数理解析研究所講究録
第 1864 巻 2013 年 12-19

Optimal Stopping Rule for the Full-Information Duration Problem With Random Horizon

愛知大学・経営学部　玉置　光司 (Mitsushi Tamaki)
Faculty of Business Administration, Aichi University

1 Introduction

In the classical duration problem as a variation of the secretary problem studied by Gilbert and Mosteller (1966) extensively, a *known* number n of rankable objects appear sequentially in random order and we must find a stopping rule that maximizes the expected duration of holding a relatively best object. At each stage, we observe only the relative rank of the current object with respect to its predecessors. We select a relatively best object, receiving a payoff of 1 plus the number of future observations before a new relatively best object appears or until the final stage n is reached. This *no-information* version is one of several duration problems studied by Ferguson et al.(1992), along with the *full-information* analogue: given n independent and identically distributed random variables $X_1, X_2, \ldots, X_n$ with a known continuous (to avoid ties) distribution F, find the stopping rule which maximizes the expected duration of holding a relative maximum. Since the distribution is known and continuous, F is assumed, without loss of generality, to be the uniform on the interval $(0,1)$.

It is known that the optimal rule in the no-information problem lets approximately $n/e^2 \approx 0.1353n$ objects go by and then selects the first relatively best object, if any. The optimal *proportional payoff* (= payoff/n) converges to $2e^{-2} \approx 0.2707$. In the full-information version, there exists a sequence of non-decreasing thresholds $b_m, m = 0, 1, \ldots$ and the optimal rule with n observations stops at the first k such that the kth observation is a relative maximum and $X_k \geq b_{n-k}$. The optimal proportional payoff converges to 0.4352. A bivariate integral expression for this value was given as

$$\int_0^1 e^{-c^*/u} \left[\int_0^u \left(\frac{e^{c^* v/u} - 1}{v} + \frac{e^{c^* v/u}}{1-v} \right) dv - 1 \right] du$$

by Mazalov and Tamaki (2003), which is shown to be equivalent to

$$\left(e^{c^*} - 1\right) I(c^*) + \left(e^{-c^*} - c^* I(c^*)\right) J(c^*)$$

by Samuels (2004, Sec. 13.2) and Gnedin (2004)(see also Mazalov and Tamaki (2006)), if the exponential-integral functions are defined as

$$I(c) = \int_c^\infty \frac{e^{-x}}{x} dx, \quad J(c) = \int_0^c \frac{e^x - 1}{x} dx$$

and $c^* \approx 2.1198$ as a solution c of the equation

$$1 + J(-c) = e^{-c}(1 - J(c)).$$

In this paper, we introduce uncertainty about the number N of the actually available objects into the above full-information problem. N is assumed to be a bounded random variable, independent of the sequence $X_1, X_2, \ldots$, and have a prior distribution $p_k = P\{N = k\}$ such that $\sum_{k=1}^n p_k = 1$ and $p_n > 0$ for a known upper bound n (thus the classical problem corresponds to the case where $p_n = 1$ and $p_k = 0$ for $1 \le k < n$). The objective of maximizing the expected duration remains unchanged. We henceforth refer to this problem as the random horizon duration problem (abbriviated to RHDP). Two models, MODEL 1 and MODEL 2, can be considered for the RHDP according to whether the final stage of the planning horizon is N or n. More specifically, if the chosen object is the last relative maximum prior to N, we hold it until stage N in MODEL 1, whereas until stage n in MODEL 2. For the corresponding no-information problem, see Tamaki (2013). See also Gnedin (2005) for the similarity between the RHDP and the best choice problem with random horizon (see, for the latter, Presman and Sonin (1972), Petruccelli (1983), Samuel-Cahn (1996) and Tamaki (2011)).

In Section 2, the structure of the optimal rule is examined, and a necessary and sufficient condition for it to be of the form

$$\tau = \min\{k : X_k = \max(X_1, X_2, \ldots, X_k) \ge a_k\}$$

for a monotone sequence $a_1 \ge a_2 \ge \cdots \ge a_n$ is given. The stopping rule is said to be *monotone* in this case. We evaluate the optimal proportional payoff. The case of uniform distribution for N is studied in detail both for MODELs 1 and 2.

