BULLETIN OF THE
INSTITUTE OF MATHEMATICS
ACADEMIA SINICA
Volume 17, Number 2, June 1989

ON SOME OPTIMAL STOPPING AND ITS APPLICATIONS TO SOME STATISTICAL INFERENCES

BY

WEN-TAO HUANG (黃文濤) AND CHING-YUN WANG (王清雲)

Abstract. We consider n populations rankable according to some quantity related to its distribution functions instead of usual secretary problem of n rankable candidates. Under this situation we can only observe relative ranks of some statistics of observations from each population. Some optimal stopping rules which minimize the expected absolute rank of the population selected are studied. We also consider some types of loss function for the usual formulation of the secretary problem. The solution for the infinite secretary problem is also provided by differential equations.

1. **Introduction.** A finite secretary problem is usually defined in the following way. A number of n rankable candidates appear sequentially in random order. We observe only their relative ranks. On the basis of their observed relative ranks and some loss function, each candidate is either selected or passed by and a passed candidate can never be recalled (we do not consider the case of a recall). If the $(n-1)$-th candidate is passed, the last candidate must be selected.

Let C be the set of all stopping rules, and let X_τ be the absolute rank of candiate selected applying the stopping rule $\tau \in C$. In 1961, Lindley [3] proposed a stopping rule to minimize the risk $E(X_\tau)$. And it is then shown that (see [1]) as $n \to \infty$

$$\min_{\tau \in C} E(X_\tau) = \prod_{j=1}^{\infty} \left(\frac{j+2}{j}\right)^{1/j+1} \simeq 3.8695.$$

Let $\mu_1, \mu_2, \cdots, \mu_n$ be n different populations such that μ_i is associated with the i-th population and $\mu_{[1]} < \mu_{[2]} < \cdots < \mu_{[n]}$ be the

Received by the editors January 8, 1987.
Key words: Secretary problem, Statistical relative rank, Loss function, Selection and ranking.
AMS Subject Classification: 60G40.

ordered values by the ascending order. Then, let these populations be randomly ordered. Let the parameter associated with the i-th population (after randomly ordered) be denoted by θ_i. We have

$$\Pr\{\theta_1 = \mu^{-1}_{\sigma(1)}, \cdots, \theta_n = \mu^{-1}_{\sigma(n)}\} = \frac{1}{n!},$$

where σ is any permutation of $\{1, 2, \cdots, n\}$.

Let T_i denote some statistics of the i-th population, $i = 1, \cdots, n$. We assume each sample is of size m, and for each pair of T_i, T_j, there is no tie. At time i, we observe i relative ranks of the first i statistic of $T_1, T_2, \cdots, T_i$ according to their descending order, and we either select the i-th population or pass on to time $i+1$. If the former case occurs, the process of selection terminates, and for the latter case, we continue sampling and observe its relative rank with respect to the preceding $(i+1)$ values. Let X_τ denote the absolute rank of the population selected applying some stopping rule $\tau \in C$, then X_τ takes values in $\{1, 2, \cdots, n\}$ with some probability determined by τ. Now, the question is which stopping rule τ_0 in C would minimize $E(X_\tau)$, the expected absolute rank of the population selected.

For our convenience, we restrict ourselves for the time being to the normal case with mean μ_i and variance σ^2, common unknown. We can see later on that this restriction is not necessary. Let T_i denote the i-th sample mean and its corresponding population mean is denoted by θ_i. Let Z_i denote the relative rank of T_i among $T_1, T_2, \cdots, T_i$ and let Y_i denote the unobservable relative rank of θ_i among $\theta_1, \theta_2, \cdots, \theta_i$. At time i, Z_i is observed. Since at each time i, we must either accept the present population or take another observation, the backward argument of Lindley [3] can thus be applied.

For $i = 0, 1, \cdots, n-1$, define c_i to be the minimal expected absolute rank of population mean selected confining ourselves to the set of stopping rules such that $\tau \geq i+1$. We can also obtain the "cut off point" rule which will be defined later (see [4]).

Let the minimal expected absolute rank in the finite secretary problem given by [3] be denoted by c_0^*. In our problem, observing the sample means, we want to choose the population with mean associated with minimal expected rank. We see that as sample size m increases to infinity, we have $c_0^* = c_0$ a. s. by the Strong Law of Large Numbers. From our practical point of view, we are interested in knowing how large the sample size is sufficient for practical application. In this paper some tables are given to show some idea for determining suitable sample size for our problem.

