TOWARD A MORE REALISTIC FORMULATION OF THE SECRETARY PROBLEM

by

Thomas Jay Lorenzen

Department of Statistics
Division of Mathematical Sciences

Mimeograph Series #487
May, 1977

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Lorenzen, Thomas Jay. Ph.D., Purdue University, May, 1977
Toward a More Realistic Formulation of the Secretary Problem.
Major Professor: S. M. Samuels.

This thesis considers several variations in the classical formulation of the so-called Secretary Problem. These variations are presented to make the problem more realistic.

Chapter II contains a new formulation of the infinite secretary problem, not limited by the classical formulation. The solution is obtained and general results are given. These results are useful in both of the variations considered in Chapters III and IV.

The first variation contained in Chapter III, is the addition of a cumulative interview cost. The finite problem is formulated and solved using backward induction. In the non-trivial case, the finite problem tends to an infinite problem defined on the unit interval. The solution is derived and explicitely given for the linear cost rank problem.

In the final chapter, backward solicitation is considered. In this variation, you are allowed to solicit a candidate who has previously been interviewed. However, the solicitation may not be successful. It is shown that the exact solution is not easily obtainable in general. A specific class can be solved and the limiting solution derived. For any problem the limiting solution can be found by considering the unique member of this specific class.

CHAPTER I
INTRODUCTION AND BACKGROUND

This thesis deals with several generalizations of the so-called Secretary Problem. These generalizations are presented to make the formulation of the problem more realistic.

The version referred to as the Secretary Problem is formulated in the following way: n applicants apply for a secretarial position. Some criterion is used to rank the candidates from 1 (best) to n (worst) with no ties. The applicants arrive in random order and when the $i^{\underline{th}}$ candidate appears (time i) she is ranked relative to the i-1 previous candidates. On the basis of her relative rank she is either hired (and the process stops) or passed by. Once a candidate is passed, she can never be recalled. If the $n\text{-}1^{\underline{st}}$ candidate is passed, the last candidate must be hired.

The problem as stated has been known under other names, beauty contest, dowry problem, marriage problem, and google just to mention a few. This problem can be applied to any process where the decision to stop is based solely on relative ranks. For this reason, the number of applications is virtually unlimited.

The first time a problem of this form appeared was in 1960 in Martin Gardner's column: Mathematical games [5] and [12]. It is said to have an earlier origin for Gilbert and Mostellar [8] claim to have heard of the problem in 1955 from Andrew Gleason who is said to have heard of it elsewhere!

Classically, the problem has had two forms, the best choice problem and the rank problem.

1. Best Choice Problem

In the best choice problem, we are interested in finding a stopping rule which maximizes the probability of choosing the best candidate. This is the form of the problem as it appeared in 1960 under the provocative name google. Google works as follows: n slips of papers have n different numbers written on them. The numbers range anywhere from a small fraction to a "google." (A google is a one followed by one hundred zeros.) They are mixed and placed face down. An opponent turns them up one at a time and announces when he believes he has just turned up the largest number. What strategy should he use and what is the probability of success?

Various methods are available to solve this problem. The optimal solution has been shown to be of the form: let k individuals go by then stop and hire the first candidate of relative rank 1. It is easily shown that for a given n the optimal k is

(1.1) $$k(n) = \inf \{r \mid \frac{1}{r} + \frac{1}{r+1} + \ldots + \frac{1}{n-1} \leq 1\},$$

(1.2) $$k(n)/n \to e^{-1} \quad \text{as } n \to \infty, \text{ and}$$

(1.3) $$P(\text{best choice}) \to e^{-1} .$$

2. Rank Problem

Under the hypothesis of the rank problem we are interested in minimizing the expected rank of the candidate chosen. So the loss for hiring the $i^{\underline{th}}$ (absolute) ranked candidate is i. For notation let

(1.4) $$V_n = \text{minimal expected loss.}$$

The best known method of solution is backward induction, sometimes called dynamic programming. (See Lindley [11].) That is, for any given n compute successively $C_{n-1}(n)$, $C_{n-2}(n), \ldots, C_0(n)$ where

(1.5) $$C_k(n) = \text{minimal expected loss given you have not stopped by time } k.$$

Then $V_n = C_0(n)$.

It is easy to show the stopping rule consists of a cutoff point rule. That is, there is an increasing sequence α_i which gives the optimal policy. The optimal policy is to stop on the first candidate of relative rank $\leq i$ arriving at or after time α_i. Take the last candidate if you have not previously stopped.

Chow, Moriguti, Robbins, and Samuels (CMRS) [3] have shown

(1.6) $$V_n \to \prod_{j=1}^{\infty} \left(\frac{j+2}{j}\right)^{\frac{1}{j+1}} \approx 3.87 \quad \text{as } n \to \infty \quad \text{and}$$

(1.7) $$\frac{\alpha_i}{n} \to \prod_{j=i}^{\infty} \left(\frac{j}{j+2}\right)^{\frac{1}{j+1}} .$$

3. General Loss Problem

From the rank problem a more general problem can be stated. Suppose $q(\cdot)$ is increasing and

(1.8) $$q(i) = \text{loss for stopping on candidate of (absolute) rank } i.$$

We are interested in minimizing the expected loss over all stopping rules. (Note $q(1) = 0$, $q(i) = 1$ for $i > 1$ is the best choice problem while $q(i) = i$ is the rank problem.) Again backward induction gives the solution

(1.9) $$C_{r-1}(n) = \frac{1}{r} \sum_{k=1}^{r} \min(Q(r,k),\ C_r(n)) \quad \text{where}$$

(1.10) $Q(r,k)$ = expected loss for taking a candidate of relative rank k at time r

$$= \sum_{\ell=k}^{n-(r-k)} q(\ell) \frac{\binom{\ell-1}{k-1}\binom{n-\ell}{r-k}}{\binom{n}{r}} .$$

$V_n = C_0(n)$ and it can be shown that the optimal policy is again a cutoff point rule. But, in general there is no simple relation between $C_k(n)$ and $C_k(n+1)$ so there is no simple way to evaluate $\lim V_n$ as $n \to \infty$.

CMRS [3] have suggested a heuristic method for obtaining this limiting solution. This is to write (1.9) as a difference equation, normalizing and letting this tend to a differential equation. For a certain class of $q(\cdot)$'s Mucci [13] and [14] showed that the differential equation actually is the limit of the difference equations. So if V is the solution to the resulting differential equation, we know $V_n \to V$ for these cases.

4. Infinite Problem

To further consider the limit of this general problem, Gianini and Samuels [7] following a suggestion of Rubin [17], considered the infinite secretary problem. The formulation was motivated by the following reasoning: Suppose we shrink the interval [0,n] to [0,1]. Then the candidates arrive randomly at times $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}$. If we consider the best candidate, she arrives uniformly among the points $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}$. The second best arrives uniformly among the remaining n-1 points. And so forth. Letting $n \to \infty$ we conclude the best candidate arrives uniformly in [0,1], as does the second best, and so on. That is, if

(1.11) U_i = arrival time of $i^{\underline{th}}$ absolute best

then U_i are identically independently distributed uniform [0,1]. At each arrival time T the decision whether to stop or go on is based solely on the observed relative ranks. If we reach time 1, pay sup q(n). Some care must be taken since the set of arrival times is almost surely dense. In other words, it is meaningless to say "the next candidate." If we let

(1.12) V = minimal expected loss and

(1.13) $f(x)$ = minimal expected loss given we have not stopped at time x

then Gianini and Samuels have shown $f(0) < \infty$ implies $f(x) < \infty$ for all x in [0,1] and f satisfies

(1.14) $$f'(x) = \frac{1}{x} \sum_{k=1}^{\infty} [f(x) - R_k(x)]^+ \quad \text{where}$$

(1.15) $$R_k(x) = \lim_{\frac{r}{n} \to x} Q(r,k) = \sum_{\ell=k}^{\infty} q(\ell) \binom{\ell-1}{k-1} x^k (1-x)^{\ell-k} .$$

This is the heuristically postulated differential equation. In addition

(1.16) $$V = f(0)$$

and the optimal rule is a cutoff point rule with α_i satisfying

(1.17) $$f(\alpha_i) = R_i(\alpha_i).$$

Gianini [6] showed by using intermediate models that

(1.18) $$\lim_{n \to \infty} V_n = V.$$

Thus, whenever $\lim V_n < \infty$, $\lim V_n = V = f(0)$ where f satisfies (1.14).

5. Unknown Number of Applicants

Recently, generalizations of the secretary problem have appeared make the formulation more realistic. One generalization considers the number of secretaries (n) stochastic. Instead of being fixed, a known distribution is assumed.

In 1975 Rasmussen and Robbins [16] consider the best choice problem with a random number of applicants. They showed that for n uniformly distributed on [1,2,..., N], P(best choice) $\rightarrow 2e^{-2}$ as $N \rightarrow \infty$. Earlier, two Russians, Presman and Sonin [15] showed the same result using different methods. In addition, they showed for n distributed Poisson (λ), P(best choice)$\rightarrow e^{-1}$ as $\lambda \rightarrow \infty$. They also showed the solution need not be a cutoff point rule.

To show a simple example, consider n = 5 with probability .999 and n = 1000 with probability .001. You must act as if there are 5 candidates so solicit the $4^{\underline{th}}$ or $5^{\underline{th}}$ candidate if she has relative rank 1. But, if you don't solicit and a $6^{\underline{th}}$ candidate comes along, you know there will be 1000 candidates. Therefore skip about 1000 e^{-1} candidates before considering stopping.

Gianini [7] considered the rank problem with the number of applicants random. She showed the optimal rule need not be a cutoff point rule. In addition, if n is uniformly distributed on 1,2,..., N, then $V_N \rightarrow \infty$ as $N \rightarrow \infty$.

The problem is still unsolved for the general loss q(•) problem even for uniform n.

6. Backward Solicitation

Another type of generalization has been termed backward solicitation. The difference is once a candidate is passed, she is not necessarily lost forever. Instead, you have some probability p of going back and successfully hiring (soliciting) her. Of course the longer you wait the smaller the probability is that she will accept. Once an applicant refuses, she will always refuse. (Presumably she has taken another position.) Formally we will present Yang's formulation [20]. Let p(•) be given with p(0) = 1 and p(•) decreasing. If we are currently at time k the probability that the candidate interviewed at time i will accept is $p(k-i)$ or 0 depending if she has previously refused or not.

Yang [20] considered this formulation with the best choice problem. The problem is solved using a set of recursion equations. For $p(0) = 1$, $p(i) = p$ $i = 1,2,\ldots$ he showed $P(\text{best choice}) \to e^{p-1}$ and for $p(i) = p^i$ $P(\text{best choice}) \to e^{-1}$. (Both limits as $n \to \infty$.) Smith [18] extended these results to show $P(\text{best choice}) \to e^{p^*-1}$ where $p^* = \lim_{i\to\infty} p(i)$.

The problem for general $q(\cdot)$ and backward induction is considered in Chapter IV. It is shown that backward induction requires too many steps for general use. Nevertheless, we can show that the limiting minimal risk, as $n \to \infty$, depends solely on $q(\cdot)$ and $p^* = \lim_{i\to\infty} p(i)$. For the particular recall function $p(0) = 1$, $p(i) = p^*$ for $i \geq 1$, this limiting risk is obtained from a differential equation similiar to (1.14). This recall function also yields procedures which are asymptotically optimal for general $p(\cdot)$.

7. Interview Cost

A third generalization has been considered with the addition of an interview cost. (Or sampling cost.) It is assumed that for each additional interview a certain known cost must be paid. Then you must minimize the expected cost plus the expected loss for the candidate chosen.

Bartoszynski and Govindarajulu [1] have considered this generalization for the best choice problem. They showed for "high" cost, you should take the first candidate while for "low" cost the optimal policy is similiar to the classical best choice problem. For "low" cost, estimates of $k(n)$ are given for large n and linear or quadratic cost.

(Recall the optimal policy is to let k(n) applicants go by and take the first relatively best to arrive.)

Govindarajulu [9] also considered the rank problem with interview cost. Using backward induction, a set of recursive equations is derived. This set gives the solution to the finite problem. To evaluate the limit, approximate differential equations are given. Hopefully these give solutions which are good approximations to the actual limiting risk. But as Govindarajulu himself points out, the optimal policy remains unknown and the risks are only approximate. As an example, for zero interview cost his best approximation is 4 while the true limiting risk has been shown to be approximately 3.87. (CMRS [3])

Chapter III is devoted to the general q(•) problem with interview cost. For the finite case, the simple recursive equations are derived. The optimal procedure is no longer a cutoff point rule. Instead, the optimal rule is described by island rules. That is, the intervals of stopping can be thought of as islands in the interval [0,n]. An example of the existence of an island is given.

To evaluate the limit, the heuristic method suggested in CMRS is used. To do this, a special form of the interview cost is assumed. Under this special form, the infinite formulation of Gianini and Samuels makes sense and can be solved. The solution turns out to be the heuristic differential equation. It is shown that lim $V_n \geq V$ and if the interview cost behaves nicely, lim $V_n = V$.

For interview costs not of this special form, the limiting solution is always trivial. Either ignore the interview cost or always take the first applicant.

The rank problem with fixed sample cost a per interview is considered. This problem falls in the trivial class so $\lim V_n \to \infty$. However, for fixed n the finite problem is approximated by the infinite problem with linear cost and the rate of convergence can be estimated. This example illustrates the usefulness of the infinite problem.

8. More General Loss Structure

In all of the previous work, a loss function $q(\cdot)$ was specified and from $q(\cdot)$ we calculated $Q(r,k)$ in the finite problem and $R_k(x)$ in the infinite problem. For the moment, let $R_k^{(n)}\left(\frac{r}{n}\right) = Q(r,k)$. From the calculations we actually forget the $q(\cdot)$'s and work only with the $R_k(\cdot)$'s. This suggests that we drop the $q(\cdot)$'s altogether and instead of using an expected loss, use an actual loss. That is, let

(1.19) $\quad A_k(t)$ (or $A_k^n(t)$) = loss for selecting a candidate of relative rank k at time t

where the $A_k(\cdot)$'s are not necessarily derived from some $q(\cdot)$.

In fact, both Chapters III and IV can be regarded in this form. Chapter III has

(1.20) $\quad A_k(t) = R_k(t) + h(t)$ where h is the cumulative interview cost

while the most important special case in Chapter IV has

(1.21) $$A_k(t) = p\,R_1(t) + p(1-p)\,R_2(t) + \ldots + p(1-p)^{k-2}\,R_{k-1}(t) + (1-p)^{k-1}\,R_k(t)$$

We ask the usual questions about this generalized loss structure problem.

1) Does the differential equation $f'(t) = \frac{1}{t} \sum_{k=1}^{\infty} [f(t) - A_k(t)]^+$ still hold?

2) What is the form of the optimal procedure? Is it a cutoff point rule?

3) For unbounded loss does the optimal risk for the truncated loss converge to the optimal risk for the untruncated loss?

4) Do the optimal risk and procedures for the finite problem converge to those for the infinite problem?

The answers depend on the properties imposed on $A_k(\cdot)$. If we demand they have all of the following properties:

1) Non-decreasing in k for each $t \in (0,1)$;

2) continuous on (0,1), right continuous at 0 and left continuous at 1 for each k;

3) $A_1(0) = A_2(0) = \dots$;

4) $\lim_{k\to\infty} A_k(t)$ non-decreasing in t and continuous on (0,1), left continuous at 0 and right continuous at 1;

5) $\sup_{1\leq k\leq i\leq n} \left| A_k^{(n)}\left(\frac{i}{n}\right) - A_k\left(\frac{i}{n}\right) \right| \to 0$ as $n\to\infty$;

6) $A_k(t)$ non-increasing in t for each k;

then the answer to all four questions is yes! The classical formulation satisfies all six of these criteria. Moreover, the proofs presented in Gianini and Gianini and Samuels go through with very little additional effort.

Formula (1.21) satisfies properties 1) through 6) so the special case of the recall problem is no more complex than the classical problem.

Properties 1) through 4) are sufficient to guarantee affirmative answers to questions 1) through 3) but without property 6) the optimal procedure may be an island rule as we have previously noted. The islands are specified by a differential equation.

The main difficulty when property 6) does not hold is with question 4). If all losses are bounded, we can still get an affirmative answer

CHAPTER II
A MORE GENERAL LOSS STRUCTURE

The first part of this chapter will consist of an analysis of the infinite secretary problem (see Gianini and Samuels [7]) under the more general loss structure defined in the introduction. The second considers the relations between the finite and infinite problems.

The first section contains many results similar to those in Gianini [6] and Gianini and Samuels [7]. If the proofs are straightforward extensions improvements or when old proofs do not extend readily. Among the results, it will be shown that the minimal risk and optimal procedure is given by a differential equation with a right boundary condition. It is also shown that the risk for the truncated loss problem tends to the risk for the non-truncated problem. (Three methods of truncation are considered.)

In the second section, it is shown that for truncated loss, the risk for the finite problem tends to the risk for the infinite problem. In general, it can be shown that the limiting risk for the finite problem is at least as large as the infinite problem risk.

The main results of Chapter II will be used in both Chapters III and IV, where the interview cost and the recall problems are considered. In Chapter IV it is shown that the limiting risk actually equals the infinite problem risk. In Chapter III, conditions are given that guarantee the equality of the two risks.

using equicontinuity. For unbounded losses, we can get the difficult inequality $\underline{\lim}\, V_n \geq V$. But the other inequality, a simple consequence of property 6) can no longer be asserted in general. For Chapter III, it does hold whenever the interview cost is bounded.

In Chapter II, proofs are given for only two reasons, either they improve previous proofs or they do not follow immediately. Otherwise we merely indicate why the proofs still work.

1. The Differential Equation for f

Using notation similar to Gianini and Samuels let

(2.1) $A_k(x)$ = loss for stopping at time x on a candidate of relative rank k,

(2.2) U_i = arrival time of $i^{\underline{th}}$ best (U_i are IID Uniform [0,1]),

(2.3) X_t = absolute rank of candidate arriving at time t

(2.4) Y_t = relative rank of candidate arriving at time t (w.r.t. [0,t])

(2.5) τ = stopping rule satisfying

1) $0 \leq \tau \leq 1$

2) τ is a function of the $Y\frac{s}{t}$

3) $\tau \in \{U_1, U_2, \ldots\} \vee \{0,1\}$ (If $\tau = 0$, we pay $A_1(0)$ and if $\tau = 1$ we pay $\sup_k A_k(1) = A_\infty(1)$.)

(2.6) $V = \inf_\tau E(A_{Y_\tau}(\tau))$ = minimal expected loss where $Y_0 = 1$ and $Y_1 = \infty$.

(2.7) $f(t) = \inf_{\tau > t} E(A_{Y_\tau}(\tau))$

Let us assume certain properties on the loss function $A_k(x)$.

(2.8) $A_k(x) \leq A_{k+1}(x)$ for all k and x

(2.9) $A_k(x)$ is finite and continuous on $(b_k, 1)$ for some b_k, and infinite on $[0, b_k]$

(2.10) $A_1(0^+) = A_2(0^+) = \ldots$

(2.11) $A_\infty(x) \equiv \lim_{k \to \infty} A_k(x)$ is non-decreasing in x and continuous.