2 MODEL 1

We simply call a relative maximum *candidate* and denote by (k, x) the *state*, where we have just observed the *kth* object to be a candidate having value x, $1 \le k \le n$. Let $S_k(x)$ and $C_k(x)$ represent the expected payoff earned by stopping with the current candidate in state (k, x) and by continuing observations in an optimal manner respectively. Then $V_k(x) = \max\{S_k(x), C_k(x)\}$ represents the optimal expected payoff, provided that we start from state (k, x). Define $\pi_k = \sum_{j=k}^n p_j, 1 \le k \le n$. Then we have

$$S_k(x) = \frac{\sum_{i=k}^n \pi_i x^i}{\pi_k x^k}$$

and

$$C_k(x) = \sum_{i=k+1}^{n} \left(\frac{\pi_i}{\pi_k}\right) x^{i-k-1} \int_x^1 V_i(y)dy.$$

Since, for a given k, $S_k(x)$ is increasing in x, while $C_k(x)$ decreasing, it is optimal to stop in (k, x) for $x \geq a_k$, where

$$a_k = \min\{x : S_k(x) \geq C_k(x)\}.$$

Lemma 2.1. *A necessary and sufficient condition for the sequence $\{a_k\}$ to be monotone is*

$$1 \leq \sum_{j=1}^{n-k} \frac{\pi_{j+k}}{\pi_k} \left(\frac{1 - a_{k+1}^j}{j}\right)$$

for each k, where a_k is a unique root x of the equation

$$\sum_{i=k}^{n} \pi_i x^i = \sum_{i=k}^{n-1} \pi_i x^i \sum_{j=1}^{n-i} \frac{\pi_{j+i}}{\pi_i} \left(\frac{1 - x^j}{j}\right).$$

For the purposes of most applications, the following corollary is useful.

Corollary 2.1. *A sufficient condition for the optimal rule to be monotone is that*

$$\frac{\pi_{j+k}}{\pi_k} \quad \text{is non} - \text{increasing in } k$$

for each possible value of j.

Corollary 2.1 is applicable to the following distributions.

Example 1 (N degenerates to n): $p_n = 1$ and $p_k = 0, 1 \leq k < n$.

Example 2 (uniform): $p_k = 1/n, 1 \leq k \leq n$.

Example 3 (generalized uniform):

$$p_k = \begin{cases} 0, & \text{if } 1 \leq k < T \\ \frac{1}{n-T+1}, & \text{if } T \leq k \leq n, \end{cases}$$

for a given parameter $T = 1, 2, \ldots, n$.

Example 4 (curtailed geometric) : $p_k = (1-q)q^{k-1}/(1-q^n), 1 \leq k \leq n$ for a given parameter $0 < q < 1$.

The explicit expression for the proportinal payoff is given as follows:

Lemma 2.2. *Let* $h_k = \sum_{j=1}^{k} 1/j$ *for* $k \geq 1$ *with* $h_0 = 0$. *Then the expected proportional payoff, when the optimal rule is monotone, can be calculated as*

$$v_n^* = \frac{1}{n} \sum_{k=1}^{n} v_k p_k$$

where

$$v_k = h_k + \sum_{j=1}^{k} \sum_{i=j}^{k} \frac{1}{i} \left(h_{k-i} - h_{i-j} - 1 \right) a_j^i.$$

When N is uniform on $\{1, 2, \ldots, n\}$, the main results can be summarized as follows (see Mazalov and Tamaki (2006)).

