数理解析研究所講究録
978 巻 1997 年 71-78

SINGLE-LEVEL STRATEGIES FOR FULL-INFORMATION BEST-CHOICE PROBLEMS, I

*名古屋商科大 坂口 実 (Minoru SAKAGUCHI)
**ブロッワウ工科大 K. シャヨウスキ (Krzysztof SZAJOWSKI)

ABSTRACT. Some best choice problems with full information and imperfect observation and their extension to two player competitive situation are discussed. Players observe one by one sequentially a sequence of *iid* random variables from a known continuous distribution with the objective of choosing one of the k largest. The observations of the random variables are imperfect and the player (or players) is (are) informed only of whether it is larger than or less than a previously determined decision level. The problem is to find the optimal decision level that maximizes the probability of achieving his objective. The solution of this one person game is derived. In the two player competitive situation two typical types of the optimal stopping games for choosing the best observation are formulated, transformed into continuous games on the domain $[0,\infty)^2$, and solutions of them are given. It is shown that our optimal stopping games are easy-to-state but hard-to-solve.

1. INTRODUCTION

The subject of the paper is a class of optimal stopping problems for a sequence of *iid rv*'s with full-information (FI) but imperfect observation to guarantee the maximal probability of achieving some objects. In 1975 Enns [1] first considered and solved the problem where the objective is to choose the best *rv*. Some generalization are posed and solved by Porosiński [8]. In Sakaguchi [11, 13], some more general problems are discussed. In Section 2 we formulate and solve the optimal stopping problem for finding the optimal threshold strategy that maximizes the probability of selecting one of the k largest *rv*'s. It is shown that the asymptotically optimal threshold strategy is to stop at the earliest *rv* that is larger than $e^{\frac{a_k}{n}}$, where a_k is determined by the equation

$$\sum_{j=k}^{\infty}\frac{a^j}{j(j+1)!}=\frac{1}{k}.$$

In the subsequent two sections the optimal stopping problem for selecting the best *rv*, i.e. the case $k=1$, solved in Section 2, is extended to optimal stopping games where two players compete in selecting the best *rv*. Two typical types of the optimal stopping games, where each player's objective and information condition are, by the notation used in Sakaguchi [15],

1°): Selecting best/ Players priority/ Common/ Zero-sum, with FI

and

2°): Selecting best/Earlier stop/ Each/ Non-zero-sum, with FI,

respectively in Section 3 and 4.

In 1^0) Players 1 and 2 observe a common *iid* sequence of *rv*'s, which are sampled one by one sequentially. Facing each *rv*, each player should accept the earliest *rv* that exceeds his decision level. If one player accepts a *rv* and the other rejects it, the game is left thereafter as the other player's one person game. If both players stop simultaneously at a *rv*, then Player 1 accepts it, by his priority, and the game is left, thereafter as Player 2's one person game. A single player who accepts the best *rv* is the winner, getting reward of one unit from the opponent. Player 1(2) wants to maximize (minimize) the expected payoff to Player 1. Related interesting results can be found in Neumann, Porosiński & Szajowski [7], Porosiński & Szajowski [9] and Majumdar [5]. The no-information case priority games have been considered by Enns & Ferenstein [2], Sakaguchi [12] and Szajowski [17].

In 2^0) there are two independent *iid* sets of *rv*'s. Players observe the private *iid* sequence of *rv*'s. Facing each *rv*, each player should accept the earliest *rv* that exceeds his own decision level. A player who is the first to stop at the best one in his set of *rv*'s is the winner. The case, where both players stop at the best one in each set of *rv*'s, is a draw. Players' aims are to maximize his own the winning probability. An interesting related work is Mazalov [6], where players' objective are slightly different from the one in the present paper. Other related works, but in the null information case, were done by Presman & Sonin [10], Fushimi [3] and Sakaguchi [12].

For both of cases, 1^0) and 2^0), the game is transformed into a continuous game on the domain $[0,\infty)^2$. Numerical solutions of them are given.