The first question must concern the relation between f and V. It is obvious that $V \leq f(t)$ for all t but it is not true that every procedure is greater than some t greater than zero. Therefore, it is not obvious that $V > f(0^+)$. However, as in Gianini and Samuels

Prop. 2.1 $V = f(0^+)$.

Proof:

$V \leq f(0^+)$ is immediate.

Consider any interval $[0,t]$. The minimal possible risk is $\inf_{0 \leq s \leq t} A_1(s)$ and $f(t)$ is the minimal risk for all procedures stopping after time t so

$$V \geq \min(f(t), \inf_{0 \leq s \leq t} A_1(s))$$

Since this holds for all t

$$V \geq \min(f(0^+), A_1(0^+)) \tag{2.12}$$

By considering any procedure that stops in $(t, t+\delta)$ and letting $\delta \to 0$ it follows that $f(t) \leq A_\infty(t)$. Now letting $t \to 0$,

$$f(0^+) \leq A_\infty(0^+) = A_1(0^+) = A_1(0). \tag{2.13}$$

Inequality (2.12) together with (2.13) give $V \geq f(0^+)$. //

If $V = A_1(0)$, we are in a trivial case, i.e., the game is not worth starting. So let us assume $V < A_1(0)$. (Equivalently $f(0^+) < A_1(0)$.) Then $f(t_0) < \inf_{0 \leq s \leq t_0} A_1(s)$ for some t_0 and we get $V \equiv f(t)$ for all t sufficiently small.

Of course, the main result of this section has to be the derivation of the function f. It is little surprise that this function takes the same form as the function given in Gianini and Samuels.

Thm. 2.1 $f(t)$ as defined by (2.7) is continuous and satisfies the differential equation

$$f'(t) = \frac{1}{t} \sum_{k=1}^{\infty} [f(t) - A_k(t)]^+ \tag{2.14}$$

subject to $f(1) = A_\infty(1)$.

Proof:

First I will prove continuity. Using the reasoning in proposition 2.1

$$f(s) \geq \inf_{s \leq x \leq t} A_{Y_x}(x) \wedge f(t) \qquad \text{for all } s < t. \tag{2.15}$$

Let $s < \theta \le t$. Then $Y_\theta \uparrow \infty$ as $t-s \to 0$ so

(2.16) $$f(s) \ge A_\infty(s) \wedge f(s^+) \qquad \text{for all } s.$$

But $f(t) \le A_\infty(t)$ implies $f(s^+) \le A_\infty(s)$ and $f(s) \le f(s^+)$ so

(2.17) $$f(s) = f(s^+).$$

The same reasoning gives $f(t^-) = f(t)$.

Let t be such that $A_1(t) < f(t)$. Then define

(2.18) $$r = \max\{j \mid A_j(t) < f(t)\}$$ and s close enough to t that

(2.19) $$A_r(u) < f(u) \text{ for all } u \in (s,t).$$

If $r = \infty$, then we will stop by time t with probability one (see next proposition) so assume $r < \infty$. Define a stopping rule σ by

(2.20) $$\sigma = \inf \{U_i \in (s,t) \mid Y_{u_i} \le r\}$$

1 if the set is empty

$\sigma I_{\{\sigma \le t\}} + \tau I_{\{\sigma > t\}}$ is an acceptable rule so we let τ approach optimality in the set of all rules greater than or equal to t and conclude

(2.21) $$f(s) \le E(A_{Y_\sigma}(\sigma)\, I_{\{\sigma \le t\}} + f(t)\, I_{\{\sigma > t\}})$$

$$\le E([\sup_{s<x\le t} A_{Y_\sigma}(x)] \wedge f(t))$$

Defining θ so $A_{Y_\theta}(\theta) = \inf_{s \le x \le t} (A_{Y_x}(x))$ for $s < \theta \le t$ we get

$$f(s) \ge E(A_{Y_\theta}(\theta) \wedge f(t))$$

Now $Y_\sigma \le r$ iff $Y_\theta \le r$, $P(Y_\theta > r) = (s/t)^r$, and $P(Y_\sigma = k) = 1/r$ for $k = 1,2,\ldots, r$ on $\{\sigma < 1\}$ so

(2.22) $$E(\sup_{s<x\le t} A_{Y_\sigma}(x) \wedge f(t)) = \frac{(1-(s/t)^r)}{r} \sum_{k=1}^{r} (\sup_{s<x\le t} A_k(x)) + \left(\frac{s}{t}\right)^r f(t)$$

and for s close enough to t

$$(2.23)\quad E(A_{Y_\theta}(\theta)\wedge f(t)) = \sum_{k=1}^{r} \left(\frac{s}{t}\right)^{k-1} (1-\frac{s}{t})\, A_k(\theta) + (s/t)^r\, f(t).$$

Letting $s \to t$, using continuity, and noting $\frac{1-(s/t)^r}{s-t} \to \frac{r}{t}$ we have

$$\frac{1}{t}\sum_{k=1}^{\infty}[f(t)-A_k(t)]^+ = \lim_{s\uparrow t} \frac{f(t) - E(A_{Y_\theta}(\theta)\wedge f(t))}{t-s}$$

$$\geq \lim_{s\uparrow t} \frac{f(t)-f(s)}{t-s}$$

$$= f'(t)$$

$$\geq \lim_{s\uparrow t} \frac{f(t) - E(\sup_{s<x\leq t} A_{Y_\sigma}(x)\wedge f(t))}{t-s}$$

$$= \frac{1}{t}\sum_{k=1}^{\infty}[f(t)-A_k(t)]^+ .$$

But obviously when $A_1(t) > f(t)$, $f'(t) = 0$ so the differential equation also holds.//

The optimal procedure is also a consequence of the differential equation as will be proved in the next theorem. The procedure bears the same relationship to the differential equation as the classical case but loses many nice properties. In the classical case you are to stop at time x on a candidate of relative rank k if $R_k(x) \leq f(x)$ and you have not previously stopped. Since $R_k(x)$ is decreasing and $f(x)$ is increasing it follows immediately that there are unique α_k satisfying $R_k(\alpha_k) = f(\alpha_k)$. The optimal rule is then to stop the first time after α_k an arrival of relative rank k appears. These rules have been called cutoff point rules.

Like the classical case, the optimal procedure is to stop at time x on a candidate of relative rank k if $A_k(x) \leq f(x)$. However, since

$A_k(x)$ is not necessarily decreasing, there need not be unique X_k for which $A_k(x_k) = f(x_k)$. Despite this fact, there are still fixed closed intervals where $R_k(x) \leq f(x)$. These intervals can be thought of as islands.

Def. τ is an island rule if τ stops the first time a candidate of relative rank k arrives in I_k where $[0,1] \geq I_1 \geq I_2 \geq \ldots \geq \{1\}$.

The restriction $I_k \geq I_{k+1}$ is not a restriction at all since $A_{k+1}(x) \leq f(x)$ implies $A_k(x) \leq f(x)$.

We will refer to cutoff point rules as single island rules.

Perhaps the nicest property of island rules is that they are fixed. That is, I_k is defined independent of the arrival times of the candidates and can be computed prior to the start of the process.

Of course, the most important thing about island rules is that some island rule is optimal. That is,

Thm. 2.2 The island rule τ^* given by

(2.24) $I_k = \{x | A_k(x) \leq f(x)\}$ is optimal whenever $v < \infty$.

Proof:

Let τ be any stopping rule with $E(A_{Y_\tau}(\tau)) < \infty$. I will show τ can be improved on the sets $\{\tau < \tau^*\}$ and $\{\tau > \tau^*\}$.

Assume $P(\tau < \tau^*) > 0$. If we let $I_0 = [0,1]$, then there exists k with $P(\{\tau < \tau^*\} \cap \{\tau \in I_{k-1} - I_k\}) > 0$. Since $A_k(x)$ is continuous, for each $x \in I_{k-1} - I_k$ there exists ϵ_x so that

$$A_k(y) > f(x) \quad \text{for } y \in (x-\epsilon_x, x).$$

Then, using the finite subcovering theorem, for some $B_x = (x-\epsilon_x, x)$, $P(\{\tau < \tau^*\} \cap \{\tau \in I_{k-1} - I_k\} \cap B_x) > 0$.

But on $\{\tau<\tau^*\}\cap\{\tau\in I_{k-1}-I_k\}\cap B_x$, $E(A_{Y_\tau}(\tau))\geq A_k(\tau)>f(x)$.
So τ can be strictly improved by waiting until x and using a rule with expected loss sufficiently close to f(x).

So $\max(\tau,\tau^*)$ is at least as good as τ for all τ.

Assume $P(\tau>\tau^*)>0$. It is easy to see for any stopping rule σ $E(A_{Y_\sigma}(\sigma)\, I_{\{\sigma>t\}}\geq f(t)\, I_{\{\sigma>t\}}$. As in Gianini and Samuels this can be extended to

$$E(A_{Y_\sigma}(\sigma))\, I_{\{\sigma>\tau'\}}\geq f(\tau')\, I_{\{\sigma>\tau'\}} \quad \text{a.s.}$$ This and (2.24)

gives $E(A_{y_\tau}(\tau))\; I_{\{\tau>\tau^*\}}\geq f(\tau^*)I_{\{\tau>\tau^*\}}$

$$\geq A_{Y_{\tau^*}}(\tau^*)\; I_{\{\tau>\tau^*\}} .$$

This implies $\min(\tau,\tau^*)$ is at least as good as τ for all τ.

Let τ_n be such that $E(A_{Y_{\tau_n}}(\tau_n))\to V$ and $\tau_n'=\min(\tau_n,\tau^*)$. Since $\tau^*=\max(\tau^*,\tau_n')$,

$$\begin{aligned} V &\leq E(A_{Y_{\tau^*}}(\tau^*)) \\ &\leq E(A_{Y_{\tau_n'}}(\tau_n')) \\ &\leq E(A_{Y_{\tau_n}}(\tau_n)). \end{aligned}$$

But $E(A_{Y_{\tau_n}}(\tau_n))\to V$ so we have equality.//

<u>Cor. 1</u> $f(t)=E(A_{Y_{\tau_t^*}}(\tau_t^*))$ for all t where τ_t^* is the island rule given by $I_k=\{x\mid A_k(x)\leq f(x)\}\cap(t,1]$.

<u>Proof</u>:

Exactly as above.//

2. The Truncated Problem

To fully understand the infinite secretary problem it is often necessary to consider the truncated problem. In the classical situation, a natural truncation is to truncate the increasing function

$q(\cdot)$. This gives both an upper bound on the loss and a greatest acceptable relative rank.

Corresponding closely to this, we consider the truncation

(2.25) $$A_k^M(x) = A_k(x) \wedge A_M(x) \wedge M$$

of course there is nothing forcing us to consider an upper bound of M. We could just as easily consider

(2.26) $$A_k^M(x) = A_k(x) \wedge A_M(x) \wedge h(M)$$

for some increasing function h. Two other truncations of interest are

(2.27) $$B_k^M(x) = A_k(x) \wedge A_M(x) \qquad \text{and}$$

(2.28) $$C_k^M(x) = A_k(x) \wedge M.$$

Each is interesting in itself and is squeezed between the non-truncated problem and (2.25). So if we prove the risk for loss (2.25) tends to the risk for non-truncated loss then we have the same relation holding for (2.27) and (2.28).

The main result of this section is to show the risk for loss (2.25) tends to V. That is, if we let $\hat{\tau}$ be the optimal procedure for loss function (2.25) then we will show $E(A_{Y_{\hat{\tau}}}(\hat{\tau})) \to V$ as $M \to \infty$. This gives similar results for (2.26), (2.27), and (2.28).

To prove this result, I need the following lemma:

Lemma 1. Assume $A_\infty(1) < \infty$. Let $\tau^{(n)}$ and $\tau^{(\infty)}$ be island rules defined by $I_k^{(n)}$ and $I_k^{(\infty)}$. If $I_k^{(n)} \to I_k^{(\infty)}$ and $\inf I_1^{(\infty)} > 0$ then $E(A_{Y_{\tau(n)}}(\tau^{(n)})) \to E(A_{Y_{\tau(\infty)}}(\tau^{(\infty)}))$.

Proof:

The proof is an extension of that in Gianini and Samuels. Let $t_k^{(n)} = \inf(I_k^{(n)})$, $t_k^{(\infty)} = \inf(I_k^{(\infty)})$ for all k. Define

$$E_n = \{\tau^{(\infty)} < t_k^{(\infty)},\ \tau^{(m)} \neq \tau^{(\infty)} \text{ for some } m > n\}.$$

$E_n \neq \phi$ implies an arrival of relative rank i has appeared in $F_i^n = \bigcup_{m>n} I_i^{(m)} - \bigcap_{m>n} I_i^{(m)}$, for some i = 1,2,..., k. But $F_i^n \to 0$ so $P(E_n) \to 0$ as long as F_i^n is bounded away from 0. Inf $I_1^{(\infty)} > 0$ is sufficient.

The rest of the proof follows immediately.//

A sufficient condition for inf $I_1^{(\infty)} > 0$ is

Lemma 2 $\quad t_1^{(\infty)} > 0$ if $\overline{\lim}\ E(A_{Y_{\tau(n)}}(\tau^{(n)})) < A_1(0)$.

Proof:

Since $I_k^{(n)}$ is closed for all k, n, $t_1^{(\infty)} = \underline{\lim}\ t_1^{(n)}$. $E(A_{Y_{\tau(n)}}(\tau^{(n)})) \leq A_1^n(0)$ so $\overline{\lim}\ E(A_{Y_{\tau(n)}}(\tau^{(n)})) \leq A_1(0)$. If $\underline{\lim}\ t_1^{(n)} = 0$, then since $A_1(0) = A_2(0) = \ldots$, $\lim E(A_{Y_{\tau(n)}}(\tau^{(n)})) = A_1(0)$ by the reasoning in Gianini and Samuels [7].//

Thm. 2.3 $\quad$ Let V^M be the minimal risk for truncated loss (2.25). Then $V^M \to V$.

Proof:

Obviously $V^M \leq V^{M+1}$ and $V^M \leq V$ so $\lim V^M \leq V$. If $\lim V^M = A_1(0)$, we are done since $V \leq A_1(0)$. So assume $\lim V^M < A_1(0)$. Let $I_k^M = \{x \mid A_k^{(M)}(x) \leq f^M(x)\}$ be the island rule specified by τ^M the optimal rule for loss (2.25). Choose a subsequence M_k that converges as in lemma 1. For convenience relabel this subsequence and let $\tau^{(\infty)}$ be the limiting island rule. For each N apply lemma 1 and proposition 2.1 to show

$$V^M = E(A^M_{Y_{\tau M}}(\tau^M)) \geq E(A^{M \wedge N}_{Y_{\tau M}}(\tau^M)) \to E(A^N_{Y_{\tau(\infty)}}(\tau^{(\infty)})) \text{ as } M \to \infty.$$

By applying the monotone convergence theorem as $N \to \infty$ we get

$$\lim V^M \geq E(A^N_{Y_{\tau(\infty)}}(\tau^{(\infty)})) \to E(A_{Y_{\tau(\infty)}}(\tau^{(\infty)}))$$

$$\geq \inf_{\tau} E(A_{Y_\tau}(\tau)) = V. //$$

This also proves the same result for the three other truncations considered. Note also that the proof applies whether $V < \infty$ or $= \infty$.

3. Finite Problem and its Limit

In the classical formulation of the secretary problem a loss sequence $q(\cdot)$ is specified. From this sequence, the expected loss for stopping is computed for a finite number (n) of applicants and an infinite number. A natural relationship between the two exist. The expected loss for stopping at time i (out of n) in the finite problem tends to the expected loss for stopping at time x in the infinite problem as i/n tends to x. The difference goes to zero uniformly in i as $n \to \infty$.

We have already presented a loss function on the infinite problem, so let us consider

(2.29) $A_k^{(n)}(i)$ = loss for stopping on the $i^{\underline{th}}$ candidate (out of n) with relative rank k; where

(2.30) $$\sup_{\substack{1 \le i \le n \\ 1 \le k \le n}} \left| A_k^{(n)}(i) - A_k\left(\frac{i}{n}\right) \right| \to 0 \text{ as } n \to \infty.$$

Then we have defined a finite secretary problem and can obtain the solution using backward induction. That is, let

$C^n(i)$ = minimal risk given we have not stopped by time i,

V^n = minimal risk = $\min_{\tau} E(A^{(n)}_{Y_\tau}(\tau))$,

then backward induction gives

$$C^n(n-1) = \frac{1}{n} \sum_{k=1}^{n} A_k^{(n)}(n) \quad \text{and}$$

$$C^n(i-1) = \frac{1}{i} \sum_{k=1}^{i} \min(A_k^{(n)}(i) - C^n(i)).$$

Using heuristics, we write the recursion equation as a difference equation, normalize to [0,1], and let this tend to a differential equation, we arrive at $f'(x) = \frac{1}{x}\sum_{k=1}^{\infty}[f(x) - A_k(x)]^+$ which is immediately recognized as (2.14), the solution to the infinite problem.

The main result of this section is that the limiting risk for the finite problem equals the risk for the infinite problem when the loss is truncated. The method used is suggested by Gianini [6]; imbed the finite problem in the infinite problem.

Two intermediary models will be considered. The first is called the finite memory model and the second the full memory model. In both the unit interval is divided into n equal parts. The stopping process is altered so it can only stop on the best arrival in each interval after observing the entire interval. In the finite memory model the stopping decision is based solely on the ranks of these best arrivals relative to themselves. In the full memory model, the stopping decision is based upon the ranks of these best arrivals relative to all previous arrivals.

The notation used will be

(2.31) T_r = arrival time of best candidate among all candidates arriving in $\left[\frac{r-1}{n}, \frac{r}{n}\right]$

(2.32) Y_r = relative rank of individual arriving at T_r among those arriving at $T_1, T_2, \ldots, T_r$.

(2.33) Y_r^n = relative rank of individual arriving at T_r among all individuals arriving in $\left[0, \frac{r}{n}\right]$.

(2.34) V = minimal risk for infinite problem

V_n = minimal risk for finite problem

V_n^{**} = minimal risk for finite memory problem

V_n^{*} = minimal risk for full memory problem.

At first, we will consider only the truncated loss (2.25). That is, for all k and x, $A_k(x) \leq A_M(x)$ and $A_k(x) \leq M$. An essential property that is often used is equicontinuity.

We consider the four types of problems in pairs.

Prop. 2.2 For truncated loss $|V_n - V_n^{**}| \to 0$ as $n \to \infty$

Proof:

For the finite memory problem the decision to stop or go on at time T_r is based solely on $y_1, \ldots, y_r$. But it is easy to see that $y_1, \ldots, y_r$ are distributed exactly as the relative ranks in the finite problem. The difference is the time scale. If we let time T_r in the finite memory problem be equivalent to time r in the finite problem, then rules are interchangeable.

