О НЕКОТОРЫХ ЗАДАЧАХ ОПТИМАЛЬНОГО ВЫБОРА РЕКОРДНЫХ ВЕЛИЧИН

Бельков Игорь Владимирович

Санкт-Петербургский государственный университет, Санкт-Петербург

Аннотация: В работе рассматриваются оптимизационные задачи, связанные с рекордными величинами. Имеется n независимых случайных величин, имеющих равномерное распределение на интервале [0, 1]. Получая последовательно наблюдаемые значения этих величин, нужно в какой-то момент остановиться на одной из них, приняв ее как начальную для отсчета верхних рекордных величин. Показано, как надо сделать правильный выбор начальной точки отсчета рекордов, чтобы максимизировать математическое ожидание суммы значений нижних рекордных величин или общей суммы верхних и нижних рекордных величин, полученных в результате такой процедуры. Библ. — 9 назв.

Ключевые слова: рекордные моменты, рекордные величины, суммы рекордных величин, среднее число рекордов, равномерное распределение, задача оптимального выбора.

Теория рекордов составляет важную часть теории вероятностей и математической статистики.

Пусть присутствует последовательность случайных величин (с. в.) X_1, X_2, … с непрерывной функцией распределения $F(x)$. Определим для этой последовательности верхние рекордные моменты и величины:

$L(1)=1$, $X(1)= X_1$, $L(\mathrm{n})=\min\{j{:}X_j>X(n-1)\}$, $X(n)=X_{L(n)}$, $n=2, 3,\ldots$

Аналогично определяются нижние рекордные моменты и величины:

$l(1)=1$, $x(1)=X_1$, $l(\mathrm{n})=\min\{j{:}X_j>X(n-1)\}$, $x(n)=X_{l(\mathrm{n})}$, $n=2, 3,\ldots$

Существует класс задач, посвященных оптимальному выбору величин.

Различные результаты можно найти в монографиях [5, 7, 8].

Одной из таких задач является классическая задача оптимального выбора (проблема секретаря) [3, 4, 9].

Пусть существует набор из n случайных величин X_1, X_2, …, X_n и соответствующее случайное число $M=M(n)$ верхних рекордов:

$X(1)=X_1<X(2)<\ldots<X(M)=\max\{X_1, X_2, \ldots, X_m\}$.

Последовательно получаем наблюдаемые значения исходных величин.

В статье [6] была рассмотрена задача об увеличении числа рекордов в последовательности случайных величин X_1, X_2, …, X_n при применении процедуры, использованной а задаче о разборчивой невесте.

В статье [1] была рассмотрена задача о максимизации математического ожидания суммы верхних рекордных величин в случае, когда исходные величины U_1, U_2, …, U_n имеют равномерное распределение U([0, 1]). Приведены таблицы значений величин x_n, V_n, $T(n)$ и $d(n)$, где x_n — значения корней уравнения $T(n-1)=T_n(x)$; V_n — максимизированное математическое ожидание суммы рекордных величин в случае, если отсчет начинать не с первой, а с другой величины; $T(n)=\frac{1}{2}+\ldots+\frac{1}{n}+\frac{1}{n+1}$ — математическое ожидание суммы рекордных величин в наборе U_1, U_2, …, U_n; $d(n)=V_n-T(n)$; $T_n(x)=$ $\frac{1}{2}+\ldots+\frac{1}{n}+x-\frac{x^2}{2}-\frac{x^3}{3}-\ldots-\frac{x^n}{n}$ — математическое ожидание в наборе x, U_2, …, U_n.

Рассмотрим теперь аналогичную задачу для нижних рекордных величин.

Задача 1. Пусть величины X_1, X_2, … имеют равномерное распределение U([0, 1]).

Пусть $L=L(n)$ – число нижних рекордов среди величин U_1, U_2, …, U_n. Эти рекорды будем обозначать $u(1)=U_1>u(2)>\ldots>u(L)$. Обозначим сумму этих рекордов через W_n и сумму их математических ожиданий через $B(n)$.

Из леммы в статье [6] следует, что

$$B(n)=\mathrm{E}W_n = \sum_{k=1}^{n} \int_0^1 u(1-u)^{k-1}du = n/(n+1)=1-1/(n+1), \quad n=1,2,\ldots .$$

Рассмотрим также суммы всех рекордных значений (как верхних, так и нижних) среди величин U_1, U_2, …, U_n. Первую величину будем считать за один рекорд (хотя он является и верхним, и нижним).

Рассмотрим математическое ожидание этой суммы:

$$A(n)=T(n)+B(n)-\frac{1}{2}=\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}+\frac{1}{n+1}+1-\frac{1}{n+1}-\frac{1}{2}=\frac{1}{2}+(\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}).$$

Пусть $B_n(x)$ и $A_n(x)$ обозначают соответственно математическое ожидание суммы нижних рекордов и математическое ожидание суммы всех (верхних и нижних) рекордных значений в наборе $x,\ U_2,\ldots,U_n$, причем в выражении для $A_n(x)$ величина x входит один раз.

