数理解析研究所講究録 1252 巻 2002 年 27-33

OPTIMAL STOPPING GAMES BY EQUAL-WEIGHT PLAYERS FOR POISSON-ARRIVING OFFERS *

MINORU SAKAGUCHI ** (坂口 実)

Abstract. Two players observe a Poisson stream of offers. The offers are i.i.d. r.v.s from $U_{[0,1]}$ distribution. Each player wishes to accept one offer in the interval $[0,T]$ and each aims to select an offer as large as possible. Offers arrive sequentially and decisions to accept or reject must be made immediately after the offers arrive. Players have equal weights, so if both players want to accept a same r.v., a lottery is used to the effect that each player can get it with probability $1/2$. If one player accepts a r.v. and the other doesn't, the game goes on as one-person game for the latter. Player who fails to accept any r.v. before T gets a reward of zero. Each player wants to maximize his expected reward. The normal form of the game is formulated. By introducing a Riccati differential equation, the explicit solution is given to this game to calculate the Nash value and the equilibrium strategies. The bilateral-move version of the game is also analysed and the explicit solution is derived. It is shown that the second-mover stads unfavorable, on the contrary to the case in multi-round poker.

1. Optimal Stopping Game for Poisson-arriving Offers.

Players I and II must make a decision to accept (A) or reject (R) an offered job at each offer presentation. The offers arrive during time interval $[0, T]$ as a Poisson process with rate λ. The offered jobs have random sizes being i.i.d. random variables from a uniform distribution on $[0,1]$. Whenever an offer with size x arrives it is presented to both players simultaneously, and players must choose either A or R. If the players' choice-pair is A-R or R-A then the player who chooses A gets x dropping out from the game thereafter, and the other player continues his (or her) one-person game. If the choice-pair is A-A then a lottery is used to the effect that A-R or R-A is enforced to the players with equal probability $\frac{1}{2}$. If the choice-pair is R-R, then the current sample x is rejected and the game passes on to the time when a new job arrives next. Player who cannot accept any offer until time T gets a reward of zero. Each player aims to maximize his expected reward.

Define state (x,t) to mean that (1) both players remain in the game, and (2) an offer with size x has just arrived at time $T-t$ (i.e. the remaining time until horizon is t). Let $\varphi(x,t)$ $(\psi(x,t))$, be the probability of choosing A by player I (II) in state (x,t). Also let $V_i(t,\varphi,\psi)$ be the expected reward for player i at time t left to go, if players employ strategies φ and ψ. Then the game is descrived by the following differntial equations (if one considers the possible events when the residual time decreases from t to $t-\Delta t$ and takes the limit as $\Delta t \to 0$).

$$(1.1)\qquad \lambda^{-1}(V_1'(t), V_2'(t)) = -(V_1(t), V_2(t)) + \int_0^1 (\bar{\varphi}, \varphi)\underline{M}(x,t)(\bar{\psi}, \psi)^T dx$$

with the initial conditions $V_1(0)=V_2(0)=0$ and

$$(1.2)\qquad \underline{M}(x,t) = (\mathrm{I})\left\{\begin{array}{c|cc} & \multicolumn{2}{c}{(\mathrm{II})} \\ & R & A \\ \hline R & V_1(t),\ V_2(t) & U(t),\quad x \\ A & x,\ U(t) & \frac{1}{2}(x+U(t)),\ \frac{1}{2}(x+U(t)) \end{array}\right.$$

Here $V_i(t,\varphi,\psi)$, $\varphi(x,t)$ and $\psi(x,t)$ are abreviated by $V_i(t)$, φ and ψ, respectively, and

$$(1.3)\qquad U(t) = \frac{\lambda t}{2+\lambda t}, \qquad 0 \leq t \leq T$$

is the optimal reward for the one-person game at time t left to go. The optimal strategy for this one-person game is to choose A (R) if $x > (<)\ U(t)$. One obtains (1.3) by solving the differential equation $\lambda^{-1}U'(t) = -U(t) + \int_0^1 x \vee U(t)\,dx = \frac{1}{2}(1-U(t))^2$, $U(0)=0$.

