CHAPTER 11

STOPPING PROBLEMS

Stopping problems are a simple but important class of learning problems. In this problem class, information arrives over time, and we have to choose whether to view the information or stop and make a decision. In this setting, "learning" means to continue to continue to receive information. We do not choose which information to observe, as we have done in the past, but we do have to decide whether to continue observing or stop and make a decision.

In this chapter we consider two classical stopping problems. The first is the sequential probability ratio test, where we have to quickly identify when a signal from some source is changing. The second is a classic problem known as the secretary problem, where we have to find the best out of a sequence of offers which arrive one at a time.

This chapter will illustrate learning using very different tools than we have seen up to now. Optimal stopping problems are both a very important problem class, as well as being a very special case which offers structure that we can exploit to derive optimal policies, rather than the well tuned approximations that have been the focus of the book until now.

Optimal Learning. By Warren B. Powell and Ilya O. Ryzhov

11.1 SEQUENTIAL PROBABILITY RATIO TEST

A fundamental problem arises when we need to decide as quickly as possible when something is changing. For example, we may think we are observing data being generated by the sequence

$$W^n = \bar{\mu}^0 + \epsilon^n \tag{11.1}$$

where we assume that ϵ is normally distributed with mean 0 and variance σ^2. But we are aware that the mean might change to $\bar{\mu}^1$, which means the observations would come from the model

$$W^n = \bar{\mu}^1 + \epsilon^n. \tag{11.2}$$

We would like to design a method that allows us to determine when the mean has changed as quickly as possible.

A more general statement of the problem would be that the observations W^n are coming from an initial model that we refer to as the null hypothesis H_0. We then want to determine when the data is coming from a different model that we refer to as the alternative hypothesis, H_1. We start with a prior probability $\rho_0 = \rho_0^0$ that H_0 is true. Similarly, ρ_1^0 is our prior probability that H_1 is true. After we observe W^1, we would update our prior using Bayes' rule to obtain

$$\begin{aligned}\rho_0^1 &= P(H_0|W^1) \\ &= \frac{P(W^1|H_0)P(H_0)}{P(W^1)},\end{aligned}$$

where $P(W^1) = \rho_0 P(W^1|H_0) + \rho_1^0 P(W^1|H_1)$. The quantity $\rho_1^1 = P(H_1|W^1)$ would be worked out similarly. For example, if our data is normally distributed, we would write

$$P(W^1 = w|H_0) = \frac{1}{\sqrt{2\pi}\sigma} \exp -\frac{1}{2}\left(\frac{(w - \bar{\mu}^0)}{\sigma}\right)^2.$$

Let $p_0(w^n) = P(W^n = w^n|H_0)$. After n observations, we can write the posterior probability as

$$\begin{aligned}\rho_0^n &= \frac{\rho_0 \Pi_{k=1}^n p_0(w^k)}{\rho_0 \Pi_{k=1}^n p_0(w^k) + \rho_1^0 \Pi_{k=1}^n p_1(w^k)} \\ &= \frac{\rho_0 \lambda^n(w^1, \ldots, w^n)}{\rho_0 + \rho_1^0 \lambda^n(w^1, \ldots, w^n)}\end{aligned}$$

where

$$\lambda^n(w^1, \ldots, w^n) = \Pi_{k=1}^n \frac{p_0(w^k)}{p_1(w^k)}.$$

We can write $S^n = (W^1, W^2, \ldots, W^n)$ as being the set of all experiments, or it could be a sufficient statistic (such as the mean and variance of the normal distribution) that captures everything we need to know from previous experiments. Later, we let $\lambda^n = \lambda^n(S^n) = \lambda^n(w^1, \ldots, w^n)$.

To solve our problem, we need to make two decisions. The first is whether to continue to observe W^n, or to stop and make a decision. If we decide to stop, we then have to choose between H_0 and H_1. We let

$$X^\pi(S^n) = \begin{cases} 1 & \text{if we decide to stop and make a decision,} \\ 0 & \text{if we continue observing.} \end{cases}$$

$$Y^\pi(S^n) = \begin{cases} 1 & \text{if we decide } H_1 \text{ is correct,} \\ 0 & \text{if we decide } H_0 \text{ is correct.} \end{cases}$$

