THE UNIVERSITY OF CALGARY

THE RELATION BETWEEN THE LENGTH OF A SEQUENTIAL PROCESS AND THE PROBABILITY OF WINNING

by

AARON RON RENERT

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS AND STATISTICS

CALGARY, ALBERTA
MARCH, 1993

ISBN 0-315-83241-X

THE UNIVERSITY OF CALGARY
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a dissertation entitled "*The relation between the length of a sequential process and the probability of winning*", submitted by **Aaron Ron Renert** in partial fulfillment of the requirements for the degree of Master of Science.

Supervisor, Dr. E.G. Enns
Dept. of Mathematics & Statistics

Dr. P.F. Ehlers
Dept. of Mathematics & Statistics

Dr. R. Schulz
Faculty of Management

25th March, 1993

ABSTRACT

N independent and identically distributed random variables are sampled sequentially from a known distribution. When trying to locate the maximum of the sequence, there seems to be a very simple relation between the probability of finding the maximum and the number of variables sampled. It will be shown how this relation holds true under different information sources when optimal policy is used.

ACKNOWLEDGMENTS

I would like to thank my supervisor Dr. E. G. Enns for his continuous support and guidance throughout the writing of the thesis. Many thanks to Dr. P. F. Ehlers and to Thi Dinh for their insightful remarks. Last but not least, I thank my family for backing me all along.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I
INTRODUCTION

1.1 Definitions

Let us define a standard "secretary problem" as follows:

1.1.1 n independent and identically distributed random variables are sampled sequentially from a known continuous distribution (the continuity guarantees that there are no "ties") and all $n!$ permutations are equally likely.

1.1.2 n, the total number of random variables available, is known.

1.1.3 After each random variable is drawn, the player must either choose it, thereby ending the game, or reject it permanently.

1.1.4 The player is satisfied with nothing but the very best (a payoff of 1 if the largest random variable is selected, a payoff of 0 otherwise).

1.1.5 The decision at every stage is based only on the information provided from some source to the player regarding the random variable drawn.

1.1.6 If the first $n-1$ random variables are rejected, the last one is automatically chosen.

The obvious application to choosing the best applicant for a job gives the problem its name, although in the literature the problem is referred to by other names such as dowry, marriage, and beauty contest.

1.2 Terminology and notations

1.2.1 The distribution of sampled variables

The distribution of the random variables sampled as specified by the general secretary problem is arbitrary. However, for simplicity we will assume throughout that the distribution is the uniform distribution on $[0,1]$. We may make this assumption since the distribution is known (by 1.1.1). This implies that the cumulative probability function,

$$F(y) = P(Y \leq y)$$

is known as well. Therefore there exists an obvious order preserving transformation that assigns to each sampled random variable Y_i its cumulative probability,

$$X_i = F(Y_i)$$

Under this transformation the probability of selecting the maximum of the Y_i's is identical to selecting the maximum of the corresponding X_i's.

1.2.2 Optimal strategies

This paper discusses optimal strategies for different types of secretary problems. Since the objectives may vary depending upon the type of problem at hand, an optimal strategy, S^*, is one that maximizes the probability of reaching the given objective. In general, the only criterion of S^* used to decide whether to choose the ith observation, once it is drawn, is the condition:

$$P(\text{choosing } X_i \text{ and winning}) >$$
$$P(\text{not choosing } X_i\text{, playing optimally thereafter and winning}) \qquad (1.1)$$

If that condition is met, X_i should be chosen.

For the standard secretary problem described in 1.1, the objective (as outlined in 1.1.4) is to find the maximum of the sampled values. S^* will therefore be a strategy that locates $\max\{X_1 \ldots X_n\}$ with the highest probability.

Given any information source, γ, and the number of observations available, n, then clearly there exists an optimal strategy, S^*, that maximizes the probability of correctly selecting the largest observation. In the case that S^* is not unique, the following tie-breaker criterion will be used.

1.2.3 Tie-breaker criterion for choosing among optional strategies

Suppose $S_1^*, S_2^* \cdots S_j^*$ are all optimal strategies. The preferred optimal strategy then is the one that minimizes $E(K_n)$. Here K_n is the location of the stopping variable when the sample is of size n.

The process of sequential sampling usually involves costs that are directly related to the number of times we have sampled. In light of this, the above criterion is readily understood.

1.2.4 Information sources

Four types of information source will be discussed:

2.4.1 Extreme no-information source (γ_E) — The source gives no information at all regarding the observation sampled.

2.4.2 No information source (γ_R) — The source reveals the rank of each observation only as it is being drawn (rank among the ones observed so far).

2.4.3 Full information source (γ_F) — The source reveals the actual value of X_i as it is being sampled.

2.4.4 Partial information source (γ_P) — All other sources fall into this category.

The following notations will be used throughout the paper:

X_i — the ith observation.

$X_{\max}$ — the largest observation.

M_n — the event that the maximum of a sample of size n is chosen.

K_n — the number of random variables sampled (including the one selected) in a sample of size n.

C_{X_i} — the event that the ith observation is chosen.

M_{X_i} — the event that the ith observation is the maximum.

γ — the information source available to the decision maker. A source may give no information whatsoever, partial information of some sort or full information.

In order to avoid cumbersome notations, the information source and the strategy will always be clearly specified. This will alleviate the need to include them as part of the notation. Additional notations will be defined later as they are needed.

1.3 Variations on the standard secretary problem

The standard secretary problem, as described in 1.1, is concise and is amenable to precise theoretical analysis. However, most real life applications do not conform to the restrictions of the standard secretary problem. Variations on the problem can be obtained by relaxing or altering one or more of the restrictions 1.1.1 through 1.1.6. Let us now review the list of restrictions once more and mention the common variations of each.

1.3.1 The distribution of the X_i's may be:

— continuous or discrete. In the case of discrete distributions ties may occur. When ties occur, the $n!$ permutations are not all equally likely.

— fully specified, partially specified or not specified at all. Partial specification may arise for instance in a case where X_i is exponentially distributed with some unknown parameter λ.

1.3.2 The total number of observations may be known (as in the standard case). Alternatively, the total number of observations may follow some distribution, $p_i = P(n = i),\ i = 1\ldots k$. If N does follow a certain distribution, this distribution may be specified in full, partially, or not at all.

1.3.3 In the standard case, no back solicitation is allowed. A whole class of problems arises if the player is allowed to backtrack and choose an observation that had previously been rejected. Back solicitation usually has an associated cost which may take the form of a fixed amount. Alternatively, the cost may take the form of restrictions on the random variables that may be back sampled.

1.3.4 The objective of the game, to pick "nothing but the best", may not be appropriate in many practical scenarios. Trying to minimize the expected value (or some utility functions that depend on the rank or the value of the observations chosen) can prove much more suitable. Frequently the objective of picking "nothing but the best" is replaced by the objective of picking "one of the best k observations". For instance, if $k = 2$, the object now becomes that of picking "nothing but the best or the second best."

1.3.5 Different information sources were already discussed in (1.2.4). The most commonly used sources are the no-information and full-information sources.

1.3.6 If back solicitation is permitted, reaching the last sample does not imply it must be selected.

By altering requirements 1.1.1-1.1.6 we can construct numerous variations on the standard secretary problem. A typical one may read as follows:

> A decision maker gets N offers sequentially, where $N \sim \text{Poisson}(3)$ and the offers are $\text{Normal}(12000, 2000)$. As every offer comes along the decision maker decides whether to accept or not. He can stop the process at any stage. Back solicitation is allowed but costs 3000 for every backtracking step. The objective of the game is to maximize the expected payoff.

1.4 Applications

A large variety of real life problems can be regarded as instances of the secretary problem. We shall consider two examples.

1.4.1 The house seller's problem

A person is offered a job in another city and he wishes to sell his house. Time is of the essence since he has only two months to relocate. If the person does not meet the selling deadline, he loses $1,140 per month in property maintenance costs and potential rent income. A short study of house sales in the same neighbourhood reveals that:

4.1.1 The number of offers that sellers get during the first month of listing is Poisson distributed with a mean of 3.6.

4.1.2 The number of offers that sellers get during the second month of listing is Poisson distributed with a mean of 1.6.

4.1.3 The size of the offers in the first month of listing follows the normal distribution $\text{Normal}(.95A,.012A)$. In this case A is the appraised value of the house as provided by an independent appraiser. In our example $A = \$154,900$.

4.1.4 The size of the offers in the second month of listing follows the normal distribution $\text{Normal}(.935A,.012A)$.

What should be the selling strategy if the seller wishes to maximize the expected payoff?

(*Remark: Note that this instance of the house seller's problem is based on real data provided by Royal LePage for house sales in West Thorncliffe and Upper Northaven in the summer of 1992. This data only encompasses sales by realtors, and does not include private sales by owners.*)

This is clearly a secretary type problem. However, it deviates from the restrictions imposed by the standard secretary problem as follows:

— The offers are not independent, since every buyer can obtain information about previous offers. Such information is most likely to affect the buyer's offer.

— The number of offers is not fixed, but follows a certain given distribution.

— The assumption of "no back sampling" still holds. Since in this example the term "offer" refers to the final offer of a prospective buyer, once an offer is rejected, the buyer departs for good.

— The objective of the game changes from "nothing but the best" to maximizing expected value.

— Full information is given regarding the distribution of the offers and the exact value of each offer.

— There is a fixed cost ($1,140 per month) that should be taken into consideration when deriving the optimal strategy.

This example demonstrates how a fairly simple real life application, once analyzed, becomes a complicated secretary type problem.

1.4.2 Kepler's wife selection

The entertaining method employed by Johannes Kepler in choosing his second wife provides us with the next example. This story is discussed at length in [9] and summarized in [4]. After Johannes Kepler's first wife died of cholera, the great German astronomer resolutely ventured to find an appropriate substitute. In two years he interviewed no fewer than 11 women. After careful consideration he decided to marry the fifth prospect. His friends objected strongly to his choice (the prospect was an orphan of a lower social rank) and convinced him to marry the fourth candidate. Kepler indeed proposed to candidate #4 only to find out that he had waited too long. At that time candidate #4 was no longer interested in his advances. Kepler went on to propose marriage to his initial choice, candidate #5. She gladly accepted the offer, bore him seven children, and ran a neat and efficient household. Subsequent letters, written by Kepler to his friends in which the new wife was portrayed very favourably, suggest that the choice was indeed successful.

