The Annals of Probability
1982, Vol. 10, No. 3, 745–753

OPTIMAL STOPPING ON AUTOREGRESSIVE SCHEMES[1]

BY MARK FINSTER

University of Wisconsin-Madison and Cornell University

For $\cdots \varepsilon_{-1}, \varepsilon_0, \varepsilon_1 \cdots$ i.i.d. random variables the autoregression $X_n = \varepsilon_n + a_1X_{n-1} + a_2X_{n-2} + \cdots$ yields a payoff $\gamma^n \sum_{-\infty}^{n} w_kX_k$ when stopped at time n, $0 < \gamma < 1$ being the discount factor. The optimal rule is characterized and under certain restrictions is the first passage time $t = \inf\{n : X_n \ge c\}$. As $c \to \infty$ the distributions of t and the remainder term $R_t = X_t - c$ are asymptotically independent and determined for exponential and algebraic tailed distributions on ε_n. An asymptotic expression for the optimal payoff is given and $c = c(\gamma)$ is calculated so that t yields a payoff asymptotically optimal and asymptotic to c as $\gamma \to 1$.

1. Introduction and summary. Let $\cdots \varepsilon_{-1}, \varepsilon_0, \varepsilon_1 \cdots$ be integrable real valued i.i.d. random variables and suppose $\cdots X_{-1}, X_0, X_1 \cdots$ is a stationary process satisfying the autoregressive equation

$$X_n = \varepsilon_n + a_1X_{n-1} + a_2X_{n-2} + \cdots \tag{1.1}$$

which has extensive applications in engineering (cf. Lapidus and Luus: 1967), econometrics (cf. Grenander and Rosenblatt, 1957, page 36) and control (cf. Box and Jenkins, 1976). Here $a = (a_1, a_2, \cdots)'$ lies within the unit disk D of ℓ^1; that is, $\sum |a_k| < 1$. All elements of ℓ^1, the absolutely summable sequences, are considered as infinite column vectors with the norm of any vector being the absolute sum of its coordinates and $'$ denotes transpose. Let $Z_n = (X_n, X_{n-1}, \cdots)'$ represent the past values or chart of $\{X_k\}_{k\le n}$ and let $w \in \ell^1$. This research investigates the optimal stopping problem $\{\gamma^n w'Z_n, \mathscr{F}_n\}$ when, as in Dubins and Teicher (1967), the future is discounted by a factor γ, $0 < \gamma < 1$. $\{\mathscr{F}_n\}$ is any increasing sequence of sigma algebras containing $\mathscr{F}(\varepsilon_k : k \le n)$ and independent of $\mathscr{F}(\varepsilon_k : k > n)$, the smallest sigma algebra generated by $\{\varepsilon_k : k > n\}$.

This model can be viewed as a first order approximation to the general problem of optimally stopping a sequence $\{X_n\}$ satisfying a stationarity equation

$$X_n = f(Z_{n-1}) + \varepsilon_n$$

when the undiscounted payoff or utility at time n is $u(Z_n)$. The linear coefficients of the Taylor expansions of the functions f and u are given by $\{a_k\}$ and $\{w_k\}$ respectively.

Specifically we seek the value $V(z, w, a) = \sup E(\gamma^t w'Z_t \mid Z_0 = z)$ where the supremum is taken over all stopping times t, t being a stopping time if $P(t \in \mathbb{N} \cup \{\infty\}) = 1$ and $\{t \le n\} \in \mathscr{F}_n \ \forall\, n$. On $t = \{\infty\}$ the payoff is zero. We also determine the existence and nature of an optimal stopping time σ yielding the extreme payoff $E(\gamma^\sigma w'Z_\sigma \mid Z_0 = z) = V(z, w, a)$.

Under specific assumptions the optimal rule is shown to be the first passage time $t_c = \inf\{n \ge 1 : X_n \ge c\}$ for some barrier c determined by Browder's Fixed Point Theorem. Under less stringent conditions we exhibit in Theorem 5.1 a barrier $c = c(\gamma) \to \infty$ as $\gamma \to 1$ which gives the stopping rule t_c a payoff that is asymptotically optimal for two different classes of distributions on ε_n. The right tail of a distribution function F dominates if $F(-x) = o(1 - F(x))$ as $x \to \infty$. A distribution function F is Paretian with exponent constant $\alpha > 1$, denoted $F \in \mathscr{P}_\alpha$, if F has a dominant right tail and varies regularly at ∞; that is, if there exists a function of $L(x)$ varying slowly at ∞ (Feller, 1966, page 276) such that $1 - F(x) = x^{-\alpha}L(x)$. A distribution function f is exponential with exponent constant

Received January 1980; revised January 1981.

[1] Research supported by NSF Grant No. MCS 75-22481

AMS 1970 *subject classifications.* Primary 62L15, 60G40; secondary 60G10, 60J05, 62P20, 90A20.
Key words and phrases. Autoregression, optimal payoff, first passage time, asymptotic distribution.