We want to solve the problem: $(V_1(T,\varphi,\psi), V_2(T,\varphi,\psi)) \longrightarrow$ Nash eq. $(\varphi(\cdot), \psi(\cdot))$

The explicit solution to the simultaneous-version of the game (1.1)~(1.3) is given in Section 2.

2. Simultaneous-move Game.

We prove

Theorem 1. Let $m(t)$ be the solution of the Riccati differential equation

$$(2.1)\qquad \lambda^{-1}m'(t) = \frac{1}{4}(1+2U-3U^2) - (1-U)m + k(U-m)^2 \quad \text{with } m(0)=0,$$

where $k=2(1-\log 2) \doteqdot 0.6137$, and t of m(t) and U(t) are omitted in the r.h.s. Then the strategy-pair

$$(2.2)\qquad \varphi_*(x,t) = \psi_*(x,t) = \begin{cases} 0, & \text{if } 0 \leq x < m(t), \\ \dfrac{x-m(t)}{\frac{1}{2}(x+U(t))-m(t)}, & \text{if } m(t) < x < U(t), \\ 1, & \text{if } U(t) < x \leq 1, \end{cases}$$

is in equilibrium with a common equlibrium value $m(T)$.

To solve the Riccati differential equation (2.1) is not desparate, because we can find one particular solution

(2.5) $$m_0(t)=(1+\gamma)U(t)-\gamma,$$

where $\gamma=(\sqrt{4k+1}-1)/(4k)=\dfrac{\sqrt{9-8\log 2}-1}{8(1-\log 2)}\doteqdot 0.3498.$

Corollary 1.1

$$m(t)=m_0(t)+(p(t)+\gamma^{-1})^{-1},$$

where $m_0(t)$ is given by (2.5), or equivalently,

(2.9) $$m_0(t)=1-(1+\gamma)(1+\tfrac{1}{2}\lambda t)^{-1},$$

and p (t) is given by

(2.8) $$p(t)=\frac{1}{2\gamma(1-\gamma)}\left\{(1+\tfrac{1}{2}\lambda t)^{\gamma^{-1}}-1+(\tfrac{1}{2}-\gamma)\lambda t\right\}$$

where $\gamma^{-1}=\sqrt{9-8\log 2}+1\doteqdot 2.8587$ and $\dfrac{1}{2\gamma(1-\gamma)}=\dfrac{1}{2}\left\{(4k+1)^{1/4}+(4k+1)^{-1/4}\right\}^2\doteqdot 2.1983$

We see that $1-m(T)=O((\lambda T)^{-1})$, as $\lambda T\to\infty$.

Corollary 1.2 $m(t)$ is concave and increasing with $m(0)=0$ and $m(t)\to 1$, as $\lambda t\to\infty$.

Corollary 1.3 The times until the game become one-person game has the defective pdf

(2.10) $$g(s)=\lambda p(s)\exp\left[-\lambda\int_0^s p(w)dw\right],$$

where

$$p(s)\equiv\int_0^1\{2\varphi_*(x,T-s)-\varphi_*^2(x,T-s)\}dx.$$

The probability of getting zero by both players is

$$1-\int_0^T g(s)ds=\exp\left[-\lambda\int_0^T p(s)ds\right].$$

3. Bilateral-move Version

We shall also discuss about the bilateral-move version of the game as follows: In each state (x,t) players' moves are split into two steps. Player I first decides to choose either R or A, and then player II, after being informed of the choice chosen by I, decides to choose either R or A.

