数理解析研究所講究録
第 899 巻 1995 年 202-207

OPTIMAL STOPPING GAMES FOR BIVARIATE UNIFORM DISTRIBUTION

名古屋商科大 坂口 実 (Minoru SAKAGUCHI)

Abstract We consider a class of two-person time-sequential games called optimal stopping games. Let (X_i, Y_i), $i=1, \cdots, n$, be an *iid* sequence of r.v.'s sampled from bivariate uniform distribution on [0, 1]. At each time $i=1,2, \cdots$, each of two players I and II is dealt with a hand X_i and Y_i, respectively. After looking at his hand *privately*, each player can then choose either to accept (A) his hand or to reject (R) it. If the players' choice pair is A–A, then the game ends with the predetermined payoffs to the players. If the choices are R–R, then the current sample is rejected and the game continues to facing a next sample (X_{i+1}, Y_{i+1}). If the choices are A–R(R–A)then a lottery is used to the effect that either A–A or R–R is enforced to the players with probability p_1 (p_2) and $\overline{p}_1$ ($\overline{p}_2$), respectively, where $\overline{p}_i = 1-p_i$. Each player wants to maximize his expected payoff at the termination time of the game. We explicitly derive the solutions of (1) zero-sum game, where the terminal payoff to I is $X_\tau - Y_\tau$, (2) non-zero-sum game, where the terminal payoffs are $E(X_\tau)$-$E(Y_\tau)$, where τ is the time at which the game is stopped.

1. Introduction and Summary.

A sequence of n *iid* random variable (X_i, Y_i), $i=1, \cdots, n$, is sampled sequentially one by one from bivariate uniform distribution on the unit square $[0, 1]^2$.

We consider a class of two-person time- sequential games called optimal stopping games (OSG), with the underlying *r.v.'s* $\{(X_i, Y_i)\}_{i=1}^{n}$.
Each of two players (I and II) draws a number (x and y, respectively) according to a bivariate uniform distribution. Player I (II) knows his hand x (y) only, and doesn't know his opponent's hand. After observing his number each player can then choose, simultaneously and independently of opponent's choice, either to accept (A) his number or to reject (R) it. If the players' choice-pair is A–A, then the game ends with payoffs $x-y$ to I and $y-x$ to II. If the players' choice-pair is R–R, then the current sample (x, y) is rejected, the game continues to the next stage, and the new r.v. (X', Y') is sampled. If the players' choice-pair is A-R(R-A), then a lottery is used to the effect that either A–A or R–R is enforced [by Umpire, so-to-speak] to the players with probability p_1 (p_2) and $\overline{p}_1$ ($\overline{p}_2$), respectively, where $\overline{p} = 1-p$. If $0 \leq p_2 < p_1 \leq 1$, then I(II) is a *strong (weak)* player. I and II are *equal* players if $p_1 = p_2$. Player I(II) wants to maximize (minimize) I's expected payoff.

We should note that the games discussed in this paper have the following *Players' Self-willing Property* (PSP): Each player preferrs to accept high hands and reject low hands. Player I(II) wants to accept his high hand, whereas his opponent wants to reject his low-hand at the same time, then he is possible to carry his self-will through with probability p_1 (p_2).

Let $\Gamma(n, p_1, p_2)$ denote the two-person zero-sum time-sequential game described above. Also let u_n represent the value of the game $\Gamma(n, p_1, p_2)$. Then u_n satisfies the Optimality Equation (OE) of dynamic programming

$$(1.1)\quad u_n = \text{pval}\left\{ \begin{matrix} & A & R \\ A & x-y & p_1(x-y)+\bar{p}_1 u_{n-1} \\ R & p_2(x-y)+\bar{p}_2 u_{n-1} & u_{n-1} \end{matrix} \right\}$$

$$= u_{n-1} + \text{pval}\left\{ (x-y-u_{n-1}) \begin{bmatrix} 1 & p_1 \\ p_2 & 0 \end{bmatrix} \right\}, \quad (n=1,2,\cdots;\ u_0 \equiv 0).$$

Here we have used the notion of "poker value" abbreviated by *pval*, and defined, for any 2×2 matrix $A(x, y)$ involving r.v.(x, y), by

$$\max_{0 \le \alpha(\cdot) \le 1} \min_{0 \le \beta(\cdot) \le 1} \int_0^1 \int_0^1 (\alpha(x), \bar{\alpha}(x))\, A(x, y) \begin{bmatrix} \beta(y) \\ \bar{\beta}(y) \end{bmatrix} dx\,dy,$$

where $\alpha(x)(\beta(y))$, with values in [0, 1], represents the probability of choosing A when player I(II)'s hand is $x(y)$. (See Sakaguchi [6]).

