Санкт-Петербургский государственный университет

Математика

Теория вероятностей и математическая статистика

Фоминых Сергей Сергеевич

Дискретные аналоги рекордов с ограничениями

Выпускная квалификационная работа

Научный руководитель:

д. ф.-м. н., профессор Невзоров В.Б.

Рецензент:

д. ф.-м. н., профессор Розовский Л.В.

Санкт-Петербург

2017

SAINT-PETERSBURG STATE UNIVERSITY

Mathematics

Probability Theory and Mathematical Statistics

Sergey Fominykh

Discrete analogues of records with restrictions

Graduation Project

Scientific supervisor:

Doctor of Physics and Mathematics,

Professor Valery Nevzorov

Reviewer:

Doctor of Physics and Mathematics,

Professor Leonid Rozovsky

Saint-Petersburg

2017

Содержание

1 Введение

1.1 Общие сведения о теории рекордов

Рассмотрим случайные величины $X_1, X_2, \ldots$ и определим рекордные моменты $L(n)$ и рекордные величины $X(n), n = 1, 2, \ldots$ следующим образом: первый рекордный момент $L(1)$ всегда равен 1, а последующие определяются следующим соотношением:

$$L(n) = \min\{j > L(n-1) : X_j > X(n-1)\},$$
$$X(n) = X_{L(n)} = \max\{X_1, X_2, \ldots, X_{L(n)}\}, \quad n = 2, 3, \ldots.$$

Эти соотношения задают верхние рекордные моменты и верхние рекордные величины. Соответственно, последовательности случайных величин, определяемые равенствами

$$l(1) = 1, \quad x(1) = X_1, \quad l(n) = \min\{j > l(n-1) : X_j < x(n-1)\},$$
$$x(n) = X_{L(n)} = \min\{X_1, X_2, \ldots, X_{L(n)}\}, \quad n = 2, 3, \ldots,$$

представляют собой нижние рекордные моменты $l(n), n = 1, 2, \ldots$ и нижние рекордные величины $x(n), n = 1, 2, \ldots$.

Рассмотрение величин $Y_1 = -X_1, Y_2 = -X_2, \ldots$ вместо исходных X-ов позволяет перейти от верхних рекордов к нижним, поэтому достаточно ограничиться только исследованием верхних рекордов.

В теории рекордов обычно по отдельности рассмативают два случая: случай, при котором исходные случайные величины имеют непрерывное распределение, и случай, когда они имеют дискретное распределение. В дискретном случае, чтобы с вероятностью единица были определены все величины $X(n), n = 1, 2, \ldots$, необходимо, чтобы у исходного распределения не было последней точки роста, т.е. такой точки a, что

$$\mathsf{P}\{X < a\} < \mathsf{P}\{X \leqslant a\} = 1.$$

Для дискретных распределений рассматривают слабые рекорды: в этом случае повторение предыдущего рекордного значения засчитывается как новый рекорд.

$$L_\omega(1) = 1$$
$$L_\omega(n) = \min\{j > L_\omega(n-1) : X_j \geqslant X(n-1)\},$$
$$X_\omega(n) = X_{L_\omega(n)} = \max\{X_1, X_2, \ldots, X_{L_\omega(n)}\}, \quad n = 2, 3, \ldots.$$

Примеры слабых рекордов могут дать, например, те виды спорта (стрельба, легкая атлетика), где спортсмен, повторивший рекордное достижение, также объявляется рекордсменом. Такая классическая рекордная модель и ее обобщения, в которых, как правило, рассматриваются независимые одинаково распределенные случайные величины $X_1, X_2, \ldots$, хорошо изучены [1].

Менее изученными остаются ситуации, в которых исходные случайные величины могут иметь разные распределения, например в F^α-схеме, где функции распределения $F_k(x) = \mathsf{P}\{X_k < x\}$ связаны соотношениями

$$F_k(x) = F^{\alpha(k)}(x), \quad k = 1, 2, \ldots,$$

где $\alpha(1), \alpha(2), \ldots$ — произвольные положительные константы, а $F(x)$ — некоторая функция распределения. Также рассматриваются рекордные схемы, в которых исходные случайные величины независимы и одинаково распределены, а рекордные величины определяются "неклассически" , например, рекорды с δ-превышением, введенные в [2], или рекорды с ограничениями, рассмотренные в [3].

1.2 Рекорды с ограничением в случае непрерывных распределений

Рассмотрим первую модель рекордов с ограничениями, в которой мы будем игнорировать все наблюдения, которые превышают последнее рекордное значение более чем на некоторую фиксированную положительную константу C. Пусть $X_1, X_2, \ldots$ — последовательность независимых одинаково распределенных величин, а C — фиксированная положительная константа. В качестве первой рекордной величины и первого рекордного момента возьмем $X(1) = X_1$ и $L(1) = 1$. Последующие рекордные величины и моменты определим как

$$L(n) = \min\{j > L(n-1) : X(n-1) < X_j \leqslant X(n-1) + C\}, \tag{1}$$

$$X(n) = X_{L(n)} \quad n = 2, 3, \ldots.$$

Если снова обратиться к результатам спортивных соревнований, то слишком существенные превышения предыдущих рекордов могут вызвать опасения, что данный результат достигнут нечестными путями. Например, могли быть использованы нестандартные спортивные снаряды или спортсменами употреблялись запрещенные препараты. В статье [3] рассмотрена модель рекордов с ограничениями в случае, когда исходные случайные величины имеют непрерывную функцию распределения $F(x)$ и плотность $p(x)$. Были получены следующие результаты:

- Выражение для $p_n(x_n \mid x_1, x_2, \ldots, x_{n-1})$ — плотности распределения рекордной величины $X(n)$ при условии, что зафиксированы значения

  $$X(1) = x_1, \quad X(2) = x_2, \ldots, \quad X(n-1) = x_{n-1},$$

  где $0 < x_j - x_{j-1} \leqslant C, j = 2, 3, \ldots, n-1$:

  $$p_n(x_n \mid x_1, x_2, \ldots, x_{n-1}) = \frac{p(x_n)}{F(x_{n-1} + C) - F(x_{n-1})},$$

  если $x_{n-1} < x_n \leqslant x_{n-1} + C$, и $p_n(x_n \mid x_1, x_2, \ldots, x_{n-1}) = 0$ иначе.

- Выражение для совместной плотности распределения $p_n(x_1, x_2, \ldots, x_n)$ рекордных величин $X(1), X(2), \ldots, X(n)$:

  $$p_n(x_1, x_2, \ldots, x_n) =$$
  $$p(x_1)\frac{p(x_2)}{F(x_1 + C) - F(x_1)} \cdots \frac{p(x_n)}{F(x_{n-1} + C) - F(x_{n-1})},$$

  если $x_1 < x_2 \leqslant x_1 + C, \quad x_2 < x_3 \leqslant x_2 + C, \ldots \quad x_{n-1} < x_n \leqslant x_{n-1} + C$, и $p_n(x_1, x_2, \ldots, x_n) = 0$ иначе.

Также, если исходные случайные величины ограничены снизу некоторой константой $a > -\infty$, то

$$p_n(x_1, x_2, \ldots, x_n) =$$
$$\frac{p(x_1)}{F(a+C)} \frac{p(x_2)}{F(x_1+C) - F(x_1)} \cdots \frac{p(x_n)}{F(x_{n-1}+C) - F(x_{n-1})},$$

если $a < x_1 \leqslant a + C, \quad x_1 < x_2 \leqslant x_1 + C, \ldots \quad x_{n-1} < x_n \leqslant x_{n-1} + C$, и $p_n(x_1, x_2, \ldots, x_n) = 0$ иначе.