The considered problems are related to the full information best choice problem considered by Enns [1], Sakaguchi [11], Porosiński [8] and the game version of the problem considered by Neumann, Porosiński and Szajowski [7], and Sakaguchi and Saario [13].

2. Best choice problems for choosing one of the k-bests

Let $X_1, X_2, \ldots, X_n$ be a sequence of *iid* *rv*'s obeying uniform distribution on the unit interval $[0,1]$. The X's are sequentially observed one by one, but the observation is imperfect and we can only know whether the observed *rv* is greater than or less than a prescribed level $z \in [0,1]$. After X_t is observed, we have to either accept or reject the observation. Our aim is to accept one of the k bests among all *rv*'s. Neither recall nor uncertainty of selection is allowed. In this paper we restrict ourselves to the strategies which reject the *rv*'s less than z and accept the earliest *rv* greater than z. We call a win the event in which we accept a *rv* satisfying the objective. The event in which either we fail to accept any *rv* or accept *rv* dissatisfying his objective is a loss. We are looking for the decision level z which gives the maximum probability of win.

Let $P^{(k)}(z)$ be the probability of win for selecting one of the k bests under the strategy with the decision level z. Then since the winning event by accepting an *rv* on or before the $(n-k)$-th is

$$\bigcup_{m=1}^{k}\bigcup_{j=1}^{n-k}\left[\{X_1, X_2, \ldots, X_{j-1} \le z < X_j\} \cap \{\text{exactly } m-1 \text{ } rv \text{ thereafter } > X_j\}\right],$$

we have

$$P^{(k)}(z) = \sum_{m=1}^{k} P^{(k,m)}(z) + z^{n-k}(1 - z^k) \tag{2.1}$$

where

$$\begin{aligned} P^{(k,m)}(z) &= \sum_{j=1}^{n-k} z^{j-1} \int_z^1 \binom{n-j}{m-1} (1-x)^{m-1} x^{n-j-m+1} dx \\ &= z^n \sum_{r=k}^{n-1} z^{-(r+1)} \binom{r}{m-1} \int_z^1 (1-x)^{m-1} x^{r-m+1} dx \end{aligned} \tag{2.2}$$

is the probability that the accepted *rv* on or before the $(n-k)$-th is the m-th best with $1 \leq m \leq k$.

We use the following abbreviations in Theorem 1

- p.w. is the probability of win with the decision level z;
- a.d.l. (a.p.w.) is the asymptotic decision level (probability of win), when the decision level is such that $z = e^{-a/n}$, with $a > 0$ and $n \to \infty$.

Let us define functions

$$H_n(z) = \sum_{m=1}^{n} \frac{z^{-m} - 1}{m}, \tag{2.3}$$

$$G_n(z) = \sum_{j=n}^{\infty} \frac{z^j}{j(j+1)!}. \tag{2.4}$$

Theorem 1. *(i) For $k = 1$ (i.e. selecting the best) the probability of win is $P^{(1)}(z) = z^n H_n(z)$ and the optimal strategy is to choose z_0, which is determined by the equation*

$$H_n(z) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{z^i} \tag{2.5}$$

and the maximum of probability of win is $n^{-1} \sum_{i=0}^{n-1} z_0^i$. Therefore

$$\begin{aligned} a.d.l. &= \exp(-\frac{a_1}{n}) \\ a.p.w. &= \frac{(1 - e^{-a_1})}{a_1} \cong 0.5174 \end{aligned} \tag{2.6}$$

where $a_1 \cong 1.5029$ is a unique root in $(1, \infty)$ of the equation

$$\int_0^a \frac{e^t - 1}{t} dt = \frac{e^a - 1}{a}, \tag{2.7}$$

i.e. $G_1(a) = 1$.

