Optimal Stopping Related to the Relatively Best Objects

Mitsushi Tamaki

Faculty of Business Administration,

Aichi University

1 Introduction

A fixed known number n of rankable objects (1 being the best and n the worst) appear one at a time with all $n!$ permutations equally likely. An object is called *candidate* if it is relatively best. Let $X_i = 1$ if the *ith* object is a candidate and $X_i = 0$, otherwise. It is well known that $X_1, X_2, \ldots, X_n$ are independent random variables with $P\{X_i = 1\} = 1/i$, $1 \le i \le n$. Denote by C_k the *kth* to last candidate, $k = 1, 2, \ldots$. Hence, C_1 is the last candidate. As each candidate appears, we must decide either to choose it, or reject it and continue observations until the next candidate appears, if any.

In Section 2, we present a formulation of the one-choice problems, where if the chosen candidate is C_k, a non-negative reward α_k is earned. If no candidate is chosen, we earn nothing. The objective of the problem is to seek a stopping rule that maximizes the expected reward of the chosen candidate, given a reward sequence $\{\alpha_k\}_{k \ge 1}$. Our main concerns are to examine the limiting behaviors of the problems as n tends to ∞. The desired setting for this turns out to be a non-homogeneous Poisson process (NHPP).

In Section 3, we consider the two-choice problems, where candidate can be chosen up to two and the reward earned depends on the combination of the chosen candidate(s). That is, the reward is $\alpha_{i,j}$ if C_i and C_j are jointly chosen, and α_i if only C_i is chosen for $1 \le i < j$. Nothing can be earned with no choice. The objective is to maximize the expected reward of the chosen candidate(s), when reward sequences $\{\alpha_{i,j}\}_{1 \le i < j}$ and $\{\alpha_k\}_{1 \le k}$ are given. Two simple problems are examined in detail. One is the problem of choosing exactly two candidates from the last three candidates C_1, C_2 and C_3 and the other is the problem of choosing either one or two candidate(s) from the last two candidates C_1 and C_2. These two problems can be solved explicitly. We also consider the problem with *additive payoff*, i.e. $\alpha_{i,j} = \alpha_i + \alpha_j$, $1 \le i < j$. This study is motivated by Gilbert and Mosteller(1966), Bruss(2000) and Tamaki(2010).

2 One-choice problem

Let $N_j = \sum_{i=j+1}^n X_i$ be the number of candidates that appear after time j exclusive. We start with giving the distribution of N_j, i.e. $P_j(k) = P\{N_j = k\}$.

Lemma 2.1. *We have, for $0 \leq k \leq n - j$,*

$$P_j(k) = \left(\frac{j}{n}\right) \sum_{j+1 \leq t_1 < t_2 < \cdots < t_k \leq n} \prod_{i=1}^k \left(\frac{1}{t_i - 1}\right), \quad 0 < j < n.$$

with the interpretation of $P_0(k) = P_1(k - 1)$.

In this section, we consider the one-choice problem having a reward sequence $\{\alpha_k\}_{k \geq 1}$. Denote by j the state, where we have just observed the *jth* object to be a candidate and no choice has been made so far, $1 \leq j \leq n$. Let $S(j)$ and $C(j)$ represent the expected reward earned by stopping with the current candidate in state j and by continuing observations in an optimal manner, respectively. Then $V(j) = \max\{S(j), C(j)\}$ represents the optimal expected reward, provided that we start from state j. We obviously have

$$S(j) = \sum_{k=1}^{n-j+1} \alpha_k P_j(k - 1), \quad C(j) = \sum_{k=j+1}^n \frac{j}{k(k-1)} V(k),$$

with the boundary condition $C(n) = 0$. Let

$$\tilde{C}(j) = \sum_{k=j+1}^n \frac{j}{k(k-1)} S(k), \quad 1 \leq j < n.$$

Then $\tilde{C}(j)$ represents the expected payoff by stopping with the first candidate, if any, after leaving state j. In other words, $\tilde{C}(a - 1)$ is the expected payoff under N_a, where N_a is a threshold rule with threshold a.

The following lemma is just a restatement of Lemma 2.1 of Ferguson et al.(1992).

