DEMONSTRATIO MATHEMATICA
Vol. XI No 3 1978

Beniamin Gołdys

THE SECRETARY PROBLEM – THE CASE WITH MEMORY FOR ONE STEP

1. Introduction

Let $K = \{1,2,\ldots,n\}$, let W be the set of all permutations of K, F – the family of all subsets of W, P – a probability measure on W such that for all $w \in W$

$$P(\{w\}) = \frac{1}{n!} .$$

In that way a probability space (W,F,P) is given. Let X_i $(i = 1,2,\ldots,n)$ be a random variable on W defined as follows: $X_i(w)$ is equal to the i-th element of the permutation w.

Let Y_i $(i = 1,2,\ldots,n)$ be the second family of random variables on W such that $Y_i = 1 +$ the number of $X_1, X_2, \ldots, X_{i-1}$ which are less than X_i.

Let $F_i = \sigma(X_1, X_2,\ldots,X_i)$ be the σ-field generated by $X_1,\ldots,X_i$.

In the general case for a given probability space (W,F,P) and for a given increasing family of σ-fields $\{F_n\}_{n=1}^{\infty}$ the stopping rule t is defined as a random variable $t : W \to \{1,2,\ldots,\infty\}$ which fulfills the following conditions

$$P(t<\infty) = 1 ,$$

$$\{t = n\} \in F_n .$$

Let C be the set of all stopping rules on $(W, F, \{F_i\}_{i=1}^{n}, P)$. Let

$$V(n) = \inf_{t \in C} E(\min(X_t, X_{t-1})).$$

In our case t takes values in $\{1, 2, \ldots, n\}$. It is shown in [1] that in the finite case the optimal stopping rule i.e. the rule which realizes the infimum given above always exists. It was shown by Govindarajulu in [2] in a slightly different form that the optimal rule t is equal to i if

$$Y_i \leqslant \min\left(\frac{i+1}{n+1} c_i, Y_{i-1}\right)$$

or

$$Y_{i-1} \leqslant \min\left(\frac{i+1}{n+1} c_i, Y_i - 1\right),$$

where coefficients c_i are given by the following relations

(1) $$c_{n-1} = E\left[\min(Y_n, E(X_{n-1} | Y_{n-1}, Y_n))\right]$$

(2) $$c_{i-1} = E\left[\min(E(X_i | Y_i), E(X_{i-1} | Y_{i-1}, Y_i), c_i)\right],$$

$$i = n-1, \ldots, 2.$$

It is easy to show that it is the well-known backward induction method and $V(n) = c_1(n)$.

The aim of this paper is to prove the following theorem.

T h e o r e m . If n tends to infinity then the limit of $c_1(n)$ exists and

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

In this proof the method used by CMRS in [3] is extended.

<u>Remark:</u> This problem is well-known as the secretary problem.

Let us take n candidates that apply for a certain vacant position. They appear one by one in a random order. Then X_i is an absolute rank of the i-th candidate and Y_i is a respective rank. We can observe only respective ranks. The integer 1 corresponds to the best candidate and n to the worst one. When the i-th candidate appears we have three possibilities:

a) to take the i-th candidate and stop the process,

b) to take the immediately preceding candidate and stop the process,

c) to reject both i-th and (i-1)-th candidates and wait for the next one. Then the (i-1)-th candidate cannot be recalled.

We must select one of the n candidates. We are looking for such a stopping rule t that minimizes the expectation of the rank of the selected candidate.

2. Basic inequalities

Govindarajulu showed that

$$c_{n-1} = \frac{n+1}{3} \tag{1}$$

and if we put

$$s_i = \left[\frac{i+1}{n+1} c_i\right], \tag{2}$$

then

$$c_{i-1} = \frac{(i-s_i-1)(i-s_i)}{i(i-1)} c_i - \frac{(n+1)s_i(1+s_i)(3i-2s_i-1)}{3i(i-1)(i+1)} . \tag{3}$$

