Chapter 6. MAXIMIZING THE RATE OF RETURN.

In stopping rule problems that are repeated in time, it is often appropriate to maximize the average return per unit of time. This leads to the problem of choosing a stopping rule N to maximize the ratio $\mathrm{E}Y_N/\mathrm{E}N$. The reason we wish to maximize this ratio rather than the true expected average per stage, $\mathrm{E}(Y_N/N)$, is that if the problem is repeated independently n times with a fixed stopping rule leading to i.i.d. stopping times $N_1,\ldots,N_n$ and i.i.d. returns $Y_{N_1},\ldots,Y_{N_n}$, the total return is $Y_{N_1}+\cdots+Y_{N_n}$ and the total time is $N_1+\cdots+N_n$, so that the average return per unit time is the ratio $(Y_{N_1}+\cdots+Y_{N_n})/(N_1+\cdots+N_n)$. If both sides of this ratio are divided by n and if the corresponding expectations exist, then this ratio converges to $\mathrm{E}Y_N/\mathrm{E}N$ by the law of large numbers. We call this ratio the rate of return. We wish to maximize the rate of return.

In the first section of this chapter, we describe a method of solving the problem of maximizing the rate of return by solving a sequence of related stopping rule problems as developed in the earlier chapters. There are a number of applications that are treated in subsequent sections and in the exercises. In Section 6.2, the main ideas are illustrated using the house-selling problem. In Section 6.3, application is made to problems where the payoff is a sum of discounted returns. This provides a background for the treatment of bandit problems in Chapter 7. In Section 6.4, a simple maintenance model is considered to illustrate the general method of computation. Finally in Section 6.5, a simple inventory model is treated.

§**6.1 Relation to Stopping Rule Problems.** We set up the problem more generally by allowing different stages to take different amounts of time. There are observations $X_1, X_2, \ldots$ as before, but now there are two sequences of payoffs, $Y_1, Y_2, \ldots$ and $T_1, T_2, \ldots$ with both Y_n and T_n assumed to be $\mathcal{F}_n$-measurable, where $\mathcal{F}_n$ is the sigma-field generated by $X_1, \ldots, X_n$. In this formulation, Y_n represents the return for stopping at stage n and T_n represents the total time spent to reach stage n. Throughout this chapter, we assume that the T_n are positive and nondecreasing almost surely,

(1) $$0 < T_1 \leq T_2 \leq \ldots \quad \text{a.s.}$$

We restrict attention to stopping rules that take at least one observation and note that $\mathrm{E}(T_N) \geq \mathrm{E}(T_1) > 0$ for every stopping rule $N \geq 1$. Thus, in forming the ratio $\mathrm{E}Y_N/\mathrm{E}T_N$, we avoid the problem of dealing with $0/0$. To avoid the troublesome $\pm\infty/+\infty$, we

restrict attention to stopping rules such that $\mathrm{E}T_N < \infty$. Thus, we let $\mathcal{C}$ denote the class of stopping rules,

(2) $$\mathcal{C} = \{N : N \geq 1, \mathrm{E}T_N < \infty\}$$

and we seek a stopping rule $N \in \mathcal{C}$ to maximize the rate of return, $\mathrm{E}Y_N / \mathrm{E}T_N$.

Without entering into the question of the existence of a stopping rule that attains a finite supremum of the above ratio, we can relate the solution of the problem of maximizing the rate of return to the solution of an ordinary stopping rule problem with return $Y_n - \lambda T_n$ for some λ.

Theorem 1. *(a) If for some λ, $\sup_{N \in \mathcal{C}} \mathrm{E}(Y_N - \lambda T_N) = 0$, then $\sup_{N \in \mathcal{C}} \mathrm{E}(Y_N)/\mathrm{E}(T_N) = \lambda$. Moreover, if $\sup_{N \in \mathcal{C}} \mathrm{E}(Y_N - \lambda T_N) = 0$ is attained at $N^* \in \mathcal{C}$, then N^* is optimal for maximizing $\sup_{N \in \mathcal{C}} \mathrm{E}(Y_N)/\mathrm{E}(T_N)$.*
(b) Conversely, if $\sup_{N \in \mathcal{C}} \mathrm{E}(Y_N)/\mathrm{E}(T_N) = \lambda$ and if the supremum is attained at $N^* \in \mathcal{C}$, then $\sup_{N \in \mathcal{C}} \mathrm{E}(Y_N - \lambda T_N) = 0$ and the supremum is attained at N^*.

