AN OVERVIEW OF OPTIMAL STOPPING TIMES FOR VARIOUS DISCRETE TIME GAMES

THESIS

Presented in Partial Fulfillment of the Requirements for

the Degree of Master of Science in the Graduate

School of the Ohio State University

By

Tyrus Hunter Berry, B.S.

* * * * *

The Ohio State University
2008

Approved by

Thesis Committee:

Professor Boris Pittel, Advisor

Professor Fanyang Zheng

Advisor
Graduate Program in Mathematics

ABSTRACT

The best strategy for three discrete time games is shown to be equivalent to a stopping time. These stopping times are found and uniqueness is considered. The games considered have sequences of random rewards which the player observes one at a time. In the first game the player must pay for each observation but can quit and take the highest reward he has seen at any time. In the second game the player can only take the last reward seen and there are only finitely many rewards to view. In the final game the player can only pick one reward but he wins only if he has chosen the highest reward of all the draws (so he must beat all the draws that come after his choice).

To Dr. Boris Pittel

VITA

2006-Present	Graduate Teaching Associate, The Ohio State University
2006	B.S. in Mathematics, The University of Virginia
2006	B.A. in Physics, The University of Virginia

FIELDS OF STUDY

Major Field: Mathematics

Specialization: Probability Theory and Stochastic Processes

TABLE OF CONTENTS

CHAPTER 1

OPTIMAL STOPPING

Imagine a game or process in which we seek hidden rewards X_i with known distribution. We may only have one reward and the rewards are uncovered one at a time in order but we have to pay a price c_i to see the i^{th} reward. In the first version of the game we may quit at any time and we may take the highest reward we have seen so far. We shall call this game the Simple Hat Game With Memory. Formally, we define the return upon quitting at time n as:

$$Y_n := \max_{1 \leq m \leq n} \{X_m\} - \sum_{i=1}^{n} c_i$$

We want to find the best strategy for this game in the sense of having the highest expected return at the lowest risk. Note that at each stage of the game we only need to decide whether to quit or continue playing, and we can only quit once. Thus a strategy for this game is a function which assigns to each sequence of events $X_1, X_2, ...$ a quitting time. Furthermore, since we cannot see the future, a decision to quit at time n can only depend upon $X_1, ..., X_n$ and $c_1, ..., c_n$. Thus a strategy is equivalent to a stopping time, so among all stopping times τ we must find that which maximizes $E[Y_\tau]$ and if this stopping time is not unique we should also try to minimize $\text{var}[Y_\tau]$.

We first determine the maximum expected return of any stopping strategy τ; meaning we find an upper bound for $E[Y_\tau]$. In what follows we assume $\{X_i\}$ are independent, identically distributed (i.i.d.) random variables, with $E[X_i] < \infty$, which are also independent from $\{c_i\}$ which are also i.i.d. random variables. First note that $(X_m - \alpha)^+ + \alpha = \max\{X_m, \alpha\} \geq X_m$ which gives us the following upper bound on Y_n:

$$Y_n \leq \max_{1 \leq m \leq n} \{(X_m - \alpha)^+\} + \alpha - \sum_{i=1}^{n} c_i \leq \alpha + \sum_{m=1}^{n} [(X_m - \alpha)^+ - c_m]$$

From the above assumptions it is clear that $Z_m := (X_m - \alpha)^+ - c_m$ are i.i.d. random variables, and thus (setting $S_n := \sum_{m=1}^{n} Z_m$) for any stopping time τ we have $E[S_\tau] = E[Z_1]E[\tau]$ which gives the following bound:

$$E[Y_\tau] \leq \alpha + E[S_\tau] = \alpha + E[Z_1]E[\tau]$$

This leads us to the following result:

Lemma 1.0.1. *If there is an α such that $E[(X_1 - \alpha)^+] = E[c_1]$ then $E[Y_\tau] \leq \alpha$ for any strategy τ such that $E[\tau] < \infty$.*

Proof. Since α was chosen such that $E[Z_1] = 0$ the result follows from the above bound. $\square$

This is clearly unsatisfying if there is no such α, so before going on we should establish the following lemma:

