1996年度日本オペレーションズ・リサーチ学会
秋季研究発表会

A Continuous Time Duration Problem With an Unknown Number of Options

Mitsushi TAMAKI — Aichi University
* Muci HU — Nagoya Institute of Technology
Katsuhisa OHNO — Nagoya Institute of Technology

1 Introduction

We consider a sequential observation and selection problem called the duration problem, which is a variation of the classical secretary problem which has been treated extensively by Ferguson, Hardwick and Tamaki[3]. The framework of the basic duration problem is described as follows: n applicants 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 an applicant appears, we only observe the rank of the applicant relative to those preceding him/her and decide, based solely on the observed rank, whether to choose the applicant or not. The reward to us is the duration of a selection, that is, the length of time we are in possession of a relatively best applicant, referred to as a *candidate*. Thus we will only select a *candidate*, receving one unit of reward as we do so and an additional one for each new observation as long as the selected candidate remains the relatively best. Our objective is to find a policy which will maximize the expected duration.

Let T_k be defined as the arrival time of the first candidate after time k if there is one, and as $n+1$ if there is none. Then the reward is $T_k - k$ if a candidate is selected at time k. Instead of maximizing the expected duration of a selection, we choose to maximize the expected proportion of time one is in the possession of a candidate, in order to make the solution more easily comparable to the model considered here. Thus the reward is not $T_k - k$ but $(T_k - k)/n$. Ferguson, Hardwick and Tamaki[3] showed that the optimal strategy of the basic duration problem is to pass over a certain number $k^*(n)$ of applicants and then to accept the first candidate. Where $k^*(n)/n \to e^{-2}$ and the corresponding reward tends to $2e^{-2}$ as $n \to \infty$.

In this paper, we consider a duration problem under the framework used by Bruss[1], referred to as *the continuous time duration problem.* Bruss's framework can be descubed as follows: Let $F(z)$ be a distribution function on the real time interval $[0, T]$ and let $Z_1, Z_2, \ldots$ be independent random variables each having continuous distribution function F . Also let N be a non-negative integer-valued random varible having its probablity mass function g. N is assumed to be independent of all Z_k's. Z_k is thought of as the arrival time of applicant k and N represents the total number of applicants that appear. Associated with each applicant is a different quality. We suppose that, given $N = n$, each arrival order of ranks $\langle 1\rangle, \langle 2\rangle, \ldots, \langle n\rangle$ has the probability $1/n!$.

Where Bruss[1] showed that the so called e^{-1}-law plays a key role in the best choice problem, we here show the similar result also holds in the duration problem, i.e., the e^{-2} -law plays an important role both in the basic duration problem and in the continuous time duration problem.

2 Waiting Time Policy

For the following, it is convenient to introduce a change of time

$$x = F(z), \qquad z \in [0, T] \tag{1}$$

such that, in the x time scale time runs from 0 to 1 and such that each $X_k = F(Z_k)$ is uniform on [0,1]. Since no generality will be lost by this supposition, we will discuss the continuous time duration problem under this supposition. As in Bruss[1], we confine our attentions to the class of the waiting time police defined as follows.

DEFINITION. The $x-$*strategy* on [0,1] is a waiting time policy to act as follows:
1. To observe and rank all incoming applicants up to time x without accepting an applicant.
2. To accept the first *candidate* arriving after time x, i.e. the first to be superior to the best of those which arrived in $[0,x]$ if it exists and to refuse all applicants if not.

The time x will be called waiting time.

THEOREM. For any distribution g with $P(N > 0) > 0$, there exists a waiting time z^* maximizing the expected duration of the x-strategy. Moreover, for all $\varepsilon > 0$, there exists integer m such that $N \geq m$ implies $z^* \in [e_F^{-2} - \varepsilon, e_F^{-2}]$, where $e_F^{-2} = inf\{z \mid F(z) = e^{-2}\}$.

PROOF. Let $T(x)$ be the time at which the first *candidate* appears after time x if there is one and 1 if there is none. The duration $D(x)$ under the $x-$*strategy* is defined as $D(x) = T(T(x)) - T(x)$. If $N = 0$, the duration is assumed to be 1 for convenience. If $N = 1$, then the *candidate* will be accepted if she arrives after time x, thus $x_1 = 0$ and the expected duration equals $\frac{1}{2}$. Suppose now $N = n$ $(n = 2, 3, \ldots)$. Let S denote the absolute rank of the best applicant that arrives in $[0, x]$ if applicant appears in $[0, x]$ and $n + 1$ if no applicant appears in $[0, x]$. The event $S = s$ occurs if and only if the $\langle s\rangle$ appears in $[0, x]$ and $s - 1$ best ones in $[x, 1]$. Thus , from the model assumption

$$P(S=s)=\begin{cases} x(1-x)^{s-1} & \text{if } s=1,2,\ldots,n \\ (1-x)^n & \text{if } s=n+1. \end{cases} \qquad (2)$$

Let $Y(x)=T(x)-x$. Then, when $T(x)<1$,i.e., $S\geq 2$, $Y(x)$ represents the elapsed time measured from time x until the first candidate appears.

