OPTIMAL STOPPING GAMES FOR BIVARIATE UNIFORM DISTRIBUTION (Optimization Theory and its Applications in Mathematical Systems)

Sakaguchi, Minoru

Number of pages: 202-207

Publication: 数理解析研究所講究録 (1995), 899

URL: http://hdl.handle.net/2433/84501

Departmental Bulletin Paper

Textversion: publisher

Kyoto University
OPTIMAL STOPPING GAMES FOR BIVARIATE UNIFORM DISTRIBUTION

Abstract We consider a class of two-person time-sequential games called optimal stopping games. Let \((X_i, Y_i), i=1, \ldots, n\), be an iid sequence of r.v.'s sampled from bivariate uniform distribution on \([0, 1]\). At each time \(i=1, 2, \ldots\), 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 \(\overline{p}_1 = 1-p_1\). 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 \(\tau\) is the time at which the game is stopped.

1. Introduction and Summary.

A sequence of \(n\) iid random variable \((X_i, Y_i), i=1, \ldots, n\), 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 \((x, y)\), 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 \(\overline{p}_1 = 1-p_1\). 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 prefers 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 \( \mu \) represent the value of the game \( \Gamma(n, p_1, p_2) \). Then \( \mu \) satisfies the Optimality Equation (OE) of dynamic programming

\[
\mu = \max \left\{ A \left[ \begin{array}{cc} \sigma & p_1 (x-y) \beta_1 (x-y) \beta \mu \mu_{n-1} \\ 0 & \mu_{n-1} \end{array} \right] \right\}
\]

\[
= \mu_{n-1} + p_1 \max \left\{ (x-y) \mu_{n-1} \right\} \left[ \begin{array}{c} 1 \\ p_2 \end{array} \right] \right\}, (n=1,2,\ldots; \mu_0 = 0).
\]

Here we have used the notion of "poker value" abbreviated by \( p\nu \), and de fined, for any \( 2 \times 2 \) matrix \( A(x, y) \) involving \( r.v.(x, y) \), by

\[
\mu(A) = \int_0^1 \int_0^1 \mu_0(x, y) \cdot A(x, y) \cdot \left[ \begin{array}{c} \alpha(y) \\ \beta(y) \end{array} \right] \cdot dx \cdot dy,
\]

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 \( p_1 = p_2 = 0 \) 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 \( \Delta = p_1 + p_2 - 1 \) 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 \( \alpha^* \), \( \beta^* \), depending on \( n, p_1 \) and \( p_2 \), at the first stage of the game, where \( \alpha^*(x), \beta^*(y) \) is determined by a unique root in \([0, 1]\) 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 \( \alpha^2 + 2 \beta \alpha - \beta = 0 \).

In Section 3 we will discuss about the non-zero-sum-game version \( \Gamma(n, p, p_2) \), where the objective for player I(II) is to maximize \( E(X_1|e(X_e)) \) if the game is stopped at time \( \tau \). 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 \( \gamma \), \( \gamma \), \( 1 \), is a given constant. The solution of the game \( \Gamma(n, p_1, p_2) \) is derived, and it is shown that the equilibrium values \( (\mu_n, v_n) \) 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 \( P_i = P_2 = p \), we
obtain \( u_n = v_n \) for all \( n \), and if \( T = 0 \) additionally, then \( \{u_n\} \) converges, as \( n \to \infty \),
to a unique root \( u_{\infty} \) in \([0, 1]\) of the equation
\[
u^2 \Delta + (2 - p) u - 1 = 0.\]
Moreover we obtain in case of \( P_i = T = 0 \), a two-person non-zero-sum-game version of the
well-known Moser's sequence of numbers \( V_n = (\frac{\nu}{2}) (1 + \nu_{n-1}) \), \( n \geq 1, V_0 = 0 \).
(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 \( \sigma \) and \( \tau \) and the terminal payoff is \( \text{sgn}(X_\sigma - Y_\tau) \).

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 \( \Gamma(n, p, p_2) \) with OE(1.1),
the optimal strategy-pair at the first stage is
\[
\alpha_n^*(x) = 0, \quad \text{if} \quad x < \alpha_n^* ;
\]
\[
\beta_n^*(y) = 0, \quad \text{if} \quad y < \beta_n^* ;
\]
where the threshold level-pair \( \alpha_n^* - \beta_n^* \) is determined by a unique root of the
simultaneous equation in \([0, 1]\)
\[
\begin{align*}
\left\{ \begin{array}{l}
\alpha_n = u_{n-1} + \frac{1}{2} \left( \frac{\bar{P}_1 a + \bar{P}_2 b}{\bar{P}_1 + \bar{P}_2} \Delta \right), \\
\beta_n = u_{n-1} + \frac{1}{2} \left( \frac{\bar{P}_1 a + \bar{P}_2 b}{\bar{P}_1 + \bar{P}_2} \Delta \right),
\end{array} \right. \\
(2.1)
\end{align*}
\]

The value of the sequential game is given by the recursion
\[
u_n = \left( \bar{P}_1 a + \bar{P}_2 b + ab \Delta \right) u_{n-1} + \frac{1}{2} \left( \bar{P}_1 a - \bar{P}_2 b + ab (\alpha_n - \beta_n) \Delta \right),
\]
with \( a \) and \( b \) replaced by \( \alpha_n^* \) and \( \beta_n^* \), respectively.