But, for the finite problem the loss for stopping at time r with $y_r = k$ is $A_k^{(n)}(r)$ while the finite memory problem loss is $A_{Y_r^n}(T_r) \geq A_k(T_r)$. Using equicontinuity and $\frac{r-1}{n} \leq T_r \leq \frac{r}{n}$, for each ϵ we can guarantee

$$V_n \leq V_n^{**} + \epsilon \quad \text{for n sufficiently large.}$$

Let τ_n be the optimal procedure for the finite problem and apply it to the finite memory problem. Ignoring the difference between $A_k^{(n)}(r)$ and $A_k(T_r)$ for the moment the two procedures give the same expected loss when

(i) $Y_{\tau_n}^n = y_{\tau_n}$ or

(ii) $y_{\tau_n} \geq M$ which implies $Y_{\tau_n}^n \geq M$.

Now considering the possible difference between $A_k^{(n)}(r)$ and $A_k(T_r)$ we get

$$V_n^{**} - V_n \leq M\ P(Y_k^{(n)} > y_k \text{ for some } k \text{ and } y_k < M) + \epsilon \text{ for } n \text{ sufficiently large.}$$

Gianini [6] showed $P(Y_k^{(n)} > y_k$ for some k and $y_k < M) \to 0$ so $-\epsilon \leq V_n - V_n^{**} \leq \epsilon$ for all n sufficiently large. Since ϵ was arbitrary we have proved the proposition.//

So, in the limit the finite problem and the finite memory problem have the same minimal risk. The same can be shown for the full memory problem and the infinite problem.

Prop. 2.3 $\quad V_n^* \to V$ as $n \to \infty$ for truncated loss.

Proof:

Let τ be the optimal procedure for the infinite problem. Consider the procedure on the full memory problem that stops on T_r precisely when τ stops in $\left[\frac{r-1}{n}, \frac{r}{n}\right]$. We immediately get

$$V_n^* \leq V.$$

Recalling Y_r^n it can be shown that $P(Y_r^n = k) = \frac{1}{r}\left(\frac{r-1}{r}\right)^{k-1}$ for $k = 1,2,\ldots$. Since only n decisions must be made backward induction can be used to solve the problem. Use this solution to induce an island rule on the infinite problem. That is, if you would stop at T_r on a candidate of relative rank k, then $\left[\frac{r-1}{n}, \frac{r}{n}\right] \subseteq I_k$. Let σ_n denote the policy on the full memory problem and τ_n the induced policy on the infinite problem. There are several ways in which the procedures can differ. One is to have the point of minimal risk different than the point of minimal rank in some $\left[\frac{r-1}{n}, \frac{r}{n}\right]$. That is, perhaps $T_r \neq U_{i_0}$ where $i_0 = \inf\{i \mid U_i \in \left[\frac{r-1}{n}, \frac{r}{n}\right]\}$. By equicontinuity, the difference is less than ϵ for n sufficiently large. So suppose

$T_r = U_{i_0}$, i_0 defined above. Then σ and τ_n always stop in the same interval. Then the risks can only differ if $Y_{\sigma_n} > Y_{\tau_n}$ and $\tau_n < M$. (Same reasoning as the previous proof.) But the probability $Y_{\sigma_n} > Y_{\tau_n}$ and $\tau_n < M$ goes to zero as in the previous theorem and our loss is bounded so

$$V \leq \underline{\lim}\, V_n^* + \epsilon.$$

But this holds for all $\epsilon > 0$ so the proposition is proved.//

One leg still remains. We must show the finite memory and the full memory problems have the same limit.

Prop. 2.4 For truncated loss, $V_n^* - V_n^{**} \to 0$ as $n \to \infty$.

Proof:

$V_n^* \leq V_n^{**}$ is obvious.

Let τ_n^* be the optimal procedure for the full memory problem. Induce a procedure σ_n^{**} on the finite memory problem from τ_n^*. That is, if τ_n^* would solicit at time r if a candidate had rank relative to all previous candidates $\leq k$, then σ_n^{**} would solicit at time r if a candidate has rank $\leq k$ relative to only $T_1, T_2, \ldots, T_r$. Obviously $\sigma_n^{**} \leq \tau_n^*$. Since the loss for soliciting the M^{th} relative best is equivalent to soliciting the $(M+1)^{st}$, $(M+2)^{nd}$, etc., the two procedures differ only when $\sigma^{**} < \tau_n^*$ and $y_{\sigma^{**}} < M$. Therefore,

$$V_n^{**} - V_n^* \leq E(A_{Y_{\sigma_n^{**}}}(\sigma_n^{**})) - V_n^*$$

$$\leq M\, P(\sigma^{**} < \tau_n^* \text{ and } y_{\sigma^{**}} < M)$$

$$\to 0 \qquad \text{as in propositions 2.2 and 2.3.}$$

This and the other inequality prove the proposition.//

We can now put the three propositions together to yield

Thm. 2.4 For truncated loss (2.25) $V_n \to V$ as $n \to \infty$.

Proof:

$$|V_n - V| = |V_n - V_n^{**} + V_n^{**} - V_n^* + V_n^* - V|$$

$$\leq |V_n - V_n^{**}| + |V_n^{**} - V_n^*| + |V_n^* - V| \to 0$$

by propositions 2.2, 2.4, and 2.3 respectively.//

Cor. 1 If $A_k(x)$ is given with $A_\infty(1) < \infty$, $V_n \to V$ as $n \to \infty$.

Proof:

Since $A_\infty(x)$ is continuous and increasing, $A_n(x) \to A_\infty(x)$ uniformly in x. Let $\epsilon > 0$ be given. Choose M_1 so $n > M$, implies $A_\infty(x) \geq A_n(x) \geq A_\infty(x) - \epsilon$ for all x. Choose M_2 so $A_\infty(1) < M_2$.

For $M \geq (M_1 \vee M_2)$, $A_k^M(x)$ as in (2.25) we have

$$A_k(x) - \epsilon \leq A_k^M(x) \leq A_k(x) \quad \text{for all k and x.}$$

Let V_n^M and V^M be the risks for the finite and infinite truncated problems. Then

$$V_n - \epsilon \leq V_n^M \leq V_n$$

$$V - \epsilon \leq V^M \leq V$$

$$\underline{\lim}\, V_n \geq \underline{\lim}\, V_n^M = V^M \geq V - \epsilon$$

$$\overline{\lim}\, V_n \leq \overline{\lim}\, V_n^M + \epsilon = V^M + \epsilon \leq V + \epsilon.$$

Since ϵ was arbitrary,

$$\underline{\lim}\, V_n = \overline{\lim}\, V_n = V.//$$

Cor. 2 For any $A_k(x)$, $\underline{\lim}\, V_n \geq V$.

Proof:

Let $A_k^M(x)$, V_n^M, and V^M be as in corollary 1. Clearly $V_n^M \leq V_n$ so $V^M = \underline{\lim}\, V_n^M \leq \underline{\lim}\, V_n$ for all M. Since $V^M \to V$ by theorem 2.3,

$$V \leq \underline{\lim}\, V_n.//$$

One gap still exists. Is it always true that $\overline{\lim}\, V_n \leq V$? For $A_k(x)$ decreasing in x for all k, the answer is trivially shown to be yes. For Chapter III, the interview cost problem, the answer is yes whenever the cost is bounded. But in total generality, the question still remains.

CHAPTER III

THE SECRETARY PROBLEM WITH SAMPLING COST

This chapter consists of three main sections. The first contains the exact solution to the finite secretary problem with a known sampling cost. Few details are given since the backward induction argument is essentially the same as Govindarajulu [9].

The second section contains the exact solution to the infinite secretary problem with known sampling cost. Again few details are given since this is merely an application of Chapter II. However, more detail is given to the relationships between the finite and infinite problems since some results were missing in Chapter II.

Finally, in the third section, the rank problem with fixed cost a per interview is considered. The behavior of the solution is analyzed quite closely.

1. Why a Sampling Cost?

A secretary problem with an additional sampling cost is certainly more realistic. If no cost per interview were considered, the hiring plan would be the same for both a nationwide and a local search. The cost for a nationwide search will certainly become prohibitive since the number of candidates will certainly be large. But significant gains in quality will probably not be obtained. Thus, we will want a rule that will take the cost per interview into consideration.

In this chapter we consider the problem in total generality, that is for arbitrary payoff functions $q(\cdot)$ and cumulative sampling cost $h(\cdot)$. The relative values of $q(\cdot)$ and $h(\cdot)$ determine the stopping rule.

Of course we add additional complexities into the form of the solution. Instead of the well-known cutoff point rules we get island rules. (Recall definition given in the previous chapter.) A simple example illustrates this point.

Let $n = 1000$, $q(i) = i$, $h(i) = \begin{cases} 0 & i = 1,2,\ldots, 50 \\ 100 & i = 51,52,\ldots, 1000. \end{cases}$
From the $51^{\underline{st}}$ candidate on, we can treat the problem as if there is no interview cost. Using the algorithm given in CMRS [3] we conclude no candidate is to be solicited prior to the $258^{\underline{th}}$ position. So in particular you should not solicit the $51^{\underline{st}}$ candidate.

The algorithm also states that the expected loss for stopping at time i on a candidate of relative rank j is $\left(\frac{n+1}{i+1}\right)$ j. So you should solicit the $50^{\underline{th}}$ candidate if her relative rank is less than or equal to 5. (Losing at most 98.51 which is certainly better than paying 100 for one more observation.) So under the optimal rule, you could solicit the $50^{\underline{th}}$ candidate but not the $51^{\underline{st}}$!

2. The Finite Problem

We will assume the following are known prior to the start of the interviewing:

A) n - the total number of applicants

B) $q(\cdot)$ - the payoff function

C) $h(\cdot)$ - the cumulative sampling cost.

Following CMRS [3] we let

(3.1) X_i = absolute rank of $i^{\underline{th}}$ candidate to appear

(3.2) Y_i = relative rank of $i^{\underline{th}}$ candidate to appear

(3.3) τ = a stopping rule on [1,2,..., n] which is

1) a function of the $Y_i^{\underline{s}}$

2) can stop only on the present candidate

3) must take the $n^{\underline{th}}$ candidate if $\tau > n-1$.

(3.4) $$V_n = \inf_{\tau} E(q(X_\tau) + h(\tau)) \quad \text{the optimal expected loss}$$

(3.5) $$C^n(r) = \inf_{\tau>r} E(q(X_\tau) + h(\tau)) \quad \text{for } r = 0,1,2,\ldots, n-1.$$

(3.6) $$Q_k^n(r) = E(q(X_r)|Y_r = k) = \sum_{\ell=k}^{n-(r-k)} q(\ell) \frac{\binom{\ell-1}{k-1}\binom{n-\ell}{r-k}}{\binom{n}{r}}$$

= expected loss for stopping at time r with candidate of relative rank k.

Backward induction yields the optimal solution easily:

(3.7) $$C^n(n-1) = \frac{1}{n} \sum_{k=1}^{n} [q(k) + h(n)] \quad ,$$

(3.8) $$C^n(r-1) = \frac{1}{r} \sum_{k=1}^{r} \min(Q_k^n(r) + h(r), C^n(r)).$$

The equation $V_n = C^n(0)$ gives the optimal risk and the optimal rule is to stop with the $r^{\underline{th}}$ candidate if her relative rank k is such that $Q_k^n(r) + h(r) \leq C^n(r)$. (Provided you have not previously stopped.)

As was previously stated, the general solution is an "island rule." For simplicity define island rule as follows:

<u>Def</u>. An island rule consists of the following:

1) $I_k \subseteq [1,2,\ldots, n]$ with $I_{k+1} \subseteq I_k$.

2) Stop on the first arrival of a candidate of relative rank k in I_k.

Then clearly the optimal rule is an island rule with

$$I_k = \{r \mid Q_k^n(r) \leq C^n(r) - h(r)\}. \tag{3.9}$$

3. The Infinite Problem

Heuristically speaking we can get the solution to the infinite problem from the finite problem. The now familiar argument goes as follows:

Normalize [1,2,..., n] to [0,1] and write (3.8) as a difference equation. Letting $n \to \infty$ we get a differential equation which is "the solution" to the infinite secretary problem. In brief, the details are

$$C^n\left(\frac{r}{n}\right) - C^n\left(\frac{r-1}{n}\right) = \frac{1}{r} \sum_{k=1}^{r} \left[C^n\left(\frac{r}{n}\right) - Q_k^n\left(\frac{r}{n}\right) - h\left(\frac{r}{n}\right)\right]^+ \tag{3.10}$$

Letting $\frac{r}{n} \to x$ we have

$$C^n\left(\frac{r}{n}\right) \to f(x) \tag{3.11}$$

$$Q_k^n\left(\frac{r}{n}\right) \to R_k(x) = \sum_{\ell=k}^{\infty} q(\ell) \binom{\ell-1}{k-1} x^k (1-x)^{\ell-k} \tag{3.12}$$

$$h\left(\frac{r}{n}\right) \to h(x). \tag{3.13}$$

Dividing (3.10) by $\frac{1}{n}$ and letting $n \to \infty$ we get

$$f'(x) = \frac{1}{x} \sum_{k=1}^{\infty} [f(x) - R_k(x) - h(x)]^+ \tag{3.14}$$

(For no interview cost h=0 and we get the differential equation obtained by Mucci [14].)

It is through this line of heuristics our section differs significantly from Govindarajulu's [9]. Govindarajulu considered the case $q(i) = i$ and made approximations to get a different set of difference equations. Then, letting these tend to a differential equation, a limiting risk was estimated. No attempt at a limiting solution was made.

Unfortunately, even for the zero interview cost problem, his methods are too crude. For two different estimations, limiting solutions of 2 and 4 are given. CMRS showed the true limiting solution is $\prod_{j=1}^{\infty} \left(\frac{j+2}{j}\right)^{\frac{1}{j+1}} \approx 3.87$. Incidently, this is the quantity given by (3.14) with $h \equiv 0$.

The differential equation (3.14) also gives the optimal procedure. The procedure is to stop at time x on a candidate of relative rank k provided $R_k(x) + h(x) \leq f(x)$ and you have not previously stopped. If 1 is reached, pay $\sup_k [R_k(1)] + h(1)$.

That both the optimal risk and the optimal procedure are given by (3.14) is a consequence of Chapter II with

(3.15) $$A_k(x) = R_k(x) + h(x).$$

It is only necessary to verify (2.8) through (2.11) for $A_k(x)$ defined by (3.15).

We only need to assume

(3.16) $h(x)$ is continuous in x and

(3.17) $h(x) < \infty$ on $[0,1)$.

Since Gianini and Samuels [7] have established (2.8) through (2.11) for $R_k(x)$, it follows trivially that

$A_k(x)$ is continuous for all k (verifying (2.8)),

$A_k(x) < \infty$ on $(b_k, 1)$ where b_k is defined by the same relationship for $R_k(x)$ (verifying (2.9)),

$A_1(0^+) = A_2(0^+) = \ldots = R_1(0^+) + h(0)$ (verifying (2.10)), and
$A_\infty(x) = R_\infty(x) + h(x) = q(\infty) + h(x)$ which is non-decreasing (verifying (2.11)).

So the results of Chapter II can be applied here and we have proved that if $V < \infty$ (recall Chapter II) then (3.14)

$$f'(x) = \frac{1}{x} \sum_{k=1}^{\infty} [f(x) - R_k(x) - h(x)]^+ ,$$

(3.18) $$\inf_\tau E(R_{Y_\tau}(\tau) + h(\tau)) = f(0), \text{ and}$$

(3.19) $$E(R_{Y_{\tau^*}}(\tau^*) + h(\tau^*)) = f(0) \text{ where } \tau^* \text{ is the island rule}$$

(recall Chapter II) given by

(3.20) $$I_k = \{x \mid R_k(x) + h(x) \leq f(x)\}$$

But we can go further into the form of the island rules in the application considered in this chapter. One characteristic is that τ^* stops earlier than the corresponding procedure for the zero interview cost problem.

To briefly review the optimal procedure for the zero interview cost problem let

(3.21) $$g'(x) = \frac{1}{x} \sum_{k=1}^{\infty} [g(x) - R_k(x)]^+ .$$

Then the optimal policy is given by the single island rule (recall Chapter II) $I_k = [\alpha_k, 1]$ with α_k uniquely satisfying

(3.22) $$R_k(\alpha_k) = g(\alpha_k).$$

Formally we have

<u>Prop. 1</u> If $V < \infty$, $h(t) < \infty$ on $[0,1)$ then for I_k defined by (3.20) and α_k defined by (3.22) we have $[\alpha_k, 1] \subseteq I_k$ for all k.

Proof:

Let $x \geq \alpha_k$. Then recalling the definition of f and g,

$$f(x) = \inf_{\tau > x} E(R_{Y_\tau}(\tau) + h(\tau))$$

$$\geq \inf_{\tau > x} E(R_{Y_\tau}(\tau)) + h(x)$$

$$= g(x) + h(x)$$

$$\geq R_k(x) + h(x) \quad \text{since } x \geq \alpha_k.$$

Therefore $x \in I_k$ and our proposition is proved.//

The easiest rules to work with are single island rules. This is because a solution for one cutoff point uniquely gives the next cutoff point. We will derive sufficient conditions on R_k and h to guarantee single island rules. Specifically these will be used later to solve a wide class of problems.

Prop. 2 Let $f(0) < R_1(0) + h(0)$, $q(k) < q(\infty)$, h be differentiable, and let

(3.23) $$W_k(x) = \frac{1}{x} \sum_{j=1}^{k} (R_{k+1}(x) - R_j(x)) - h'(x).$$

If, for each k, $W_k(x)$ has at most one sign change; from + to -, then the optimal rule is a single island rule.

Proof:

Define $g_k(x) = f(x) - R_k(x) - h(x)$. We immediately conclude g_k is continuous, g_k is differentiable, $g_k(0) < 0$, and $g_k(1) = R_\infty(1) - R_k(1) > 0$. Let $I = [0,1]$, $I_k = \{x | g_k(x) \geq 0\}$, and $x_0 = \sup(I - I_k)$. By continuity $g_k(x_0) = 0$ while by definition $g_k(x) \geq 0 \; \forall \; x > x_0$ and $g(y) < 0$ for some $y < x_0$. Therefore $g_k'(x_0) \geq 0$. Assume $x_1 < x_0$ exists so that $g_k(x_1) > 0$. Then by the mean value theorem we can find

$y_1 \in [x_1, x_0] \cap I_k$ satisfying $g_k'(y_1) < 0$. We will show that this inequality contradicts our hypothesis.

Let $z \in I_k$. Then, since $R_k'(z) = \frac{k}{z}(R_k(z) - R_{k+1}(z))$,

$$\begin{aligned} g_x'(z) &= f'(z) - R_k'(z) - h'(z) \\ &= \frac{1}{z}\sum_{j=1}^{\infty}[f(z) - R_j(z) - h(z)]^+ + \frac{k}{z}(R_{k+1}(z) - R_k(z)) - h'(z) \\ &\geq \frac{1}{z}\sum_{j=1}^{k}[f(z) - R_j(z) - h(z)] + \frac{k}{z}(R_{k+1}(z) - R_k(z)) - h'(z) \\ &\geq \frac{1}{z}\sum_{j=1}^{k}[R_k(z) + h(z) - R_j(z) - h(z)] + \frac{k}{z}(R_{k+1}(z) - R_k(z)) - h'(z) \\ &= W_k(z). \end{aligned}$$

At x_0, we can show equality so $g_k'(x_0) \geq 0$ implies $W_k(x_0) \geq 0$. But $W_k(y_1) \leq g_k'(y_1) < 0$ and $y_1 < x_0$ so $W_k(x)$ must have at least two sign changes and we have contradicted our hypothesis. So we can conclude that there exists β_k so that

$$(3.24) \qquad \begin{aligned} f(x) &\geq R_k(x) + h(x) \quad \text{for } x \in [\beta_k, 1] \\ f(x) &\leq R_k(x) + h(x) \quad \text{for } x \in [0, \beta_k]. \end{aligned}$$

Since it doesn't matter what we do at x if $R_k(x) + h(x) = f(x)$, we can conclude $I_k = [\beta_k, 1]$ for all k and our proposition is proved.//

In a later section we consider the solution to differential equation (3.14) for a large class of $q^{\underline{s}}$ and $h^{\underline{s}}$. For this class we will use proposition 2 to show the optimal rules are single island rules and derive a set of recurrsion equations. These equations are solved to give both the minimal risk and the optimal procedure.