Kepler's process of finding a spouse is clearly one of sequential sampling. Again we shall examine how a practical application conforms to the conditions of the theoretical secretary problem.

— The marriage candidates are not independent of one another. Failure with one candidate probably lead Kepler to avoid others who were of the same "type".

— The total number of candidates is very difficult to determine. This number is influenced by factors that are too numerous to specify.

— The information available is partial because no one may know everything there is to know about the virtues of the opposite sex. Nevertheless, male arrogance is such that every man thinks he knows quite a lot.

— The importance of the decision leads one to believe that the goal of the game is "nothing but the best."

— The last candidate is not necessarily chosen automatically if the first n-1 candidates are rejected. In some cases, staying single may be the preferred approach.

In the next chapter some of the better known extensions will be discussed, along with their solutions.

1.5 Historical review

The earliest known variation of the secretary problem appeared in the *British Educational Times* in 1875. It was submitted by none other than the distinguished English mathematician Arthur Cayley. The problem deals with a lottery of n tickets representing $1\ldots n$ pounds. A person draws a ticket, looks at it, and decides whether he wishes to keep it or draw again (out of the remaining $n-1$ tickets), drawing in all not more than k times. His payoff is the value of the last ticket drawn and he wishes to maximize his expected payoff. Note that although Cayley's problem and the standard secretary problem are highly similar there is one important difference: Cayley's payoff is some numerical value depending on the ticket selected, which clearly violates condition 1.1.4.

Ever since then, different variations of the secretary problem have appeared frequently in the literature. The origins of the standard secretary problem itself are somewhat obscure. Gleason posed the problem in 1955, mentioning he heard it from someone else. Gardener (1960) was most likely the first to publish a statement of the no-information case, attributing it to Marnie and Fox (1958). A solution to the problem was given in the March 1960 issue of Scientific American by Moser and Pounder.

Although statisticians started to show interest in the problem in the early to mid-1950s, it was only later, in 1960-1961, that it made its way to scientific journals. In his 1961 paper, Lindley not only solves the no-information case but also considers minimizing the expected rank of the applicant chosen. Bissinger and Siegel (1963) posed the no-information case once more, this time considering the $n = 1000$ case. A year later a solution was proposed by Bosch and 12 others.

A real breakthrough occurred in 1966 due to the outstanding basic paper by Gilbert and Mosteller. In "Recognizing the maximum of a sequence", the two start by giving a comprehensive summary of the problem and then derive elegant solutions for the no-information case as well as the full-information case. They also: extend and solve the no-information case where r choices are allowed; consider the problem where the goal is to obtain the best or second best; solve the "full information case" under the minimum rank criterion; and also discuss asymptotic theories (n approaches infinity) for all the above cases. This paper, more than any other, is the foundation of what grew into a wide field of study within probability optimization. Ever since then, numerous extensions of the problem have been discussed. Although the literature now contains hundreds of papers, it seems that the field continues to grow rapidly.

CHAPTER II
ENNS'S CONJECTURE IN THE SIMPLE CASES

2.1 The extreme no-information case

Recall that γ_E is the source that gives the player no information whatsoever about X_i as it is being drawn. This is a trivial case, in which for all strategies S, the probability of correctly locating $X_{\max}$ is:

$$P(M_n) = \sum_{i=1}^{n} P(M_{X_i} \cap C_{X_i})$$

$$= \sum_{i=r}^{n} P(C_{X_i} | M_{X_i}) P(M_{X_i})$$

$$= \frac{1}{n} \sum_{i=1}^{n} P(C_{X_i}) = \frac{1}{n}$$

Each strategy is as "optimal" as any other and yields $P^*(M_n) = \frac{1}{n}$. Using the tie-breaking criterion discussed in (1.2) we select the strategy that chooses X_1 with probability 1. Under this selection of S^*, we have:

$$P^*(M_n) = \frac{1}{n} \quad \text{and} \quad E^*(K_n) = 1$$

Before we go on to the next case note that:

$$nP^*(M_n) = E^*(K_n) \tag{2.1}$$

2.2 The classical secretary problem

This case is by far the most discussed. Here the player is under the γ_R source which provides him with the rank of the observation drawn (among the observations drawn so far) as he proceeds.

2.2.1 Deriving the Optimal Strategy

Let us think of ourselves as the player and find the form of the optimal strategy. Based on (1.3), once we are given the rank of the ith observation, it should only be chosen if:

$$P(\text{choosing } X_i \text{ and winning}) >$$
$$P(\text{ not choosing } X_i, \text{ playing optimally thereafter and winning }) \quad (2.2)$$

Define a "candidate" to be an observation, the value of which exceeds the values of all earlier draws. Clearly, the ith draw should be considered only if it is a candidate, or otherwise the game would be lost.

Observe that:

$$P\left(C_{X_i} \cap M_{X_i}\right) = P\left(X_{\max} = X_i \mid X_i = \max\{X_1 \ldots X_i\}\right)$$
$$= P\left(X_{\max} \in \{X_1 \cdots X_i\}\right) = \frac{i}{n}$$

which is a strictly increasing function of i .

The right hand side of inequality (2.2) is a decreasing function of i, since we can always get to a later point in the sequence and then use whatever strategy is available. In other words, being "young" in the game does not rule out any strategies that can be employed later on. Consequently, the optimal strategy takes the form of passing the first $r-1$ draws, for some r, and choosing the first candidate thereafter. Such a strategy will be referred to as an $s(r)$ rule.

2.2.2 Deriving $P^*(M_n)$

S^* in this case is an $s(r)$ rule. Under such a scheme, the probability of winning is given by:

$$P(M_n) = \sum_{i=r}^{n} \left(C_{X_i} \cap M_{X_i}\right)$$

$$= \sum_{i=r}^{n} P\left(C_{X_i} \middle| M_{X_i}\right) P\left(M_{X_i}\right)$$

$$= \frac{1}{n} \sum_{i=r}^{n} \frac{r-1}{i-1} \tag{2.3}$$

The last step of the derivation is true since $P\left(M_{X_i}\right) = \frac{1}{n}, \forall i$, and given that X_i is the maximum of the whole sequence, it will be selected only if $\max\{X_1 \cdots X_{i-1}\}$ falls within the first $r-1$ observations (or else we will never get to sample X_i). The probability of the above occurring is $\frac{r-1}{i-1}$ and (2.3) follows.

For any given n, we can calculate r^*, the value that maximizes $P(M_n)=\frac{1}{n}\sum_{i=r}^{n}\frac{r-1}{i-1}$. Optimal values of r^*, together with their corresponding $P^*(M_n)$ are presented in Table 2.1 for various values of n. It is interesting to note that $r^*=5$ when $n=11$. Recall that in Kepler's case (1.4.2) 11 candidates were interviewed and the fifth one was chosen. If the great astronomer were to use the above model, he would have ended up with the same wife without having to interview the last six candidates at all.

n	r^*	$P^*(M_n)$	n	r^*	$P^*(M_n)$
1	1	1.00000	15	6	0.38940
2	1,2	0.50000	20	8	0.38420
3	2	0.50000	30	12	0.37865
4	2	0.45833	40	16	0.37574
5	3	0.45833	50	19	0.37427
6	3	0.42777	60	23	0.37320
7	3	0.41428	70	27	0.37239
8	4	0.40982	80	30	0.37185
9	4	0.40595	90	34	0.37142
10	4	0.39869	100	38	0.37104
11	5	0.39841	1000	369	0.36819
12	5	0.39551	∞	n/e	$1/e$
13	6	0.39226			
14	6	0.39171			

Table 2.1

Optimal r^* values together with optimal probabilities of winning for the classical secretary problem.

2.2.3 Deriving $E^*(K_n)$.

Let us derive the expected location of the stopping variable. Since an $s(r)$ rule is used, we get:

$$E(K_n) = \sum_{i=1}^{n} (i) P(C_{X_i})$$

$$= \left[\sum_{i=1}^{n-1} (i) P(C_{X_i}) \right] + nP(C_{X_n})$$

$$= \sum_{i=1}^{n-1} (i) P(\text{no choice was made among } \{X_1 \cdots X_{i-1}\} \text{ and } X_i \text{ is a candidate})$$
$$+ (n) P(\text{no choice was made among } \{X_1 \cdots X_{n-1}\})$$

$$= \left[\sum_{i=r}^{n-1} (i) \left(\frac{r-1}{i-1} \right) \left(\frac{1}{i} \right) \right] + n \left(\frac{r-1}{n-1} \right)$$

$$= \left[\sum_{i=r}^{n} \left(\frac{r-1}{i-1} \right) \right] + (r-1) \qquad (2.4)$$

To calculate the value of $E^*(K_n)$, all one must do is determine the value of r^* and substitute it into (2.4).

Combining (2.3) and (2.4), it follows immediately that:

$$nP(M_n) = E(K_n) - (r-1) \qquad (2.5)$$

Recall that in the extreme no-information case, the two quantities $P(M_n)$ and $E(K_n)$ were also very closely related (see 2.1).

2.3 Enns's Conjecture and Its Immediate Extensions

E.G. Enns was the first to note that in secretary-type problems there seems to be a close relation between $P(M_n)$ and $E(K_n)$. In his unpublished paper "The Role of Information in a Sequential Decision Problem" (1974), he derives $P(M_n)$ and $E(K_n)$ for various information sources under optimal strategies, points out the simple relationship between the two quantities, and poses the following conjecture:

Enns's Conjecture (EC) :

> **When trying to locate the maximum of a sequence of length n under any given information source, the following holds true when playing optimally:**
>
> $$E^*(K_n)-(r^*-1)=nP^*(M_n).$$
>
> **Here, r^* is the smallest integer satisfying $P\left(C_{X_{r^*}}\right)\neq 0$ when playing optimally.**

As will be seen in later chapters, most information sources provide us with enough information regarding X_1 so that $P\left(C_{X_1}\right)\neq 0$ when playing correctly. In such cases, $r^*=1$ and the conjecture reduces to the simple relation $E^*(K_n)=n\,P^*(M_n)$.