(ii) For $k = 2$ (i.e. selecting one of the two best) the probability of win is

$$P^{(2)}(z) = z^n [2H_n(z) - n(\frac{1}{z} - 1)], \; n \geq 2, \tag{2.8}$$

and the optimal strategy is to choose z_0 which is determined by the equation

$$H_n(z) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{z^i} + \frac{n-1}{2} - \frac{n}{2}. \tag{2.9}$$

The maximum of $P^{(2)}(z)$ is

$$P^{(2)}(z_0) = \frac{2}{n}\sum_{i=0}^{n-1} z_0^i - z_0^{n-1}. \tag{2.10}$$

Moreover we have

$$\begin{aligned} a.d.l. &= \exp(-\frac{a_2}{n}) \\ a.p.w. &= \frac{2(1-e^{-a_2})}{a_2} \cong 0.7265 \end{aligned} \tag{2.11}$$

$-e^{-a_2}$

where $a_2 \cong 2.0177$ *is a unique root in* (a_1, ∞) *of the equation*

$$\int_0^a \frac{e^t - 1}{t} dt = \frac{e^a - 1}{a} + \frac{a-1}{2}, \tag{2.12}$$

i.e. $2G_2(a) = 1$.

(iii) For $k = 3$ *(i.e. selecting one of the three best*s*) the probability of win is*

$$P^{(3)}(z) = z^n[3H_n(z) - \frac{n(n-1)}{2}\frac{z^{-2}-1}{2} + \frac{n(n-5)}{2}(\frac{1}{z} - 1)], \; n \geq 3, \tag{2.13}$$

and the optimal strategy is to choose z_0 *which is determined by the equation*

$$H_n(z) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{z^i} + \frac{(n-1)(n-2)}{6}\frac{1}{z^2} - \frac{(n-1)(n-5)}{6}\frac{1}{z} - \frac{n+1}{3}. \tag{2.14}$$

The maximum of $P^{(3)}(z)$ *is*

$$P^{(3)}(z_0) = \frac{3}{n}\sum_{i=0}^{n-1} z_0^i + \frac{(n-1)(n-4)}{4}(z_0^{n-2} - z_0^n) + \frac{n-5}{2} z_0^{n-1}. \tag{2.15}$$

Moreover we have

$$\begin{aligned} a.d.l. &= \exp(-\frac{a_3}{n}) \\ a.p.w. &= \frac{3(1-e^{-a_3})}{a_3} - (2 + \frac{a_3}{2})e^{-a_3} \cong 0.8355 \end{aligned} \tag{2.16}$$

where $a_3 \cong 2.4934$ *is a unique root in* (a_2, ∞) *of the equation*

$$\int_0^a \frac{e^t - 1}{t} dt = \frac{e^a - 1}{a} - \frac{2}{3} + \frac{a}{2} + \frac{a^2}{12}, \tag{2.17}$$

i.e. $3G_3(a) = 1$.

3. A zero-sum best choice game where players' priority is given

A zero-sum game version of the discrete-time, full information and imperfect observation best choice problem is considered in this section. Two Players 1 and 2 observe sequentially a sequence of n *iid rv*'s from uniform distribution on $[0, 1]$, with each objective of choosing the largest *rv*. Let Player 1(2) chose a single level z (w), and he rejects *rv*'s as long as they are less than z (w) and accepts the earliest *rv*'s that is $\geq z(w)$. If the earliest *rv* that is $\geq z \vee w$ appears, Player 1 is given the priority to accept it and drops out from the game thereafter.

The player accepting the largest *rv* wins the game and he is paid a unit reward from the opponent. Player 1(2) wants to maximize (minimize) the player I's expected payoff.

Let $M(z, w)$ stand for the payoff to Player 1 when the levels z and w are chosen. Then for

$z < w$ we have

$$(3.1)\quad M(z,w) = \sum_{s=1}^{n} z^{s-1}\int_z^1 x^{n-s}dx$$
$$-\sum_{s=1}^{n-1} z^{s-1}\sum_{t=s+1}^{n} w^{t-s-1}\left\{(w-z)\int_w^1 y^{n-t}dy+\int_w^1 dx\int_x^1 y^{n-t}dy\right\}.$$
$$= z^n\sum_{i=1}^{n}\frac{z^{-i}-1}{i}+w^n\sum_{j=1}^{n}\left\{\sum_{s=1}^{n-j}\left(\frac{z}{w}\right)^s\right\}\left\{j^{-1}(w^{-j}-1)-\frac{w}{z}\frac{w^{-j-1}-1}{j+1}\right\}.$$