Proof. If $\sup_{N \in \mathcal{C}} \mathrm{E}(Y_N - \lambda T_N) = 0$, then for all stopping rules $N \in \mathcal{C}$, $\mathrm{E}(Y_N - \lambda T_N) \leq 0$ so that $\mathrm{E}(Y_N)/\mathrm{E}(T_N) \leq \lambda$. If, for some $\epsilon \geq 0$, the rule $N^* \in \mathcal{C}$ is ϵ-optimal, so that $\mathrm{E}(Y_{N^*} - \lambda T_{N^*}) \geq -\epsilon$, then

$$\mathrm{E}(Y_{N^*})/\mathrm{E}(T_{N^*}) \geq \lambda - \epsilon/\mathrm{E}(T_{N^*}) \geq \lambda - \epsilon/\mathrm{E}(T_1),$$

so that N^* is $(\epsilon/\mathrm{E}(T_1))$-optimal for maximizing $\mathrm{E}(Y_N)/\mathrm{E}(T_N)$.

Conversely, suppose $\sup_{N \in \mathcal{C}} \mathrm{E}(Y_N)/\mathrm{E}(T_N) = \lambda$, and suppose the supremum is attained at $N^* \in \mathcal{C}$. Then, $\mathrm{E}Y_{N^*} - \lambda \mathrm{E}T_{N^*} = 0$ and for all stopping rules $N \in \mathcal{C}$, $\mathrm{E}Y_N - \lambda \mathrm{E}T_N \leq 0$. ∎

The optimal rate of return, λ, may also be considered as the "shadow" cost of time measured in the same units as the payoffs. This is because, when λ is the optimal rate of return, we search for the stopping rule that maximizes $\mathrm{E}(Y_N - \lambda T_N)$. It is as if we are being charged λ for each time unit. This is the mathematical analog of the aphorism, "Time is money".

Sometimes an extra argument may be provided to show that the limiting average payoff cannot be improved using rules for which $\mathrm{E}T_N = \infty$. (See §6.2.)

COMPUTATION. Many of the good applications require heavy computation to reach the solution and so we mention a fairly effective method suggested by G. Michaelides. We use part (a) of Theorem 1 to approximate the solution to the problem of computing the optimal rate of return. To use this theorem, we first solve the ordinary stopping rule problem for stopping $Y_n - \lambda T_n$ with arbitrary λ, and find the value. Ordinarily, this value will be a decreasing function of λ going from $+\infty$ at $\lambda = -\infty$ to $-\infty$ at $\lambda = \infty$. We then search for that λ that makes the value equal to zero.

To be more specific, let us make the assumption that for each λ there exists a rule $N(\lambda) \in \mathcal{C}$ that maximizes $\mathrm{E}(Y_N - \lambda T_N)$, and let $V(\lambda)$ denote the optimal return,

$$V(\lambda) = \sup_{N \in \mathcal{C}} [\mathrm{E}(Y_N) - \lambda \mathrm{E}T_N] = \mathrm{E}(Y_{N(\lambda)}) - \lambda \mathrm{E}(T_{N(\lambda)}).$$

Lemma 1. $V(\lambda)$ *is decreasing and convex.*

Proof. Let $\lambda_1 < \lambda_2$. Then

$$V(\lambda_2) = \mathrm{E}Y_{N(\lambda_2)} - \lambda_2 \mathrm{E}T_{N(\lambda_2)} < \mathrm{E}Y_{N(\lambda_2)} - \lambda_1 \mathrm{E}T_{N(\lambda_2)} \leq \mathrm{E}Y_{N(\lambda_1)} - \lambda_1 \mathrm{E}T_{N(\lambda_1)} = V(\lambda_1),$$

so $V(\lambda)$ is decreasing in λ. To show convexity, let $0 < \theta < 1$, fix λ_1 and λ_2 in $\mathcal{C}$, and let $\lambda = \theta\lambda_1 + (1-\theta)\lambda_2$. Then,

$$
\begin{aligned}
V(\lambda) &= \mathrm{E}Y_{N(\lambda)} - (\theta\lambda_1 + (1-\theta)\lambda_2)\mathrm{E}T_{N(\lambda)} \\
&= \theta(\mathrm{E}(Y_N) - \lambda_1 \mathrm{E}T_N) + (1-\theta)(\mathrm{E}(Y_N) - \lambda_2 \mathrm{E}T_N) \\
&\leq \theta V(\lambda_1) + (1-\theta)V(\lambda_2). \quad \blacksquare
\end{aligned}
$$