Corollary 1.1 In Theorem 1, let \( p_i = p_2 = p \). Then we have \( u_n = 0 \)
\[
\alpha_n = \beta_n = \alpha, \quad \text{for all} \; n, \; \text{and}
\]
\[
\alpha = -a \pm \sqrt{a^2 + q}, \quad \text{(with \text{sign, if} q > 0, \; i.e., \; if p > < \; \text{\textbf{\text{2}}} \; \text{\textbf{\text{2}}})}
\]
where \( q = \frac{p}{1 - 2p} \). The game reduces to the one-shot game with the payoff matrix
\[
\begin{pmatrix}
\frac{1}{p} & \frac{p}{b} \\
0 & b
\end{pmatrix}.
\]

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.](image)


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

\[
A(x, y) = 1 + 1 - z(1 - 2x)(1 - 2y), \quad 0 \leq x, y \leq 1,
\]

where \( z, |z| \leq 1 \), 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 \( \frac{1}{2} \). A class of pdf's with the given marginal pdf's \( f(x) \) and \( g(y) \) is given by

\[
h(x, y) = f(x)g(y) \left[ 1 + z(1 - 2F(x))(1 - 2G(y)) \right], \quad |z| \leq 1,
\]

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 \( z \).

For any 2 x 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
where
\[ M_i(\alpha, \beta) = \int_0^1 \int_0^1 (\alpha(x), \beta(y)) \left[ \frac{\beta(y)}{\beta(y)} \right] h(x, y) \, dx \, dy, \quad (i=1, 2) \]

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 \(E_x[E(Y|Y)]\) if the game is stopped at "time" \(\tau\). Then \((u_n, v_n)\) satisfies the OE

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

\[
(n=1, 2, \ldots; \quad u_0 = v_0 = 0)
\]

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 \(0 \leq y \leq 1\), the equilibrium values \((u_n, v_n)\) satisfy the recurrence relation

\[
\begin{align*}
(3.3a) & \quad u_n = a + \frac{1}{2} \left[ (1-p_1 \beta) a^2 - p_2 \beta a^2 \right] + \gamma \, b b \left[ \frac{1}{6} \bar{p}_1 + \left( \frac{1}{2} a_1 - \frac{1}{2} a^2 \right) \Delta \right] \\
(3.3b) & \quad v_n = b + \frac{1}{2} \left[ (1-p_2 a) b^2 - p_1 \alpha a^2 \right] + \gamma \, a a \left[ \frac{1}{6} \bar{p}_2 + \left( \frac{1}{2} b_1 - \frac{1}{2} b^2 \right) \Delta \right] \\
(n=1, 2, \ldots; \quad u_0 = v_0 = 0, \quad a_0 = b_0 = 0)
\end{align*}
\]

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

The equilibrium strategy-pair at the first stage is

\[
\begin{align*}
(3.4a) & \quad \alpha^*(x) = 0, \quad \text{if } x < u_{n-1}; \quad = 1, \quad \text{if } x > u_{n-1} \\
(3.4b) & \quad \beta^*(y) = 0, \quad \text{if } y < v_{n-1}; \quad = 1, \quad \text{if } y > v_{n-1}.
\end{align*}
\]

**Corollary 2.1** In Theorem 2, let \(p_1 = p_2 = \frac{1}{2}\) and \((a, b) = (u_{n-1}, v_{n-1})\). Then we have \(u_n = v_n\) for all \(n\), and
\[
\begin{align*}
\begin{bmatrix}
\alpha + \frac{t}{2} \left( \tilde{a}^2 - a \tilde{a} (\alpha + \tilde{a}) \right) + \gamma \tilde{a} \alpha \left( \frac{1}{6} \tilde{a} + \frac{1}{3} a^2 \tilde{a}^2 \right) \\
\alpha + \frac{t}{2} \tilde{a}^3 + \gamma a \tilde{a} \left( \frac{1}{6} \frac{1}{2} a^2 \tilde{a}^2 \right)
\end{bmatrix}
\end{align*}
\]

if \( p \in [0,1] \)

\[
\begin{align*}
\begin{bmatrix}
a + \frac{t}{2} \left( \tilde{a}^2 - a \tilde{a} (\alpha + \tilde{a}) \right) + \gamma \tilde{a} \alpha \left( \frac{1}{6} \tilde{a} + \frac{1}{3} a^2 \tilde{a}^2 \right) \\
\alpha + \frac{t}{2} \tilde{a}^3 + \gamma a \tilde{a} \left( \frac{1}{6} \frac{1}{2} a^2 \tilde{a}^2 \right)
\end{bmatrix}
\end{align*}
\]

if \( p = 0 \) (AND)

\[
\begin{align*}
\begin{bmatrix}
a + \frac{t}{2} \left( \tilde{a}^2 - a \tilde{a} (\alpha + \tilde{a}) \right) + \gamma \tilde{a} \alpha \left( \frac{1}{6} \tilde{a} + \frac{1}{3} a^2 \tilde{a}^2 \right) \\
\alpha + \frac{t}{2} \tilde{a}^3 + \gamma a \tilde{a} \left( \frac{1}{6} \frac{1}{2} a^2 \tilde{a}^2 \right)
\end{bmatrix}
\end{align*}
\]

if \( p = 1 \) (OR)

The sequence \( \{ u_n \} \) is increasing for all \( 0 \leq \gamma \leq 1 \), when \( p = 0 \).

以下：田多様

REFERENCES