Finally, we consider the secretary problem with loss function $q(X_\tau)\,h(\tau)$, where $q(\cdot)$ is positive monotone increasing and $h(\cdot)$ is some cost function, Our goal is to find a stopping rule in C' that attains the minimal expected loss, where

$$C' = \{\tau \mid E[q(X_\tau)\,h(\tau)] < \infty\}\,.$$

2. **A basic Lemma.** It is well-known that $Y_1, Y_2, \cdots, Y_n$ are independent with $\Pr\{Y_i = j\} = 1/i$ $(j = 1, 2, \cdots, i)$. For any realization of $\theta_1, \theta_2, \cdots, \theta_{i-1}$ which divide the real line into i intervals, the probability of θ_i falling in each of the i intervals is equal to $1/i$. At time i, the random variables $T_1, T_2, \cdots, T_{i-1}$ divide the real line into i intervals and we desire to evaluate the probability that T_i falls in each interval, i. e. the distribution of Z_i given $Z_1, Z_2, \cdots, Z_{i-1}$.

LEMMA 2.1. *For any i $(1 \le i \le n)$,*

$$\Pr\{Z_i = j \mid Z_1, Z_2, \cdots, Z_{i-1}\} = \Pr\{Z_i = j\}$$
$$= \frac{1}{i}, \quad j = 1, 2, \cdots, i.$$

Proof. For fixed i and j, define

$$A = \{\tilde{\sigma} \mid \tilde{\sigma} \text{ is a permutation of } \{1, 2, \cdots, i\} \text{ such the that } T_{\tilde{\sigma}(i)} < \cdots < T_{\tilde{\sigma}(j)} = T_i < \cdots < T_{\tilde{\sigma}(1)}\}\,.$$

$$B = \{(b_1, b_2, \cdots, b_i) \mid b_r \ne b_s,\ b_t \in \{1, 2, \cdots, n\}\}\,.$$

Then, the cardinality of B is $|B| = \binom{n}{i}$ and for any $\underset{\sim}{b} \in B$, $\Pr\{\theta_1 = \mu_{b_1}, \cdots, \theta_i = \mu_{b_i}\} = 1\Big/\binom{n}{i}$.

For given $\underset{\sim}{b} = (b_1, b_2, \cdots, b_i) \in B$, let

$$C_{\underset{\sim}{b}} = \{\sigma \mid \sigma \text{ is a permutation of } (b_1, b_2, \cdots, b_i)\}\,.$$

Then, for any $\tilde{\sigma} \in A$,

$$\sum_{\sigma\in \underset{\sim}{C_b}} \Pr\{T_{\tilde{\sigma}(i)} <\cdots< T_{\tilde{\sigma}(j)} = T_i <\cdots< T_{\tilde{\sigma}(1)} | (\theta_1, \theta_2, \cdots, \theta_i) = (\mu_{\sigma(b_1)}, \cdots, \mu_{\sigma(b_i)})\} = 1.$$

Since

$$\Pr\{(\theta_1, \theta_2, \cdots, \theta_i) = (\mu_{\sigma(b_1)}, \cdots, \mu_{\sigma(b_i)})\} = 1\Big/\left\{\binom{n}{i}\cdot i\right\}$$

and $|A| = (i-1)!$, we have

$$\begin{aligned}\{Z_i = j\} &= 1\Big/\left\{\binom{n}{i}\cdot i!\right\} \sum_{\tilde{\sigma}\in A} \sum_{\underset{\sim}{b}\in B} \sum_{\sigma\in \underset{\sim}{C_b}} \Pr\{T_{\tilde{\sigma}(i)} <\cdots< T_i = T_{\tilde{\sigma}(j)} <\cdots< T_{\tilde{\sigma}(1)} | (\theta_1, \theta_2, \cdots, \theta_i) = (\mu_{\sigma(b_1)}, \cdots, \mu_{\sigma(b_i)})\} \\ &= \left\{\binom{n}{i}(i-1)!\right\}\Big/\left\{\binom{n}{i}\cdot i!\right\} \\ &= 1/i.\end{aligned}$$

Follow the analogous argument and applying the mathematical induction, we conclude that

$$\Pr\{Z_i = j | Z_1, Z_2, \cdots, Z_{i-1}\} = \Pr\{Z_i = j\} = \frac{1}{i}.$$

We finally conclude the following

COROLLARY 2.1. *For each i $(1 \le i \ge n)$ and a, b $(1 \le a, b \le i)$,*

$$\Pr\{Y_1 = b | Z_i = a\} = \Pr\{Z_i = a | Y_i = b\}.$$

Proof.