Lemma 1.0.2. *If $0 < E[c_1] \leq E[X_1]$ then there exists and α such that $E[(X_1 - \alpha)^+] = E[c_1]$.*

Proof. Let $f(\alpha) = E[(X_1 - \alpha)^+]$. Note that $f(\alpha) \leq \epsilon$ for α large enough, and $f(0) = E[X_1]$ and furthermore f is a continuous function of α and hence for any $c \in (0, E[X_1])$ there is a solution to $f(\alpha) = c$. $\square$

The following theorem establishes a surprisingly simple strategy that is optimal in the sense of expected returns; we simply quit as soon as we observe an X_i that exceeds α from the above lemma.

Theorem 1.0.3. *Let α be as defined by Lemma 0.1 and assume that $P(X_1 > \alpha) \neq 0$. Set $T := \inf\{n : X_n > \alpha\}$, then $E[Y_T] = \alpha$ and thus T achieves the maximum expected return established above.*

Proof. The proof is simply a computation of $E[Y_T]$. Note that since T is the first time that $X_i > \alpha$ we have

$$\max_{1 \leq i \leq T} \{X_i\} = X_T$$

Thus, setting $C_n := \sum_{i=1}^{n} c_i$, we can compute:

$$E[Y_T] = E[X_T - C_T] = E[X_T] - E[C_T] = E[X_T] - E[c_1]E[T]$$

Let $p = P(X > \alpha)$ and $q = 1 - p = P(X \leq \alpha)$. We can now compute:

$$E[X_T] = \sum_{n=1}^{\infty} E[X_T 1_{T=n}] = \sum_{n=1}^{\infty} E[X_n; X_n > \alpha, \forall (i < n)(X_i \leq \alpha)]$$

$$= \sum_{i=1}^{\infty} E[X_n; X_n > \alpha] q^{n-1} = \frac{1}{p} E[X_1; X_1 > \alpha]$$

and

$$E[T] = \sum_{n=1}^{\infty} n P(X_n > \alpha, \forall (i < n)(X_i \leq \alpha)) = p \sum_{n=1}^{\infty} n q^{n-1} = \frac{1}{p}$$

Thus combining these computations we may now use the choice of α such that $E[c_1] = E[(X_1-\alpha)^+]$ which lets us compute:

$$E[Y_T] = \frac{1}{p}(E[X_1; X_1 > \alpha] - E[c_1]) = \frac{1}{p}(E[X_1; X_1 > \alpha] - E[(X_1-\alpha)^+])$$

$$= \frac{1}{p}(E[X_1; X_1 > \alpha] - E[X_1 - \alpha; X_1 > \alpha]) = \frac{1}{p}E[\alpha; X_1 > \alpha] = \alpha$$

Thus the simple stopping time T achieves the maximal expected return $E[Y_T] = \alpha$. $\square$

At this point the α parameter is becoming somewhat mysterious, since it is defined by the equation $E[(X_1-\alpha)^+] = c$ which can be very difficult to solve. Thus I provide a simple example to ground the discussion which follows and which will help establish non-uniqueness of the optimal solution.

Example 1: Suppose that for \$1 we may draw from a hat containing papers labeled \$1,\$2,...,\$10 each equally probable. After being drawn the number is recorded and replaced and we may either take the dollar amount listed or pay another \$1 to draw again. We may then continue to draw as many times as we like, paying \$1 per draw, and when we quit we are paid the highest dollar amount that was drawn. In this example, the cost is $c = 1$ and the rewards $X_i \in \{1, 2, ..., 10\}$ are uniformly distributed, thus to find the optimal stopping time from Theorem 0.3 we must solve the following equation for α:

$$1 = E[(X_1-\alpha)^+] = \frac{1}{10}\sum_{X_1=1}^{10}(X_1-\alpha)^+$$

Since we know that the solution is unique, we may assume α is integer valued and our computation will validate this fact in the end. This allows us to simplify the above expression as:

$$= \frac{1}{10}\sum_{X_1=\alpha}^{10}(X_1-\alpha) = \frac{1}{10}\sum_{X_1=\alpha}^{10}X_1-\frac{\alpha}{10}(10-\alpha+1) = \frac{1}{10}\left(\frac{10*11}{2}-\frac{(\alpha-1)\alpha}{2}\right)-\frac{11\alpha}{10}+\frac{\alpha^2}{10}$$

