

A NATURAL VARIATION OF THE STANDARD SECRETARY PROBLEM

Shoou-Ren Hsiau and Jiing-Ru Yang

National Changhua University of Education

Abstract: We consider a natural variation of the standard secretary problem: N groups of applicants are to be interviewed sequentially (in groups) by a manager and the manager wants to find a strategy which maximizes the probability of selecting the best applicant. We use the usual backward induction method to find the optimal strategy, which can be easily described and is a natural extension of the solution for the standard secretary problem.

Key words and phrases: Backward induction, optimal stopping, secretary problem.

1. Introduction

In the standard secretary problem, N rankable people apply for one secretary position and are interviewed sequentially in random order by a manager. It is assumed that the manager knows the relative ranks of the present and previous applicants at each stage and he must decide immediately whether to accept or to reject the present applicant. No recall for previous applicants is permitted and the interviews continue until one of the applicants is accepted. The strategy which maximizes the probability of selecting the best of the N applicants is well-known, and is to reject the first $r - 1$ applicants and then accept the next one who is preferable to all his/her predecessors, where

$$r = \min\left\{n \mid \frac{1}{n} + \cdots + \frac{1}{N-1} \leq 1\right\}. \quad (1.1)$$

Under this strategy, the probability of selecting the best applicant is

$$\frac{r-1}{N} \sum_{k=r}^N \frac{1}{k-1}. \quad (1.2)$$

It is easy to see that for large N , $r \approx N/e$ and the probability of selecting the best one is $\approx 1/e$. The first published solution is due to Lindley (1961).

In this paper, we consider the following natural variation. There are $l_1 + \cdots + l_N$ rankable applicants who are divided randomly into N groups, the i th group containing l_i members. At the i th stage, the manager interviews all members of

the i th group and then he must decide immediately whether to select the best member in the i th group. If so, then the interview ends, otherwise he continues to interview the members of the $(i + 1)$ th group. At each stage the manager is assumed to know the relative ranks of the applicants who have been interviewed so far, but no recall is permitted. The objective is to find a strategy which maximizes the probability of selecting the best applicant. It is not surprising that the optimal strategy is to reject the applicants in the first $r - 1$ groups (for some r) and accept the best one in the next group which contains the one who is preferable to all the predecessors. But it came as a surprise to us that the threshold r is determined by the following simple formula

$$r = \min\{n \mid \sum_{k=n+1}^N \frac{l_k}{b_{k-1}} \leq 1\}, \text{ where } b_k = \sum_{i=1}^k l_i.$$

Our variation seems more realistic than the standard version since interviews are often held in groups. There are many other variations of the standard secretary problem in the literature. For example, Gilbert and Mosteller (1966) studied the consequences of allowing the manager to choose more than one of the N applicants; Dynkin and Yushkevich (1969) outlined a proof of the optimal strategy for the variation in which the manager aims to select an applicant whose overall rank is less than or equal to s , $s = 2$; and for $s = 3$, Quine and Law (1996) and Yang (1998) derived exact results by different methods. The interested reader is referred to the review papers by Freeman (1983) and Ferguson (1989) for additional references.

We organize the paper as follows. In Section 2, we review the usual backward induction method, which is developed extensively by Chow, Robbins and Siegmund (1971). Section 3 is devoted to proving the optimal strategy. In Section 4, we consider some interesting examples and asymptotic properties.

2. Preliminaries

In this section, we shall establish some basic general results. Suppose that there is given a probability space (Ω, \mathcal{F}, P) , an increasing sequence $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_N$ of sub- σ -algebras of \mathcal{F} , and a sequence X_1, \dots, X_N of integrable random variables such that X_n is \mathcal{F}_n -measurable, $n = 1, \dots, N$. A *stopping rule* T is a random variable taking values in $\{1, \dots, N\}$ such that the event $\{T = n\} \in \mathcal{F}_n$, $n = 1, \dots, N$. Denote by C the class of all stopping rules and define $V = \sup_{T \in C} E(X_T)$. We shall say that a stopping rule S is optimal in C if $S \in C$ and $E(X_S) = V$. The following theorem provides a method (usually referred to as backward induction) to find an optimal stopping rule for $\{X_n, \mathcal{F}_n\}_1^N$; for a proof, see Chow, Robbins and Siegmund (1971).

