

УДК 519.833

ИГРА N ЛИЦ С ОПТИМАЛЬНОЙ ОСТАНОВКОЙ

А. А. Ивашко, Е. Е. Ивашко

*Институт прикладных математических исследований
Карельского научного центра РАН*

Рассматривается следующая игра n лиц с оптимальной остановкой. Каждый из n игроков независимо от остальных получает в качестве очков значение случайной величины, равномерно распределенной на $[0, b]$, $b \geq 1$. Игрок должен принять решение: остановиться после первого шага или продолжить и получить дополнительно значение следующей независимой случайной величины, которое прибавляется к полученным ранее очкам. Игроки не знают ни значений наблюдений, ни решений, принятых другими игроками. Побеждает тот игрок, чья сумма очков окажется наиболее близкой, но не превышающей 1. Если сумма очков каждого из игроков превысила 1, то побеждает игрок, получивший наименьшее число очков. Каждый игрок стремится максимизировать вероятность своего выигрыша. Найдены оптимальные стратегии игроков в данной задаче. Рассмотрено обобщение игры на случай разладки.

Ключевые слова: игры с оптимальной остановкой, пороговая стратегия, разладка.

A. A. Ivashko, E. E. Ivashko. N -PERSON OPTIMAL STOPPING GAME

The following n -person optimal stopping game is considered. Each of n players gets a random number from $[0, b]$, $b \geq 1$ as a score. A player has to decide whether to stop after the first draw or to continue to a second draw, in which case the value of a second random variable is added to their score. Players do not know the scores or the decisions of other players. The player with the largest total score not exceeding 1 wins. In the case the total scores of each player exceeds 1, the player with the lowest score closest to 1 wins. The objective of each player is to maximize the probability of one's win. Optimal strategies of the players are presented. The disorder version of the game is also investigated.

Key words: optimal stopping games, threshold strategy, disorder.

ВВЕДЕНИЕ

Рассматривается следующая игра n лиц с оптимальной остановкой. Игрок i независимо от остальных участников получает в качестве очков значение случайной величины Y_1^i , $i = 1, \dots, n$. Игрок должен принять решение: *остановиться* после первого шага или *продолжить* и получить дополнительно значение

случайной величины Y_2^i , $i = 1, \dots, n$, которое прибавляется к полученным ранее очкам. Y_1^i и Y_2^i — независимые равномерно распределенные на отрезке $[0, b]$, $b \geq 1$, случайные величины. Побеждает тот игрок, чья сумма очков окажется наиболее близкой, но не превышающей 1. Если сумма очков каждого из игроков превысила 1, то побеждает игрок, получивший

наименьшее число очков. Игроки не знают ни значений наблюдений, ни решений, принятых другими игроками. Каждый игрок стремится максимизировать вероятность своего выигрыша.

Данная задача является моделью популярного в США телевизионного игрового шоу «The Price is Right». В работе [Каунар, 2009] было представлено решение этой задачи при $b = 1$, основанное на перечислении всех вероятностей выигрыша игрока. Такой метод приводит к необходимости перебора большого числа вариантов и построению объемных формул при большом числе игроков. Как следствие, в указанной статье представлено решение только для случаев двух и трех игроков. В данной статье решение получено методом динамического программирования в общем виде для случая n игроков и при различных значениях параметра равномерного распределения наблюдаемых случайных величин. Также рассмотрено обобщение задачи на случай разладки, в котором распределение поступающих случайных величин может измениться в случайный момент времени.

Другие задачи с оптимальной остановкой были исследованы в работах [Сое, Butterworth, 1995; Tijms, 2007]. Метод динамического программирования для решения задач с оптимальной остановкой применялся в работах [Mazalov, 1996; Sakaguchi, 2005, 2007]. В работе [Mazalov, Ivashko, 2010] рассмотрена многошаговая игра с оптимальной остановкой.

РАВНОМЕРНОЕ РАСПРЕДЕЛЕНИЕ НАБЛЮДЕНИЙ НА ОТРЕЗКЕ $[0, 1]$

Рассмотрим случай $b = 1$, т. е. наблюдаемые игроками случайные величины независимы и равномерно распределены на отрезке $[0, 1]$.

Предположим, что каждый из игроков i , $i = 1, \dots, n$ для принятия решения об остановке/продолжении использует следующую однопороговую стратегию: остановиться на первом шаге, если полученное значение случайной величины Y_1^i больше или равно пороговому значению, и продолжить в противном случае. Пусть Игроки $1, \dots, n-1$ используют пороговую стратегию $0 < t < 1$, а Игрок n ищет свою наилучшую ответную стратегию $t < v < 1$.

