数理解析研究所講究録
第 1912 巻 2014 年 202-210

On the Optimal Stopping Problems With Monotone Thresholds

Mitsushi Tamaki
Faculty of Business Administration, Aichi University

1 Introduction

We first review the (full-information) *best-choice problem* originally studied by Gilbert and Mosteller (1966, Sec.3) as a variation of the secretary problem. A known number, n, of objects appear one at a time. Let $X_k, 1 \leq k \leq n$, denote the value of the kth object and suppose that $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables with a known continuous distribution F. We can assume without loss of generality that $X_1, X_2, \ldots, X_n$ are uniformly distributed on the interval (0, 1). As each object appears, we observe its value and decide either to select or reject it based on the values observed so far. Once an object is chosen, the process terminates. The objective of the problem is to find a stopping rule which maximizes the probability of choosing the best, i.e. stopping with the largest of $X_1, X_2, \ldots, X_n$ and compute the probability of choosing the best under an optimal stopping rule.

Let $L_k = \max(X_1, \ldots, X_k), 1 \leq k \leq n$, and call the kth object (or X_k) *candidate* if it is relatively best, i.e. $X_k = L_k$. Obviously an optimal stopping rule only stops with a candidate except for the last stage. Consider now a class of stopping rules of the form

$$\tau_n = \tau_n(\mathbf{a}) = \min\{k : X_k \geq a_k, X_k = L_k\} \wedge n, \tag{1}$$

where $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ is a given sequence of thresholds satisfying the monotone condition $1 \geq a_1 \geq a_2 \geq \cdots \geq a_n \geq 0$. This rule is simply referred to as a *monotone rule* (with thresholds $\mathbf{a}$). Gilbert and Mosteller (1966) showed that, in their Theorem 4, the probability of choosing the best object under a monotone rule $\tau_n(\mathbf{a})$ is calculated as

$$v_n(\mathbf{a}) = \frac{1-a_1^n}{n} + \sum_{j=1}^{n-1}\left[\sum_{k=1}^{j}\frac{a_k^j}{j(n-j)} - \sum_{k=1}^{j}\frac{a_k^n}{n(n-j)} - \frac{a_{j+1}^n}{n}\right]$$

and that the optimal stopping rule is within the class of monotone rules and, in effect, the monotone rule with particular thresholds $\mathbf{a}^* = (a_1^*, a_2^*, \ldots, a_n^*)$

gives an optimal stopping rule, where $a_n^* = 0$ and $a_k^*,\ k < n$, is a unique root $x \in (0,1)$ of the equation $\sum_{j=1}^{n-k} \left(x^{-j} - 1\right)/j = 1$. An explicit expression for the limiting optimal probability was given by Samuels (1982). If we introduce the exponential-integral functions

$$I(c) = \int_1^\infty \frac{e^{-cx}}{x} dx, \quad J(c) = \int_0^1 \frac{e^{cx} - 1}{x} dx.$$

and define $c^*(\approx 0.80435)$ as a solution c to the equation $J(c) = 1$, then

$$v^* = \lim_{n\to\infty} v_n^* = e^{-c^*} + (e^{c^*} - c^* - 1)I(c^*) \approx 0.580164.$$

Besides the best-choice problem, there are many optimal stopping problems having the property that the selection is resricted to a candidate except for the last stage and the optimal rule falls under the class of monotone rules. The problem along this line is henceforth referred to as a *candidate-choice problem* (CCP). See Section 2 for more detail of the CCP. Denote by (k, x) the state of the process of a given CCP, where we have just observed the kth object to be a candidate having value x, i.e. $X_k = L_k = x, 1 \le k \le n,\ 0 < x < 1$. The problem is specified by $p_k(x)$ defined as the payoff earned by stopping in state (k, x). One of the aims of this note is to give unified formulae for calculating $v_n(\mathbf{a})$ of the CCP.

