ON OPTIMAL STOPPING[1]

By Joseph A. Yahav

University of California, Berkeley

1. Introduction and summary. This paper considers the following problem. One takes independent and identically distributed observations from a population obeying a probability law $F_\theta(x)$. However, one does not know $F_\theta(x)$. What is known is a family of distribution functions $\Theta = \{F_\theta(x)\}$ and one assumes that there exists a prior probability measure $\mu(d\theta)$ on Θ.

At each stage $n = 1, 2, \cdots$ of the sampling one may stop. If one does stop he gets a payoff: $m_n - nC = Y_n$, where $m_n =$ maximum $(X_1, \cdots, X_n)$ and $C > 0$.

One is interested in a procedure which is optimal.

In Section 2 we define "procedure" and "optimal procedure." We show that under some conditions an "optimal procedure" exists. The optimal procedure turns out to be: Stop at stage j if $Y_j = \alpha(j)$ where $\alpha(j)$ is a function of $X_1, X_2, \cdots, X_j \mu(d\theta)$. The function $\alpha(j)$, although easy to describe, is quite difficult to calculate, hence we give in Section 3 two other functions which again are functions of $X_1, \cdots, X_j, \mu(d\theta)$ and are somewhat easier to calculate. These functions, denoted by $\beta_1(j)$, $\beta_2(j)$ have the following properties: $\beta_1(j) \leqq \alpha(j) \leqq \beta_2(j)$ and $\beta_2(j) - \beta_1(j) \to 0$ almost surely as $j \to \infty$.

Furthermore, we give an example for which

$$Y_j < \alpha(j) \Leftrightarrow Y_j < \beta_1(j) \Leftrightarrow Y_j < \beta_2(j).$$

The case for which $\mu(d\theta)$ is degenerate, namely the case for which $F_\theta(x)$ is known to the sampler was solved in [2]. The work done in [1], [2], [3] helped us to obtain the results in Sections 2 and 3. The work done in [4] and [5] deals with similar problems but the approach is different. All relations in this work are understood to hold almost surely unless otherwise specified.

2. Existence. Let the pair $(\Omega, \mathfrak{A})$ be an abstract space and a σ-field of subsets. Let $\{X_n : n = 1, 2, \cdots\}$ be a sequence of random variables on $(\Omega, \mathfrak{A})$. Let Θ be an abstract space such that for each $\theta \varepsilon \Theta$ corresponds a probability P_θ on Ω, so that $\{X_n\}$ are independent and identically distributed with respect to P_θ. Let $\mathfrak{F}$ stand for a σ-field of sets of Θ and let μ be a probability on $\mathfrak{F}$. Consider the measure $P = P_\theta \cdot \mu(d\theta)$ to be the probability measure on the product σ-field of $\Omega \cdot \Theta$. Let $F_\theta(x)$ be the distribution of X_1 with respect to P_θ. We impose

Received 5 September 1963; revised 15 August 1965.

[1] This research formed a portion of the author's doctoral dissertation submitted to the University of California, Berkeley, and was written with the partial support of the National Science Foundation, Grant GP-10. It was revised with the partial support of the Ford Foundation Grant admistered by the Center for Research in Management Science, University of California, Berkeley.

the condition $F_\theta(x)$ is measurable jointly in (θ, x). We assume that $E[|X_1|] < \infty$.

To define a procedure for stopping amounts to defining sets $B_1, B_2, B_3, \cdots$ with the following properties:

I. $B_1 \subset B_2 \subset B_3 \subset \cdots$

II. $B_i \, \varepsilon \, \mathcal{B}_i(X_1 \cdots X_i)$ where $\mathcal{B}_i(X_1 \cdots X_i)$ is the σ-field generated by $X_1 \cdots X_i$.

III. $P[\bigcup_{i=1}^\infty B_i] = 1$.

We say stop at stage n if $(X_1, \cdots, X_n, \cdots) \, \varepsilon \, B_n$ and $(X_1, \cdots, X_j, \cdots) \not\varepsilon \, B_j$ for $j = 1, 2, \cdots, n - 1$.