Обозначим X_i — сумма очков, полученных игроком i ($i = 1, \dots, n-1$) в результате использования стратегии t . X_i равно Y_1^i при остановке на первом шаге или $Y_1^i + Y_2^i$ при продолжении.

Для вычисления ожидаемого выигрыша Игрока n рассмотрим вероятности следующих

событий:

$$P(X_i \leq y) = y - t + \int_0^t (y - x)dx, \quad t \leq y \leq 1;$$

$$P(X_i > 1) = \int_0^t xdx;$$

$$P_{y>1}(X_i > y) = \int_{y-1}^t (1 - (y - x))dx, \quad y > 1,$$

где $i = 1, \dots, n-1$. Так как игроки независимы друг от друга, то получим, что рассмотренные вероятности одинаковы для всех игроков.

Обозначим $H(v, t)$ ожидаемый выигрыш Игрока n при остановке на первом шаге в случае, когда общее число его очков равно v . Тогда

$$H(v, t) = \sum_{k=0}^{n-1} C_{n-1}^k \left(P(X_1 \leq v) \right)^{n-1-k} \left(P(X_1 > 1) \right)^k = (v - t + vt)^{n-1},$$

так как $P(X_1 \leq v) = P(X_2 \leq v) = \dots = P(X_{n-1} \leq v)$, и $P(X_1 > 1) = P(X_2 > 1) = \dots = P(X_{n-1} > 1)$.

$H'(v, t)$ — ожидаемый выигрыш Игрока n , если он продолжит игру, и число его очков на первом шаге равно v . Получим

$$H'(v, t) = \int_v^1 H(y, t)dy + \int_1^{v+1} (P_{y>1}(X_1 > y))^{n-1} dy = \frac{1 - (v - t + vt)^n}{n(t+1)} + \frac{t^{2n-1} - (t-v)^{2n-1}}{2^{n-1}(2n-1)},$$

т.к. $P_{y>1}(X_1 > y) = P_{y>1}(X_2 > y) = \dots = P_{y>1}(X_{n-1} > y)$.

Нетрудно показать, что для любого $0 < t < 1$ функция $H(v, t)$ возрастает, а функция $H'(v, t)$ убывает по v . Значение оптимального порога t таково, что ожидаемый выигрыш Игрока n при продолжении равен ожидаемому выигрышу при остановке, т. е. $H(v, t) = H'(v, t)$.

В силу симметрии задачи положим $v = t$. Окончательно получим, что значением оптимального порога t для принятия первого наблюдения является корень уравнения:

$$t^{2(n-1)} = \frac{1 - t^{2n}}{n(t+1)} + \frac{t^{2n-1}}{2^{n-1}(2n-1)}.$$

При $n = 2$ эта формула дает то же решение, что и представленное в работе [Каунар, 2009]: $2t^3 + 3t^2 + 3t - 3 = 0$.

Значения оптимальных порогов для различных значений n представлены в табл. 1.

Таблица 1. Оптимальные пороги для различных значений n

n	2	3	4	5	6	7	8	9	10	100	500
t	0,563	0,661	0,718	0,757	0,785	0,806	0,823	0,837	0,849	0,974	0,993

РАВНОМЕРНОЕ РАСПРЕДЕЛЕНИЕ НАБЛЮДЕНИЙ НА ОТРЕЗКЕ $[0, b]$, ГДЕ $b > 1$

Рассмотрим случай, когда наблюдаемые случайные величины распределены на отрезке $[0, b]$, где $b > 1$. Отметим, что изменение параметра распределения уменьшит значение оптимального порога t по сравнению с вариантом задачи, рассмотренной в предыдущем пункте. Это является следствием того, что при принятии решения о продолжении увеличивается вероятность превышения 1.

Действуя методом, использованным ранее, получаем, что выигрыши игроков $H(v, t)$ и $H'(v, t)$ при остановке и при продолжении, соответственно, имеют следующий вид:

$$\begin{aligned} H(v, t) &= \sum_{k=0}^{n-1} C_{n-1}^k (P(X_1 \leq v)^{n-k-1} (P(X_1 > 1))^k) \\ &= (P(X_1 \leq v) + P(X_1 > 1))^{n-1} \\ &= \left[\frac{bv - t + tv + b(b-1)}{b^2} \right]^{n-1}, \end{aligned}$$