From (1.1) and (1.2) we have

$$c_1 \leqslant c_2 \leqslant \dots \leqslant c_{n-1} = \frac{n-1}{3} \tag{4}$$

and

(5) $$s_1 \leqslant s_2 \leqslant \dots \leqslant s_{n-1} = \left[\frac{n}{3}\right].$$

Put, for $i = 1,\dots,n-1$,

(6) $$t_i = \frac{i+1}{n+1} c_i .$$

Then

(7) $$t_1 < t_2 < \dots < t_{n-1} = \frac{n}{3}.$$

This yields the equation (from (3) and (6))

(8) $$t_{i-1} = \frac{(i-s_i)(i-s_i-1)}{i-1} t_i - \frac{s_i(1+s_i)(3i-2s_i-1)}{3(i^2-1)} .$$

L e m m a 1. For $i = 1,\dots,n-1$ we have

(9) $$t_i \leqslant \frac{2n}{n-i+3}$$

P r o o f : It is shown in [3] that

$$t_i' \leqslant \frac{2n}{n-i+3} ,$$

where

$$t_{n-1}' = \frac{n}{2} \qquad t_i' = \frac{i+1}{n+1} c_i'$$

and

$$c_{n-1}' = \frac{n+1}{2} \qquad c_{i-1}' = E\min\left[E(X_i|Y_i),\ c_i'\right]$$

but $c_{i-1} \leqslant c_{i-1}'$ for $i = 1,\dots,n-1$ and hence (9) is obvious.

Corollary 1.

(10) $$c_1 < 8 .$$

Proof. Let $i = \left[\frac{n}{2}\right]$. Then we have

$$c_1 \leqslant c_i = \frac{n+1}{i+1} t_i \leqslant \frac{n+1}{i+1} \cdot \frac{2n}{n-i+3} < 8 .$$

Lemma 2. For $i = 2,\dots,n-1$ we have

(11) $$t_{i-1} \geqslant \frac{i t_i}{i+1}\left(1 - \frac{t_i}{i-1}\right) .$$

It is easy to show this by putting in (8)

$$t_i = [s_i] + a \qquad \text{for} \quad a \in \langle 0, 1) .$$

Lemma 3. For $n \geqslant 5$ $4 \leqslant i \leqslant n-1$ we have

(12) $$t_i \geqslant \frac{i+1}{4(n-i+2)} .$$

Proof. Let $T(x) = x\left(1 - \frac{x}{i-1}\right)$. T is increasing for $x \leqslant \frac{1}{2}(i-1)$, because $t_{n-1} = \frac{n}{3}$ the lemma is true for $i = n-1$. Suppose that it is true for i. Then we have

$$t_i \leqslant \frac{i+1}{3} \leqslant \frac{i-1}{2}$$

and from (11) we infer that

$$t_{i-1} \geqslant \frac{iT(t_i)}{i+1} \geqslant \frac{i}{4(n-i+3)} .$$

Let us define for any positive integer k and each $n \geqslant 2k$,

i_k = the smallest integer $j \geqslant 1$ such that $s_j \geqslant k$.

We note that $s_{i-1} = 0$ and hence

(13) $$c_1 = c_2 = \ldots = c_{i-1} .$$

From the lemmas given above the following corollaries follow.

Corollary 2.

(14) $$\liminf_{n\to\infty} \frac{i_1}{n} \geqslant \frac{1}{8} .$$

The proof is easy from (10).

Corollary 3. On every set

$$\left\{a \leqslant \frac{i}{n} \leqslant b; \qquad 0 < a < b < 1\right\} ,$$

we have

(15) $$\lim_{n\to\infty} (t_i - t_{i-1}) = 0 .$$

Proof. From the inequalities (9) and (11) we have

$$0 \leqslant t_i - t_i - 1 \leqslant t_i - \frac{it_i}{i+1}\left(1 - \frac{t_i}{i-1}\right) \leqslant \frac{1}{i}\left(\frac{2n}{n-i} + \frac{i}{i-1}\left(\frac{2n}{n-i}\right)^2\right) <$$

$$< \frac{1}{n}\frac{1}{a}\left(\frac{2}{1-b} + 2\left(\frac{2}{1-b}\right)^2\right) \xrightarrow[n\to\infty]{} 0 .$$