- Если исходные случайные величины имеют стандартное $\mathsf{E}(1)$ экспоненциальное распределение, то полученные формулы переписываются в виде:

$$p_n(x_1, x_2, \ldots, x_n) = \frac{e^{-x_n}}{(1 - e^{-C})^n},$$

если $0 < x_1 \leqslant C, \quad x_1 < x_2 \leqslant x_1 + C, \ldots \quad x_{n-1} < x_n \leqslant x_{n-1} + C$, и $p_n(x_1, x_2, \ldots, x_n) = 0$ иначе.

- Одинаково распределенными являются случайные векторы

$$\{X(1), X(2), \ldots, X(n)\}$$

и

$$\{W_1, W_1 + W_2, \ldots, W_1 + W_2 + \ldots + W_n\},$$

где независимые случайные величины $W_1, W_2, \ldots, W_n$ имеют плотность распределения

$$g(x) = \frac{e^{-x}}{1 - e^{-C}},$$

если $0 \leqslant x \leqslant C$, и $g(x) = 0$ иначе.

Также была рассмотрена вторая модель рекордов с ограничениями: снова зафиксируем положительную константу C, но в этом случае значение, превышающее последний рекорд $X(n)$ больше чем на C, не теряется, и новой рекордной величиной объявляется $X(n+1) = X(n) + C$. Если же наблюдение попало в интервал $(X(n), X(n) + C]$, то оно и становится новой рекордной величиной X(n+1).
Для нее, в случае экспоненциально распределенных исходных случайных величин, было получено, что одинаково распределенными являются векторы

$$\{X(1), X(2), \ldots, X(n)\}$$

и

$$\{V_1, V_1 + V_2, \ldots, V_1 + V_2 + \ldots + V_n\},$$

где независимые случайные величины $V_1, V_2, \ldots, V_n$ имеют функцию распределения

$$\mathsf{P}\{V(k) < x\} = \begin{cases} 1 - e^{-x} & , \quad 0 \leqslant x \leqslant C \\ 1 & , \quad x > C. \end{cases}$$

В этой работе мы получим аналогичные результаты для рекордных схем с ограничениями в случае дискретных распределений.

1.3 Классические дискретные рекорды

Приведем несколько результатов для классических дискретных рекордов из [1], для того чтобы далее провести аналогию с полученными результатами.

Теорема 1 *Пусть* $X_1, X_2, \ldots$ *— независимые случайные величины, принимающие значения* $1, 2, \ldots$ *с вероятностями*

$$p_k = \mathsf{P}\{X_j = k\} = (1-p)p^{k-1}, \quad j = 1, 2, \ldots, \quad k = 1, 2, \ldots.$$

Тогда случайные величины $X(1), X(2)-X(1), X(3)-X(2), \ldots$ *независимы, и имеют то же геометрическое распределение, что и исходные случайные величины* $X_1, X_2, \ldots$.

Теорема 2 *В условиях теоремы* 1 *для любого* $n = 1, 2, \ldots$ *справедливо равенство*

$$X(n) \stackrel{d}{=} X_1 + X_2 + \ldots + X_n.$$

Поскольку случайные величины $X_1, X_2, \ldots$ *независимы, одинаково распределены и имеют математические ожидания* $a = 1/(1-p)$ *и дисперсии* $\sigma^2 = p/(1-p)^2$*, мы получаем что*

$$\frac{X(n) - na}{\sigma n^{1/2}} = \frac{(1-p)X(n) - n}{(np)^{1/2}} \underset{n\to\infty}{\longrightarrow} \mathsf{N}(0,1),$$

где $\mathsf{N}(0,1)$ *— стандартное нормальное распределение.*

1.4 Рекорды с $\delta-$превышением и случайные блуждания

В 1996 г. в работе [2] были введены так называемые рекорды с δ-превышением. В этой схеме для объявления очередного наблюдения рекордным требовалось, чтобы оно превышало предыдущее рекордное достижение более чем на некоторое значение δ. Для последовательности независимых случайных величин $X_1, X_2, \ldots,$ имеющих одинаковую функцию распределения, рекордные моменты и рекордные величины определяются следующим образом:

$$L(n) = \min\{j > L(n-1) : X_j > X(n-1) + \delta\},$$

$$X(n) = X_{L(n)} \quad n = 2, 3, \ldots.$$

Рекорды с $\delta-$превышением возникают, например, в счетчиках частиц второго типа. Счетчик выставляют в радиусе действия источника излучения частиц (например, радиоактивный материал). Частицы попадают в счетчик, но в силу некоторых физических ограничений конструкции прибора не все частицы регистрируются. В счетчиках второго типа после прибытия частицы (зарегистрированного или нет) новые частицы перестают регистрироваться в период времени δ, причем новые частицы, прибывающие в этот период восстановления счетчика, продливают его. Таким образом, только те частицы, которые не попали в период восстановления, регистрируются.

В работе [4] показано, что количество зарегистрированных частиц до момента t является количеством рекордов с $\delta-$превышением для величин с некоторой функцией распределения. Рекорды с $\delta-$превышением рассмотрены в [5] [6] в случаях $\delta \geqslant 0$ и $\delta \leqslant 0$ и в [7] в дискретном случае.

В работах [8] [9] [10] рассматриваются рекордные величины в случайных блужданиях. В этой работе мы постараемся обобщить некоторые результаты для дискретных случайных блужданий на случай рекордов с превышениями.

2 Основные результаты

Мы хотим рассмотреть рекордные схемы с ограничениями, в случае когда исходные случайные величины имеют дискретное распределение. Ниже приведем результаты, полученные в работе в этом случае.

2.1 Рекорды с ограничением I

Рассмотрим первую модель рекордов с ограничениями, в которой мы будем игнорировать все наблюдения, которые превышают последнее рекордное значение более чем на некоторую положительную константу C. Пусть $X_1, X_2, \dots$ — последовательность независимых одинаково распределенных величин, а C_k — положительная постоянная, зависящая от номера рекорда. В качестве первой рекордной величины и первого рекордного момента возьмем $X(1) = X_1$ и $L(1) = 1$. Последующие рекордные величины и моменты определим как

$$L(n) = \min\{j > L(n-1) : X(n-1) < X_j \leqslant X(n-1) + C_{n-1}\}, \tag{2}$$

$$X(n) = X_{L(n)} \quad n = 2, 3, \dots.$$

Для начала докажем следующую лемму:

Лемма 1 *При любом $n = 1, 2, \dots$ случайные величины*

$$Y_1 = X_{L(n)+1}, \quad Y_2 = X_{L(n)+2}, \quad \dots$$

не зависят от $X(1), X(2), \dots, X(n)$, независимы между собой и имеют общую функцию распределения F.