Справедливы следующие равенства:

Для любого n=2,3,…

$$B_n(x)=x+\int_0^x v\,dv+\int_0^x v(1-v)\,dv+\ldots+\int_0^x v(1-v)^{n-2}dv=2x-\frac{1}{n}+\frac{(1-x)^n}{n},$$

$$A_n(x)=2x+1/2+1/3+\ldots+1/(n-1)+(1-x)^n/n-x^2/2-x^3/3-\ldots-x^n/n,\ n=3,4,\ldots,$$

и $$A_2(x)=2x+(1-x)^2/2-x^2/2.$$

Чтобы максимизировать математическое ожидание суммы всех нижних рекордов путем выбора начальной точки отсчета, необходимо сравнить при каждом n=2,3,… и любом x из интервала (0; 1) величины $B_n(x)$ и $B(n-1)$.

Для этого необходимо найти для каждого n корни уравнения

$B_n(x)=B(n-1)$, или $2x-1/n+(1-x)^n/n=1-1/n$, что приводит к равенствам

$$2x+(1-x)^n/n=1,\ n=2,3,\ldots \qquad (1)$$

В таблице 1 даны значения корней уравнений (1), лежащие внутри интервала (0; 1), с точностью до 4, 6 или 12 знаков после запятой.

Таблица 1

Корни уравнений в задаче 1

n	x_n
2	0,4142
3	0,4760

n	x_n
4	0,4961
5	0,4967
6	0,4986
7	0,4994
8	0,499755
9	0,499891
10	0,499951
11	0,499978
12	0,499990
20	0,499999976158
30	0,499999999984

Предельное значение корня при $n \to \infty$ равно 0,5.

Таким образом, если $x \geq x_n$, то значение $U_1 = x$ следует брать как первый нижний рекорд. Если $x < x_n$, то отбрасываем значение x и уже рассматриваем серию $U_2, \ldots, U_n$. И так далее.

Пусть W_n обозначает полученное при такой процедуре математическое ожидание суммарного числа нижних рекордов в наборе $X_1, X_2, \ldots, X_n$.

Отметим, что $W_1 = \frac{1}{2}$.

Справедливы рекуррентные соотношения:

$$W_n = x_n W_{n-1} + \int_{x_n}^{1} B_n(u)\,du = x_n W_{n-1} + 1 - x_n^2 - \frac{1}{n} + \frac{x_n}{n} + \frac{(1-x_n)^{n+1}}{n(n+1)}.$$

В таблице 2 даны значения величин W_n, $B(n)$ и $d_1(n) = W(n) - B(n)$ с точностью до 4 знаков после запятой.

Таблица 2

Значения некоторых величин в задаче 1

n	W_n	$B(n)$	$d_1(n)$
2	0,7761	0,6667	0,1095
3	0,9745	0,7500	0,2245
4	1,1130	0,8000	0,3130
5	1,2060	0,8333	0,3727
6	1,2694	0,8571	0,4122
7	1,3131	0,8750	0,4381
8	1,3440	0,8889	0,4551
9	1,3664	0,9000	0,4664

n	W_n	$B(n)$	$d_1(n)$
10	1,3832	0,9091	0,4741
20	1,4470	0,9524	0,4946
30	1,4564	0,9677	0,4977

В пределе получаем 1,5 для W_n и 0,5 для $d_1(n)$.

Рассмотрим теперь другую задачу.

Задача 2. Максимизация математического ожидания суммы всех рекордов (верхних и нижних) в наборе $U_1, U_2, \ldots, U_n$. Как и в задаче 1, величина U_1 считается за один рекорд.

Чтобы максимизировать математическое ожидание суммы рекордов (верхних и нижних) путем выбора начальной точки отсчета, необходимо сравнить при каждом n=2,3,… и любом x из интервала (0; 1) величины $A_n(x)$ и $A(n-1)$.

Нужно найти для каждого n корни уравнений

$$A_n(x)=A(n-1),$$

или

$$2x+1/2+1/3+\ldots+1/(n-1)+(1-x)^n/n-x^2/2-x^3/3-\ldots-x^n/n = ½+(1/2+1/3+\ldots+1/(n-1)),$$

что сводится к соотношениям

$$2x+(1-x)^n/n - x^2/2 - x^3/3-\ldots- x^n/n=½. \quad (2)$$

В таблице 3 даны корни на (0; 1) уравнений (2) с точностью до 4 знаков после запятой.

Таблица 3

Корни уравнений в задаче 2

n	$(x_n)_1$	$(x_n)_2$
2	0	
3	0,1569	

n	$(x_n)_1$	$(x_n)_2$
4	0,2166	
5	0,2427	
6	0,2556	
7	0,2625	0,9752
8	0,2665	0,9498
9	0,2688	0,9330
10	0,2703	0,9213
20	0,2728	0,8884

Заметим, что при $n \geq 7$ получаем два корня. Тогда величина $A_n(x) - A(n-1)$ будет положительна, если x лежит между корнями, и отрицательна в противном случае.

