House-Hunting Without Second Moments

Thomas S. Ferguson, University of California, Los Angeles
Michael J. Klass[1], University of California, Berkeley

Abstract: In the house-hunting problem, i.i.d. random variables, $X_1, X_2, \ldots$ are observed sequentially at a cost of $c > 0$ per observation. The problem is to choose a stopping rule, N, to maximize $\mathrm{E}(X_N - Nc)$. If the X's have a finite second moment, the optimal stopping rule is $N^* = \min\{n \geq 1 : X_n > V^*\}$, where V^* satisfies $\mathrm{E}(X - V^*)^+ = c$. The statement of the problem and its solution requires only the first moment of the X_n to be finite. Is a finite second moment really needed? In 1970, Herbert Robbins showed, assuming only a finite first moment, that the rule N^* is optimal within the class of stopping rules, N, such that $\mathrm{E}(X_N - Nc)^- > -\infty$, but it is not clear that this restriction of the class of stopping rules is really required. In this paper it is shown that this restriction is needed, but that if the expectation is replaced by a generalized expectation, N^* is optimal out of all stopping rules assuming only first moments.

AMS 2000 Subject Classification: 60G40; 90A80.

Key Words: Optimal Stopping; Selling an Asset; Job Search.

1. Introduction. The house-hunting problem, also called the problem of selling an asset or the job search problem, was introduced and solved almost simultaneously in papers by MacQueen and Miller (1960), Derman and Sacks (1960), Chow and Robbins (1961) and Sakaguchi (1961). This problem may be described as follows. Independent, identically distributed random variables, $X_1, X_2, \ldots$ with common distribution function, $F(x)$, are observed sequentially at a cost of c per observation, where $c > 0$. We always assume that the expectation of the positive part of X is finite : $\mathrm{E}X^+ < \infty$, where X has distribution $F(x)$.

You must take at least one observation. If you stop after $n \geq 1$ observations, you receive X_n as a payoff, so your net return is $X_n - nc$. If you never stop, your payoff is defined to be $-\infty$ since $X_n - nc \to -\infty$ a.s. as $n \to \infty$ when $\mathrm{E}X^+ < \infty$.

The problem is to choose a stopping rule, N, to maximize $\mathrm{E}(X_N - Nc)$. It is customary to assume $F(x)$ has a finite second moment, or more generally that $\mathrm{E}\{(X_1^+)^2\} < \infty$. Under this assumption, the stopping rule

$$N^* = \min\{n \geq 1 : X_n > V^*\} \tag{1}$$

maximizes $\mathrm{E}(X_n - nc)$ among all stopping rules, where V^* satisfies

$$\int_{V^*}^{\infty} (x - V^*)\, dF(x) = c. \tag{2}$$

[1] Partially supported by NSF Grant DMS-0205054.

In addition, $V^* = \mathrm{E}(X_{N^*} - N^*c)$ is also the optimal expected return.

The statement of the problem and its solution requires only the first moment of $F(x)$ to be finite. In particular, the stopping rule (1) still gives expected return V^*. Yet the proofs of optimality of N^* seem to require that the second moment of $F(x)$ be finite. Is a finite second moment really needed? By an elegant direct argument based on Wald's equation and using only the assumption that $\mathrm{E}X^+ < \infty$, Robbins (1970) shows that the rule N^* given by (1) with V^* given by (2) is still optimal. However, he uses a slightly different definition of optimality, namely, he defines N^* to be optimal if it maximizes $\mathrm{E}(X_N - Nc)$ within the class of stopping rules, N, such that $\mathrm{E}(X_N - Nc)^- > -\infty$.

This seems innocuous enough. Who would like to accept a random reward the expectation of whose negative part is $-\infty$? The trouble is that this restriction also excludes payoffs the expectation of whose positive part is $+\infty$. If $\mathrm{E}(X_N - Nc)^- = -\infty$ and $\mathrm{E}(X_N - Nc)^+ < \infty$, then you don't need to exclude N. The rule N^* is definitely better. So a slightly stronger definition of optimality would be to restrict consideration to stopping rules N such that $\mathrm{E}(X_N - Nc)^+ < \infty$. This looks more questionable. Why should one exclude rules with infinite positive expectation?

Is a finite second moment necessary for N^* to be optimal out of all stopping rules?