Theorem 2.1. *Let N be a fixed positive integer. Define successively $\gamma_N, \dots, \gamma_1$ by setting $\gamma_N = X_N$, and $\gamma_n = \max\{X_n, E(\gamma_{n+1}|\mathcal{F}_n)\}$, $n = N-1, \dots, 1$. Let S be the smallest n such that $X_n = \gamma_n$. Then S is optimal in C and $V = E(X_S) = E(\gamma_1)$.*

Let Z_1, \dots, Z_N be integer-valued random variables with each Z_n being \mathcal{F}_n -measurable. Set $X_n = P\{Z_n = 1|\mathcal{F}_n\}$. Then for any stopping rule T , we have $E(X_T) = P\{Z_T = 1\}$. Hence choosing a stopping rule T to maximize the probability $P\{Z_T = 1\}$ is equivalent to solving the optimal stopping problem for $\{X_n, \mathcal{F}_n\}_1^N$ with $X_n = P\{Z_n = 1|\mathcal{F}_n\}$. In Section 3, Z_n will be the absolute rank of the best member in the n th group.

3. The Optimal Strategy

Let A_k be the absolute rank of the k th applicant for $k = 1, \dots, b_N$. Setting $M = b_N$, by assumption, (A_1, \dots, A_M) is a random permutation of the integers $1, \dots, M$, all $M!$ permutations being equally likely. Let Y_k be the relative rank of the k th applicant among the first k people. Let $Z_n = \min\{A_{b_{n-1}+1}, \dots, A_{b_n}\}$, the absolute rank of the best member in the n th group, and $\mathcal{F}_n = \sigma(Y_1, \dots, Y_{b_n})$. By Theorem 2.1, the optimal strategy is determined by the optimal stopping rule for $\{X_n, \mathcal{F}_n\}_1^N$ with $X_n = P\{Z_n = 1|\mathcal{F}_n\}$. We state two simple lemmas without proof. (The first one is due to Rényi (1962) and the second one an easy consequence of the first.)

Lemma 3.1. *Y_1, \dots, Y_M are independent random variables and for each $k = 1, \dots, M$,*

$$P\{Y_k = j\} = 1/k, \quad j = 1, \dots, k. \tag{3.1}$$

Lemma 3.2. *For any $n = 1, \dots, N$,*

$$\begin{aligned} X_n &= P\{Z_n = 1|\mathcal{F}_n\} \\ &= \begin{cases} c_n = b_n/M, & \text{if } Y_k = 1 \text{ for some } k, \quad b_{n-1} + 1 \leq k \leq b_n \\ 0, & \text{if } Y_k \neq 1 \text{ for all } k, \quad b_{n-1} + 1 \leq k \leq b_n, \end{cases} \end{aligned} \tag{3.2}$$

where we use the convention $b_0 = 0$.

From Lemma 3.2, note that X_n is a function of $Y_{b_{n-1}+1}, \dots, Y_{b_n}$. Now, as in Theorem 2.1, define

$$\begin{aligned} \gamma_N &= X_N \\ \gamma_n &= \max\{X_n, E(\gamma_{n+1}|\mathcal{F}_n)\}, \quad n = N-1, \dots, 1. \end{aligned}$$

Since the Y 's are independent, \mathcal{F}_{N-1} and $Y_{b_{N-1}+1}, \dots, Y_{b_N}$ are independent. This implies $E(\gamma_N|\mathcal{F}_{N-1}) = E(\gamma_N)$, since $\gamma_N = X_N$ is a function of $Y_{b_{N-1}+1}, \dots, Y_{b_N}$.