We diverge momentarily to develop the conditions for which we can conclude $V < \infty$. For these conditions, we refer to the zero interview cost problem and use conditions developed by Gianini [6] and Gianini and Samuels [7]. The known results are summed up in the following proposition.

<u>Prop. 3</u> $V = \infty$ if A) $\sum_{k=1}^{\infty} \frac{\ln q(k)}{k^2} = \infty$.

$V < \infty$ if B) $h(x) \leq B < \infty$ and $\sum \frac{p_k R_k(x_k)}{k} < \infty$

for choice of $0 \leq x_1 \leq \ldots \leq 1$ and $p_k = \prod_{j=1}^{k-1} \left(\frac{x_j}{x_{j+1}} \right)^j$

or C) $q(\cdot)$ grows like a polynomial.

<u>Proof</u>:

A) is immediate since Gianini showed this leads to infinite risk in the zero interview cost problem.

B) is also immediate by a proof in Gianini.

C) follows from a result in Gianini and Samuels no matter how fast h grows!

Gianini and Samuels showed that for zero interview cost and $q(\cdot)$ growing like a polynomial (i.e., $q(k+1)/q(k) = 1 + 0\left(\frac{1}{k}\right)$) and for any T ($0 \leq T \leq 1$), there exists $\tau_T \leq T$ with $E(q(x_{\tau_T})) < \infty$. Applying the same τ_T to the problem with interview cost gives $V \leq E(q(x_{\tau_T})) + h(T)$ which is finite for T chosen so that $h(T) < \infty$. //

There is still a gap. Does there exist a q and h with

$$\inf_{\tau} E(q(x_\tau)) < \infty \text{ and}$$

$$\inf_{\tau} E(q(x_\tau) + h(\tau)) = \infty \text{ ?}$$

That is, if there is a finite solution with no interview cost is there always a finite solution no matter how fast the interview cost grows?

4. The Infinite Problem as the Limit of the Finite Problem

In the heuristics of the earlier part of this section we normalized the interval [0,n] to [0,1] and normalized the appropriate sampling cost function h. Thus, we are essentially letting h be defined on [0,1] and taking $h^n(i) = h\left(\frac{i}{n}\right)$, for i = 1,2,..., n. This may seem unrealistic. So let us consider an arbitrary sequence

(3.25) $$0 \leq h(1) \leq h(2) \leq \ldots$$

(3.26) $$h(\infty) = \lim h(n) \leq \infty.$$

Then take $h^n(i) = h(i)$ for i = 1,2,..., n. Of course $q(\cdot)$ is defined in the usual way. Let

(3.27) $$q(\infty) = \lim q(n) \leq \infty .$$

It is shown that for h defined by (3.25) $\lim_{n\to\infty} V_n$ (where V_n is the optimal risk for the finite n-girl problem) is always trite.

Thm. 3.1 If $V = \inf_{\tau} E(q(x_\tau))$ is the minimal risk for the infinite problem with zero interview cost, then

(3.28) $$\lim_{n\to\infty} V_n = \min(V + h(\infty),\ q(\infty) + h(1)).$$

That is, in the limit, either take the first candidate and obtain a risk of $q(\infty) + h(1)$ or ignore the interview cost and obtain a risk of $V + h(\infty)$.

Proof:

The inequality $\lim_{n\to\infty} V_n \leq \min(V + h(\infty),\ q(\infty) + h(1)$ follows from the remark following the theorem.

Recall (3.6). It is easy to show

(3.29) $$\lim_{n\to\infty} Q_k^n(r) = q(\infty) \quad \text{for any fixed k and r.}$$

Let V_n, τ_n be the optimal risk and rule for the n-girl problem. In addition, let V_n^*, τ_n^* be the optimal risk and rule for the n-girl zero interview cost problem. Then, we get

$$V_n = E(q(x_{\tau_n}) + h(\tau_n))$$

$$\geq [Q_1^n(N) + h(1)]\ P(\tau_n < N) + E(q(x_{\tau_n}) + h(\tau_n)) I_{(\tau_n \geq N)}$$

$$\geq [Q_1^n(N) + h(1)]\ P(\tau_n < N) + E(q(x_{\tau_n})) I_{(\tau_n \geq N)} + h(N)\ P(\tau_n \geq N)$$

$$\geq [Q_1^n(N) + h(1)]\ P(\tau_n < N) + E(q(x_{\tau^*_n})) I_{(\tau_n \geq N)} + h(N)\ P(\tau_n \geq N)$$

$$= [Q_1^n(N) + h(1)]\ P(\tau_n < N) + [V_n^* + h(N)]\ P(\tau_n \geq N)$$

$$\geq \min([Q_1^n(N) + h(1)],\ [V_n^* + h(N)])$$

for n so big that $\tau_n^* \geq N$.

Letting $n \to \infty$ and applying (3.29), we have $\lim V_n \geq \min([q(\infty) + h(1)], [V + h(N)])$. Since this holds for all N, and $h(N) \to h(\infty)$, $\lim V_n \geq \min(V + h(\infty),\ q(\infty) + h(1))$. //

We may still be interested in the case $h(\infty) = \infty$, $q(\infty) = \infty$, the only interesting case. We know from the above theorem that $V_n \to \infty$ but not the rate. No clues are given by the above arguments. A general method is to approximate the n-girl problem with the infinite problem having appropriately chosen interview cost h^n. Then let h^n vary and consider the corresponding solutions. This should give a good approximation to the true rate.

An example is given in a later section with $h(i) = ai$ and $q(i) = i$ for some positive constant a. (Linear cost rank problem.)

So, without loss of generality, let us consider the loss structure

(3.30) $$h^n(i) = h\left(\frac{i}{n}\right), \quad h:[0,1] \to [0,\infty],\ h \uparrow .$$

We are interested in showing $V_n \to V$. That is, we want the infinite problem to be the limit of the finite problems.

Using (3.6) and (3.12) it is easy to show

$$\sup_{\substack{1 \leq i \leq n \\ 1 \leq k \leq i}} \left| Q_k^n(i) - R_k\left(\frac{i}{n}\right) \right| \to 0 \text{ as } n \to \infty.$$

So, invoking corollary 2 of theorem 2.4, we conclude

(3.31) $\qquad \underline{\lim}\, V_n \geq V.$

We must obtain the other inequality. For this, let us assume h is uniformly continuous on [0,1]. Undoubtedly this condition can be weakened near 1 since the process has only a remote chance of actually being near 1; but the relationship is in general too complex to work with.

Prop. 4 If $V < \infty$ and h is uniformly continuous on [0,1], then $V_n \to V$.

Proof.

Let $\epsilon > 0$ be given.

Choose k so that $h(1) \cdot P(\sup \Phi \geq k) < \epsilon/5$ for Φ distributed as Brownian Bridge. Choose N so large that $|x - y| \leq \frac{1}{N} \Rightarrow |h(x) - h(y)| \leq \epsilon/5$ and $|x - y| \leq \frac{2k}{\sqrt{N}} \Rightarrow |h(x) - h(y)| \leq \epsilon/5$. Consider 3 processes, labelled by a *, ', and ". The star process is a finite problem with $h^n(i) = h\left(\frac{i}{n+1}\right)$. For $n > N$, $V_n \leq V_n^* + \sup_i \left| h\left(\frac{i}{n}\right) - h\left(\frac{i}{n+1}\right) \right| \leq V_n^* + \epsilon/5$. The prime process is an infinite problem where it is known if a candidate is one of the n best or not. Clearly $V_n' \leq V$. The " problem is exactly the ' problem except you must stop on one of the n best candidates, V_n'' being the corresponding minimal risk.

Let y_i be the (random) arrival time of the i^{th} candidate in processes ' and " (y_i being one of the n best). It follows that $V_n'' \geq V_n^* - \left(\max_{1 \leq i \leq n} \left| h(y_i) - h\left(\frac{i}{n+1}\right) \right| \right)$. But $E(y_i) = \frac{i}{n+1}$ and by applying the Kolmogorov-Smirnov statistic to the uniform [0,1] distribution (recall that a given candidate arrives uniformly on [0,1]) we get for $n \geq N$, $\sup|y_i - \frac{i}{n+1}| \leq \frac{k}{\sqrt{N}}$ with probability $\geq 1-P$ where P satisfies $h(1) \cdot P \leq \epsilon/5$! So for $n > N$, $V''_n \geq V^*_n - \epsilon/5 - h(1) \cdot P$

$$\geq V^*_n - 2\,\epsilon/5 .$$

And by the same reasoning, $\sup\limits_{i=1,\ldots,n-1} (y_{i+1} - y_i) \leq \frac{2k}{\sqrt{N}}$ with probability 1-P so $V'_n \geq V''_n - \frac{2\epsilon}{5}$ for $n > N$. (The difference being the time until the next candidate (out of best n) arrives.) Putting all these inequalities together, for $n > N$ we have

$$\begin{aligned} V_n &\leq V^*_n + \epsilon/5 \\ &\leq V''_n + 3\,\epsilon/5 \\ &\leq V'_n + \epsilon \\ &\leq V + \epsilon . \end{aligned}$$

Since this holds for all ϵ, $V_n \to V$.//

Cor. 1 If $V < \infty$, h is uniformly continuous on [0,1], and $h^n(i)$ can be written as $h(\alpha(i,n))$ where $\sup\limits_{x \in [0,1]} |\alpha([nx],n) - x| \to 0$ as $n \to \infty$, then $V_n \to V$. (According to the hypothesis, if $\frac{i}{n} \to x$ implies $\alpha(i,n) \to x$ uniformly in x, then $V_n \to V$.)

Proof:

$\sup\limits_{1 \leq i \leq n} |\alpha(i,n) - \frac{i}{n+1}| \to 0$ so above proof goes through verbatum.//

In particular, $h^n(i) = h\left(\frac{i-1}{n}\right)$ satisfies the conditions in the corollary. This turns out to be helpful in comparing the finite and infinite linear cost problem.

5. Solving the Differential Equation

Suppose the optimal procedure is a single island rule and h is differentiable. Then on $[\beta_M, \beta_{M+1})$ (recall I_k is of the form

$[\beta_k, 1])$ our differential equation becomes

$$f'(x) = \frac{1}{x} \sum_{k=1}^{M} [f(x) - R_k(x) - h(x)].$$

This can be rewritten as

$$\left(\frac{f}{X^M}\right)' = \frac{1}{M-1}\left(\frac{1}{X^M}\sum_{k=1}^{M-1}(R_k(x) + h(x))\right)' - \frac{h'(x)}{X^M}$$

Integrating from β_M to β_{M+1}, substituting $R_M(\beta_M) + h(\beta_M)$ for $f(\beta_M)$, and $R_{M+1}(\beta_{M+1}) + h(\beta_{M+1})$ for $f(\beta_{M+1})$ we obtain the following recurrsive equation:

(3.32)
$$\frac{1}{\beta_{M+1}^M}\sum_{k=1}^{M-1}[R_{M+1}(\beta_{M+1}) - R_k(\beta_{M+1})] = \frac{1}{\beta_M^M}\sum_{k=1}^{M-1}[R_M(\beta_M) - R_k(\beta_M)] - (M-1)\int_{\beta_M}^{\beta_{M+1}} \frac{h'(y)}{y^M}\,dy$$

(Note: for general island rules, the integration is performed over $I_M - I_{M+1}$. Of course $I_M - I_{M+1}$ can be written as a discrete set of intervals a.s. but the problem is still complex computationally.)

Let us consider a class of $q(\cdot)^{\underline{s}}$ and $h(\cdot)^{s}$ defined by

(3.33)
$$q_\xi(k) = \frac{\Gamma(\xi + k - 1)}{\Gamma(\xi)\,\Gamma(k)} \qquad \xi \geq 2$$

(3.34)
$$h_m(x) = \sum_{i=1}^{m} a_i x^i \qquad a_i \geq 0.$$

In particular, when $\xi = 2$, $m = 1$, we are considering the rank problem with linear cost. This particular problem is considered in some detail in the final section of this chapter.

We will use proposition 2 to verify the optimality of single-island rules.

Mucci [14] has shown

(3.35)
$$R_k(x) = q_\xi(k)\, x^{1-\xi}.$$

Therefore $W_k(x)$ becomes

$$W_k(x) = x^{-\xi}\left(\sum_{j=1}^{k}[q_\xi(k+1) - q_\xi(j)]\right) - \sum_{i=1}^{M} i\ a_i\ x^{i-1}.$$

This equation has the same sign as

$$\hat{W}_k(x) = \sum_{j=1}^{k}\ [q_\xi(k+1) - q_\xi(j)] - \sum_{i=1}^{M} i\ a_i\ x^{i+\xi-1}$$

for $x \in [0,1]$.

Now, $\hat{W}_k(0) > 0$ and is differentiable so if $\hat{W}_k(x)$ changes sign more than once in $(0,1)$, $\hat{W}_k(x)$ has a local minimum in $(0,1)$. We know that no local minimum exist since

$$\hat{W}_k'(x) = -\sum_{i=1}^{M} i(i+\xi-1)\ a_i\ x^{i+\xi-2} < 0 \qquad x \in (0,1),\ \xi \geq 2.$$

(At a local minimum the derivative has value 0.)

So we conclude that $W_k(x)$ changes sign at most once, and from + to -. We also know $f(0) < \infty$ and $R_k(0) = \infty$ so all conditions in proposition 2 are met. We have thus proven that the optimal rules are single island rules.

If we rewrite the recurrsion equation (3.32) as

$$\beta_M^{\xi+M-1} = \left[\frac{\left(\frac{\xi-1}{\xi}\right)(M-1)\ q_\xi(M) - \sum_1^m i\ q_i\left(\frac{M-1}{M-i}\right)\beta_M^{\xi+i-1}}{\frac{M+\xi}{M}\left(\frac{\xi-1}{\xi}\right)(M-1)\ q_\xi(M) - \sum_1^m i\ q_i\left(\frac{M-1}{M-i}\right)\beta_{M+1}^{\xi+i-1}}\right]\beta_{M+1}^{\xi+M-1}, \quad (M > m)$$

and use the fact that $\beta_M \to 1$ as $M \to \infty$, we can get

$$\text{(3.36)} \qquad \beta_M \geq \prod_{\ell=M}^{\infty}\left(\frac{\ell}{\ell+\xi} - \sum_1^m \frac{i\ q_i\ \ell\ \xi}{(\ell-i)(\ell+\xi)(\xi-1)\ q_\xi(\ell)}\right)^{1/\xi+\ell-1}, \quad (M > m).$$

If we let α_M be the cutoff point for the zero interview cost problem and use a result of Mucci, we get

$$\text{(3.37)} \qquad \beta_M \leq \alpha_M = \prod_{\ell=M}^{\infty}\left(\frac{\ell}{\ell+\xi}\right)^{1/\xi+\ell-1}.$$

We have upper and lower bounds for β_M that squeeze together as $M\to\infty$. We make an approximation on β_M and use (3.32) to successively get approximations on β_{M-1}, $\beta_{M-2},\ldots,\beta_1$. When the upper and lower approximations on β_1 are sufficiently close to each other, we can find $V=f(0)=f(\beta_1)=R_1(\beta_1)+h(\beta_1)$, the minimal limiting risk.

We can then solve the recurrsive equation forward to obtain the optimal procedure. As an example, the linear cost rank problem is solved in the next section.

6. Linear Cost Rank Problem

Consider now the rank problem $(q(i)=i)$ with linear interview cost $h_a(i) = ai$ for fixed $a > 0$. Theorem 3.1 tells us $V_n\to\infty$ but no more. Since this is probably the most realistic situation, we will look carefully into the solution. Various questions that arise are: When should you solicit the first candidate? If you don't solicit the first, then what? How fast does V_n tend to infinity?

These questions and others will be answered in the final pages of this section.

The first question is simply answered.

Prop. 5 If $a>\frac{n+1}{6}$, solicit the first candidate.

That is $\tau\equiv 1$.

(Note: An obvious inequality is obtained by considering when the cost for interviewing one more candidate is prohibitive. This leads to the inequality $a+\frac{n+1}{2} < 2a+1$ or $a > \frac{n-1}{2}$. But the above inequality, a slight improvement, will be useful later.)

Proof:

Let S_i be the minimal acceptable rank at time i. Consider the only three cases $S_2 = 0$, $S_2 = 1$, and $S_2 = 2$.

A) $S_2 = 2$. Then $C^n(1) = \frac{n+1}{2} + 2a$ so

$$S_1 = \left[\frac{2}{n+1}\,(C^n(1) - h(1))\right] = \left[1 + \frac{2a}{n+1}\right] \geq 1.$$

B) $S_2 = 0$. Then $C^n(1) = C^n(2) \geq \min\left(\frac{n+1}{4} + 3a,\ 1 + 4a\right)$

$$\geq \frac{n+1}{3} + 2a$$

So $S_1 = \left[\frac{2}{n+1}\,(C^n(1) - h(1))\right] \geq \left[\frac{2}{n+1}\left(\frac{n+1}{3} + a\right)\right] \geq \left[\frac{2}{3} + \frac{2a}{n+1}\right] \geq 1.$

C) $S_2 = 1$. Then $C^n(1) = \frac{1}{2}\left(\frac{n+1}{3} + 2a\right) + \frac{1}{2}\,C^n(2) \geq \frac{n+1}{3} + 2a.$

As in B), $S_1 \geq 1$.

$\therefore$ $S_1 = 1$, that is stop immediately.//

A partial converse to this theorem we have

Prop. 6 If $a < (.082)(n+1)$, then $S_1 = 0$.

That is, do not solicit the first candidate.

Proof:

Consider the procedure τ described by $S_i = i-1$.

We have $C_0 \leq E(q(x_\tau) + h(\tau))$

$$= \sum_{i=2}^{n-1} \left[\left(\frac{n+1}{i+1}\right)\frac{(i-1)}{2} + a(i-1)\right]\frac{1}{(i-1)!} + \left[\frac{n+1}{2} + na\right]\frac{1}{(n-1)!}$$

Now $S_1 = 0$ if $C_0 \leq \frac{n+1}{2} + a$ = risk for stopping immediately.

A sufficient condition is to have

$$E(q(x_\tau) + h(\tau)) < \frac{n+1}{2} + a$$

which is equivalent to

$$a < \frac{\frac{n+1}{2}\left(\frac{2}{3} - \sum_{i=3}^{n} \frac{1}{(i+1)(i-2)!}\right) - \frac{1}{(n-1)!}}{\sum_{i=3}^{n+1} \frac{1}{(i-2)!}}$$

$\approx .081977\ (n+1) \quad$ for $n \geq 8$.