Ferguson et al.(1992, Section 3.1 and Section 3.2) showed that both the *duration problem* and the *best-choice duration problem* are the CCP when no recall is allowed. The duration problem is concerned with maximizing the expected duration of holding a candidate. That is, if we stop with a candidate, we receive a payoff of 1 plus the number of future observations before a new candidate appears or until the final stage n is reached. On the other hand, the best-choice duration problem is concerned with maximizing the expected duration of holding the best object (i.e. the last candidate). However, instead of maximizing the expected duration, we choose to maximize the expected proportion of time we are in possession of a candidate in order to make the solution easily comparable to the best-choice problem and other related problems. Hence, we have $p_k(x) = \sum_{j=0}^{n-k} x^j/n$ for the duration problem and $p_k(x) = (n - k + 1)x^{n-k}/n$ for the best-choice duration problem respectively. For these two problems, Ferguson et al.(1992) were mainly concerned with finding the optimal stopping rule, implying that the corresponding expected payoff was left unsolved. See Mazalov and Tamaki (2006) for further investigation of the duration problem. For the best-choice duration problem, we will examine further and give some additional results both in the case without recall, where the most recently observed object may be selected and in the case with recall, where any of the previously observed objects may be selected. We are also interested in the limiting optimal payoff v^*. In Section 3, we will obtain $v^* \approx 0.310965$ for the best-choice duration problem without recall. Interestingly this value can be obtained directly from Samuel-Cahn (1996) if we recognize the equivalence between the best-choice duration problem without recall and the Samuel-Cahn's best-choice problem with uniform freeze. The same result can be obtained via the PPP. We show that, when recall is allowed, the limiting optimal payoff increases up to 0.335360.

2 Formulae for calculating the expected payoff

2.1 Formula related to τ_n

We start with deriving the distribution of the stopping time $\tau_n(\mathbf{a})$ defined in (1).

Lemma 2.1. Assume $n \geq 2$ and define, for a given monotone sequence of thresholds $\mathbf{a} = (a_1, a_2, \ldots, a_n)$,

$$A_k(\mathbf{a}) = \frac{1}{k}\sum_{i=1}^{k} a_i^k, \quad 1 \leq k < n$$

with $A_0(\mathbf{a}) \equiv 1$ and $A_n(\mathbf{a}) \equiv 0$ for convention. Then

$$\begin{aligned} &(a) \quad P\{\tau_n(\mathbf{a}) > k\} = A_k(\mathbf{a}), \quad 0 \leq k \leq n \\ &(b) \quad P\{\tau_n(\mathbf{a}) = k\} = A_{k-1}(\mathbf{a}) - A_k(\mathbf{a}), \quad 1 \leq k \leq n \\ &(c) \quad E[\tau_n(\mathbf{a})] = \sum_{k=0}^{n-1} A_k(\mathbf{a}). \end{aligned}$$

We give a unified formula for calculating $v_n(\mathbf{a})$ of the CCP.

Theorem 2.1. Let $p_k(x)$ be the payoff earned by stopping in state (k, x). Then the expected payoff $v_n(\mathbf{a})$ under the monotone rule $\tau_n = \tau_n(\mathbf{a})$ is calculated from, for $n \geq 2$,

$$v_n(\mathbf{a}) = \int_{a_1}^{1} p_1(x)dx + \sum_{k=2}^{n}\sum_{j=1}^{k-1}\int_{a_k}^{1} p_k(x)\frac{[\min(x, a_j)]^{k-1}}{k-1}dx. \tag{2}$$

Suppose that we are in state (k, x) of a CCP with $p_k(x), 1 \leq k \leq n$. If we stop with the current candidate, we receive $p_k(x)$, while if we leave this state and stop with the next candidate, if any, we can expect to receive the payoff