$\alpha > 0$, denoted $F \in \mathscr{E}_\alpha$, if F has a dominant right tail and $1 - F(x) \sim e^{-\alpha x}$ as $x \to \infty$. Throughout this paper F refers exclusively to the distribution function of ε_n and L refers to the slowly varying function associated with F.

Finster (1982) has shown that for such distributions $t = t_c$ and the remainder term $R_t = X_t - c$ properly normalized are asymptotically independent as $c \to \infty$. Furthermore the normalized stopping time converges in distribution to the exponential law $[1 - e^{-x}]I_{(0,\infty)}(x)$. In the exponential case R_t also converges in distribution to this exponential law and in the Paretian case the distribution of $R_t c^{-1}$ converges to the Pareto law $[1 - (x + 1)^{-\alpha}]I_{(0,\infty)}(x)$. Under certain restrictions the limits of Et and ER_t were obtained asymptotically via uniform integrability.

Utilizing these results, asymptotic limits for the optimal payoff are determined as $\gamma \to 1$. If ε_k is Pareto in nature $V(z, w, a) \sim e_1' w\phi(\gamma)$ where $\phi(\gamma) = \inf\{x : P(\varepsilon_k \geq x) \leq (\alpha - 1)\lambda\}$, $\lambda = -\log \gamma$, and $\{e_k\}$ is the standard basis for $\mathbb{R}^\infty$. The first passage time $t = \inf\{n : w'Z_n \geq e_1' w\phi\}$ is shown to be asymptotically optimal in that $V(z, w, a) \sim E(\gamma^t w' Z_t \mid Z_0 = z)$. If ε_k is exponentially distributed, $V(z, w, a) = \psi(\gamma) + o(1)$ where $\psi(\gamma) = -\log \lambda - \log\log \lambda + \log \beta$ for a given constant β. The payoff under t_ψ is shown to differ from $V(z, w, a)$ by $o(1)$ terms.

Dubins and Teicher (1967) as well as Darling, Liggett, and Taylor (1972) have investigated the random walk stopping problem $\{\gamma^n(x + S_n)\}$ where $S_n = \sum_1^n \varepsilon_k$. Since $\{x + S_n\}$ corresponds to the autoregression (1.1) when $X_0 = x$ and a is the boundary point e_1 of D, one might expect similar results. In Section 7 our results are shown to differ markedly; the optimal payoff of the random walk problem is asymptotic to $\mu/e\lambda$ as $\gamma \to 1$ provided $\mu = E\varepsilon_n > 0$ and $E\,|\,\varepsilon_n\,|^m < \infty$ for some $m > 1$.

2. The optimal stopping rule. Let $a = (a_1, a_2, \cdots)' \in \mathbb{R}^\infty$ and define $\{b_n\}$ inductively by $b_0 = 1$ and

$$b_n = a_1 b_{n-1} + \cdots + a_n b_0 \quad \text{for} \quad n > 0$$

so that $b = (b_0, b_1, \cdots)' = \sum_0^\infty S^k(a^{k*})$ where the $*$ product is convolution (cf. Finster, 1982, Section 2) and $S(a) = (0, a_1, a_2, \cdots)'$ is the right shift operator on $\mathbb{R}^\infty$. Throughout this paper a lies in the unit disc D of ℓ^1; that is $\sum |\,a_k\,| < 1$. Hence $b \in \ell^1$ since $|\,b\,| = \sum |\,b_k\,| < (1 - |\,a\,|)^{-1}$ and

$$X_n = \textstyle\sum_0^\infty b_k \varepsilon_{n-k} \tag{2.1}$$

is the unique stationary time series satisfying the autoregression (1.1). An alternate formulation of b is

$$b * (e_1 - S(a)) = e_1. \tag{2.2}$$

Define the transpose matrix $A' = (a, e_1, e_2, \cdots)$ and $r_n = e_1\varepsilon_n$ so that

$$Z_n = AZ_{n-1} + r_n = A^j Z_{n-j} + \textstyle\sum_{k=0}^{j-1} A^k r_{n-k} \tag{2.3}$$

is a stationary Markov sequence under the transition probabilities determined by $P_z(Z_1 \in C) = P(r_1 \in C - Az)$ for $C \in \beta(\Lambda)$—the Borel sets of the state space Λ. One possible state space is found by defining $\eta_n = (\varepsilon_n, \varepsilon_{n-1}, \cdots)'$ so that $X_n = b'\eta_n$ and $Z_n = B\eta_n$ with the upper triangular matrix $B = (b_{ij})$ defined having entries $b_{ij} = b_{j-i}(j \geq i)$. Note that for $w \in \ell^1$, $w'Z_n = (w * b)'\eta_n$ converges a.e. Hence a state space is $\Lambda = \Lambda(a, w) = \theta_n \cap \theta_w$ where

$$\theta_v = \{B\eta : \eta \in \mathbb{R} \quad \text{and} \quad \lim_{n\to\infty}\sup\, T^n(\bar{v} * \bar{b})'\bar{\eta} < \infty\}.$$

Here $\bar{v} = (|\,v_1\,|, |\,v_2\,|, \cdots)'$ for $v = (v_1, v_2, \cdots)' \in \mathbb{R}^\infty$ and $T(v) = (v_2, v_3, \cdots)'$ is the left shift operator. It is easy to show $P(Z_0 \in \theta_v) = 1\ \forall\ v \in \ell^1$.