Theorem 2.1 *(a) Optimal stopping rule: The thresholds value* a_{n-m} *is given as a unique root* $x \in (0,1)$ *to the equation*

$$\sum_{j=0}^{m} \sum_{k=0}^{j} x^k = \sum_{k=0}^{m-1} x^k \sum_{i=1}^{m-k} \sum_{j=1}^{i} \left(1 - x^j\right) / j.$$

(b) Optimal proportional payoff:

$$v_n^* = \frac{1}{n^2} \sum_{k=1}^{n} v_k.$$

(c) Asymptotics: Let $c^* (\approx 3.6925)$ *be the unique root* c *to the equation*

$$2(e^{-c} + c - 1) = e^{-c} J(c) - (c-1) J(-c). \tag{1}$$

Then v_n^* *converges, as* $n \to \infty$, *to*

$$\begin{aligned} v^* &= \left(e^{c^*} + \frac{1}{c^* - 1} \right) I(c^*) + \frac{1}{2} J(c^*) \left(e^{-c^*} - \frac{(c^*)^2}{c^* - 1} I(c^*) \right) \\ &\approx 0.2022\ldots \end{aligned} \tag{2}$$

Proof. A brief sketch of (c) by PPP (see Samuels (2004)): Let

T=the arrival time of the first (leftmost) atom that lies below the threshold curve $y = c/(1-t)$.

S=the time when the value of the best (lowest) atom above threshold is now equal to the threshold.

V=a uniform random variable on $(0,1)$.

Then,

$$\begin{aligned} f_T(t) &= c(1-t)^{c-1}, \quad 0<t<1 \\ f_S(s) &= \frac{cs}{(1-s)^{c+2}} e^{-\frac{cs}{1-s}}, \quad 0<s<1. \end{aligned}$$

Let (t,y) be the state on PPP and denote by $p(t,y)$ and $q(t,y)$ the expected payoff earned by stopping in state (t,y) and by continuing and stopping with the next candidate respectively. Then, by letting $D(t,y)$ represent the stopping payoff in state (t,y), we have

$$\begin{aligned} p(t,y) &= \int_0^{1-t} P\{D(t,y) > x\} dx \\ &= \int_0^{1-t} \frac{1-t-x}{1-t} e^{-yx} dx \\ &= \frac{c-1+e^{-c}}{cy} \end{aligned}$$

$$\begin{aligned} q(t,y) &= \int_t^1 \left\{ \int_0^y p(s,z) \frac{1}{y} dz \right\} f_S(s) P\{V > s \mid V > t\} dt ds \\ &= \frac{1}{cy} \left((1-c-e^{-c}) + e^{-c} J(c) + (1-c) J(-c) \right) \end{aligned}$$

$p(t,y) = q(t,y)$ yields (1). Moreover,

$$\begin{aligned} v^* &= \int_0^1 \int_0^t (1-s) p\left(s, \frac{c^*}{1-s}\right) f_S(s) f_T(t) ds dt \\ &+ \int_0^1 \int_0^s \left[\int_0^{c^*/(1-t)} (1-t) p(t,y) \frac{1-t}{c^*} dy \right] f_T(t) f_S(s) dt ds, \end{aligned}$$

which yields (2) by straightforward calculations.

3 MODEL 2

We have

$$S_k(x) = \frac{\sum_{i=k}^n \sigma_i x^i}{\pi_k x^k},$$

where $\sigma_i = \pi_i + (n-i)p_i$. The analogous results to Lemma 2.1 and Corollary 2.1 can be given to MODEL 2 as well by simply replacing π_k by σ_k.

Lemma 3.1. *A necessary and sufficient condition for the sequence $\{a_k\}$ to be monotone is*

$$1 \le \sum_{j=1}^{n-k} \frac{\sigma_{j+k}}{\sigma_k} \left(\frac{1 - a_{k+1}^j}{j} \right)$$

for each k, where a_k is a unique root x of the equation

$$\sum_{i=k}^{n} \sigma_i x^i = \sum_{i=k}^{n-1} \sigma_i x^i \sum_{j=1}^{n-i} \frac{\sigma_{j+i}}{\sigma_i} \left(\frac{1 - x^j}{j} \right).$$

Corollary 3.1. *A sufficient condition for the optimal rule to be monotone is that*

$$\frac{\sigma_{j+k}}{\sigma_k} \quad \text{is non} - \text{increasing in } k$$

for each possible value of j.