$$\begin{aligned}\{Y_i = b | Z_i = a\} &= \frac{\Pr\{Y_i = b, Z_i = a\}}{\Pr\{Z_i = a\}} = \frac{\Pr\{Y_i = b, Z_i = a\}}{1/i} \\ &= \frac{\Pr\{Y_i = b, Z_i = a\}}{\Pr\{Y_i = b\}} = \Pr\{Z_i = a | Y_i = b\}.\end{aligned}$$

For our convenience, we define

$$p^{(i)}(y, z) = \Pr\{Z_i = z | Y_i = y\} \tag{2.3}$$

$$h^{(i)}(z, y) = \Pr\{Y_i = y | Z_i = z\} \tag{2.4}$$

$$P^{(i)} = \begin{bmatrix} p^{(i)}(1, 1), \cdots, p^{(i)}(1, i) \\ \vdots \\ p^{(i)}(i, 1), \cdots, p^{(i)}(i, i) \end{bmatrix} \tag{2.5}$$

$$H^{(i)} = \begin{bmatrix} h^{(i)}(1, 1), \cdots, h^{(i)}(1, i) \\ \vdots \\ h^{(i)}(i, 1), \cdots, h^{(i)}(i, i) \end{bmatrix}. \tag{2.6}$$

We note that both matrices $P^{(i)}$ and $H^{(i)}$ are stochastic matrices and $H^{(i)} = (P^{(i)})^T$, the transpose of $P^{(i)}$.

3. **An optimal stopping rule.** Let X_i denote the absolute rank of θ_i. We note that $X_i = Y_i + k_i$, where k_i equals to the number of elements in $\{\theta_{i+1}, \cdots, \theta_n\}$ which is greater than θ_i. Since the distribution of X_i is independent of Z_i if Y_i is given, we have

LEMMA 3.1. $\Pr\{X_i = x \mid Y_i = y, Z_i = z\} = \Pr\{X_i = x \mid Y_i = y\}$, *for* $x \geq y$.

For each i $(1 \leq i \leq n)$, define

$$g^{(i)}(y, x) = \Pr(X_i = x \mid Y_i = y) \tag{3.1}$$

$$G_n^{(i)} = \begin{bmatrix} g^{(i)}(1, 1), \cdots, g^{(i)}(1, n) \\ g^{(i)}(2, 1), \cdots, g^{(i)}(2, n) \\ \vdots \\ g^{(i)}(i, 1), \cdots, g^{(i)}(i, n) \end{bmatrix}. \tag{3.2}$$

We can show easily that

$$\sum_{x=y}^{n} x \cdot \Pr\{X_i = x \mid Y_i = y\} = ((n+1)/(i+1))y, \quad \text{i.e.}$$

$$(g(y, 1), \cdots, g(y, n)) \begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix} = ((n+1)/(i+1))y \tag{3.3}$$

for $y = 1, 2, \cdots, i$.

We therefore obtain the following

THEOREM 3.1. *For* $z = 1, 2, \cdots, i$

$$E[X_i \mid Z_i = z] = ((n+1)/(i+1))\, \underset{\sim}{h}^{(i)}(z) \cdot \underset{\sim}{i} \tag{3.4}$$

where $\underset{\sim}{h}^{(i)}(z) = (h^{(i)}(z, 1), \cdots, h^{(i)}(z, i))$, *the* z-th *row vector of* $H^{(i)}$

defined by (2.6) *and* $\underset{\sim}{i}=(1, 2,\cdots, i)'$.

Proof. Note that

$$\begin{aligned}
&\Pr\{X_i=x|Z_i=z\}\\
&\quad=\sum_{y=1}^{i}\Pr\{X_i=x|Y_i=y,\ Z_i=z\}\cdot\Pr\{Y_i=y|Z_i=z\}\\
&\quad=\sum_{y=1}^{i}\Pr\{X_i=x|Y_i=y\}\cdot\Pr\{Y_i=y|Z_i=z\}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{(by Lemma 3.1)}\\
&\quad=(h^{(i)}(z,1),\ \cdots,\ h^{(i)}(z,i))(g^{(i)}(1,x),\cdots,\ g^{(i)}(i,x))'\\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{(defined by (3.2))}.
\end{aligned}$$

Hence,

$$\begin{aligned}
E(X_i|Z_i=z)&=\sum_{x=1}^{n}(h^{(i)}(z,1),\cdots,\ h^{(i)}(z,i))\\
&\qquad\cdot(g^{(i)}(1,x),\cdots,\ g^{(i)}(i,x))'\cdot x\\
&=\underset{\sim}{h}^{(i)}(z)\cdot G_n^{(i)}\cdot\underset{\sim}{n}\qquad\text{(defined by (3.2))}\\
&=\underset{\sim}{h}^{(i)}(z)((n+1)/(i+1),\\
&\qquad 2(n+1)/(i+1),\cdots,\ i(n+1)/(i+1))'\\
&\qquad\qquad\qquad\qquad\text{(by (3.3))}\\
&=((n+1)/(i+1))\,\underset{\sim}{h}^{(i)}(z)\cdot\underset{\sim}{i}.
\end{aligned}$$