'he procedure described above defines a stopping time t which is a random variable $t(w) = n$ on $B_n - B_{n-1}$ or $t(w) \leqq n$ for all w in B_n, where $w \, \varepsilon \, \Omega$. Let T denote the collection of all stopping times which are defined through the above method. For any subcollection $T' \subset T$, s is optimal in T' if $s \, \varepsilon \, T'$, $E[Y_s]$ exists and $E[Y_s] \geqq E[Y_t]$ for all $t \, \varepsilon \, T'$, for which $E[Y_t]$ exists. Let T_N be the collection of stopping times $T_N = \{t \, \varepsilon \, T : t \leqq N\}$ where N is any positive integer.

It follows from [1] that in the set T_N there is a stopping time s_N which is optimal in T_N. Furthermore $E[Y_t]$ exists for all $t \, \varepsilon \, T_N \cdot s_N$ can be constructed in the following way:

Define inductively:

$$\alpha(N, N) = Y_N,$$

$$\alpha(N - n, N) = \max(Y_{N-n}, E[\alpha(N - n + 1, N) \mid X_1 \cdots X_{N-n}]) \qquad \text{for } n = 1, 2, \cdots, N - 1.$$

Set $S_{j,N} = \{x : Y_j = \alpha(j, N);\ Y_i < \alpha(i, N) \text{ for } i = 1, \cdots, j - 1\}$ for $j = 1, 2, \cdots, N$. Then the sets $\bigcup_{j=1}^n S_{jN}$ for $n = 1, 2, \cdots, N$ define the stopping time $s_N \, \varepsilon \, T_N$ which is optimal in T_N, namely $s_N = j$ on $S_{j,N}$.

It is interesting to note that the $\alpha(j, N)$ form a martingale on the set of continuation of sampling and

$$E\{Y_{s_N} \mid X_1 \cdots X_j) = \alpha(j, N) \qquad \text{on the set} \qquad (\bigcup_{i=1}^{j-1} S_{i,N})^c = \bigcup_{i=j}^N S_{i,N}.$$

Furthermore

$$\alpha(j, N) = \max_{t \varepsilon T_{j,N}} E[Y_t \mid X_1 \cdots X_j] \tag{2.1}$$

where $T_{j,N} = \{t \, \varepsilon \, T, j \leqq t \leqq N\}$. One might hope that the optimal stopping time in the class of all stopping times will be constructed by a limit operation on the optimal stopping times in the sequence $\{T_N\}$. Unfortunately this is not true in general and we will show that it is true for our problem under some conditions.

Let t be any stopping time. Define $t^{(N)} = t$ for $t \leqq N$ and $t^{(N)} = N$ for $t > N$. Then $t^{(N)}$ is the stopping time t truncated at N. Let $Y_n^* = m_n - n(C - \epsilon)$ for some $0 < \epsilon < C$.

LEMMA 2.1. *If* $\sup_{t \varepsilon T} E[Y_t^*] < \infty$ *then for any* $t \, \varepsilon \, T$

$$E[Y_t] = -\infty \Leftrightarrow E[t] = \infty.$$

PROOF. Suppose $E[t] = \infty$, then $E[Y_t] = E[Y_t^*] - \epsilon E[t] = -\infty$. Suppose $E[Y_t] = -\infty$, then we have $-\infty = E[Y_t] \geqq E[X_1] - CE[t]$, which proves the lemma.

THEOREM 2.1. *If* $\sup_{t\varepsilon T} E[Y_t^*] < \infty$ *then* $E[Y_{t^{(N)}}] \to E[Y_t]$ *as* $N \to \infty$.

PROOF. Case (i). $E[Y_t] = -\infty$.

