



# Теория вероятностей

Конспекты лекций и семинаров

ЛЕКТОР: Д.А. ШАБАНОВ

Конспекты вели Денис Беляков, Никита Попов и Алексей Хачиянц

НИУ ВШЭ, 2016-2017



# Оглавление

<b>1</b>	<b>Дополнительные главы теории вероятностей</b>	<b>5</b>
1.1	Лекция от 24.01.2017	5
1.1.1	Случайные процессы	5
1.1.2	Случайные блуждания	6
1.1.3	Процесс восстановления	6
1.1.4	Простые случайные блуждания	9
1.1.5	Возвращение в ноль в простом случайном блуждании	10
1.2	Лекция от 31.01.2017	11
1.2.1	Числа Каталана. Рекуррентное соотношение. Производящая функция	11
1.2.2	Вероятность возвращения	13
1.2.3	Многомерный случай	13
1.2.4	Числа Каталана через биномиальные коэффициенты	14
1.2.5	Математическое ожидание первого момента возвращения в ноль	15
1.2.6	Среднее время в нуле	15
1.3	Лекция от 07.02.2017	16
1.3.1	Среднее время в нуле	16
1.3.2	Геометрия траекторий. Закон повторного логарифма	18
1.4	Лекция от 14.02.2017	20
1.4.1	Закон повторного логарифма	20
1.4.2	Следствие из ЦПТ и теоремы Берри-Эссеена	20
1.4.3	Доказательство ЗПЛ	22
1.4.4	Ветвящиеся случайные процессы	25
1.4.5	Физическая модель	25
1.4.6	Математическая модель	25



# Глава 1

## Дополнительные главы теории вероятностей

### 1.1 Лекция от 24.01.2017

«На прошлой лекции Дмитрий Александрович запнулся, на этой посмотрел записи. Ну всё, действительно сложный ТеорВер начался.»

---

Один из слушателей

#### 1.1.1 Случайные процессы

**Определение 1.** Пусть есть множество  $T$  (неформально оно обозначает время). Набор случайных величин  $\{x_t, t \in T\}$  будем называть *случайным процессом*.

*Примечание.* То, что написано в определении, на самом деле, называется *случайной функцией*, но для определенности оставим определение в таком же виде, потому что в основном будем использовать  $T \subseteq \mathbb{R}$ , что уже действительно является случайным процессом по определению во многих учебниках.

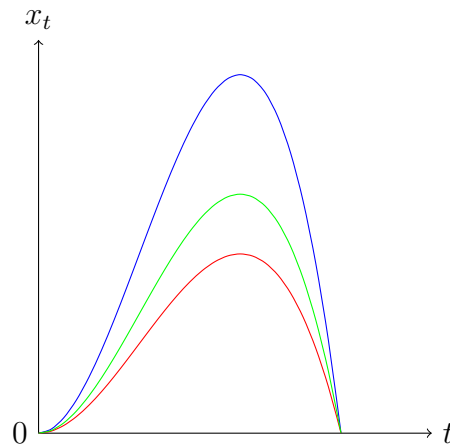
**Определение 2.** Классифицируем случайные процессы:

- если  $T = \mathbb{N}, \mathbb{Z}, \mathbb{Z}_+$  — случайный процесс с дискретным временем.
- если  $T = [a, b], \mathbb{R}, \mathbb{R}_+$  — случайный процесс с непрерывным временем.
- если  $T \subseteq \mathbb{R}^d, d > 1$ , тогда случайный процесс будет случайным полем.

Сейчас уделим больше внимания дискретным процессам. Будем считать, что у нас задана тройка Колмогорова  $(\Omega, \mathcal{F}, \mathbb{P})$  и все  $x_t : \Omega \rightarrow \mathbb{R}$ .

