1996年度日本オペレーションズ・リサーチ学会
春季研究発表会

An Optimal Selection and Replacement Policy Related to the Duration Problem

01303783 愛知大学 玉置光司 TAMAKI Mitsushi

1.Introduction

We consider here a variation of the sequential observation and selection problem called the duration problem treated by Ferguson, Hardwick, and Tamaki. The basic framework of the no information duration problem is described as follows. n candidates appear one at a time in random order with all n! permutations equally likely. If we could observe them all, we could rank them absolutely with no ties from best to worst. However, each time a candidate appears, we only observe the rank of the candidate relative to those preceding her and decide, based solely on the observed rank, whether to choose a candidate or not. The reward to the decision maker(abbreviated by DM) is the length of time he is in possession of a relatively best candidate, who is called an eligible candidate. Thus DM will only select an eligible candidate, receiving one unit of reward as he does so and an additional one for each new observation as long as the selected candidate remains eligible. The objective of DM is to maximize the expected reward earned. Before considering our problem, we give a lemma for later use.

LEMMA 1

Assume that $r(x)$, $0\le x\le 1$, is a differentiable unimodal function with end condition $r(0)=r(1)\le 0$ and attains its peak at a, where$r(a)>0$. Also let α and β be respectively the smaller and the larger root of the equation $r(x)=0$.

Then, if $\lambda_{x,y} = x/y^2$, $0<x<y<1$, the functional equation

$$f(x) = \max\left\{ r(x), \int_x^1 f(y)\lambda_{x,y}dy \right\}, \quad 0\le x\le 1$$

has a continuous solution of the form

$$f(x) = \left\{ \begin{array}{ll} r(b), & 0\le x\le b \\ r(x), & b\le x\le \beta \\ 0, & \beta\le x\le 1 \end{array} \right\},$$

where $b(\alpha\le b\le a)$ is the unique root x of the equation

$$r(x) = \int_x^{\beta} r(y)\lambda_{x,y}dy \ .$$

2. Optimal selection and replacement problem

The problem we consider here is composed of two sub problems, selection problem and replacement problem. DM earns the reward proportional to time duration spent with an eligible candidate, thus the selection problem arises when DM has no candidate - DM must decide when to select an eligible candidate that appears. As soon as a new eligible candidate appears after selection, the present candidate turns out to be ineligible and no further reward is earned. Thus the replacement problem arises when DM has a candidate - DM must decide whether or not to release the present candidate by paying her consolation money $c(\ge 0)$ and simultaneously get a new eligible candidate. When DM decides not to release, he receives no reward until the final time or the next eligible candidate appears, in which case, DM again faces the replacement problem. The objective is to find a procedure which maximizes the expected net reward.

For simplicity, we consider here the asymptotic case, letting n tend to infinity. Let $\bar{x}(\hat{x})$ denote the state where DM has no candidate(keeps a candidate) and facing an eligible candidate at time x. Also define $f(x)(g(x))$ to be the supremum of the expected net reward from time x onward, starting from state $\bar{x}(\hat{x})$.

Then we have the following equations.

$$f(x) = \max\left\{\int_x^1 (y - x + g(y))\lambda_{x,y}dy + (1 - x)\left[1 - \int_x^1 \lambda_{x,y}dy\right], \int_x^1 f(y)\lambda_{x,y}dy\right\},$$

$$g(x) = \max\left\{-c + \int_x^1 (y - x)\lambda_{x,y}dy + (1 - x)\left[1 - \int_x^1 \lambda_{x,y}dy\right], \int_x^1 g(y)\lambda_{x,y}dy\right\}$$

These are simplified to

$$f(x) = \max\left\{r_1(x), \int_x^1 f(y)\lambda_{x,y}dy\right\}, \quad g(x) = \max\left\{r_2(x), \int_x^1 g(y)\lambda_{x,y}dy\right\},$$

where

$$r_1(x) = -x\log x + \int_x^1 g(y)\lambda_{x,y}dy, \quad r_2(x) = -c - x\log x .$$

These can be solved to yield the following lemmas.

LEMMA 2

Assume that DM is at state $\hat{x}$. Then two cases are distinguished depending on the value of c.

Case 1($c \geq e^{-1}$) : No replacement occurs.

Case 2($0 \leq c < e^{-1}$) : Let $\beta(> e^{-1})$ be the larger root of the equation $r_2(x) = 0$, and define $x^* = e^{-2}/\beta$, then replacement occurs if and only if $x^* \leq x \leq \beta$.

LEMMA 3

Assume that DM is at state $\hat{x}$. Then we have

Case 1($c \geq e^{-1}$) : Selection is made if and only if $e^{-2} \leq x \leq 1$. The expected net reward is $2e^{-2}$.

Case 2($0 \leq c < e^{-1}$) : Let $\lambda = 1 + \log\beta = 1 - c/\beta$, and define $x^{**} = \exp\left(-\left(1 + \sqrt{1 + 4\lambda^3/3}\right)\right)$, then selection is made if and only if $x^{**} \leq x \leq 1$. The expected net reward is $-\left(x^{**}\log x^{**} + x^*\log x^* + c\right)$.

Summarizing the above lemmas, we have the main result.

THEOREM 4

Case 1($c \geq e^{-1}$) : The optimal procedure passes over the candidates that appear before e^{-2} and then select the first eligible candidate. No replacement occurs and the expected net reward is $2e^{-2}$.

Case 2($0 \leq c < e^{-1}$) : The optimal procedure passes over the candidates that appear before x^{**} and then select the first eligible candidate. After selection, replacement occurs when an eligible candidate appears in (x^*, β) (neither too early or too late). The expected net reward is $-\left(x^{**}\log x^{**} + x^*\log x^* + c\right)$.

Remark : When consolation money is zero, i.e., $c = 0$, $\beta = 1$, $x^* = e^{-2} \cong 0.1353$, and $x^{**} = \exp(-(1 + \sqrt{7/3})) \cong 0.0799$. Thus the maximum net reward is $2e^{-2} + (1 + \sqrt{7/3})\exp(-(1 + \sqrt{7/3})) \cong 0.4725$.

Reference

T. S. Ferguson, J. P. Hardwick and M. Tamaki (1992) "Maximizing the duration of owning a relatively best object" Contemporary Mathematics 125, 37-57.