2. Necessity of $\mathrm{E}\{(X^+)^2\} < \infty$. Robbins' result certainly provides an extension of the optimal property of the rule N^* that is valid even if $\mathrm{E}(X^+)^2 = \infty$. However, there are difficulties of interpretation that arise because of the restriction to stopping rules that satisfy $\mathrm{E}(X_N - Nc)^+ < \infty$. Restricting attention to such rules seems to say that any rule, N, with $\mathrm{E}|X_N - Nc| < \infty$, no matter how bad, is better than a rule whose expected payoff does not exist because $\mathrm{E}(X_N - Nc)^- = -\infty$ and $\mathrm{E}(X_N - Nc)^+ = +\infty$. For what W do you prefer a gamble giving you a payoff of \$$W$ to a gamble giving you \$$Z$, where Z is chosen from a standard Cauchy distribution? Answering such questions seems to require an extension to standard utility theory.

One important question that arises is whether or not there are some distributions of X with $\mathrm{E}(X^+)^2 = \infty$ for which all stopping rules have $\mathrm{E}(X_N - Nc)^+ < \infty$. Then, at least for some distributions with infinite second moment one could say that N^* is optimal among all stopping rules. It will be shown that for all distributions F with $\mathrm{E}(X^+)^2 = \infty$, there exist stopping rules, N, such that $\mathrm{E}(X_N - Nc)^- = -\infty$ and $\mathrm{E}(X_N - Nc)^+ = +\infty$.

Theorem 1. *If $X, X_1, X_2, \ldots$ are i.i.d. with $\mathrm{E}X^+ < \infty$ and $\mathrm{E}(X^+)^2 = \infty$, then the stopping rule,*

$$N = \min\{n \geq 1 : X_n \geq 2nc\} \tag{3}$$

satisfies $\mathrm{E}(X_N - Nc)^+ = \infty$.

Thus no new examples of optimality within the class of *all* stopping rules may be found using Robbins' result. Proofs are deferred to the appendix.

3. A Stronger Result. As it stands, Theorem 1 would not impress Robbins. Robbins requires that stopping rules stop with probability one. This contrasts with others that allow stopping rules, N, such that $\mathrm{P}(N = \infty) > 0$ (see for example the electronic text, Ferguson (2006)). This allows treatment of more general problems, for example bandit problems, but it requires specifying what the payoff will be if $N = \infty$. Generally, one may force the decision maker to use rules, N, that stop with probability one by choosing the payoff to be $-\infty$ if $N = \infty$.

However in this problem, the restriction to stopping rules that stop with probability one is very reasonable. All Theorem 1 really says is that if $\mathrm{E}(X^+)^2$ is infinite, a prophet can get an infinite expected return. He simply stops at N if there exists an n such that $X_n > 2nc$ and stops at $n = 1$ otherwise. It seems that those of us without superpowers must be satisfied with V^* or risk not stopping and so receiving infinite loss.

Therefore to satisfy Robbins, we need to answer the question: For what distributions of X with finite first moment and $\mathrm{E}(X^+)^2 = \infty$ is it true that there exists stopping rules N that stop with probability one and for which $\mathrm{E}(X_N - Nc)^+ = \infty$? The answer is contained in the following theorem.

Theorem 2. *If $X, X_1, X_2, \ldots$ are i.i.d. with $\mathrm{E}X^+ < \infty$ and $\mathrm{E}(X^+)^2 = \infty$, then there exists a stopping rule, N, with $\mathrm{P}(N < \infty) = 1$ such that $\mathrm{E}(X_N - Nc)^+ = \infty$, for all $c < \infty$.*

Thus there are no distributions with infinite positive second moment for which we may dispense with Robbins' restriction to stopping rules, N, such that $\mathrm{E}(X_N - Nc)^+ < \infty$.

It is interesting to note that the stopping rule, N, of Theorem 2 does not depend on c. The proof is constructive. In addition, the stopping rule has the simple form, $N = \min\{n : X_n > a_n\}$ for some sequence of constants, a_n.

4. Optimality of N^* among all stopping rules. To extend Robbins' result to make it valid for all stopping rules, we must therefore find some way to compare two payoff distributions whose first moments don't exist. Certainly if given the choice between two Cauchy distributions with the same interquartile range, we would prefer the one with the higher median. More generally, we would prefer F to G if F stochastically dominates G (i.e. if $F(x) < G(x)$ for all x). There are many ways to extend this idea further. One sufficient for the problem at hand is the following.