Lemma 2.2. *If $S(j)$ is unimodal in the sense that for some integer M, $S(j) \leq S(j+1)$ for $j < M$ and $S(j) \geq S(j+1)$ for $j \geq M$, then there exists an optimal stopping rule among the threshold rules N_a for $a \leq M$.*

The main concerns in this paper are to examine the limiting behavior of the optimal rule and the corresponding payoff as n tends to ∞. Divide the time interval $(0, 1]$ into n equal spaces and let n objects appear at times $1/n, 2/n, \ldots, n/n$. Also let n tend to ∞ and denote by $N(a, b)$ the number of candidates that appear on time interval $(a, b), 0 < a < b \leq 1$. Then since $\{N(a, b), 0 < a < b \leq 1\}$ has independent increments with intensity function

$\lambda(x) = 1/x$, $0 < x \le 1$ at time x, this counting process becomes a NHPP with mean value function

$$m(a,b) = \int_a^b \lambda(x)dx = \log(b/a).$$

Hence,

$$P\{N(a,b) = k\} = e^{-m(a,b)}\frac{\{m(a,b)\}^k}{k!} = \left(\frac{a}{b}\right)\frac{\{\log(b/a)\}^k}{k!}.$$

Note that, as $n \to \infty$ with $j/n = t$, $P_j(k) \to P\{N(t,1) = k\}$. Let $T(t)$ be the arrival time of the first candidate after time t. We have the following result.

Corollary 2.1. *The density of $T(t)$ is given by*

$$f_{T(t)}(y) = \frac{t}{y^2}, \quad t < y \le 1. \tag{1}$$

The limiting problem considered on the NHPP is referred to as a *continuous problem* (with reward sequence $\{\alpha_k\}_{k \ge 1}$). For the continuous problem, we denote by t the state, where we have just observed a candidate at time t and no choice has been made so far, $0 < t \le 1$. Let $s(t)$ and $c(t)$ represent the expected reward earned by stopping with the current candidate in state t and by continuing in an optimal manner after leaving state t, respectively. Then $v(t) = \max\{s(t), c(t)\}$ represents the optimal expected reward, provided that we start from state t. Since $N(t,1)$ is a Poisson random variable with parameter $-\log t$, we have

$$
\begin{aligned}
s(t) &= E[\alpha_{N(t,1)+1}] = t\sum_{k=1}^{\infty} \alpha_k \frac{(-\log t)^{k-1}}{(k-1)!} \\
c(t) &= E[v(T(t))] = \int_t^1 \frac{t}{y^2}v(y)dy.
\end{aligned}
$$

Thus we have the following result as a continuous version of Lemma 2.2.

Theorem 2.1. *Suppose that $s(t)$ is unimodal.*
(i) The optimal rule: There exists a threshold $a^* = e^{-\sigma^*}$ such that the optimal rule stops in state t if $t \ge a^*$, where σ^* is the unique root σ of the equation

$$\sum_{k=0}^{\infty} \frac{\alpha_k - \alpha_{k+1}}{k!}\sigma^k = 0.$$

(ii) The optimal payoff: Let v^* denote the optimal payoff. Then

$$v^* = a^*\sum_{k=1}^{\infty} \frac{\alpha_k}{k!}(\sigma^*)^k.$$

3   Two-choice problem

Suppose that we can choose at most two candidates and the payoff earned depends on the combination of the candidates chosen. If we choose both C_i and C_j, we earn the payoff $\alpha_{i,j}$ for $i < j$. If we only choose C_k, we earn α_k. If we fail to choose a candidate, we earn nothing.

Imagine a situation where we observe a candidate at time t in the continuous problem. Then this situation is described as state (i,t) if we have already selected the *ith* last to the current candidate, $1 \leq i$ and $0 < t \leq 1$. Thus, if the current candidate is C_k, the previously selected one is C_{k+i}. For convenience, we denote by $(0,t)$ the state, if no selection has been made previously. Let $s_i(t)$ and $c_i(t)$ respectively represent the stopping reward and the optimal continuation reward, provided that the process starts from state (i,t). Then the optimality equations are given in a similar manner as in the one-choice problem if we let $v_i(t) = \max\{s_i(t), c_i(t)\}$ for $0 \leq i$ (omitted to save the space). The two problems of interest are treated in Sections 3.1 and 3.2. In Section 3.3, we consider a problem with additiva payoff.

3.1   Choosing exactly two candidates from C_1, C_2 and C_3

We are required to choose exactly two candidates from the last three candidates C_1, C_2 and C_3, so the payoff is specified by

$$\alpha_{1,2} = \alpha, \quad \alpha_{1,3} = \beta, \quad \alpha_{2,3} = \gamma,$$

depending on which pair is chosen, where $\alpha \geq \beta \geq \gamma \geq 0$ with $\alpha > 0$. All other payoffs are zero. Since the optimal rule evidently stops in state $(2,t)$, it suffices to consider the optimal decision in states (i,t) for $i = 0, 1$.