We let π denote a policy consisting of the two functions (X^π, Y^π). Given a policy π, there are two mistakes we can make. The first is a false alarm, which means we stop and conclude that H_1 is true when in fact H_0 is true, and a miss, which means that H_1 is true, but we did not pick it up and we still think H_0 is true. We define the probability of these two events using

$$\begin{aligned} P_F^\pi &= \text{the probability we conclude } H_1 \text{ is true (the false alarm) given } H_0 \\ &\quad\ \text{when using policy } \pi \\ &= \mathbb{E}[Y^\pi(S^n)|H_0], \\ P_M^\pi &= \text{the probability we conclude } H_0 \text{ is true given } H_1 \text{ when using policy} \\ &\quad\ \pi \\ &= \mathbb{E}[1 - Y^\pi(S^n)|H_1]. \end{aligned}$$

Now let ρ_0 be the prior probability that H_0 is true. We can define the overall probability of an error using

$$P_e = (1 - \rho_0)P_F^\pi + \rho_0 P_M^\pi.$$

Of course, we can minimize this error by running many experiments. The number of experiments is given by

$$N^\pi = \min\{n | X^\pi(S^n) = 1\}.$$

N^π is a random variable that depends on our policy (the decision function X^π) and the observations $(W^1, W^2, \ldots, W^n)$. We can assign a "cost" c to each experiment, giving us a utility function

$$U^\pi(c) = P_e + c\mathbb{E}N^\pi.$$

Here, c is not a true cost in the sense of being measured in units of dollars per experiment. Rather, it is a scaling coefficient that allows us to combine the probability of being wrong with the number of experiments. Our challenge is to find a policy π that minimizes the utility function $U^\pi(c)$.

The problem can be solved (approximately) by exploiting some nice structural properties. We first write the conditional risk as

$$\begin{aligned} r_0^\pi &= P_F^\pi + c\mathbb{E}[N^\pi|H_0], \\ r_1^\pi &= P_M^\pi + c\mathbb{E}[N^\pi|H_1], \\ r^\pi &= \rho_0 r_0^\pi + (1 - \rho_0) r_1^\pi. \end{aligned}$$

We would like to find a policy π that minimizes the risk, given by

$$R^0(\rho_0) = \min_{\pi} r^{\pi}.$$

We start by observing that if $\rho_0 = 1$ (which means we are positive that H_1 is true), then we can stop and choose $Y = 1$ with no risk of a false positive (and $N = 0$), which means that the risk is $R^0(1) = 0$. The same reasoning tells us that $R^0(0) = 0$. It is also possible to show that $R^0(\rho)$ is concave in ρ.

Assume that we stop after running no experiments. If we choose $Y = 0$ then the risk is $R^0(\rho_0|Y = 0) = \rho_0$ (which is the same thing as saying that if I have to choose now with no information, my probability of being right is my original prior that H_0 is true). Similarly, if we choose $Y = 1$ then the risk is $R^0(\rho_0|Y = 1) = 1 - \rho_0$. Or, we could choose to make a single experiment (sort of like choosing curtain number 3). In this case, we want to choose the experimental policy π that solves

$$R^1(\rho_0^0) = \min_{\{\pi, N>0\}} \rho_0^0 r_0^{\pi} + (1 - \rho_0^0) r_1^{\pi}.$$

Here, we are solving the same problem as we were at time 0, but we are now forcing ourselves to take at least one experiment. So, our policy will be to run a single experiment (which means $N > 0$), and then we have to do the best we can choosing between $Y = 0$, $Y = 1$ or running yet another experiment.

So, we want the smallest of ρ_0^0 (corresponding to stopping and choosing H_0), $1 - \rho_0^0$ (corresponding to stopping and choosing H_1) and $R^1(\rho^0)$ (which means take another observation and repeat the process). The problem is depicted in Figure 11.1. Here, we plot the lines ρ_0^0 and $1 - \rho_0^0$, and the concave function $R^1(\rho^0)$, all as a function of ρ_0^0. If c is large enough, it is possible that the midpoint of $R^1(\rho^0)$ is greater than .5, in which case the best choice is to stop right away ($N = 0$) and choose between H_0 and H_1. Now assume that the maximum of $R^1(\rho^0)$ is less than .5. In this case, run one experiment and compute the posterior ρ_0^1. We then divide the horizontal axis into three regions: $\rho_0 < \rho^L$, $\rho^L \le \rho_0 \le \rho^U$, and $\rho_0 > \rho^U$. If $\rho_0 < \rho^L$, then the best choice is to choose $Y = 0$. If $\rho_0 > \rho^U$ then we stop and choose $Y = 1$. If $\rho^L \le \rho_0 \le \rho^U$, then we run another experiment and repeat the process. After each experiment, we face the same problem, where the only change is that we have a new prior.