We write $V(z, w, a)$ for the optimal payoff under P_z when the autoregression is determined by a and the payoff is scaled by w. By (2.3)

$$V(z, w, a) = \sup E(\gamma^t w' A^t z + \gamma^t w' \textstyle\sum_{k=0}^{t-1} A^k r_{t-k}) \tag{2.4}$$

where the supremum is taken over all stopping times for which the expectation above is defined and $\gamma^t(a'Z_t) = 0$ on $\{t = \infty\}$. If a and w are understood we simply write $V(z)$.

The stationary Markovian structure of $\{Z_n\}$ facilitates the selection of the 'best' optimal rule (more than one may exist). Faced at the nth stage with the decision of stopping or continuing, one would logically stop if the value of $\gamma^n w' Z_n$ is at least as great as, say, $V_n(Z_n)$—the optimal payoff if one knows the past Z_n and is to make at least one more observation X_{n+1}. In light of (2.3) we have when $Z_n = z$,

$$V_n(z) = \sup E(\gamma^t w' A^{t-n} z + \gamma^t w' \textstyle\sum_{k=0}^{t-n-1} A^k r_{t-k})$$

where the supremum is over all stopping times t stopping after time n. Thus $V_n(z) = \gamma^n V(z)$ and the natural prospect for an optimal stopping time is

$$\sigma = \inf\{n : w'Z_n \geq V(Z_n)\}. \tag{2.5}$$

Since $E_z(\sup \gamma^n(w'Z_n)^+) < \infty$, rigorous justification for σ's optimality and preference is given by Chow, Robbins, and Siegmund (1971) in Theorem 5.2, Theorem 4.5′, and Lemma 4.6 Here E_z represents expectation under P_z and $c^+ = \max\{c, 0\}$.

THEOREM 2.1. *If* $|a| < 1$, $w \in \ell^1$ *and* $E|\varepsilon_1| < \infty$, *then* $\sigma = \inf\{n \geq 1 : w'Z_n \geq V(Z_n)\}$ *is optimal.*

3. Autoregressions of finite order. If all but the first p coordinates of a and w are zero, consider them as elements of $\mathbb{R}^p$. In this case the definitive equation identical to (2.3) in form and use arises when z, w, r and $A = (a_{ij})$ are replaced by the vectors of $\mathbb{R}^p$ formed with their first p coordinates and by the $p \times p$ matrix $(a_{ij} : i \leq p, j \leq p)$ respectively. We do not distinguish between these representations and $\mathbb{R}^p$ is so embedded in ℓ^1. In this case the natural state space is $\mathbb{R}^p$ and if $z = (z_1, \cdots, z_p)' \in \mathbb{R}^p$, we write $V(z)$ for the value function corresponding to $V(\sum_1^p z_k e_k)$.

Now assume $w = e_1$ and $a = \rho e_1$ where $|\rho| < 1$ so that

$$X_n = \rho X_{n-1} + \varepsilon_n = \rho^n X_0 + \textstyle\sum_0^{n-1} \rho^k \varepsilon_{n-k} \tag{3.1}$$

is itself a stationary Markov sequence. The value function (2.4) becomes

$$V(x) = \sup E(\gamma^t \rho^t x + \gamma^t \textstyle\sum_{k=0}^{t-1} \rho^k \varepsilon_{t-k}), \qquad x \in \mathbb{R} \tag{3.2}$$

and the optimal rule (2.5) is $\sigma = \inf\{n : X_n \geq V(X_n)\}$. From (3.2) we see

$$|V(x) - V(y)| \leq \gamma |x - y| \tag{3.3}$$

and the contraction mapping principle produces a unique fixed point $c = V(c)$. Setting $y = c$ in (3.3) forces

$$V(x) \geq \gamma x + (1 - \gamma)c \quad \text{iff} \quad x \leq c.$$

Thus

$$\{x : x \geq V(x)\} = \{x : x \geq c\}$$

and

$$\sigma = \inf\{n : X_n \geq c\}.$$

If $a \in \mathbb{R}^p$ and $w = e_1$ then (3.3) holds for $x, y \in \mathbb{R}^p$ and the unique Borel measurable function $c = c(y_1, \cdots, y_{p-1})$ defined by $c = V(c, y_1, \cdots, y_{p-1})$ determines the optimal rule

$$\sigma = \inf\{n : X_n \geq c(X_{n-1}, \cdots, X_{n-d-1})\}.$$

Later we will need the average value function

$$U(w, a) = \sup E\gamma^t w' Z_t, \qquad w \in \ell^1, \qquad a \in D \tag{3.4}$$

the supremum being over all stopping times. Since σ is optimal under each P_z, $U(w, a) = E\gamma^\sigma w' Z_\sigma$. Note that if $a \in \mathbb{R}^p \cap D$ then

$$| U(e_1, a) - V(z, a)| \leq E \mid V(Z_0, a) - V(z, a)| \leq pE \mid X_0 \mid + \mid z \mid. \tag{3.5}$$