With this result, we may describe a simple iterative method of approximating the optimal rate of return and the optimal stopping rule. This method is a variation of Newton's method and so converges quadratically. Let λ_0 be an initial guess at the optimal value. At λ_0, the line $y = V(\lambda_0) - \mathrm{E}T_{N(\lambda_0)}(\lambda - \lambda_0)$ is a supporting hyperplane. This follows because $V(\lambda_0) - \mathrm{E}T_{N(\lambda_0)}(\lambda - \lambda_0) = \mathrm{E}Y_{N(\lambda_0)} - \lambda\mathrm{E}T_{N(\lambda_0)} \leq V(\lambda)$. Therefore, in Newton's method, $\lambda_{n+1} = \lambda_n - V(\lambda_n)/V'(\lambda_n)$, we may replace the derivative of $V(\lambda)$ at λ_n with $-\mathrm{E}T_{N(\lambda_n)}$. This gives the iteration for $n = 0, 1, 2, \ldots,$

(3)
$$\lambda_{n+1} = \lambda_n + \frac{V(\lambda_n)}{\mathrm{E}T_{N(\lambda_n)}} = \frac{\mathrm{E}Y_{N(\lambda_n)}}{\mathrm{E}T_{N(\lambda_n)}}.$$

For any initial value, λ_0, this sequence will converge quadratically to the optimal rate of return. It is interesting to note that the convergence is quadratic even if the derivative of $V(\lambda)$ does not exist everywhere. See §6.4 for an example.

§**6.2. House-Selling.** Consider the problem of selling a house without recall and with i.i.d. sequentially arriving offers of $X_1, X_2, \ldots$ dollars, constant cost $c \geq 0$ dollars per observation and with return $X_n - cn$ for stopping at n. When the house is sold, you may construct a new house to sell. Construction cost is $a \geq 0$ dollars and construction time is $b \geq 0$ time units, measured in units of time between offers. Thus your return for one cycle is $Y_n = X_n - a - cn$, and the time of a cycle is $T_n = n + b$. Note that in this formulation, the cost of living, c, is not assessed while the house is being built. We assume the cost of living while building is included in the cost a.

To solve the problem of maximizing the rate of return, $\mathrm{E}(Y_N)/\mathrm{E}(T_N)$, we solve the related stopping rule problem with return for stopping at n taken to be $Y_n - \lambda T_n =$

$X_n - a - cn - \lambda n - \lambda b$, and then choose λ so that the optimal return is zero. If we assume that the X_n have finite second moment, $E(X^2) < \infty$, this is the problem solved in §4.1 with return X_n replaced by $X_n - a - \lambda b$ and cost c replaced by $c + \lambda$. The solution found there requires $c + \lambda > 0$. The optimal rule is to accept the first offer $X_n \geq V^* + a + \lambda b$, where V^* satisfies

$$(4) \qquad\qquad E(X - a - \lambda b - V^*)^+ = c + \lambda.$$

The value of λ that gives $V^* = 0$ is then simply the solution of (4) with $V^* = 0$:

$$(5) \qquad\qquad E(X - a - \lambda b)^+ = c + \lambda.$$

If $b > 0$, the left side is a continuous decreasing function of λ from $E(X - a + bc)^+$ at $\lambda = -c$, to zero at $\lambda = \infty$, and the right side is continuous increasing from 0 at $\lambda = -c$, to ∞. If $b = 0$, the left side is constant. In either case, if $E(X - a + bc)^+ > 0$, there is a unique root, λ^*, of (5) such that $\lambda^* > -c$. The optimal rule is to accept the first offer of $a + \lambda^* b$ or greater:

$$(6) \qquad\qquad N^* = \min\{n \geq 1 : X_n \geq a + \lambda^* b\}.$$

This rule is optimal for maximizing the limiting average payoff out of all rules N such that $EN < \infty$, provided $E(X^+)^2 < \infty$ and $E(X - a + bc)^+ > 0$.

If $E(X - a + bc)^+ = 0$, then $\lambda^* = -c$ and $N^* \equiv \infty$. In other words, if $P(X > a - bc) = 0$, we never sell the house and our expected rate of return is $-c$. This makes sense since stopping can only make the rate of return less than $-c$, but since we have not defined a limiting average payoff for continuing forever, we make the assumption that $E(X - a + bc)^+ > 0$.