We say that a lottery from a distribution G is preferred to a 0 payoff if for i.i.d. $Z_1, Z_2, \ldots$ from G we have $(1/n)\sum_1^n Z_i \xrightarrow{a.s.} \gamma$ for some $0 < \gamma \leq \infty$. For a distribution G with finite mean, this just means that a lottery from G is preferred to 0 if the mean of G is positive. For a distribution G whose mean does not exist, it can still happen that $(1/n)\sum_1^n Z_i \xrightarrow{a.s.} +\infty$, in which case we prefer G to 0. This can happen if the mass on the positive axis dominates the mass on the negative axis, even though the expectation of the positive and negative parts are both infinite.

Similarly, we prefer 0 to a lottery from G if for i.i.d. $Z_1, Z_2, \ldots$ from G we have $(1/n)\sum_1^n Z_i \xrightarrow{a.s.} -\infty$. More generally, we prefer a lottery for G_1 to a lottery for G_2 if for i.i.d. $Y_1, Y_2, \ldots$ from G_1 and independent i.i.d. $Z_1, Z_2, \ldots$ from G_2, we have $(1/n)\sum_1^n (Y_i - Z_i) \xrightarrow{a.s.} \gamma$ for some $0 < \gamma \leq \infty$.

Using this extension of the definition of preference between lotteries, one can show that Robbins' result is true without restricting the stopping rules one considers. In the paper of Robbins and Samuel (1966), an extension of the definition of mathematical expectation is given which is useful in this context.

Definition. *For a random variable X, we say that the extended expectation of X exists and is equal to μ, in symbols $\hat{\mathrm{E}}X = \mu$, if*

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mu \quad \text{a.s.} \tag{5}$$

when $X_1, X_2, \ldots$ are i.i.d. with the same distribution as X.

If $\mathrm{E}X$ exists, then $\hat{\mathrm{E}}X = \mathrm{E}X$, including the case where $\mathrm{E}X = \pm\infty$. However, if $\mathrm{E}X$ does not exist, it still may happen that $(1/n)\sum_1^n X_i$ converges almost surely to $+\infty$ or $-\infty$. Thus, $\hat{\mathrm{E}}$ is indeed an extension of the notion of expectation.

Using this notion, we can state the optimality of the stopping rule (1) for the house hunting problem.

Theorem 3. *In the house hunting problem with finite first moment, the stopping rule N^* of (1) is optimal in the sense that for all stopping rules N, $\hat{\mathrm{E}}(X_N - Nc) \leq V^*$.*

In other words, if $\mathrm{E}(X^+) < \infty$, then for any stopping rule, N, either $X_N - Nc$ has finite expectation less than or equal to V^*, or its extended expectation is $-\infty$.

5. A Near Counterexample. If the first moment of X barely exists in the sense that $\mathrm{E}X^+ < \infty$ and $\mathrm{E}X^+ \log^+(X^+) = \infty$, then there is a stopping rule that looks as if it might be a counterexample to Theorem 3.

Theorem 4. *If $\mathrm{E}X^+ < \infty$ and $\mathrm{E}X^+ \log^+(X^+) = \infty$, then there exists a stopping rule of the form $N = \min\{n \geq 1 : X_n > a_n\}$ for some sequence $a_n \to \infty$ such that $\mathrm{P}(N < \infty) = 1$, and*

$$\sum_{n=1}^{\infty} \mathrm{E}\{(X_n - nc)\,\mathrm{I}(N = n)\} = \sum_{n=1}^{\infty} \mathrm{P}(N > n-1)\mathrm{E}\{(X - nc)\,\mathrm{I}(X > a_n)\} = \infty \tag{6}$$

for all $c > 0$.

Note that, again, N does not depend on c.

To see how curious this result is, examine the second sum in (6). The stopping rule N stops with probability 1, and when it stops at stage n, the conditional payoff given $N = n$

is simply $\mathrm{E}\{(X - nc)\,\mathrm{I}(X > a_n)\}$, a fairly large positive number, even though stopping may occur at negative values of $X_n - nc$. This has to be multiplied by the probability of reaching that stage which is $\mathrm{P}(N > n-1) = \prod_{i=1}^{n-1} F(a_i)$. The product of this and $\mathrm{E}\{(X - nc)\,\mathrm{I}(X > a_n)\}$ is the summand of the second sum, and is a fairly small number. Nevertheless, if you add up all these small numbers, you get $+\infty$. Doesn't that seem better than receiving V^* as the payoff?