Consequently, $\gamma_{N-1} = \max\{X_{N-1}, E(\gamma_N)\}$ is a function of X_{N-1} , and hence a function of $Y_{b_{N-2}+1}, \dots, Y_{b_{N-1}}$. By a similar argument and backward induction, it is not difficult to show that γ_n is a function of $Y_{b_{n-1}+1}, \dots, Y_{b_n}$ and $E(\gamma_{n+1}|\mathcal{F}_n) = E(\gamma_{n+1}) \equiv V_{n+1}$ for $n = 1, \dots, N - 1$. Then if $V_{N+1} = 0$,

$$\gamma_n = \max\{X_n, V_{n+1}\}, \quad n = 1, \dots, N. \tag{3.3}$$

Observe that $E(\gamma_1) \equiv V_1 \geq V_2 \geq \dots \geq V_N \geq V_{N+1} = 0$ and $0 \equiv c_0 < c_1 < c_2 < \dots < c_N = b_N/M = 1$. So we can find a unique positive integer $r, 1 \leq r \leq N$, such that $c_i \geq V_{i+1}$ whenever $r \leq i \leq N$ and $c_{r-1} < V_r$.

The optimal stopping rule S defined in Theorem 2.1 can be described as

$$S = \text{the smallest } n \geq r \text{ such that } Y_k = 1 \text{ for some } k, b_{n-1} + 1 \leq k \leq b_n.$$

By (3.2), (3.3) and the definition of r , for each $n = r, \dots, N$,

$$\gamma_n = \max\{X_n, V_{n+1}\} = \begin{cases} b_n/M, & \text{if } Y_k = 1 \text{ for some } k, b_{n-1} + 1 \leq k \leq b_n, \\ V_{n+1}, & \text{if } Y_k \neq 1 \text{ for all } k, b_{n-1} + 1 \leq k \leq b_n. \end{cases}$$

Now for $r \leq n \leq N$, the V_n satisfy the following recursive formula with $V_{N+1} = 0$:

$$\begin{aligned} V_n &= E(\gamma_n) = (b_n/M)P\{Y_k = 1 \text{ for some } b_{n-1} + 1 \leq k \leq b_n\} \\ &\quad + V_{n+1}P\{Y_k \neq 1 \text{ for all } b_{n-1} + 1 \leq k \leq b_n\} \\ &= (b_n/M)(l_n/b_n) + V_{n+1}(b_{n-1}/b_n) = (l_n/M) + (b_{n-1}/b_n)V_{n+1}. \end{aligned} \tag{3.4}$$

Solving (3.4) yields

$$V_n = \frac{b_{n-1}}{M} \sum_{k=n}^N \frac{l_k}{b_{k-1}}, \quad r \leq n \leq N. \tag{3.5}$$

(If $r = 1$ then, from (3.4), $V_1 = l_1/M$, though some ambiguity arises at (3.5).) In view of (3.5) and the definition of r , we have

$$\begin{aligned} \frac{b_r}{M} = c_r &\geq V_{r+1} = \frac{b_r}{M} \sum_{k=r+1}^N \frac{l_k}{b_{k-1}} \\ \text{and} \quad \frac{b_{r-1}}{M} = c_{r-1} &< V_r = \frac{b_{r-1}}{M} \sum_{k=r}^N \frac{l_k}{b_{k-1}}. \end{aligned}$$

Equivalently, $1 \geq \sum_{k=r+1}^N (l_k/b_{k-1})$ and $1 < \sum_{k=r}^N (l_k/b_{k-1})$. Hence

$$r = \min\{n \mid \sum_{k=n+1}^N \frac{l_k}{b_{k-1}} \leq 1\}. \tag{3.6}$$

On the other hand, from (3.3) and the definition of r , for each $n = 1, \dots, r-1$, $\gamma_n = \max\{X_n, V_{n+1}\} = V_{n+1}$, which implies that

$$V_1 = V_2 = \dots = V_r. \tag{3.7}$$

Now by (3.1), (3.5), (3.7) and Theorem 2.1, we have

$$P\{Z_S = 1\} = E(X_S) = E(\gamma_1) = V_1 = V_r = \frac{b_{r-1}}{M} \sum_{k=r}^N \frac{l_k}{b_{k-1}}. \tag{3.8}$$

This establishes the following theorem.