REMARK 3.1. *If we take* $h^{(i)}(z,y)=\begin{cases}1 & \text{if } y=z\\ 0 & \text{if } y\neq z\end{cases}$, *then it becomes the classic secretary problem.*

COROLLARY 3.1. *For* $i=1, 2,\cdots, n$, *let*

$$E(X_i|Z_i=z)=f_i(z)\qquad(1\le z\le i), \tag{3.5}$$

then f *is an increasing function of* z.

For any stopping rule τ, the expected absolute rank of the population selected is given by

$$\begin{aligned}E(X_\tau)=E[((n+1)/(\tau+1))\cdot(h^{(\tau)}(Z_\tau,1),\cdots,\ h^{(\tau)}(Z_\tau,\tau))\\ \cdot(1,2,\cdots,\tau)'].\end{aligned} \tag{3.6}$$

We are desirable to find some $\tau\in C$ to minimize $E(X_\tau)$.

We use the backward induction argument to find an optimal stopping. Let c_i denote the expected absolute rank of population selected if we confine ourselves to stopping such that $\tau \geq i+1$ $(i=0, 1, 2, \cdots, n-1)$. Clearly, c_0 is the minimal expected absolute rank among all stoppings. Note that $c_{n-1}=E(X_n)=(n+1)/2$ since $\theta_1, \theta_2, \cdots, \theta_n$ are in random order. Also

$$\begin{aligned} c_{i-1} &= E[\min(E(X_i|Z_i=z), c_i)] \\ &= E[\min(f_i(z), c_i)] \\ &= \sum_{z=1}^{i} \min[f_i(z), c_i]/i \\ &\leq c_i. \end{aligned}$$

Define

$$f_i(0)=0$$

and

(3.7) $$s_i = \max\{z|f_i(z) \leq c_i, z=0, 1, 2, \cdots, i\}.$$

Then by (3.5)

(3.8) $$c_{i-1} = \frac{1}{i}\left\{\sum_{z=1}^{s_i} f_i(z) + (i-s_i)\, c_i\right\}.$$

Since $f_1(1)=(n+1)/2$, hence, $c_0=E\{\min((n+1)/2, c_1)\}=c_1$. Clearly,

(3.9) $$\begin{aligned} &c_0 = c_1 \leq \cdots \leq c_{n-1} = (n+1)/2, \\ &0 = s_1 \leq s_2 \leq \cdots \leq s_{n-1} = [n/2]. \end{aligned}$$

We note that if the second last population is not selected and passed, the last population must be selected according to our rule. Hence, $s_n=n$ and we conclude the following stopping rule of τ_0 to minimize $E(X_\tau)$. The expected absolute rank of the population selected using the stopping defined by Theorem 3.2 is thus c_0.

THEOREM 3.2. *In order to minimize $E(X_\tau)$ defined by* (3.6), *the optimal choice of τ_0 is defined as follows.*

(3.10) $$\tau_0 = \min\{i \geq 1 | Z_i \leq s_i\}$$

where the increasing sequence of $\{s_i\}_1^{n-1}$ is defined by (3.7) *and the associated increasing sequence of $\{c_i\}$ is defined by* (3.8).

Note that $f_i(z)=((n+1)/(i+1))\, \underset{\sim}{h}^{(k)}(z) \cdot \underset{\sim}{i}$ which is given by

Theorem 3.1. Therefore, as long as the matrix $H^{(i)}$ (defined by (2.6)) is given, numerical computation of τ_0 is always possible.

REMARK 3.2. Note that for finding the optimal τ_0 in Theorem 3.2, the assumption of normality of random observation can be discarded. Specification of *cdf* of random observations involves only the quantity of $H^{(i)}$ defined by (2.6). By some theories of statistics, the quantity of $H^{(i)}$ can always be computed or approximated.

Define

$$\delta(n) = c_0(n) - c_0^*(n) \geq 0 \tag{3.11}$$

where c_0^* is the minimum expected absolute rank given in [3]. In [1], a direct proof is given that $c_0^*(n)$ is strictly increasing in n. By the same argument as given in [1], we have the following

COROLLARY 3.2. *If all quantities* $\mu_1, \mu_2, \cdots, \mu_{n+k}$ *are different, then* $c_0(n) < c_0(n+k)$ *for* $k = 1, 2, \cdots$.