4. A lower bound on V. Letting $c = c(\gamma) \to \infty$ as $\gamma \to 1$ and utilizing the asymptotic joint distribution of $t = t_c$ with its overshoot $R_t = X_t - c$ as given in Finster (1982), the asymptotic payoff under $t = t_c$ is easily calculated and determines asymptotic lower bounds for the optimal payoff. Precisely if $\xi(1 - F(c)) \sim \lambda$ for $\xi > 0$ and $\lambda = -\log \gamma$ then Theorem 5.1 of Finster (1982) implies

$$c^{-1} E\gamma^t X_t \geq E\gamma^t = Ee^{-\lambda t} \to \frac{\beta}{\beta + \xi}, \qquad F \in \mathscr{E}_\alpha \tag{4.1}$$

where $\beta = E(\alpha a' Z_0)$. Noting that Theorem 4.1, Remark 4.1 and Corollary 4.2 of Finster (1982) apply to $\bar{t} = \inf\{n : | X_n | \geq c\}$ and $R_{\bar{t}} = X_{\bar{t}} - c$ gives

$$\begin{aligned} c^{-1} E_z \gamma^t X_t &\geq E_z \gamma^{\bar{t}} + c^{-1} E_z \gamma^{\bar{t}} R_{\bar{t}}, \qquad F \in \mathscr{P}_\alpha \\ &\sim E_z e^{-\lambda \bar{t}}(1 + E_z R_{\bar{t}} c^{-1}) \to \alpha/(\alpha - 1)(1 + \xi). \end{aligned} \tag{4.2}$$

If $a = \rho e_1 \in D$ and $w = e_1$, the first passage time $t = t_c$ was shown to be optimal providing $V(c) = E_c \gamma^t X_t = c$. Choosing ξ so that the payoff is asymptotically c gives a lower bound which later is shown to be nearly sharp as $\gamma + 1$, $\forall\, a \in D$.

LEMMA 4.1. *Set $\lambda = \log \gamma$ and suppose $\gamma \to 1$.*

a. $\liminf c^{-1} E_z \gamma^t X_t \geq 1$ *if* $\lambda(\alpha - 1) \sim 1 - F(c)$ *and* $F \in \mathscr{P}_\alpha$.
b. $\liminf c^{-1} E\gamma^t X_t \geq 1$ *if* $\lambda = o(1 - F(c))$ *and* $F \in \mathscr{E}_\alpha$.

PROOF. This follows directly from (4.1) and (4.2).

For $F \in \mathscr{E}_\alpha$ the asymptotic results can be improved and a payoff of c up to terms that are $o(1)$ is obtained for $c = \psi(\gamma)$ where

$$\psi(\gamma) = -\alpha^{-1} \log \lambda - \alpha^{-1} \log \log \lambda^{-1} + \alpha^{-1} \log \beta. \tag{4.3}$$

LEMMA 4.2. *Suppose $F \in \mathscr{E}_\alpha$, $F(0) = 0$, and each $a_k \geq 0$.*

a. *If $c = \psi + o(1)$ then $E\gamma^t X_t \geq \psi - \alpha^{-1} + o(1)$.*
b. *If, in addition, $a \in \mathbb{R}^p$ then $E\gamma^t X_t = \psi + o(1)$.*

PROOF. Set $s = e^{-\alpha} ct$ and assume $a \in \mathbb{R}^p$. Note that

$$\lambda = \beta e^{-\alpha c} (-\log \lambda)^{-1} e^{o(1)} \tag{4.4}$$

implies $\lambda e^{\alpha c} \to 0$. Theorem 5.1 of Finster (1982) indicates $E\gamma^t = E \exp(-\beta e^{\alpha c} s) \to 1$ and, together with Corollary 5.1 of Finster (1982) and (4.4), shows

$$\begin{aligned} E\gamma^t X_t &= cE\gamma^t + E\gamma^t R_t = \psi E e^{-\lambda t} + \alpha^{-1} + o(1) \\ &= \psi - \psi E(1 - \exp[-\beta(-\log \lambda)^{-1} e^{o(1)} s]) + \alpha^{-1} + o(1). \end{aligned} \tag{4.5}$$

Since $\sup\{x^{-1} | 1 - e^{-x} - x | : x > 0\} < \infty$ and $\beta \alpha^{-1}(-\log \lambda)^{-1} s$ converges in distribution to zero, the expectation in (4.5) equals

$$\beta(-\log \lambda)^{-1} e^{o(1)} E(s) + o([-\log \lambda]^{-1} E(s)).$$

Corollary 5.1 of Finster (1982) gives $E(s) \to \beta^{-1}$ and because $\alpha\psi \sim -\log \lambda$ we have established (b). By deleting the term $E\gamma^t R_t$, (a) follows as above.