The rest of the game rule is the same as in the simultaneous-move version. So, the game in state (x,t) is described by

Players	1st step	2nd step	Payoffs
I: (x,t)	R		
	A		
II: (x,t)	after R	R	$(V_1(t-\Delta t), V_2(t-\Delta t))$
		A	$(U(t-\Delta t),\ x)$
	after A	R	$(x,\ U(t-\Delta t))$
		A	$(\frac{1}{2}(x+U(t-\Delta t)), \frac{1}{2}(x+U(t-\Delta t)))$

Let $\psi_R(x,t)$ ($\psi_A(x,t)$) be the probability that II chooses A, after he is informed of the fact that I has chosen R (A). Also let $\varphi(x,t)$ be the probability that one chooses A.

Denote $V_1(t,\varphi,\psi_{R*},\psi_{A*})$ and $V_2(t,\varphi_*,\psi_R,\psi_A)$, $V_i(t,\varphi_*,\psi_{R*},\psi_{A*})$ simply by $V_1(t)$, $V_2(t)$ and $V_{i*}(t)$, respectively. Our problem is now

$$(V_1(T,\varphi,\psi_R,\psi_A), V_2(T,\varphi,\psi_R,\psi_A)) \longrightarrow \underset{(\varphi,\ \psi_R\text{-}\psi_A)}{\text{Nash eq.}}$$

Then since player II's behavior after being informed of I's $\left\{\begin{matrix} R \\ A \end{matrix}\right\}$ is evidently to choose A (R) if $x > (<) \left\{\begin{matrix} V_{2*}(t) \\ U(t) \end{matrix}\right\}$ we have

$$\psi_{R*}(x,t)=I(x>V_{2*}(t)) \text{ and } \psi_{A*}(x,t)=I(x>U(t)). \tag{1.6}$$

We obtain for the game

$$\lambda^{-1}V_1'(t)+V_1(t)=\int_0^1\Big[\bar{\varphi}(\bar{\psi}_{R*}V_1+\psi_{R*}U) + \varphi\{\bar{\psi}_{A*}x+\tfrac{1}{2}(x+U)\psi_{A*}\}\Big]dx, \tag{1.7}$$

$$\lambda^{-1}V_2'(t)+V_2(t)=\int_0^1\Big[\bar{\varphi}_*(\bar{\psi}_R V_2+\psi_R x) + \varphi_*\{\bar{\psi}_A U+\tfrac{1}{2}(x+U)\psi_A\}\Big]dx, \tag{1.8}$$

$$\lambda^{-1}V_{1*}'(t)+V_{1*}(t)=\int_0^1\Big[\bar{\varphi}_*(\bar{\psi}_{R*}V_{1*}+\psi_{R*}U)+\varphi_*\{\bar{\psi}_{A*}x+\tfrac{1}{2}(x+U)\psi_{A*}\}\Big]dx, \tag{1.9}$$

$$\lambda^{-1}V_{2*}'(t)+V_{2*}(t)=\int_0^1\Big[\bar{\varphi}_*(\bar{\psi}_{R*}V_{2*}+\psi_{R*}x)+\varphi_*\{\bar{\psi}_{A*}U+\tfrac{1}{2}(x+U)\psi_{A*}\}\Big]dx, \tag{1.10}$$

with the initial conditions $V_i(0)=V_{i*}(0)=0$, $i=1,2$. In the r.h.s.of Eq. (1.7)~(1.10) simplified notat ions for $\varphi(x,t)$, $V_i(t)$, etc. are used.

In the bilateral-move version of the game, symmetry for the players' role disappears. Also, since (x, t) is the common knowledge for both players, the disadvantage of information for the first-mover doesn't exist, and therefore the relation $V_{1*}(t) \leqq V_{2*}(t)$ is not assured.

Theorem 2. The solution to the bilateral-move version of the game given by (1.7)~(1.10) is as follows: Let $V_{1*}(t)$, and $V_{2*}(t)$ satisfy the simultaneous differential equation

(3.4) $$\lambda^{-1}(V_{1*}(t)-V_{2*}(t))' = -(V_{1*}-V_{2*})(1-V_{2*})+\frac{1}{2}(U-V_{2*})^2,$$

(3.5) $$\lambda^{-1}V_{2*}'(t) = -V_{2*}+\frac{1}{2}V_{2*}^2+\frac{1}{4}(1+2U-U^2)$$

with $V_{i*}(0) = 0,\ i=1,2$ (In the r.h.s. of (3.4)-(3.5) the argument t is omitted).