$$q_k(x) = \sum_{j=k+1}^{n} x^{j-k-1}\int_{x}^{1} p_j(y)dy. \tag{3}$$

Now let

$$G = \bigcup_{k=1}^{n} \{(k, x) : p_k(x) \geq q_k(x)\}.$$

Hence, G represents the set of states for which stopping immediately is at least as good as waiting for the next candidate to appear and then stopping. The stopping rule which stops as soon as the state enters the set G is called *one-stage look-ahead rule* (1-sla rule). It is well known (see, e.g. Ferguson (2006) or Chow et.al (1971)) that the stopping rule $\tau_n(\mathbf{a}^*)$, which is also 1-sla, is optimal if there exists a monotone sequence $\mathbf{a}^* = (a_1^*, a_2^*, \dots, a_n^*)$ such that G can be expressed as $G = \cup_{k=1}^n \{(k,x) : x \geq a_k^*\}$. To be more precise, we refer to the stopping problem as a CCP when it has an optimal 1-sla rule.

We give some examples and comments. The best-choice duration problem will be examined in detail in Section 3.

Example 1. (Best-choice problem): $p_k(x) = x^{n-k}$.

$$v_n(\mathbf{a}) = \frac{1}{n}\left[1 + \sum_{k=1}^{n-1}\sum_{j=k}^{n-1}\frac{1}{j}a_k^j + \sum_{k=1}^{n-1}\left\{\sum_{j=k}^{n-1}\frac{1}{n-j}a_k^j - (1+h_{n-k})\,a_k^n\right\} - a_n^n\right],$$

where $h_k = \sum_{j=1}^k 1/j,\ k \geq 1$ with $h_0 = 0$.

Example 2. (Duration problem without recall): $p_k(x) = \sum_{j=0}^{n-k} x^j/n$.

$$v_n(\mathbf{a}) = \frac{1}{n}\left[h_n + \sum_{k=1}^{n}\sum_{j=k}^{n}\frac{1}{j}\left(h_{n-j} - h_{j-k} - 1\right)a_k^j\right].$$

Example 3. (Version of the best-choice problem): The problem considered in Section 4 of Tamaki (2010) is concerned with maximizing the probability of stopping with any of the last m candidates, where m is a predetermined positive integer. Let $r_n(k)$ be defined recursively as follows.

$$r_n(k) = \frac{1}{n}r_{n-1}(k-1) + \left(1 - \frac{1}{n}\right)r_{n-1}(k), \quad 1 \leq k \leq n,\ 2 \leq n$$

with $r_1(1) = 1$ and $r_n(k) = 0$ for $k = 0$ or $k > n$ and define $d_n = \sum_{k=1}^{m-1} r_n(k)$ for $n \geq m$ and $d_n = 1$ for $n < m$. Then

$$p_k(x) = \sum_{j=0}^{n-k} d_j \binom{n-k}{j}(1-x)^j x^{n-k-j}.$$

2.2 Formula related to σ_n

Besides (2), we can give another expression for $v_n(\mathbf{a})$. Let $\sigma_n(\mathbf{a})$ be the time at which the largest value observed so far initially exceeds the threshold, i.e.

$$\sigma_n = \sigma_n(\mathbf{a}) = \min\{k : L_k \geq a_k\} \wedge n.$$

Lemma 2.2. Assume $n \geq 2$. Then, for a given monotone sequence $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ with the interpretation that $a_0 = 1$ and $a_n^n = 0$, we have

$$\begin{aligned} &(a) \quad P\{\sigma_n(\mathbf{a}) > k\} = a_k^k, \quad 0 \leq k \leq n \\ &(b) \quad P\{\sigma_n(\mathbf{a}) = k\} = a_{k-1}^{k-1} - a_k^k, \quad 1 \leq k \leq n \\ &(c) \quad E[\sigma_n(\mathbf{a})] = \sum_{k=0}^{n-1} a_k^k. \end{aligned}$$

Another expression for $v_n(\mathbf{a})$ is given as follows.