Доказательство Заметим, что события $C_{n,m} = \{L(n) = m\}$ при каждом m определяются только лишь случайными величинами $X_1, X_2, \dots, X_m$ и не зависят от $X_{m+1}, X_{m+2}, \dots$. Вероятность $\mathsf{P}\{AB\}$, где B — произвольное событие порожденное рекордными величинами $X(1), X(2), \dots, X(n)$, а $A = \{Y_1 < x_1, \dots, Y_k < x_k\}$, перепишем в виде:

$$\mathsf{P}\{AB\} = \sum_{m=0}^{\infty} \mathsf{P}\{ABC_{n,m}\} = \sum_{m=0}^{\infty} \mathsf{P}\{A \mid BC_{n,m}\}\mathsf{P}\{BC_{n,m}\} =$$
$$= \sum_{m=0}^{\infty} \mathsf{P}\{X_{m+1} < x_1, \quad \dots, \quad X_{m+k} < x_k \mid BC_{n,m}\}\mathsf{P}\{BC_{n,m}\}.$$

Поскольку событие $BC_{n,m}$ определеяется случайными величинами $X_1, X_2, \dots, X_m$, которые не зависят от $X_{m+1}, \dots X_{m+k}$, получаем, что

$$\mathsf{P}\{AB\} = \sum_{m=0}^{\infty} \mathsf{P}\{X_{m+1} < x_1, \quad \dots, \quad X_{m+k} < x_k \mid BC_{n,m}\}\mathsf{P}\{BC_{n,m}\} =$$
$$= F(x_1) \dots F(x_k) \sum_{m=0}^{\infty} \mathsf{P}\{BC_{n,m}\} = F(x_1) \dots F(x_k)\mathsf{P}\{B\} \qquad \blacksquare$$

Если предполагать, что исходные случайные величины имеют некоторую дискретную функцию распределения $F(x)$, то выполнена следующая теорема:

Теорема 3 *Пусть зафиксированы значения*

$$X(1) = x_1, \quad \ldots, \quad X(n) = x_n,$$

причем $0 < x_j - x_{j-1} \leqslant C_{j-1}, \quad j = 2, 3, \ldots, n-1.$
Тогда

$$\mathsf{P}\{X(n+1) = x \mid X(1) = x_1, \ldots, X(n) = x_n\} =$$
$$= \frac{\mathsf{P}\{X_1 = x\}}{\mathsf{P}\{x < x_n + C_n + 1\} - \mathsf{P}\{x < x_n + 1\}} = \frac{\mathsf{P}\{X_1 = x\}}{F(x_n + C_n) - F(x_n)},$$

если $x_n < x \leqslant x_n + C_n$, *и*

$$\mathsf{P}\{X(n+1) = x \mid X(1) = x_1, \ldots, X(n) = x_n\} = 0$$

иначе.

Доказательство Представим случайную величину $X(n+1)$ как $X_{L(n)+\tau(X(n))}$, где $\tau(u)$ – индекс первой в последовательности $X_{u+1}, X_{u+2}, \ldots$ случайной величины, для которой верно $X_u < X_{\tau(u)} \leqslant X_u + C_u$.
Тогда

$$\mathsf{P}\{X(n+1) = x \mid X(1) = x_1, \ldots, X(n) = x_n\} =$$
$$= \mathsf{P}\{X_{L(n)+\tau(x_n)} = x \mid X(1) = x_1, \ldots, X(n) = x_n\} =$$
$$= \mathsf{P}\{X_{L(n)+\tau(x_n)} = x\}$$

по лемме 1.
Обозначим за $Y_i = X_{L(n)+i}, i = 1, 2, \ldots$. Тогда

$$\mathsf{P}\{X_{L(n)+\tau(x_n)} = x\} = \mathsf{P}\{Y_{\tau(x_n)} = x\} =$$
$$= \mathsf{P}\{Y_1 = x\} + \mathsf{P}\{Y_1 \leqslant x_n, Y_1 > x_n + C_n, Y_2 = x\} + \ldots +$$
$$+ \mathsf{P}\{Y_1 \leqslant x_n, Y_1 > x_n + C_n, \ldots, Y_k = x\} + \ldots =$$
$$= \mathsf{P}\{Y_1 = x\} \Big(1 + (F(x_n) + 1 - F(x_n + C_n)) + (F(x_n) + 1 - F(x_n + C_n))^2 + \ldots\Big) =$$
$$= \frac{\mathsf{P}\{X_1 = x\}}{F(x_n + C_n) - F(x_n)} = \frac{\mathsf{P}\{X_1 = x\}}{\mathsf{P}\{x < x_n + C_n + 1\} - \mathsf{P}\{x < x_n + 1\}},$$

если $x_n < x \leqslant x_n + C_n$, и $\mathsf{P}\{X(n+1) = x \mid X(1) = x_1, \ldots, X(n) = x_n\} = 0$ иначе. ■

Тогда совместная вероятность

$$\mathsf{P}\{X(n) = i_n, \ldots, X(1) = i_1\} =$$
$$= \mathsf{P}\{X_1 = i_n\} \prod_{j=1}^{n-1} \frac{\mathsf{P}\{X_1 = i_j\}}{\mathsf{P}\{X_1 < i_j + C_j + 1\} - \mathsf{P}\{X_1 < i_j + 1\}},$$

где $i_1 < i_2 < \ldots < i_n$ и $i_k - i_{k-1} \leqslant C_{k-1}$, $k = 2, 3, \ldots, n$, так как из теоремы 3 последовательность $X(1), X(2), \ldots$ образует цепь Маркова.
Если мы имеем дело со случайными величинами, множество значений которых ограничено снизу некоторым целым неотрицательным a, удобно немного изменить предложенную

схему, введя дополнительно $X(0) = a, L(0) = 0$, и определяя $X(1)$ и $L(1)$ равенством (2). В этом случае совместная вероятность перепишется как

$$\mathsf{P}\{X(n) = i_n, \ldots, X(1) = i_1\} =$$

$$= \frac{\mathsf{P}\{X_1 = i_n\}}{\mathsf{P}\{X_1 < a + C_0 + 1\}} \prod_{j=1}^{n-1} \frac{\mathsf{P}\{X_1 = i_j\}}{\mathsf{P}\{X_1 < i_j + C_j + 1\} - \mathsf{P}\{X_1 < i_j + 1\}}, \tag{3}$$

если $a < i_1 \leqslant a + C_0, i_1 < i_2 \leqslant i_1 + C_1, \ldots, i_{n-1} < i_n \leqslant i_{n-1} + C_{n-1}$, иначе

$$\mathsf{P}\{X(n) = i_n, \ldots, X(1) = i_1\} = 0.$$

2.2 Рекорды с ограничением I для геометрически распределенных величин

Полученные выше соотношения упрощаются, если исходные случайные величины имеют геометрическое распределение и все константы C_n одинаковы и равны C. Рассмотрим геометрическое распределение, сосредоточенное на множестве $\{1, 2, \ldots\}$. Пусть

$$\mathsf{P}\{X_1 = n\} = (1-p)p^{n-1}, \quad n = 1, 2, \ldots.$$

Тогда

$$\mathsf{P}\{X_1 \geqslant n\} = p^{n-1}, \quad n = 1, 2, \ldots.$$

Из равенства (3) при $a = 0$ получаем, что в этом частном случае

$$\mathsf{P}\{X(n) = i_n, \ldots, X(1) = i_1\} =$$

$$= \frac{p^{i_n - 1}(1-p)}{1 - p^C} \prod_{j=1}^{n-1} \frac{p^{i_j - 1}(1-p)}{1 - p^{i_j + C} - 1 + p^{i_j}} = \frac{p^{i_n - n}(1-p)^n}{(1-p^C)^n}, \tag{4}$$

если $0 < i_1 \leqslant C, i_1 < i_2 \leqslant i_1 + C, \ldots, i_{n-1} < i_n \leqslant i_{n-1} + C$, иначе

$$\mathsf{P}\{X(n) = i_n, \ldots, X(1) = i_1\} = 0.$$

Введем разности

$$V(1) = X(1) - X(0), V(2) = X(2) - X(1), \ldots, V(n) = X(n) - X(n-1).$$

Из (4) для них верно

$$\mathsf{P}\{V(n) = v_n, \ldots, V(1) = v_1\} = \frac{p^{(v_1 + v_2 + \ldots + v_n) - n}(1-p)^n}{(1-p^C)^n}, \tag{5}$$

$0 < v_1 \leqslant C, \ldots, 0 < v_n \leqslant C$.