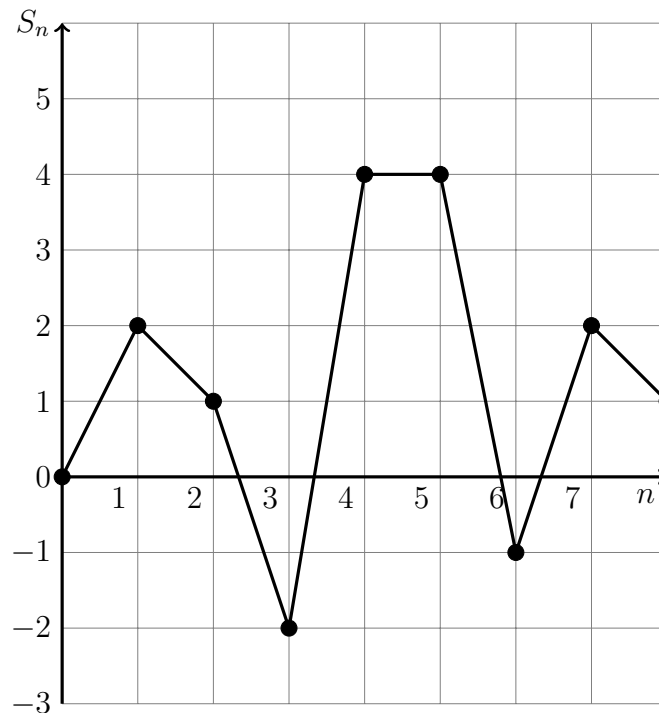
**Определение 3.** Для фиксированного  $\omega_0 \in \Omega$  функция  $\tilde{x}_t(t) = x_t(\omega_0), t \in T$  является *траекторией* (или *реализацией*) случайного процесса.

*Пример.*  $x_t = f(t)\xi$ , где  $f(t)$  — какая-то функция,  $\xi$  — случайная величина. Приведем пример траекторий для некоторых  $\omega_0$ .



### 1.1.2 Случайные блуждания

**Определение 4.** Пусть  $\{\xi_n, n \in \mathbb{N}\}$  — независимые случайные величины. Определим  $S_0 = 0$ , а  $S_n = \xi_1 + \dots + \xi_n, n \in \mathbb{N}$ . Тогда  $\{S_n, n \in \mathbb{Z}_+\}$  называют *случайным блужданием*.



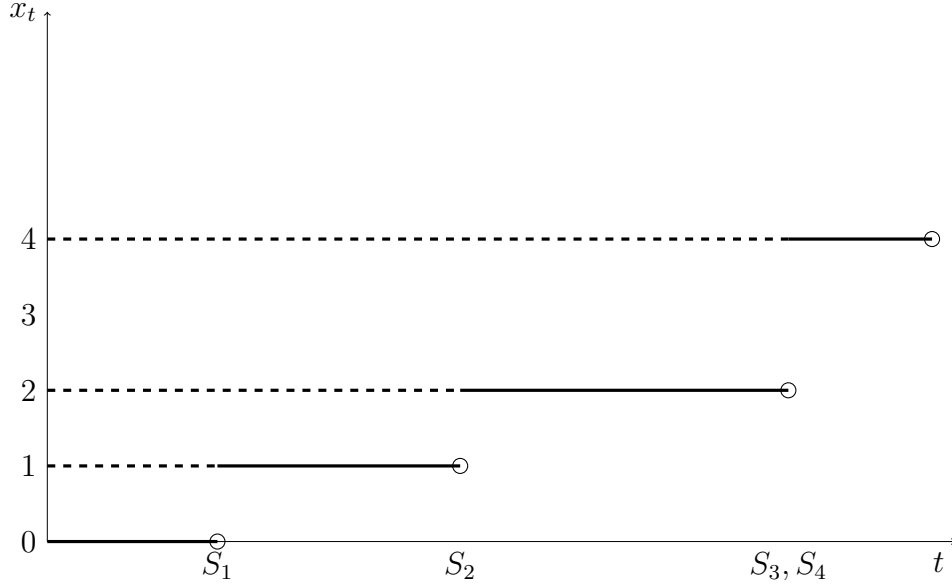
*Пример.* Физической моделью, соответствующую случайному блужданию, могут являться прыжки кузнечика.