Let's also define R as the absolute rank of the *candidate* that appears at time $T(x)$, when $T(x)<1$.

It is easy to see that, given $S(\geq 2)$, random variables R and $Y(x)$ are conditionally independent and so the joint distribution of R and $Y(x)$ is given by

$$\begin{aligned} &P(R=r, Y(x)\in(y,y+dy)\mid S) \\ &= P(R=r\mid S)f(y\mid S)dy \ , \end{aligned}$$

where

$$P(R=r\mid S) = \frac{1}{S-1}, \quad r=1,2,\ldots,S-1 \qquad (3)$$

$$\begin{aligned} &f(y\mid S) \\ &= \left(\frac{S-1}{1-x}\right)\left(\frac{1-x-y}{1-x}\right)^{S-2}, 0\leq y\leq 1-x. \qquad (4) \end{aligned}$$

In the above, (3) is immediate and (4) is obtained from the argument of the order statistics because $Y(x)$,conditional on S, can be interpreted as the smallest value among $S-1$ independent and identically distributed random variables each having uniform density on $[0,1-x]$.

A simple calculation leads to

$$E[Y(x)\mid S] = \int_0^{1-x} yf(y\mid S)dy = \frac{1-x}{S} \ .$$

Analogously, we have

$E[D(x)\mid R, Y(x)] = \frac{1-(x+Y(x))}{R}$.

Therefore,

$$\begin{aligned} E[D(x)\mid S] &= E[E[D(x)\mid R,Y(x)]\mid S] \\ &= \left(\frac{1-x}{S}\right)\sum_{r=1}^{S-1}\left(\frac{1}{r}\right) \ , \qquad (5) \end{aligned}$$

where the third equatity follows from the independence of R and Y(x), conditional on S . Let $p_n(x)$ denote the expected duration, conditional on $N=n$, under the $x-$ *strategy* .

Then, setting

$A_0=0, \quad A_s=\frac{1}{s+1}\sum_{r=1}^{s}\frac{1}{r} \ , \quad s\geq 1$,

we can now calculate $p_n(x)$ from (2) and (5) as follows.

$$\begin{aligned} p_n(x) &= E[D(x)]=E[E[D(x)\mid S]] \\ &= \sum_{s=1}^{n}(A_s-A_{s-1})(1-x)^{s+1} \ . \qquad (6) \end{aligned}$$

Since A_s is decreasing in s , for $s\geq 2$,

$$p_n(x)-p_{n+1}(x)=(A_n-A_{n+1})(1-x)^{n+2}\geq 0 \ .$$

Thus, as $n\to\infty$, $p_n(x)$ converges from the above to

$$\begin{aligned} p(x) &= \sum_{s=1}^{\infty}(A_s-A_{s-1})(1-x)^{s+1} \\ &= \frac{1}{2}x\log^2 x \ , \qquad (7) \end{aligned}$$

where the last equality follows from the well known formula (see, e.g., [2]). The function $p(x)$ is maximized by $x=e^{-2}$.

We have from (6)

$\frac{dp_n(x)}{dx}=-(1-x)+\sum_{s=3}^{n}(s+1)(A_{s-1}-A_s)(1-x)^s$.

Thus, $\frac{dp_n(x)}{dx}=0$ is equivalent to

$$\sum_{s=3}^{n}(s+1)(A_{s-1}-A_s)(1-x)^{s-1} = 1 \ . \qquad (8)$$

Since the left hand side of (8) is decreasing in x and increasing in n, if we define

$$n^* = min\{n\geq 3 : \ \textstyle\sum_{s=3}^{n}(s+1)(A_{s-1}-A_s)>1\},$$

$p_n(x)$ has a unique maximum x_n which can be defined as a unique root x of the equation (8) for $n\geq n^*$ and

$$x_n \nearrow e^{-2} \ . \qquad (9)$$

Now let

$$G_m(x) = \sum_{n\geq m} p_n(x)P(N=n) \ . \qquad (10)$$

It follows from (9) that, if there exists a value x^* which maximizes $G_m(x)$, then necessarily $x^*\in[x_m,e^{-2}]$. However, the convergence in (7) and (9) is uniform and so $G_m(x)$ is continuous, i.e., x^* exists . Using (1) and the continuity of F completes the proof.

References

[1] F. Thomas Bruss,(1984) "A Unified Approach To A Class Of Best Choice Problems With An Unknown Number Of Options," *Ann. Probab.*, vol.12, No.3, pp.882-889

[2] S.Moriguchi, K.Udagawa and S.Hitotsumatsu,1987 "Mathematics Formula II ," *Iwanami Shoten, INC.*, pp.144

[3] T.S.Ferguson, J.P.Hardwick and M.Tamaki,(1992) "Maximizing The Duration Of Owning A Reletively Best Object," *American Mathematical Society,* 0217-4132 pp.1-25