If $b = 0$, then (5) has a simple solution. The optimal rule for maximizing the rate of return is to *accept the first offer greater than the construction cost*, and the optimal rate of return becomes

$$\lambda^* = E(X - a)^+ - c.$$

If the offers are a.s. greater than a, this means that we accept the first offer that comes in, so that N is identically equal to 1.

Can we do better with rules N such that $ET_N = \infty$? From §4.1, it follows that N^* is optimal for maximizing $E(Y_N - \lambda^* T_N)$ out of all stopping rules provided we define the payoff for not stopping to be $-\infty$. We have not defined Y_∞ or T_∞ for this problem but we can extend the optimality of the rule N^* for maximizing the limiting average payoff to the class of rules N such that $P(N < \infty) = 1$ as follows. Let N be such a rule. As in the first paragraph of this chapter, we consider the problem repeated independently n times using the same stopping rule each time. Let the i.i.d. stopping times be denoted by $N_1, \ldots, N_n$, the corresponding i.i.d. returns by $Y_{N_1}, \ldots, Y_{N_n}$ and the corresponding i.i.d. reward times by $T_{N_1}, \ldots, T_{N_n}$. From §4.1, it follows that for any rule N with $P(N < \infty) = 1$, we have $E(Y_N - \lambda^* T_N) \leq 0$, possibly $-\infty$, so that from the strong law of large numbers,

$$\frac{1}{n} \sum_{i=1}^{n} (Y_{N_i} - \lambda^* T_{N_i}) \xrightarrow{a.s.} E(Y_N - \lambda^* T_N) \leq 0.$$

Also from the strong law of large numbers,

$$\frac{1}{n}\sum_{i=1}^{n} T_{N_i} \xrightarrow{a.s.} \mathrm{E}T_N > \mathrm{E}T_1 > 0,$$

so that

(7)
$$\limsup_{n\to\infty} \frac{\sum_{i=1}^{n} Y_{N_i}}{\sum_{i=1}^{n} T_{N_i}} - \lambda^* \le 0.$$

From the Fatou-Lebesgue Lemma, the expected value of the $\limsup$ of the average return is also nonpositive. This shows that N^* achieves the optimal rate of return out of all stopping rules for which $\mathrm{P}(N < \infty) = 1$, provided $\mathrm{E}(X^+)^2 < \infty$ and $\mathrm{E}(X - a + bc)^+ > 0$.

It is interesting to note that we can now make use of the observation of Robbins (1970) to weaken the condition $\mathrm{E}(X^+)^2 < \infty$ to requiring only $\mathrm{E}X^+ < \infty$. Under this weaker condition, Robbins shows that the rule N^* is optimal for stopping $Y_n - \lambda T_n$ out of all rules N such that $\mathrm{E}(Y_N - \lambda T_N)^- > -\infty$. But one can show that when $\mathrm{E}X^+ < \infty$ and $\mathrm{E}(Y_N - \lambda^* T_N)^- = -\infty$, we still have $(1/n)\sum_1^n (Y_{N_i} - \lambda^* T_{N_i}) \xrightarrow{a.s.} -\infty$ even though $\mathrm{E}(Y_N - \lambda T_N)^+ = \infty$. One may conclude that if $\mathrm{E}X^+ < \infty$, then N^* is optimal for maximizing the rate of return in the sense that if N is any stopping rule with $\mathrm{P}(N < \infty) = 1$ and $N_1, N_2, \ldots$ are i.i.d. with the distribution of N, then (7) holds, and equality is achieved if $N = N^*$.

§**6.3. Sum of Discounted Returns.** Let $X_1, X_2, \ldots$ represent your returns for working on days $1, 2, \ldots$. It is assumed that the X_j have some known joint distribution. For example, it might be assumed that $X_1, X_2, \ldots$ are daily returns from some mining operation or from studying some new mathematical problem. It may be that the returns indicate that the mine or problem is not likely to be very profitable, and so you should switch to a different mine or problem. The future is discounted by $0 < \beta < 1$ so that your total return for working n days has present value $Y_n = \sum_1^n \beta^{j-1} X_j$. In considering time, we should also discount, so the total time used earning this reward has present value $T_n = \sum_1^n \beta^{j-1}$. The problem of maximizing the rate of return is the problem of finding a stopping rule N to achieve the supremum in

(8)
$$V^* = \sup_{N \ge 1} \frac{\mathrm{E}\sum_1^N \beta^{j-1} X_j}{\mathrm{E}\sum_1^N \beta^{j-1}}.$$

We assume that the expectations of the X_j exist and are uniformly bounded above, $\sup_n \mathrm{E}X_n < \infty$. It may be noted that in this problem, we may allow N to assume the value $+\infty$ with positive probability; both sums in (8) will still be finite.