2.3.1 EC in the extreme no-information case

In the extreme no-information case, discussed in (2.1), EC holds for the optimal strategy with $r=1$. However, EC also holds for other strategies that are not optimal. In particular, every strategy that chooses X_i with probability 1, yields:

$$P(M_n) = \frac{1}{n} \text{ and } E(K_n) = i$$

regardless of the value of i chosen. EC follows, since here $r^* = i$.

It is important to note here that EC may also be satisfied in cases where a non-optimal strategy is being used. It follows that optimality is clearly not a necessary condition for EC to hold.

2.3.2 EC in the classical secretary problem.

The following theorem is a direct consequence of (2.5):

Theorem 1

Whenever an $s(r)$ rule is used for finding the maximum of a sequence of length n, the following holds:

$$E(K_n) - (r-1) = nP(M_n)$$

where r is the smallest integer satisfying $P\left(C_{X_r}\right) \neq 0$.

Looking at the derivation of (2.5) we can see that the above theorem is true regardless of the value of r chosen. In particular, given a classical secretary problem with n observations, we can calculate r^* (the r that maximizes (2.3)), and $P^*(M_n)$, the probability of winning under S^*. The expected location of the stopping variable can then be calculated by using:

$$E^*(K_n) = nP(M_n) + (r^* - 1) \tag{2.6}$$

An intuitive argument can explain why (2.5) holds for any choice of r in the classical case. Consider, a new "game" defined as follows: an $s(r)$ strategy is used for finding $X_{\max}$ (just as before), but if the strategy ever leads us to choose X_n, this event will be considered as an instant win, regardless of whether X_n is the largest of the whole sequence or not. We will be using the superscript ^{O} for the "old" $s(r)$ procedure, and ^{N} for the modified one.

Clearly, $E^O(K_n) = E^N(K_n)$, since the choice of the stopping variable is based on the same $s(r)$ rule in both cases. However, $P^O(M_n) < P^N(M_n)$ because the effect of the new game is to take the losing event of $\left(C_{X_n} \cap \overline{M_{X_n}}\right)$ and to redefine it as a win.

Therefore,

$$P^N(M_n) = P^O(M_n) + P\left(C_{X_n} \cap \overline{M_{X_n}}\right)$$

and since

$$P\left(C_{X_n} \cap \overline{M_{X_n}}\right) = P\left(C_{X_n} \middle| \overline{M_{X_n}}\right) P\left(\overline{M_{X_n}}\right)$$

$$= \left(\frac{r-1}{n-1}\right)\left(\frac{n-1}{n}\right)$$

$$=\frac{r-1}{n}$$

we get:

$$P^{O}(M_n)=P^{N}(M_n)-\left(\frac{r-1}{n}\right) \tag{2.7}$$

Now note that in the new game:

$$P^{N}\left(M_n \middle| C_{X_i}\right)=\frac{i}{n} \qquad \forall i \geq r \tag{2.8}$$

Multiplying both sides of (2.8) and summing, we get:

$$\sum_{i=r}^{n} P^{N}\left(M_n \middle| C_{X_i}\right)P\left(C_{X_i}\right)=\sum_{i=r}^{n}\left(\frac{i}{n}\right)P\left(C_{X_i}\right)$$

and so,

$$\sum_{i=r}^{n} P^{N}\left(M_n \cap C_{X_i}\right)=\left(\frac{1}{n}\right)E^{N}(K_n)$$

or,

$$P^{N}(M_n)=\frac{E^{N}(K_n)}{n}=\frac{E^{O}(K_n)}{n} \tag{2.9}$$

Combining (2.9) and (2.7), the desired result follows.

2.3.2.1 The Limiting Case.

No discussion of the classical problem can be complete without mentioning the limiting case (as $n \to \infty$) and its extraordinary solution. For our purpose, the following derivation suffices, as it yields the correct answer. A more thorough derivation of this result can be found in Gilbert and Mosteller (1966). First, we want to find the optimal r that maximizes $P(M_n) = \frac{1}{n}\sum_{i=r}^{n} \frac{r-1}{i-1}$, as $n \to \infty$. Then we shall calculate the probability of winning under this r^* value. The argument is as follows:

$$P(M_n) = \frac{1}{n}\sum_{i=r}^{n} \frac{r-1}{i-1}$$

$$= \left(\frac{r-1}{n}\right)\sum_{i=r}^{n} \frac{1}{i-1}$$

$$= \left(\frac{r-1}{n}\right)\sum_{i=r}^{n} \left(\frac{n}{i-1}\right)\left(\frac{1}{n}\right)$$

Now, we let n tend to infinity and use:

$$x = \lim_{n\to\infty} \frac{r}{n}, \quad t = \frac{i}{n} \quad \text{and} \quad dt = \frac{1}{n}.$$

Under these substitutions, the above summation becomes a Reimann approximation for the integral:

$$x\int_x^1 \frac{1}{t}dt = -x\ln x \tag{2.10}$$

The x value that maximizes (2.10) is obtained by using basic calculus:

$$x_{optimal} = \frac{1}{e}, \qquad \lim_{n\to\infty} \frac{r^*}{n} = \frac{1}{e} \quad \text{and} \quad \lim_{n\to\infty} P^*(M_n) = \frac{1}{e}.$$

And so, for large n, the optimal strategy will pass approximately 37% of the observations without making a decision, and then select the first candidate thereafter.

Taking limits on both sides of (2.6) we get:

$$\lim_{n\to\infty}\left(\frac{1}{n}\right)E^*(K_n) \qquad = \lim_{n\to\infty}\left(\frac{1}{n}\right)\left[nP^*(M_n)+\left(r^*-1\right)\right]$$

$$= \frac{1}{e}+\frac{1}{e}=\frac{2}{e}$$

Therefore, $E^*(K_n) \approx \frac{2n}{e} \approx 0.73575n$ for large n.

This result seems strikingly high and it may be impractical to use a strategy that samples about 74% of the observations before making a decision.

This is particularly true, considering that n is large and that usually there is a cost involved in sampling.

2.4 Extending Theorem 1

2.4.1 A word on $s(r)$ rules

We have by now established that EC holds whenever an $s(r)$ strategy is being used. It therefore seems natural to try to find the conditions under which an $s(r)$ rule is optimal (and therefore applicable). In searching for these conditions, one first has to realize that a rule that automatically rejects the first $r-1$ observations can never be optimal if we are given information about the actual magnitude of the X_i's. For example, consider the following simple case, where the player is given the exact values of the X_i's (the full information case). Let $n=10$, and suppose that the first observation drawn is relatively large, say $X_1 = 0.95$. The probability of winning if we choose X_1 is $(0.95)^9 \approx 0.63$. The probability of winning if we continue has to be less than 0.37. This is true, since once we pass X_1, the only way we can win is if the following two events occur simultaneously:

(a) at least one of $X_1 \cdots X_9$ is greater than 0.95.

(b) we manage to identify $\max\{X_1 \cdots X_9\}$ correctly.

Since the probability of the first event is 0.37, the probability of both occurring can not exceed 0.37. Therefore, X_1 should be picked in this case if we play optimally. Moreover, regardless of how large n is, we can always find a value, d_1, close enough to 1, such that:

$$P\left(\text{win } \middle| X_1 > d_1 \text{ and we choose } X_1\right)$$
$$> P\left(\text{win } \middle| X_1 > d_1, \text{ we do not choose } X_1 \text{ and play optimally thereafter}\right)$$

Based on the above example, we can easily construct a threshold for X_1 and say that whenever $X_1 > (0.5)^{\left(\frac{1}{n-1}\right)}$, it should be chosen. Thus a strategy that has $P\left(C_{X_1}\right) = 0$ as one of its properties is clearly not optimal here. This argument can be extended to include any mixture of known distributions. By this we mean the following: let $X_i \sim U\left[0, a_j\right]$ with probability p_j $(1 \le j \le k)$, where $\sum_{j=1}^{k} p_j = 1$. The X_i's in this case are not identically distributed as in the full information case. Although the player is given the exact values of the X_i's, he does not have nearly as much information as in the previous case. Nevertheless, here again we can note that the probability of choosing X_1 when playing optimally should not be zero, for if

$$X_1 > \left(\frac{1}{2\max\left\{a_j\right\}}\right)^{\left(\frac{1}{n-1}\right)}, \text{ it should be chosen.}$$

2.4.2 An extension of Theorem 1

In 2.4.1 we saw that a process for which an $s(r)$ rule is optimal must involve a source that provides minimal information. Let us, therefore, restrict our attention to the ranking source γ_R. Variations of the model will be obtained by altering requirement 1.1.2. Instead of assuming that N, the total number of observations, is known, we let N be a discrete random variable with density:

$$P(N = k) = p_k \qquad k = 1 \ldots n$$

The optimal case where $P(N = n) = p_n = 1$ is nothing but the classical secretary problem for which an $s(r)$ rule is optimal (as we have already seen) with:

$$r^* = \min\left\{ l \geq 1 : \sum_{m=l}^{n-1} \frac{1}{m} \leq 1 \right\}$$

The above is an alternative way of describing the r value that maximizes:

$$P(M_n) = \frac{1}{n} \sum_{i=r}^{n} \frac{r-1}{i-1}$$

Rasmussen and Robbins (1975) consider the problem when N is a bounded random variable. They concluded that an $s(r)$ rule is optimal regardless of the distribution of N. Petrucelli (1983) has proved the last statement to be wrong by finding an error in lemma 3.2 of Rasmussen and Robbins. Irle (1980) shows by a counter-example that whether an optimal rule is an $s(r)$ rule *does* depend on the distribution of N. He generalizes the formulation of the problem and derives several sufficient conditions on the distribution of N for the optimal rule to be an $s(r)$ rule. One of these conditions is:

$$p_k > \sum_{m=k+1}^{n} \frac{p_m}{m} \quad \rightarrow \quad p_l > \sum_{m=l+1}^{n} \frac{p_m}{m} \quad \text{whenever} \qquad l > k$$

The above is also a special case of Presman and Sonin (1972). The sufficient condition is weakened somewhat in Theorem 2.2 of Petrucelli (1983).