Examples 1, 2 and 4 satisfy the sufficient condition in Corollary 3.1. When N is uniform on $\{1, 2, \ldots, n\}$, the asymptotic proportional payoff is $2v^* \approx 0.4044$. This is intuitively clear because $\sigma_i = 2\pi_i - 1/n$, and so, as $n \to \infty$, $S_k^{(2)}(x)/S_k^{(1)}(x) = \sum_{i=k}^{n} \sigma_i x^i / \sum_{i=k}^{n} \pi_i x^i \to 2$ where $S_k^{(i)}(x)$ is just the $S_k(x)$ for MODEL i.

4 Remark

Consider now a class of stopping rules having an identical threshold value $b (0 < b < 1)$, i.e., a rule which chooses the first observation whose value exceeds b. How about the asymptotic performance of such a simple rule ? Let K be the number of observations that exceed b. Then K is a binomial random variable with parameters n and $1-b$ and the expected proportional payoff is given by

$$f_n(b) = \frac{1}{n} \sum_{k=1}^{n} \left(\frac{n}{k+1} h_k \right) \binom{n}{k} (1-b)^k b^{n-k}.$$

Let n tend to infinity and b to 1 in such a manner that $n(1-b)$ tends to some constant λ. Then we can approximate the binomial random variable K by the Poisson random variable with parameter λ, implying that

$$\begin{aligned} f_n(b) \longrightarrow f(\lambda) &= \sum_{k=1}^{\infty} \left(\frac{1}{k+1} h_k \right) e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \frac{-J(-\lambda) - e^{-\lambda} J(\lambda)}{\lambda}, \end{aligned}$$

where the last equality follows from Gnedin (2006). $\frac{df(\lambda)}{d\lambda} = 0$ is equivalent to

$$(1 + \lambda)e^{-\lambda}J(\lambda) + J(-\lambda) = 0,$$

which has a unique solution $\lambda^* \approx 2.83970$. Thus $f(\lambda)$ is maximized at $\lambda = \lambda^*$ yielding

$$f(\lambda^*) = e^{-\lambda^*}J(\lambda^*) \approx 0.42632.$$

Surprizingly large ! Compare $0.42632/0.43517 \approx 0.98$ with $0.51735/0.58016 \approx 0.89$ for the best-choice problem.

References

[1] Ferguson, T. S., Hardwick, J. P. and Tamaki, M. (1992). Maximizing the duration of owning a relatively best object. In *Strategies for Sequential Search and Selection in Real Time*(Contemp. Math. 125), American Mathematical Society, Providence, RI, pp.37-57.

[2] Gilbert, J.P. and Mosteller, F. (1966). Recognizing the maximum of a sequence. *J. Amer. Stat. Assoc.* **61**,35-73.

[3] Gnedin, A.V. (2004). Best choice from the planar Poisson process. *Stoch. Process. Appl.* **111**, 317-354.

[4] Gnedin, A.V. (2005). Objectives in the best-choice problems. *Sequential Analysis* **24**, 177-188.

[5] Mazalov, V. V. and Tamaki, M. (2003). Explicit solutions to th duration problem. *Aichi Keiei Ronsyu* **147**, 69-92.

[6] Mazalov, V. V. and Tamaki, M. (2006). An explicit formula for the optimal gain in the full-information problem of owning a relatively best object. *J. Appl. Prob.* **43**, 87-101.

[7] Petruccelli, J.D. (1983). On the best-choice problem when the number of observations is random, *J. Appl. Prob.* **20**, 165-171.

[8] Presman, E. L. and Sonin, I. M. (1972). The best choice problem for a random number of objects, *Theor. Prob. Appl.* **17**, 657-668.

[9] Samuel-Cahn, E. (1996). Optimal stopping with random horizon with application to the full-information best-choice problem with random feeze. *J. Amer. Stat. Assoc.* **91**,357-364.

[10] Samuels, S. M. (2004). Why do these quite different best-choice problems have the same solutions ? *Adv. Appl. Prob.* **36**,398-416.

[11] Tamaki, M. (2011). Maximizing the probability of stopping on any of the last m successes in independent Bernoulli trials with random horizon. *Adv. Appl. Prob.* **43**, 760-781.

[12] Tamaki, M. (2013). Optimal stopping rule for the no-information duration problem with random horizon. to appear in *Adv. Appl. Prob.* **45**.