The problem given by (8) can be justified by a method similar to that used in the first paragraph of this chapter. We assume the original problem can be repeated independently as many times as desired. The kth repetition yields the sequence $X_{k1}, X_{k2}, \ldots$, each sequence being an independent sample from the original joint distribution of $X_1, X_2, \ldots$.

At any time n, after observing $X_{11}, \ldots, X_{1n}$ and earning $\sum_1^n \beta^{j-1} X_{1j}$, you may ask to start the problem over using the second sequence, but these returns will be discounted by an extra β^n because they start at time $n+1$. Similarly for any k, while observing sequence k, you may call for a restart and begin to observe the sequence $k+1$, etc. This is called the **restart problem** in Katehakis and Vienott (1987). Suppose the same stopping rule N is used in each restarted problem, yielding i.i.d. random variables $N_1, N_2, \ldots$. Then the total discounted return is

$$V = \sum_1^{N_1} \beta^{j-1} X_{1j} + \beta^{N_1} \sum_1^{N_2} \beta^{j-1} X_{2j} + \beta^{N_1+N_2} \sum_1^{N_3} \beta^{j-1} X_{3j} + \cdots.$$

Its expected return is

$$\mathrm{E}V = \mathrm{E} \sum_1^{N_1} \beta^{j-1} X_{1j} + \mathrm{E}\beta^{N_1} \mathrm{E} \left[\sum_1^{N_2} \beta^{j-1} X_{2j} + \beta^{N_2} \sum_1^{N_3} \beta^{j-1} X_{3j} + \cdots \right]$$

$$= \mathrm{E} \sum_1^{N_1} \beta^{j-1} X_{1j} + \mathrm{E}\beta^{N_1} \mathrm{E}V.$$

Solving for $\mathrm{E}V$, we find

(9)
$$\mathrm{E}V = \frac{\mathrm{E} \sum_1^N \beta^{j-1} X_j}{1 - \mathrm{E}\beta^N} = \frac{\mathrm{E} \sum_1^N \beta^{j-1} X_j}{(1-\beta)\mathrm{E} \sum_1^N \beta^{j-1}}.$$

Thus, the optimal rate of return given in (8) is equal to $1 - \beta$ times the optimal value of the restart problem.

To take an example of the computation of (8), assume that $X_1, X_2, \ldots$ are i.i.d. given a parameter $\theta > 0$ with distribution

$$\mathrm{P}(X = 0|\theta) = 1/2$$
$$\mathrm{P}(X = \theta|\theta) = 1/2$$

for all $\theta > 0$. We assume that the prior distributin of θ on $(0, \infty)$ is such that $\mathrm{E}\theta < \infty$. To find the supremum in (8), we first solve the associated stopping rule problem of finding a stopping rule to maximize $\mathrm{E}Y_N - \lambda \mathrm{E}T_N = \mathrm{E}(\sum_1^N \beta^{j-1}(X_j - \lambda))$. Let $V^* = V^*(\lambda)$ denote this maximum value. We must take at least one observation. With probability $1/2$, $X_1 = 0$, we lose λ and gain no information. In this case, the future looks as it did at the initial stage (except that it is now discounted by β), so we would continue if $V^* > 0$ and stop otherwise. With probability $1/2$, $X_1 = \theta$, we receive $\theta - \lambda$ and we would have complete information. In this case, if $\theta/2 > \lambda$, we would continue forever and expect to receive $\sum_2^\infty \beta^{j-1}(\theta/2 - \lambda) = (\beta/(1-\beta))(\theta/2 - \lambda)$, while if $\theta/2 \le \lambda$, we would stop now and receive nothing further. Combining this, we arrive at the following equation for V^*.

$$V^* = (1/2)(-\lambda + \beta \max(0, V^*)) + (1/2)(\mathrm{E}(\theta - \lambda) + \mathrm{E}((\theta/2 - \lambda)^+)(\beta/(1-\beta))).$$

Therefore,

$$V^* = [-2\lambda + (\beta/(1-\beta))\mathrm{E}(\theta/2 - \lambda)^+ + \mathrm{E}\theta]/(2-\beta) \quad \text{if} \quad V^* > 0$$
$$= [-2\lambda + (\beta/(1-\beta))\mathrm{E}(\theta/2 - \lambda)^+ + \mathrm{E}\theta]/2 \qquad \text{if} \quad V^* \le 0.$$

To find the maximal rate of return, we choose λ so that $V^* = 0$. This gives the optimal rate of return, λ^*, as the root of the equation,

$$2\lambda = \mathrm{E}(\theta) + (\beta/(1-\beta))\mathrm{E}(\theta/2 - \lambda)^+.$$

The left side is increasing from $-\infty$ to $+\infty$, and the right side is nonincreasing, so there is a unique root. The optimal rule is: take one observation; if $X_1 > 2\lambda^*$, then continue forever; otherwise stop.