Theorem 2.2 (Petrucelli, 1983):

Let $r = \min\left\{ l \geq 1 \ : \ \sum_{m=l}^{n-1} \frac{1}{m} \leq 1 \right\}$. Then a sufficient condition for an optimal rule to be an $s(r)$ rule is:

$$p_k \Big/ \sum_{m=k+1}^{n} \frac{p_m}{m} > 1 \quad \rightarrow \quad p_l \Big/ \sum_{m=l+1}^{n} \frac{p_m}{m} > 1 \qquad\qquad k \leq l \leq r-1$$

The highlight of this last paper is Theorem 2.3 in which the necessary and sufficient condition is given for an $s(r)$ rule to be optimal:

Theorem 2.3 (Petrucelli, 1983):

Let $c(m,l) = \left(\sum_{j=l+1}^{m} \frac{1}{j-1} \right) \Big/ m$

and

$$r = \max\left\{ l \geq 1 \; : \; \sum_{m=l}^{n} \frac{p_m}{m} < \sum_{m=l+1}^{n} c(m,l) p_m \right\}$$

Then the optimal rule is an $s(r)$ rule if and only if:

$$l \sum_{m=l}^{n} \frac{p_m}{m} < r \sum_{m=r+1}^{n} c(m,r) p_m \qquad\qquad 1 \leq l \leq r$$

Furthermore, if the last condition holds, the optimal rule passes over observations $1 \ldots r-1$ and chooses the first candidate thereafter.

Using this last result we can extend theorem 1 to obtain the following:

Theorem 1 (extended)

Suppose the maximum of a sequence of length N is to be found under the ranking source γ_R. Let N be a discrete random variable with:

$$P(N = k) = p_k \qquad\qquad k = 1 \ldots n$$

Define:

$$c(m,l) = \left(\sum_{j=l+1}^{m} \frac{1}{j-1} \right) \Big/ m$$

and

$$r = \max\left\{ l \geq 1 : \sum_{m=l}^{n} \frac{p_m}{m} < \sum_{m=l+1}^{n} c(m,l) p_m \right\}$$

Then, if the condition

$$l \sum_{m=l}^{n} \frac{p_m}{m} < r \sum_{m=r+1}^{n} c(m,r) p_m \qquad 1 \leq l \leq r$$

holds, then under optimal strategy

$$E^*(K_n) = nP^*(M) + (r-1)$$

also holds.

CHAPTER III
THE FULL-INFORMATION CASE

3.1 Deriving the optimal strategy

As was previously mentioned, γ_F is the source that provides the player with the exact values of the x_i's as they are being drawn. The assumption that the x_i's are uniformly distributed on $[0,1]$ is still valid. For convenience, in this chapter we will use lower case x_i's to denote the ith observation and reserve the X_i notation for a different purpose.

The optimal strategy in this case belongs to a family of strategies which are indexed by a sequence of n monotonically decreasing "decision numbers". For example, consider the situation where $n=5$ and $x_1 = .9$. The stopping criterion of S^* is given by (1.3) and so in this case x_1 is selected, because the probability that the other four observations are smaller than the first is $(0.9)^4 \approx 0.656$. Thus in the full-information case (as opposed to the classical secretary problem) no "buildup of experience" is needed and a choice can be made starting at x_1.

Furthermore, the decision whether or not x_i should be chosen is solely based on the magnitude of x_i. In fact, x_i should be selected only if it is a candidate and is greater than some "decision number" d_i. Here d_i satisfies the following:

$$\Pr\left(win \mid C_{x_i} \text{ and } x_i = d_i\right) =$$

$$\Pr(win \mid \overline{C}_{x_i}, x_i = d_i \text{ and we play optimally from } x_{i+1} \text{ on})$$

In general the stopping rule is: choose the first candidate that is greater than its decision number. The d_i's obviously form a decreasing sequence because with fewer draws to go, the less chance there is of winning. Since one of the x_i's must be chosen, let $d_n \equiv 0$.

3.1.1. Deriving the decision numbers

For a fixed n, let d_{n-i} be the decision number for the $(n-i)$th observation (after which there are i draws left). This number satisfies

$\Pr(win \mid C_{x_{n-i}} \text{ and } x_{n-i} = d_{n-i}) =$

$\Pr\left(win \mid \overline{C}_{X_{n-i}}, x_{n-i} = d_{n-i} \text{ and we play optimally from } x_{n-i+1} \text{ on}\right)$

The left hand side is the probability that all the remaining draws will be less than d_{n-i}, which is clearly $(d_{n-i})^i$. The right hand side is obtained by conditioning on the number of draws among the last i that are greater than d_{n-i}. If there is only one such draw, it is the maximum of the whole sequence as well as the only candidate left, and the game will be won. The probability of this occurring is $\binom{i}{1}(d_{n-i})^{i-1}(1-d_{n-i})$. In general, if there are i draws left, and j of them are greater than d_{n-i} we will win only if the first candidate sampled is the maximum of the whole sequence. The probability of this event is $\left(\frac{1}{j}\right)\binom{i}{j}(d_{n-i})^{i-j}(1-d_{n-i})^j$, and so the $(n-i)$th decision number satisfies:

$$(d_{n-i})^i = \sum_{j=1}^{i}\left(\frac{1}{j}\right)\binom{i}{j}(d_{n-i})^{i-j}(1-d_{n-i})^j \tag{3.1}$$

The result of simplifying the above is:

$$\sum_{j=1}^{i}\binom{i}{j}\left(\frac{z^j}{j}\right)=1 \text{ where } z=\frac{1-d_{n-i}}{d_{n-i}} \tag{3.2}$$

So, for instance for the case mentioned before where $n=5$ the last condition generates the levels:

$$d_1 = 0.8246$$
$$d_2 = 0.7758$$
$$d_3 = 0.6898$$
$$d_4 = 0.5000$$
$$d_5 = 0.0000$$

Note that condition (3.2) can also take the form:

$$\sum_{j=1}^{i}\frac{(d_{n-i})^{-j}-1}{j}=1 \qquad i=1\ldots n-1\,, \qquad d_n \equiv 0 \tag{3.2a}$$

3.2 The probability of winning the full-information game.

Claim: The probability of winning a full-information game when using a strategy indexed by n decreasing "decision numbers" is given by:

$$P(M_n)=\frac{(1-d_1)^n}{n}+\sum_{r=1}^{n-1}\left[\sum_{i=1}^{r}\left(\frac{(d_i)^r}{(r)(n-r)}-\frac{(d_i)^n}{(n)(n-r)}\right)-\frac{(d_{r+1})^n}{n}\right] \tag{3.3}$$

Proof: Let us assume that $x_{\max}$ is in the $(r+1)$st position. To win the game we must reject $x_1 \ldots x_r$ and also have $x_{r+1} > d_{r+1}$ (we already know that

x_{r+1} is a candidate). Suppose that $x_i = \max\{x_1 \ldots x_r\}$ and note that all of $x_1 \ldots x_r$ will be rejected if $x_i < d_i$. This is true because none of the earlier draws $x_1 \ldots x_{i-1}$ can exceed its own decision number and therefore can not be chosen. On the other hand, $x_{i+1} \ldots x_r$ are not candidates and hence are not chosen either. It is also clear that if $x_i > d_i$, then it is selected (and so $x_1 \ldots x_r$ are not all rejected). The following are therefore equivalent:

$$P(x_i = \max\{x_1 \ldots x_r\} \text{ and no selection is made among} x_1 \ldots x_r)$$

$$= P(x_i = \max\{x_1 \ldots x_r\} \text{ and } x_i < d_i) = \left(\frac{1}{r}\right)(d_i)^r$$

Also, $P(x_i = \max\{x_1 \ldots x_n\}) = \left(\frac{1}{n}\right)(d_i)^n$, and so the difference,

$$\left(\frac{1}{r}\right)(d_i)^r - \left(\frac{1}{n}\right)(d_i)^n \tag{3.4}$$

gives the probability of rejecting the first r observations and knowing that $x_{\max}$ had not been sampled yet. Since $x_{\max} \in \{x_{r+1} \ldots x_n\}$, the probability it is in the $(r+1)$st position is $\left(\frac{1}{n-r}\right)$. And so the probability of no draws among the first r, and $x_{\max}$ is in the $(r+1)$st position, is given by:

$$\sum_{i=1}^{r}\left(\frac{1}{n-r}\right)\left(\left(\frac{1}{r}\right)(d_i)^r - \left(\frac{1}{n}\right)(d_i)^n\right) \tag{3.5}$$

The summation is taken since any of the first r can be $\max\{x_1 \ldots x_r\}$.

The probability that we will not select x_{r+1} upon sampling when it is indeed the maximum of the whole sequence is:

$$P(x_{r+1} < d_{r+1}) = \frac{1}{n}(d_{r+1})^n \tag{3.6}$$

Combining (3.5) and (3.6), the probability of choosing the winning x_{r+1} is:

$$\left[\sum_{i=1}^{r}\left(\frac{1}{n-r}\right)\left(\left(\frac{1}{r}\right)(d_i)^r - \left(\frac{1}{n}\right)(d_i)^n\right)\right] - \left(\frac{1}{n}\right)(d_{r+1})^n \quad 1 \le r \le n-1 \tag{3.7}$$

For x_1 we have a special case: $P(M_{x_1}) = \frac{1}{n}$ and $P(\overline{C}_{x_1} \mid M_{x_1}) = \frac{(d_1)^n}{n}$.