In (1.1) if $p_1 = p_2 = 0$ [$p_1 = p_2 = 1$] then the game is an AND [OR] game in which both players [at least one player] must choose A in order for an END of the game to occur. Later we observe that $\Delta \equiv p_1 + p_2 - 1$ has an important role in the analysis of the game.

In Section 2 the solution of the game $\Gamma(n, p_1, p_2)$ is derived. It is shown that player I(II) should accept his hand if and only if it is higher than a certain level a_n^* (b_n^*), depending on n, p_1 and p_2, at the first stage of the game, where(a_n^*, b_n^*) is determined by a unique root in $[0, 1]^2$ of a certain simultaneous equation. It is also shown that in the special case of $p_1 = p_2 = p$, the game reduces to a one-shot game, in which the value of the game is zero and the optimal strategy of each player is to accept his hand if and only if it is higher than a unique root in [0, 1] of the equation $a^2\Delta + 2\bar{p}a - \bar{p} = 0$.

In Section 3 we will discuss about the non-zero-sum-game version $G(n, p_1, p_2)$, where the objective for player I[II] is to maximize $E(X_\tau)[E(Y_\tau)]$ if the game is stopped at time τ. We consider as the underlying distribution, bivariate uniform distribution on the unit square $[0, 1]^2$, with pdf

$$h(x, y) = 1 + \gamma(1-2x)(1-2y),$$

where γ, $|\gamma| \le 1$, is a given constant. The solution of the game $G(n, p_1, p_2)$ is derived, and it is shown that the equilibrium values (u_n, v_n) of the game are determined by a certain simultaneous recursive relation, and that player I(II), at the first stage of the game, should accept his hand if and only if it is higher than u_{n-1} (v_{n-1}). It is also shown that, in the special case of $p_1=p_2=p$, we obtain $u_n=v_n$ for all n, and if $\gamma=0$ additionally, then $\{u_n\}$ converges, as n→∞, to a unique root u_∞ in [0, 1] of the equation $u^2\Delta+(2-p)u-1=0$. Moreover we obtain in case of $p=\gamma=0$, a two-person non-zero-sum-game version of the well-known Moser's sequence of numbers $V_n=(1/2)(1+V_{n-1}^2)$, $(n\geqq 1, V_0=0)$, (See Moser [4] and Gilbert and Mosteller [2; Section 5b]).

The notion of the poker value was first introduced by Karlin [3; chapter 9] and was applied to one -shot exchange games by Brams, Kilgour and Davis [1] and Sakaguchi [7, 8], and to multistage poker games by Sakaguchi [6], and Sakaguchi and Sakai [9]. Moreover a very recent research closely related to this paper is Mazalov [5], in which a sequential zero-sum game over a given bivariate distribution with cdf F(x) G(y) is solved, but the rule of the game is different from that assumed in this paper. There, the players are allowed to freely choose their stopping times σ and τ and the terminal payoff is $\mathrm{sgn}(X_\sigma-Y_\tau)$.

2. Zero-Sum Sequential Game. within some specified classes

We now go to deriving the solution of the zero-sum game described by the OE (1. 1). Players' Self-willing Property mentioned in the previous section leads to the conjecture that each player should choose A if and only if his hand is higher than some determined number, and using this conjecture we can show the following result.

Theorem 1. *For the zero-sum sequential game* $\Gamma(n,p_1,p_2)$ *with* OE(1.1), *the optimal strategy-pair at the first stage is*

$$\alpha_n^*(x)=0,\ \text{if } x<a_n^*\ ;\ =1,\ \text{if } x>a_n^*$$
$$\beta_n^*(y)=0,\ \text{if } y<b_n^*\ ;\ =1,\ \text{if } y>b_n^*,$$

where the threshold level-pair $a_n^*-b_n^*$ *is determined by a unique root of the simultaneous equation in* [0, 1]

$$(2.1)\qquad \begin{cases} a_n=u_{n-1}+\dfrac{1}{2}\left(\dfrac{\bar{p}_2+b_n^2\Delta}{\bar{p}_2+b_n\Delta}\right), \\ b_n=-u_{n-1}+\dfrac{1}{2}\left(\dfrac{\bar{p}_1+a_n^2\Delta}{\bar{p}_1+a_n\Delta}\right), \end{cases} \qquad (n=1,2,\cdots;\ u_0\equiv 0).$$