Theorem 3.1. *Let r be defined as in (3.6). Then the strategy which maximizes the probability of selecting the best applicant is as follows: the manager should reject the applicants in the first $r - 1$ groups and accept the best one in the next group which contains the one who is preferable to all his/her predecessors. Under this strategy, the probability of selecting the best applicant is*

$$\frac{b_{r-1}}{M} \sum_{k=r}^N \frac{l_k}{b_{k-1}}.$$

As pointed out by a referee, this theorem can also be derived elegantly from the generating function approach of Bruss (1984a). The original version of our paper was completed in January, 1998. The main result, Theorem 3.1, was later generalized by Bruss (1998) in the following elegant form.

Theorem (Bruss (1998)). *Let I_1, \dots, I_N be a sequence of independent indicator functions with $p_j = E(I_j)$. Let $q_j = 1 - p_j$ and $r_j = p_j/q_j$. Then an optimal rule τ_N for stopping on the last success exists and stops on the first index (if any) k with $I_k = 1$ and $k \geq s$, where*

$$s = \sup\{1, \sup\{1 \leq k \leq N : \sum_{j=k}^N r_j \geq 1\}\}.$$

The optimal reward is given by $V(N) = \prod_{j=s}^N q_j \sum_{j=s}^N r_j$.

While Bruss's result includes Theorem 3.1 as a special case, it can be proved by observing that Theorem 3.1 is essentially equivalent to Bruss's result with the p_j 's rational. The general case can then be approximated by the rational case.

4. Examples and Approximations

In this section, we apply the previous result to three examples. Incidentally, we find that b_r/b_N and $P\{Z_S = 1\}$ have the same asymptotic behavior under some weak conditions.

Example 1. If $l_k = l$ for each $k = 1, \dots, N$, then $b_k = \sum_{i=1}^k l_i = kl$ and $r = \min\{n \mid \sum_{k=n+1}^N (k-1)^{-1} \leq 1\}$, as in the standard secretary problem.

Example 2. If $l_k = k$, $1 \leq k \leq N$, then $b_k = \sum_{i=1}^k l_i = k(k+1)/2$ and

$$r = \min\{n \mid \sum_{k=n+1}^N \frac{1}{k-1} \leq \frac{1}{2}\}.$$

Under the optimal strategy, the probability of selecting the best one is

$$P\{Z_S = 1\} = \frac{2r(r-1)}{N(N+1)} \sum_{k=r}^N \frac{1}{k-1}.$$

When N is large, we get $r \approx N/\sqrt{e}$, $b_r/b_N \approx 1/e$ and $P\{Z_S = 1\} \approx 1/e$, as in the standard one.

(As a referee pointed out, this stability of the $1/e$ answer in solutions of various versions of the secretary problem was already observed in the case of an unknown number of candidates (Bruss (1984b)) as well as in several examples of Samuel-Cahn (1995) for the model of hiring freeze.) In fact, for our problem we have the following result which can explain the asymptotic behaviors of b_r/b_N and $P\{Z_S = 1\}$ in the above two examples.

Theorem 4.1. *Let r be defined as in (3.6). Then*

$$b_r/b_N \geq P\{Z_S = 1\} \geq 1/e. \tag{4.1}$$

Moreover, setting $l = \max_{r \leq i \leq N} l_i$, then if $l/b_N < 1/2e$, we have

$$0 \leq P\{Z_S = 1\} - \frac{1}{e} \leq \frac{b_r}{b_N} - \frac{1}{e} \leq \left(\frac{2e-1}{e}\right)\left(\frac{l}{b_N}\right). \tag{4.2}$$

Proof. By (3.6), $1 \geq \sum_{k=r+1}^N (l_k/b_{k-1})$. Hence

$$\ln \frac{b_N}{b_r} = \int_{b_r}^{b_N} \frac{1}{x} dx \leq \sum_{k=r+1}^N \frac{l_k}{b_{k-1}} \leq 1,$$

which implies that $b_r/b_N \geq 1/e$. Next we prove that $P\{Z_S = 1\} \leq b_r/b_N$. By (3.8) and (3.6) it follows that

$$\begin{aligned} P\{Z_S = 1\} &= \frac{b_{r-1}}{M} \sum_{k=r}^N \frac{l_k}{b_{k-1}} = \frac{b_{r-1}}{b_N} \left(\frac{l_r}{b_{r-1}} + \sum_{k=r+1}^N \frac{l_k}{b_{k-1}} \right) \\ &= \frac{b_r}{b_N} + \frac{b_{r-1}}{b_N} \sum_{k=r+1}^N \frac{l_k}{b_{k-1}} \\ &\leq \frac{l_r}{b_N} + \frac{b_{r-1}}{b_N} (1) = \frac{b_r}{b_N}. \end{aligned}$$