EXAMPLE 3.1. Consider $n = 3$ and μ_1, μ_2, μ_2 denote the respective means of three normal populations with common variance 1. Suppose $\mu_{[3]} - \mu_{[2]} = \mu_{[2]} - \mu_{[1]} = 1$. Let the three populations be randomly ordered. To find an optimal stopping rule τ_0 for our problem, it suffices to find (s_1, s_2, s_3). For $m = 1$, we take T_i to be our observation, then we have

$$P^{(2)} = \begin{pmatrix} 0.814 & 0.186 \\ 0.186 & 0.814 \end{pmatrix} \quad \text{(defined by (2.5))}.$$

Hence,

$$H^{(2)} = \begin{pmatrix} 0.814 & 0.186 \\ 0.186 & 0.814 \end{pmatrix} \quad \text{(defined by (2.6))}.$$

Thus,

$$E(X_1 | Z_1 = 1) = 2$$
$$E(X_2 | Z_2 = 1) = f_2(1) = 1.581$$

and

$$E(X_2 | Z_2 = 2) = f_2(2) = 2.419.$$

We get

$$s_1 = 0,\ s_2 = 1,\ s_3 = 3$$

and

$$c_0 = c_1 = 1.791,\ c_2 = 2.$$

$\delta(3) = c_0(3) - c_0^*(3) = 1.791 - 1.667 = 0.124$. Our optimal rule is given by $(s_1, s_2, s_3) = (0, 1, 3)$.

4. **Some Monte Carlo results.** Practically, it's very tedious to compute the matrix $P^{(i)}$ (defind by (2.5)) for n greater than 3. For example, let $n=4$, $m=1$, and $\mu_{[i]} = i$ for $i = 1,\cdots, 4$. Then even computation for the term $p^{(3)}$ (2, 2) is not easy. In fact, we have $\Pr\{Z_3 = 2 | Y_3 = 2\} = \Pr\{T_1 < T_3 < T_2 | \theta_1 < \theta_3 < \theta_2\} + \Pr\{T_2 < T_3 < T_1 | \theta_1 < \theta_3 < \theta_2\} + \Pr\{T_1 < T_3 < T_2 | \theta_2 < \theta_3 < \theta_1\} + \Pr\{T_2 < T_3 < T_1 | \theta_2 < \theta_3 < \theta_1\}$. The term, $\Pr\{T_1 < T_3 < T_2 | \theta_1 = 1, \theta_2 = 3, \theta_3 = 2\}$ equals to integral

$$\int_{-\infty}^{\infty}\int_{x_1}^{\infty}\int_{x_3}^{\infty} \frac{1}{(2\pi)^{3/2}}\, e^{-((x_1-1)^2+(x_3-3)^2+(x_2-2)^2)/2}\, dx_2\, dx_3\, dx_1.$$

Therefore explicit form for $P^{(i)}$ is not easy to obtain. Hence, for a practical point of view, we usually get it by simulation.

For given n populations and fixed sample size m, let the associated values of s_i's defined by (3.7) be $(s_1(m),\cdots, s_n(m))$. Assume that the statistics T defined in section 1 is consistent so that as $m \to \infty$, $T_i \to \theta_i$ a.s. $(i = 1, 2,\cdots, n)$, and thus $Y_i = Z_i$. Under this limit situation, our problem becomes the classic secretary problem, and the optimal stopoping rule is the same as $(s_1^*,\cdots, s_n^*)$, denoted by τ^*, which is given by Lindley [3].

Realisitically, we are unable to take m to be infinite. Therefore, for $\varepsilon > 0$, we are desirable to know some behavior of m so that the expected absolute rank with stopping rule τ^* is bounded by $E_m(X_{\tau^*}) \le c^* + \varepsilon$. Using CDC CYBER 172 system, we sample from each normal population of size m and make a selection according to stopping rule $\tau^* = (s_1^*,\cdots, s_n^*)$ given by [3]. We repeat the same process 5000 times and take the average.