Similarly for $z > w$ we have

$$(3.2)\qquad M(z,w) = \sum_{s=1}^{n} w^{s-1}\int_z^1 x^{n-s}dx+\sum_{s=1}^{n-1} w^{s-1}\sum_{t=s+1}^{n} z^{t-s-1}(z-w)\int_z^1 x^{n-t}dx$$
$$-\sum_{s=1}^{n} w^{s-1}\int_w^z y^{n-s}dy-\sum_{s=1}^{n-1} w^{s-1}\sum_{t=s+1}^{n} w^{t-s-1}\int_z^1 dx\int_x^1 y^{n-t}dy.$$
$$= w^n\sum_{i=1}^{n} i^{-1}w^{-i}(1-2z^i+w^i)$$
$$+z^n\left(\frac{z}{w}-1\right)\sum_{j=1}^{n}\left\{\sum_{s=1}^{n-j}\left(\frac{w}{z}\right)^s\right\}\frac{z^{-j}-1}{j}$$
$$+w^n\sum_{j=1}^{n}(n-j)\left(\frac{z}{w}\right)^{j+1}\left\{\frac{z^{-j}-1}{j}-\frac{z^{-j-1}-1}{j+1}\right\}.$$

Note that (3.1) and (3.2) coincide at $z = w$ and so the game is a continuous game on the unit square $0 \le z, w \le 1$.

Define the function for $a \in (0,\infty)$,

$$\Phi(a) \cong \int_0^1 \frac{e^{at}-1}{t}dt = \sum_{j=1}^{\infty}\frac{a^j}{j\cdot j!}.$$

This function is strictly increasing and convex function with $\Phi(0) = 0$, $\Phi'(a) = a^{-1}(e^a - 1)$ and $\Phi''(a) = a^{-2}\{1-(1-a)e^a\} = \sum_{k=2}^{\infty}\frac{a^{k-2}}{k(k-2)!} > 0$.

Now let $z = e^{-\frac{a}{n}}$, $w = e^{-\frac{b}{n}}$, with $a, b > 0$ and $n \to \infty$. Then (3.1′) - (3.2′) becomes a continuous game on $[0,\infty)^2$, with the payoff function

$$(3.3)\qquad M(a,b) = \begin{cases} \frac{b}{a-b}(e^{-b}-e^{-a})+e^{-a}\Phi(a-b), & \text{if } 0 < b < a,\\ e^{-b}\left[(b-a)(\Phi(b)-\Phi(b-a))-\Phi(b)-a(\frac{1}{b}-1)\right] & \\ \quad +e^{-a}(\Phi(a)+1)+\frac{a}{b}-1, & \text{if } 0 < a < b, \end{cases}$$

Note that (3.3) is a continuous function on $[0,\infty)^2$ with value ae^{-a} on the halfline $a = b > 0$ and we have

$$(3.4)\qquad \begin{aligned} M(a,0) &= e^{-a}\Phi(a) = e^{-a}\int_0^a \frac{e^t-1}{t}dt,\\ M(0,b) &= -e^{-b}\Phi(b), \end{aligned}$$

both of which are reasonable facts since setting z or $w = 1$ means that the game is actually one-person game.

Computer result, on the basis of (3.3)~(3.6), gives the following

Theorem 2. *For the zero sum game on* $[0,\infty)^2$ *with the payoff function* (3.3), *a saddle point* (a_0, b_0), *with* $0 < a_0 < b_0$, *exists and the saddle value is*

$$e^{-b_0}\left[(b_0 - a_0)(\Phi(b_0) - \Phi(b_0 - a_0) - \Phi(a_0) - a_0(b_0^{-1} - 1)\right] + e^{-a_0}(\Phi(a_0) + 1) + \frac{a_0}{b_0} - 1 \cong 0.3233,$$

where $(a_0, b_0) \cong (1.57205, 2.99628)$ *is determined by a simultaneous equation*

$$e^{-a}(\Phi(a) - \Phi'(a)) + e^{-b}(\Phi(b) - \Phi'(b)) = e^{-b}\Phi(b-a)$$

$$(3.7) \qquad (\Phi(b) - \Phi(b-a)) + (\Phi(b) - \Phi'(b)) = (b-a)\{\Phi(b) - \Phi'(b) - \Phi(b-a) + \Phi'(b-a)\} + a^{-1}(e^b - 1 - b + b^2).$$