This seems like a pretty simple rule. The only challenge is finding ρ^L and ρ^U. We begin by computing the likelihood ratio

$$L^n(S^n) = \Pi_{k=1}^n \frac{p_1(W^k)}{p_0(W^k)} = \Pi_{k=1}^n L(W^k)$$

where $L(W^k) = \frac{p_1(W^k)}{p_0(W^k)}$ is the likelihood ratio for a single observation. We next use Bayes' rule to compute the posterior $\rho_0^{n+1}(W^n)$ as follows

$$\begin{aligned}
\rho_1^{n+1}(S^n) &= \frac{p_1(S^n)(1 - \rho_0^n)}{\rho_0^n p_0(S^n) + (1 - \rho_0^n) p_1(S^n)} \\
&= \frac{L^n(S^n)}{L^n(S^n) + \rho_0^n/(1 - \rho_0^n)} \\
&= f(L^n(S^n), \rho_0^n/(1 - \rho_0^n)),
\end{aligned}$$

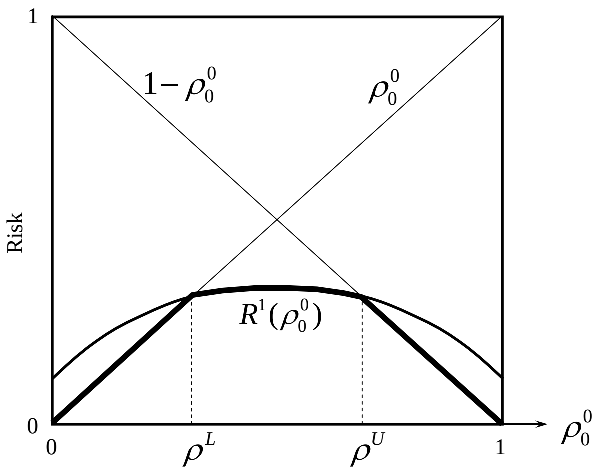

Figure 11.1 The expected risk as a function of the prior probability that H_0 is true.

where $f(\ell, \beta) = \ell/(\ell + \beta)$. For $\beta \geq 0$, $f(\ell, \beta)$ is strictly increasing for $\ell \geq 0$. This means that if $0 < \rho_0^n < 1$, $\rho_0^{n+1}(S^n)$ is strictly increasing with $L^n(S^n)$. This means that determining if $\rho_0^{n+1}(S^n) \leq \rho^L$ or $\rho_0^{n+1}(S^n) \geq \rho^U$ is the same as testing if $L^n(S^n) \leq A$ or $L^n(S^n) \geq B$, where A and B satisfy

$$\begin{aligned} f\left(A, \rho_0^n/(1-\rho_0^n)\right) &= \rho^L, \\ f\left(B, \rho_0^n/(1-\rho_0^n)\right) &= \rho^U. \end{aligned}$$

We can solve for A and B, which gives us

$$\begin{aligned} A &= \frac{\rho_0^n \rho^L}{(1-\rho_0^n)(1-\rho^L)}, \\ B &= \frac{\rho_0^n \rho^U}{(1-\rho_0^n)(1-\rho^U)}. \end{aligned}$$

This means that we can write our policy in the form

$$L^n(S^n) = \begin{cases} \geq B & \text{stop and choose } Y^n = 1 \\ \leq A & \text{stop and choose } Y^n = 0 \\ \text{otherwise} & \text{take an additional observation.} \end{cases}$$

Hence, it is easy to see why this rule is known as the *sequential probability ratio test* ($SPRT$). The $SPRT$ is controlled by the parameters A and B, and hence we refer to the rule as $SPRT(A, B)$.