$= .056\ (n+1) \quad n=3$

$= .075\ (n+1) \quad n=4$

$= .080\ (n+1) \quad n=5$

$= .0817\ (n+1) \quad n=6$

$= .0819\ (n+1) \quad n=7$.

But for $n=2,3,\ldots, 7$, we can check the proposition.//

Now, if you do not stop on the first candidate, it is conceivable that the cost will become so great that you will stop at a certain later time regardless of the candidate's relative rank. That is, conceivably $\tau \leq k < n$ a.s. could hold for some k. The next proposition proves this is not the case.

Prop. 7 If $\tau > 1$, then $P(\tau=n) > 0$, where τ is the optimal procedure. So we have a dichotomy on τ, either stop immediately or you may not stop until the last candidate.

Proof:

$\tau > 1$ implies $S_1 = 0$ which by proposition 5 implies $a < \frac{n+1}{6}$.

It suffices to show $S_i < i$ for all $i < n$.

Clearly $C^n(i) \leq \frac{n+1}{2} + a(i+1)$ so

$$S_i = \left[\frac{i+1}{n+1}\,(C^n(i) - h(i))\right] \leq \left[\frac{i+1}{n+1}\left(\frac{n+1}{2} + a(i+1) - a(i)\right)\right]$$
$$= \left[\frac{i+1}{2} + \frac{a(i+1)}{n+1}\right].$$

But for $i \geq 2$,

$$a < \frac{n+1}{6} = \frac{n+1}{2} - \frac{n+1}{3} \leq \frac{n+1}{2} - \frac{n+1}{i+1} = \frac{(n+1)(i-1)}{2(i+1)}$$

So $\frac{i+1}{2} + \frac{a(i+1)}{n+1} < i \Rightarrow S_i \leq i-1$ for $i \geq 2$.//

Let us now compare the finite and infinite problems computationally. Let

(3.38) $h(x) = Kx.$

The finite problem was solved directly using backward induction and equation (3.32) was used to solve the infinite problem. Since $\prod_{\ell=1}^{n}\left(\frac{\ell}{\ell+2}\right)^{\frac{1}{\ell+1}} - \prod_{\ell=1}^{\infty}\left(\frac{\ell}{\ell+2}\right)^{\frac{1}{\ell+1}} = 0\left(\frac{1}{n}\right)$ it is hard to get good computational accuracy on the bounds for β_M by directly taking the product. Therefore, the Euler Maclaurin sum formula was used to approximate the log of the infinite product (14 terms of the sum formula give at least 10 place accuracy). The interval halving method was used to solve the iterative equation. Computations were made on the CDC 6500 at Purdue University. The comparisons are summarized in the following chart:

Table 3.1 Minimal risks for the linear cost rank problem

K\n	10	100	1000	10000	∞
0	2.5579	3.6032	3.8324	3.8649	3.8695
1	3.1415	4.1189	4.3320	4.3619	4.3661
5	4.9110	5.9592	6.1505	6.1641	6.1674
10	6.5000*	7.8553	8.0489	8.0716	8.0745
100	15.5000*	23.0564	24.6013	24.7520	24.7688
500	55.5000*	46.0512	54.5161	55.2973	55.3845
1000	105.5000*	60.5000*	76.5554	78.1496	78.3255
5000	505.5000*	100.5000*	166.0788	174.2509	175.1412
10000	1005.5000*	150.5000*	229.4733	245.9021	247.6871

* Select first candidate.

To get a visual picture of the rate of convergence of the finite problem to the infinite problem several computer graphs were constructed. the following two illustrate the convergence for K = 100 and K = 1000.

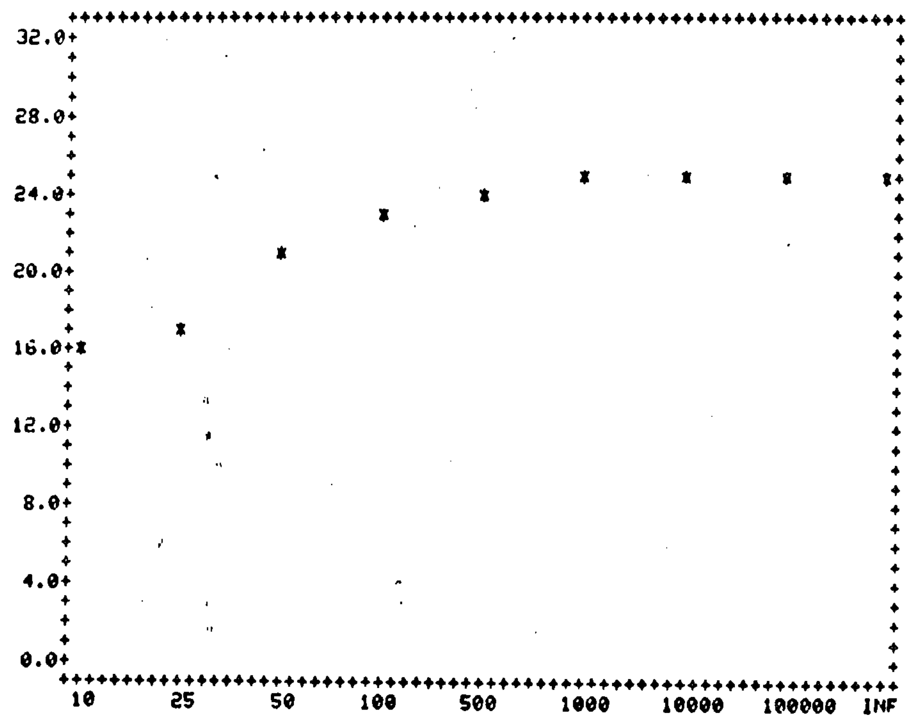
Please notice the bottom scale, chosen to show the actual rate of convergence. In the two graphs the cost per interview is 100/N and 1000/N respectively.

THE SECRETARY PROBLEM

WHERE THE TOTAL COST FOR OBSERVING ALL N SECRETARIES IS 100

COST PER INTERVIEW IS THEREFORE 100/N
WHERE N=NUMBER OF APPLICANTS.

GRAPH OF EXPECTED RISK VS. N




```
THE MINIMAL RISK FOR N=10 IS           15.50000
THE MINIMAL RISK FOR N=25 IS           17.00000
THE MINIMAL RISK FOR N=50 IS           21.35323
THE MINIMAL RISK FOR N=100 IS          23.05640
THE MINIMAL RISK FOR N=500 IS          24.43248
THE MINIMAL RISK FOR N=1000 IS         24.60135
THE MINIMAL RISK FOR N=10000 IS        24.75201
THE MINIMAL RISK FOR N=100000 IS       24.76709
THE MINIMAL RISK FOR N=INFINITY IS     24.76875
```

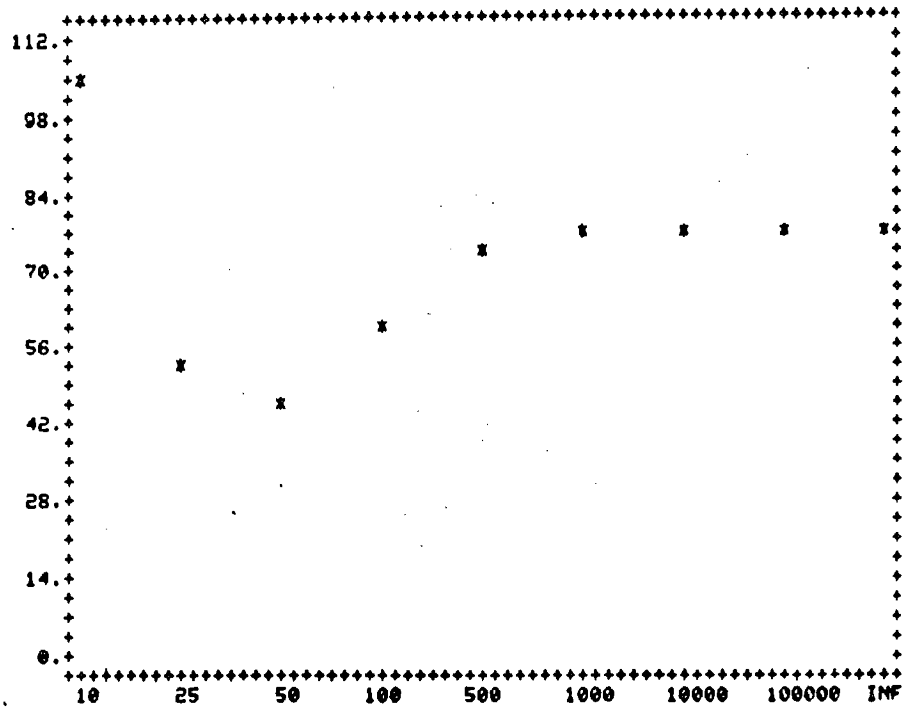
Figure 3.1 Rank problem, linear cost, total cost = 100.

THE SECRETARY PROBLEM

WHERE THE TOTAL COST FOR OBSERVING ALL N SECRETARIES IS 1000

COST PER INTERVIEW IS THEREFORE 1000/N
WHERE N=NUMBER OF APPLICANTS.

GRAPH OF EXPECTED RISK VS. N



THE MINIMAL RISK FOR N=10 IS 105.50000
THE MINIMAL RISK FOR N=25 IS 53.00000
THE MINIMAL RISK FOR N=50 IS 45.50000
THE MINIMAL RISK FOR N=100 IS 60.50000
THE MINIMAL RISK FOR N=500 IS 74.76363
THE MINIMAL RISK FOR N=1000 IS 76.55537
THE MINIMAL RISK FOR N=10000 IS 78.14963
THE MINIMAL RISK FOR N=100000 IS 78.30794
THE MINIMAL RISK FOR N=INFINITY IS 78.32553

Figure 3.2 Rank problem, linear cost, total cost = 1000.

Notice the strange looking dip in the second plot (K = 1000). The first four points are a result of stopping immediately. $\left(a > \frac{n+1}{6}\right)$ Because the cost per interview (1000/N) is so high, the minimal risk must be high. To alleviate this situation, it has been suggested we let $h^n(i) = h\left(\frac{i-1}{n}\right)$. Recalling corollary 1 to proposition 4 we know than our new V_n still tends to V. As a matter of fact, the difference in the V_n^s is exactly $\frac{K}{n}$, $\frac{1000}{n}$ for this case.

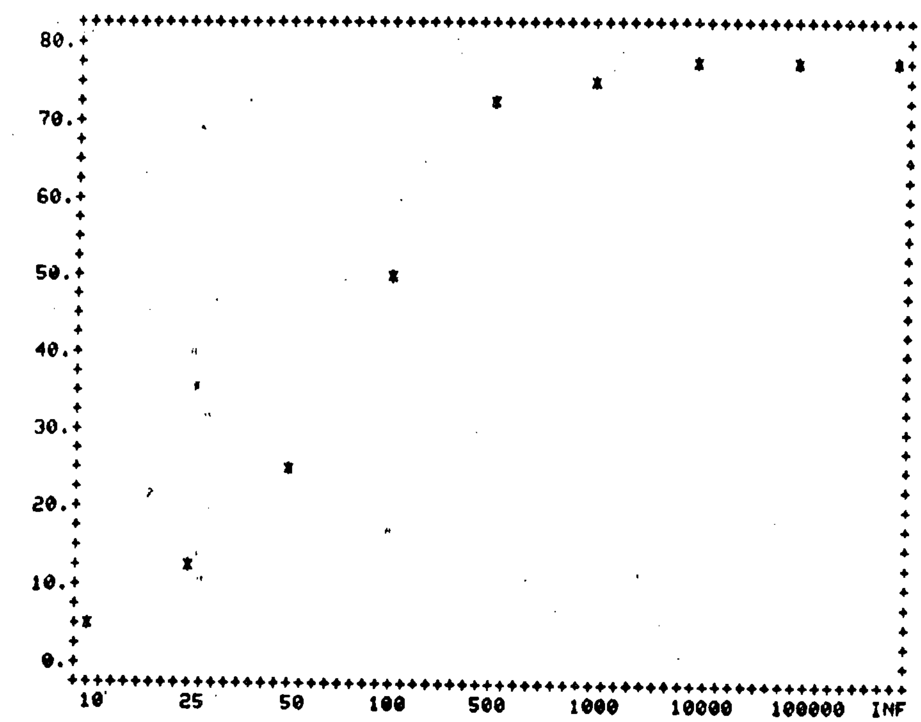
The new result is plotted on the next graph. Now the risk is increasing in n, a property which was proven to hold in the zero interview cost problem. (See CMRS [3].)

THE SECRETARY PROBLEM

WHERE THE TOTAL COST FOR OBSERVING ALL N SECRETARIES IS 1000-1/N

NO COST FOR FIRST INTERVIEW
COST PER SUCCEEDING INTERVIEW IS THEREFORE 1000/N
WHERE N=NUMBER OF APPLICANTS.

GRAPH OF EXPECTED RISK VS. N



THE MINIMAL RISK FOR N=10 IS		5.50000
THE MINIMAL RISK FOR N=25 IS		13.00000
THE MINIMAL RISK FOR N=50 IS		25.50000
THE MINIMAL RISK FOR N=100 IS		50.50000
THE MINIMAL RISK FOR N=500 IS		72.76363
THE MINIMAL RISK FOR N=1000 IS		75.55537
THE MINIMAL RISK FOR N=10000 IS		78.04964
THE MINIMAL RISK FOR N=100000 IS		78.29794
THE MINIMAL RISK FOR N=INFINITY IS		78.32553

Figure 3.3 Rank problem, linear cost, total cost 1000 - 1/N.

A close look at the last column in Table 3.1 indicates $V_K \approx 2.477\sqrt{K}$ for K large. This observation motivated the next proposition.

Prop. 8 Let $h_K(x) = Kx$, $q(i) = i$, $V_K = \inf_{\tau} E[q(x_\tau) + h_K(\tau)]$. Then $V_K = O(\sqrt{K})$ and not $o(\sqrt{K})$.

Proof:

Let τ_K be the single island rule τ applied to $[0, \frac{1}{\sqrt{K}}]$, where τ is the optimal stopping rule for zero interview cost. That is $\beta_i = \frac{1}{\sqrt{K}} \prod_{j=i}^{\infty} \left(\frac{1}{j+2}\right)^{\frac{1}{j+1}} \leq \frac{1}{\sqrt{K}}$, and $\tau_K \leq \frac{1}{\sqrt{K}}$. The arrival distributions on $[0, \frac{1}{\sqrt{K}}]$ are uniform and play the same role as the arrival distributions on [0,1]. Also, $R_i(x) = \frac{i}{x}$ so we can conclude

$$
\begin{aligned}
V_K &\leq E(q(x_{\tau_K}) + h(\tau_K)) \\
&\leq E(q(x_{\tau_K})) + h\left(\frac{1}{\sqrt{K}}\right) \\
&= \sqrt{K}\, E(q(x_\tau)) + \frac{K}{\sqrt{K}} \\
&= \sqrt{K}\left(\prod_{j=1}^{\infty} \left(\frac{j+2}{j}\right)^{\frac{1}{j+1}} + 1\right) \leq 4.87\sqrt{K}
\end{aligned}
$$

$$V_K = O(\sqrt{K}).$$

Suppose now we could observe the entire unit interval and choose the place to stop. Obviously you would choose $\tau = t$ where $X_t = i_0$, i_0 minimizing $(i+KU_i)$ and U_i is the arrival time of the $i^{\underline{th}}$ best candidate (w.r.t. [0,1]).

Then, $V_K \geq E(\min_i(i + K\, U_i))$

$$= \int_1^{K+1} P(\min_i(i + K\, U_i) > y)\, dy$$

$$= \int_1^{K+1} \prod_{i=1}^{[y]} \left(1 + \frac{i-y}{K}\right) dy$$

$$\geq \int_1^{K+1} \left(1 + \frac{1-y}{K}\right)^{[y]} dy$$

$$= \sum_{i=1}^{K} \int_i^{i+1} \left(1 + \frac{1-y}{K}\right)^i dy$$

$$\geq \sum_{i=1}^{[\sqrt{K}]} \left(1 - \frac{i}{K}\right)^i$$

$$\geq [\sqrt{K}] \left(1 - \frac{[\sqrt{K}]}{K}\right)^{[\sqrt{K}]} \to e^{-1} \sqrt{K}\ .$$

V_K is not $o(\sqrt{K})$. //

We can show that $V_K/\sqrt{K}$ actually has a limit by using the following trick:

Let $g_K(x) = \dfrac{f_K(x)}{\sqrt{K}}$ where $f_K(x) < \infty$ on $[0,1)$, $f_K(1) = \infty$, and $f_K'(x) = \dfrac{1}{x} \sum_{i=1}^{\infty} (f_K(x) - \frac{i}{x} - Kx)^+$. Letting $y = \sqrt{K}\ x$ we substitute into g_K to obtain

$$(3.39) \qquad g_K'(y) = \frac{1}{y} \sum_{i=1}^{\infty} (g_K(y) - \frac{i}{y} - y)^+, g_K(y) < \infty \text{ on } [0,\sqrt{K})$$

$$g_K(\sqrt{K}) = \infty.$$

Clearly $g_K(0) = \dfrac{V_K}{\sqrt{K}}$ and $g_K(0)$ is decreasing in K. Since $\dfrac{V_K}{\sqrt{K}} > e^{-1}$, $\dfrac{V_K}{\sqrt{K}}$ has a limit. If we define $g_\infty(x)$ by

$$(3.40) \qquad g_\infty'(x) = \frac{1}{x} \sum_{i=1}^{\infty} (g_\infty(x) - \frac{i}{y} - y)^+, g_\infty(x) < \infty \text{ on } [0,\infty)$$

$$g_\infty(\infty) = \infty$$

then hopefully $g_\infty(0)$ is unique and equals $\lim V_K/\sqrt{K}$. We can prove even more, we can prove $\lim_{k\to\infty} g_K(x)$ satisfies (3.40) for all x. Since $g_K(x)$

is decreasing in K for each x and $g_K(x) > 0$, $\lim_{k\to\infty} g_K(x)$ exists. Let h(x) be this limit. Consider any interval [0,z]. Choose K_0 so $\sqrt{K} > z$. From an earlier section we know that there are unique x_{Ki} satisfying $g_K(x_{Ki}) = \frac{i}{x_{Ki}} + x_{Ki}$ and since $g_K(x)$ is decreasing in K for fixed x, x_{Ki} is increasing in K for fixed i. Choose I so large that $x_{K_0 I} > z$. Then for all $K \geq K_0$, $x \in [0,z]$, $g_K(x) < \frac{I}{z} + z$ and we can replace (3.39) with

$$g'_K(y) = \frac{1}{y} \sum_{i=1}^{I} \left(g_K(y) - \frac{i}{y} - y\right)^+ \quad K \geq K_0,\ y \leq z. \tag{3.41}$$

Let us now consider $g'_k(y)$. On $[0,x_{11}]$, $g'_K(y) = 0$ for all K so zero will not bother us. Since $g_K(y)$ is continuous in y and decreasing in K so is $g'_K(y)$. For all $K \geq K_0$, $x \in [0,z]$, $0 \leq g'_K(x) \leq \frac{I}{x_{11}}\left(\frac{I}{z} + z\right)$ so g_K, $K \geq K_0$ is an equi-Lipschitzian family. Therefore $g_K(x)$ converges uniformly to its limit h(x) for $x \in [0,z]$ and h(x) is Lipschitzian. Using (3.41) it is easy to see $g'_k(y)$ is equicontinuous so $\lim_{k\to\infty} g'_K(y)$ exists and equals h'(y), $y \in [0,z]$. Since z was chosen arbitrarily, we get

$$h'(x) = \lim_{K\to\infty} g'_K(x) = \lim_{K\to\infty} \frac{1}{x} \sum_{i=1}^{\infty} \left(g_K(x) - \frac{i}{x} - x\right)^+ = \frac{1}{x} \sum_{i=1}^{\infty} \left(h(x) - \frac{i}{x} - x\right)^+$$

and since $h(x) < \infty$ for $x < \infty$, $h(\infty) = \infty$, h(x) satisfies (3.40).