5. First order autoregressions. Throughout this section $w = e_1$, $F(0) = 0$, and $a = \rho e_1$ where $0 < \rho < 1$. Let $\sigma = \inf\{n : X_n \geq c\} = t_c$ where $c = V(c)$; that is, σ is optimal. Set $\lambda = -\log \gamma$ and define $\phi(\gamma) = \inf\{x : 1 - F(x) \leq (\alpha - 1)\lambda\}$. We begin by showing that as $\gamma \to 1$, $t = t_\phi$ is asymptotically optimal with a payoff asymptotic to ϕ in that

$$\phi(\gamma) \sim c \sim V(x) \sim E_x\gamma^t X_t, \qquad \forall x \in \mathbb{R}, \qquad F \in \mathscr{P}_\alpha. \tag{5.1}$$

Given arbitrary sequences $\{\lambda_k = \lambda\}$, $\{\sigma_k = \sigma\}$ and $\{c_\sigma = c\}$, suppose $c \to \infty$ as $\gamma \to 1$ in such a manner that $\lambda c^\alpha / L(c) \to \xi$ for some ξ, $0 \leq \xi \leq \infty$. The slow variation of L implies

$$c\phi^{-1} \to (\alpha - 1)^{1/\alpha}\xi^{1/\alpha}. \tag{5.2}$$

Our strong positivity assumptions ensure that (4.2) remains valid when the inequality is replaced by equality. Coupled with (5.2), this yields for finite ξ

$$\phi^{-1}V(x) = (c/\phi)c^{-1}E_x\gamma^\sigma X_\sigma \to \alpha(\alpha - 1)^{(1-\alpha)/\alpha}\xi^{1/\alpha}(1 + \xi)^{-1}. \tag{5.3}$$

Since $\sigma = t_c$ is the optimal first passage time, this last limit must attain its maximum value of one, which occurs when $\xi = (\alpha - 1)^{-1}$ or equivalently when $c \sim \phi$. Replacing σ by $t = t_\phi$ in (5.3) completes (5.1) for finite ξ.

If ξ is infinite, (5.2) gives $\phi = o(c)$. Define $M_n = \max\{X_1, \cdots, X_n\}$. Since

$$\begin{aligned} P_x(\sigma \leq n) = P_x(M_n \geq c) &= \textstyle\sum_1^n P_x(X_k = \rho X_{k-1} + \varepsilon_k \geq c, M_{k-1} < c) \\ &\leq n[1 - F(c - \rho c)] \end{aligned} \tag{5.4}$$

we deduce that for sufficiently large c

$$E_x e^{-\lambda\sigma} \leq \lambda \int_0^\infty P_x(\sigma \leq y)e^{-\lambda y}\,dy \leq 2(c - c\rho)^{-\alpha}L(c)\lambda^{-1}$$

and so

$$\begin{aligned} \phi^{-1}E_x\gamma^\sigma X_\sigma \sim c\phi^{-1}E_x(e^{-\lambda\sigma}) &\leq 2(1 - \rho)^{-\alpha}c^{1-\alpha}L(c)\phi^{-1}\lambda^{-1} \\ &\sim 2(1 - \rho)^{-\alpha}(c/\phi)^{1-\alpha}L(c)/L(\phi) = o(1) \end{aligned}$$

contradicts σ's optimality via the lower bound given in Lemma 4.1a with $c = \phi$. Thus (5.1) is established.

The analogue of (5.1) for exponential F is

$$c = \psi(\gamma) + o(1) = V(z) + o(1), \qquad F \in \mathscr{E}_\alpha, \qquad z \in \mathbb{R} \tag{5.5}$$

where $\psi(\gamma)$ has been given in (4.3). We begin by showing

$$U(e_1, a) = \psi + o(1) = E\gamma^t X_t \quad \text{iff} \quad t = t_c \quad \text{with} \quad c = \psi + o(1) \tag{5.6}$$

where $U = E\gamma^\sigma X_\sigma$ is the average value function defined in (3.4). Note that σ's optimality and Lemma 4.2b assert

$$E\gamma^\sigma X_\sigma \geq E\gamma^t X_t = \psi + o(1) \quad \text{if} \quad t = t_\psi. \tag{5.7}$$

To prove (5.6), assume that for sequences $\{c_k = c\}$, $\{\lambda_k = \lambda\}$, $\{\psi_k = \psi\}$ and $\{v_k = v\}$ we have $c = \psi + v$. Set $s = e^{-\alpha c}\sigma$.

First suppose

$$\xi = \lambda e^{\alpha c} = \beta(-\log \lambda)^{-1}e^{\alpha v} \to 0 \tag{5.8}$$

so that by Theorem 5.1 and Corollary 4.1 of Finster (1982)

$$E(1 - e^{-\xi s}) = E(\xi s) + o(\xi) = \xi/\beta + o(\xi)$$

and

(5.9) $\quad E\gamma^\sigma X_\sigma = c - cE(1 - e^{-\lambda\sigma}) + Ee^{-\lambda\sigma}R_\sigma = (\psi + v) - (\psi + v)\xi\beta^{-1}\Delta + \alpha^{-1} + o(1)$

where $\Delta \to 1$. Thus (5.7) forces

$$v - (\psi + v)\xi\beta^{-1}\Delta + \alpha^{-1} \geq o(1)$$

or equivalently

$$\Delta \leq (\psi + v)^{-1}(-\log \lambda)e^{-\alpha v}[v + 1/\alpha + o(1)]$$

which can occur only with equality and only when $v \to 0$; that is, only when $c = \psi + o(1)$ and (5.6) holds.