Finding A and B exactly is difficult, but we can find good estimates using *Wald's approximation* which gives us

$$\begin{aligned} P_F^\pi &\approx \frac{(1-A)}{B-A}, \\ P_M^\pi &\approx \frac{A(B-1)}{B-A}. \end{aligned}$$

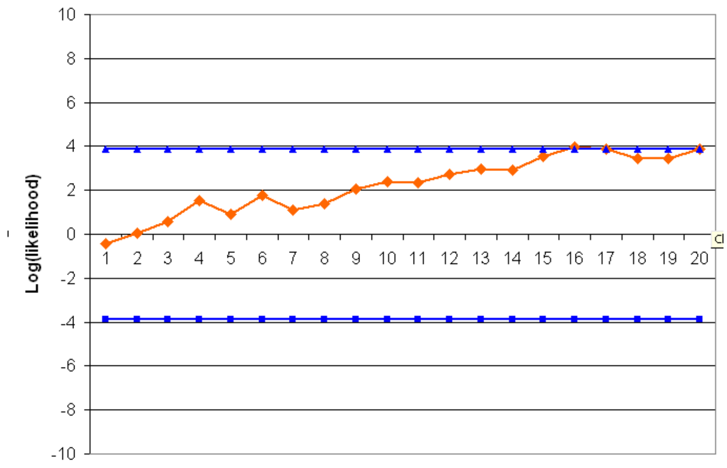

Figure 11.2 Sample path of ln of SPRT likelihood function

We then choose the acceptable probability of a false positive, P_F^π, and the probability of a miss, P_M^π, and then solve for A and B, giving us

$$\begin{aligned} A &= \frac{P_M^\pi}{1 - P_F^\pi}, \\ B &= \frac{1 - P_M^\pi}{P_F^\pi}. \end{aligned}$$

This rule is (approximately) optimal in the sense that it meets these goals for the probability of a false positive and the probability of missing a change with the fewest number of experiments. We note that it is customary to work in terms of the logarithm of L.

Figure 11.2 plots the log of the likelihood for a set of sample observations. After 16 observations, we conclude that H_1 is true.

11.2 THE SECRETARY PROBLEM

The so-called secretary problem is one of the first (formally defined) learning problems. The motivation of the problem is determining when to hire a candidate for a job (presumably a secretarial position), but it can also be applied to reviewing a series of offers for an asset (such as selling your house or car). The problem involves the tradeoff between observing candidates (which allows us to collect information) and making a decision (exploiting the information). As with the sequential probability ratio test in the previous section, our decision is when to stop reviewing new offers and accept the most recent offer. In contrast with our work on Bayesian models, at the heart of the secretary problem is that we assume that we know absolutely nothing about the distribution of offers.

11.2.1 Setup

Assume that we have N candidates for a secretarial position (you can also think of these as offers to purchase an asset). Each candidate is interviewed in sequence and assigned a score that allows us to compare him or her to other candidates (if we are trying to sell an asset, these scores are the offers to purchase the asset). While it may be reasonable to try to maximize the expected score that we would receive, in this case, we want to maximize the probability of accepting the highest score out of all that have been, or might be, offered. We need to keep in mind that if we stop at candidate n, then we will not interview candidates $n+1, \ldots, N$. Also, we only have the option of accepting the last candidate or interviewing the next one. Once we have turned down a candidate, we cannot return to that candidate at a later time.

Let

$$
\begin{aligned}
W^n &= \text{Score of the } n\text{th candidate.} \\
S^n &= \begin{cases} 1 & \text{If the score of the } n\text{th candidate is the best so far} \\ 0 & \text{If the score of the } n\text{th candidate is not the best so far} \\ \Delta & \text{If we have stopped already} \end{cases} \\
\mathcal{S} &= \text{State space, given by } (0, 1, \Delta)\text{, where the states 0 and 1 mean that} \\
&\quad\ \text{we are still searching, and } \Delta \text{ means we have stopped the process.} \\
\mathcal{X} &= \{0 \text{ (continue)}, 1 \text{ (stop)}\}, \text{ where "stop" means that we hire the last} \\
&\quad\ \text{candidate interviewed.}
\end{aligned}
$$

Because the decision function uses the most recent piece of information, we define our history as

$$h^n = \{W^1, \ldots, W^n\}.$$

To describe the system dynamics, it is useful to define an indicator function

$$I^n(h^n) = \begin{cases} 1 & \text{if } W^n = \max_{1 \le m \le n}\{W^m\} \\ 0 & \text{otherwise.} \end{cases}$$

which tells us if the last observation is the best. Our transition function can now be given by

$$S^{n+1} = \begin{cases} I^n(h^n) & \text{if } x^n = 0 \text{ and } S^n \neq \Delta \\ \Delta & \text{if } x^n = 1 \text{ or } S^n = \Delta. \end{cases}$$