We have $E[Y_{t^{(N)}}] \leqq E[Y^*_{t^{(N)}}] - \epsilon E[t^{(N)}]$. By Lemma 2.1 $E[t^{(N)}] \uparrow \infty$ and hence the theorem holds.

Case (ii). $E[Y_t] > -\infty$.

In this case we can write

$$E[Y_t] = E[m_t] - CE[t]$$

and

$$E[Y_{t^{(N)}}] = E[m_{t^{(N)}}] - CE[t^{(N)}].$$

Since $m_{t^{(N)}} \uparrow m_t$, $t^{(N)} \uparrow t$ and $E[t] < \infty$ the theorem follows.

Since $T_{j,N} \subset T_{j,N+1}$ it follows from (2.1) that $\alpha(j, N)$ is a nondecreasing sequence and hence we can introduce $\alpha(j) = \lim_N \alpha(j, N)$. By the monotone convergence theorem and the definition of $\alpha(j, N)$ we have

$$\alpha(j) = \max [Y_j, E(\alpha(j+1) \mid X_1, \cdots, X_j)]. \tag{2.2}$$

We define the random variable s in the following way: $s = j$ on the set $S_j = \{x: Y_j = \alpha(j), Y_i < \alpha(i), i = 1, \cdots, j-1\}$, $s = \infty$ on the complement of $\bigcup_{j=1}^{\infty} S_j$.

LEMMA 2.2. *If* $\sup_{t\varepsilon T} E[Y_t^*] < \infty$ *and there exists a sequence of random variables* $\{\beta_2(j)\}$ *such that:*

(a) $\beta_2(j)$ *is* $\mathcal{B}_j$ *measurable,*

(b) $\alpha(j) \leqq \beta_2(j)$,

(c) $\beta_2(j) = Y_j$ *for some* j *(where* j *may depend on* w*),*

(d) $\lim_{N\to\infty} E[\beta_2(N) - Y_N] = 0$,

then s *is a stopping time and*

$$\lim_{N\to\infty} E[Y_{s_N}] = E[Y_s]. \tag{2.3}$$

PROOF. It follows from (c), (b) and the definition of $\alpha(j)$ that $\alpha(j) = Y_j$ for some j (where j may depend on w) and so $p(s < \infty) = 1$. On the other hand, it is easily verified that $\{s \leqq j\} \varepsilon \mathcal{B}_j$ so that s is a stopping time. To prove (2.3) we note first that on the set $\bigcup_{i=j+1}^{\infty} S_i$ we have $\alpha(j) > Y_j$ and so by (2.2) $\alpha(j) = E[\alpha(j+1) \mid X_1 \cdots X_j]$. So we can write

$$\begin{aligned}
\lim_N E[Y_{s_N}] &= E[\alpha(1)] = \int \alpha(1)\, dP = \int_{\{s=1\}} \alpha(1)\, dP + \int_{\{s>1\}} \alpha(1)\, dP \\
&= \int_{\{s=1\}} \alpha(1)\, dP + \int_{\{s\geqq 2\}} \alpha(2)\, dP \\
&= \int_{\{s=1\}} \alpha(1)\, dP + \int_{\{s=2\}} \alpha(2)\, dP + \int_{\{s>2\}} \alpha(2)\, dP \\
&= \textstyle\sum_{i=1}^{N} \int_{\{s=i\}} Y_i\, dP + \int_{\{s>N\}} \alpha(N)\, dP \\
&= \textstyle\sum_{i=1}^{N} \int_{\{s=i\}} Y_i\, dP + \int_{\{s>N\}} Y_N\, dP + \int_{\{s>N\}} \alpha(N) - \int_{\{s>N\}} Y_N\, dP \\
&= \lim_N E[Y_{s^{(N)}}] + \lim_N [\int_{\{s>N\}} \alpha(N)\, dP - \int_{\{s>N\}} Y_N\, dP].
\end{aligned}$$

By Theorem 2.1 and the assumptions on $\beta_2(N)$ the lemma follows.