Now (5.8) must hold, for if ξ has a positive or infinite limit, then

$$\liminf E(1 - e^{-\xi s}) > 0$$

so that as in (5.9)

$$E\gamma^\sigma X_\sigma < qc + o(c) \quad \text{for some} \quad q < 1.$$

This contradicts σ's optimality via (5.7) since

$$\limsup c/\psi \leq 1.$$

To obtain this bound, choose sequences $\{c_k = c\}$ etc. as above and set $\ell = -\alpha c/\log \lambda$. If $\ell \sim c/\psi$ has a limit greater than one then

$$P(\sigma \leq y) \leq P(M_y \geq c) \leq yP(X_1 \geq c) \sim y\beta e^{-\alpha c}$$

and an integration by parts implies

$$\psi^{-1}E\gamma^\sigma X_\sigma = \psi^{-1}cE\gamma^\sigma O(1) \sim O(1)\ell\lambda \int_0^\infty e^{-\lambda y}P(\sigma \leq y)\, dy$$

$$\leq O(1)\beta\ell e^{-sc}\lambda^{-1} = O(1)\beta\ell\lambda^{\ell-1} = o(1)$$

which contradicts (5.7).

We have established (5.6). To obtain (5.5) from (5.6), note that for $x \leq c$, (3.2) implies

$$0 \leq c - V(x) = V(c) - V(x) \leq E_c\rho^\sigma(c - x) = o(1).$$

The last equality follows by (5.4) and

$$E_c\rho^\sigma c = c \int_0^\infty (-\log \rho)\rho^y P_c(\sigma \leq y)\, dy$$

$$\leq c[1 - F(c - \rho c)](-\log \rho) \int_0^\infty \rho^y y\, dy = o(1).$$

Collecting the above results, we have the following.

THEOREM 5.1. *Let $F(0) = 0$ and $a = \rho e_1$ where $0 < \rho < 1$. If $c = V(c)$ then $\sigma = \inf\{n : X_n \geq c\}$ is optimal. Furthermore $\forall\, x \in \mathbb{R}$*

a. $V(x) \sim E_x\gamma^t X_t \sim \phi \sim c$ *if* $F \in \mathscr{P}_\alpha$ *and* $t = t_\phi$.
b. $c = \psi + o(1) = V(x) + o(1)$ *and* $U = E\gamma^t X_t$ *if* $F \in \mathscr{E}_\alpha$ *and* $t = t_\psi$.

6. The general case. The notation of Section 2 is used here extensively. If $a \in D$, the sequence $\{w'Z_n\}$ is itself autoregressive for certain $w \in \ell^1$. One such w, as indicated by (2.3), is any eigenvector $w = w(\rho)$ of A associated with an eigenvalue ρ, $|\rho| < 1$. Specifically, for the eigenvector

$$w(\rho) = (1, \rho, \rho^2, \cdots)' * (e_1 - S(a)) \tag{6.1}$$

of A, (2.3) implies that when $Z_0 = z$

$$w'Z_n = \rho^n(w'z) + \textstyle\sum_{k=0}^{n-1} \rho^k \varepsilon_{n-k}$$

is the simple autoregression (3.1) with the starting point $X_0 = w'z$. Hence the results of Sections 3 and 5 apply here. Restating Theorem 5.1 gives the following.

COROLLARY 6.1. *If $w = w(\rho)$, $|\rho| < 1$, and $t_c = \inf\{n : w'Z_n \geq c\}$ then $\sigma = t_c$ is optimal provided $c = V(c, \rho e_1, e_1)$. If, in addition, $F(0) = 0$ and $0 < \rho < 1$ then $\forall\, z \in \ell^\infty$*

a. *$V(z, w, a) \sim E_z\gamma^t w'Z_t \sim \phi \sim c$ if $t = t_\phi$ and $F \in \mathscr{P}_\alpha$.*
b. *$V(z, w, a) = c + o(1) = \psi + o(1)$ and $U(w, a) = E\gamma^t w'Z_t$ if $F \in \mathscr{E}_\alpha$ and $t = t_\psi$.*

Without loss of generality assume $e_1'w = 1$. In general $w'Z_n = (w * b)'\eta_n$ has the form of (2.1) and corresponds to an autoregression (1.1) with $a \in D$ replaced by $d \in D$ if there exists $d \in D$ such that

$$w * b = \textstyle\sum_1^\infty S^k(d^{k*}).$$

An algebraic manipulation of (2.2) shows this to be equivalent to

$$\hat{w} * a - T(\hat{w}) = d \in D \tag{6.2}$$

where $\hat{w} \in \mathbb{R}^\infty$ is defined by $\hat{w} * w = e_1$ and T is the left shift operator. For example, the eigenvector $w(\rho)$ corresponds to $d = \rho e_1$ and $w = e_1$ corresponds to $d = a$.