We have

(3.6) $$V_{2*}(t) \leqq V_{1*}(t), \quad (0 \leqq t \leqq T),$$

and

(3.7) $$\varphi_*(x,t)=\psi_{A*}(x,t)=I(U(t)<x\leqq 1),$$
$$\psi_{R*}(x,t)=I(V_{2*}(t)<x\leqq 1).$$

is in equilibrium with the payoffs $(V_{1*}(T),\ V_{2*}(T))$.

Corollary 2.1 The solution to (3.4)-(3.5) is ; Let $a=\frac{1}{2}(\sqrt{3}-1)\doteqdot 0.36603$. Then

(3.8) $$V_{2*}(t)=(1+a)U(t)-a+(q(t)+a^{-1})^{-1},$$

(3.9) $$V_{1*}(t)-V_{2*}(t)=\int_0^t \frac{1}{2}(U(\tau)-V_{2*}(\tau))^2 \exp\left\{-\int_\tau^t (1-V_{2*}(s))ds\right\} d\tau,$$

where

(3.10) $$q(t)=\left(1+\frac{5}{6}\sqrt{3}\right)\left\{\left(1+\frac{1}{2}\lambda t\right)^{1+\sqrt{3}}-1\right\}+\frac{\sqrt{3}}{12}\lambda t.$$

We see that $1-V_{2*}(T)=O((\lambda T)^{-1})$ as $\lambda T\to\infty$.

Corollary 2.2. $V_{2*}(t)$ is concave and increasing with $V_{2*}(0)=0$ and $V_{2*}(t)\to 1$, as $\lambda t\to\infty$.

Corollary 2.3 $V_{1*}(t)$ is increasing with $V_{1*}(0)=0$ and $V_{1*}(t)\to 1$, as $\lambda t\to\infty$

Corollaly 2.4. The time s until the game becomes one-person game has the defective p.d.f.

(3.8) $$g(s) = \lambda q(s)\exp\left[-\lambda\int_0^s q(w)dw\right],$$

where $q(s) = 1-V_{2*}(T-s)$.

The probability of getting zero by both players is $1-\int_0^T g(s)ds = \exp\left[-\lambda\int_0^T q(s)ds\right]$.

Theorem 2 shows that the first-mover behaves just as if he plays his one-person game without val, but, in reality, he gets less (i.e. $V_{1*}(t) < U(t)$) because of the fact that the second-over prevents him from getting U (t).

4. Remarks.

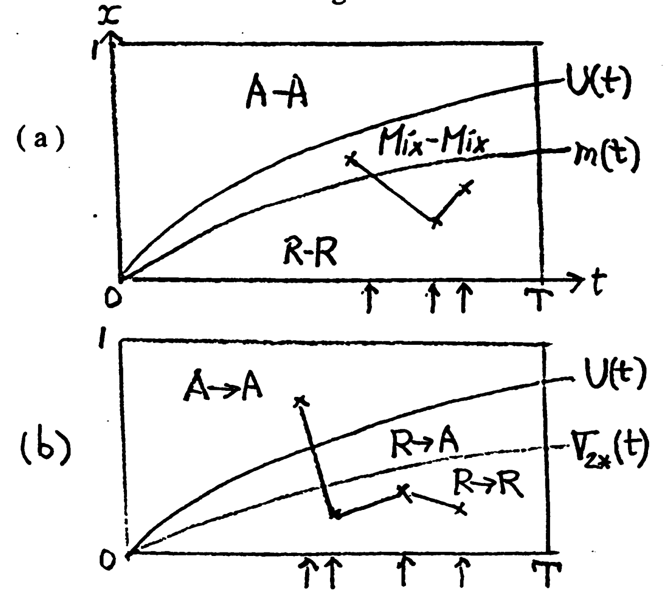
1°) Table 1 gives values of U(T), m(T), and $V_{2*}(T)$, computed from the equations in Corollaries 1.1 and 2.1 for some values of λT. Values of $V_{1*}(T)$ in Theorem 2 are not given for its troublesome computations.