To complete the proof of (4.1), it remains to show $P\{Z_S = 1\} \geq 1/e$. This is a simple consequence of the fact that $P\{Z_S = 1\}$ is at least as large as the probability of selecting the best one among $l_1 + \dots + l_N$ applicants for the standard secretary problem and the well-known fact that, under the standard version, the (optimal) probability of selecting the best one decreases to $1/e$ as the number of applicants increases to infinity.

Furthermore, when $l/b_N < 1/2e$, since $b_r/b_N \geq 1/e$, we have $b_{r-1} - l = (b_r - l_r) - l \geq b_N/e - 2l > 0$. It follows that $0 < b_n - l \leq b_{n-1}$ for all $n = r, \dots, N$. Then we have $1 < \sum_{k=r}^N (l_k/b_{k-1}) \leq \sum_{k=r}^N [l_k/(b_k - l)] \leq \int_{b_{r-1}}^{b_N} (x - l)^{-1} dx$, and consequently $1 \leq \ln[(b_N - l)/(b_{r-1} - l)]$, i.e., $b_N \geq e(b_{r-1} - l) + l$. It follows that

$$\begin{aligned} \frac{b_r}{b_N} - \frac{1}{e} &= \frac{eb_r - b_N}{eb_N} \leq \frac{eb_r - e(b_{r-1} - l) - l}{eb_N} \\ &\leq \frac{el_r + el - l}{eb_N} \leq \frac{(2e - 1)}{e} \left(\frac{l}{b_N}\right). \end{aligned}$$

Acknowledgement

The authors wish to thank the referee and Associate Editor for their comments and suggestions that improved the paper and made it more readable.

References

- Bruss, F. Thomas (1984a). Patterns of relative maxima in random permutations. *Ann. Soc. Scient. de Bruxelles* **98**, 19-28.
- Bruss, F. Thomas (1984b). A unified approach to a class of best choice problems with an unknown number of options. *Ann. Probab.* **12**, 882-889.
- Bruss, F. Thomas (1998). Sum the odds to one and stop. *Research report ISRO 1998-32*, No 155, Universite Libre de Bruxelles.
- Chow, Y. S., Robbins, H. and Siegmund, D. (1971). *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin, New York.
- Dynkin, E. B. and Yushkevich, A. (1969). *Markov Processes: Theorems and Problems*. 1st edition. Plenum, New York.
- Freeman, P. R. (1983). The secretary problem and its extensions—a review. *Int. Statist. Rev.* **51**, 189-206.
- Ferguson, T. S. (1989). Who solved the secretary problem? *Statist. Sci.* **4**, 282-296.
- Gilbert, J. and Mosteller, F. (1966). Recognizing the maximum of a sequence. *J. Amer. Statist. Assoc.* **61**, 35-73.
- Lindley, D. V. (1961). Dynamic programming and decision theory. *Appl. Statist.* **10**, 35-51.
- Quine, M. P. and Law, J. S. (1996). Exact results for a secretary problem. *J. Appl. Probab.* **33**, 630-639.
- Rényi, A. (1962). Théorie des éléments saillants dans une suite d'observations, *Proc. Coll. Comb. Methods* (Aarhus Universitet), 104-115.
- Samuel-Cahn, E. (1995). The best-choice secretary problem with random freeze on jobs. *Stochastic Processes and Their Applications* **55**, 315-327.

Yang, J. R. (1998). A generalized secretary problem. (In Chinese). *J. Chinese Statist. Assoc.* **36**, 79-92.

Department of Mathematics, National Changhua University of Education, Changhua, 500, Taiwan.

E-mail: srhsiau@math.ncue.edu.tw

E-mail: m8521204@math.ncue.edu.tw czw2@stat.sinica.edu.tw

(Received January 1999; accepted August 1999)