Setting this expression equal to $c = 1$ and simplifying the quadratic we have:

$$0 = \alpha^2 - 21\alpha + 90 = (\alpha - 6)(\alpha - 15)$$

Thus $\alpha = 6$ is the unique solution which validates our assumption that α should be integer valued. This implies that the optimal strategy is to wait until we see a value that is strictly greater than \$6 and then quit. Theorem 0.3 tells us that the expected return for this strategy is $\alpha =$\$6 and that no other strategy has a higher expected return. However Theorem 0.3 does not assert uniqueness of the optimal strategy and thus other strategies may have the same optimal expected return of \$6. In fact other such strategies exist although they are not very different, we will show later that there is a unique optimal strategy with minimal variance (risk), but first:

Theorem 1.0.4. *The optimal stopping time of Theorem 0.3 is not unique.*

Proof. Consider the stopping time T^* defined by:

$$T^* = \begin{cases} 1 & X_1 = 6 \\ \inf\{n : X_n > 6\} & \text{else} \end{cases}$$

This is a valid stopping time because $\{T^* = 1\}$ is measurable with respect to X_1 and otherwise T^* returns the normally optimal stopping time. To see that T^* is an optimal stopping time, we will need $E[Y_1; X_1 > 6]$ which we can compute by applying the result from Theorem 1 which says $E[Y_T] = 6$ and thus:

$$6 = E[Y_T] = \sum_{n=1}^{\infty} E[Y_n; X_n > 6] P(X_1 \le 6)^{n-1} = E[Y_1; X_1 > 6] \frac{10}{4}$$

Where in the last equality we used the fact that $P(X_1 \le 6) = 6/10$. This calculation shows that $E[Y_n; X_n > 6] = 12/5$. Now we can compute $E[Y_{T^*}]$:

$$E[Y_{T^*}] = E[Y_1; T^* = 1] + \sum_{n=2}^{\infty} E[Y_n; X_n > 6] P(X_1 \le 5) P(X_2 \le 6) \cdots P(X_{n-1} \le 6)$$

$$= E[Y_1; X_1 > 5] + \frac{5}{10} E[Y_1; X_1 > 6] \frac{10}{4} = 3 + \frac{5}{4}\frac{12}{5} = 6$$

Thus T^* is an optimal stopping time. □

Intuitively this new strategy should have a lower variance since it 'stops sooner' than our simple strategy. We will now formalize the connection between variance of a strategy and the expected stopping time, which will require Wald's second equation:

Theorem 1.0.5 (Wald's Second Equation)**.** *Let $\{X_i\}$ be i.i.d. with $E[X_i] = 0$ and $E[X_i^2] < \infty$. If T is a stopping time with $E[T] < \infty$ then $E[S_T^2] = E[X_i^2]E[T]$.*

Proof. We will compute $E[S_{T\wedge n}^2]$ inductively, thus first note that when $T < n$ we have $T \wedge n = T = T \wedge (n-1)$, thus we have:

$$E[S_{T\wedge n}^2] = E\left[\sum_{i,j=1}^{n} X_i X_j\right] = E\left[S_{T\wedge(n-1)}^2 + 1_{\{T\ge n\}}\left(2X_n \sum_{i=1}^{n-1} X_i + X_n^2\right)\right]$$

Thus we can compute the difference between successive cutoff times, noticing that the event $\{T \ge n\} = \{T < n\}^c \in \sigma(X_1, ..., X_{n-1})$ and hence is independent of X_n.

$$E[S_{T\wedge n}^2] - E[S_{T\wedge(n-1)}^2] = E[1_{\{T\ge n\}} 2X_n S_{n-1}] + E[1_{\{T\ge n\}} X_n^2] = P(T \ge n) E[X_n^2]$$