3. The functions $\beta_1(j)$, $\beta_1'(j)$, $\beta_2(j)$, $\beta_2'(j)$. Consider the following situation. One is told in stage j that in stage $j+1$ he will be informed which $F_\theta(x)$ he is sampling from. Then he faces the decision whether to stop or to take the $j+1$ observation. If he does take the $j+1$ observation, he may continue from stage $j+1$ and on, according to the optimal way given in [2].

Denote by $\beta_2'(j)$ the expected value of the following stopping rule. Take the $j+1$ observation. Then use the knowledge of $F_\theta(x)$ to determine the optimal stopping rule as is given in [2]. (There the optimal stopping rule is given subject to $E_\theta[X^2] < \infty$.) Let t_θ denote the stopping time which corresponds to the above stopping rule so that

$$(3.1) \qquad \beta_2'(j) = E[E_\theta(Y_{t_\theta}|X_1, \cdots, X_{j+1})|X_1, \cdots, X_j].$$

Let

$$\beta_2(j) = \max(Y_j, \beta_2'(j)).$$

Lemma 3.1. *$\beta_2'(j) \geqq E[\alpha(j+1)|X_1, \cdots, X_j]$ and $E[\beta_2(1)] \geqq \sup_{t\varepsilon T} E[Y_t]$.*

Proof. $E[\alpha(j+1, N) \mid X_1, \cdots, X_j]$ is the expected payoff for a stopping rule that requires one observation or more beyond stage j. $\beta_2'(j)$ is the expected payoff for the procedure that requires one observation or more beyond stage j but uses the knowledge of $F_\theta(x)$ at stage $j+1$ and on. In other words, for each θ there exists a stopping time t_θ which is a function of $X_1, \cdots, X_{j+1}, \theta$ which is best for this θ at stage $j+1$. Hence $E_\theta[Y_{t_\theta} \mid X_1, \cdots, X_{j+1}] \geqq E_\theta[Y_t \mid X_1, \cdots, X_{j+1}]$ for all $\theta \varepsilon \Theta$ and $j+1 \leqq t, t \varepsilon T$. Using (3.1) we have for any $t \varepsilon T$ and $t \geqq j+1$, $E[Y_t \mid X_1, \cdots, X_j] = E[E_\theta[Y_t \mid X_1, \cdots, X_{j+1}] \mid X_1, \cdots, X_j] \leqq \beta_2'(j)$. Since $E[\alpha(j+1, N) \mid X_1, \cdots, X_j]$ is the expected value of a stopping procedure for a certain $t \varepsilon T$ and $t \geqq j+1$, namely, stop for the first i such that $Y_i = \alpha(i, N)$ and $i \geqq j+1$, we obtain $\beta_2'(j) \geqq \lim_N E[\alpha(j+1, N) \mid X_1, \cdots, X_j] = E[\alpha(j+1) \mid X_1, \cdots, X_j]$. It follows by an argument similar to the one above that $E[\beta_2(1)] \geqq \sup_{t\varepsilon T} E[Y_t]$.

Lemma 3.2. *If $E_\theta[X^2] < \infty$ for all $\theta \varepsilon \Theta$ and $\int \gamma(\theta)\mu(d\theta) < \infty$ then $Y_j = \beta_2(j)$ for some j (where j may depend on w) and $\lim_{N\to\infty} E[\beta_2(N) - Y_N] = 0$.*

Proof. By [2] if $E_\theta[X^2] < \infty$, then there exists an optimal stopping rule for the case $F_\theta(x)$ known to the sampler, namely, stop for the first n such that $m_n \geqq \gamma(\theta)$ where $\gamma(\theta)$ is the unique solution of $\int_{\{x>\gamma(\theta)\}} (x - \gamma(\theta))\, dF_\theta(x) = C$. Moreover, $E_\theta[Y_s] = \gamma(\theta)$, hence $E[Y_s \mid X_1 \cdots X_j, m_j < \gamma(\theta), \theta] = \gamma(\theta) - jC$. So