### 1.1.3 Процесс восстановления

**Определение 5.** Пусть  $\{\xi_n, n \in \mathbb{N}\}$  — одинаково распределенные неотрицательные случайные величины. Положим  $S_0 = 0, S_n = \xi_1 + \dots + \xi_n, n \in \mathbb{N}$ , а также для каждого  $t \geq 0$  рассмотрим такие случайные величины:  $x_t = \max\{n : S_n \leq t\}$  (если максимума не существует, положим  $x_t = +\infty$ ).

Процесс  $(x_t, t \geq 0)$  называется *процессом восстановления* для случайных величин  $\{\xi_n, n \in \mathbb{N}\}$ .

График можно описать как-то так для какого-то  $\omega_0$ :



Где возможны склеивания, как в  $S_3, S_4$  — это лишь означает, что  $\xi_4 = 0$ .

*Пример.* Физической моделью, соответствующую процессу восстановления, может служить процесс горения лампочки, где  $\xi_n$  — общее время горения  $n$ -ой лампочки, а  $x_t$  тогда будет количеством поменянных лампочек до времени  $t$ .

Пара бы уже что-то доказать. Действительно, покажем, что мы не можем часто убегать в бесконечность. На самом деле, такие ситуации очень плохи в реальной жизни. Происходит «перенасыщение» чего-то. В примере с лампочкой, мы можем менять бесконечное число лампочек за время ноль. Такая ситуация очень плоха, поэтому, чтобы успокоиться, докажем следующее утверждение:

**Теорема 1.** *Процесс восстановления конечен с вероятностью один, если  $P(\xi_i = 0) < 1$ .*

*Доказательство.* Пусть  $t > 0$  фиксировано.  $P(x_t = +\infty) = P(\forall n : S_n \leq t)$ .

Заметим, что  $S_n \leq S_{n+1}$  из-за неотрицательности  $\xi_{n+1}$ .

Выпишем тривиальное равенство пределов:

$$\lim_{n \rightarrow +\infty} P(S_n \leq t) = \lim_{n \rightarrow +\infty} P\left(\frac{S_n}{n} \leq \frac{t}{n}\right)$$

Есть 2 случая:

- $E[\xi_i]$  конечно и равно  $A > 0$  (больше нуля, так как  $P(\xi_i = 0) < 1$ ). Тогда по закону больших чисел имеем, что  $\frac{S_n}{n} \xrightarrow{P} A$  (на самом деле мы немного лукавим, потому что в основном курсе эта теорема была доказана с использованием, что все моменты до четвёртого конечны, но ЗБЧ работает и при конечности средней величины).

Тогда продолжим равенство пределов:

$$= \lim_{n \rightarrow +\infty} P\left(\frac{S_n}{n} \leq \frac{t}{n}\right) \leq \lim_{n \rightarrow +\infty} P\left(\frac{S_n}{n} \leq \frac{A}{2}\right)$$

Действительно, с какого-то момента  $\frac{t}{n} < \frac{A}{2}$ , так как  $A > 0$ . Далее, из закона больших чисел получаем:

$$= \lim_{n \rightarrow +\infty} P\left(A \leq \frac{A}{2}\right) = 0.$$

Последнее равенство идёт из-за неотрицательности  $A$ .

Поэтому  $P(x_t = +\infty) = 0$  в этом случае.

- $E[\xi_i] = +\infty$ .

Рассмотрим  $\hat{\xi}_i = \min(\xi_i, 1) \leq \xi_i$ . Откуда сразу получаем, что  $\hat{S}_n = \hat{\xi}_1 + \dots + \hat{\xi}_n \leq S_n$ .

Заметим, что  $E[\hat{S}_n]$  конечно, так как матожидание каждого из  $\hat{\xi}_i$  конечно (так как  $\hat{\xi}_i \leq 1$ ).

А значит, что  $P(S_n \leq t) \leq P(\hat{S}_n \leq t)$  (так как  $S_n \geq \hat{S}_n$ , а значит, что  $\hat{S}_n$  принимает меньшие значения).

Но мы уже всё доказали для конечного матожидания  $\hat{\xi}_i$ , поэтому получаем, что  $P(\hat{S}_n \leq t) \rightarrow 0$ , откуда и  $P(S_n \leq t) \rightarrow 0$ , что нам и требуется.