Where there first term in the second expression is zero since X_n is independent of S_{n-1} and $\{T \geq n\}$ and $E[X_n] = 0$. Now that we know the difference between successive cutoffs, we can compute $E[S^2_{T\wedge n}]$ as a telescoping sum:

$$E[S^2_{T\wedge n}] - E[S^2_{T\wedge 1}] = \sum_{i=2}^{n}(E[S^2_{T\wedge i}] - E[S^2_{T\wedge(i-1)}]) = \sum_{i=2}^{n} E[X_1^2]P(T \geq i)$$

Note that $E[S^2_{T\wedge 1}] = E[S_1^2] = E[X_1^2] = E[X_1^2]P(T \geq 1)$, so adding this term to both sides of the previous equality we have:

$$E[S^2_{T\wedge n}] = E[X_1^2]\sum_{i=1}^{n} P(T \geq i) \to E[X_1^2]E[T]$$

Were we know that the right hand side converges to $E[X_1^2]E[T]$ as $n \to \infty$ since $E[T] = \sum_{i=1}^{\infty} P(T \geq i) < \infty$ by assumption. Intuitively, as $n \to \infty$ the left hand side should converge to $E[S^2_{T\wedge n}] \to E[S_T^2]$, however we need to know that this expression does in fact converge before we can claim this. Note that we cannot use dominated convergence since $E[S_n^2]$ is unbounded and boundedness of $E[S_T^2]$ is what we are trying to establish; thus we will show that $S_{T\wedge n}$ is a Cauchy sequence in L^2. Let $n \geq m \geq M$, we need to bound:

$$E[(S_{T\wedge n} - S_{T\wedge m})^2] = E[1_{\{T\geq n\}}(S_n - S_m)^2 + 1_{\{m<T<n\}}(S_T - S_m)^2 + 1_{\{T\leq m\}}(S_T - S_T)^2]$$
$$= \sum_{i=n+1}^{\infty} E[1_{\{T=i\}}(S_n - S_m)^2] + \sum_{i=m+1}^{n} E[1_{\{T=i\}}(S_i - S_m)^2]$$

Where the last term, with $T \leq m$, is clearly zero. Note that $(S_i - S_m)^2 = \sum_{k,j=m+1}^{i} X_k X_j$ and since $1_{\{T=i\}} \leq 1$ and $E[X_k X_j] = 0$ when $k \neq j$ we eliminate all these terms, leaving:

$$\leq \sum_{i=n+1}^{\infty} E\left[1_{\{T=i\}}\sum_{j=m+1}^{n} X_j^2\right] + \sum_{i=m+1}^{n} E\left[1_{\{T=i\}}\sum_{j=m+1}^{i} X_j^2\right]$$

$$\leq \sum_{i=m+1}^{\infty} E\left[1_{\{T=i\}} \sum_{j=m+1}^{n} X_j^2\right]$$

Thus switching the order of summation (since the terms are all positive) we have:

$$= \sum_{j=m+1}^{\infty} E\left[X_j^2 \sum_{i=j}^{\infty} 1_{\{T=i\}}\right]$$

$$= \sum_{j=m+1}^{\infty} E[X_j^2 1_{\{T \geq j\}}] = E[X_1^2] \sum_{j=m+1}^{\infty} P(T \geq j) \to_{M \to \infty} 0$$

Where the limit is zero since $E[T] < \infty$. So we conclude $S_{T \wedge n}$ is Cauchy in L^2 and hence $E[S_{T \wedge n}^2] \to E[S_T^2]$ as $n \to \infty$ and so finally we may conclude that $E[S_T^2] = E[X_1^2]E[T]$. $\square$

Thus, Wald's Second Equation tells us that the variance of a stopped sum is proportional to the expected stopping time, this will let us choose the minimal variance strategy among all optimal strategies but first we must characterize all strategies which are optimal in the sense of expected return:

Theorem 1.0.6. *Let T with $E[T] < \infty$ be an optimal stopping time in the sense that $E[Y_T] = \alpha$. Then with probability 1,* $\inf\{n : X_n \geq \alpha\} \leq T \leq \inf\{n : X_n > \alpha\}$.

Note: this means that an optimal strategy can stop anytime after observing an $X_n = \alpha$ or continue, but an optimal strategy must stop if $X_n > \alpha$.