Theorem 3.1.*(Optimal decision in state $(1,t)$)*
There exists a threshold $a_1 = \exp(-\sigma_1)$ such that the optimal rule stops in state $(1,t)$ if $t \geq a_1$, where

$$\sigma_1 = \frac{\alpha}{\beta - \gamma}.$$

Theorem 3.2.*(Optimal decision in state $(0,t)$)*
There exists a threshold $a_0 = \exp(-\sigma_0)$ such that the optimal rule stops in state $(0,t)$ if $t \geq a_0$.

Theorems 3.1 and 3.2 state that the optimal rule of the problem is, as a whole, specified by two thresholds (a_0, a_1) in the sense that it chooses, in state (i,t), the current candidate if $t \geq a_i$ for each i. Hence, what is left is to determine a_0 and then examine the boundary condition between two possible cases; $a_1 \leq a_0$ and $a_0 < a_1$, because the feature of the optimal rule quite differs between these two cases. When $a_1 \leq a_0$, the optimal rule is simply described as choosing two candidates that appear after time a_0

successively. Obviously this rule concentrates on choosing either pair of C_1C_2 or C_2C_3, excluding the possibility of choosing C_1C_3. However, when $a_0 < a_1$, the optimal rule does not exclude this possibility because it chooses the first and the third candidates that appear after time a_0 if the number of candidates on time interval (a_0, a_1) is more than one.

Define

$$\beta^*(\gamma) = \gamma + \frac{3(1-\gamma) + \sqrt{9 + 6\gamma + 9\gamma^2}}{12},$$

as a function of $\gamma \in [0, 1]$. Now we can summarize the main results as follows.

Theorem 3.3. *The threshold a_0 and the optimal payoff v^* are given as follows according to two cases.*
Case 1: $\gamma \leq \beta \leq \beta^*(\gamma)$ *for $\gamma \in [0, 1]$, or equivalently, $a_1 \leq a_0$.*
Let

$$\sigma_0 = \frac{-3(1-\gamma) + \sqrt{9 + 6\gamma + 9\gamma^2}}{2\gamma}.$$

Then

$$a_0 = \exp(-\sigma_0), \quad v^* = \frac{a_0\sigma_0}{2}(2 + \gamma\sigma_0).$$

Case 2: $\beta^*(\gamma) < \beta \leq 1$ *for $\gamma \in [0, 1]$, or equivalently, $a_0 < a_1$.*
Let σ_0 be the unique root x of the equation

$$\beta x^3 - 3\beta x^2 + 3\sigma_1 x - \sigma_1(3 + \sigma_1) = 0.$$

Then

$$a_0 = \exp(-\sigma_0), \quad v^* = \frac{a_0}{2}(\sigma_1 + \beta\sigma_0^2).$$

3.2 Choosing either one or two candidates from C_1 and C_2

W consider a problem specified by

$$\alpha_{1,2} = \alpha, \quad \alpha_1 = \beta, \quad \alpha_2 = \gamma, \quad \alpha_{1,k} = \beta, \quad \alpha_{2,k} = \gamma \quad for \; k \geq 3,$$

where $\alpha \geq \beta \geq \gamma \geq 0$ with $\alpha > 0$. All other payoffs are zero. Obviously the serious decision must be made in states (i, t) for $i = 0, 1, 2$.
Theorem 3.4.*(Optimal decision in state $(2, t)$)*
There exists a threshold $a_2 = \exp(-\sigma_2)$ such that the optimal rule stops in state $(2, t)$ if $t \geq a_2$, where

$$\sigma_2 = \frac{\sqrt{\beta^2 + \gamma^2} - (\beta - \gamma)}{\gamma}.$$

Theorem 3.5. *(Optimal decision in state $(1, t)$)*
There exists a threshold $a_1 = \exp(-\sigma_1)$ such that the optimal rule stops in state $(1, t)$ if $t \geq a_1$, where σ_1 is given as follows depending on whether $\alpha \geq \beta + \gamma$ or $\alpha < \beta + \gamma$.
Case 1: *$\alpha \geq \beta + \gamma$, or equivalently, $a_1 \leq a_2$.*
σ_1 is the unique root $x(\geq \sigma_2)$ of the equation

$$(\alpha - \gamma + \gamma x)e^{-x} = K,$$

where $K = (\beta + \gamma \sigma_2)a_2$.
Case 2: *$\alpha < \beta + \gamma$, or equivalently, $a_2 < a_1$.*
σ_1 is given by

$$\sigma_1 = \frac{\sqrt{(\beta - \gamma)^2 + 2\gamma(\alpha - \gamma)} - (\beta - \gamma)}{\gamma}.$$

We can give the following result.
Theorem 3.6. *(Optimal decision in state $(0, t)$)*
There exists a threshold $a_0 = \exp(-\sigma_0)$ such that the optimal rule stops in state $(0, t)$ if $t \geq a_0$.