□

Приведём более мощный пример, обобщим эту модель.

*Пример.* Пусть  $\{\xi_n, n \in \mathbb{N}\}$  — независимые и одинаково распределенные случайные величины,  $\{\eta_n, n \in \mathbb{N}\}$  — независимые и одинаково распределенные случайные величины, притом независимы с  $\{\xi_n, n \in \mathbb{N}\}$ .

Пусть  $(x_t, t > 0)$  — процесс восстановления, то есть  $x_t = \max\{n : \xi_1 + \dots + \xi_n \leq t\}$ .

Для  $y_0, c > 0 \in \mathbb{R}$  введём

$$Y_t = y_0 + ct - \sum_{k=1}^{x_t} \eta_k$$

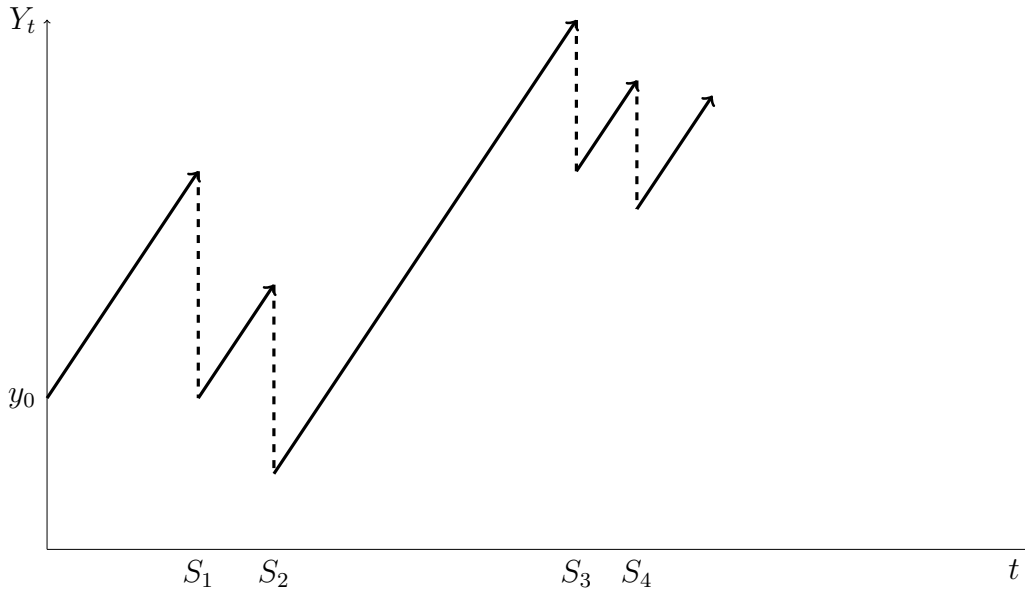
Эту модель называют моделью страхования Спарре-Андресена. Поясним, что значит каждая введенная переменная/величина.

- $y_0$  — начальный капитал.
- $c$  — скорость поступления страховых взносов. Для простоты считают, поступление линейно, что примерно одинаково «бьются» машины в любое время года.
- $\eta_k$  — размер выплаты с номером  $k$ .
- $\xi_k$  — время между  $k-1$ -й и  $k$ -й выплатой.
- $x_t$  — количество выплат к моменту времени  $t$ .
- $\sum_{k=1}^{x_t} \eta_k$  — общий объём выплат по времени.
- И понятно, что тогда  $Y_t$  — текущий капитал.



В будущем, когда в нашем курсе мы затронем мартингалы, мы сможем понять и оценить, какова вероятность, что компания разорится.

Ясно, что тогда график капитала от времени будет выглядеть примерно таким образом:



#### 1.1.4 Простые случайные блуждания

**Определение 6.** Пусть  $\{\xi_n, n \in \mathbb{N}\}$  независимые одинаково распределенные случайные величины такие, что для какого-то  $p \in [0, 1]$

$$P(\xi_n = 1) = p, P(\xi_n = -1) = 1 - p = q$$

Положим  $S_0 = 0, S_n = \xi_1 + \dots + \xi_n, n \in \mathbb{N}$ . Тогда  $\{S_n, n \in \mathbb{Z}_+\}$  называют *простым случайным блужданием на прямой*.