To compute the one-step transition matrix, we observe that the event the $(n+1)$st applicant is the best has nothing to do with whether the nth was the best. As a result, we can write the conditional probability that $I^{n+1}(h^{n+1}) = 1$ using

$$\mathbb{P}[I^{n+1}(h^{n+1}) = 1 | I^n(h^n)] = \mathbb{P}[I^{n+1}(h^{n+1}) = 1].$$

This simplifies the problem of finding the one-step transition probabilities. By definition we have

$$\mathbb{P}[S^{n+1} = 1 | S^n, x^n = 0] = \mathbb{P}[I^{n+1}(h^{n+1}) = 1].$$

$I^{n+1}(h^{n+1}) = 1$ if the $(n+1)$st candidate is the best out of the first $n+1$, which clearly occurs with probability $1/(n+1)$. So

$$\begin{aligned}\mathbb{P}(S^{n+1} = 1|S^n, x^n = 0) &= \frac{1}{n+1},\\ \mathbb{P}(S^{n+1} = 0|S^n, x^n = 0) &= \frac{n}{n+1}.\end{aligned}$$

Our goal is to maximize the probability of hiring the best candidate. So, if we do not hire the last candidate, then the probability that we hired the best candidate is zero. If we hire the nth candidate, and the nth candidate is the best so far, then our reward is the probability that this candidate is the best out of all N. This probability is simply the probability that the best candidate out of all N is one of the first n, which is n/N. So, the conditional reward function is

$$C^n(S^n, x^n|h^n) = \begin{cases} n/N & \text{if } S^n = 1 \text{ and } x^n = 1,\\ 0 & \text{otherwise.}\end{cases}$$

With this information, we can now set up the optimality equations

$$V^n(s^n) = \max_{x^n \in \mathcal{X}} \mathbb{E}\{C^n(s^n, x^n|h^n) + V^{n+1}(S^{n+1})|s^n\}.$$

11.2.2 Solution

The solution to the problem is quite elegant, but the technique is unique to this particular problem. Readers interested in the elegant answer but not the particular proof (which illustrates dynamic programming but otherwise does not generalize to other problem classes) can skip to the end of the section.

Let $V^n(s)$ be the probability of choosing the best candidate out of the entire population, given that we are in state s after interviewing the nth candidate. Recall that implicit in the definition of our value function is that we are behaving optimally from time period t onward. The terminal reward is

$$\begin{aligned}V^N(1) &= 1\\ V^N(0) &= 0\\ V^N(\Delta) &= 0\end{aligned}$$

Let

$$\begin{aligned}C^{stop,n} &= \left(C^n(1, \text{stop}) + V^{n+1}(\Delta)\right),\\ C^{continue,n} &= \left(C^n(1, \text{continue}) + \sum_{s' \in \{0,1\}} p(s'|s)V^{n+1}(s')\right).\end{aligned}$$

The optimality recursion for the problem is given by

$$V^n(1) = \max\left\{C^{stop,n}, C^{continue,n}\right\}.$$

Noting that:

$$\begin{aligned} C^n(1, \text{continue}) &= 0, \\ C^n(1, \text{stop}) &= \frac{n}{N}, \\ V^{n+1}(\Delta) &= 0, \\ p(s'|s) &= \begin{cases} 1/(n+1) & s' = 1 \\ n/(n+1) & s' = 0 \end{cases}. \end{aligned}$$

We get

$$V^n(1) \quad = \quad \max\left\{\frac{n}{N}, \frac{1}{n+1}V^{n+1}(1) + \frac{n}{n+1}V^{n+1}(0)\right\}. \tag{11.3}$$

Similarly, it is easy to show that

$$\begin{aligned} V^n(0) &= \max\left\{0, \frac{1}{n+1}V^{n+1}(1) + \frac{n}{n+1}V^{n+1}(0)\right\} \qquad (11.4) \\ &= \frac{1}{n+1}V^{n+1}(1) + \frac{n}{n+1}V^{n+1}(0). \end{aligned}$$

Comparing (11.4) and (11.3), we can rewrite (11.3) as

$$V^n(1) \quad = \quad \max\left\{\frac{n}{N}, V^n(0)\right\}. \tag{11.5}$$

From this we obtain the inequality

$$V^n(1) \geq V^n(0) \tag{11.6}$$

which seems pretty intuitive (you are better off if the last candidate you interviewed was the best you have seen so far).