Theorems 3.4, 3.5 and 3.6 state that the optimal rule is specified by three thresholds (a_0, a_1, a_2) in the sense that it chooses, in state (i, t), the current candidate if $t \geq a_i$ for each i. What is left is to determine a_0. Hereafter we only consider Case 1. Let

$$A = \frac{\gamma}{2}\sigma_1^2 + (\alpha - 2\gamma)\sigma_1 - (\alpha - \beta - \gamma).$$

Then Case 1 is distinguished into two cases Case 1(a) and Case 1(b) depending on whether or not $A \geq \gamma(1 - \sigma_1)$. Define two constants

$$P = \gamma - A = -\frac{\gamma}{2}\sigma_1^2 - (\alpha - 2\gamma)\sigma_1 + (\alpha - \beta)$$

$$Q = -\frac{\gamma}{3}\sigma_1^3 - \frac{\alpha - 2\gamma}{2}\sigma_1^2 - \beta.$$

Then a_0 and v^* are given as follows.
Theorem 3.7. *Assume $\gamma > 0$. Then the threshold a_0 and the optimal payoff v^* can be given as follows in each case. Case 1(a) is further distinguished into two cases; Case 1(a1) and Case 1(a2) depending on whether $\Phi(\sigma_1) \leq 0$ or $\Phi(\sigma_1) > 0$, where*

$$\Phi(x) = \frac{\gamma}{6}x^3 + \frac{\alpha - \gamma}{2}x^2 - (\alpha - \beta)x - \beta, \quad 0 < x.$$

Case 1(a1): *$\Phi(\sigma_1) \leq 0$, or equivalently, $a_0 \leq a_1$.*
Let

$$\sigma_0 = \frac{P + \sqrt{P^2 - 2\gamma Q}}{\gamma}.$$

Then

$$a_0 = \exp(-\sigma_0), \quad v^* = K + (A + \gamma\sigma_0)a_0.$$

Case 1(a2): $\Phi(\sigma_1) > 0$, *or equivalently,* $a_1 < a_0$.
Let σ_0 *be the unique root* $x \in (0, \sigma_1)$ *of the equation* $\Phi(x) = 0$. *Then*

$$a_0 = \exp(-\sigma_0), \quad v^* = (\frac{\gamma}{2}\sigma_0^2 + \alpha\sigma_0 + \beta)a_0.$$

Case 1(b): $a_0 < a_1$.
σ_0 *has the same expression as in Case 1(a1), that is,*

$$\sigma_0 = \frac{P + \sqrt{P^2 - 2\gamma Q}}{\gamma}.$$

Then

$$a_0 = \exp(-\sigma_0), \quad v^* = K + (A + \gamma\sigma_0)a_0.$$

3.3   Additive payoff

Suppose that the payoff is additive, i.e, $\alpha_{i,j} = \alpha_i + \alpha_j$ for $1 \leq i < j$. Then we have the following result under some condition on $\{\alpha_{i,j}\}$. Denote by $< i, t >$ the state where we are facing a candidate at time t with i choice(s) left yet to be made. For ease of description, a^*, σ^*, and v^* in Theorem 2.2 are described as a_1, σ_1, and v_1 respectively.

Theorem 3.8.   *Suppose that the sequence* $\{\alpha_k\}$ *is non-increasing in* k.
(i)   The optimal rule: There exist two thresholds $a_i = e^{-\sigma_i}$ *for* $i = 1, 2$ *such that* $a_2 \leq a_1$ *and the optimal rule stops in state* $< i, t >$ *if* $t \geq a_i$, *where* σ_2 *is the unique root* σ *of the equation*

$$\sum_{k=1}^{\infty} \frac{\alpha_{k-1} - \alpha_k}{k!}\sigma_1^k + \sum_{k=0}^{\infty} \frac{\alpha_k - \alpha_{k+1}}{k!}\sigma^k = 0.$$

(ii)   The optimal payoff: Let v_2 *denote the optimal payoff. Then*

$$v_2 = a_1 \sum_{k=1}^{\infty} \frac{\alpha_k}{k!}\sigma_1^k + a_2 \sum_{k=0}^{\infty} \frac{\alpha_{k+1}}{k!}\sigma_2^k.$$

References

[1] Bruss, F. T. (2000a). Sum the odds to one and stop. *Ann. Prob.* 28, 1384-1391.

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

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