Из (5) следует, что межрекордные разности $V(1), V(2), \ldots$ независимы и имеют распределение вида

$$\mathsf{P}\{V = x\} = \frac{p^{x-1}(1-p)}{(1-p^C)}, \tag{6}$$

если $0 < x \leqslant C$, и $\mathsf{P}\{V = x\} = 0$ иначе. Отсюда получаем, что одинаково распределенными являются случайные векторы

$$\{X(1), X(2), \ldots, X(n)\}$$

и

$$\{W_1, W_1 + W_2, \dots, W_1 + W_2 + \ldots + W_n\},$$

где независимые случайные величины $W_1, W_2, \dots, W_n$ имеют распределение (6).
Таким образом, рекордная величина $X(n)$ представима в виде суммы n независимых одинаково распределенных случайных величин.
Математическое ожидание и дисперсия величин W_k задается равенствами

$$a(C) = \mathsf{E}W_k = \frac{1}{(1-p)(1-p^C)}$$
$$\sigma^2(C) = \mathsf{D}W_k = \frac{p^{2C} - 2p^{C+1} + p^{2C+1} + p}{(1-p)^2(1-p^C)^3}.$$

Отсюда получаем асимптотическую нормальность величины $X(n)$

$$\frac{X(n) - na(C)}{\sigma(C)\sqrt{n}} \underset{n\to\infty}{\longrightarrow} \mathsf{N}(0,1).$$

Заметим, что классическая рекордная модель является частным случаем схемы, рассмотренной при $C = \infty$, и при таком значении C полученные результаты совпадают с соответствующими результатами для классической рекордной модели в случае дискретных распределений.

2.3 Рекорды с ограничением II

Теперь рассмотрим вторую модель рекордов с ограничениями. Снова зафиксируем положительные целые константы C_k, зависящие от номера рекорда, но в этом случае значение, превышающее последний рекорд $X(n)$ больше чем на C_n, не теряется, и новой рекордной величиной объявляется $X(n+1) = X(n) + C_n$. Если же наблюдение попало в интервал $(X(n), X(n) + C_n]$, то оно и становится новой рекордной величиной X(n+1).
Для второй рекордной модели верен аналог леммы 1.
Пусть исходные случайные величины имеют некоторую дискретную функцию распределения $F(x)$. Тогда докажем следующее утверждение:

Теорема 4 *Пусть зафиксированы значения*

$$X(1) = x_1, \quad \dots, \quad X(n) = x_n,$$

причем $0 < x_j - x_{j-1} \leqslant C_{j-1}, \quad j = 2, 3, \dots, n-1.$
Тогда

$$\mathsf{P}\{X(n+1) = X(n) + C_n \mid X(1) = x_1, \dots, X(n) = x_n\} = \frac{\mathsf{P}\{X_1 \geqslant x_n + C_n\}}{\mathsf{P}\{X_1 \geqslant x_n + 1\}},$$
$$\mathsf{P}\{X(n+1) = x \mid X(1) = x_1, \dots, X(n) = x_n\} = \frac{\mathsf{P}\{X_1 < x\} - \mathsf{P}\{X_1 \leqslant x_n\}}{1 - \mathsf{P}\{X_1 \leqslant x_n\}},$$

если $x_n < x < x_n + C_n$, *и*

$$\mathsf{P}\{X(n+1) = x \mid X(1) = x_1, \dots X(n) = x_n\} = 0$$

иначе.

Доказательство Представим случайную величину $X(n+1)$ как $X_{L(n)+\tau(X(n))}$, где $\tau(u)$ – индекс первой в последовательности $X_{u+1}, X_{u+2}, \ldots$ случайной величины, для которой верно $X_u < X_{\tau(u)}$.
Тогда

$$\mathsf{P}\{X(n+1) = X(n) + C_n \mid X(1) = x_1, \ldots, X(n) = x_n\} = \mathsf{P}\{X_{L(n)+\tau(x_n)} = x_n + C_n\},$$

и

$$\mathsf{P}\{X(n+1) = x \mid X(1) = x_1, \ldots, X(n) = x_n, x_n < x < x_n + C_n\} =$$
$$= \mathsf{P}\{X_{L(n)+\tau(x_n)} = x \mid x_n < x < x_n + C_n\}.$$

Отсюда

$$\mathsf{P}\{X(n+1) = X(n) + C_n \mid X(n) = x_n\} =$$
$$= \mathsf{P}\{Y_1 \geqslant x_n + C_n\} + \mathsf{P}\{Y_1 \leqslant x_n, Y_2 \geqslant x_n + C_n\} + \ldots +$$
$$+ \mathsf{P}\{Y_1 \leqslant x_n, Y_2 \leqslant x_n, \ldots, Y_k \geqslant x_n + C_n\} + \ldots =$$
$$= \mathsf{P}\{X_1 \geqslant x_n + C_n\}(1 + \mathsf{P}\{X_1 \leqslant x_n\} + (\mathsf{P}\{X_1 \leqslant x_n\})^2 + \ldots) =$$
$$= \frac{\mathsf{P}\{X_1 \geqslant x_n + C_n\}}{1 - \mathsf{P}\{X_1 \leqslant x_n\}} = \frac{\mathsf{P}\{X_1 \geqslant x_n + C_n\}}{\mathsf{P}\{X_1 \geqslant x_n + 1\}}.$$

И аналогично

$$\mathsf{P}\{X(n+1) < x \mid X(n) = x_n, x_n < x < x_n + C_n\} =$$
$$\mathsf{P}\{Y_1 < x, Y_1 > x_n\} + \mathsf{P}\{Y_1 \leqslant x_n, Y_2 < x, Y_2 > x_n\} + \ldots =$$
$$= \frac{\mathsf{P}\{X_1 < x\} - \mathsf{P}\{X_1 \leqslant x_n\}}{1 - \mathsf{P}\{X_1 \leqslant x_n\}},$$

если $x_n < x < x_n + C_n$, и $\mathsf{P}\{X(n+1) = x \mid X(1) = x_1, \ldots X(n) = x_n\} = 0$ иначе. ■

Будем полагать, не умаляя общности, что наши дискретные случайные величины принимают целые неотрицательные значения. Тогда в случае геометрического распределения сосредоточенного на множестве $\{1, 2, \ldots\}$:

$$\mathsf{P}\{X_1 = n\} = (1-p)p^{n-1}, \quad n = 1, 2, \ldots,$$
$$\mathsf{P}\{X_1 \geqslant n\} = p^{n-1}, \quad n = 1, 2, \ldots.$$

межрекордные разности

$$V(1) = X(1) - X(0), V(2) = X(2) - X(1), \ldots, V(n) = X(n) - X(n-1).$$

будут независимы и будут иметь распределение

$$\mathsf{P}\{V(k) < x\} = \begin{cases} 1 - p^{x-1} & , \quad 0 < x \leqslant C_{k-1} \\ 1 & , \quad x > C_{k-1}. \end{cases}$$

Тогда

$$\mathsf{P}\{V(k) = x\} = \begin{cases} p^{x-1}(1-p) & , \quad 0 < x < C_{k-1} \\ p^{C_{k-1}-1} & , \quad x = C_{k-1}. \end{cases} \tag{7}$$

Таким образом, математическое ожидание $a(C_{k-1})$ и дисперсия $\sigma^2(C_{k-1})$ величин V_k задаются равенствами:

$$
\begin{aligned}
a(C_{k-1}) = \mathsf{E}V_k = C_{k-1}p^{C_{k-1}-1} + \sum_{k=1}^{C_{k-1}-1} k(1-p)p^{k-1} = \\
= C_{k-1}p^{C_{k-1}-1} + \frac{1 - C_{k-1}p^{C_{k-1}-1} + (C_{k-1}-1)p^{C_{k-1}}}{1-p} = \\
= \frac{1-p^{C_{k-1}}}{1-p}, \\
\sigma^2(C_{k-1}) = \mathsf{D}V_k = \frac{p - 2Cp^{C_{k-1}} + p^{C_{k-1}-1} + 2Cp^{C_{k-1}+1} - p^{C_{k-1}+1} - p^{2C_{k-1}}}{(1-p)^2}.
\end{aligned}
$$

Тогда распределения векторов $\{X(1), X(2), \ldots, X(n)\}$ и
$\{v_1, v_1 + v_2, \ldots, v_1 + v_2 + \ldots + v_n\}$ совпадают, где v_i — независимые случайные величины с распределением (7), и при $C_k = C$ для $k = 0, 1, 2, \ldots$

$$\frac{X(n) - na(C)}{\sigma(C)\sqrt{n}} \underset{n\to\infty}{\longrightarrow} Z,$$

где $Z \sim \mathsf{N}(0,1)$ - стандартная нормальная величина.