Theorem 2.2. Let $q_k(x)$ be as defined in (3). Then we have that

$$v_n(\mathbf{a}) = \sum_{k=1}^{n} \left[\int_{a_k}^{1} p_k(x) \left[\min(x, a_{k-1})\right]^{k-1} dx + (k-1) \int_{a_k}^{a_{k-1}} q_k(x) x^{k-1} dx \right].$$

2.3 Limiting expected values of τ_n/n and σ_n/n

To make explicit the dependence on n, we here write $\mathbf{a}(n) = (a_1(n), a_2(n), \ldots, a_n(n))$ instead of $\mathbf{a} = (a_1, a_2, \ldots, a_n)$. Let $\{b_k\}_{k=0}^{\infty}$ be an infinite sequence such that

$$\begin{aligned} &(i) \quad 0 \leq b_0 \leq b_1 \leq b_2 \leq \cdots \leq 1 \\ &(ii) \quad \lim_{n \to \infty} n(1 - b_n) = c < \infty \end{aligned}$$

and define $a_k(n) = b_{n-k}, 1 \leq k \leq n$, i.e. $\mathbf{a}(n) = (b_{n-1}, b_{n-2}, \ldots, b_0), n \geq 1$. Then we have the following limiting results.

Lemma 2.3. For $\mathbf{a}(n)$ satisfying the above properties (i) and (ii), we have

$$\begin{aligned} &(a) \quad \lim_{n \to \infty} E\left[\frac{\tau_n(\mathbf{a}(n))}{n}\right] = e^{-c} + (e^c - c - 1) I(c) \\ &(b) \quad \lim_{n \to \infty} E\left[\frac{\sigma_n(\mathbf{a}(n))}{n}\right] = 1 - c e^c I(c). \end{aligned}$$

3 Best-choice duration problem

3.1 Sampling without recall

This problem is a CCP and the next lemma yields the expected payoff.

Lemma 3.1. (Expected payoff): $p_k(x) = (n - k + 1) x^{n-k}/n$.

$$(a) \quad v_n(\mathbf{a}) = \frac{1}{n}\left[1 + \sum_{k=1}^{n-1} \sum_{j=k}^{n-1} \frac{a_k^j}{j} - \frac{1}{n} \sum_{k=1}^{n} (2(n-k)+1) a_k^n \right] \tag{4}$$

$$(b) \quad v_n^* = \frac{1}{n}\left[1 + \sum_{k=1}^{n-1} \sum_{j=k}^{n-1} \left(\frac{1}{j} - \frac{1}{n}\right) (a_k^*)^j \right],$$

where $a_n^* = 0$ (*i.e.* $r = n$) and $a_k^*, k < n$, is a unique solution $x \in (0,1)$ of the equation

$$(2(n-k)+1)\,x^n = \sum_{i=k}^{n-1} x^i. \tag{5}$$

Samuel-Cahn (1996) generalized the best-choice problem by introducing a random "freeze-time" variable N which makes it impossible to make a selection after time N. The goal of stopping with the largest of $X_1, X_2, \ldots, X_n$ remains unchanged. In the case of N uniform on $\{1, 2, \ldots, n\}$, Samuel-Cahn showed that the optimal rule is a monotone rule with $\mathbf{a}^* = (a_1^*, a_2^*, \ldots, a_n^*)$, where a_k^* is determined from (5) and the optimal probability is given by $v_n(\mathbf{a}^*)$, where $v_n(\mathbf{a})$ is given by (4). That is, the best-choice duration problem without recall is equivalent to the Samuel-Cahn's best-choice problem with N uniform on $\{1, 2, \ldots, n\}$. From this equivalence, we can give an integral expression for v^*.

Lemma 3.2. (Limiting optimal payoff)

$$\begin{aligned} v^* &= \int_0^1 \frac{1}{x}\left[\int_0^x e^{-\frac{c^* x}{1-y}}dy\right]dx - 2\int_0^1 ye^{-\frac{c^*}{y}}dy \\ &\approx 0.310965, \end{aligned}$$

where $c^* (\approx 1.25643)$ is a unique solution $c (> 0)$ of the equation

$$e^c = 1 + 2c.$$

Another simpler expression for v^* in terms of $I(c)$ is given as follows.

