\left[a + \frac{d}{\sigma}\sqrt{1 + mG_m\left(\frac{m(m+1)d^2}{a^2\sigma^2}\right)}\right] - 1$$

$$\rightarrow 2\Phi(a) - 1 = \gamma, \text{ as } d \rightarrow 0.$$

To obtain a reverse inequality, we have by Lemma 1 for any $\alpha > 1$

$$P(\mu \not\in I_{N_0}) \leq 2E(1 - \Phi(a\sqrt{\bar{x}_N})) = 2E((1 - \Phi(a\sqrt{\bar{x}_N}))^{\frac{1}{\alpha}})^{\alpha}$$

$$\leq 2E(1 - \Phi(a\sqrt{\bar{x}_m}))\left(\frac{\alpha}{\alpha - 1}\right)^{\alpha}.$$

Letting $\alpha \rightarrow \infty$ we obtain

$$P(\mu \in I_{N_0}) \geq 2eE(\Phi(a\sqrt{\bar{x}_m})) - 2e + 1 = 2eF_m(a) - 2e + 1,$$

where $F_m(\cdot)$ denotes the Student t distribution function with m degrees of freedom. If m_0 is large, this states that *uniformly for all* $d > 0$

$$P(\mu \in I_{N_0}) > \gamma - 1.72(1 - \gamma);$$

e.g., for $\gamma = 0.99$, $P(\mu \in I_{N_0}) > 0.9728$ for all $d > 0$.

[1] Chow, Y. S., and H. Robbins, "On the asymptotic theory of fixed-width confidence intervals," *Ann. Math. Statist.*, **36**, 457–462 (1965).

[2] Doob, J. L., *Stochastic Processes* (New York: John Wiley & Sons, 1953).

[3] Simons, G., "On the cost of not knowing the variance when making a fixed-width confidence interval for the mean," *Stanford Univ., Statistics Dept. Tech. Rept.*, no. 28 (1967).

[4] Starr, N., "The performance of a sequential procedure for the fixed-width interval estimation of the mean," *Ann. Math. Statist.*, **37**, 36–50 (1966).

[5] Starr, N., "On the asymptotic efficiency of a sequential procedure for estimating the mean," *Ann. Math. Statist.*, **37**, 1173–1185 (1966).

[6] Starr, N., and M. B. Woodroofe, "On sequential fixed-width confidence intervals for the mean," *Carnegie-Mellon Univ., Statistics Dept., Tech. Rept.*, no. 8 (1967).