2.4 Рекорды с ограничением III

Рассмотрим следующую модель рекордов с ограничениями. Зафиксируем положительные целые константы C_k, зависящие от номера рекорда. Значение, превышающее последний рекорд $X(n)$ больше чем на C_n, объявляется новой рекордной величиной $X(n+1)$. Если же наблюдение попало в интервал $(X(n), X(n) + C_n]$, то новой рекордной величиной объявляется $X(n+1) = X(n) + C_n$.

Для третьей рекордной модели верен аналог леммы 1.

Пусть исходные случайные величины имеют некоторую дискретную функцию распределения $F(x)$. Тогда докажем следующее утверждение:

Теорема 5 *Пусть зафиксированы значения*

$$X(1) = x_1, \quad \ldots, \quad X(n) = x_n,$$

причем $x_j - x_{j-1} \geqslant C_{j-1}, \quad j = 2, 3, \ldots, n-1.$
Тогда

$$
\begin{aligned}
\mathsf{P}\{X(n+1) = X(n) + C_n \mid X(1) = x_1, \ldots, X(n) = x_n\} = \frac{\mathsf{P}\{X_1 \leqslant x_n + C_n\} - \mathsf{P}\{X_1 \leqslant x_n\}}{1 - \mathsf{P}\{X_1 \leqslant x_n\}}, \\
\mathsf{P}\{X(n+1) < x \mid X(1) = x_1, \ldots, X(n) = x_n\} = \frac{\mathsf{P}\{X_1 < x\} - \mathsf{P}\{X_1 \leqslant x_n + C_n\}}{1 - \mathsf{P}\{X_1 \leqslant x_n\}},
\end{aligned}
$$

если $x > x_n + C_n$, *и*

$$\mathsf{P}\{X(n+1) < x \mid X(1) = x_1, \ldots X(n) = x_n\} = 0$$

иначе.

Доказательство Представим случайную величину $X(n+1)$ как $X_{L(n)+\tau(X(n))}$, где $\tau(u)$ – индекс первой в последовательности $X_{u+1}, X_{u+2}, \ldots$ случайной величины, для которой верно $X_u < X_{\tau(u)}$.
Тогда

$$\mathsf{P}\{X(n+1) = x \mid X(1) = x_1, \ldots, X(n) = x_n, x > x_n + C_n\} =$$
$$= \mathsf{P}\{X_{L(n)+\tau(x_n)} = x \mid x > x_n + C_n\},$$

и

$$\mathsf{P}\{X(n+1) = X(n) + C_n \mid X(1) = x_1, \ldots, X(n) = x_n\} = \mathsf{P}\{X_{L(n)+\tau(x_n)} = x, x \leqslant x_n + C_n\}.$$

Обозначим за $Y_i = X_{L(n)+i}, i = 1, 2, \ldots$. Тогда

$$\mathsf{P}\{X(n+1) = X(n) + C_n \mid X(n) = x_n\} =$$
$$= \mathsf{P}\{Y_1 > x_n, Y_1 \leqslant x_n + C_n\} + \mathsf{P}\{Y_1 \leqslant x_n, Y_2 > x_n, Y_2 \leqslant x_n + C_n\} + \ldots +$$
$$+\mathsf{P}\{Y_1 \leqslant x_n, Y_2 \leqslant x_n, \ldots, Y_k > x_n, Y_k \leqslant x_n + C_n\} + \ldots =$$
$$= \mathsf{P}\{X_1 > x_n, X_1 \leqslant x_n + C_n\}(1 + \mathsf{P}\{X_1 \leqslant x_n\} + (\mathsf{P}\{X_1 \leqslant x_n\})^2 + \ldots) =$$
$$= \frac{\mathsf{P}\{X_1 > x_n, X_1 \leqslant x_n + C_n\}}{1 - \mathsf{P}\{X_1 \leqslant x_n\}} = \frac{\mathsf{P}\{X_1 \leqslant x_n + C_n\} - \mathsf{P}\{X_1 \leqslant x_n\}}{1 - \mathsf{P}\{X_1 \leqslant x_n\}}.$$

И аналогично

$$\mathsf{P}\{X(n+1) < x \mid X(n) = x_n, x > x_n + C_n\} =$$
$$\mathsf{P}\{Y_1 > x_n + C_n, Y_1 < x\} + \mathsf{P}\{Y_1 \leqslant x_n, Y_2 > x_n + C_n, Y_2 < x\} + \ldots =$$
$$= \frac{\mathsf{P}\{X_1 < x\} - \mathsf{P}\{X_1 \leqslant x_n + C_n\}}{1 - \mathsf{P}\{X_1 \leqslant x_n\}},$$

если $x > x_n + C_n$, и $\mathsf{P}\{X(n+1) < x \mid X(1) = x_1, \ldots X(n) = x_n\} = 0$ иначе. ■

Будем полагать, не умаляя общности, что наши дискретные случайные величины принимают целые неотрицательные значения. Тогда в случае геометрического распределения сосредоточенного на множестве $\{1, 2, \ldots\}$:

$$\mathsf{P}\{X_1 = n\} = (1-p)p^{n-1}, \quad n = 1, 2, \ldots,$$
$$\mathsf{P}\{X_1 \geqslant n\} = p^{n-1}, \quad n = 1, 2, \ldots.$$

результат предыдущей теоремы перепишется как

$$\mathsf{P}\{X(n+1) = X(n) + C_n \mid X(1) = x_1, \ldots, X(n) = x_n\} = \frac{\mathsf{P}\{X_1 \leqslant x_n + C_n\} - \mathsf{P}\{X_1 \leqslant x_n\}}{1 - \mathsf{P}\{X_1 \leqslant x_n\}} =$$
$$= \frac{\mathsf{P}\{X_1 \geqslant x_n + 1\} - \mathsf{P}\{X_1 \geqslant x_n + C_n + 1\}}{\mathsf{P}\{X_1 \geqslant x_n + 1\}} = \frac{p^{x_n+1} - p^{x_n+C_n+1}}{p^{x_n+1}} = 1 - p^{C_n},$$
$$\mathsf{P}\{X(n+1) < x \mid X(1) = x_1, \ldots, X(n) = x_n\} = \frac{\mathsf{P}\{X_1 < x\} - \mathsf{P}\{X_1 \leqslant x_n + C_n\}}{1 - \mathsf{P}\{X_1 \leqslant x_n\}} =$$
$$= \frac{\mathsf{P}\{X_1 \geqslant x_n + C_n + 1\} - \mathsf{P}\{X_1 \geqslant x\}}{\mathsf{P}\{X_1 \geqslant x_n + 1\}} = \frac{p^{x_n+C_n+1} - p^x}{p^{x_n+1}} = p^{C_n} - p^{x-x_n-1},$$

если $x > x_n + C_n$, и

$$\mathsf{P}\{X(n+1) < x \mid X(1) = x_1, \ldots X(n) = x_n\} = 0$$

иначе.