At this point, we are going to suggest a policy that seems to be optimal. We are going to interview the first $\bar{n}$ candidates, without hiring any of them. Then, we will stop and hire the first candidate who is the best we have seen so far. The decision rule can be written as

$$x^n(1) \quad = \quad \begin{cases} 0 \text{ (continue)} & n \leq \bar{n} \\ 1 \text{ (quit)} & n > \bar{n} \end{cases}.$$

To prove this, we are going to start by showing that if $V^m(1) > m/N$ for some m (or alternatively if $V^m(1) = m/N = V^m(0)$), then $V^{m'}(1) > m'/N$ for $m' < m$. If $V^m(1) > m/N$, then it means that the optimal decision is to continue. We are going to show that if it was optimal to continue at set m, then it was optimal to continue for all steps $m' < m$.

Assume that $V^m(1) > m/N$. This means, from equation (11.5), that it was better to continue, which means that $V^m(1) = V^m(0)$ (or there might be a tie, implying that $V^m(1) = m/N = V^m(0)$). This allows us to write

$$\begin{aligned} V^{m-1}(0) &= \frac{1}{m}V^m(1) + \frac{m-1}{m}V^m(0) \\ &= V^m(1) \qquad (11.7) \\ &\geq \frac{m}{N}. \qquad (11.8) \end{aligned}$$

Equation (11.7) is true because $V^m(1) = V^m(0)$, and equation (11.8) is true because $V^m(1) \geq m/N$. Stepping back in time, we get

$$\begin{aligned} V^{m-1}(1) &= \max\left\{\frac{m-1}{N}, V^{m-1}(0)\right\} \\ &\geq \frac{m}{N} && (11.9) \\ &> \frac{m-1}{N}. && (11.10) \end{aligned}$$

Equation (11.9) is true because $V^{m-1}(0) \geq m/N$. We can keep repeating this for $m-1, m-2, \ldots$, so it is optimal to continue for $m' < m$.

Now we have to show that if $N > 2$, then $\bar{n} \geq 1$. If this is not the case, then for all n, $V^n(1) = n/N$ (because we would never continue). This means that (from equation (11.4)):

$$\begin{aligned} V^n(0) &= \left(\frac{1}{n+1}\right)\left(\frac{n+1}{N}\right) + \left(\frac{n}{n+1}\right) V^{n+1}(0) \\ &= \frac{1}{N} + \left(\frac{n}{n+1}\right) V^{n+1}(0). && (11.11) \end{aligned}$$

Using $V^N(0) = 0$, we can solve (11.11) by backward induction:

$$\begin{aligned} V^N(0) &= 0, \\ V^{N-1}(0) &= \frac{1}{N} + \frac{N-1}{N-1+1} V^N(0) \\ &= \frac{1}{N}, \\ V^{N-2}(0) &= \frac{1}{N} + \frac{N-2}{N-2+1}\left(\frac{1}{N}\right) \\ &= \frac{N-2}{N}\left(\frac{1}{N-2} + \frac{1}{N-1}\right). \end{aligned}$$

In general, we get

$$V^m(0) = \frac{m}{N}\left[\frac{1}{m} + \frac{1}{m+1} + \cdots + \frac{1}{N-1}\right].$$

We can easily see that $V^1(0) > \frac{1}{N}$; since we were always quitting, we had found that $V^1(1) = \frac{1}{N}$. Finally, equation (11.6) tells us that $V^1(1) \geq V^1(0)$, which means we have a contradiction.

This structure tells us that for $m \leq \bar{n}$

$$V^m(0) = V^m(1),$$

and for $m > \bar{n}$

$$\begin{aligned} V^m(1) &= \frac{m}{N}, \\ V^m(0) &= \frac{m}{N}\left[\frac{1}{m} + \frac{1}{m+1} + \cdots + \frac{1}{N-1}\right]. \end{aligned}$$

It is optimal to continue as long as $V^m(0) > m/N$, so we want to find the largest value for m such that

$$\frac{m}{N}\left[\frac{1}{m}+\frac{1}{m+1}+\cdots+\frac{1}{N-1}\right] \quad > \quad \frac{m}{N},$$

or

$$\left[\frac{1}{m}+\frac{1}{m+1}+\cdots+\frac{1}{N-1}\right] \quad > \quad 1.$$

If $N = 5$, then we can solve by enumeration:

$$\begin{array}{lll}
\bar{n}=1 & \frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4} & =2.08 \\
\bar{n}=2 & \frac{1}{2}+\frac{1}{3}+\frac{1}{4} & =1.08 \\
\bar{n}=3 & \frac{1}{3}+\frac{1}{4} & =0.58
\end{array}$$

So for $N = 5$, we would use $\bar{n} = 2$. This means interview (and discard) two candidates, and then take the first candidate that is the best to date.