Если n=3, 4, 5, 6, то первое наблюдаемое значение, если оно больше x_n, выбираем в качестве рекордного; если же оно меньше x_n, то его отбрасываем.

Если $n \geq 7$, то в случае $(x_n)_1 \leq x \leq (x_n)_2$ величину x можно брать в качестве первого рекорда; иначе ее отбрасываем и переходим к следующей с. в.

Если $n \to \infty$, то в пределе уравнение (2) превращается в $3x+\ln(1-x)=1/2$.

Его корни: x_1=0,2728994620…; x_2=0,883613816…

Пусть Z_n — максимизированное математическое ожидание, полученное при такой процедуре. Справедливо рекуррентное соотношение:

$$Z_n = \left((x_n)_2 - (x_n)_1\right) Z_{n-1} + \int_{(x_n)_1}^{(x_n)_2} A_n(u)\,du = \left((x_n)_2 - (x_n)_1\right) Z_{n-1} + \Big(x^2 + \frac{x}{2} + \ldots + \frac{x}{n-1} - \frac{(1-x)^{n+1}}{n(n+1)} - \frac{x^3}{2\cdot 3} - \frac{x^4}{3\cdot 4} - \ldots - \frac{x^n}{n(n}$$

При n от 2 до 6 вместо $(x_n)_2$ следует брать 1, при этом $Z_2=1$.

Приводим значения Z_n, $A(n)$ и $d_2(n)=Z_n-A(n)$ с точностью до 4 знаков после запятой.

Таблица 4

Значения некоторых величин в задаче 2

n	Z_n	$A(n)$	$d_2(n)$
2	1,0000	1,0000	0
3	2,0328	1,3333	0,6994
4	2,9149	1,5833	1,3316
5	3,6445	1,7833	1,8612
6	4,2520	1,9500	2,3020
7	4,6231	2,0929	2,5302

8	4,7724	2,2179	2,5546
9	4,8199	2,3290	2,4909
10	4,8276	2,4290	2,3986

Аналогичную задачу можно рассматривать для максимизации разностей математических ожиданий сумм верхних и нижних рекордов. Приводим уравнения для соответствующих корней и максимизированных математических ожиданий:

$$\sum_{k=1}^{n}\left(x^k/k\right) \quad +(1\text{-}x)^n/n=1,$$

$$Y_n=x_n Y_{n-1}+\int_0^{x_n}\left(T_n(u)-B_n(u)\right)du=x_n Y_{n-1}+x_n\left(1+\frac{1}{2}+\ldots+\frac{1}{n}\right)+\frac{x_n^2}{2}-\frac{x_n^3}{2\cdot 3}-\frac{x_n^4}{3\cdot 4}-\ldots-\frac{x_n^{n+1}}{n(n+1)}-x_n^2+\frac{x_n}{n}+\frac{(1-x}{n+1}$$

При этом Y_1=0.

Предельное уравнение в этой задаче таково: $-\ln(1-x)=1$, откуда $x=1-\frac{1}{e}$.

Также подобную задачу можно рассматривать не для суммы, а для суммарного числа верхних и нижних рекордов. Соответствующие результаты даны в статье [2].

В данной статье исследованы задачи оптимального выбора для нижних рекордов и для сумм и разностей верхних и нижних рекордов. Во всех случаях получаем, что путем надлежащего выбора начальной точки отсчета можно увеличить математическое ожидание суммы.

Список литературы

1. Бельков И. В., Невзоров В. Б. Об одной задаче оптимального выбора рекордных величин. // Записки научных семинаров ПОМИ. — 2017. Т. 466. — С. 30-37.

2. Бельков И. В., Невзоров В. Б. Об одной проблеме оптимального выбора рекордных величин. // Вестник СПбГУ, серия 1. — 2018. Т. 5 (63), вып. 2. — С. 179–188.

3. Гусейн-Заде С. М. Разборчивая невеста. — М.: МЦНМО, 2003.

4. Дынкин Е. Б. Оптимальный выбор момента остановки марковского процесса. // ДАН СССР. — 1963. Т. 150, №2. С. 238–240.
5. В. Б. Невзоров. Рекорды. Математическая теория. — М.: Фазис, 2000.
6. В. Б. Невзоров, С. А. Товмасян. О максимальном значении среднего числа рекордов. // Вестник СПбГУ, серия 1. — 2014. Т. 1 (59), вып. 2. С. 196–200.
7. Ahsanullah M., Nevzorov V. B.. Records via probability theory. — Atlantis Press, 2015.
8. Arnold B. C., Balakrishnan N., Nagaraja H. N. Records. — NY: John Wiley & Sons, 1998.
9. Gardner M. Mathematical Games. A fifth collection of “brain-teasers”. // Sci. Amer. — 1960. Vol. 202, No. 2. P. 150–154.