The difference $\frac{1-(d_1)^n}{n}$ is therefore the probability of choosing x_1 and winning. The validity of (3.3) follows.

For obtaining $P^*(M_n)$, the probability of winning under S^*, all one has to do is substitute the optimal decision numbers given by (3.2a) into (3.3).

3.3 The expected location of the stopping variable $E(K_n)$.

Claim: When using a strategy of decreasing decision numbers the expected location of the stopping variable is given by:

$$E(K_n) = (1-d_1) + \left\{\sum_{r=2}^{n-1}\left[(r)\left(\sum_{i=1}^{r-1}\frac{(d_i)^{r-1}}{r-1} - \sum_{i=1}^{r}\frac{(d_i)^r}{r}\right)\right]\right\} + \left(\frac{n}{n-1}\right)\sum_{i=1}^{n-1}(d_i)^{n-1} \tag{3.8}$$

Proof: As has already been mentioned, $\sum_{i=1}^{r-1}\frac{(d_i)^{r-1}}{r-1}$ and $\sum_{i=1}^{r}\frac{(d_i)^{r}}{r}$ are the probabilities of rejecting the first r and the first $r+1$ observations respectively. The difference $\sum_{i=1}^{r-1}\frac{(d_i)^{r-1}}{r-1}-\sum_{i=1}^{r}\frac{(d_i)^{r}}{r}$ is therefore the probability of x_r being chosen for $1 \le r \le n-2$. The first and the last draws are treated separately as special cases. The probability that x_1 is chosen is clearly $1-d_1$. The probability of x_n being chosen is $\sum_{i=1}^{n-1}\frac{(d_i)^{n-1}}{n-1}$ because choosing x_n is equivalent to rejecting $x_1 \ldots x_{n-1}$. The distribution of the stopping variable K_n is:

$$
\begin{aligned}
\Pr(K_n = r) &= 1-d_1 && r=1 \\
&= \sum_{i=1}^{r-1}\frac{(d_i)^{r-1}}{r-1}-\sum_{i=1}^{r}\frac{(d_i)^{r}}{r} && 2 \le r \le n-1 \qquad (3.9) \\
&= \sum_{i=1}^{n-1}\frac{(d_i)^{n-1}}{n-1} && r=n
\end{aligned}
$$

and equation (3.5) follows.

A simpler expression for $E(K_n)$ when $2 \le r \le n-1$ can be derived as follows:

$$\sum_{r=2}^{n-1}(r)\left(\sum_{i=1}^{r-1}\frac{(d_i)^{r-1}}{r-1}-\sum_{i=1}^{n}\frac{(d_i)^{r}}{r}\right) =$$

$$2\sum_{i=1}^{1}\frac{(d_i)^1}{1}-2\sum_{i=1}^{2}\frac{(d_i)^2}{2}+3\sum_{i=1}^{2}\frac{(d_i)^2}{2}-3\sum_{i=1}^{3}\frac{(d_i)^3}{3}+4\sum_{i=1}^{3}\frac{(d_i)^3}{3}$$

$$+\cdots+(n-1)\sum_{i=1}^{n-2}\frac{(d_i)^{n-2}}{n-2}-(n-1)\sum_{i=1}^{n-1}\frac{(d_i)^{n-1}}{n-1}$$

$$= d_1+\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\frac{(d_i)^j}{j}-n\sum_{i=1}^{n-1}\frac{(d_i)^{n-1}}{n-1} \qquad (3.10)$$

The result of substituting (3.10) into (3.8) is:

$$E(K_n)=1+\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\frac{(d_i)^j}{j} \qquad (3.11)$$

3.4 The validity of the conjecture in the full-information case

Next we want to prove that EC holds in this case as well. One approach is to take the optimal decision numbers given by (3.2a), substitute them into (3.3) and (3.11) to obtain $P^*(M_n)$ and $E^*(K_n)$, and then show that $nP^*(M_n) = E^*(K_n)$. However, I would like to use an alternative and algebraically much simpler approach.

Let $X_k=(x_1\ldots x_k)$ be the k-tuple of observed values and I_k be the information from X_k available to the player (in the full-information case I_k is X_k itself).

$$D_k(I_k)=\Pr\left(C_{X_k}\mid \overline{C}_{x_1}\cap\cdots\cap\overline{C}_{x_{k-1}},I_k\right) \qquad k=1\ldots n$$

Under this formulation $D=\{D_1(I_1),\ldots,D_n(I_n)\}$ is the strategy to be followed and:

$$P^*(M_n)=\max_D P(M_n)=P_{D^*}(M_n)$$

where the subscript D indicates that strategy D is followed. One can write:

$$P_D(M_n)=\int_0^1 P(M_n|C_{x_1},X_1)\,D_1(I_1)dx_1 \;+\; \int_0^1(1-D_1(I_1))dx_1\int_0^1 P(M_n|C_{x_2},\overline{C}_{x_1},X_2)D_2(I_2)dx_2+\cdots$$

$$\cdots+\int_0^1(1-D_{n-1}(I_{n-1}))dx_{n-1}\int_0^1 P(M_n|C_{x_n},\overline{C}_{x_{n-1}},\ldots\overline{C}_{x_1},X_n)\,D_n(I_n)dx_n \quad (3.12)$$

Since exactly one variable must be selected, define $D_n(I_n)\equiv 1$. In the full information case we have:

$$\begin{aligned} P(M_n|C_{x_k},\overline{C}_{x_{k-1}}\ldots,\overline{C}_{x_1},X_k) &= 0 \text{ if } \quad x_k<\max\{x_1\ldots x_{k-1}\} \\ &= x_k^{n-k} \quad x_k>\max\{x_1\ldots x_{k-1}\} \quad (3.13) \end{aligned}$$

Substituting (3.13) into (3.12) yields:

$$P(M_n)=\sum_{k=1}^{n}\left[\prod_{j=1}^{k-1}\left(\int_0^1(1-D_j(I_j))dx_j\right)\left(\int_{\max\{x_1\ldots x_{k-1}\}}^{1} x_k^{n-k}D_k(I_k)dx_k\right)\right] \quad (3.14)$$

where by convention $\prod_{j=1}^{0}\cdot\equiv 1$, and $\max\{x_1,x_0\}\equiv 0$.

The optimal strategy here, as discussed before, takes the form:

$$D_k(I_k) = 1 \quad \text{if } x_k > \max\{\max\{x_1 \ldots x_{k-1}\}, d_k\} \tag{3.15}$$
$$0 \quad \text{otherwise}$$

The kth decision number, d_k, was previously shown in (3.2a) to be the unique root in [0,1] of:

$$\sum_{j=1}^{n-k} \frac{(d_k)^{-j} - 1}{j} = 1$$

Substitution of (3.15) into (3.14) gives $P^*(M_n)$, and after simplifying we get:

$$P^*(M_n) = \left(\frac{1}{n}\right)\left(1 + \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} \frac{(d_i)^j}{j}\right) \tag{3.16}$$

Comparing this result to (3.11) results in $nP^*(M_n) = E^*(K_n)$. Therefore, Enns's Conjecture also holds in the full-information case.

Below is a table of optimal decision numbers for the full-information case, together with the probability of winning for various n values:

n	d_i^*	$P^*(M_n)$
1	0.0000	1.0000
2	0.5000	0.7500
3	0.6899	0.6843
4	0.7758	0.6554
5	0.8246	0.6392
6	0.8560	0.6288
7	0.8778	0.6215
8	0.8939	0.6161
9	0.9062	0.6120
10	0.9160	0.6087
15	0.9448	0.5990
20	0.9589	0.5942
30	0.9728	0.5895
40	0.9797	0.5871
50	0.9838	0.5857
∞	1.0000	0.5802

Table 3.1

Optimal decision numbers for the full-information case, together with the probability of winning for various n values

CHAPTER IV
PARTIAL-INFORMATION CASES

4.1 Introduction

Various types of partial-information models have been analyzed by Enns (1974), Stewart (1978), Samuels (1981), Campbell (1981), and others. In such a model the player is not given the exact value of the observation drawn. Instead he receives partial information regarding the magnitude of the draws. Consider the problem introduced by Enns (1974). Here a decision maker determines a specific "level" before each draw to which the observed draw will be compared. As the random variable is sampled, information is supplied as to whether the value observed is greater or less than the predetermined level. Two situations where this might occur are: 1) where measuring the exact values of the sampled variables is very expensive or time consuming, or 2) where destructive sampling is used. In the first case, classifying values into one of two categories, greater or less than the specified level, is much simpler than finding exact values. In the case of destructive sampling, suppose that the strongest bolt in a sample of n must be chosen. An application of the partial-information case arises if each bolt is subjected to a predetermined level of stress, and the first to survive is accepted as the strongest in the sample.

4.2 The optimal strategy

The logic that was used for deriving S^*, the optimal strategy for the full-information case, can be employed here as well. Therefore S^* is a strategy indexed by n "levels" $l_1 \geq l_2 \ldots \geq l_n$. The first observation, x_1, is measured at level l_1 and is accepted if $x_1 > l_1$. Note that, unlike before, the only criterion for choosing x_1 is if $x_1 > l_1$ regardless of whether it is a candidate or not (information which is unavailable). There is a possibility of having $l_1 > x_1 > x_2 > l_2$ where x_1 is rejected and x_2 accepted, although $x_2 < x_1$. Since one random variable must be accepted, the task must be to define $l_n \equiv 0$.