We can numerically show (3.40) is unique and evaluate it by approximating with cutoff points as was done earlier. Let $x_{\infty i}$ be the cutoff point. A lower bound for x_{Ki} is also a lower bound for $x_{\infty i}$ so we have already obtained lower bounds. An upper bound is obtained by setting $g_\infty(N) = N$. This corresponds to the problem where, if you have not stopped by time N, the loss is N, which is the limit of the corresponding finite problems. These give upper and lower bounds on all possible $g_\infty(0)$, which computationally is unique to seven decimal places.

(3.42) $$\lim V_K/\sqrt{K} = g_\infty(0) = 2.4768709.$$

It is interesting to note that 2.477 appears as a solution to a secretary type problem presented in an entirely different format. This problem was presented by Professor H. Robbins while visiting Purdue University during the summer of 1976. Its solution is given by the rank problem with linear cost. Briefly stated, the problem is:

You are interested in obtaining the best possible item where each item is drawn at random from an infinite population having an unknown continuous cumulative distribution function F(x). (Best meaning lowest possible x value.) You pay an amount c per observation and cannot recall any items. Upon stopping, pay an amount equivalent to the expected percentile of F.

In statistical terms, if Y_i is the relative rank of the $i^{\underline{th}}$ item, then the expected percentile is $Y_i/(i+1)$. So we are interested in

(3.43) $$V_c = \inf_\tau E\left(\frac{Y_\tau}{\tau+1} + c\,\tau\right).$$

Robbins showed $\sqrt{2c} \leq V_c \leq 2\sqrt{2c}$, so wanted to know the value of $\lim_{c\to 0} \frac{V_c}{\sqrt{c}}$. We will show this limit is $g_\infty(0) \approx 2.477$.

Choose K. For any sequence $c_m \to 0$, let N_m be chosen so that $\frac{(N_m+1)\sqrt{c_m}}{\sqrt{K}} \to 1$. Then if V_K^N is the finite N-girl secretary problem with linear cost Kx, we will use continuity of the $V^{\underline{s}}$ to conclude

$$\lim_{m\to\infty} \frac{V_{c_m}}{\sqrt{cm}} = \lim_{m\to\infty} \inf_\tau E\left[\frac{Y_\tau}{\sqrt{c_m}(\tau+1)} + \sqrt{c_m}\,\tau\right]$$

$$\leq \lim_{m\to\infty} \inf_{\tau\leq N_m} \frac{1}{\sqrt{c_m}(N_m+1)} E\left[\left(\frac{N_m+1}{\tau+1}\right) Y_\tau + [N_m(N_m+1)c_m]\frac{\tau}{N_m}\right]$$

$$= \lim_{m\to\infty} \frac{1}{\sqrt{c_m}(N_m+1)} V^{N_m}_{N_m(N_m+1)c_m}$$

$$= \frac{V_K}{\sqrt{K}} \leq g_\infty(0).$$

For the other inequality let us consider the kickback problem as considered in Gianini and Samuels [7]. Take the rank problem with linear cost Kx and the loss for $\tau = 1$ equal to K. It is easy to see that the risk divided by $\sqrt{K}$ is given by $h_K(0)$ where h_K satisfies differential equation (3.40) with side condition $h_K(\sqrt{K}) = \sqrt{K}$.

So for the same sequences c_m and N_m,

$$\lim_{m\to\infty} \frac{V_{c_m}}{\sqrt{c_m}} = \lim_{m\to\infty} \inf_{\tau} E\left[\frac{Y_\tau}{\sqrt{c_m}(\tau+1)} + \sqrt{c_m}\,\tau\right]$$

$$\geq \lim_{m\to\infty} \inf_{\tau} E\left[\frac{Y_\tau}{\sqrt{c_m}(\tau+1)} + \sqrt{c_m}\,\tau\right] I_{\{\tau < N_m\}} + \sqrt{c_m}\, N_m I_{\{\tau \geq N_m\}}$$

$$= \lim_{m\to\infty} \inf_{\tau \leq N_m} E\left[\frac{Y_\tau}{\sqrt{c_m}(\tau+1)} + \sqrt{c_m}\,\tau\right] I_{\{\tau < N_m\}} + \sqrt{c_m}\, N_m I_{\{\tau = N_m\}}$$

$$= \lim_{m\to\infty} \inf_{\tau \leq N_m} \frac{1}{(N_m+1)\sqrt{c_m}}\left[E\,\frac{N_m+1}{\tau+1}\, Y_\tau + [N_m(N_m+1)c_m]\frac{\tau}{N_m}\right] I_{\{\tau \leq N_m\}}$$

$$+ N_m(N_m+1)c_m I_{\{\tau = N_m\}}$$

$$= h_K(0).$$

Since this is true for all K, and by computation $h_K(0) \to g_\infty(0)$, $\lim_{c\to 0} \frac{V_c}{\sqrt{c}} = g_\infty(0) = \lim_{k\to\infty} h_K(0) \approx 2.477.$

CHAPTER IV

THE SECRETARY PROBLEM WITH BACKWARD SOLICITATION

In this chapter the secretary problem is considered under less restrictive criteria. We assume the number of candidates n is fixed, the candidates are rankable and arrive in a random fashion. However, we do not make the restriction that the present candidate must be selected now or lost forever. Instead, we assume that a candidate who has been passed can be recalled. However, the longer you wait, the smaller the probability of successfully hiring the candidate. We assume that once a candidate refuses an offer she will always refuse. (Presumably she has accepted another offer.)

This new formulation makes the secretary problem much more realistic. For example, in the application formally known as the secretary problem, a candidate is generally available until she accepts another offer. Or, in the application concerning the purchase of a house, the house is available until someone else buys it; the longer you wait, the smaller the probability of its availability.

As is typical, the more realistic the problem, the more difficult it is to solve. In the classical case, only the rank of the present candidate is important. (Since this is the only candidate you could possibly hire at the present time.) In our new formulation, all of the previous candidates can possibly be hired, so all must be considered.

This fact discourages the use of the well known method known as backward induction. For the classical problem with n applicants, n steps are required to solve the problem with backward induction. For the new problem, if we are presently at time i, we must consider the position of all i candidates and whether or not they have previously refused an offer. Since there are i! ways in which the numbers 1 through i can be ordered, backward induction will require $1! + 2^1 \cdot 2! + \ldots + 2^{n-1} n!$ steps. Even for n as small as 10, we need approximately two billion steps to solve the problem. For $n = 50$, we require greater than 10^{79} iterations. To give some idea how large this number is, it is estimated that, in all the universe, there are only 10^{74} particles! And 50 applicants is certainly a reasonable number.

To make matters worse, we no longer have independence of the past and future. For example, if I have previously attempted to solicit the fifth candidate to arrive, then in the future, this candidate can never accept. If I have never attempted to recall, then the probability she will accept is specified by a probability function, and need not be zero.

1. Formulation

Let us use a formulation similar to Yang [20]. Let

(4.1) X_i = absolute rank of $i^{\underline{th}}$ candidate to arrive

(4.2) $Y_i(k)$ = relative rank of $i^{\underline{th}}$ candidate among first k to arrive

(4.3) $q(i)$ = loss incurred for hiring candidate of absolute rank i, $q(\cdot)$ increasing

(4.4) $p(i)$ = probability of successfully hiring the candidate interviewed i candidates back given she has never previously refused. Let $1 = p(0) \geq p(1) \geq p(2) \geq \ldots$ and $\lim_{n \to \infty} p(n) = p(\infty) \geq 0$. We assume $p(0) = 1$ because $p(0) < 1$ gives a positive probability all candidates will refuse.

(4.5) $Z_i(k)$ = time of arrival of candidate of relative rank i with respect to the first k. $(i \leq k)$

(4.6) $t_i(k)$ = $k - Z_i(k) \geq 0$ if $i^{\underline{th}}$ ranked candidate has not been previously solicited

= $Z_i(k) - k \leq 0$ if $i^{\underline{th}}$ ranked candidate has been unsuccessfully solicited. t is the "time" elasped since candidate of current relative rank i has been interviewed

(4.7) $R_i(k) = \sum_{\ell=i}^{n-(k-i)} q(\ell) \frac{\binom{\ell-1}{i-1}\binom{n-\ell}{k-i}}{\binom{n}{k}}$, the risk (expected loss) if the candidate of relative rank i (w.r.t. first k) accepts.

(4.8) V_n = minimal risk for given $p(\cdot)$, $q(\cdot)$ and n.

For the best choice problem, $q(1) = 0$, $q(i) = 1$ $i > 1$, Yang [20] and later Smith [18] considered this formulation. Since they are only interested in obtaining the best candidate, then at time k they only considered the candidate of current relative rank 1. (Obviously no other candidate could possibly be absolute best.) The ordering of the remaining candidates is irrelevant. Since at time i, there are i possible positions for the relative best, backward induction can work. It requires $1 + 2 + \ldots + n = n(n+1)/2$ steps. This is certainly feasible.

Yang solved the best choice problem by giving the set of recursive equations. He considered some special cases and studied the limiting behavior. Smith generalized his work by considering the general limiting risk and limiting optimal procedures. Smith showed the limiting solution depends solely on $p(0)$ and $p(\infty)$. In particular, if $p(0) = 1$ and $p(\infty) = 0$, then the limiting risk is e^{-1}, the limit with no backward solicitation. Also, the classical procedure is asymptotically optimal. If $p(\infty) > 0$, procedures similar to the classical procedure are asymptotically optimal. These procedures are described by skipping r_n candidates and taking the first relatively best candidate arriving after r_n. (If there is one.) If none appear by the time n is reached, go back and solicit the best.

In this chapter, we will use a technique similar to Smith. A certain class C of $p(\cdot)\underline{s}$ will be considered. For this class, backward induction is feasible and is used to solve the problem. The limiting solution is easily derived using results from Chapter 2.

Next, the asymptotic solution for a general $p(\cdot)$ is considered. It turns out that the class C is asymptotically essentially complete in the sense that the limiting risk, as $n \to \infty$, depends only on $p(\infty)$ and can be computed using the unique member of C with the same value $p(\infty)$.

Both the procedure and the limiting risk are obtained at least as easily as the classical case. That is, they are given by a differential equation with a right boundary condition.

Specifically, the results for the best choice and rank problem are obtained. The best choice result agrees with Smith.

2. Non-optimal Procedures

Suppose we have just interviewed the $k^{\underline{th}}$ candidate to arrive. (We are at time k.) Our available procedures are to attempt to hire (solicit) any or all of the previously interviewed candidates (and in any order) or go on. Let us denote each of these procedures by an ordered set $J_k = \langle j_1, j_2, \ldots, j_\ell \rangle$ for some $\ell < k$, $j_n \neq j_m$ and $1 \leq j_n \leq k$. By J_k we mean the following procedure: First solicit the candidate of current relative rank j_1. If she refuses, solicit the candidate of current relative rank j_2. If both refuse, solicit the candidate of current relative rank j_3. And so on. If all ℓ refuse, go on. (That is, interview the next candidate.) $J_k = \emptyset$ is interpreted as going on without solicitation.

In general, the optimal choice of J_k is random. It depends on both the arrival times and the availabilities of all previously interviewed candidates. J_k cannot be specified prior to the start of the process.

The $t_i(k)$ (recall (4.6)) contain information on both availability and arrival times of all previously interviewed candidates. $t_i(k) < 0$ means the $i^{\underline{th}}$ relative best is unavailable and $t_i(k) \geq 0$ tell how long ago she has been interviewed. Let us represent the past by $T_k = (t_1(k), t_2(k), \ldots, t_k(k))$. Then we know the optimal procedure is some $J_k^*(T_k)$. That is, if T_k were known, the optimal procedure could be found. But there are $2^k\, k!$ different $T_k^{\underline{s}}$ so we cannot possibly specify all $J_k^*(T_k)^{\underline{s}}$.

Despite this difficulty, it is not too hard to show that certain procedures are inadmissible, and dominated by an essentially complete class.

The first result is obvious, do not solicit a candidate who has previously refused an offer. Once an individual has refused, she will always refuse so the additional solicitation is useless. However, there is no solicitation cost and it may be notationally convenient to include this individual in the solicitation set.

A true reduction in the class of admissible procedures can be made using the following theorem.

Thm. 4.1 An essentially complete class C_0 is given by

$C_0 = \{J_k \mid$ with probability one the co-ordinates of J_k are in increasing order$\}$.

Proof:

The proof is a consequence of the following lemma:

Lemma 1 If $X_1 \leq X_2 \leq \cdots \leq X_n$, $0 \leq p_i \leq 1$ for $i = 1,2,\ldots, n$ and σ is a permutation on $1,2,\ldots, n$, then

$$\sum_{i=1}^{n} p_i \left(\prod_{j=1}^{i-1} (1-p_j) \right) X_i \leq \sum_{i=1}^{n} p_{\sigma(i)} \left(\prod_{j=1}^{i-1} (1-p_{\sigma(j)}) \right) X_{\sigma(i)} \ .$$

Proof:

Choose ℓ so that $\sigma(\ell+1) < \sigma(\ell)$.

Let $\sigma^1(i) = \sigma(i) \qquad i \neq \ell, \ell+1$

$\sigma^1(\ell) = \sigma(\ell+1)$

$\sigma^1(\ell+1) = \sigma(\ell)$.

Then

$$\sum_{i=1}^{n} p_{\sigma^1(i)} \left(\prod_{j=1}^{i-1} (1-p_{\sigma^1(j)}) \right) X_{\sigma^1(i)} - \sum_{i=1}^{n} p_{\sigma(i)} \left(\prod_{j=1}^{i-1} (1-p_{\sigma(j)}) \right) X_{\sigma(i)} =$$

$$\prod_{j=1}^{\ell-1} (1-p_{\sigma(j)}) \Big[p_{\sigma^1(\ell)} X_{\sigma^1(\ell)} + (1-p_{\sigma^1(\ell)}) p_{\sigma^1(\ell+1)} X_{\sigma^1(\ell+1)} - p_{\sigma(\ell)} X_{\sigma(\ell)} - (1-p_{\sigma(\ell)}) p_{\sigma(\ell+1)} X_{\sigma(\ell+1)} \Big] =$$

$$\prod_{j=1}^{\ell-1} (1-p_{\sigma(j)}) \left[p_{\sigma(\ell)} p_{\sigma(\ell+1)} (X_{\sigma(\ell+1)} - X_{\sigma(\ell)}) \right] \leq 0$$

since $X_{\sigma(\ell+1)} \leq X_{\sigma(\ell)}$.

Applying this result for every ℓ where $\sigma(\ell+1) < \sigma(\ell)$, we get successive inequalities that prove the lemma.//

The proof of the theorem follows easily since $j_1 < j_2$ implies $R_{j_1}(k) < R_{j_2}(k)$ (recall (4.7)) and the going on risk is the same regardless of the order of solicitation.//

So we only need consider procedures that solicit in increasing order of rank.

An interesting thought arises next. Suppose you are at time k and an "acceptable" procedure calls for soliciting an individual of relative rank j. Is it always better to solicit a candidate of lower rank first? That is, is there a situation where you want to solicit a candidate of rank j but not one of rank i where i < j? Intuitively the answer is no. If you are willing to take someone who is fifth best, then you should at least try to get someone second best. Unfortunately, there are no clear relationships between the going on risks given a candidate is available and given she is not. I will state the conjecture in general and prove it in a special case that applies to the asymptotically essentially complete class.

<u>Conjecture 1</u> An essentially complete class C_1 is given by

$$C_1 = \{J_k | J_k = < 1,2,\ldots, \ell > \text{ for some } \ell\}.$$

The order of J_k is guaranteed by theorem 4.1. If the conjecture were true, then we have a much simpler situation than we had imagined. The $J_k^{\underline{s}}$ are still random but the number of possibilities has been reduced to k+1 available strategies, namely $\ell = 0,1,2,\ldots, k$.

A special case of the conjecture is the following theorem:

Thm. 4.2 On $\{p(t_{j_m}(k)) = 1$ for some $j_m \in J_k\}$, we can replace J_k by $J_k' = < 1,2,\ldots, j_m>$ without increasing the risk. We may actually decrease the risk.

Proof:

Let $j_1, j_2,\ldots, j_m$ be the first m components of J_k. Theorem 4.1 assures us that $j_1 < j_2 < \ldots < j_m$. Assume $j_\ell < \hat{j} < j_{\ell+1}$ for some integers ℓ and $\hat{j}$. (If this is not the case the proof is complete.)

I claim that $\hat{J}_k = < j_1,\ldots, j_\ell, \hat{j}, j_{\ell+1},\ldots, j_m>$ is at least as good as J_k. Let L_k and $\hat{L}_k$ be the respective expected losses using J_k and $\hat{J}_k$ given we have reached time k. A direct calculation gives

$$L_k - \hat{L}_k = \prod_{n=1}^{\ell-1} [1 - p(t_{j_n}(k))] \Bigg[\sum_{i=\ell+1}^{m} p(t_{j_i}(k)) \left(\prod_{n=\ell}^{i-1} [1 - p(t_{j_n}(k))] \right)$$

$$+ R_{j_i}(k) - p(t_{\hat{j}}(k)) R_{\hat{j}}(k) - [1 - p(t_{\hat{j}}(k))]$$

$$+ \left[\sum_{i=\ell+1}^{m} p(t_{j_i}(k)) \left(\prod_{n=\ell}^{i-1} [1 - p(t_{j_n}(k))] \right) R_{j_i}(k) \right] \Bigg]$$

$$= \sum_{i=\ell+1}^{m} p(t_{j_i}(k)) \left(\prod_{n=1}^{i-1} [1 - p(t_{j_n}(k))] \right) [R_{j_i}(k) - R_{\hat{j}}(k)] \geq 0.$$

Successively applying this result at all ℓ for which $j_\ell < j_{\ell+1} - 1$, and prior to j_1 if $j_1 \neq 1$, we arrive at our theorem.//

Conjecture 1 cannot be proved in the same way because we typically have a going on risk. Using the k-hat procedure we have a smaller risk if we stop but a larger risk if we go on. (Since one less candidate is available in the future.) Nevertheless, it is reasonable to assume that, if it is better to take the candidate of rank j_m than go on, it is even better to take the candidate of rank $\hat{j} < j_m$.