Table 1. Values of U(T), m(T) and $V_{2*}(T)$.

λT	1	2	6	10	16	22
U(T)	1/3	1/2	3/4	5/6	0.8889	0.9167
m(T)	0.2252	0.3830	0.6710	0.7777	0.8509	0.8879
$V_{2*}(T)$	0.2170	0.3744	0.6677	0.7754	0.8492	0.8866

2°) A remarkable feature contained in this work is that in the simultaneous-move version of th game, the equilibrium strategy uses some randomization between R and A, whereas , in the bilateral-move version of the game, they employ non-randomized strategies only. See Figure 1.

Figure 1. Sampie paths of equilibrium plays when both players Remain in the game



In Theorem 1, Mix means a random choice $(R, A; \bar{\varphi}_*, \varphi_*)$

In Theorem 2, R→A, for example, means choice-pair of R by I first, and A followed by II.

3°) In the bilateral-move game, the first-mover stands at advantage than the second-mover. See (3.6) in Theorem 2. This is different from earlier works on single and multi-round poker, where the first mover stands at disadvantage, because players' hands x for I and y for II are private informations, and I leaks some information to his rival about his x by moving first. (See Garnaev [3] and Sakaguchi [7])

4°) It is interesting to investigate some open problems as follows :

(a) Solve the version where the game is played under winning probability (WP) maximization. Player wins if he gets the offer larger than that his rival gets. Each player aims to maximize his probability of winning. For the case where the offers arrive with the unit pace ⟵⟶ deterministically, see Sakaguchi [9].

(b) In (a), player wins if he gets the largest offer among those arrived and will arrive before time T. Each player aims to maximize his probability of winning.

(c) Solve no-information version of the game. Players do not know the size-distribution of Poisson arriving offers, but can only observe the relative rank among those arrived so far of the offer. Each player wants to minimize the expected absolute rank of the offer he gets. The best (worst) has the absolute rank 1 (n, if the n-th is the last offer arrived before time T).

REFERENCES

[1] M.T. Dixon, Equilibrium Points for Three Games based on the Poisson Process, J. Appl. Prob., 30 (1963), 627-638.

[2] E.G.Enns and E.Z. Ferenstein, A Competitive Best Choice Problem with Poisson Arrivals, J. Appl.Prob., 27 (1989), 333-342,

[3] A.Y.Garnaev, Value of Information in Optimal Stopping Games, Game Th. and Applications, Vol.V (2000) 55-64.

[4] K. Ravindran and E.G. Enns, A Two Poisson Stream Competitive Decision Problem, Math. Japonica 37 (1992), 723-731.

[5] 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.

[6] M. Sakaguchi, A Simplified Two- Person Mulistage Poker with Optional Stopping, Math. Japonica, 28 (1983), 287-303.

[7] M.Sakaguchi, Best-Choice Games with Random priority on a Two-Poisson Stream, Math. Japonica, 36 (1991), 731-745.

[8] M. Sakaguchi, Optimal stopping games—A review, *Math. Japonica* 42 (1995), 343-351.

[9] M. Sakaguchi, Optimal stopping games where players have weighted privilege, *Game Theory and Applications*, Vol. VI. (2001), 116-131.

[10] D. Ramsey and K. Szajowski, Three-person stopping game with players having privileges, *to appear* in *Theory of Prob. and its Appl.*

* Full paper is to appear in GTA vol VII (2002).

** 3-26-4 Midorigaoka, Toyonaka, Osaka 560-0002 , JAPAN
Fax +86-6856-2314