The optimal strategy is the n-tuple $l_1^* > l_2^* ... > l_n^*$ that maximizes the probability of winning given by :

$$P(M_n) = \sum_{i=1}^{n} P\left(c_{x_i} \cap m_{x_i}\right) \tag{4.1}$$

Now,

$$P\left(c_{x_i} \cap m_{x_i}\right) = \int_{l_i^*}^{1} x_i^{n-i} dx_i \times \int_{0}^{\min(x_i,\, l_{i-1})} dx_{i-1} \times \int_{0}^{\min(x_i,\, l_{i-2})} dx_{i-2} \ldots \times \int_{0}^{\min(x_i,\, l_1)} dx_1 \tag{4.2}$$

The first integral gives the probability of winning, given that x_i is chosen. The product of all the others is the probability of x_i being chosen (or equivalently $\{x_1 \ldots x_{i-1}\}$ being rejected). Conditioning on the exact value of x_i one can reduce (4.2) to:

$$P\left(c_{x_i} \cap m_{x_i}\right) = \sum_{r=1}^{i}\left[\left(\frac{l_{i-r}^{n+1-r} - l_{i-r+1}^{n+1-r}}{n+1-r}\right)\left(\prod_{k=1}^{r}\frac{l_{i-k+1}}{l_i}\right)\right] \tag{4.3}$$

Substituting (4.3) into (4.1) yields

$$P(M_n) = \sum_{i=1}^{n}\left\{\sum_{r=1}^{i}\left[\left(\frac{l_{i-r}^{n+1-r} - l_{i-r+1}^{n+1-r}}{n+1-r}\right)\left(\prod_{k=1}^{r}\frac{l_{i-k+1}}{l_i}\right)\right]\right\}$$

$$= \sum_{i=1}^{n}\left[\left(\frac{1}{i}\right)\left(\prod_{r=0}^{n-i} l_r\right)\right] - \left(\frac{1}{n-1}\right)\sum_{i=1}^{n} l_i^n - \sum_{i=1}^{n-2}\frac{1}{(i)(i+1)}\sum_{r=1}^{i} l_r^{i+1}\prod_{j=r+1}^{n-i+r-1} l_j \quad (n \geq 3) \tag{4.4}$$

Obviously $P(M_1)=1$, and we can easily find that $P(M_2)=\frac{1}{2}+l_1-l_1^2$. Therefore when n=2, $l_1=\frac{1}{2}$ maximizes the probability of winning (in fact, $P^*(M_2)=0.75$). This is also intuitively clear. Unlike the situation in the previous chapter, where the n-game is a mere extension of the n-1 game (all that is necessary is to add a new decision number to existing numbers), here we are not as fortunate. The strategy for n-levels employs n levels which can not be determined by knowing any of the optimal policies with less than n draws. No recursive relation exists between the optimal levels for n-1 draws and the levels for n draws.

However, Enns (1974) manages to obtain the optimal strategies for $n \leq 25$. The technique used is an n-dimensional Newton's method. The table below lists optimal policies up to n=12. The probabilities of winning for all these strategies, as well as for some larger n-values, were found to be:

	$n=2$	$n=3$	$n=4$	$n=5$
1	0.500	0.673	0.757	0.808
2		0.546	0.692	0.768
3			0.587	0.712
4				0.622

Table 4.1

Optimal decision levels for n=1..12

(continued next page)

	n=6	n=7	n=8	n=9
1	0.8406	0.8603	0.8814	0.8949
2	0.8138	0.8447	0.8669	0.8836
3	0.7791	0.8209	0.8495	0.8703
4	0.7309	0.7901	0.8280	0.8544
5	0.6524	0.7473	0.8004	0.8349
6		0.6773	0.7618	0.8097
7			0.6985	0.7746
8				0.7169

	n=10	n=11	n=12
1	0.9056	0.9144	0.9216
2	0.8965	0.9069	0.9154
3	0.8861	0.8985	0.9084
4	0.8739	0.8888	0.9006
5	0.8593	0.8775	0.8915
6	0.8414	0.8640	0.8810
7	0.81825	0.8474	0.8684
8	0.78602	0.8260	0.8523
9	0.7329	0.7962	0.8331
10		0.7470	0.8054
11			0.7595
12			

Table 4.1

Optimal decision levels for n=1..12

The probabilities of winning for all these strategies, as well as for some larger n-values are presented in the following table (Table 4.2).

n	$P^*(M_n)$	n	$P^*(M_n)$
1	1	15	0.5829
2	0.7500	20	0.5775
3	0.6798	25	0.5744
4	0.6474	30	0.5723
5	0.6289	50	0.5681
6	0.6169	75	0.5661
7	0.6085	100	0.5651
8	0.6024	200	0.5635
9	0.5976	500	0.5626
10	0.5939	∞	0.5620

Table 4.2

Probabilities of winning for various n values.

It can also be shown that for large n the approximation:

$$P^*(M_n) = 0.56203 + {0.30127}/{n} + {0.17749}/{n^2}$$

is accurate to four decimal places as long as $n \geq 10$. Note also how close these values are for the probabilities of winning the full information game. This suggest that by classifying each of the observation into one of the two classes (instead of measuring its exact value) we don't really sacrifice much. The gain, however, is considerable because as mentioned before measuring exact values is often costly and time consuming.

4.3 The location of the stopping variable

Let K_n be the location of the stopping variable just as before. Trivially,

$$P(K_n = k) = (1 - l_k)\prod_{i=1}^{k-1} l_i \quad k = 1\ldots n \tag{4.5}$$

The expected value of K_n is therefore:

$$E(K_n) = \sum_{k=1}^{n} k\left[(1 - l_k)\prod_{i=1}^{k-1} l_i\right] = \sum_{k=1}^{n}\prod_{i=1}^{n-k} l_i \tag{4.6}$$

4.4 The relationship between $E^*(K_n)$ and $P^*(M_n)$ in the partial-information case

By differentiating (4.4) with respect to the different levels $l_1 \ldots l_n$ we can obtain the partial derivatives $\dfrac{\partial\, P(M_n)}{\partial\, l_i}$ $i = 1\ldots n$. These derivatives satisfy the relation:

$$\sum_{i=1}^{n}\left(\frac{l_i}{n}\right)\left(\frac{\partial P(M_n)}{\partial l_i}\right) = P(M_n) - \frac{E(K_n)}{n} \tag{4.7}$$

When S^* is employed the levels are adjusted such that $\dfrac{\partial\, P(M_n)}{\partial\, l_i} = 0 \;\; \forall i$. And so under S^* equation (4.7) reduces to $E(K_n) = nP(M_n)$ and the validity of the conjecture is established for this model.

4.5 A variation on the model — reducing the number of decision levels

Let us assume that the decision maker wants to reduce the complexity involved when working with n different levels. One way to do it is by selecting only t levels ($t < n$) and using some or all of them more than once. To illustrate this, let n=6 and t =3 and denote the three reduced levels by L_1, L_2, L_3. The decision maker may then decide to have

$$\begin{aligned} L_1 &= l_1 = l_2 \\ L_2 &= l_3 = l_4 = l_5 \\ L_3 &= l_6 = 0 \end{aligned} \qquad \text{or alternatively,} \qquad \begin{aligned} L_1 &= l_1 = l_2 = l_3 \\ L_2 &= l_4 = l_5 \\ L_3 &= l_6 = 0 \end{aligned}$$

or any other reasonable scheme.

If one is restricted to t levels $L_1, L_2 \ldots L_t$ then the probability of obtaining the maximal random variable is just (4.4) with the l_i's replaced by the appropriate $L_1, L_2 \ldots L_t$. Note also that once the levels $L_1, L_2 \ldots L_t$ are set, it is still up to the decision maker to decide how many l-levels correspond to each L-level. Since $P(M_n)$ is a polynomial in l_i $(i = 1 \ldots n)$, then for any arbitrary $L_u \left(= l_{u_1} = l_{u_2} \ldots = l_{u_k}\right)$ we obtain:

$$\sum_{i=1}^{k} \frac{\partial\, P(M_n)}{\partial\, l_i} = \frac{\partial\, P(M_n)}{\partial\, L_u}.$$

This enables us to rewrite (4.7) as:

$$\sum_{u=1}^{t} \frac{L_u}{n} \cdot \frac{\partial\, P(M_n)}{\partial\, L_u} = P(M_n) - \frac{E(K_n)}{n} \tag{4.8}$$

When using optimal strategy the levels $L_1, L_2 \ldots L_t$ satisfy $\dfrac{\partial\, P(M_n)}{\partial\, L_u} = 0$ $\forall u = 1 \ldots t$. It follows that the conjecture is valid in this case, when the number of levels is restricted.

4.6 Inductive Games

When trying to prove EC in general when the information source is not specified, an inductive approach may be taken. Let us assume for a moment that the player tries to locate the maximum of a sequence of length $n+1$. He tests x_1 at level l_1, and if $x_1 > l_1$, it is selected and the game is over. Otherwise,, the player concentrates on finding $\max\{x_2 \ldots x_{n+1}\}$, completely ignoring the fact that x_1 had ever been sampled. Now make the inductive assumption that EC is true whenever we try to find the maximum of a sequence of length n. We will show that it also holds true for the $n+1$ case. Under the above scheme, extending from the n-game to the $n+1$ game is done via the addition of a new level l_1 at which x_1 is tested where:

$$P\left(C_{x_1}\right) = \begin{array}{ll} 1 & x_1 \geq l \\ 0 & x_1 < l_1 \end{array}$$

Clearly,

$$\begin{aligned} E(K_{n+1}) &= (1-l_1)(1) + (l_1)(E(K_n)+1) \\ &= 1 + l_1 E(K_n) \end{aligned} \tag{4.9}$$

because if $x_1 < l_1$, the game reduces to an n-game and we still have to sample $E(K_n)$ more observations on the average (in addition to x_1 that had already been sampled). On the other hand we have:

$$P(M_{n+1}) = P(C_{x_1} \cap M_{x_1}) + P(\overline{C_{x_1}} \cap M_{n+1})$$

$$= P(M_{x_1}) \cdot P(C_{x_1} | M_{x_1}) + \left(l_1 P(M_n) - \frac{l_1^{n+1}}{n+1} \right)$$

$$= \left(\frac{1}{n+1} \right)(1 - l_1^{n+1}) + \left(l_1 P(M_n) - \frac{l_1^{n+1}}{n+1} \right) \tag{4.10}$$

Note that when calculating the probability of not choosing x_1 and winning, $\frac{l_1^{n+1}}{n+1}$ is subtracted from $l_1 P(M_n)$. This is so since there exists a possibility that x_1 will not be selected, $\max\{x_2 \ldots x_{n+1}\}$ will be correctly identified, but we will still lose. It occurs when x_1 is the largest random variable and all of $x_1 \ldots x_{n+1}$ are less than l_1. The probability of this event is $\frac{l_1^{n+1}}{n+1}$.