Let $W = W(a)$ be the set of w satisfying (6.2). It is now easy to extend the results of Corollary 6.1 to arbitrary $a \in D$ and $w \in W(a)$. Lemmas 4.1 and 4.2 provide lower bounds on the optimal payoff. To establish upper bounds the autoregression of concern is simply compared with one whose payoff is larger and known.

First consider the stopping problem $\{\gamma^n X_n\}$ where $F(0) = 0$ and $a \in D$ has nonnegative coordinates. For $|a| < \rho < 1$ the coordinates of $w = w(\rho)$ defined in (6.1) are positive and therefore $X_n \leq w'Z_n$. Thus $U(e_1, a) \leq U(w, a)$ and $V(z, e_1, a) \leq V(z, w(\rho), a)$ $\forall\, z \in \Lambda(e_1, a) \cap \Lambda(w, a)$. That this holds for arbitrary $z \in \Lambda(e_1, a)$ follows since

$$V(z_1, e_1, a) - V(z_2, e_1, a) = O(1) \qquad \forall\, z_1, z_2 \in \Lambda(e_1, a). \tag{6.3}$$

Hence Corollary 6.1 extends to

COROLLARY 6.2. *Let $F(0) = 0$, $a \in D$ and $e_k'a \geq 0$ $\forall\, k$.*

a. *If $F \in \mathscr{P}_\alpha$ and $t = t_\phi$ then $V(z, e_1, a) \sim \phi \sim E_z\gamma^t X_t$.*
b. *If $F \in \mathscr{E}_\alpha$ and $t = t_\psi$ then*

$$\psi - \beta^{-1} + o(1) \leq E\gamma^t X_t \leq U(e_1, a) \leq \psi + o(1) \quad \text{and} \quad V(z, e_1, a) \leq \psi + o(1).$$

c. *If, in addition to* (b), *$a \in \mathbb{R}^p$ then*

$$V(z, e_1, a) + O(1) = U(e_1, a) = \psi + o(1) = E\gamma^t X_t + o(1).$$

Note that in (c), $O(1) \leq pE|X_0| + |z|$ by (3.5).

When $a \in D$ and $F(0)$ are arbitrary, an upper bound is obtained through comparison of $\{X_n\}$ with the positive autoregression $\{\xi_n\}$ satisfying

$$\xi_n = |\varepsilon_n| + |a_1|\xi_{n-1} + |a_2|\xi_{n-2} + \cdots.$$

Fix the values of ε_k for $k \leq 0$ so that the corresponding $Z_0 = z$ lies in $\Lambda(e_1, \bar{a})$; then $X_n \leq \xi_n$ and thus

$$E_z\gamma^t X_t \leq E\gamma^t \xi_t, \quad t \text{ any stopping time}$$

which together with (6.3) and Corollary 6.2 gives an upper bound $\forall\, z \in \Lambda$. An analogous comparison holds for $U(e_1, a)$ giving the following.

COROLLARY 6.3. *Let $a \in D$, $w \in W(a)$ and $t_c = \inf\{n : w'Z_n \geq c\}$.*

a. *If $F \in \mathscr{P}_\alpha$ and $t = t_\phi$ then $V(z, w, a) \sim \phi \sim E_z\gamma^t w'Z_t \;\forall\, z \in \Lambda$.*
b. *If $F \in \xi_\alpha$ and $t = t_\psi$ then*

$$\psi + o(\psi) \leq E\gamma^t w'Z_t \leq U(w, a) \leq \psi + o(1) \quad \text{and} \quad V(z, w, a) \leq \psi + o(1).$$

c. *If, in addition to* (b), $a \in \mathbb{R}^p$ *then*

$$\psi + o(\psi) = E\gamma^t w'Z_t + o(\psi) \leq V(z, w, a) + O(1) = U(w, a) \leq \psi + o(1).$$

REMARK 6.1. If one could exhibit the uniform integrability of $t \;\forall\, a \in D$ the proof of Lemma 4.2 would yield $E\gamma^t X^t = \psi + o(1)$ and $o(\psi)$ could be replaced by $o(1)$ in (b). Also, for $a \in \mathbb{R}^p$, Fatou's lemma implies $V(Z_0, w, a) - E_{Z_0}\gamma^t X_t$ converges in distribution to zero. Hence a uniform integrability argument would give a payoff by t_ψ under P_z that is optimal up to $o(1)$ terms. The author has not yet been able to show this.

7. A Comparison of results. The random walk stopping problem $\{\gamma^n(x + S_n)\}$ where $S_n = \sum_1^n \varepsilon_k$ corresponds to the autoregression (1.1) when $Z_0 = xe_1$ and a is the boundary point e_1 of D. Define $t_c = \inf\{n : S_n \geq c\}$ and

$$V(x) = V(xe_1, e_1, e_1) = \sup E\gamma^t(x + S_t).$$

Dubins and Teicher (1967) proved the following.