Corollary 4. $(\forall k \in N)(\forall n > 32k)$

(16) $$\frac{i_k}{n} \geqslant 1 - \frac{2}{k} ,$$

(17) $$\frac{i_k}{n} \leqslant 1 - \frac{1}{32k} .$$

P r o o f. From inequality (9) we have

$$s_{i_k} \geqslant k \Longrightarrow t_{i_k} \geqslant k \Rightarrow \frac{2n}{n-i_k} \geqslant k \Rightarrow \frac{i_k}{n} \geqslant 1 - \frac{2}{k}$$

what proves the formula (16).

The inequality (17) is true if $i_k \leqslant \left[\frac{1}{2}n\right]$. If $i_k > \left[\frac{1}{2}n\right]$, then it is true by (12).

C o r o l l a r y 5. For every $k,m \in N$, $k>m$, we have

(18) $$\lim_{n\to\infty} t_{i_k} = \lim t_{i_k-m} = k .$$

P r o o f . Take a,b such that $0<a<\frac{1}{8}$, $1 - \frac{1}{32k}<b<1$. Then by (14) and (17) we have

$$a < \frac{i_k-m}{n} < \frac{i_k}{n} < b$$

for sufficently large n, so by Corollary 3 we have

$$\lim_{n\to\infty}\left(t_{i_k} - t_{i_k-m}\right) = 0 .$$

But $t_{i-m} < k \leqslant t_{i_k-m}$, which ends the proof.

C o r o l l a r y 6 . For $k = 1,2,\ldots$ we have

(19) $$s_{i_k} = k \qquad \text{for sufficently large } n$$

and

(20) $$\lim_{n\to\infty} (i_{k+1} - i_k) = \infty .$$

P r o o f . $k \leqslant s_{i_k} \leqslant t_{i_k}$. Thus by (18) the equality (19) is proved. Moreover, we have

$$\lim_{n\to\infty}\left(t_{i_{k+1}} - t_{i_k}\right) = 1 \quad \text{and} \quad \lim_{n\to\infty}\left(t_{i_k+1} - t_{i_{k+1}-m}\right) = 0$$

so (20) is true.

3. The proof of the theorem

Let k be a fixed positive integer. Let n be so large that

$$s_{i_k} = k, \qquad s_{i_{k+1}} = k+1 .$$

For $i_k \leqslant i < i_{k+1}$ we have by (2.8)

$$t_i = \frac{i^2-1}{(i-k)(i-k-1)}\, t_{i-1} - \frac{k(k+1)(3i-2k-1)}{3(i-k-1)(i-k)} .$$

Let

$$t_i = v_i + \frac{k(k+1)}{2k+1} .$$

Then we get

$$v_i = \frac{i^2-1}{(i-k)(i-k-1)}\, v_{i-1} + \frac{a_k}{(i-k)(i-k-1)} ,$$

where

$$a_k = \frac{k(k+1)\left(3\frac{1+k+k^2}{2k+1} + 2k+1\right)}{3} .$$

Now we have

$$v_i = \frac{(i+1)i\ldots(i_k+2)}{(i-k)(i-k-1)\ldots(i_k-k+1)} \cdot \frac{(i-1)(i-2)\ldots i_k}{(i-k-1)\ldots(i_k-k)}\, v_i +$$

$$+ \frac{a_k}{(i-k)(i-k-1)}$$

but

$$\frac{(i+1)\dots(i+1-k)(i-k)\dots(i_k+2)}{(i_k-k+1)\dots(i_k+1)(i_k+2)\dots(i-k)} = \frac{(i+1)\dots(i+1-k)}{(i_k+1)\dots(i_k+1-k)} =$$

$$= \prod_{j=1}^{k+1}\left(\frac{i+j-k}{i_k+j-k}\right)$$

and in the same way we obtain

$$\frac{(i-1)\dots i_k}{(i-k-1)\dots(i_k-k)} = \prod_{j=1}^{k}\left(\frac{i+j-k-1}{i_k+j-k-1}\right)$$

so we have

$$v_i = v_{i_k}\prod_{j=1}^{k+1}\left(\frac{i+j-k}{i_k+j-k}\right)\prod_{j=1}^{k}\left(\frac{i+j-k-1}{i_k+j-k-1}\right) + \frac{a_k}{(i-k)(i-k-1)} .$$