If conjecture 1 holds, we can concisely specify the optimal procedure. At each time k, we need only specify $\ell_k = \ell_k(T_k)$ indicating the following procedure: Solicit the candidate of relative rank 1 if she is available. If she refuses, solicit the candidate of relative rank 2 if she is available. If both refuse, etc. If all ℓ_k refuse, go on. ($\ell_k = 0$ means go on without solicitation.)

Of course ℓ_k is random (depending on the past) so even if conjecture 1 were true, backward induction is not generally useful. But this type of procedure is of sufficient importance that we maintain the notation for later use.

Our next theorem is intuitively obvious but still must be carefully proved. As the probability of successful solicitation increases, the minimal expected loss decreases. Formally,

<u>Thm. 4.3</u> If $p^1(i) \geq p^2(i)$ for i = 1,2,..., n and q(•) is fixed, then $V_n^1 \leq V_n^2$ where V_n^i is the optimal risk corresponding to p^i, $i = 1,2$.

<u>Proof</u>:

Even though backward induction is too lengthy to use in practice, it still gives the theoretical solution and is useful in this proof.

Let $V^i_{n|T_k}$ $(i = 1,2)$ be the optimal going on risk given we are at time k with the "past" T_k known. It suffices to show $V^1_{n|T_k} \leq V^2_{n|T_k}$ for all k, T_k.

At time n, we must select a candidate so both theorem 1 and theorem 2 apply. Therefore, we solicit in order of increasing rank. Since q(•) is non-decreasing,

$$V^1_{n|T_n} \leq V^2_{n|T_n} \quad \text{for each } T_n.$$

Assume the same is true at time k, i.e.,

$$V^1_{n|T_k} \leq V^2_{n|T_k} \quad \text{for each } T_k.$$

Consider the optimal procedure for the $p^2(\cdot)$ problem applied at time k-1. (Knowing T_{k-1}) Necessarily, the risk for stopping is smaller than the going on risk. (Otherwise it is not optimal.) The same procedure, applied to the $p^1(\cdot)$ problem has both a higher probability of stopping and a lower going on risk. And this procedure is not necessarily optimal for $p^1(\cdot)$ given T_{k-1}. Therefore, we have reasoned

$$V^1_{n|T_{k-1}} \leq V^2_{n|T_{k-1}} \quad \text{for each } T_{k-1}$$

and by induction our theorem is proved.//

3. A Special Class of $p(\cdot)$'s

In this section we consider a special class of $p(\cdot)$ functions. This class was chosen for many reasons. One, mentioned earlier, is the property of asymptotic essential completeness. Another is the feasibility of backward induction. The form of the optimal solution is easy to write down and lends itself readily to the analysis developed in chapter 2. The limiting solution is obtained through a differential equation with a given right boundary condition. The class C considered is

$$(4.9) \qquad \begin{aligned} p(i) &= 1 && i = 0,1,2,\ldots, r \\ &= p^* && i = r+1, r+2,\ldots \qquad (0 \leq p^* < 1). \end{aligned}$$

We can immediately classify the form of the optimal risk.

Prop. 4.1 If $p(\cdot)$ is of the form (4.9) then the minimal risk is attained using only procedures J_k satisfying

(a) $J_k = \emptyset$ or $J_k = < 1,2,\ldots, Y_{k-r}(k) >$

(b) On $\{J_k = \emptyset$ for all $k < k^* < n$, $J_{k*} \neq \emptyset$,

$$Y_{k-r}(k) = \min_{s \leq r} Y_{k-s}(k)\}$$

(c) If $r = 0$ for each k there exists a constant a_k so that

$$J_k = \emptyset \quad \text{if } Y_k(k) > a_k$$
$$= <1,2,\ldots, Y_k(k)> \text{ if } Y_k(k) \leq a_k.$$

In other words (a) if you solicit then the procedure stops with probability 1, (b) you solicit only when the best among the previous r is precisely r units back, and (c) for $r = 0$ the stopping rules are no longer random. For $r = 0$, wait until the first time $Y_k(k) \leq a_k$. At that time, solicit candidates in increasing order of rank up to an including rank $Y_k(k)$ if necessary. Since $Y_k(k)$ is the present candidate, the process will stop at this time with certainty.

<u>Proof</u>:

Let $k < n$. If $k = n$, theorems 4.1 and 4.2 apply so we know we must solicit in order of increasing rank. This is equivalent to $J_n = < 1,2,\ldots, Y_{n-r}(n) >$.

(a) Suppose $J_k \neq \emptyset$ and $Y_{k-r}(k) \notin J_k$. Let us assume $J_k = < j_{k,1}, j_{k,2}, \ldots, j_{k,\ell} >$, $J_{k+1} = < j_{k+1,1}, j_{k+1,2}, \ldots, j_{k+1,m} >$. (Recall these are typically random.) From time k to time k+1, only the probability of recalling the candidate who arrived at time k-r is changed so we obviously obtain the same risk using $J'_k = \emptyset$, $J'_{k+1} = (J_k, J_{k+1}) = < j_{k,1}, \ldots, j_{k,\ell}, j_{k+1,1}, \ldots, j_{k+1,m} >$. We can repeat this argument on J'_{k+1} provided $Y_{k+1-r}(k+1) \notin J'_{k+1}$. Suppose $Y_{k-r}(k) \in J_k$ for some k. Then $p(t_{k-r}(k)) = 1$ so theorems 4.1 and 4.2 apply and we rewrite J_k as $< 1,2,\ldots, Y_{k-r}(k) >$.

(b) Suppose $J_k = \{1,2,\ldots, Y_{k-r}(k)\}$. (By (a) these are the conditions given in (b).) Let S_0 be such that $Y_{k-S_0} = \min_{S \leq r} \{Y_{k-S}(k)\}$. If $S_0 = r$, we have nothing to prove. If $S_0 < r$, then $p(t_{k-S_0}(k)) = 1$ so by theorems 4.1 and 4.2 we can rewrite J_k as $J_k' = \{1,2,\ldots, Y_{k-S_0}(k)\}$. But then part (a) of the proof applies and we can modify the procedures until $S_0 = r$.

(c) Let $r = 0$. Since the minimal going on risk is independent of the past and present and the stopping risk is independent of the past given k and $Y_k(k)$, we can use backward induction. (This will be done later.) For a fixed k then, the going on risk is fixed while $< 1,2,\ldots, Y_k(k) >$ gives an increasing risk as the value of $Y_k(k)$ increases. This implies the existance of this fixed constant a_k satisfying the conditions of (c).//

So for $p(\cdot)$ of the form (4.9) we can make many simplifications. One mentioned earlier is that $J_k(T_k) = \ell_k(T_k)$. This is simplified further since either $\ell_k(T_k) = 0$ or $\ell_k(T_k) = Y_{k-r}(k)$. We know that no solicitations are to be made until we are ready to stop, and then we will solicit in order. Once we reach the time when we start soliciting, one of the solicited candidates will accept the offer with probability one. That is, once we reach a point where it is not optimal to immediately go on, we will stop with probability one.

For $r = 0$, backward induction works easily so let us set up the recursion equations. As in the previous chapters let

(4.10) $\quad n$ = number of applicants

(4.11) $\quad V_n(p^*) = \inf_{\tau} E(q(x_\tau))$ = minimal risk over all procedures

(4.12) $\quad C_{p^*}(k) = \inf_{\tau > k} E(q(x_\tau))$ = minimal risk over all procedures that interview at least k+1 candidates.

Then applying (c) from proposition 4.1 we get

$$(4.13)\qquad C_{p*}(n-1) = \frac{1}{n}\sum_{k=1}^{n}\left(\sum_{i=1}^{k-1} p*(1-p*)^{i-1}\, q(i) + (1-p*)^{k-1} q(k)\right)$$

$$= \frac{1}{n}\sum_{k=1}^{n} S_n(k)$$

$$(4.14)\qquad C_{p*}(r-1) = \frac{1}{r}\sum_{k=1}^{r}\min\left(\sum_{i=1}^{k-1} p*(1-p*)^{i-1} Q(r,i) + (1-p*)^{k-1}\, Q(r,k),\ C_{p*}(r)\right)$$

$$= \frac{1}{r}\sum_{k=1}^{r}\min(S_r(k),\ C_{p*}(r))$$

where

$$(4.15)\qquad Q(r,i) = \sum_{\ell=i}^{n-(r-i)} q(\ell)\,\frac{\binom{\ell-1}{i-1}\binom{n-\ell}{r-i}}{\binom{n}{r}}$$

and

$$(4.16)\qquad S_r(k) = \sum_{i=1}^{k-1} p*(1-p*)^{i-1} Q(r,i) + (1-p*)^{k-1} Q(r,k).$$

Implicit in these relations are the facts that $C_{p*}(\cdot)$ depends on n and $S_r(k)$ depends on p*. I will not introduce these into the notation until later when they are needed.

Clearly we have $C_{p*}(r)$ increasing in r and $S_r(k)$ decreasing in r and increasing in k. So we can employ the arguments from the classical secretary problem to prove that we have cutoff point rules. That is, in addition to the fixed a_k we have a fixed sequence of cutoff points α_i which specify the optimal procedure. That is, we are to stop the first time a candidate of relative rank i arrives after time α_i. Upon stopping, solicit the candidates in order of increasing rank until a candidate accepts. (One must accept with probability one.) If we reach n, then solicit in order of increasing rank.

Prop. 4.2 $C_{p*}(r)$ is continuous in p* for all n and r.

Proof:

Fix n and use induction on r. From (4.16), $S_r(k)$ is continuous in p* for all r and k. Therefore $C_{p*}(n-1)$ is continuous in p*.

Assume $C_{p*}(m)$ is continuous. Using relation (4.14) and the fact that the minimum and sum of continuous functions are continuous, $C_{p*}(m-1)$ is continuous. This completes the induction hypothesis.//

The heuristics of the recursive equations (4.14) are straightforward. Normalize $\{1,2,\ldots, n\}$ to $\{\frac{1}{n},\frac{2}{n},\ldots, 1\}$ and let $r/n \to x$. Then

(4.17) $$C_{p*}(r/n) \to f_{p*}(x)$$

(4.18) $$S_r(k) \to T_k(x) = \sum_{i=1}^{k-1} p^*(1-p^*)^{i-1}R_i(x) + (1-p^*)^{k-1}R_k(x)$$

where $Q(r,k) \to R_k(x) = \sum_{\ell=k}^{\infty} q(\ell)\binom{\ell-1}{k-1}x^k(1-x)^{\ell-k}$.

Writing (4.14) as a difference equation and letting this tend to a differential equation we get

(4.19) $$f'_{p*}(x) = \frac{1}{x}\sum_{k=1}^{\infty}[f_{p*}(x) - T_k(x)]^+ \qquad \text{subject to}$$

$$f_{p*}(1) = \sum_{i=1}^{\infty} q(i)\; p^*(1-p^*)^{i-1} \qquad p^* \neq 0$$

$$= \sup q(i) \qquad p^* = 0.$$

Remark: For $p^* = 0$, $T_k(x) = R_k(x)$ and we have the classical differential equation $f'_0(x) = \frac{1}{x}\sum_{k=1}^{\infty}[f_0(x) - R_k(x)]^+$ subject to $f_0(1) = \sup q(i)$.

Chapter II developed a method formalizing the heuristics of such an argument. The results can be applied here to show that the difference equation indeed tends to the differential equation and the risk V_n for the finite problem tends to $f_{p*}(0)$, the corresponding risk for the infinite problem.

It is easy to verify properties (2.1), (2.8), (2.9), (2.10), and (2.11) hold for $T_k(x)$ for each value of p*. So we immediately conclude that (4.19) holds for the infinite problem with loss function $T_k(x)$.

Assume p* is such that $f_{p*}(0) < \infty$. That is, assume a solution to the differential equation exists for this value of p*. We want to show $V_n(p*) \to f_{p*}(0)$ as $n \to \infty$. The result follows from Chapter II.

Thm. 4.4 If $f_{p*}(0) < \infty$, then $V_n(p*) \to f_{p*}(0)$.

Proof:

Since $\sup_{k,r} \left| S_r(k) - T_k\left(\frac{r}{n}\right) \right| \to 0$ as $n \to \infty$, we can quote corollary 2 of theorem 2.4 to prove

(4.20) $$\underline{\lim}\ V_n(p*) \geq f_{p*}(0).$$

For the other inequality we use a trick of Gianini [6]. Let us look at the infinite problem where we are given additional information. Assume we are told whether a candidate is one of the best n or not. Obviously we will only consider stopping on one of these best n. Let V_n' be the associated minimal risk. Two things are obvious. First this problem is identical to the finite problem i.e.,

$$V_n' = V_n(p*) \quad \text{for all } n.$$

Second, since additional information is supplied, this risk is smaller than the risk for the infinite problem with no additional information i.e.,

$$V_n' \leq f_{p*}(0) \quad \text{for all } n.$$

These, along with (4.20), prove equality.//

For p* increasing, $T_k(x)$ is decreasing for each k and x so we can trivially conclude $f_{p*}(x)$ is decreasing for all $x \in [0,1]$. So if $f_0(0) < \infty$, we know $f_{p*}(0) < \infty$ for all p*. So we get immediate conditions on $q(\cdot)$ to guarantee $f_{p*}(0) < \infty$ for all p*. Furthermore, for all $q(\cdot)$ $f_1(0) = q(1)$ so we can conclude $f_{p*}(0) < \infty$ on $[0,t)$ for some t.

By inspection, it can be seen that $f_{p*}(1)$ is continuous in p*. It would seem logical that $f_{p*}(0)$ is also continuous in p*, no matter what the loss function q(•). A direct proof is not tractable since, in general, it is not possible to "bring the continuity down the unit interval."

Instead of a direct proof, the following probabilistic proof works:

Prop. 4.3 If q(•) is truncated at M, then $f_{p*}(0)$ is continuous in p*.

Proof:

Let $p_n^* \uparrow p^*$. Since $f_{p*}(0)$ is decreasing in p*, $\underline{\lim}\, f_{p_n^*}(0) \geq f_{p*}(0)$.

Let $0 < x_1 \leq x_2 \leq \cdots \leq x_M = 1$ be the optimal cutoff points for the p* problem. Apply this rule to the p_n^* problem. Suppose the process stops at time x. Then the number of unsuccessful solicitations by the p_n^* and p* processes are respectively $N_1 \wedge Y_x$ and $N_2 \wedge Y_x$ where N_1 and N_2 are geometrically distributed with parameters p_n^* and p*. Rewrite N_1 as

$$N_1 = N_2 \qquad \text{with probability } p_n^*/p^*$$

$$= N_2 + N_3 \qquad \text{with probability } 1-(p_n^*/p^*) \text{ where } N_3 \sim \text{geometric } (p_n^*)$$

and N_3 is dependent of N_2.

Since for every x, $\tau = x \implies Y_x \leq M$, we conclude

$$f_{p_n^*}(0) \leq E_{p_n^*} q(x_\tau) \leq E_{p^*} q(x_\tau) + [1-(p_n^*/p^*)]\ q(M)$$

$$= f_{p*}(0) + [1-(p_n^*/P^*)]\ q(M).$$

The right hand side goes to zero as $n \to \infty$ so we get equality. The same process can used for $p_n^* \downarrow p^*$. ///

In general we have

Thm. 4.5 $f_{p*}(0)$ is right continuous in p*.

Proof:

Let $p_n^* \downarrow p^*$. We immediately have $\overline{\lim}\, f_{p_n^*}(0) \le f_{p^*}(0)$. Let $f_{p_n^*}^M(0)$ and $f_{p^*}^M(0)$ be the minimal risks corresponding to the truncated loss

(4.21) $$q^M(i) = q(i) \wedge q(M).$$

Using proposition 4.3 we have

$$\underline{\lim}_{n\to\infty}\; f_{p_n^*}(0) \ge \underline{\lim}_{n\to\infty}\; f_{p_n^*}^M(0) = f_{p^*}^M(0) \quad \text{for all M.}$$

But applying theorem 2.3 from Chapter II

$$f_{p^*}^M(0) \to f_{p^*}(0) \quad \text{as } M \to \infty \text{ so}$$

$$\underline{\lim}_{n\to\infty}\; f_{p_n^*}(0) \ge f_{p^*}(0).$$

Therefore we have equality.//

For the main theorem of this chapter, right continuity is sufficient. In general, it would seem like left continuity should also hold. For the rank problem, solved at the end of the chapter, left continuity can be shown. But in general, the technique used for that proof will not work.

We have totally solved the problem with $r = 0$. A similar but more complex method can be used on the problem with $r \ge 1$. Since we are only interested in the limiting solution, we consider only the next theorem. This theorem states that the limits are the same for $r = 0$ and any $r \ge 1$.

Thm. 4.6 Let $p^1(0) = 1$ and $p^2(i) = 1 \quad i = 1,2,\ldots, r$

$p^1(i) = p^*$ for $i \ge 1 \qquad p^2(i) = p^* \quad i \ge r+1$.

Then for r fixed,

$$\lim_{n\to\infty} V_n^2(p^*) = \lim_{n\to\infty} V_n^1(p^*) = f_{p^*}(0).$$

The intuitive reasoning behind the theorem is quite simple. Suppose we can solicit r individuals back with probability one of acceptance. For r fixed, $r/n \to 0$ as $n \to \infty$ so in the limit the $p^2(\cdot)$ problem should behave as the $p^1(\cdot)$ problem.

The actual proof is more complicated.

Proof:

$p^1(i) \le p^2(i)$ so by Theorem 4.3

$$\overline{\lim}\, V_n^2(p^*) \le \lim V_n^1(p^*) = f_{p^*}(0). \tag{4.22}$$

For the other inequality, consider the truncated loss (4.21). Let $V_n^{M,1}(p^*)$ and $V_n^{M,2}(p^*)$ be the risk corresponding to (4.21) for $p^1(\cdot)$ and $p^2(\cdot)$. It suffices to show

$$\underline{\lim}\, V_n^{2,M}(p^*) \ge \lim V_n^{1,M}(p^*) = f_{p^*}^M(0) \tag{4.23}$$

because we can then conclude $\underline{\lim}\, V_n^2(p^*) \ge \underline{\lim}\, V_n^{2,M}(p^*) \ge f_{p^*}^M(0) \to f_{p^*}(0)$. So let us verify (4.23). Consider the $p^2(\cdot)$ problem. Proposition 4.1 applies and we will use (a) and (b). These two can be used to show the optimal rule is to solicit when (1) $Y_{k-r}(k) < Y_{k-s}(k)$ for $s = 0,1,2,\ldots, r-1$ and (2) $S_{Y_{k-r}(k)}(k) \le C_k(T_k)$ where $C_k(T_k)$ is the going on risk. It is easy to verify that $C_k(T_k)$ is independent of the first k-r arrivals. That is

$$C_k(T_k) \equiv C_k(Y_{k-r+1}(k),\ Y_{k-r+2}(k), \ldots,\ Y_k(k)),$$

i.e., $C_k(T_k)$ depends only on the relative ranks of the last r-1 candidates. For notation ease, let $T_k - T_{k-r} = (Y_{k-r+1}(k), \ldots, Y_k(k))$. Let

$$C^k = \sup_{T_k - T_{k-r}} \{C(T_k - T_{k-r}) \text{ where } Y_{k-s}(k) > Y_{k-r}(k) \text{ for } s = 0,1,\ldots, r-1\}.$$

Then clearly $C(T_k - T_{k-r}) < C^k \Longrightarrow Y_{k-s}(k) < M$ for some $s = 0,1,\ldots, r-1$. (Recall q is truncated at M.) Define constants b_{k-r} by

$$b_{k-r} = \max\{i \mid S_i(k) \leq C^k \text{ where } Y_{k-r}(k) = i\}$$
$$= 0 \quad \text{if the above set is empty.}$$

Using the optimal rule, there is a positive probability of stopping at time k if $Y_{k-r}(k) \leq b_{k-r}$. If $Y_{k-r}(k) > b_{k-r}$, go on.

