<table>
<thead>
<tr>
<th>Title</th>
<th>Optimal Stopping Rule for the Full-Information Duration Problem With Random Horizon (Stochastic Decision Analysis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Tamaki, Mitsushi</td>
</tr>
<tr>
<td>Citation</td>
<td>数理解析研究所講究録 (2013), 1864: 12-19</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2013-11</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/195357">http://hdl.handle.net/2433/195357</a></td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Textversion</td>
<td>publisher</td>
</tr>
</tbody>
</table>

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

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 ($= \text{payoff}/n$) 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 $k$th 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 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 \( 1 \leq k < n \)). The objective of maximizing the expected duration remains unchanged. We henceforth refer to this problem as the random horizon duration problem (abbreviated 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

\[ \tau = \min \{k : X_k = \max(X_1, X_2, \ldots, X_k) \geq a_k\} \]

for a monotone sequence \( a_1 \geq a_2 \geq \cdots \geq a_n \) 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 \( k \)th object to be a candidate having value \( x \), \( 1 \leq k \leq n \). 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 \( \pi_k = \sum_{j=k}^{n} p_j \), \( 1 \leq k \leq n \). 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
\[ \frac{\pi_{j+k}}{\pi_k} \]
is non-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): \( p_k = \frac{1}{n}, 1 \leq k \leq n \).

**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 proportional 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

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

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

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

(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

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

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

$$v^* = (e^{c^*} + \frac{1}{c^*-1}) I(c^*) + \frac{1}{2} J(c^*) \left( e^{-c^*} - \frac{(c^*)^2}{c^*-1} I(c^*) \right)$$

$$\approx 0.2022 \ldots$$

**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,

$$f_T(t) = c(1-t)^{c-1}, \quad 0 < t < 1$$

$$f_S(s) = \frac{cs}{(1-s)^{c+2}}e^{-\frac{cs}{1-s}}, \quad 0 < s < 1.$$ 

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

$$p(t, y) = \int_0^{1-t} P\{D(t, y) > x\} dx$$

$$= \int_0^{1-t} \frac{1-t-x}{1-t} e^{-yx} dx$$

$$= \frac{c-1+e^{-c}}{cy}$$

$$q(t, y) = \int_0^1 \left\{ \int_0^y p(s, z) \frac{1}{y} dz \right\} f_S(s) P\{V > s \mid V > t\} dt ds$$

$$= \frac{1}{cy} \left( (1 - c - e^{-c}) + e^{-c} J(c) + (1 - c) J(-c) \right)$$

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

$$v^* = \int_0^1 \int_0^t (1-s) p\left( s, \frac{c^*}{1-s} \right) f_S(s) f_T(t) ds dt$$

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

which yields (2) by straightforward calculations.

### 3 MODEL 2

We have

$$S_k(x) = \frac{\sum_{t=k}^{n} \sigma_t x^t}{\pi_k x^k},$$
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 $\pi_k$ by $\sigma_k$.

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

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

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

$$\frac{\sigma_{j+k}}{\sigma_k}$$

is non–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 $\sigma_{i+k} = \pi_{i+k} + (n-k)p_{i+k}$, the asymptotic proportional payoff is given by $2v^* \approx 0.4044$. 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 $\lambda$. Then we can approximate the binomial random variable $K$ by the Poisson random variable with parameter $\lambda$, implying that

$$f_n(b) \to f(\lambda) = \sum_{k=1}^{\infty} \left( \frac{1}{k+1} h_k \right) e^{-\lambda} \frac{\lambda^k}{k!} = \frac{-J(-\lambda) - e^{-\lambda} J(\lambda)}{\lambda},$$
where the last equality follows from Gneden (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

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

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

References