For a specific example, suppose θ has a uniform distribution on the interval $(0, 1)$. Then $\mathrm{E}(\theta) = 1/2$ and $\mathrm{E}(\theta/2 - \lambda)^+ = \lambda^2 - \lambda + 1/4$, so that λ^* is the root of $2\lambda = 1/2 + (\beta/(1-\beta))(\lambda^2 - \lambda + 1/4)$ between $1/4$ and $1/2$, namely

$$\lambda^* = [2 - \beta - \sqrt{2(2-\beta)(1-\beta)}]/(2\beta).$$

§**6.4. Maintenance.** A machine used in production of some item will produce a random number of items each day. As time progresses, the performance of the machine deteriorates and it will eventually need to be overhauled entailing a cost for the service and a loss of time for use of the machine. Suppose that if the machine has just been overhauled it produces X_1 items where X_1 has a Poisson distribution with mean μ. Suppose also that deterioration is exponential in time so that the number of items produced on the nth day after overhaul, X_n, is Poisson with mean μq^{n-1}, where q is a given number, $0 < q < 1$. Let $c > 0$ denote the cost of the overhaul and suppose that the service takes one day. The problem of finding a time at which to stop production for overhaul in order to maximize the return per day is then the problem of finding a stopping rule $N \ge 1$ to maximize $\mathrm{E}(S_N - c)/\mathrm{E}(N+1)$, where $S_n = X_1 + \ldots + X_n$.

To solve this problem, we first consider the problem of finding a stopping rule to maximize $\mathrm{E}(Y_N - \lambda T_N)$, where $Y_n = S_n - c$ and $T_n = n+1$. Let us see if the 1-sla is optimal. If we stop at stage n, we gain $S_n - c - \lambda(n+1)$; if we continue one stage and stop, we expect to gain

$$S_n + \mathrm{E}X_{n+1} - c - \lambda(n+2) = S_n + \mu q^n - c - \lambda(n+2).$$

Therefore, the 1-sla is

(10) $\qquad N_1 = \min\{n > 0 : \lambda \ge \mu q^n\} = \min\{n > 0 : n \ge \log(\mu/\lambda)/\log(1/q)\}.$

The problem is monotone and the 1-sla is an optimal stopping rule of a fixed sample size, $N_1 = m$, where $m = 1$ if $\lambda \ge \mu$ and $m = \lceil \log(\mu/\lambda)/\log(1/q) \rceil$ if $\lambda < \mu$. Its expected return is simply

(11)
$$V(\lambda) = \mathrm{E}(S_m - c - \lambda(m+1)) = \mu(1 + q + \ldots + q^{m-1}) - c - \lambda(m+1)$$
$$= \mu(1 - q^m)/(1-q) - c - \lambda(m+1).$$

We set this expression to zero and solve for λ, which looks easy until we remember that m depends on λ. We illustrate the general method of solving for λ suggested in §6.1 on a simple numerical example.

Suppose $\mu = 3$, $q = .5$, and $c = 1$. As an initial guess at the optimal rate of return, let us take $\lambda_0 = 1.5$. The iteration involved in (3) requires that we iterate the following two equations in order:

$$m = \left\lceil \frac{\log(\mu/\lambda)}{\log(1/q)} \right\rceil \qquad \text{and} \qquad \lambda = \frac{\mu\frac{1-q^m}{1-q} - c}{m+1}.$$

On the first iteration, we find $m = \lceil 6.579 \rceil = 7$ and $\lambda_1 = 1.813$. Applying the iteration again, we find $m = \lceil 4.686 \rceil = 5$ and $\lambda_2 = 1.881$. On the third iteration we find $m = \lceil 4.431 \rceil = 5$, and we must therefore have $\lambda_3 = \lambda_2$. The iteration has converged in a finite number of steps. We overhaul every sixth day ($m = 5$) and find as the average return per day, $\lambda^* = 1.881$.

In this problem, the iteration converges in a finite number of steps whatever be the values of μ, q and c, because the value function, $V(\lambda)$, is piecewise linear. It is only in very simple problems that this will be the case.