For large N, we can find a neat approximation. We would like to find m such that:

$$\begin{aligned}
1 \quad &\approx \quad \frac{1}{m}+\frac{1}{m+1}+\cdots+\frac{1}{N-1} \\
&\approx \quad \int_M^N \frac{1}{x}dx \\
&= \quad \ln N - \ln m \\
&= \quad \ln\left(\frac{N}{m}\right).
\end{aligned}$$

Solving for m means finding $\ln(N/m) = 1$ or $N/m = e$ or $m/N = e^{-1} = 0.368$. So, for large N, we want to interview 37 percent of the candidates, and then choose the first candidate that is the best to date.

The secretary problem is a classic, partly because it illustrates an interesting information collection problem, and partly because it yields such an elegant solution. In real applications, we can translate the result to a rough rule that says “look at a third of the candidates, and then choose the first candidate that is better than all the others.”

11.3 BIBLIOGRAPHIC NOTES

Section 11.1 - The sequential probability ratio test is due to Wald & Wolfowitz (1948).

Section 11.2 - The secretary problem was first introduced in Cayley (1875). Our presentation is based on Puterman (1994). Vanderbei (1980) provides an elegant generalization of the secretary problem to one of finding the best subset. See also Bruss (1984) for extensions.

PROBLEMS

11.1 Download a spreadsheet illustrating the sequential probability ratio test from

`http://optimallearning.princeton.edu/exercises/SPRT.xls`

In the initial spreadsheet, the standard deviation of an experimental observation has been set to 5.

a) Translate the cells in row 12 to mathematics. Identify the equation in the book corresponding to each cell starting in column D.

b) Change the probability of missing from .02 to .10. How does this change the hypotheses $H0$ and $H1$? Now change the probability of a false alarm from .02 to .10? How does this change the hypotheses? When you are done this question, restore both probabilities back to .02.

c) The sequential probability ratio test stops with a conclusion that either hypothesis $H0$ is true or $H1$ is true when the red line crosses one of the blue lines. If it does not cross either line within the experimental budget, then the test is inconclusive. Hit the F9 key 20 times, and count the number of times the red line crosses one of the blue lines.

d) Now change the standard deviation (cell B9) to 3. Again perform 20 simulations and count how many times the red line crosses one of the blue lines. You should see that with a lower standard deviation, SPRT is more effective at declaring that one of the two hypotheses is true. Why does a smaller standard deviation make it easier to come to a conclusion?

11.2 Assume you think you can look at up to 20 bids for the house you are selling. Completely unknown to you, the bids can be modeled as being drawn from a uniform distribution between $380,000 and $425,000. Use the policy that you are going to reject the first seven bids (which is approximately 37 percent of 20), and then accept the first bid that is better than these seven. If none are better, you have to accept the very last bid. Start the process by generating all 20 bids in advance, so that when you are done, you can compare the bid you accepted against the best that you might have accepted with perfect foresight.

a) Repeat this policy 100 times and report on: i) how many times you accepted the very best bid out of the 20, and ii) the bid you accepted as a percentage of the very best bid.

b) Repeat (a), but now reject only the first four bids before you are ready to accept the bid.

c) Finally, repeat (a), but now you decide to reject the first 10 bids before you are ready to accept a bid.

d) Compare the results of the policies you simulated in (a), (b) and (c). Do you see evidence that one policy outperforms the others?

11.3 You have to choose the best out of up to 30 bids, where the i^{th} bid, R_i, follows an exponential distribution given by

$$f_R(y) = .02e^{-.02y}.$$

You can generate random observations from an exponential distribution using

$$R = -50 \ln U$$

where U is a random variable that is uniformly distributed between 0 and 1.

a) Use the policy where you look at the first 11 bids, and then pick the best bid which outperforms all previous bids. Repeat this policy 100 times and report on: i) how many times you accepted the very best bid out of the 20, and ii) the bid you accepted as a percentage of the very best bid.

b) Repeat (a), but now reject only the first six bids before you are ready to accept the bid.

c) Finally, repeat (a), but now you decide to reject the first 16 bids before you are ready to accept a bid.

d) Compare the results of the policies you simulated in (a), (b) and (c). Do you see evidence that one policy outperforms the others?