In Table 1, for various cases of normal populations, we tabulate the associated average of absolute rank of the population selected,

Table 1. Assume $\mu_1 < \mu_2 < \cdots < \mu_n$

	$n=10$, $c_0^*(10)=2.56$; $\mu_{i+1}-\mu_i=1$, $i=1,\cdots,9$; $\sigma_i^2=1$		$n=20$, $c_0^*(20)=3.00$; $\mu_{i+1}-\mu_i=1$, $i=1,\cdots,19$; $\sigma_i^2=1$		$n=10$, $c_0^*(10)=2.56$; $\mu_{i+1}-\mu_i=1$, $i=1,\cdots,4$; $=\frac{1}{2}$, $i=5,\cdots,9$; $\sigma_i^2=1$		$n=20$, $\sigma_i^2=1$; $\mu_{i+1}-\mu_i=1$, $i=1(1)6$; $=\frac{1}{2}$, $i=7(1)13$; $=\frac{1}{5}$, $i=14(1)19$		$n=5$, $\sigma_i^2=\frac{1}{2}$; $\mu_{i+1}-\mu_i=1$, $i=1,2$; $=\frac{1}{2}$, $i=3,4$; $c_0^*(5)=2.05$	
m	$\overline{X}_{\tau^*}$	MSE($\overline{X}_{\tau^*}$)	$\overline{X}_{\tau^*}$	MSE($\overline{X}_{\tau^*}$)	$\overline{X}_{\tau^*}$	MSE($\overline{X}_{\tau^*}$)	$\overline{X}_{\tau^*}$	MSE($\overline{X}_{\tau^*}$)	$\overline{X}_{\tau^*}$	MSE($\overline{X}_{\tau^*}$)
2	2.63	3.30	3.09	6.32	2.80	3.46	3.77	7.13	2.33	1.36
4	2.57	3.06	3.00	5.96	2.68	3.27	3.48	6.89	2.23	1.31
6	2.58	3.18	3.04	6.33	2.63	3.16	3.29	6.29	2.19	1.27
8	2.55	3.13	3.00	5.79	2.63	3.24	3.33	6.44	2.17	1.22
10	2.57	3.00	3.02	6.03	2.59	3.18	3.28	6.77	2.15	1.19
12	2.51	3.03	2.94	5.60	2.59	3.16	3.25	6.34	2.13	1.21
14	2.58	3.17	3.00	6.22	2.54	3.01	3.20	6.34	2.11	1.23
16	2.55	3.00	3.00	6.36	2.59	3.27	3.22	6.54	2.12	1.23
18	2.55	3.13	3.00	5.81	2.55	3.17	3.21	6.34	2.10	1.19
20	2.56	3.14	2.97	5.81	2.56	3.02	3.07	5.75	2.09	1.16

$\bar{X}_{\tau^*}$ and its associated mean square error (MSE($\bar{X}_{\tau^*}$)). These tables propose some idea for determining conservative m for the selection problem of selecting the population associated with the largest mean among n normal populations. In general, in order to satisfy $E_m(X_{\tau^*}) \leq c^* + \varepsilon$ for fixed $\varepsilon > 0$, we note that the smaller the value of $|\mu_i - \mu_j|$, the larger the integer m.

5. **Some type of loss function.** For a usual finite secretary problem, an increasing function $q(\cdot)$ is assigned so that $q(x)$ is considered as a loss for stopping when a candidate of absolute rank x is selected. We derive the expected loss of $q(X_i)$ given $Y_i = y$. For notational convenience, we denote $Q(i, y) = E(q(X_i) | Y_i = y)$. The solution τ^* for minimizing $E(q(X_\tau))$ depends only through the function $Q(i, y)$.

In this section we consider the situation that we observe the ith candidate directly from population instead of observing the statistic T_i in previous sections.

Let X_i and Y_i denote respectively the absolute and relative ranks of the i-th candidate ($i = 1, 2, \cdots, n$). Lorenzen [4] generalized the secretary problem with a larger class of loss function. In particular, he [5] solved the problem of minimizing expected rank with a linear interview cost function. The risk he considered consists of two parts, a loss function $q(\cdot)$ based on absolute ranks and a cost function $h_n(\cdot)$. His goal is to find a stopping rule τ^* such that $E[q(X_{\tau^*}) + h(\tau^*)]$ attains its minimum over a set of suitable τ. Realistically, the loss function $q(X_\tau) \cdot h_n(\tau)$ would be more meaningful under some situations. Suppose τ_0 is a stopping which selects the candidate with absolute rank 1 with total observing time τ_0 while some other stopping τ_1, selects candidate of absolute rank 2 with observing time τ_1 such that $\tau_1 < \frac{1}{2}\tau_0$. In many occasions, we consider τ_1 more efficient than τ_0 since it stops earlier than half of τ_0 even though it chooses the second best especially, when τ_0 is large and the cost is expensive. Thus we are interested in considering the type of loss function given by $q(X_\tau) \cdot h_n(\tau)$.