§**6.5. An Inventory Problem.** A warehouse can hold W items of a certain stock. Each day a random number of orders for the item are received. The items are sold up to the number of items in stock; orders not filled are lost. Each item sold yields a net profit of $C_1 > 0$ (selling price of the item minus its cost). The warehouse may be restocked to capacity at any time by paying a restocking fee of $C_2 > 0$. On day n, orders for X_n items are received, $n = 1, 2, \ldots$, where $X_1, X_2, \ldots$ are independent and all have the same distribution, $f(x) = P(X_n = x)$ for $x = 0, 1, 2, \ldots$. The problem is to find a restocking time N (a stopping rule) to maximize the rate of return,

$$(12) \qquad \qquad \qquad \mathrm{E}(\min(S_N, W)C_1 - C_2)/\mathrm{E}(N),$$

where $S_n = X_1 + \ldots + X_n$.

If C_2 were zero, we would restock every day ($N = 1$) and have a rate of return equal to $\mathrm{E}(\min(X_1, W))C_1$. Since $C_2 > 0$, it may be worthwhile to wait until the number of items on stock gets low before reordering.

To find the optimal restocking time, we first solve the stopping rule problem for maximizing the return

$$\begin{aligned} Y_n &= \min(S_n, W)C_1 - C_2 - n\lambda \\ &= (W - (W - S_n)^+)C_1 - C_2 - n\lambda. \end{aligned}$$

The one-stage look-ahead rule is

$$\begin{aligned} N_1 &= \min\{n \geq 1 | Y_n \geq \mathrm{E}(Y_{n+1}|\mathcal{F}_n)\} \\ (13) \qquad &= \min\{n \geq 1 | (W - S_n)^+ - \mathrm{E}((W - S_n - X_{n+1})^+|\mathcal{F}_n) \leq \lambda/C_1\} \\ &= \min\{n \geq 1 | W - S_n \leq z\}, \end{aligned}$$

where,

(14) $$z = \max\{u \mid u - \mathrm{E}(u - X)^+ \leq \lambda/C_1\}.$$

The 1-sla is monotone and the theorems of Chapter 5 show that it is optimal. Thus, the optimal restocking rule has the form: *restock the warehouse as soon as the inventory has z items or less.* The optimal value of z may be found from (14) when λ is chosen to make $\mathrm{E}(Y_{N_1}) = 0$. In this problem, it may be simpler to calculate the ratio (12) for all stopping rules $N(z)$ of the form $N(z) = \min\{n \geq 1 \mid W - S_n \leq z\}$ and find z that makes this ratio largest.

We carry out the computations when f is the geometric distribution,

$$f(x) = (1 - p)p^x \quad \text{for} \quad x = 0, 1, \ldots.$$

First, we compute the numerator of (12) using $N(z)$. The distribution of $S_{N(z)} - (W - z)$ is the same geometric distribution, f , so that

$$\mathrm{E}\min(S_{N(z)}, W) = W - z + \mathrm{E}\min(X, z)$$
$$= W - z + (1 - p^z)p/(1 - p).$$

Second, to compute the denominator of (12), note that the geometric random variable X_n may be considered as the number of heads in a sequence of tosses of a coin with probability p of heads tossed until the first tail occurs. Therefore, $N(z) - 1$ represents the number of tails observed before $W - z$ heads, and so $N(z) - 1$ has a negative binomial distribution with probability of success $1 - p$ and $W - z$ fixed failures. Hence,

$$\mathrm{E}(N(z)) = 1 + (W - z)(1 - p)/p.$$

Combining these two expectations and letting $\lambda(z)$ represent the ratio (1) when $N = N(z)$, we find,

$$\lambda(z) = [W - z + (1 - p^z)p/(1 - p) - C_2]/[1 + (W - z)(1 - p)/p],$$

where without loss of generality we have taken $C_1 = 1$, since the optimal rule depends only on the ratio C_2/C_1. After some wonderfully exciting and beautiful algebraic manipulations, we find that

$$\lambda(z - 1) > \lambda(z) \quad \text{if and only if} \quad C_2 > p^z(W - z).$$

Since the right side of this inequality is decreasing in z, we find that the maximum of $\lambda(z)$ occurs at $z = 0$ if $C_2 \geq p(W - 1)$ and at $z = n$ if

$$p^{n+1}(W - n - 1) \leq C_2 \leq p^n(W - n) \quad \text{for} \quad n = 1, \ldots, W - 1.$$

As a numerical example, suppose that $C_1 = 1$, $W = 10$, and $p = 2/3$. Then if $C_2 > 6$, we have $z = 0$; we wait until we run out completely before reordering. (If

$C_2 > 10$, we are operating at a loss.) If $3.555 < C_2 < 6$, we have $z = 1$; we reorder when there is at most one left. Similarly down to: if $0 < C_2 < .026$, we have $z = 9$; we reorder as soon as at least one item is sold.