The theorem shows that Player 1 sets his decision-level $z_0 = e^{-a_0/n}$ higher than the opponent in the optimal play, and the saddle value is positive, reflecting Player 1's advantage over his opponent due to the higher priority in playing the game.

4. A NON-ZERO SUM BEST CHOICE GAME WHERE WINNING REQUIRES EARLIER STOP

A non-zero-sum game version of the discrete time, full information and imperfect observation best choice problem is considered in this section. We first state the problem as follows:

(1) There are two Players 1 and 2, and a sequence of n *iid* bivariate *rv*'s $\{(X_i, Y_i)\}_{i=1}^n$ from independent uniform distribution on $[0,1]^2$. Player 1(2) observes $X_t(Y_t)$'s sequentially one by one.

(2) Let Player 1(2) choose a *single level* $z(w)$. He rejects $X_t(Y_t)$'s as long as they are less than $z(w)$ and he accepts the earliest *rv* that is $X_\sigma \geq z$ $(Y_\tau \geq w)$, where σ and τ are the stopping times of the players.

(3) If one of the players stops (accepts), the other player is not informed of this fact and continues playing. We call a "win" for each player the event in which he gets to be the first to stop at the best one in his set of *rv*'s. If the two players stop simultaneously at the best one in each player's set of *rv*'s, both of them are the winners.

(4) The aim of each player in the game is to determine his decision level by which the probability that he becomes a *single winner* is maximized.

Let $M_1(z,w)$ $(M_2(z,w))$ stand for the probability of winning for Player 1(2) when the levels z and w are chosen. Then we have

$$(4.1) \qquad M_1(z,w) = \sum_{s=1}^{n} z^{s-1}\int_z^1 x^{n-s}dx\left\{w^s + \sum_{t=1}^{s-1} w^{t-1}\int_w^1 (1 - y^{n-t})dy\right\}.$$

$$= (zw)^n \sum_{j=1}^{n} \frac{z^{-j} - 1}{j}\left\{w^{-n} - w^{-j} + w^{-j+1} - \sum_{i=j+1}^{n} \frac{w^{-i} - 1}{i}\right\}.$$

$M_2(z,w)$ is equal to $M_1(z,w)$ with z and w interchanged.

By letting $z = e^{-a/n}$, $w = e^{-b/n}$, with $a, b > 0$ and $n \to \infty$, we obtain a non-zero-sum continuous game on $[0,\infty)^2$ with payoff functions

$$M_1(a,b) = e^{-a}\left[\Phi(a) - be^{-b}\int_0^1 \Phi(at)\Phi'(bt)dt\right], \tag{4.2}$$
$$M_2(a,b) = e^{-b}\left[\Phi(b) - ae^{-a}\int_0^1 \Phi'(at)\Phi(bt)dt\right],$$

We note that (4.2) gives

$$M_1(a,0) = e^{-a}\Phi(a),\ M_1(0,b) = 0 \tag{4.3}$$
$$M_2(0,b) = e^{-b}\Phi(b),\ M_2(a,0) = 0$$

and

$$M_1(a,a) = M_2(a,a) = e^{-a}\Phi(a) - \frac{1}{2}(e^{-a}\Phi(a))^2 = \frac{1}{2}\left[1 - (1 - e^{-a}\Phi(a))^2\right], \tag{4.4}$$

on the halfline $a = b \geq 0$. Again setting $z = 1$ ($w = 1$) means that Player 1(2) does not play any role in the game. Furthermore we find that

$$M_1(a,b) + M_2(a,b) = 1 - (1 - e^{-a}\Phi(a))(1 - e^{-b}\Phi(b)), \tag{4.5}$$

by an evident identity

$$a\int_0^1 \Phi'(at)\Phi(bt)dt + b\int_0^1 \Phi'(bt)\Phi(at)dt = \Phi(a)\Phi(b).$$

Now we find an equilibrium for a pair of payoff functions (4.2).