The catch is, of course, that this summation is not equal to $\mathrm{E}(X_N - Nc)$, which doesn't exist. This is an example where the expectation of the sum is not the sum of the expectations. Worse, the sum of the expectations is $+\infty$ while the generalized expectation of the sum is $-\infty$; in other words, if you take a sample from the distribution of $X_N - Nc$, the average of the sample will tend to $-\infty$, a.s.

APPENDIX

Proof of Theorem 1. We show that the rule $N = \min\{n \geq 1 : X_n \geq 2nc\}$ gives $\mathrm{E}(X_N - Nc)^+ = \infty$ when $\mathrm{E}X^+ < \infty$ and $\mathrm{E}(X^+)^2 = \infty$.

$$\begin{aligned} \mathrm{E}(X_N - Nc)^+ &= \sum_{n=1}^{\infty} \mathrm{E}(X_n - nc) I(N = n) \\ &\geq \sum_{n=1}^{\infty} nc\mathrm{P}(N = n) \\ &= \sum_{n=1}^{\infty} nc\mathrm{P}(N > n-1)\mathrm{P}(X_n > 2nc) \\ &\geq \sum_{n=1}^{\infty} nc\mathrm{P}(N = \infty)\mathrm{P}(X_n > 2nc) = \infty, \end{aligned} \tag{6}$$

since $\mathrm{E}(X^+)^2 = \infty$ implies $\sum_{n=1}^{\infty} n\mathrm{P}(X_n > 2nc) = \infty$, and

$$\begin{aligned} \mathrm{P}(N = \infty) &= \mathrm{P}(X_n < 2nc \quad \text{for all} \quad n) \\ &= \prod_{n=1}^{\infty} \mathrm{P}(X_n < 2nc) = \prod_{n=1}^{\infty} (1 - \mathrm{P}(X_n \geq 2nc)) \\ &\sim \exp\{-\sum \mathrm{P}(X_n > 2nc)\} \sim \exp\{-\mathrm{E}X^+/2c\} > 0. \ \blacksquare \end{aligned} \tag{7}$$

Proof of Theorem 2. Without loss of generality, assume that $X > 0$ and that X has a continuous distribution function. (Otherwise, only stop when $X > 0$ and replace the distribution of X with the distribution of XU where U has a uniform(0,1) distribution independent of X.) Since $\mathrm{E}(X^+)^2 = \infty$, we have

$$\sum_{k=1}^{\infty} \int_k^{\infty} \mathrm{P}(X > y)\, dy = \infty. \tag{8}$$

Let

$$E_k = \int_k^\infty \mathrm{P}(X > y)\,dy \qquad \text{and} \qquad Q_k = E_1 + E_2 + \cdots + E_k, \tag{9}$$

for $k > 0$. Then, $\sum_k E_k = \infty$ and $Q_k \to \infty$. It is also true that

$$\sum_{k=1}^\infty \frac{E_k}{Q_k} = \infty. \tag{10}$$

See, for example, Rudin (1976), Problem 11(b) page 79. Let n_* be the smallest k such that $Q_k > 1$. Let a_{n_*} be defined by

$$\mathrm{P}(X \le a_{n_*}) = \frac{1}{Q_{n_*}} \tag{11}$$

and for $k > n_*$, let a_k satisfy

$$\mathrm{P}(X \le a_k) = \frac{Q_{k-1}}{Q_k}. \tag{12}$$

Let $N = \min\{k \ge n_* : X_k > a_k\}$. Notice that for $n > n_*$.

$$\begin{aligned}\mathrm{P}(N > n) &= \mathrm{P}(\cap_{k=n_*}^n \{X_k \le a_k\}) \\ &= \frac{1}{Q_{n_*}} \prod_{k=n_*+1}^n \frac{Q_{k-1}}{Q_k} = \frac{1}{Q_n} \to 0.\end{aligned} \tag{13}$$

Hence N stops with probability one. (This may also be seen using $\sum_{k>n_*+1} \mathrm{P}(X > a_k) = \sum_{k>n_*+1}(1 - (Q_{k-1}/Q_k)) = \sum_{k>n_*+1}(E_k/Q_k) = \infty$; so $\mathrm{P}(X_n > a_n \text{ i.o.}) = 1$.)

Notice that for any $1 < c < \infty$,

$$\frac{\mathrm{P}(X > a_k)}{\mathrm{P}(X > ck)} = \frac{E_k}{Q_k \mathrm{P}(X > ck)} \ge \frac{(c-1)k\mathrm{P}(X > ck)}{Q_k \mathrm{P}(X > ck)} = \frac{(c-1)k}{Q_k} \to \infty. \tag{14}$$

In particular, $Q_k < ck$ from some point on.