Тогда межрекордные разности

$$W(1) = X(1) - X(0), W(2) = W(2) - X(1), \ldots, W(n) = X(n) - X(n-1)$$

будут независимы и будут иметь распределение

$$\mathsf{P}\{W(k) < x\} = \begin{cases} 0 & , \quad 0 < x < C_{k-1} \\ 1 - p^{C_{k-1}} & , \quad x = C_{k-1} \\ p^{C_{k-1}} - p^{x-1} & , \quad x > C_{k-1}. \end{cases}$$

Тогда

$$\mathsf{P}\{W(k) = x\} = \begin{cases} 0 & , \quad 0 < x < C_{k-1} \\ 1 - p^{C_{k-1}} & , \quad x = C_{k-1} \\ p^{x-1}(1-p) + p^{C_k} - p^{C_{k-1}} & , \quad x > C_{k-1}. \end{cases} \tag{8}$$

В случае, если все константы C_k одинаковые, то есть $C_i = C, \quad i = 1, 2, \ldots,$ получаем

$$\mathsf{P}\{W(k) = x\} = \begin{cases} 0 & , \quad 0 < x < C \\ 1 - p^{C} & , \quad x = C \\ p^{x-1}(1-p) & , \quad x > C. \end{cases} \tag{9}$$

Таким образом, математическое ожидание $a(C)$ и дисперсия $\sigma^2(C)$ величин W_k задаются равенствами:

$$a(C) = \mathsf{E}W_k = C(1-p^C) + \sum_{k=C+1}^{\infty} k(1-p)p^{k-1} =$$

$$= C(1-p^C) + \frac{(1 + C(1-p))p^C}{1-p} = C + \frac{p^C}{1-p},$$

$$\sigma^2(C) = \mathsf{D}W_k = \frac{p^{1+C} + p^C - p^{2C}}{(1-p)^2}.$$

Тогда распределения векторов $\{X(1), X(2), \ldots, X(n)\}$ и $\{v_1, v_1 + v_2, \ldots, v_1 + v_2 + \ldots + v_n\}$ совпадают, где v_i — независимые случайные величины с распределением (9), и

$$\frac{X(n) - na(C)}{\sigma(C)\sqrt{n}} \underset{n\to\infty}{\longrightarrow} Z,$$

где $Z \sim \mathsf{N}(0,1)$ - стандартная нормальная величина.

2.5 Рекорды с превышением в случайных блужданиях

Рассмотрим одномерное дискретное случайное блуждание $\{X(n)\}_{n\geqslant 0}$, где $X(n)$ — положение после шага с номером n, которое определяется следующим соотношением:

$$X_0 = 0, \quad X_n = X_{n-1} + Y_n, \tag{10}$$

где Y_i — независимые одинаково распределенные случайные величины. Далее мы будем рассматривать дискретные случайные блуждания, определенные как в (10), с дополнительным условием, что блуждание возвращается в начальное положение через n шагов, то есть $X(n) = X(0) = 0$.

Для таких блужданий определелим рекорды с ограничением следующим образом. Зафиксируем положительную целую константу C. В качестве первой рекордной величины и первого рекордного момента возьмем $X(1) = X_0$ и $L(1) = 0$. Значение, превышающее последний рекорд $X(n)$ больше чем на C, объявляется новой рекордной величиной $X(n+1)$.

В статье [9] получены результаты для распределения и математического ожидания количества классических рекордов (случай $C = 0$, когда рекордом объявляется значение, превышающее предыдущий рекорд) до шага n в случайных блужданиях с симметричной функцией распределения скачков, которые возвращаются в начальное положение через n шагов.

Мы, пользуясь аналогичной техникой, постараемся обобщить результаты на случай целой константы $C > 0$ в случае, когда случайные величины Y_i будут иметь следующее распределение:

$$\mathsf{P}(Y_i = k) = \begin{cases} \frac{1}{2} & , \quad k = 1 \\ \frac{1}{2} & , \quad k = -1. \end{cases} \tag{11}$$

При таком распределении Y_i блуждание происходит на решетке целых чисел.

Нас будет интересовать математическое ожидание количества рекордов с ограничениями и значение наибольшего рекорда в дискретных случайных блужданиях, которые через n шагов возвращаются в начальное положение.

2.5.1 Математическое ожидание количества рекордов

Теорема 6 *Пусть имеется дискретное случайное блуждание X_k, такое что $X_0 = X_n = 0$ и распределение скачков Y_i задается как в (11). Пусть также $C = 1$. Обозначим за $N(n)$ количество рекордов с ограничениями до шага n.*

Тогда

$$\mathsf{E}(N(n)) = \begin{cases} 1 & , \quad n = 0 \\ \frac{2^{2k-2}}{C_{2k}^k} + \frac{1}{2} & , \quad n \text{ — четное}, \quad n = 2k, \quad k = 1, 2, 3, \ldots \\ 0 & , \quad n \text{ — нечетное}. \end{cases}$$

Доказательство Выразим $N(n)$ следующим образом:

$$N(n) = \sum_{i=0}^{n} \eta(i, n),$$

где $\eta(i, n)$ — рекордный индикатор, который принимает значение 1, если на шаге i произошел рекорд, и значение 0 иначе. В наших условиях $\eta(0, n) = 1$, так как начальное положение является рекордом, и $\eta(n, n) = 0$, так как через n шагов блуждание возвращается в начальное положение, и $X(n) = X(0) = 0$ рекордом уже не является. Тогда математическое ожидание $N(n)$ примет вид:

$$\mathsf{E}(N(n)) = \sum_{i=0}^{n} \mathsf{E}(\eta(i, n)) = \sum_{i=0}^{n} r(i, n),$$

где $r(i, n)$ — вероятность того, что на шаге i произойдет рекорд. Вычислим вероятности $r(i, n)$. Предположим, что на шаге i появляется рекордное значение величиной x. Это

значит, что блуждание из начального положения $X(0) = 0$ за i шагов впервые попало в точку x. При этом, так как блуждание происходит на целочисленной решетке и на каждом шаге можно сдвинуться только на $+1$ или -1, то все рекордные величины имеют вид $X(i) = (C+1)(i-1)$. В частности, $X(1) = 0, X(2) = C, X(3) = 2C, \ldots$.
Вероятности $r(i,n)$ выразим через следующие величины:

- $G(x_0, x, n)$ — вероятность того, что случайное блуждание попадет из x_0 в x через n шагов.
- $G_>(x_0, x, n)$ — вероятность того, что случайное блуждание попадет из x_0 в x через n шагов находясь строго ниже уровня x после всех шагов, кроме последнего шага n.