Proof. Since T is optimal, $\alpha = E[Y_T] = E[\max_{1 \leq i \leq T}\{X_i\}] - E[c_1]E[T]$. We solve this expression for $E[T]$ and then use the upper bound $\max_{1 \leq i \leq T}\{X_i\} \leq \alpha + \sum_{i=1}^{T}(X_i - \alpha)^+$

$$E[T] = \frac{E[\max_{1 \leq i \leq T}\{X_i\}] - \alpha}{E[c_1]} \leq \frac{E\left[\alpha + \sum_{i=1}^{T}(X_i - \alpha)^+\right] - \alpha}{E[c_1]} = \frac{E\left[\sum_{i=1}^{T}(X_i - \alpha)^+\right]}{E[c_1]}$$

$$= \frac{E[T]E[(X_1 - \alpha)^+]}{E[c_1]} = E[T]$$

Where in the last equality we used that α was chosen so that $E[(X_1 - \alpha)^+] = E[c_1]$. Note that since the first and last quantities are in fact equal, this implies that the inequality is in fact equality, and thus:

$$E\left[\max_{1\le i\le T}\{X_i\}\right] = E\left[\alpha + \sum_{i=1}^{T}(X_i - \alpha)^+\right]$$

thus, since the expected value of the difference is zero and since we know the right hand side is always greater than or equal to the left hand side we can conclude that:

$$P\left(\max_{1\le i\le T}\{X_i\} = \alpha + \sum_{i=1}^{T}(X_i - \alpha)^+\right) = 1$$

Thus with probability 1 there can be only one $i \in 0, ..., T$ such that $X_i > \alpha$. This implies that $T < \inf\{n : X_n > \alpha\}$ with probability 1. Furthermore, with probability one, at least one $i \in 0, ..., T$ must have $X_i \geq \alpha$, which implies $T \geq \inf\{n : X_n \geq \alpha\}$ with probability one. □

This theorem has told us what all the optimal strategies look like, and, since strategy variance is proportional to expected stopping time, we have the following corollary:

Corollary 1.0.7 (Uniqueness of Minimal Variance Strategy)**.** *The stopping time* $T = \inf\{n : X_n \geq \alpha\}$ *is the unique optimal strategy (*$E[Y_T] = \alpha$*) which minimizes* $\mathrm{var}[Y_T]$.

Proof. Since any stopping time τ with $E[Y_\tau] = \alpha$ has $\tau \geq T$, if $\tau \neq T$ then $E[\tau] > E[T]$ (since every sequence has positive probability) and therefore, by Wald's Second Equation, $\mathrm{var}[X_\tau] > \mathrm{var}[X_T]$. □

We now turn to choosing random variables with rejection, meaning we no longer have access to the maximum value; once we make an observation the previous value is permanently rejected. We first introduce the optimal stopping time for two independent choices.

Lemma 1.0.8. *Let X_1, X_2 be independent random variables with $E[X_1], E[X_2] < \infty$ then the optimal stopping time for maximizing the expected value of the stopped random variable is given by $T = 1 \Leftrightarrow X_1 \geq E[X_2]$, in the sense that for any stopping time τ we have $E[X_\tau] \leq E[X_T]$.*

Proof. Let τ be a stopping time. The key point is that, since X_1 and X_2 are independent, and since $\{\tau = 1\}$ if measurable with respect to $\sigma(X_1)$, we have X_2 and $1_{\{\tau \neq 1\}}$ are independent, thus:

$$\begin{aligned} E[X_\tau] &= E[X_1 1_{\{\tau=1\}}] + E[X_2 1_{\{\tau\neq1\}}] = E[X_1 1_{\{\tau=1\}}] + E[X_2] - E[X_2]E[1_{\{\tau=1\}}] \\ &= E[X_2] + \int_{\{\tau=1\}} (x - E[X_2]) dP_{X_1}(x) \end{aligned}$$

Since we can choose τ to maximize this sum, we want to choose $\tau = 1$ if an only if the integrand is non-negative. Thus $\tau = 1$ if and only if $X_1 \geq E[X_2]$ so $\tau = T$ is the optimal stopping time. $\square$