The value of the sequential game is given by the recursion

$$(2.2)\qquad u_n=(\bar{p}_2a+\bar{p}_1b+ab\Delta)u_{n-1}+\frac{1}{2}\{\bar{p}_2a\bar{a}-\bar{p}_1b\bar{b}+ab(a-b)\Delta\},$$

with a *and* b *replaced by* a_n^* *and* b_n^*, *respectively.*

Corollary 1.1 *In Theorem 1, let* $p_1=p_2=p$. *Then we have* $u_n=0$, $a_n=b_n=a$, *for all n, and*

$$(2.7)\qquad a=-q\pm\sqrt{q^2+q}\ ,\ (\textit{with}+[-]\textit{sign, if } q>[<]0,\ \textit{i.e, if } p>[<]\ 1/2\ .)$$

where $q=\bar{p}/(1-2\bar{p})$ *. The game reduces to the one-shot game with the payoff matrix* $(x-y)\begin{bmatrix} 1 & \bar{p} \\ p & 0 \end{bmatrix}$.

We observe that if the game is AND (i.e. $p_1=p_2=0$), the players never accept no matter how high their hands are. If the game is OR(i.e. $p_1=p_2=1$) the players never reject, no matter how low their hands are. (See Brams, Kilgour and Davis [1], and Sakaguchi [7]).

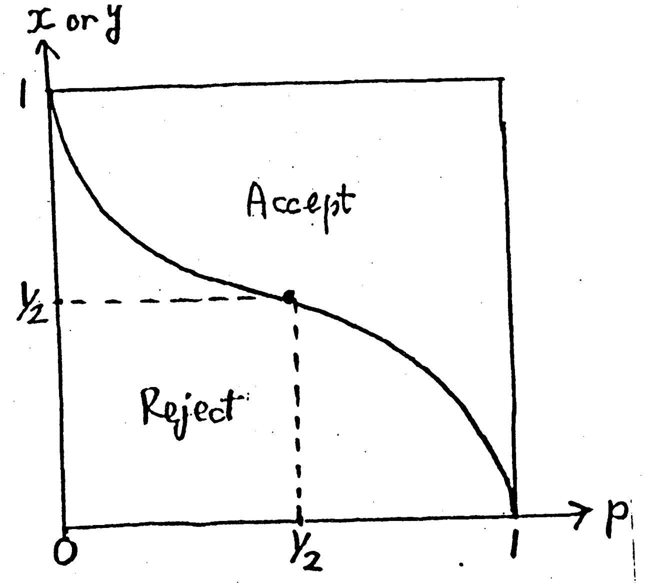

Figure 2. The function $a(p)$ bordering the two decision regions.

3. Non-Zero-Sum Sequential Game.

In this section we consider bivariate uniform distribution on the unit square $[0, 1]^2$,

(3.1) $$h(x,y)=1+\gamma(1-2x)(1-2y), \qquad 0\leq x,\ y\leq 1,$$

where γ , $|\gamma|\leq 1$, is a given constant. This bivariate pdf is one of the simplest one that has identical uniform marginals and correlated component variables. The correlation coefficient is equal to $(1/3)\gamma$.

A class of pdf's with the given marginal pdf's f(x) and g(y) is given by

$$h(x,y)=f(x)g(y)\{1+\gamma(1-2F(x))(1-2G(y))\}, \qquad |\gamma|\leq 1,$$

where F(x) and G(y) are the corresponding cdf's. If (X, Y) has this pdf, then the bivariate r.v. (F(X), G(Y)) is distributed with pdf of (3.1) with the same γ .

For any 2×2 bimatrix game $[A^1(x,y), A^2(x,y)]$ involving r.v.(x, y), the "equilibrium poker value" abbreviated by *eq. pval.* is defined by

$$\underset{(\alpha(\cdot),\,\beta(\cdot))}{\text{eq.val.}}\left(M^1(\alpha,\beta),\, M^2(\alpha,\beta)\right)$$

where

$$M^i(\alpha,\beta) \equiv \int_0^1\int_0^1 (\alpha(x), \bar{\alpha}(x)) A^i(x,y) \begin{bmatrix}\beta(y)\\ \bar{\beta}(y)\end{bmatrix} h(x,y)\,dx\,dy, \quad (i=1,2)$$

if it exists uniquely.