Fix any $c > 1$. There exists n_c such that $Q_k \le ck$ for all $k \ge n_c$. Therefore,

$$\begin{aligned}\mathrm{E}(X_N - Nc)^+ &= \sum_{k=n_c}^\infty \int_{ck}^\infty \mathrm{P}(X > y)\,dy \mathrm{P}(N > k-1) = \sum_{k=n_c}^\infty \frac{E_{ck}}{Q_{k-1}} \\ &\ge \sum_{k=n_c}^\infty \frac{E_{ck}}{Q_{ck}} > \frac{1}{2c} \sum_{j=cn_c}^\infty \frac{E_j}{Q_j} = \infty.\end{aligned} \quad \blacksquare \tag{15}$$

Proof of Theorem 3. We use the cute idea, cited in the paper of Robbins (1970) as due to David Burdick, entailed in the inequality

$$\begin{aligned} X_n - nc &= v + (X_n - v) - nc \\ &\le v + (X_n - v)^+ - nc \\ &\le v + \sum_1^n (X_i - v)^+ - nc \\ &= v + \sum_1^n ((X_i - v)^+ - c) \\ &= v + \sum_1^n W_i \end{aligned} \tag{16}$$

where $W_i = (X_i - v)^+ - c$. Now choose v to be any number greater than V^*. The W_i are i.i.d. with expectation $\mathrm{E}W_i < 0$ since $v > V^*$. Let N be an arbitrary stopping rule. We show below that $\hat{\mathrm{E}} \sum_1^N W_i < 0$; this implies that $\hat{\mathrm{E}}(X_N - Nc) < v$ and, since v is an arbitrary number greater than V^*, that $\hat{\mathrm{E}}(X_N - Nc) \le V^*$.

We now use the idea of Blackwell (1946) in his proof of Wald's Equation. Consider n stopping problems as follows. Let N_1 be the stopping rule N applied to the sequence $W_1, W_2, \ldots$, let N_2 be the stopping rule N applied to $W_{N_1+1}, W_{N_1+2}, \ldots$, etc., and let N_n be the stopping rule N applied to $W_{N_1+\cdots+N_{n-1}+1}$, $W_{N_1+\cdots+N_{n-1}+2}, \ldots$. Let the returns for these problems be denoted by $Z_1, \ldots, Z_n$ where $Z_j = W_{N_1+\cdots+N_{j-1}+1} + \cdots + W_{N_1+\cdots+N_j}$. Then, the Z_j are independent with the same distribution as $\sum_1^N W_i$, and we have

$$\frac{Z_1 + \cdots + Z_n}{n} = \frac{W_1 + \cdots + W_{N_1+\cdots+N_n}}{N_1 + \cdots + N_n} \cdot \frac{N_1 + \cdots + N_n}{n}. \tag{17}$$

From the strong law of large numbers, the first term on the right of (17), $(W_1 + \cdots + W_{N_1+\cdots+N_n})/(N_1 + \cdots + N_n)$, converges a.s. to $\mathrm{E}W_i < 0$. The second term on the right of (17), $N_1 + \cdots + N_n)/n$ converges a.s. to $\mathrm{E}N$, whether $\mathrm{E}N$ is finite or $+\infty$. Therefore, the left side of (17) converges a.s. to $\mathrm{E}W_i\mathrm{E}N < 0$. This shows that $\hat{\mathrm{E}} \sum_1^N W_i < 0$ as was to be shown. ■

Proof of Theorem 4. Again assume without loss of generality that $X > 0$ and that X has a strictly increasing continuous distribution function on $(0, \infty)$. Let

$$\varphi(b) = \mathrm{E}(X | X > b) = \frac{\mathrm{E}(X\,\mathrm{I}(X > b))}{\mathrm{P}(X > b)}.$$

Then $\varphi(b)$ is increasing and continuous and the inverse function, $b(y) = \varphi^{-1}(y)$ is well-defined by

$$\varphi(b(y)) = \mathrm{E}(X | X > b(y)) = y \tag{18}$$

for $y \geq \mathrm{E}(X)$. Clearly, $b(y)$ is increasing, $\lim_{y\to\infty} b(y) = \infty$, and $b(y) < y$. Then

$$\begin{aligned}\sum_{n=1}^{\infty} \mathrm{P}(X > b(n)) &= \sum_{n=1}^{\infty} \frac{1}{n}\mathrm{E}\{X\,\mathrm{I}(X > b(n))\} \\ &= \sum_{n=1}^{\infty}(\sum_{j=1}^{n}\frac{1}{j})\mathrm{E}\{X\,\mathrm{I}(b(n) < X \leq b(n+1))\} \\ &\geq \sum_{n=1}^{\infty}\mathrm{E}\{X\log(\varphi(X)/2)\,\mathrm{I}(b(n) < X \leq b(n+1))\}\end{aligned} \tag{19}$$