Lemma 3.3. Write c for c^* for convenience. Then

$$v^* = ce^{-c} + c(1-c)I(c). \tag{6}$$

The asymptotic result (6) is also obtained via the PPP model. According to Samuels (2004), we use a Poisson process with unit rate on the semi-infinite strip $[0,1] \times [0, \infty)$. This turns the problem upside down, making the 'best' become the 'smallest'. Suppose that an atom is identified as a point (t, y) if the atom appears at time t as a candidate (relatively best atom as in the finite problem) having value y in the PPP. Let $P(t, y)$ denote the expected payoff if we choose this point, i.e. stop on the point (t, y). Then

$$P(t, y) = (1-t)e^{-y(1-t)}. \tag{7}$$

If we do not choose this point, but choose the point related to the next candidate, if any, then we can expect to receive a payoff

$$\begin{aligned} Q(t, y) &= \int_0^{1-t}\left(\int_0^y P(t+r, x)\frac{1}{y}dx\right) ye^{-yr}dr \\ &= \frac{1}{y}\left(1 - e^{-y(1-t)}\right) - (1-t)e^{-y(1-t)}. \end{aligned}$$

Solving for the locus of point (t, y) at which $P(t, y) = Q(t, y)$ yields $y(1-t) = c^*$, where $c^* \approx 1.25643$. Since $P(t, y) \geq Q(t, y)$ implies $P(t', y') \geq Q(t', y')$ for $t' > t, y' < y$, we are in the monotone case of optimal stopping (see Ferguson (2006) or Chow et al. (1971) for the monotone case) and can conclude that the optimal rule stops with the first candidate, if any, that lies below the threshold curve $y = c^*/(1-t)$. Henceforth, we again write c instead of c^* for simplicity. Let T be the arrival time of the first (leftmost) atom that lies below the threshold curve $y = c/(1-t)$, and let S be the time when the value of the best (lowest) atom above threshold is equal to the threshold. Then it is easy to see that (see, e.g. Section 10.2 of Samuels (2004)) the limiting optimal payoff is calculated from

$$\begin{aligned} v^* &= \int_0^1 \int_0^t P\left(s, \frac{c}{1-s}\right) f_S(s) f_T(t) ds dt \\ &\quad + \int_0^1 \int_0^s \left(\frac{1-t}{c} \int_0^{c/(1-t)} P(t, y) dy \right) f_T(t) f_S(s) dt ds, \end{aligned} \tag{8}$$

where $f_T(t)$ and $f_S(s)$ are the densities of T and S given as

$$f_T(t) = c(1-t)^{c-1}, \quad f_S(s) = \frac{cs}{(1-s)^{c+2}} e^{-\frac{cs}{1-s}}.$$

Applying (7) to (8) yields

$$\begin{aligned} v^* &= e^{-c} \int_0^1 \int_0^t (1-s) f_S(s) f_T(t) ds dt \\ &\quad + \frac{1-e^{-c}}{c} \int_0^1 \int_0^s (1-t) f_T(t) f_S(s) dt ds. \end{aligned} \tag{9}$$

Moreover, the bivariate integrals can be simplified to

$$\int_0^1 \int_0^t (1-s) f_S(s) f_T(t) ds dt = c\left[(1+c)e^c I(c) - 1\right] \tag{10}$$

$$\int_0^1 \int_0^s (1-t) f_T(t) f_S(s) ds dt = c\left[1 - c e^c I(c)\right] \tag{11}$$

respectively. Substituting (10) and (11) into (9) gives

$$v^* = 1 - (1+c)e^{-c} + c(2 + c - e^c) I(c).$$

Remark 3.1. The equivalence between the best-choice duration problem without recall and the Samuel-Cahn's best-choice problem with uniform freeze is not surprizing, because this is very similar to the equivalence between the duration problem without recall and the Porosinski (1987)'s best-choice problem with uniform horizon, to which Samuels (2004) and Gnedin (2004) have given a good explanation. See also Gnedin (2005) for further generalization of the equivalence.