так как $P(X_1 \leq v) = P(X_2 \leq v) = \dots = P(X_{n-1} \leq v)$, $P(X_1 > 1) = P(X_2 > 1) = \dots = P(X_{n-1} > 1)$,

и

$$\begin{aligned} H'(v, t) &= \frac{1}{b} \left[\int_v^1 H(y, t) dy \right. \\ &\quad \left. + \int_1^{v+b} (P_{y>1}(X_1 > y))^{n-1} dy \right] \\ &= \frac{1}{b} \left[\int_v^1 H(y, t) dy + \int_1^b (P_{b>y>1}(X_1 > y))^{n-1} dy \right. \\ &\quad \left. + \int_b^{v+b} (P_{y>b}(X_1 > y))^{n-1} dy \right] \\ &= \frac{1}{b} \left[\frac{(b^2)^n - (bv - t + tv + b(b-1))^n}{n(b+t)b^{2(n-1)}} \right. \\ &\quad \left. + \frac{(t(b-1+\frac{t}{2}) + b(b-1))^n - (\frac{t^2}{2})^n}{b^{2(n-1)}(t+b)n} + \frac{t^{2n-1} - (t-v)^{2n-1}}{2^{n-1}(2n-1)b^{2(n-1)}} \right], \end{aligned}$$

так как $P_{y>1}(X_1 > y) = P_{y>1}(X_2 > y) = \dots = P_{y>1}(X_{n-1} > y)$, $P_{b>y>1}(X_1 > y) = P_{b>y>1}(X_2 > y) = \dots = P_{b>y>1}(X_{n-1} > y)$ и $P_{y>b}(X_1 > y) = P_{y>b}(X_2 > y) = \dots = P_{y>b}(X_{n-1} > y)$.

Решая уравнение $H(v, t) = H'(v, t)$, а затем полагая $v = t$, получим:

$$\begin{aligned} [bt - t + t^2 + b(b-1)]^{n-1} &= \frac{1}{b} \left[\frac{b^{2n} - (bt - t + t^2 + b(b-1))^n}{n(b+t)} \right. \\ &\quad \left. + \frac{(t(b-1+\frac{t}{2}) + b(b-1))^n - (\frac{t^2}{2})^n}{(t+b)n} + \frac{t^{2n-1}}{2^{n-1}(2n-1)} \right]. \end{aligned}$$

На рис. 1 изображены значения оптимальных порогов $t(b)$ как функции от b при $n = 2$.

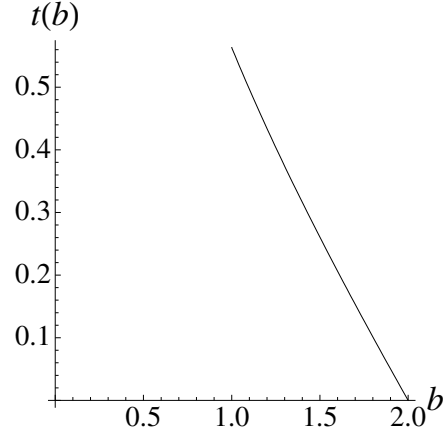


Рис. 1. Значения оптимальных порогов $t(b)$ при $n = 2$

Заметим, что оптимальный порог для $b \geq 2$ равен 0. Это означает, что игрок должен всегда останавливаться на первом шаге независимо от полученного количества очков.

КОМБИНИРОВАННЫЙ ВАРИАНТ

Интерес представляет следующий комбинированный вариант исходной задачи. Рассмотрим задачу, в которой распределение на первом шаге является равномерным на отрезке $[0, 1]$, а на втором — на отрезке $[0, b]$, $b > 1$. Тогда, учитывая различные распределения получаемых значений, ожидаемые выигрыши игроков примут вид:

$$\begin{aligned} H(v, t) &= \left[v - t + \frac{t}{b} \left(v - \frac{t}{2} \right) + \frac{t(b-1+\frac{t}{2})}{b} \right]^{n-1} \\ &= \left[v + \frac{tv}{b} - \frac{t}{b} \right]^{n-1}, \end{aligned}$$

и

$$\begin{aligned} H'(v, t) &= \frac{1}{b} \left[\frac{b^n - (vb + tv - t)^n}{nb^{n-1}(b+t)} \right. \\ &\quad \left. + \left(\frac{t}{b} \right)^{n-1} \frac{(b-1+\frac{t}{2})^n - (\frac{t}{2})^n}{n} + \frac{(t)^{2n-1} - (t-v)^{2n-1}}{2^{n-1}(2n-1)b^{2(n-1)}} \right]. \end{aligned}$$

Следовательно, значение оптимального порога будет удовлетворять уравнению

$$\begin{aligned} [tb + t^2 - t]^{n-1} &= \frac{1}{b} \left[\frac{b^n - (tb + t^2 - t)^n}{n(b+t)} \right. \\ &\quad \left. + t^{n-1} \frac{(b-1+\frac{t}{2})^n - (\frac{t}{2})^n}{n} + \frac{t^{2n-1}}{2^{n-1}(2n-1)} \right]. \end{aligned}$$

На рис. 2 представлены значения оптимальных порогов $t(b)$ для $n = 2$.