Now we can compute the return $E[X_T]$ of the optimal strategy:

$$E[X_T] = E[X_2] + \int_{x \geq E[X_2]} (x - E[X_2]) dP_{X_1}(x) = E[X_2] + E[(X_1 - E[X_2])^+]$$

and the expected stopping time for the optimal strategy $E[T]$:

$$E[T] = 1 * P(X_1 \geq E[X_2]) + 2 * (1 - P(X_1 \geq E[X_2]) = 2 - P(X_1 > E[X_2])$$

We can now generate the optimal strategy for n independent random variables inductively:

Theorem 1.0.9. *Let $X_1, ..., X_n$ be independent random variables with finite expected value. Define inductively constants $\alpha_1, ..., \alpha_n$ by:*

$$\alpha_n := E[X_n]$$

$$\alpha_{n-1} := E[X_n] + E[(X_{n-1} - E[X_n])^+]$$

$$\alpha_{n-2} := \alpha_{n-1} + E[(X_{n-2} - \alpha_{n-1})^+]$$

$$\vdots$$

$$\alpha_{n-j} := \alpha_{n-j+1} + E[(X_{n-j} - \alpha_{n-j+1})^+]$$

Then the optimal stopping time is given by:

$$T = \inf\{i : X_i \geq \alpha_i\}$$

Proof. The proof is by induction on n, the case $n = 2$ is done by the previous Lemma. Assume the result is true for $n = k$, and let $X_1, ..., X_{k+1}$ be independent, $\alpha_2, ..., \alpha_{k+1}$ be the k constants that define the optimal stopping time for the sequence $Y_1, ..., Y_k$ where $Y_i = X_{i+1}$. Then, by the inductive hypothesis, for any stopping time τ we have:

$$E[X_\tau] = E[X_1 1_{\{\tau=1\}}] + E[X_\tau 1_{\{\tau>1\}}] = E[X_1 1_{\{\tau=1\}}] + E[Y_{(\tau-2)^+ +1}]P(\tau > 1)$$

Where $(\tau - 2)^+ + 1 = \tau - 1$ for $\tau > 1$ and equals 1 for $\tau = 1$, so this is a stopping time for Y and hence

$$E[X_\tau] \leq E[X_1 1_{\{\tau=1\}}] + (\alpha_2 + E[(X_2 - \alpha_2)^+]) * (1 - P(\tau = 1))$$

So we set $\alpha_1 := \alpha_2 + E[(X_2 - \alpha_2)^+]$ as in the theorem, then we have:

$$E[X_\tau] \leq \alpha_1 + \int_{\{\tau=1\}} (x - \alpha_1) dP_{X_1}(x)$$

And as in the Lemma, this expectation is maximized when $\tau = 1 \Leftrightarrow X_1 \geq \alpha_1$, otherwise by the inductive hypothesis we have $\tau = \inf\{i > 1 : X_i \geq \alpha_i\}$ and these two facts together confirm that the optimal strategy for $X_1, ..., X_{k+1}$ is given by $\inf\{i : X_i \geq \alpha_i\}$ completing the induction. $\square$

Before moving on let us ground the discussion in another example similar to the first one but which implements the previous theorem:

Example 2: As in the first example we will take a hat with 10 slips of paper labeled with dollar amounts \$1,...,\$10 each of which is equally likely to be drawn and then replaced. Assume we will make n draws, then X_i is the uniform distribution on the set $\{1, ..., 10\}$ and so $E[X_i] = 5.5$ and thus we have $\alpha_n = 5.5$. Thus we can compute $\alpha_{n-1} = 5.5 + E[(X_{n-1} - 5.5)^+] = 5.5 + \frac{1}{10}\sum_{i=6}^{10}(i - 5.5) = 6.75$. Thus we will need a formula:

$$\alpha_{n-j-1} = \alpha_{n-j} + \frac{1}{10} \sum_{\alpha_{n-j} < i \leq 10} (i - \alpha_{n-j})$$

$$= \frac{11}{2} - \frac{\lfloor \alpha_{n-j} \rfloor (\lfloor \alpha_{n-j} \rfloor + 1)}{20} + \frac{\lfloor \alpha_{n-j} \rfloor \alpha_{n-j}}{10}$$