$$\beta_2'(j) = \int_{\gamma(\theta)\leqq m_j} \{m_j P_\theta(X \leqq m_j) + \int_{m_j<x} x\, dF_\theta(x)\}\mu_j(d\theta)$$
$$+ \int_{\gamma(\theta)>m_j} \{\gamma(\theta)P_\theta(X \leqq \gamma(\theta)) + \int_{\gamma(\theta)<x} x\, dF_\theta(x)\}\mu_j(d\theta) - (j+1)C$$

where $\mu_j(d\theta)$ is the posterior probability on θ, at stage j. To prove the lemma it would be enough to show that $\beta_2'(j) - Y_j + C \to 0$ almost surely and in the first mean. We have

$$\beta'(j) - Y_j = U_j + V_j - C$$

where $U_j = \int_{\gamma(\theta)\leqq m_j} \{\int_{m_j<x} x\, dF_\theta(x) - m_j P_\theta(X > m_j)\}\mu_j(d\theta)$ and $V_j =$

$\int_{\gamma(\theta)>m_j} \{\gamma(\theta) - m_j + C\}\mu_j(d\theta)$. We will show first that U_j converges to 0 almost surely and in L_1; we shall denote this convergence by $U_j \rightarrow 0$. To show it, set $f_j = \int_{m_j<x} x\, dF_\theta(x) - m_j P_\theta(X > m_j)$. Set $g_j = E[f_j \mid X_1 \cdots X_j]$.

Note that (1) $0 \leqq f_j \leqq f_1$. (2) For each θ, $f_j \downarrow 0$ a.s. with respect to P_θ so that $f_j \downarrow 0$. (3) The g_j's form a reversed semi-martingale sequence, namely $E[g_{j+1} \mid X_1 \cdots X_j] \leqq g_j$. (4) The condition that $\{Y_n\}$ is an integrable sequence implies that $E[f_1] < \infty$. (5) By (4) and (1) $\sup_j E[g_j] \leqq E[f_1] < \infty$. Now (1), (2) and (4) imply that $E[f_j] \rightarrow 0$ which in turn imply that $E[g_j] \rightarrow 0$ by (3), (5) and the semi-martingale convergence theorem [6] $g_j \underset{1}{\overset{\text{a.s.}}{\rightarrow}} 0$. Which proves that $U_j \underset{1}{\overset{\text{a.s.}}{\rightarrow}} 0$. It remains to show that $V_j \underset{1}{\overset{\text{a.s.}}{\rightarrow}} 0$. Set

$$f_j' = I_{\{\gamma(\theta)>m_j\}}\{\gamma(\theta) - m_j + C\}, \qquad g_j' = E[f_j' \mid X_1 \cdots X_j],$$

where I_A is the indicator function of the A, and by the same argument as above we have $g_j' \underset{1}{\overset{\text{a.s.}}{\rightarrow}} 0$ which proves the lemma.

In order to prove the main theorem we have to impose a condition on $E[\beta_2(1)]$ for the case of sampling cost smaller than C. So we write $E[\beta_2(1, C)]$ and $\gamma(\theta, C)$ to denote $E[\beta_2(1)]$ and $\gamma(\theta)$ for sampling cost C.

THEOREM 3.1. *If $E_\theta[X^2] < \infty$ for every $\theta \in \Theta$ and $\int \gamma(\theta, C - \epsilon)\mu(d\theta) < \infty$ for some $0 < \epsilon < C$ then s defined in Section 2 is optimal in T.*

PROOF. Note first that $\int \gamma(\theta, C - \epsilon)\mu(d\theta) < \infty$ and the assumption on the integrability of the sequence $\{Y_n\}$ imply that $E[\beta_2(1, C - \epsilon)] < \infty$. Now by Lemmas 3.1 and 3.2 the conditions of Lemma 2.2 are satisfied and then by applying Theorem 2.1 we get Theorem 3.1.