Смысл этого определения в том, что на каждом шаге мы выбираем с какими-то вероятностями в какую сторону пойти.

Понятное дело, что этим можно не ограничиваться и, например, ходить в 4 разные стороны на плоскости. Но давайте пока разберёмся с одномерным случаем.

**Определение 7.** Если  $p = q = \frac{1}{2}$ , то говорят, что случайное блуждание симметрично.

*Примечание.* Не лишним будет упомянуть, что в данном случае  $E[S_n] = (p - q)n$ . Действительно, ответ можно легко получить, если раскрыть по линейности, и воспользоваться совершенно ясным фактом, который получается из определения, что  $E[\xi_i] = p - q$ .

Перед нами возникают достаточно интересные вопросы:

1. Какова вероятность вернуться в ноль после ненулевого количества шагов?
2. Какое среднее время мы проведем в нуле? То есть сколько в среднем раз мы окажемся в нуле при достаточно больших  $n$ .
3. Геометрия траекторий. То есть то, как выглядит график, какие существуют зависимости. Нашей кульминацией на этот вопрос будет закон повторного логарифма, который показывает, насколько далеко в среднем мы можем отходить от нуля.

Будем на все эти вопросы отвечать. Каждый из них — отдельная история, поэтому будем отвечать постепенно.

### 1.1.5 Возвращение в ноль в простом случайном блуждании

Легко понять, что нас интересует  $P(\exists n > 0 : S_n = 0)$ .

Попробуем найти  $P(S_n = x)$ . Во-первых, ясно, что  $n$  и  $x$  должны быть одной четности, так как иначе мы не сможем из нуля прийти в  $x$ . Также, нужно, чтобы  $n \geq |x|$ , иначе мы просто не дойдём до  $x$ , но скомпенсируем в будущем это тем, что  $\binom{n}{k} = 0$  при  $k > n$ .

Пусть мы сделали  $k$  шагов вправо и  $n - k$  шагов влево. Под шагами подразумевается  $+1$  и  $-1$  соответственно. Тогда ясно, что  $k = \frac{n+x}{2}$ , так как должно выполняться равенство  $k - n + k = x$ .

Поэтому получаем, что нам надо выбрать  $k$  ходов, ответ для данной вероятности:

$$P(S_n = x) = \binom{n}{k} p^k q^{n-k} I\{n + x \text{ чётно}\}$$

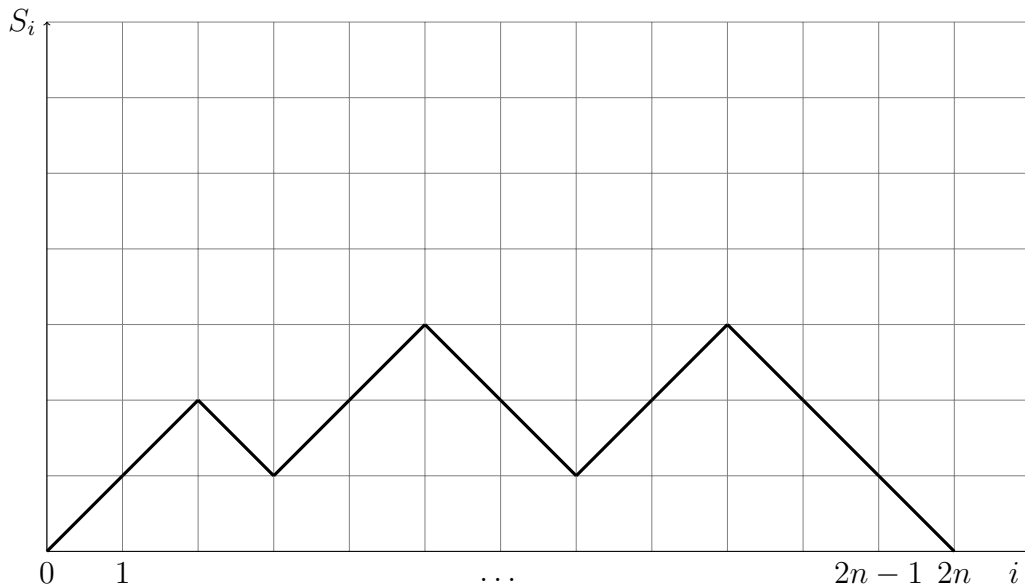
Но этого мало. Мы не сможем как-то легко посчитать вероятность существования  $n$ , что  $S_n = 0$ .