Notice that for $\alpha_{n-j} < 10$ we have $\alpha_{n-j-1} < \frac{11}{2} - \frac{9}{2} + 9\frac{\alpha_{n-j}}{10} < 10$. Furthermore, if fix $n = 10$ a short computer program yields the following ceiling values for $(\lceil \alpha_1 \rceil, ..., \lceil \alpha_{10} \rceil) = (10, 9, 9, 9, 9, 9, 8, 8, 7, 6)$, thus we should stop at the first variable if $X_1 \geq 10$ and at the second if $X_2 \geq 9$ and so forth. Finally one may also compute the approximate value of $\alpha_1 \cong 9.087$ which tells us the the expected return of this strategy is $E[X_T] = \alpha_1 + E[(X_1 - \alpha_1)^+] \cong 9.087 + \frac{1}{10}(10 - 9.087) \cong 9.18$. Recall that in the previous example the expected value of the chosen slip was $E[X_T] = 6 + E[(X_1 - 6)^+] = 7$ so in this new game we are able to achieve a much higher expected number on the slip since there is no cost for playing repeatedly.

So far we have been searching for strategies that maximize the expected value of the stopped sequence, but sometimes the goals are different. For example assume that instead of receiving the last value drawn (as in the previous theorem) the chosen value is recorded and then the game is continued to complete a total of n draws, when the game ends if the value recorded is the highest value seen we win, otherwise we lose. This new game is all or nothing, so the expected value of the stopped random variable is irrelevant. Instead all we care about is the probability of choosing the maximum of the variables, in other words we want to choose the stopping time T that maximizes:

$$P\left(X_T = \max_{1 \leq i \leq n} X_i\right) = P(X_T \geq X_1 \wedge \cdots \wedge X_T \geq X_n)$$

Again we first prove a lemma which gives the result for the case $n = 2$:

Lemma 1.0.10. *Let X_1, X_2 be independent random variables with $E[X_1], E[X_2] < \infty$ then the optimal stopping time for choosing the maximum of these two is given by*

$$T = 1 \Leftrightarrow P(X_1 \geq X_2) > \frac{1}{2}(1 + P(X_1 = X_2))$$

in the sense that for any stopping time τ we have $P(X_\tau = \max\{X_1, X_2\}) \leq P(X_T = \max\{X_1, X_2\})$.

Proof. Note that we have:

$$P(X_\tau = \max\{X_1, X_2\}) = P(1_{\{\tau=1\}}X_1 \geq X_2 + 1_{\{\tau\neq1\}}X_2 \geq X_1)$$

$$= E[1_{\{\tau=1\}}1_{\{X_1\geq X_2\}} + 1_{\{X_2\geq X_1\}} - 1_{\{\tau=1\}}1_{\{X_2\geq X_1\}}]$$

Note that since X_1 and X_2 are independent the product measure on these two variables is simply the product of the two measures, thus:

$$= P(X_2 \geq X_1) + \int_{\tau(x_1)=1}\int(2 * 1_{\{x_1\geq x_2\}} - 1_{\{x_1=x_2\}} - 1)dP_{X_2}(x_2)dP_{X_1}(x_1)$$

$$= P(X_2 \geq X_1) + \int_{\tau(x_1)=1}(2P(x_1 \geq X_2) - P(x_1 = X_2) - 1)dP_{X_1}(x_1)$$

Thus to maximize this integral we want to take $\tau(x) = 1$ if and only if $P(x \geq X_2) \geq \frac{1}{2}(1 + P(x = X_2))$. □

Thus, noting that when $X_1 = X_2$ the integrand is zero, we can compute the probability of success under the optimal strategy as:

$$P(X_T = \max\{X_1, X_2\}) = P(X_2 \geq X_1) + E[(2P(x \geq X_2) - 1)^+]\Big|_{x=X_1}$$

Where the probability is computed with respect to the random variable X_2 and x is fixed and then the expected value is computed with respect to the random variable X_1.

BIBLIOGRAPHY

[D] RICK DURRET, *Probability: Theory and Examples*, Brooks/Cole, Belmont, CA, 2005.