3.2 Sampling with recall

Ferguson et al.(1992) showed that, in the recall case, the optimal stopping rule is within the class of stopping rules $\{\sigma_n(\mathbf{a})\}$ in the sense that it stops at time σ_n with the current candidate if $L_{\sigma_n} = X_{\sigma_n}$, but with the previous objext, say, the jth object if $L_{\sigma_n} = X_j > X_{\sigma_n}$. Moreover they showed that the optimal thresholds are given by $a_k^* = 2^{-1/(n-k)}, 1 \le k \le n$. Let $u_n(\mathbf{a})$ denote the expected payoff under $\sigma_n(\mathbf{a})$ and $u_n^* = u_n(\mathbf{a}^*)$. Then we have the following results.

Lemma 3.4. (Expected payoff): $p_k(x) = (n-k+1)x^{n-k}/n$.

$$(a)\quad u_n(\mathbf{a}) = \frac{1}{n}\left[1 + \sum_{k=1}^{n} a_k^k - 2\sum_{k=1}^{n} \frac{k}{n} a_k^n\right]$$

$$(b)\quad u_n^* = \frac{1}{n}\left[1 + \sum_{k=1}^{n}\left(1 - \frac{k}{n}\right) 2^{-\frac{k}{n-k}}\right].$$

Let $u^* = \lim_{n\to\infty} u_n^*$. Then we have the following limiting result.

Lemma 3.5. Let $\tilde{c} = \log 2$. Then

$$u^* = \frac{1-\tilde{c}}{2} + (\tilde{c})^2 I(\tilde{c}) \approx 0.335360.$$

Remark 3.2. u^* is also obtained from the PPP model, i.e. u^* is just the value of v^* of (9) when $c(= c^* \approx 1.25643)$ is replaced by $\tilde{c} = \log 2 \approx 0.69315$, because, as is easly seen from the argument of the infinitesimal look-ahead stopping rule used for the duration problem with recall in Section 3.2 of Mazalov and Tamaki (2006), the optimal threshold curve in the recall case is given by $y = \tilde{c}/(1-t)$, contrasting the curve $y = c^*/(1-t)$ in the no recall case ($\tilde{c} = \log 2$ was already suggested in Ferguson et al.(1992)).

Lemma 3.6. Let $c^* \approx 1.25643$ and $\tilde{c} = \log 2$. Then

$$(a)\qquad \lim_{n\to\infty} E\left[\frac{\tau_n^*}{n}\right] = e^{-c^*} + (e^{c^*} - c^* - 1)I(c^*) \approx 0.46678$$

$$(b)\qquad \lim_{n\to\infty} E\left[\frac{\tilde{\sigma}_n}{n}\right] = 1 - \tilde{c}e^{\tilde{c}} I(\tilde{c}) \approx 0.47505.$$

References

[1] Chow, Y. S., Robbins, H. and Siegmund, D. (1971). *The Theory of Optimal Stopping*, Houghton Mifflin, Boston.

[2] 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, 37-57.

[3] Ferguson, T. S. (2006). *Optimal Stopping and Applications*, Electronic Text at http://www.math.ucla.edu/~tom/Stopping/Contents.html

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

[5] Gnedin, A. V. (1996). On the full information best choice problem, *J. Appl. Prob.*, **33**, 678-687.

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

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

[8] 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.

[9] Porosinski, Z. (1987), The full-information best-choice problem with a random number of observations, *Stoch. Process. Appl.* **24**, 293-307.

[10] 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.

[11] Samuels, S. M. (1982), Exact solutions for the full information best choice problem, *Purdue Univ. Stat. Dept. Mimeo. Series 82-17.*

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

[13] Tamaki, M. (2009). Optimal choice of the best available applicant in full-information models, *J. Appl. Prob.* **46**, 1086-1099.

[14] Tamaki, M. (2010). Sum the multiplicative odds to one and stop, *J. Appl. Prob.* **47**, 761-777.

Postal address: Department of Business Administration, Aichi University, Nagoya Campus, Hiraike 4-60-6, Nakamura, Nagoya, Aichi 453-8777, Japan.
Email address: tamaki@vega.aichi-u.ac.jp
愛知大学・経営学部　玉置　光司