§6.6 Exercises.

1. *Selling an asset.* You can buy an item at a constant cost $c > 0$ and sell it when you like. Bids for the item come in one each day, $X_1, X_2, \ldots$ i.i.d. $F(x)$ and although it does not cost anything to observe these, it takes d days to obtain a new item. The problem is to find a stopping rule to maximize $\mathrm{E}(X_N - c)/\mathrm{E}(N + d)$. Assume that $\mathrm{P}(X > c) > 0$, and that $\mathrm{E}(X^2) < \infty$. Find an optimal rule and the optimal rate of return. Specialize to the case where F is the uniform distribution on $(0, 1)$.

2. *Maintenance.* (Taylor (1975), Posner and Zuckerman (1986), and Aven and Gaardner (1987)) A machine accumulates observable damage at discrete times through a series of shocks. Shocks occur independently at times $t = 1, 2, \ldots$, with probability, q, $0 < q < 1$, independent of time. When a shock occurs, the machine will accrue a certain amount of damage, assumed to be exponentially distributed with mean μ. If X_n denotes the damage accrued at time n, it is thus assumed that the X_n are i.i.d. with distribution $\mathrm{P}(X_n > x) = q \exp\{-x/\mu\}$ for $x > 0$, and $\mathrm{P}(X_n = 0) = 1 - q$. The total damage accrued to the machine by time n is $S_n = \sum_1^n X_j$. The machine breaks down at time n if S_n exceeds a given number, $M > 0$. The time of breakdown is thus $T = \min\{n \geq 1 : S_n > M\}$. A machine overhaul costs an amount $C > 0$. If the machine breaks down, the machine must be overhauled and there is an additional cost of $K > 0$. The problem is to decide when to overhaul the machine. The cost of overhauling the machine at stage n is thus $Y_n = C + K\mathrm{I}(n = T)$. To enforce stopping at T, we may put $Y_n = \infty$ on $\{n > T\}$. We want to choose a stopping rule, N, to *minimize* the cost per unit time, $\mathrm{E}(Y_N)/\mathrm{E}(N)$. This reduces to seeking a stopping rule to minimize $\mathrm{E}(Y_N - \lambda N)$, for some $\lambda > 0$.
(a) Find the 1-sla for the latter problem and show it is optimal.
(b) Show how to solve the original problem. (Choose λ as the root of $\lambda[M - \mu \log(Kq/\lambda)] = Cq\mu$.)

3. *Foraging.* Consider an animal that forages for food in spacially separated patches of prey. He feeds at one patch for awhile and then moves on to another. The problem of when to move to a new patch in order to maximize the rate of energy intake is addressed in the papers of Allan Oaten (1977), and Richard Green (1984, 1987). As an example, take the fisherman who moves from waterhole to waterhole catching fish. Suppose that in each waterhole there are initially n fish, where n is known. Assume that each fish has an exponential catch time at rate 1, and captures are independent events. This problem is treated in Example 5 of §5.4. Suppose that the expected time it takes to move from one waterhole to another is a known constant, $\tau > 0$. The problem is to find a stopping rule N to maximize the rate of return, $\mathrm{E}(N)/\mathrm{E}(X_N + \tau)$, where X_j is the jth order statistic of a sample of size n from the exponential distribution. Find an optimal rule and the optimal rate of return. As a numerical example, take $n = 10$ and $\tau = 1$.

4. *Attaining a goal.* Let $X_1, X_2, \ldots$ be independent Bernoulli trials with probability $1/2$ of success, and let S_n denote the sum, $\sum_1^n X_j$. Your goal is to achieve $S_n = a$,

where a is a fixed positive integer. If you attain your goal you win $c_1 > 0$, but the cost is 1 per trial. You may give up at any time by paying an additional amount, c_2. The real problem, however, is to choose a stopping rule, N, to maximize the rate of return, $c_1 \mathrm{P}(N \geq a)/\mathrm{E}(N + c_2)$. Find the optimal rule and the optimal rate of return. (Refer to Exercise 4.8.)