Differentiating (4.10) with respect to l_1 and equating to zero yields:

$$-2l_1^n + P(M_n) = 0 \tag{4.11}$$

And so, the optimal value for l_1 is:

$$l_1^* = \sqrt[n]{\frac{P^*(M_n)}{2}} \tag{4.12}$$

Substituting this last result into (4.9) and (4.10), we obtain:

$$E^*(K_{n+1}) = 1 + \left(\sqrt[n]{\frac{P^*(M_n)}{2}}\right)\left(E^*(K_n)\right) \tag{4.13}$$

and

$$P^*(M_{n+1}) = \left(\frac{1}{n+1}\right)\left(1 - 2\sqrt[n]{\frac{P^*(M_n)}{2}}\right) + \left(P^*(M_n)\right)\left(\sqrt[n]{\frac{P^*(M_n)}{2}}\right) \tag{4.14}$$

respectively.

Using the inductive assumption $E^*(K_n) = nP^*(M_n)$ together with (4.13) we get:

$$E^*(K_{n+1}) = 1 + \left(\sqrt[n]{\frac{P^*(M_n)}{2}}\right)\left(nP^*(M_n)\right) \tag{4..15}$$

On the other hand:

$$(n+1)P^*(M_{n+1}) = 1 - 2\left(\sqrt[n]{\frac{P^*(M_n)}{2}}\right)^{n+1} + (n+1)P^*(M_n)\sqrt{\frac{P^*(M_n)}{2}}$$

$$= 1 + \left(\sqrt[n]{\frac{P^*(M_n)}{2}}\right)\left(nP^*(M_n)\right) \tag{4.16}$$

And so we have inductively established that:

$$E^*(K_{n+1}) = (n+1)P^*(M_{n+1})$$

Note that equations (4.11) and (4.15) together with the initial conditions $P(M_1) = 1$ generate the following results:

n	$P^*(M_n)$	l^*	$E^*(K_n)$
1	1.0000	0.0000	1.00
2	0.7500	0.5000	1.50
3	0.6395	0.6124	1.92
4	0.5780	0.6838	2.31
5	0.5390	0.7332	2.70
6	0.5122	0.7693	3.07
7	0.4928	0.7969	3.45
8	0.4780	0.8186	3.82
9	0.4664	0.8362	4.20
10	0.4570	0.8506	4.57
11	0.4494	0.8628	4.94
12	0.4430	0.8731	5.31
13	0.4376	0.8820	5.68
14	0.4329	0.8897	6.06
15	0.4289	0.8964	6.43
20	0.4148	0.9208	8.29
30	0.4007	0.9462	12.02
40	0.3937	0.9592	15.74
50	0.3896	0.9672	19.48
100	0.3814	0.9834	38.14
500	0.3749	0.9967	187.46
1000	0.3741	0.9983	374.14

Table 4.3

Probabilities of winning and optimal levels for various n values

in an inductive process

4.6.1 When should inductive procedure be used?

Let us refer back to the problem discussed in 4.1 and assume that the strongest bolt of a sample of size nine has to be identified. As mentioned before, the bolts are tested sequentially, each under a certain amount of stress. Suppose that the process controller starts receiving samples of size ten, and would like to locate the strongest bolt in each of the new samples. An optimal solution would be to readjust the set of old levels and use the optimal numbers for the $n = 10$ case as given in Table 4.1. Recall, however that one of our initial assumptions in chapter 4 is that readjusting the levels is costly and hence another solution is required.

The best one can do in such a case is to attach one new level to the set of old levels. The first bolt out of the ten is stressed at that new level, and if it fails the problem reduces to the $n = 9$ case.

Deriving the new level is done by using the following results:

1) $P^*(M_9) = .5976$ (Table 4.2)

2) $l^* = \sqrt[n]{\dfrac{P^*(M_n)}{2}}$ (4.12)

Combining the above two, we get $l^* = .8744$. The probability of correctly identifying the bolt with the maximal strength for the new extended sample can easily be obtained by using (4.10), and we get $P(M_{10}) = .57028$.

Note how close the last result is to 0.5939 (Table 4.2) which is the probability of winning the $n = 10$ case if we were to readjust all ten levels. So as it turns out in this case we do not lose much efficiency by not playing optimally.

CHAPTER V
TOWARD A NEW CONJECTURE

5.1 Introduction

In the winter of 1991, I was enrolled in a senior level statistics course as part of my undergraduate program. In one of the lectures the instructor, E. Enns, presented the class with a conjecture posed by him years earlier. This conjecture stated the strikingly simple relationship between the probability of winning a sequential game and the expected number of steps the game lasts. A year later I had to select a topic for my master's thesis. My decision was to try to prove the validity of the conjecture. It seemed to me at the time that proving a relation as simple as:

$$nP^*(M_n) = E^*(K_n)$$

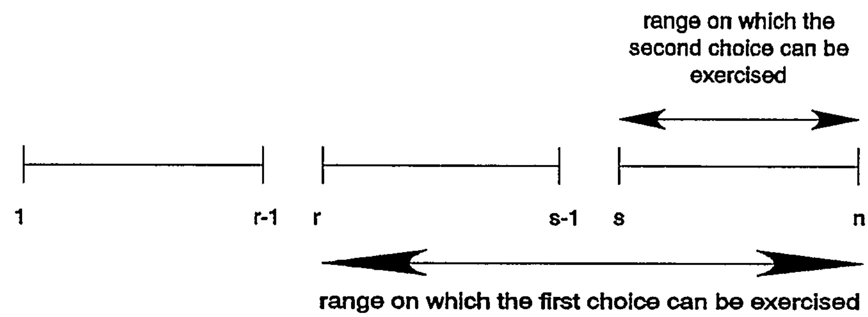
is true in general, should not be tremendously complicated. A year of intensive research has convinced me that the general proof (if it exists) may not be as simple as I had hoped. Nevertheless, as I progressed, I discovered a whole range of new problems related to topics previously discussed, most of which are well worth investigating. This paper started with the introduction of a rather interesting relationship, and as a proper conclusion I would like to introduce another curious relationship that seems to hold. This relationship will be posed as a conjecture.

5.2 The classical secretary problem with two choices

Let us consider a sequential sampling process where the source of information is γ_R (a player is given the rank of observations as he proceeds). The objective of the game is still to find the maximal observation, but in this case the player is given two choices. Note that the decision maker is not forced to exercise his second selection if he wishes to retain the first selection until the game terminates. This is nothing but a classical secretary problem with two choices, a model that was partially discussed by Gilbert and Mosteller (1966).

5.2.1 Deriving the optimal strategy

S^* here is derived in a manner similar to the one used previously for ranking observations in the game with one choice. The ith observation should be chosen if and only if it satisfies the criterion for optimality given in 1.2.1. This criterion will serve us in making both decisions. Since the right hand side and the left hand side of (1.1) are increasing and decreasing functions of i respectively, the argument that was used for the one choice game still holds. It follows that S^* will reject the first r-1 observations (for some r) and then exercise the first choice once it finds an observation of rank 1. Once the first choice is made, the game reduces to a one choice secretary problem. This is so because the objective is still finding $x_{\max}$, and we have one selection left. Therefore, the second selection will not be made before the sth observation, where s is the parameter associated with the classical secretary problem for this particular n. And so we get that S^* is a two parameter strategy. It rejects the first r-1 in any case and chooses the first candidate thereafter. The second choice is exercised only if another candidate is located among the $\{x_s \ldots x_n\}$, Graphically S^* is of the form:



5.2.2 The probability of winning

Under the scheme discussed in the previous section, the probability of winning the two choice game is:

$$P(M_n) = \left(\frac{1}{n}\right)\sum_{i=r}^{n}\left(\frac{r-1}{i-1}\right) + \left(\frac{s-r}{n}\right)\sum_{i=s}^{n}\left(\frac{1}{i-1}\right) + \left(\frac{1}{n}\right)\sum_{i=s+1}^{n}\sum_{j=s}^{i-1}\left(\frac{r-1}{j-1}\right)\left(\frac{1}{i-1}\right) \qquad (5.1)$$

The first term is the probability of winning with the first selection when the second is never used. The second term is the probability of winning with second choice and first choice was made among $\{x_r \ldots x_{s-1}\}$. The last term is the probability of winning when both choices are in $\{x_s \ldots x_n\}$.

A simple computer simulation generates the following optimal values of r and s together with probabilities of winning for a selection of n-values.

n	r^*	s^*	$P^*(M_n)$
1	1	1	1
2	1	2	1
3	1	2	.833
4	1	2	.708
5	2	3	.708
10	3	4	.646
15	4	6	.627
20	5	8	.618
30	7	12	.608
50	12	19	.601
100	13	38	.596

Table 5.1

The two choice game—optimal thresholds and probabilities of winning.

(Remark: in the special case $r^*=1$, the first summation of (5.1) should be replaced by 1.)