Let $E[Y_{j+1} \mid X_1, \cdots, X_j] = \beta_1'(j)$ and $\beta_1(j) = \max(Y_j, \beta_i'(j))$. Define the conservative stopping rule in the following way. Stop for the first time $Y_j = \beta_1(j)$. By the following (3.1) the conservative stopping rule will stop before or with the optimal stopping rule. Define the optimist stopping rule as stop for the first time for which $Y_j = \beta_2(j)$. By the following (3.1) the optimist stopping rule will stop no sooner than the optimal stopping rule. (The optimist and the conservative stopping rules determine a stopping time according to our definition; this follows from Lemma 3.2.)

THEOREM 3.2. *Under the conditions of Theorem* 3.1

$$\beta_1(j) \leqq \alpha(j) \leqq \beta_2(j), \tag{3.1}$$

$$\beta_2(j) - \beta_1(j) \underset{1}{\overset{\text{a.s.}}{\rightarrow}} 0 \quad \text{as} \quad j \rightarrow \infty. \tag{3.2}$$

PROOF. To prove the left side of the inequality note that $\beta_1(j) = \alpha(j, j+1)$ and so $\beta_1(j) \leqq \alpha(j)$. To prove the right side of the inequality one uses (2.2) and Lemma 3.1.

For the second assertion recall that $Y_j \leqq \beta_1(j) \leqq \beta_2(i)$ and that we proved in Lemma 3.2 that $\beta_2(j) - Y_j \underset{1}{\overset{\text{a.s.}}{\rightarrow}} 0$ as $j \rightarrow \infty$.

The above theorem essentially implies that for small C the conservative rule and the optimist rule will be "close" and hence so would be the optimal rule.

In the following example the optimal stopping rule, the optimist stopping rule and the conservative stopping rule are the same. Namely,

$$Y_j < \alpha(j) \Leftrightarrow Y_j < \beta_1(j) \Leftrightarrow Y_j < \beta_2(j).$$

$$\Theta = \{\theta_1, \theta_2\}, \qquad \mu(\theta_1) = \mu(\theta_2) = \tfrac{1}{2}, \qquad C = 1.$$

$$\begin{aligned} F_{\theta_1}(x) &= 0, & x &< 0, \\ &= \tfrac{1}{2}, & 0 &\leqq x < 2, \\ &= 1, & 2 &\leqq x. \\ F_{\theta_2}(x) &= 0, & x &< 0, \\ &= \tfrac{1}{2}, & 0 &\leqq x < 4, \\ &= \tfrac{3}{4}, & 4 &\leqq x < 12, \\ &= 1, & 12 &\leqq x. \end{aligned}$$

In this case $\gamma(\theta_1) = 0$ and $\gamma(\theta_2) = 8$ and the conservative rule and the optimist rule will both tell us to continue sampling until we get 2 or 12 and to stop if we get 2 or if we get 12.

4. Acknowledgment. I wish to express sincere thanks to Professor David Blackwell for his interest and guidance under which this paper was written.

REFERENCES

[1] Arrow, K. J., Blackwell, D. and Girshick, M. A. (1949). Bayes and minimax solutions of sequential decision problems. *Econometrica* **17** 213–244.

[2] Chow, Y. S. and Robbins, H. (1961). A martingale system and applications. *Proc. Fourth Berkeley Symposium Math. Statist. Prob.* **1** 93–104.

[3] Chow, Y. S. and Robbins, H. (1963). On optimal stopping rules. *Z. Warscheinlichkeitstheorie und Verw. Gebiete.* **2** 33–49.

[4] MacQueen, J. and Miller, R. C., Jr. (1960). Optimal persistence policies. *J. Oper. Res. Soc. America.* **8** 362–379.

[5] Sakaguchi, M. (1961). Dynamic programming of some sequential sampling designs. *J. Math. Anal. Appl.* **2** 446–466.

[6] Doob, J. L. (1953). *Stochastic Processes.* Wiley, New York.