Let τ be a stopping rule taking values in $\{1, 2, \cdots, n\}$ based on

Y_i and let $V_n = \inf E[q(X_\tau) h_n(\tau)]$ be the minimal expected loss. By the backward induction, the recursive equations are given as follows.

$$
\begin{aligned}
&c_n(n-1) = \sum_{y=1}^{n} q(y)\, h_n(n)/n \\
(5.1)\qquad &c_{n-1}(i-1) = \left(\frac{1}{i}\right) \sum_{y=1}^{n} \min\,[Q_n(i, y)\, h_n(i),\ c_n(i)], \\
&i = n-1, \cdots, 1,
\end{aligned}
$$

where $c_n(i)$ is the minimal cost at time i when the i-th candidate is passed and

$$
\begin{aligned}
Q_n(i, y) &= E[q(X_i) \mid Y_i = y] \\
&= \sum_{x=y}^{n} q(x) \binom{x-1}{y-1} \binom{n-x}{i-y} \Big/ \binom{n}{y}
\end{aligned}
$$

the expected loss for stopping on a candiate of relative rank y at time i. Clearly, the minimal risk is $V_n = c_n(0)$, and the optimal stopping is to stop at time i when the first time the relative rank $Y_i = y$ and $Q_n(i, y)\, h_n(i) \leq c_n(i)$. In this situation, the form of the stopping is unusual. We can usually no longer obtain a cut-off point rule. By a cut-off point rule (see [4]), we mean $\tau = \min\{i \mid i \geq \alpha_k,\ Y_i \leq k\}$ for some integer $0 < \alpha_1 \leq \alpha_2 \leq \cdots \leq n$, where α_i are all cut-off points.

As an example, consider the rank problem $q(x) = x$ with five applicants. The algorithm given in [1] gives optimal stopping $(s_1^*, s_2^*, \cdots, s_5^*) = (0, 1, 1, 2, 5)$, i.e. stop for the first time i such that $Y_i \leq s_i$. The cutoff points are $\alpha_1 = 2$, $\alpha_2 = 4$, $\alpha_3 = \alpha_4 = \alpha_5 = 5$. If we let $h_5(i) = 1$ for $i = 1, 2, 3$ and $h_5(i) = 3$ for $i = 4, 5$, it is obtained from (5.1) that $(s_1, s_2, \cdots, s_5) = (0, 1, 3, 2, 5)$. Note that $s_3 > s_4$. Let $I_y = \{i \mid Q_n(i, y)\, h_n(i) \leq c_n(i)\}$. The optimal stopping is to stop for the first time a candidate of relative rank y arrives in I for some y. This class of rules is called an island rule (see [4]). To be more precise, a rule τ is called an island rule if τ stops as soon as a candiate of relative rank y arrives in the island I_y where $I_1 \supset I_2 \supset \cdots \supset \{n\}$. In the previous example, $I_1 = \{2, 3, 4, 5\}$, $I_2 = \{3, 4, 5\}$, $I_3 = \{3, 5\}$, $I_4 = I_5 = \{5\}$.

To consider the limiting case $n \to \infty$, let $q(\cdot)$ be a nondecreasing sequence. Let $h(\cdot)$ be a nondecreasing function. We consider two types of loss function. The first one, we define $h_n(i)=h(i)$ for $i=1, 2,\cdots, n$, and the second we define $h_n(i)=h(i/n)$ where $h(\cdot)$ is defined on [0, 1]. We apply the analogous idea proposed by Lorenzen ([4], [5]) to obtain the following.

LEMMA 5.1. *Let τ^* be the optimal stopping for no-cost problem (take $h(i)=1$ for all i), then, for any fixed N, there always exists an n large enough such that $\tau^* \geq N$ with probability one.*

THEOREM 5.1. *Let $h_n(i)=h(i)$ where $h(i)$ is monotonically increasing, and let $h(\infty)=\lim_{i\to\infty} h(i)$, $q(\infty)=\lim_{i\to\infty} q(i)$. If V^* is the minimal risk for the no-cost problem, then*

$$\lim_{n\to\infty} V_n = \min\left[V^* h(\infty),\ q(\infty)\, h(1)\right].$$

Consider $h_n(i)=h(i/n)$ where $h(\cdot)$ is an increasing function defined on [0, 1]. Let $f_n(i/n)=c_n(i)$. Then (5.1) can be rewritten by

$$f_n\left(\frac{n-1}{n}\right)=\sum_{y=1}^{n} q(y)\, h(1)/n,$$

$$f_n(i/n)=\left(\frac{1}{i+1}\right)\sum_{y=1}^{i} \min\left(Q_n(i+1,\ y)\, h\left(\frac{i+1}{n}\right),\right.$$

$$\left. f_n\left(\frac{i+1}{n}\right)\Big/(i+1)\right)$$

which again, can be expressed by

$$f_n\left(\frac{n-1}{n}\right)=\sum_{y=1}^{n} q(y)\, h(1)/n,$$

$$(5.2)\qquad f_n\left(\frac{i}{n}\right)-f_n\left(\frac{i+1}{n}\right)$$

$$=\frac{1}{i+1}\left(Q_n(i+1,\ y)\, h\left(\frac{i+1}{n}\right)-f_n\left(\frac{i+1}{n}\right)\right)^{+}.$$