Событие появления нового рекорда величиной x после шага с номером i совпадает со следующим событием: блуждание после i шагов впервые попало в значение x и после n шагов вернулось в начальное положение. Вероятностью перейти из 0 в x через i шагов, оставаясь строго ниже уровня x, является $G_>(x_0, x, i)$. Вероятностью перейти из положения x в 0 за $n-i$ шагов является $G(x, 0, n-i) = G(0, x, n-i)$, в силу симметричного распределения скачков. Воспользовавшись независимостью блуждания на промежутках $[0, i]$ и $[i, n]$ и просуммировав по $x = (C+1)k$, где k — целое неотрицательное число, мы получим вероятность такого события.
Таким образом, для $n \geqslant 1$

$$r(i,n) = \frac{1}{G(0,0,n)} \sum_{k=0}^{\infty} G_>(0,(C+1)k,i) G(0,(C+1)k,n-i),$$

где делим на $G(0,0,n)$, поскольку рассматриваем блуждания возвращающиеся в 0 после n шагов.
Вычислим производящую функцию величины

$$N_0(n) = \sum_{i=0}^{n} \sum_{k=0}^{\infty} G_>(0,(C+1)k,i) G(0,(C+1)k,n-i).$$

Производящие функции для $G_>(0, x, n)$ и $G(0, x, n)$ вычислены в [5] (Appendix A) и равны

$$\tilde{G}(0,x,z) = \sum_{n=0}^{\infty} z^n G(0,x,n) = \frac{1}{\sqrt{1-z^2}} \left(\frac{1-\sqrt{1-z^2}}{z} \right)^x, \tag{12}$$

$$\tilde{G}_>(0,x,z) = \sum_{n=0}^{\infty} z^n G_>(0,x,n) = \left(\frac{1-\sqrt{1-z^2}}{z} \right)^x. \tag{13}$$

Тогда

$$\sum_{n=0}^{\infty} N_0(n) z^n = \sum_{k=0}^{\infty} \tilde{G}_>(0,(C+1)k,z) \tilde{G}(0,(C+1)k,z).$$

Подставляя значения для производящих функций и суммируя, получаем

$$\sum_{n=0}^{\infty} N_0(n) z^n = \sum_{k=0}^{\infty} \frac{1}{\sqrt{1-z^2}} \left(\frac{1-\sqrt{1-z^2}}{z} \right)^{(C+1)x} \left(\frac{1-\sqrt{1-z^2}}{z} \right)^{(C+1)x} =$$

$$= \frac{1}{\sqrt{1-z^2}\left(1-\left(\frac{1-\sqrt{1-z^2}}{z}\right)^{2(C+1)}\right)}. \tag{14}$$

В случае $C = 1$ выражение (14) перепишется как

$$\sum_{n=0}^{\infty} N_0(n) z^n = \frac{1}{\sqrt{1-z^2}\left(1-\left(\frac{1-\sqrt{1-z^2}}{z}\right)^{4}\right)} =$$

$$= \frac{1}{\sqrt{1-z^2}\left(1-\left(\frac{1-\sqrt{1-z^2}}{z}\right)^{2}\right)\left(1+\left(\frac{1-\sqrt{1-z^2}}{z}\right)^{2}\right)}. \tag{15}$$

Упростим знаменатель.

$$\frac{1}{1-\left(\frac{1-\sqrt{1-z^2}}{z}\right)^{2}} = \frac{z^2}{z^2-(1-\sqrt{1-z^2})^2} = \frac{z^2}{2(z^2-1+\sqrt{1-z^2})} = \frac{z^2}{2\sqrt{1-z^2}(1-\sqrt{1-z^2})} =$$

$$= \frac{1}{2\sqrt{1-z^2}}(1+\sqrt{1-z^2}).$$

$$\frac{1}{1+\left(\frac{1-\sqrt{1-z^2}}{z}\right)^{2}} = \frac{z^2}{z^2+(1-\sqrt{1-z^2})^2} = \frac{z^2}{2(1-\sqrt{1-z^2})} = \frac{1}{2}(1+\sqrt{1-z^2}).$$

Подставляя полученные результаты в (15), получим

$$\sum_{n=0}^{\infty} N_0(n) z^n = \frac{1}{4(1-z^2)}(1+(1-z^2)+2\sqrt{1-z^2}) = \frac{1}{4(1-z^2)}(1+\sqrt{1-z^2})^2 =$$

$$= \frac{1}{4} + \frac{1}{4(1-z^2)} + \frac{1}{2\sqrt{1-z^2}} = \frac{1}{4} + \frac{1}{4}\sum_{k=0}^{\infty} z^{2k} + \frac{1}{2}\sum_{k=0}^{\infty} \frac{C_{2k}^k}{2^{2k}} z^{2k}. \tag{16}$$

Отсюда

$$N_0(n) = \begin{cases} 1 & , \quad n = 0 \\ \frac{1}{4} + \frac{C_{2k}^k}{2^{2k+1}} & , \quad n \text{ — четное}, \quad n = 2k, \quad k = 1, 2, 3, \ldots \\ 0 & , \quad n \text{ — нечетное}. \end{cases} \tag{17}$$

Так как

$$\mathsf{E}(N(n)) = \sum_{i=0}^{n} r(i, n) = \frac{N_0(n)}{G(0,0,n)},$$

и

$$G(0,0,n) = \frac{C_n^{\frac{n}{2}}}{2^n},$$

если n — четное и 0 иначе, то математическое ожидание количества рекордов до шага n выразится как

$$\mathsf{E}(N(n)) = \begin{cases} 1 & , \quad n = 0 \\ \frac{2^{2k-2}}{C_{2k}^k} + \frac{1}{2} & , \quad n \text{ — четное}, \quad n = 2k, \quad k = 1, 2, 3, \ldots \\ 0 & , \quad n \text{ — нечетное}. \end{cases} \tag{18}$$

■

Воспользуемся асимптотикой для биномиальных коэффициентов: $C_{2k}^k \sim \frac{2^{2k}}{\sqrt{\pi n}}$.
Математическое ожидание количества рекордов до шага n при больших n в случае $C = 1$ ведет себя как

$$\mathsf{E}(N(n)) \sim \begin{cases} \frac{\sqrt{\pi}}{4\sqrt{2}}\sqrt{n} & , \quad n \text{ — четное} \\ 0 & , \quad n \text{ — нечетное.} \end{cases} \tag{19}$$

В случае $C > 1$ не удалось получить общей формулы для математического ожидания количества рекордов до шага n. Из разложения в ряд для производящей функции $N_0(n)$ (14) можем привести несколько значений для небольших n и небольших C:

	$\mathsf{E}(N(0))$	$\mathsf{E}(N(2))$	$\mathsf{E}(N(4))$	$\mathsf{E}(N(6))$	$\mathsf{E}(N(8))$	$\mathsf{E}(N(10))$
$C = 1$	1	1	$\frac{7}{6}$	$\frac{13}{10}$	$\frac{99}{70}$	$\frac{191}{126}$
$C = 2$	1	1	1	$\frac{21}{20}$	$\frac{39}{35}$	$\frac{33}{28}$
$C = 3$	1	1	1	1	$\frac{71}{70}$	$\frac{131}{126}$
$C = 4$	1	1	1	1	1	$\frac{253}{252}$

Таблица 1: Таблица значений $\mathsf{E}(N(n))$ математического ожидания количества рекордов до шага n для некоторых значений n и C.