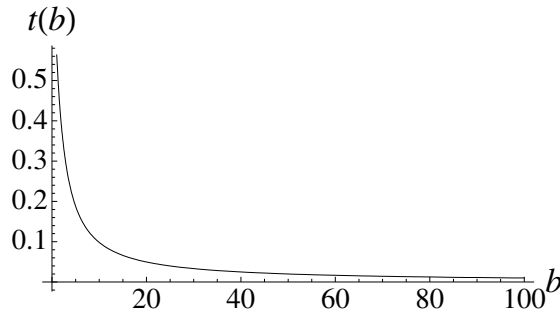


Рис. 2. Значения оптимальных порогов $t(b)$ при $n = 2$

При $b \rightarrow \infty$ значение оптимального порога стремится к нулю.

ВАРИАНТ ЗАДАЧИ С РАЗЛАДКОЙ

Рассмотрим обобщение исходной задачи на случай «разладки». Пусть изначально распределение обоих наблюдений является равномерным на отрезке $[0, 1]$. В случайный момент времени (перед первым или вторым шагом) одновременно для всех наблюдателей может произойти разладка, и распределение случайных величин изменится на равномерное на отрезке $[0, b]$, где $b > 1$. Момент разладки — это случайная величина, имеющая геометрическое распределение с параметром $1 - \alpha$ ($0 \leq \alpha \leq 1$). Игрокам известны параметры α и b , но они не обладают информацией, в какой момент произошла разладка. Будем искать пороговую стратегию, которая устанавливается игроками в начале игры и не меняется в дальнейшем.

Решение данной задачи строится на основе комбинации трех предыдущих игр. Ожидаемый выигрыш игрока зависит от того, на каком шаге произойдет разладка. С вероятностью α^2 распределение случайных величин не изменится ни перед первым шагом, ни перед вторым. В этом случае игрок, использующий стратегию t , получает выигрыш $t^{2(n-1)}$. С вероятностью $\alpha(1 - \alpha)$ разладка произойдет только перед вторым шагом, тогда игрок получит выигрыш $[tb + t^2 - t]^{n-1}$, соответствующий комбинированному варианту задачи. Наконец, с вероятностью $1 - \alpha$ разладка произойдет перед первым шагом, и тогда каждый игрок получит в качестве выигрыша величину $[bt - t + t^2 + b(b - 1)]^{n-1}$, которая равна выигрышу игрока в задаче с равномерным распределением обеих случайных величин на отрезке $[0, b]$. Аналогично можно провести рассуж-

дения для выигрыша при продолжении. Далее, приравнявая выигрыш при остановке и выигрыш при продолжении, получаем, что оптимальные пороги находятся из следующего уравнения:

$$\begin{aligned} & \alpha^2 t^{2(n-1)} + \alpha(1 - \alpha) [tb + t^2 - t]^{n-1} \\ & + (1 - \alpha) [bt - t + t^2 + b(b - 1)]^{n-1} \\ & = \alpha^2 \left[\frac{1 - t^{2n}}{n(t+1)} + \frac{t^{2n-1}}{2^{n-1}(2n-1)} \right] + \alpha(1 - \alpha) \frac{1}{b} \left[\frac{b^n - (tb + t^2 - t)^n}{n(b+t)} \right. \\ & \quad \left. + t^{n-1} \frac{(b-1 + \frac{t}{2})^n - (\frac{t}{2})^n}{n} + \frac{t^{2n-1}}{2^{n-1}(2n-1)} \right] \\ & \quad + (1 - \alpha) \frac{1}{b} \left[\frac{b^{2n} - (bt - t + t^2 + b(b-1))^n}{n(b+t)} \right. \\ & \quad \left. + \frac{(t(b-1 + \frac{t}{2}) + b(b-1))^n - (\frac{t^2}{2})^n}{(t+b)n} + \frac{t^{2n-1}}{2^{n-1}(2n-1)} \right] \end{aligned}$$