These b_i induce a stopping rule on the truncated $p^1(\cdot)$ problem. Call this rule τ_n^{1*} and the corresponding risk V_n^{1*}. The rule τ_n^{1*} is to stop the first time i where $Y_i(i) \leq b_i$. Upon stopping, solicit in order of increasing rank.

Let $\tau_n^{2,M}$ and $V_n^{2,M}$ be the optimal procedure and risk for the truncated $p^2(\cdot)$ problem.

By definition $\tau_n^{1*} \leq \tau_n^{2,M}$. Since $q(i) \leq q(M) < \infty$, $S_i(k) - S_i(k-r) \to 0$ uniformly in i, k as $n \to \infty$ so we have proved (4.23) if $P(\tau_n^{2,M} = \tau_n^{1*} + r) \to 1$ as $n \to \infty$.

Assume $\tau_n^{1*} = x$. By definition of C^k and b_{k-r}, $\tau_n^{2,M} > x+r$ only if a candidate of rank $\leq M$ arrives at $x+1$, $x+2, \ldots,$ or $x+r$. That is,

$$(4.24) \qquad P\left(\tau_n^{2,M} = x+r \mid \tau_n^{1*} = x\right) \geq \left(\frac{x+1-M}{x+1}\right)\left(\frac{x+2-M}{x+2}\right) \cdots \left(\frac{x+r-M}{x+r}\right).$$

Let $a(M,n) = \min\{i \text{ such that } b_i > 0\}$. By arguments similar to these in Chapter II we have $a(M,n) \to \infty$ as $n \to \infty$. We also have $\tau_n^{1*} \geq a(M,n)$ so we can conclude

$$P\left(\tau_n^{2,M} = \tau_n^{1*} + r\right) = \sum_{x \geq a(M,n)} P\left(\tau_n^{1*} = x\right) P\left(\tau_n^{2,M} = x+r \mid \tau_n^{1*} = x\right)$$
$$\geq \left[\frac{a(M,n) - M}{a(M,n)}\right]^r \longrightarrow 1 \quad \text{as } n \to \infty .$$

So we conclude

$$\underline{\lim}\ V_n^2(p^*) \geq \underline{\lim}\ V_n^{2,M}(p^*) = \underline{\lim}\ V_n^{1*} \geq \underline{\lim}\ V_n^{1,M}(p^*) = f_{p^*}^M(0) \to f_{p^*}(0)$$

and our theorem is proved.//

So we have obtained the limiting risk for all p(•) of the form (4.9). The limiting risk is given by $f_{p^*}(0)$ where

$$f'_{p^*}(x) = \frac{1}{x} \sum_{k=1}^{\infty} [f_{p^*}(x) - T_k(x)]^+$$

$$\text{subject to } f_{p^*}(1) = \sum_{i=1}^{\infty} q(i)\ p^*(1-p^*)^{i-1} \qquad p^* \neq 0$$

$$= \sup q(i) \qquad p^* = 0.$$

The case r = 0 gives procedures which are asymptotically optimal.

4. General Asymptotic Results

From the special class (4.9), we deduce the general asymptotic results as follows:

Thm. 4.7 For any given p(•) and q(•), $\lim_{n\to\infty} V_n = f_{p^*}(0)$

where $p^* = \lim_{n\to\infty} p(n)$, V_n = minimal risk for p(•), q(•), and given n,

$$f'_{p^*}(x) = \frac{1}{x} \sum_{k=1}^{\infty} [f_{p^*}(x) - T_k(x)]^+ \quad \text{subject to}$$

$$f_{p^*}(1) = \sum_{i=1}^{\infty} q(i)\ p^*(1-p^*)^{i-1} \qquad p^* \neq 0$$

$$= \sup q(i) \qquad p^* = 0.$$

Proof:

Let $p^1(i) = 1 \qquad i = 0$

$\qquad\qquad p^* \qquad i = 1,2,3,\ldots$

Then $p(i) \geq p^1(i)$ for all i so if V_n^1 is the minimal risk associated with $p^1(\cdot)$,

(4.25) $$\overline{\lim}\ V_n \leq \overline{\lim}\ V_n^1 = f_{p^*}(0).$$

Let $\epsilon > 0$ be given. Choose $N(\epsilon)$ so big that $n > N(\epsilon)$ implies $p(n) < p^* + \epsilon$.

$$\text{Let } p^2(i) = 1 \qquad i = 0,1,2,\ldots, N(\epsilon)$$
$$= p^* + \epsilon \qquad i = N(\epsilon) + 1, N(\epsilon) + 2, \ldots$$

Then $p(i) \leq p^2(i)$ for all i so if V_n^2 is the minimal risk associated with $p^2(\cdot)$,

$$\underline{\lim}\ V_n \geq \underline{\lim}\ V_n^2 = f_{p^*+\epsilon}(0). \tag{4.26}$$

Since $f_p(0)$ is right continuous and (4.26) holds for all $\epsilon > 0$,

$$\underline{\lim}\ V_n \geq f_{p^*}(0) \geq \overline{\lim}\ V_n$$

and we have equality.///

In general, the exact solution to the finite n-girl problem has been shown to be unobtainable. In spite of this, we have obtained the limiting value for the minimal risks. Many procedures will actually obtain this limiting value. We will present a particular solution.

<u>Cor. 1</u> The cutoff point rule given by the recursion equation (4.14)

i.e.,
$$C(r-1) = \frac{1}{r} \sum_{k=1}^{r} \min\left(S_r(k), C(r)\right)$$

where
$$S_r(k) = \sum_{i=1}^{k-1} p^*(1-p^*)^{i-1} Q(r,i) + (1-p^*)^{k-1} Q(r,k)$$

and
$$Q(r,i) = \sum_{\ell=i}^{n-(r-i)} q(\ell) \frac{\binom{\ell-1}{i-1}\binom{n-\ell}{r-i}}{\binom{n}{r}}$$

is asymptotically optimal in the sense that $|V_n - C(0)| \to 0$.

<u>Proof</u>:

We have shown $C(0) \to f_{p^*}(0)$ as $n \to \infty$ in section 3 of this chapter and our theorem states $V_n \to f_{p^*}(0)$, as $n \to \infty$ therefore $|V_n - C(0)| \to 0$ as $n \to \infty$.///

Smith [18] solved the best choice problem. We can easily show that our results agree with his general results.

Smith showed P(best choice) $\to e^{-(1-p^*)}$ and the limiting optimal procedure is a cutoff procedure with $r_n/n \to e^{-(1-p^*)}$.

Let $q(1) = 0$

$q(i) = 1 \quad i = 2,3,4,\ldots$.

Then,

$$R_1(x) = 1 - x \quad ,$$

$$R_k(x) = 1 \qquad k = 2,3,4,\ldots,$$

$$T_1(x) = 1 - x \quad ,$$

$$T_k(x) = p^*(1-x) + 1 - p^* \quad k = 2,3,4,\ldots,$$

$$f_{p^*}(1) = 1 - p^* \leq T_k(x) \text{ for } k = 2,3,4,\ldots, .$$

So t_1 is the only cutoff point $\neq 1$ and our differential equation becomes $f'_{p^*}(x) = \frac{f_{p^*}(x)}{x} - \frac{1-x}{x}$ on $[t_1, 1]$.

This is the same as

$$\left(\frac{f_{p^*}(x)}{x}\right)' = \frac{x-1}{x^2} .$$

Integrating gives

$$\frac{f_{p^*}(1)}{1} - \frac{f_{p^*}(t_1)}{t_1} = \int_{t_1}^{1} \frac{x-1}{x^2}\,dx = 1 - \ln(t_1) - \frac{1}{t_1}$$

and since $f_{p^*}(0) = f_{p^*}(t_1) = T_1(t_1) = 1 - t_1$,

$$t_1 = e^{-(1-p^*)} = \lim r_n/n \qquad \text{and}$$

$$f_{p^*}(0) = 1 - e^{-(1-p^*)} = 1 - P(\text{best choice}).$$

4. The Rank Problem

Let us now consider the rank problem $q(i) = i$. By applying the approximations considered in proposition 4.3, we can show that $f_{p*}(0)$ is both right and left continuous, as the solution will indicate.

Computations yield

$$R_k(x) = k/x$$

$$T_k(x) = \frac{1-(1-p^*)^k}{p^*}\left(\frac{1}{x}\right) \qquad p^* \neq 0$$

$$= \frac{k}{x} \qquad p^* = 0$$

$$f_{p*}(1) = \frac{1}{p^*} \qquad p^* \neq 0$$

$$= \infty \qquad p^* = 0 .$$

Since $f_{p*}(1) > T_k(1)$ for all k, there are an infinite number of cutoff points. On $[\alpha_i, \alpha_{i+1}]$ we have

$$f'_{p*}(x) = \frac{i}{x} f_{p*}(x) - \sum_{k=1}^{i} \frac{1-(1-p^*)^k}{p^* x^2} \qquad p^* \neq 0$$

$$= \frac{i}{x} f_{p*}(x) - \sum_{k=1}^{i} \frac{k}{x^2} \qquad p^* = 0.$$

Ignoring $p^* = 0$ for the moment, we get

$$\left[\frac{f_{p*}(x)}{x^i}\right]' = \frac{f_{p*}(x)'}{x^i} - \frac{i\, f_{p*}(x)}{x^{i+1}} = \frac{1-p^*(i+1)-(1-p^*)^{i+1}}{p^{*2}\, x^{i+2}} .$$

Integrating from α_i to α_{i+1} and using $f_p{*}(\alpha_k) = T_k(\alpha_k)$ we obtain

$$\alpha_i^{i+1} = \left[\frac{1-(1+ip^*)(1-p^*)^i}{1-(1+(i+1)p^*)(1-p^*)^{i+1}}\right]\alpha_{i+1}^{i+1} \qquad p^* \neq 0$$

$$= \left[\frac{i}{i+2}\right]\alpha_i^{i+1} \qquad p^* = 0.$$

But $\alpha_i \to 1$, see Mucci [13] so

$$\alpha_1 = \prod_{i=1}^{\infty} \left[\frac{1-(1+ip^*)(1-p^*)^i}{1-(1+(i+1)p^*)(1-p^*)^{i+1}}\right]^{\frac{1}{i+1}} \qquad p^* \neq 0$$

$$= \prod_{i=1}^{\infty} \left[\frac{i}{i+2}\right]^{\frac{1}{i+1}} \qquad p^* = 0.$$

Since $f_{p^*}(0) = f_{p^*}(\alpha_1) = 1/\alpha_1$,

$$f_{p^*}(0) = \prod_{i=1}^{\infty} \left[\frac{1-(1+(i+1)p^*)(1-p^*)^{i+1}}{1-(1+ip^*)(1-p^*)^i}\right]^{\frac{1}{i+1}} \qquad p^* \neq 0$$

$$= \prod_{i=1}^{\infty} \left[\frac{i+2}{i}\right]^{\frac{1}{i+1}} \qquad p^* = 0.$$

Taking the second derivative of $f_{p^*}(0)$ and using the fact that the first derivative is negative, it is possible to show $f_{p^*}(0)$ is convex in p^*. Some values and a graph is given.

Table 4.1 Some limiting values for the rank problem with recall.

p^*	0.0	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$f_{p^*}(0)$	3.87	3.34	3.01	2.53	2.19	1.92	1.70	1.52	1.36	1.22	1.11	1.00

Of course, for a specified recall function $p(\cdot)$ many asymptotically optimal rules exist. In general, the exact optimal rule takes an extremely complex form, requiring far too many steps for computation. What is given, however, is a relatively simple rule providing a good approximation to the exact risk and giving the means to examine the limiting risk.

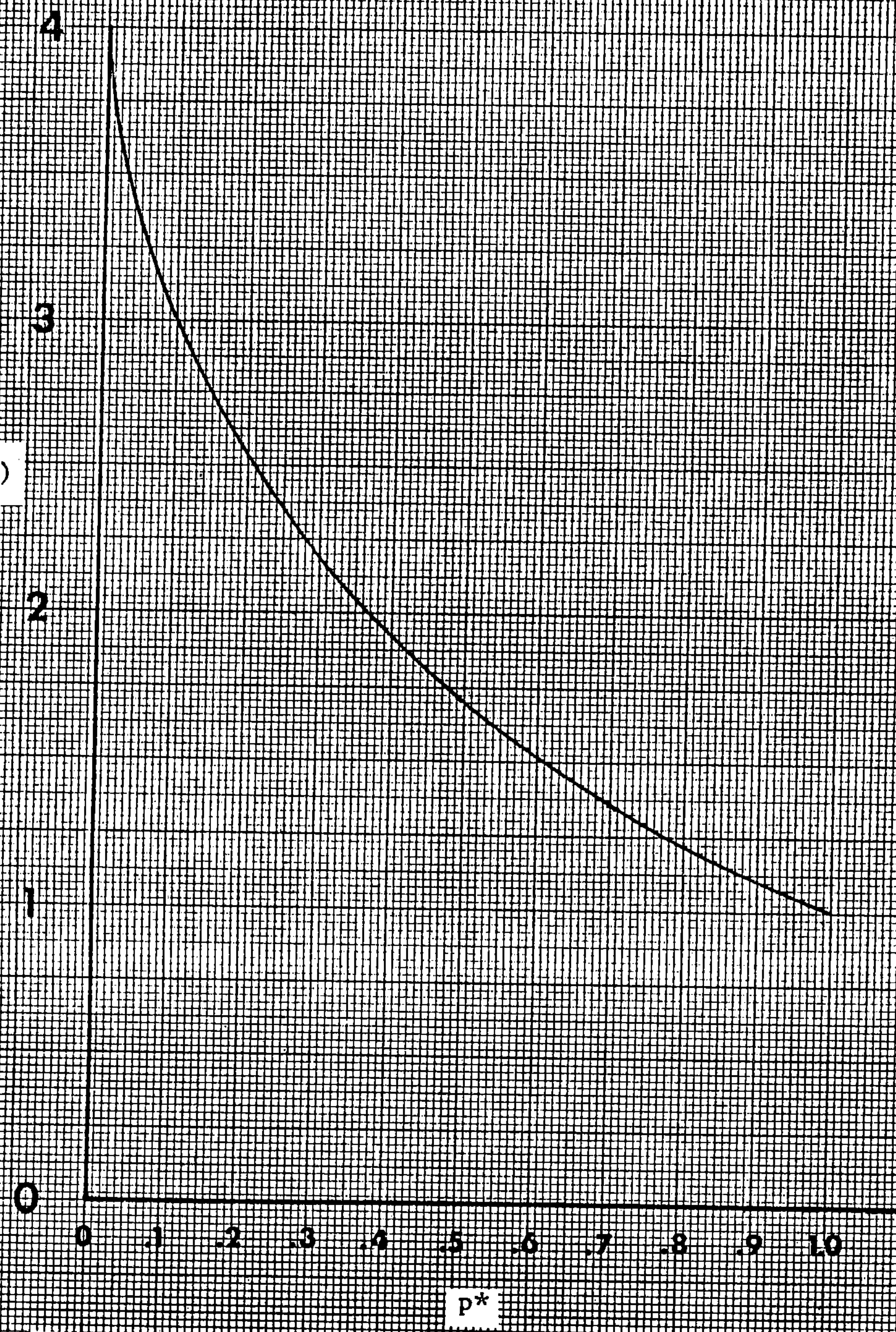


Figure 4.1 The rank problem with recall.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Bartosznski, R. and Govindarajulu, Z., The secretary problem with interview cost, Technical Report No. 71, University of Kentucky, ONR Technical Report No. 6, (1974).

[2] Cayley, A., Mathematical questions and their solutions, Educational Times, 22, (1875) pp. 18-19.

[3] Chow, Y. S., Moriguti, S., Robbins, S., Samuels, S. M., Optimal selection based on relative rank, Israel Journal of Mathematics, Vol. 2, No. 2, (1964) pp. 81-90.

[4] Dynkin, E. B. and Yushkevich, A. A., Markov Processes Theorems and Problems, Plenum Press, New York, (1969) pp. 87-145.

[5] Fox, John H. and Marnie, L. Gerald in Martin Gardner's column: Mathematical games, Scientific American, 202, (1960) pp. 150, 153.

[6] Gianini, Jacqueline, Problems of optimal selection based on relative rank, Technical Report No. 400, Purdue University (1975).

[7] Gianini, Jacqueline and Samuels, Stephen M., The infinite secretary problem, Annals of Probability, Vol. 4, No.3, (1976) pp. 418-432.

[8] Gilbert, J. P. and Mostellar, F., Recognizing the maximum of a sequence, Journal of the American Statistical Association, Vol. 61, (1966) pp. 35-73.

[9] Govindarajulu, Z., The secretary problem: optimal selection with interview cost, Technical Report No. 82, University of Kentucky, ONR Technical Report No. 11, (1975).

[10] Gusein-Zade, S. M., The problem of choice and the optimal stopping rule for a sequence of independent trials, Theory of Probability and its Applications, 11, (1966) pp. 472-476.

[11] Lindley, D. V., Dynamic programming and decision theory, Applied Statistics, 10, (1961) pp. 47-51.

[12] Moser, Leo and Pounder, J. R., in Martin Gardner's column: Mathematical games, Scientific American, 202, (1960) pp. 178, 181.

[13] Mucci, Anthony G., Differential equations and optimal choice problems, Annals of Statistics, Vol. 1, No. 1, (1973) pp. 104-113.

[14] Mucci, Anthony G., On a class of secretary problems, The Annals of Probability, Vol. 1, No. 3, (1973) pp. 417-427.

[15] Presman, E. L. and Sonin, I. M., The best choice problem for a random number of objects, Theory of Probability and its Applications, Vol. XVII, No. 4, (1972) pp. 657-668.

[16] Rasmussen, W. T. and Robbins, H., The candidate problem with unknown population size, Journal of Applied Probability, 12, (1975) pp. 692-701.

[17] Rubin, H., The "secretary" problem, Annals of Mathematical Statistics, 37, (1966) p. 544. (abstract)

[18] Smith, M. H., On the asymptotic properties of a best choice problem with backward solicitation, JAP, to appear Vol. 14, 1977.

[19] Smith, M. H. and Deely, J. J., A secretary problem with finite memory, Journal of the American Statistical Association, Vol. 70, (1975) pp. 357.

[20] Yang, Mark C. K., Recognizing the maximum of a random sequence based on relative rank with backward solicitation, Journal of Applied Probability, Vol. 11, No. 3, (1974) pp. 504-512.