2.5.2 Максимальное рекордное значение

Обозначим за $M_(n)$ — максимальное рекордное значение после шага n.
Событие $\{M(n) = (C+1)k, \quad k \in \mathbb{Z}\}$ совпадает со следующим событием: блуждание попадает из начального положения 0 в положение $(C+1)k$ после i шагов, находясь строго ниже уровня $(C+1)k$, а потом возвращается в исходное положение после n шагов, не превышая уровень $(C+1)(k+1)$, для того чтобы не появилось большего рекорда.
Просуммировав по i вероятности таких событий получим вероятность

$$\mathsf{P}\{M(n) = (C+1)k\} = \frac{1}{G(0,0,n)} \sum_{i=0}^{n} G_{>}(0,(C+1)k,i) G_{<,C}((C+1)k, 0, n-i), \tag{20}$$

где делим на $G(0,0,n)$, поскольку рассматриваем блуждания, возвращающиеся в 0 после n шагов, и $G_{<,C}(x_0, x, n)$ — вероятность того, что блуждание из x_0 попадет в $x < x_0$ за n шагов, оставаясь строго ниже уровня $x + (C+1)$. Эту вероятность можно вычислить с помощью принципа отражения:

$$G_{<,C}(x,0,n) = G(x,0,n) - G(x+2(C+1),0,n) = G(0,x,n) - G(0,x+2(C+1),n) =$$
$$= \begin{cases} \frac{1}{2^n}\left(C_n^{\frac{n+x}{2}} - C_n^{\frac{n+x}{2}+(C+1)}\right) & , \quad n+x \text{ — четное} \\ 0 & , \quad n+x \text{ — нечетное.} \end{cases}$$

Тогда пользуясь выражением для производящей функции $\tilde{G}(0,x,z)$ (12) получим

$$\tilde{G}_{<,C}(x,0,z) = \sum_{n=0}^{\infty} G_{<,C}(x,0,z) z^n =$$
$$= \frac{1}{\sqrt{1-z^2}} \left(\frac{1-\sqrt{1-z^2}}{z}\right)^x - \frac{1}{\sqrt{1-z^2}} \left(\frac{1-\sqrt{1-z^2}}{z}\right)^{x+2(C+1)} =$$

$$= \frac{1}{\sqrt{1-z^2}} \left(\frac{1-\sqrt{1-z^2}}{z} \right)^x \left(1 - \left(\frac{1-\sqrt{1-z^2}}{z} \right)^{2(C+1)} \right). \tag{21}$$

Обозначим за $P_0((C+1)k,n)$ сумму из (20)

$$P_0((C+1)k,n) = \sum_{i=0}^{n} G_{>}(0,(C+1)k,i) G_{<,C}((C+1)k,0,n-i)$$

и вычислим производящую функцию $P_0((C+1)k,n)$:

$$\sum_{n=0}^{\infty} P_0((C+1)k,n) z^n = \tilde{G}_{>}(0,(C+1)k,z) \tilde{G}_{<,C}((C+1)k,0,z).$$

Подставив выражения для производящих функций $\tilde{G}_{>}(0,(C+1)k,z)$ и $\tilde{G}_{<,C}((C+1)k,0,z)$ получим

$$\sum_{n=0}^{\infty} P_0((C+1)k,n) z^n = \frac{1}{\sqrt{1-z^2}} \left(\frac{1-\sqrt{1-z^2}}{z} \right)^{2(C+1)k} \left(1 - \left(\frac{1-\sqrt{1-z^2}}{z} \right)^{2(C+1)} \right). \tag{22}$$

Получить из (22) явное представление для $P_0((C+1)k,n)$ для произвольных C и k не удалось.
Зная $P_0((C+1)k,n)$ искомые вероятности вычисляются как

$$\mathsf{P}\{M(n) = (C+1)k\} = \frac{P_0((C+1)k,n)}{G(0,0,n)}.$$

Приведем несколько значений $\mathsf{P}\{M(n) = (C+1)k\}$ для небольших C и k.

	$n=0$	$n=2$	$n=4$	$n=6$	$n=8$	$n=10$
$C=1, k=0$	1	1	$\frac{5}{6}$	$\frac{7}{10}$	$\frac{3}{5}$	$\frac{11}{21}$
$C=1, k=1$	0	0	$\frac{1}{6}$	$\frac{3}{10}$	$\frac{27}{70}$	$\frac{55}{126}$
$C=1, k=2$	0	0	0	0	$\frac{1}{70}$	$\frac{5}{126}$
$C=2, k=0$	1	1	1	$\frac{19}{20}$	$\frac{31}{35}$	$\frac{23}{28}$
$C=2, k=1$	0	0	0	$\frac{1}{20}$	$\frac{4}{35}$	$\frac{5}{28}$
$C=2, k=2$	0	0	0	0	0	0

Таблица 2: Таблица значений $\mathsf{P}\{M(n) = (C+1)k\}$ вероятностей того, что наибольший рекорд равен $(C+1)k$ для некоторых значений k , C и n.

Еще одним результатом является то, что значение наибольшего рекорда напрямую связано с количеством рекордов: если $M(n) = (C+1)k$, то $N(n) = k+1$. Таким образом,

$$\mathsf{P}\{N(n) = k\} = \mathsf{P}\{M(n) = (C+1)(k-1)\}. \tag{23}$$

3 Заключение

В заключении подведем итоги проделанной работы. Мы хотели исследовать рекордные схемы с ограничениями в случае, когда исходные случайные величины имеют дискретное распределение, и были получены следующие результаты:

- Доказана независимость исходных случайных величин $X_{L(n)+1}, X_{L(n)+2}, \ldots$ и рекордных значений $X(1), X(2), \ldots, X(n)$.
- Найдено совместное распределение рекордных значений $X(1), X(2), \ldots, X(n)$.
- Доказано, что рекордная величина $X(n)$ представима в виде суммы n независимых одинаково распределенных величин, что позволяет в будущем получить для $X(n)$ ряд предельных теорем, справедливых для сумм независимых случайных величин.
- Рекордная схема с ограничениями была рассмотрена как некоторое обобщение классической рекордной схемы. Было замечено совпадение полученных в этой работе результатов с уже известным результатами для классической рекордной схемы, в частном случае, в котором рекордная схема с ограничениями совпадает с классической рекордной схемой.

Были получены результаты для рекордных схем с ограничениями в случае дискретно распределенных исходных случайных величин, аналогичные результатам в статье [3], полученными для рекордных схем с ограничениями в случае непрерывно распределенных исходных случайных величин.

Также были рассмотрены рекорды с превышениями для одномерных дискретных случайных блужданий, которые возвращаются в исходное положение после n шагов. В случае распределения скачков

$$\mathsf{P}(Y_i = k) = \begin{cases} \frac{1}{2} & , \quad k = 1 \\ \frac{1}{2} & , \quad k = -1. \end{cases}$$

были найдены производящие функции для математического ожидания количества рекордов и величины наибольшего рекорда, и в случае $C = 2$ математическое ожидание количества рекордов до шага n было вычислено явно.

Список литературы

[1] Невзоров В. Б., Рекорды. Математическая теория., Москва: ФАЗИС, 2000.

[2] Balakrishnan N., Balasubramanian K., Panchapakesan S. δ-exceedance records // Journal of Applied Statistical Science. 1996. Vol.4. P.123–132.

[3] Невзоров В. Б., Рекордные величины с ограничением., Вестник СПбГУ, Сер.1 (2013), выпуск 3, С. 70-74.

[4] Gouet R, Lòpez FJ, Sanz G, On δ-record observations: asymptotic rates for the counting process and elements of maximum likelihood estimation. Test 21(1):188–214 (2012).

[5] Lopez-Blazquez, F. Salamanca-Mino B., Distribution theory of δ-record values: case $\delta \geqslant 0$ TEST September 2015, Volume 24, Issue 3, pp 558–582.

[6] Lopez-Blazquez, F. Salamanca-Mino B., Distribution theory of δ-record values. Case $\delta \leqslant 0$, TEST November 2013, Volume 22, Issue 4, pp 715–738.

[7] Gouet, R., Lopez, F.J., Sanz, G. (2007). Asymptotic normality for the counting process of weak records and δ-records in discrete models. Bernoulli 13, 754–781.

[8] Su-Chan Park, Joachim Krug, δ-exceedance records and random adaptive walks //J. Phys. A: Math. Theor. 49, 315601 (2016)

[9] C. Godrèche, S. N. Majumdar and G. Schehr, Record statistics for random walk bridges, J. Stat. Mech. P07026 (2015).

[10] C. Godrèche, S. N. Majumdar and G. Schehr, Record statistics of a strongly correlated time series: random walks and Levy flights, arXiv:1702.00586 [cond-mat.stat-mech] (2017).