Thus, turning back to t_i and putting $i=i_{k+1}-1$, we get

$$t_{i_{k+1}-1} = \frac{k(k+1)}{2k+1} +$$

$$+ \left(t_{i_k} - \frac{k(k+1)}{2k+1}\right)\prod_{j=1}^{k+1}\left(\frac{i_{k+1}+j-k-1}{i_k+j-k}\right)\prod_{j=1}^{k}\left(\frac{i_{k+1}+j-k-2}{i_k+j-k-1}\right) +$$

$$+ \frac{a_k}{(i_{k+1}-k-1)(i_{k+1}-k-2)} .$$

From (18), (20) and (14), taking into consideration that i_k is a function of n we obtain

$$k+1 = \frac{k(k+1)}{2k+1} +$$

$$+\left(k - \frac{k(k+1)}{2k+1}\right)\prod_{j=1}^{k+1} \lim_{n\to\infty}\left(\frac{i_{k+1}+j-k-1}{i_k+j-k}\right)\prod_{j=1}^{k} \lim_{n\to\infty}\left(\frac{i_{k+1}+j-k-2}{i_k+j-k-1}\right) =$$

$$= \frac{k^2}{2k+1} \lim_{n\to\infty}\left(\frac{i_{k+1}}{i_k}\right)^{2k+1}$$

and consequently

$$\lim_{n\to\infty}\left(\frac{i_{k+1}}{i_k}\right) = \left(\frac{k+1}{k}\right)^{\frac{2}{2k+1}} .$$

Hence we have

$$\lim_{n\to\infty} \frac{i_1}{i_k} = \lim_{n\to\infty} \frac{i_1/n}{i_k/n} = \prod_{j=1}^{k-1}\left(\frac{j}{j+1}\right)^{\frac{2}{2j+1}} .$$

From (16) we obtain the following inequality

$$\left(1 - \frac{2}{k}\right)\prod_{j=1}^{k-1}\left(\frac{j}{j+1}\right)^{\frac{2}{2j+1}} \leqslant \lim_{n\to\infty} \inf \frac{i_1}{n} \leqslant \lim_{n\to\infty} \sup \frac{i_1}{n} \leqslant \prod_{j=1}^{k-1}\left(\frac{j}{j+1}\right)^{\frac{2}{2j+1}}$$

Tending with k to infinity we infer that

$$\lim_{n\to\infty} \frac{i_1}{n} = \prod_{j=1}^{\infty}\left(\frac{j}{j+1}\right)^{\frac{2}{2j+1}}$$

Now we have

$$1 = \lim_{n\to\infty} t_{i_1-1} = \lim_{n\to\infty} \left(\frac{i_1}{n+1} c_{i_1-1}\right) =$$

$$= \lim_{n\to\infty} \frac{i_1}{n} c_1 = \lim_{n\to\infty} c_1 \prod_{j=1}^{\infty} \left(\frac{j}{j+1}\right)^{\frac{2}{2j+1}}$$

In this way we obtain

$$\lim_{n\to\infty} c_1(n) = \prod_{j=1}^{\infty} \left(\frac{j+1}{j}\right)^{\frac{2}{2j+1}} \approx 2,57.$$

which ends the proof of the theorem.

REFERENCES

[1] Y.S. Chow, H. Robbins, D. Siegmund: Great expectations: The theory of optimal stopping. Boston 1971.

[2] Z. Govindarajulu: The secretary problem: optimal selection with interview cost, University of Kentucky, Technical Report No. 82: January 1975.

[3] Y.S. Chow, H. Robbins, S. Moriguti, S.M. Samuels: Optimal selection based on relative rank, Israel J. of Math. 2 (1964), 80-90.

INSTITUTE OF MATHEMATICS, TECHNICAL UNIVERSITY OF WARSAW
Received October 14, 1977.