или

$$\begin{aligned} & \alpha^2 t^{2(n-1)} + \alpha(1 - \alpha) [tb + t^2 - t]^{n-1} \\ & + (1 - \alpha) [bt - t + t^2 + b(b - 1)]^{n-1} \\ & = \alpha^2 \frac{1 - t^{2n}}{n(t+1)} + \alpha(1 - \alpha) \frac{1}{bn} \left[\frac{b^n - (tb + t^2 - t)^n}{(b+t)} \right. \\ & \quad \left. + t^{n-1} \left((b-1 + \frac{t}{2})^n - (\frac{t}{2})^n \right) \right] \\ & \quad + (1 - \alpha) \frac{1}{b(t+b)n} \left[b^{2n} - (bt - t + t^2 + b(b-1))^n \right. \\ & \quad \left. + (t(b-1 + \frac{t}{2}) + b(b-1))^n - (\frac{t^2}{2})^n \right] \\ & \quad + \left(\alpha^2 - \frac{\alpha^2}{b} + \frac{1}{b} \right) \frac{t^{2n-1}}{2^{n-1}(2n-1)}. \end{aligned}$$

Значения оптимальных порогов для различных n , $\alpha = 0, 1$, $b = 1, 5$ представлены в табл. 2.

Таблица 2. Значения оптимальных порогов t для различных n ($\alpha = 0, 1$, $b = 1, 5$)

n	2	3	4	5	10	20	50
t	0,277	0,424	0,516	0,579	0,735	0,839	0,920

ЗАКЛЮЧЕНИЕ

В представленной статье рассмотрена игровая задача оптимальной остановки для n игроков. Для данной задачи рассмотрены три варианта с различными параметрами равномерного распределения наблюдаемых игроками случайных величин. Представлено обобщение задачи на случай разладки. Для иллюстрации полученных решений приведены результаты численного моделирования значений оптимальных порогов.

Авторы выражают благодарность проф. В. В. Мазалову за его помощь в постановке задачи и обсуждении полученных результатов. Исследования поддержаны грантом РФФИ, проект 10-01-00089-а и Отделением математических наук РАН.

ЛИТЕРАТУРА

Coe P. R., Butterworth W. Optimal stopping in «The Showcase Showdown» // The American Statistician. 1995. N. 49. P. 271–275.

Kaynar B. Optimal stopping in a stochastic game // Probability in the Engineering and Information Sciences. 2009. N. 23. P. 51–60.

Mazalov V. V. A Game related to optimal stopping of two sequences of independent random variables having different distributions // Mathematica Japonica. 1996. 43. N. 1. P. 121–128.

Sakaguchi M. Equilibrium in two-player games of «Showcase Showdown» // Scientiae Mathematicae Japonicae. 2005. N. 61. P. 145–151.

Sakaguchi M. Players' information in two-player games of «Score Showdown» // Game Theory and Applications. 2007. N. 11 / Edit. Petrosjan and Mazalov N.Y. Nova Science Publ. P. 111–124.

Tijms H. C. Understanding probability, 2nd ed. Cambridge: Cambridge Univ. Press. 2007.

Mazalov V. V., Ivashko A. A. Equilibrium in n -person game of Showcase-Showdown // Probability in the Engineering and Informational Sciences, Cambridge Univ. Press, 2010. N. 24. P. 397–403.

СВЕДЕНИЯ ОБ АВТОРАХ:

Ивашко Анна Антоновна

научный сотрудник
Институт прикладных математических исследований
КарНЦ РАН
ул. Пушкинская, 11, Петрозаводск, Республика
Карелия, Россия, 185910
эл. почта: aivashko@krc.karelia.ru
тел.: (8142) 766312

Ивашко Евгений Евгеньевич

научный сотрудник
Институт прикладных математических исследований
КарНЦ РАН
ул. Пушкинская, 11, Петрозаводск, Республика Каре-
лия, Россия, 185910
эл. почта: ivashko@krc.karelia.ru
тел.: (8142) 766312

Ivashko, Anna

Institute of Applied Mathematical Research, Karelian
Research Centre, Russian Academy of Science
11 Pushkinskaya St., 185910 Petrozavodsk,
Karelia, Russia
e-mail: aivashko@krc.karelia.ru
tel.: (8142) 766312

Ivashko, Evgeny

Institute of Applied Mathematical Research, Karelian
Research Centre, Russian Academy of Science
11 Pushkinskaya St., 185910 Petrozavodsk, Karelia,
Russia
e-mail: ivashko@krc.karelia.ru
tel.: (8142) 766312