THEOREM 7.1. *If c satisfies $c = V(c)$ then for $t = t_{c-x}$*

$$V(x) = E\gamma^t(x + S_t).$$

Although there exists these strong similarities in the formulations of the stopping problems and in the optimal stopping times, the asymptotic results differ markedly as shown by

THEOREM 7.1. *Assume $E\varepsilon_n = \mu > 0$ and $E|\varepsilon_n|^m < \infty$ for some $m > 1$. If $c = V(c)$ and $\lambda = -\log\gamma$ then $c \sim \mu/\lambda$ and $V(x) \sim \mu/e\lambda \;\forall\, x \in \mathbb{R}$.*

PROOF. Since $|V(x) - V(y)| \leq |x - y|$ assume $x = 0$. Later we'll show

$$\limsup \lambda c < \infty \tag{7.1}$$

so suppose the sequences $\{\lambda_k = \lambda\}$, $\{c_k = c\}$ and $\{t_c = t\}$ satisfy $\lambda c/\mu \to y$ for $0 \leq y < \infty$. The strong law of large numbers indicates $\hat{t} = c^{-1}\mu t \to 1$ w.p. one and since R_t converges in distribution (Feller, 1966, page 371) a standard uniform integrability argument indicates $\lambda E\gamma^t R_t \to 0$. Hence

$$\mu^{-1}\lambda E\gamma^t S_t = \mu^{-1}c\lambda E(\exp[-\mu^{-1}\lambda c\hat{t}]) + o(1) = ye^{-y} + o(1).$$

Since t is optimal $y = 1$, $c \sim \mu/\lambda$, and $V(x) \sim \mu/\lambda e$.

To see (7.1) assume $\lambda c = \xi \to \infty$ for sequences $\{\lambda_k = \lambda\}$ and $\{c_k = c\}$. If $1 < m < 2$, a result of Von Bahr and Esseen (1965)

$$E(S_n - \mu n^m)^{1/m} = o(n),$$

along with Kolomogorov's and Holder's inequalities, give the existence of a constant K independent of n and $y > 0$ such that

$$P(M_n \geq y) \leq K(n/y)^m$$

where $M_n = \max\{S_k : k \le n\}$. See Finster (1976) for the details. Thus

$$\lambda E\gamma^t S_t = \lambda c E e^{-\lambda t} + \lambda E\gamma^t R_t = \xi\lambda \int_0^\infty e^{-\lambda y} P(t \le y)\, dy + o(1)$$

$$= \xi\lambda \int_0^\infty e^{-\lambda y} P(M_y \ge c)\, dy + o(1)$$

$$= \xi\lambda c^{-m} \int_0^\infty e^{-\lambda y} y^m\, dy + o(1) = o(1)$$

contradicting the optimality of t and completing the proof.

Note that this random walk sequential rule performs no better asymptotically than the optimal nonsequential procedure which stops at N, the closest integer to λ^{-1}, and yields an expected payoff

$$\mu/\lambda e + \gamma^\gamma x + o(1).$$

If the second moments of ε_n exist, the results of Theorem 7.2 can be sharpened to $o(1)$ terms through the use of Blackwell's Renewal Theorem. See Woodroofe (1981). However when $1 < \alpha < 2$ the second moments of $F \in \mathscr{P}_\alpha$ do not exist.

REFERENCES

[1] Box, G. E. P. and Jenkins, G. M. (1975). *Time Series Analysis.* Holden-Day, San Francisco.
[2] Chow, Y. S., Robbins, Herbert, and Siegmund, David (1971). *Great Expectations: The Theory of Optimal Stopping.* Houghton-Mifflin, Boston.
[3] Darling, D. A., Liggett, T., and Taylor, H. M. (1972). Optimal stopping for partial sums. *Ann. Math. Statist.* **43** 1363–1368.
[4] Dubins, Lester E. and Teicher, Henry (1967). Optimal stopping when the future is discounted. *Ann. Math. Statist.* **38** 601–606.
[5] Feller, William (1966). *An Introduction To Probability Theory and Its Applications, 2.* Wiley, New York.
[6] Finster, Mark (1976). *Optimal Stopping on Autoregressive Schemes.* Unpublished Ph.D. Thesis. University of Michigan, Ann Arbor.
[7] Finster, Mark (1982). The maximum term and first passage times for autoregressions. *Ann. Probability* **10** 737–744.
[8] Grenander, Ulf and Rosenblatt, Murray (1957). *Statistical Analysis of Stationary Time Series.* Wiley, New York.
[9] Lapidus, Leon and Luus, Rein (1967). *Optimal Control of Engineering Processes.* Blaisdell; Waltham, Mass.
[10] Shiryayev, A. N. (1978). *Optimal Stopping Rules.* Springer-Verlag, New York.
[11] Von Bahr, and Esseen, C. G. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \le r \le 2$. *Ann. Math. Statist.* **36** 299–303.
[12] Woodroofe, Michael (1981). Nonlinear renewal theory in sequential analysis. Technical Report #106. University of Michigan, Ann Arbor.

Department of Statistics
University of Wisconsin-Madison
1210 W. Dayton Street
Madison, Wisconsin 53706