Note that $f_n(0)=\inf_\tau E[q(X_\tau)\, h(\tau)]$

Suppose that as $n\to\infty$,

$$
(5.3)\quad \begin{cases} i/n \to \alpha \ (0 \le \alpha \le 1) \\ f_n(i/n) \to f(\alpha) \\ Q_n(i, y) \to R_y(\alpha) \equiv \sum_{x=y}^{\infty} q(x) \binom{x-1}{y-1} \alpha^y (1-\alpha)^{x-y} \\ h_n(i) \to h(\alpha). \end{cases}
$$

Under the limiting case that $h \to \infty$, the difference equations of (5.2) become the following differential equations

$$
(5.4)\quad \begin{aligned} f'(\alpha) &= \sum_{y=1}^{\infty} [f(\alpha) - R_y(\alpha) h(\alpha)]^+ / \alpha \\ f(1) &= q(\infty) h(1). \end{aligned}
$$

Note that if $h(\alpha) = 1$ for $\alpha \in [0, 1]$, the problem belongs to no-cost type and the differential equations of (5.4) become the one obtained by Mucci [6].

Gianini and Samuels [2] defined an infinite secretary problem. Here the ranks of an infinite number of candidate arrive at each time are iid uniform on [0, 1]. We show that the differential equations of (5.4) give the limiting solution for our finite problem. Let $f(\alpha)$ be the minimal cost of continuation of sampling at time α, and let $R_y(\alpha) h(\alpha)$ be the expected cost for stopping on a candidate of relative rank y at time α. By applying Theorem 2.1 and Corollary to Theorem 4.1 of Lorenzen [4], we obtain the following.

THEOREM 5.2. *If $h(\cdot)$ is continuous, then the differential equation of* (5.4) *holds and $V_n \to f(0)$, the minimal risk for the infinite problem.*

REMARK 5.1. The optimal stopping rule will stop sooner with a cost function than that without a cost function.

EXAMPLE 5.1. Consider a loss function given by $X_\tau/((n+1)-\tau)$. By (5.1), we have

$$
\begin{aligned} c_n(n-1) &= (n+1)/2 \\ c_n(i-1) &= \sum_{y=1}^{i} \min [E(X_i/(n+1-i) \mid Y_i = y), c_n(i)]. \end{aligned}
$$

Let

$$
t_i = \{(i+1)(n+1-i)/(n+1)\} c_n(i)
$$

and

$$s_i = \min\{[t_i], i\}.$$

Then

$$c_n(i-1) = \{[(n+1)s_i(s_i+1)]/ [(i+1)(n+1-i)2] + (i-s_i)c_n(i)\}/i.$$

If $n=10$, then $(s_1, s_2, \cdots, s_{10}) = (0, 1, 1, 2, 3, 3, 5, 6, 9, 10)$. Denote this rule by τ_0. Recalling the usual rank problem considered in section 3, we have, denoted by rule τ^*, $(s_1^*, s_2^*, \cdots, s_{10}^*) = (0, 0, 0, 1, 1, 2, 2, 3, 5, 10)$. Clearly, Remark 5.1 holds and $E\tau_0 = 3.11$ while $E\tau^* = 6.29$.

Acknowledgement. The authors are grateful to a referee for his helpful comments.

REFERENCES

1. Y.S. Chow, S. Moriguti, H. Robbins and S.M. Samuels, *Optimal selection based on relative rank* (the "secretary problem"), Israel J. Math., 2 (1964), 81–90.
2. J. Gianini and S.M. Samuels, *The infinite secretary problem*, Ann. Prob., 4 (1976), 418–432.
3. D.V. Lindley, *Dynamic programming and decision theory*, Appl. Statist., **10** (1961), 39–52.
4. T.J. Lorenzen, *Generalising the secretary problem*, Adv. Appl. Prob., **11** (1979), 384–396.
5. ———, *Optimal stopping with sampling cost: the secretary problem*, Ann. Prob., 9 (1981), 167–172.
6. A.G. Mucci, *Differential equations and optimal choice problems*, Ann. Statist., **1** (1973), 104–113.

Institute of Statistical Science
Academia Sinica
Taipei, Taiwan 11529
Republic of China