Therefore, using $\varphi(y) > y$,

$$\sum_{n=1}^{\infty} \mathrm{P}(X > b(n)) \geq \mathrm{E}\{X\log(X/2)\,\mathrm{I}(X > b(1))\} = \infty. \tag{20}$$

There exist constants, γ_k, increasing to infinity such that

$$\sum_{k=1}^{\infty} \mathrm{P}(X > b(k\gamma_k)) = \infty. \tag{21}$$

Let $a_k = b(k\gamma_k)$. Notice that as in the first line of (19),

$$\mathrm{P}(X > a_k) = \frac{\mathrm{E}\{X\,\mathrm{I}(X > a_k)\}}{k\gamma_k} = o(1/k). \tag{22}$$

There exists a k_* such that

$$\mathrm{P}(X > a_k) < \frac{1}{2k} \qquad \text{for all } k \geq k_*. \tag{23}$$

Let $N = \min\{k \geq k_* : X_k > a_k\}$. Then from (23) and Borel-Cantelli, $\mathrm{P}(N < \infty) = 1$. Now fix any $c > 1$ and choose k_c so that $a_k > b(2ck)$ for all $k \geq k_c$. Notice that for $k > k_*$,

$$\begin{aligned}\mathrm{P}(N > k) &= \mathrm{P}(X_j \leq a_j \text{ for } k_* \leq j \leq k) \\ &= \prod_{j=k_*}^{k}(1 - \mathrm{P}(X > a_j)) \\ &\geq \prod_{j=k_*}^{k}\frac{2j-1}{2j} \\ &> \frac{2k^* - 1}{2k} \geq \frac{1}{2k}\end{aligned} \tag{24}$$

Putting these inequalities together,

$$
\begin{aligned}
&\sum_{n=k_c}^{\infty} \mathrm{P}(N > n-1)\mathrm{E}\{(X - cn)\,\mathrm{I}(X > a_n)\} \\
&\qquad = \sum_{n=k_c}^{\infty} \mathrm{P}(N > n-1)\mathrm{E}((X - cn)\,\mathrm{I}(X > b(n\gamma_n))) \qquad \text{(def. of } a_n) \\
&\qquad = \sum_{n=k_c}^{\infty} \mathrm{P}(N > n-1)[n\gamma_n - cn]\mathrm{P}(X > b(n\gamma_n)) \qquad \text{(from (18))} \\
&\qquad > \frac{1}{2}\sum_{n=k_c}^{\infty} (\gamma_n - c)\mathrm{P}(X > b(n\gamma_n)) \qquad \text{(from (24))} \\
&\qquad = \infty \qquad\qquad \text{(since } \gamma_n \to \infty \text{ and (21))} \qquad \blacksquare
\end{aligned}
$$

References.

Blackwell, D. (1946). On an equation of Wald, *Annals of Mathematical Statistics* 17: 84-87.

Chow, Y. S., and Robbins, H. (1961). A martingale system theorem and applications, *Proceedings of the Fourth Berkeley Symposium on Mathathematical Statistics and Probability*, 1: 93-104.

Derman, C., and Sacks, J. (1960). Replacement of periodically inspected equipment (an optimal stopping rule), *Naval Research Logistics Quarterly* 7: 597-607.

Ferguson, T. S. (2006). *Optimal Stopping and Applications*, Electronic Text at http://www.math.ucla.edu/~tom/Stopping/Contents.html

MacQueen, J., and Miller, R. G., Jr. (1960). Optimal persistence policies, *Operations Research* 8: 362-380.

Robbins, H. (1970). Optimal stopping, *Americam Mathematical Monthly*, 77: 333-343.

Robbins, H., and Samuels, E. (1966). An extension of a lemma of Wald, *Journal of Applied Probability* 3: 272-273.

Rudin, W. (1976). *Principles of Mathematical Analysis, 3rd Edition*, McGraw-Hill.

Sakaguchi M. (1961). Dynamic programming of some sequential sampling design, *Journal of Mathematical Analysis and Applications* 2: 446-466.