Let (u_n, v_n) represent the equilibrium values of the non-zero-sum game $G(n, p_1, p_2)$, where player I[II]'s objective is to maximize $EX_\tau[E(Y_\tau)]$ if the game is stopped at "time" τ. Then (u_n, v_n) satisfies the OE

(3.2)

$$(u_n, v_n) = \text{eq.p val}\left\{\begin{array}{c|c|c} & A & R \\ \hline A & x,\ y & p_1 x + \bar{p}_1 u_{n-1},\ p_1 y + \bar{p}_1 v_{n-1} \\ \hline R & p_2 x + \bar{p}_2 u_{n-1},\ p_2 y + \bar{p}_2 v_{n-1} & u_{n-1},\ v_{n-1} \end{array}\right\}$$

$$(n = 1, 2, \cdots;\ u_0 = v_0 \equiv 0)$$

We shall prove the following

Theorem 2. *For the non-zero-sum sequential game* $G(n, p_1, p_2)$ *over bivariate uniform distribution* (3.1) *with* $0 \leqq \gamma \leqq 1$, *the equilibrium values* (u_n, v_n) *satisfy the recurrence relation*

(3.3a) $$u_n = a + \frac{1}{2}\left\{(1-\bar{p}_1 b)\bar{a}^2 - p_2\bar{b}a^2\right\} + \gamma\, b\bar{b}\left\{\frac{1}{6}\bar{p}_1 + \left(\frac{1}{2}a^2 - \frac{1}{3}a^3\right)\Delta\right\}$$

(3.3b) $$v_n = b + \frac{1}{2}\left\{(1-\bar{p}_2 a)\bar{b}^2 - p_1\bar{a}b^2\right\} + \gamma\, a\bar{a}\left\{\frac{1}{6}\bar{p}_2 + \left(\frac{1}{2}b^2 - \frac{1}{3}b^3\right)\Delta\right\}$$

$$(n = 1, 2, \cdots;\ u_0 = v_0 \equiv 0,\ a_0 = b_0 = 0)$$

with a *and* b *replaced by* u_{n-1} *and* v_{n-1}, *respectively.*

The equilibrium strategy-pair at the first stage is

(3.4a) $$\alpha^*(x) = 0,\ \text{if } x < u_{n-1};\ = 1,\ \text{if } x > u_{n-1}$$

(3.4b) $$\beta^*(y) = 0,\ \text{if } y < v_{n-1};\ = 1,\ \text{if } y > v_{n-1}.$$

Corollary 2.1 *In Theorem 2. let* $p_1 = p_2 = p$ *and* $(a, b) = (u_{n-1}, v_{n-1})$. *Then we have* $u_n = v_n$ *for all* n, *and*

$$u_n = \begin{cases} a + \frac{1}{2}\left\{\bar{a}^2 - a\bar{a}(pa + \bar{p}\bar{a})\right\} + \gamma a\bar{a}\left\{\frac{1}{6}\bar{p} + \left(\frac{1}{2}a^2 - \frac{1}{3}a^3\right)\Delta\right\}, & \text{if } p \in [0,1] \\ a + \frac{1}{2}\bar{a}^3 + \gamma a\bar{a}\left(\frac{1}{6} - \frac{1}{2}a^2 - \frac{1}{3}a^3\right), & \text{if } p=0 \text{ (AND)} \\ a + \frac{1}{2}\bar{a}(\bar{a} - a^2) + \gamma a\bar{a}\left(\frac{1}{2}a^2 - \frac{1}{3}a^3\right), & \text{if } p=1 \text{ (OR)} \end{cases}$$

The sequence $\{u_n\}$ *is increasing for all* $0 \leqq \gamma \leqq 1$, *when* $p=0$.

以下,略す.

REFERENCES

[1] Brams, S. J., Kilgour, D. M., and Davis, M. D., *Unraveling in games of sharing and exchange* (to appear).

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

[3] Karlin S, *Mathematical methods and theory in games, Programming and Economics, Vol II,* Pergamon Press, London 1959.

[4] L. Moser, *On a problem of Cayley,* Scripta Math. 22(1956), 289–292.

[5] V. V. Mazalov, *A game related to optimal stopping of two sequences of independent random variables having different distributions* (to appear).

[6] M. Sakaguchi, *A simplified two-person multistage poker with optional stopping,* Math. Japonica 28(1983), 287–303.

[7] ——————, *On two and three person exchange games,* Math. Japonica 38(1993), 791–801.

[8] ——————, *Information structure and perfect information in simple exchange games,* to appear in Intern. Year Book. Game Theory and Applications, 2(1994).

[9] ——————, and S. Sakai, *Multistage poker with random amount of bets,* Math. Japonica 37(1992), 827–838.