5.2.3 Asymptotic theories

Consider the limiting case as $n \to \infty$ and let us define:

$$x = \lim_{n\to\infty} \frac{r}{n} \qquad y = \lim_{n\to\infty} \frac{s}{n}$$

$$t = \frac{i}{n} \qquad dt = \frac{1}{n} \tag{5.2}$$

$$p = \frac{j}{n} \qquad dp = \frac{1}{n}$$

The terms of (5.1) in the limiting case become:

$$\frac{1}{n}\sum_{i=r}^{n}\frac{1}{i-1} = \frac{r-1}{n}\sum_{i=1}^{n}\left(\frac{1}{n}\right)\left(\frac{n}{i-1}\right) \approx x\int_x^1\left(\frac{1}{t}\right)dt = -x\ln x$$

$$\frac{1}{n}\sum_{i=s}\frac{s-r}{i-1} = \frac{s-r}{n}\sum_{i=s}^{n}\left(\frac{n}{i-1}\right)\left(\frac{1}{n}\right) = (y-x)\int_y^1\left(\frac{1}{t}\right)dt = -(y-x)\ln y \tag{5.3}$$

and finally:

$$\left(\frac{1}{n}\right)\sum_{i=s+1}^{n}\sum_{j=s}^{i-1}\left(\frac{r-1}{j-1}\right)\left(\frac{1}{i-1}\right) = \left(\frac{r-1}{n}\right)\sum_{i=s+1}^{n}\sum_{j=s}^{i-1}\left(\frac{n}{j-1}\right)\left(\frac{n}{i-1}\right)\left(\frac{1}{n}\right)\left(\frac{1}{n}\right)$$

$$\approx x\int_y^1\int_y^t\left(\frac{1}{p}\right)\left(\frac{1}{t}\right)dp\,dt$$

$$= \left(\frac{1}{2}\right)(x)(\ln y)^2. \tag{5.4}$$

The asymptotic probability of winning is therefore:

$$P(M_\infty) = x\ln x - (y-x)\ln y + \left(\frac{1}{2}\right)(x)(\ln y)^2 \tag{5.5}$$

From the single choice game we know that:

$$s^* = \frac{n}{e} \qquad \text{i.e.} \quad y^* = \frac{1}{e} \tag{5.6}$$

Substituting this result into (5.5) gives the result:

$$P(M_\infty) = x\ln x + \left(\frac{1}{e} - x\right) + \frac{x}{2} \tag{5.7}$$

Differentiating (5.7) and equating to zero yields:

$$\frac{\partial\, P(M_\infty)}{\partial\, x} = 0 \qquad \rightarrow \qquad x^* = e^{-\frac{3}{2}}$$

and we can easily obtain $r^* = ne^{-\frac{3}{2}}$.

The optimal probability of winning in the limiting case is therefore:

$$P^*(M_n) = e^{-1} + e^{-\frac{3}{2}} = .5910.$$

Note that in the limiting cases we obtain the following:

For the single choice game:

$$nP^*(M_\infty) = n\left(\frac{1}{e}\right) = s^*$$

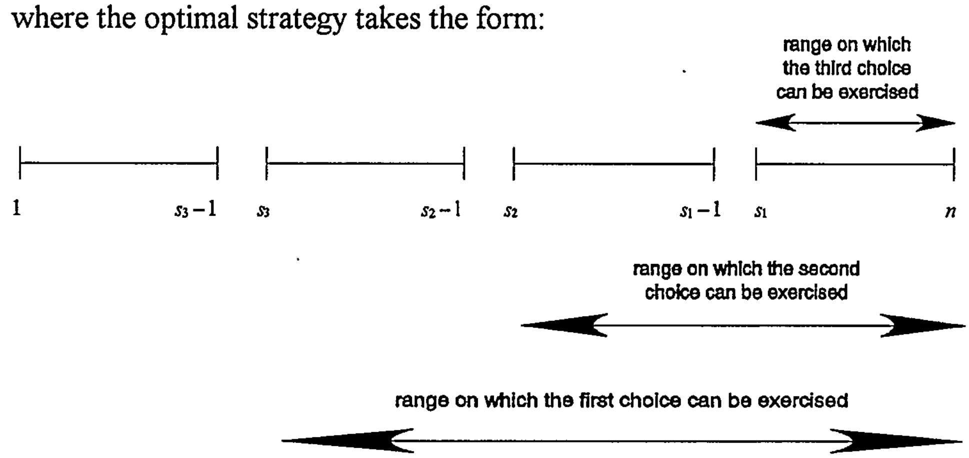
For the double choice game:

$$nP^*(M_\infty) = n\left(\frac{1}{e}\right) + ne^{-\frac{3}{2}} = s^* + r^*$$

Analysis of the three-choice game is performed in the same way, but is quite a bit more complicated. The results of the three-choice game in the limiting case are:

$$s_1^* = ne^{-1}$$
$$s_2^* = ne^{-\frac{3}{2}}$$
$$s_3^* = ne^{-2}$$

where the optimal strategy takes the form:



The probability of winning the three-choice game in the limiting case is found to be:

$$P(M_n) = e^{-1} + e^{\frac{-3}{2}} + e^{-2}$$

$$= .7263$$

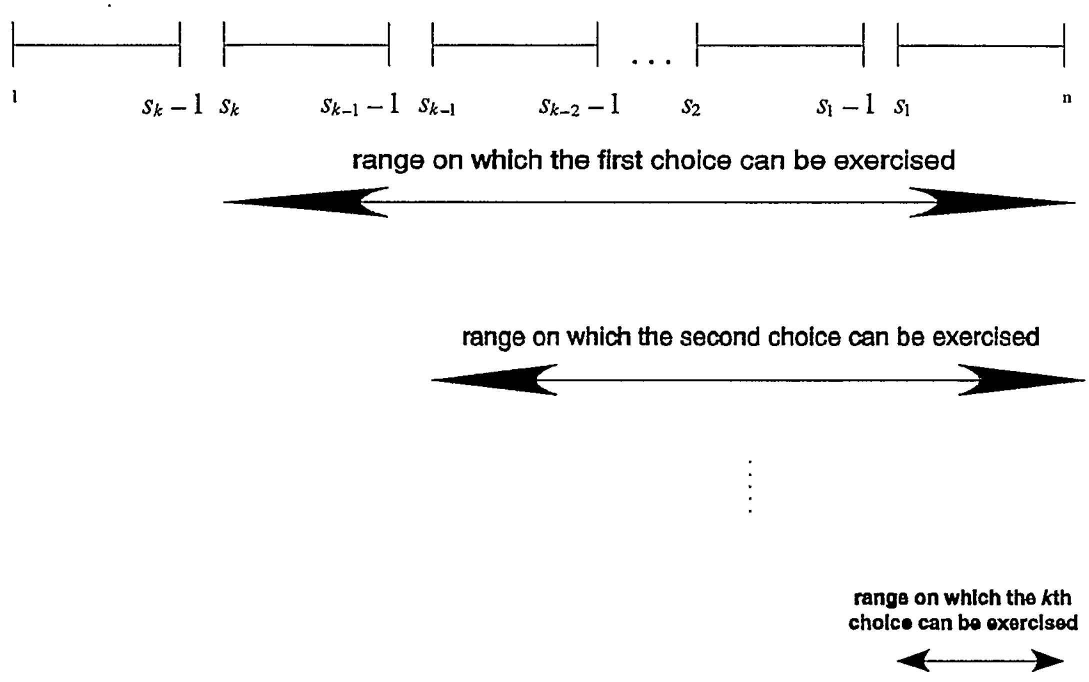
Computer simulations for cases greater than n=3 support the following claim which is posed as a conjecture:

New Conjecture

In trying to locate the maximum of a sequence of length n under a ranking source with k selections, the following holds true:

$$P^{*}(M_n) = \left(\frac{1}{n}\right)\sum_{i=1}^{k} s_i^{*}$$

where the s_i^{*}'s are the optimal threshold number when playing optimally. The optimal strategy takes the form:


1
$s_k - 1$
s_k
$s_{k-1} - 1$
s_{k-1}
$s_{k-2} - 1$
s_2
$s_1 - 1$
s_1
n
range on which the first choice can be exercised
range on which the second choice can be exercised
range on which the kth choice can be exercised

BIBLIOGRAPHY

[1] Bissinger, B.H. and Siegel, C., "Problem 5086". Am. Math. Mon., 70 (1963), pp. 336

[2] Campbell, G. and Samuels, S. M., "Choosing the best of the current crop". Adv. Appl. Prob. Vol. 13 (1981), pp. 510-532.

[3] Enns, E.G., "The role of information in a sequential decision problem". (unpublished, 1974).

[4] Enns, E.G., "Selecting the maximum of a sequence with imperfect information". JASA Vol. 70 (1975), pp. 640-643.

[5] Ferguson, T.S., "Who solved the secretary problem?" Stat. Sci. Vol. 4 (1983), pp. 282-296.

[6] Freeman, P.R., "The secretary problem and its extensions: A review". Int. Stat. Rev. Vol. 51 (1983), pp. 189-206.

[7] Gardner, M., "Mathematical games". Scientific American, 202(2) (1960), pp.152-153.

[8] Gianini-Pettitt, J., "Optimal selection based on relative ranks with a random number of individuals". Adv. Appl. Prob. Vol. 11 (1979), pp. 720-736.

[9] Gilbert, J.P. & Mosteller, F., "Recognizing the maximum of a sequence". JASA Vol. 61 (1966), pp. 35-73.

[10] Irle, A., "On the best choice problem with random population size". Z. Operat. Res. Vol. 25 (1980), pp. 177-190.

[11] Koestler, A., "The watershed. A biography of Johannes Kepler". Anchor Books, Garden City, NY (1960).

[12] Petrucelli, J.D., "On the best choice problem when the number of observations is random". J. Appl. Prob. Vol. 20 (1983). pp. 165-171.

[13] Presman, E.L. and Sonin, I.M., "The best choice problem for a random number of objects". Theory Prob. Appl. Vol. 18 (1972), pp. 657-668.

[14] Rasmussen, W.T., "A generalized choice problem". J. Optimization Theory Appl. Vol. 15 (1975), pp. 311-325.

[15] Rasmussen, W.T. & Robbins, H., "The candidate problem with unknown population size". J. Appl. Prob. Vol. 12 (1975), pp. 692-701.

[16] Rose, J.S., "Twenty years of secretary problems: a survey of developments in the theory of optimal choice". Adv. Manag. Studies Vol. 1 (1982), pp. 53-64.

[17] Samuels, S.M., "Minimax stopping rules and the underlying distribution with uniform". J. Am. Stat. Assoc. Vol. 76 (1981), pp. 188-197.

[18] Stewart, T.J., "Optimal selection from a random sequence with learning of the underlying distribution". J. Am. Stat. Assoc. Vol. 73 (1978), pp. 775-780.