Theorem 3. *For the non-zero-sum game on $[0,\infty)^2$ with the payoff functions (4.2) the equilibrium point exists and it is (a,b) with $a = b = a_0 \cong 1.6065$ where a_0 is defined by $\psi(c) = 0$. The common equilibrium value is*

$$e^{-a_0}\Phi(a_0) - \frac{1}{2}(e^{-a_0}\Phi(a_0))^2 \cong 0.3830.$$

Therefore the probability of draw of the game is 0.2339.

and

$$\psi(c) = e^a \frac{\partial M_1}{\partial a}\Big|_{a=b=c}$$
$$= e^{-c}\Big[\frac{1}{2}(\Phi(c))^2 - \frac{1}{c}(\Phi(2c) - 2\Phi(c))\Big] - \Phi(c) + \Phi'(c),$$

References

1. E.G. Enns, *Selecting the maximum of a sequence with imperfect information*, J. Amer. Statist. Assoc. **70** (1975), 640 – 643.
2. E.G. Enns and E.Z. Ferenstein, *The horse game*, J. Oper. Res. Soc. Jap. **28** (1985), 51 – 62.
3. M. Fushimi, *The secretary problem in a competitive situation*, J. Oper. Res. Soc. Jap. **24** (1981), 350–358.
4. J.P. Gilbert and F. Mosteller, *Recognizing the maximum of a sequence*, J. Amer. Statist. Assoc. **61** (1966), no. 313, 35–73.
5. A.A.K. Majumdar, *Optimal stopping for a two-person sequential game in the continuous case*, Pure and Appl. Math. Sci **22** (1985), 79–89.
6. V.V. Mazalov, *A game related to optimal stopping of two sequences of independent random variables having different distributions*, Math. Japonica **43** (1996), 121–128.
7. P. Neumann, Z. Porosiński, and K. Szajowski, *A note on two person full-information best choice problems with imperfect observation*, 1993, submitted to Game Theory and Applications.
8. Z. Porosiński, *Full-information best choice problems with imperfect observation and a random number of observations*, Zastos. Matem. **21** (1991), 179–192.
9. Z. Porosiński and K. Szajowski, *On continuous-time two person full-information best choice problem with imperfect observation*, 1994, to appear in Sankhya.
10. Eh.L. Presman and I.M. Sonin, *Equilibrium points in game related to the best choice problem*, Theory of Probab. and its Appl. **20** (1975), 770–781.
11. M. Sakaguchi, *Best choice problems with full information and imperfect observation*, Math. Japonica **29** (1984), 241–250.
12. ______, *Some two-person bilateral games in the generalized secretary problem*, Math. Japonica **33** (1988), 637–654.
13. M. Sakaguchi and V. Saario, *A best choice problem for bivariate uniform distribution*, Math. Japonica **40** (1994), 585–599, Correction in : Math. Japonica 41(1995), 231.
14. ______, *A class of best choice problems with full information*, Math. Japonica **41** (1995), 389–398.
15. M.Sakaguchi, Optimal stopping games — A review, Math. Japonica 42 (1995), 343–351
16. S.M. Samuels, *An explicit formula for the limiting optimal success probability in the full information best choice problem*, Mimeograph series, Purdue Univ. Stat. Dept., 1980.
17. K. Szajowski, *Double stopping by two decision makers*, Adv. Appl. Probab. **25** (1993), 438 – 452.

* Research Institute for Information Systems, Nagoya University of Commerce and Business Administration, Komenoki, Nisshin, Aichi 470-01, Japan

** Institute of Mathematics, Technical University of Wrocław, Wybrzeże Wyspiańskiego 27, PL-50-370 Wrocław, Poland

E-mail: szajow@im.pwr.wroc.pl