Поэтому давайте поймём, когда впервые достигнем нуля. Ясное дело, что нуля можно достигнуть только на четных ходах. Поэтому давайте посчитаем такую вероятность:

$$P(S_1 \neq 0, \dots, S_{2n-1} \neq 0, S_{2n} = 0)$$

Заметим, что все  $S_1, \dots, S_{2n-1}$  либо одновременно больше нуля, либо все одновременно меньше нуля, так как из-за дискретной непрерывности, если есть  $S_i, S_j$  разных знаков, то найдётся между ними ноль, что противоречит тому, что мы ищем. Так как нам надо сделать  $n$  ходов вправо и влево, эти случаи ничем не отличаются. Посчитаем вероятность, когда все  $S_i > 0, i \in [2n - 1]$ .

Посмотрим на траектории, которые у нас могут получиться:



Теперь мы хотим посчитать все такие пути.

Обозначим через  $\varepsilon_i$  — выбор  $\pm 1$  на  $i$ -ом шаге. Тогда путь подходит тогда и только тогда, когда  $\sum_{i=1}^{2n} \varepsilon_i = 0$  и  $\sum_{i=1}^k \varepsilon_i > 0, k \in [2n - 1]$ .

Обозначим количество таких путей через  $\tilde{C}_n$ .

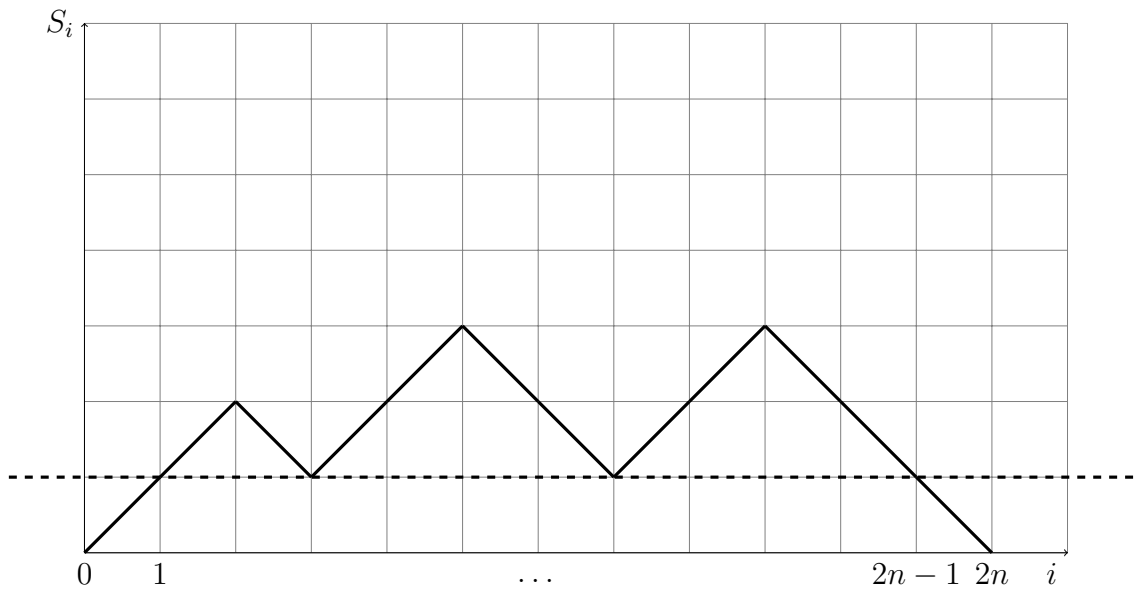
А теперь вспомним, что числа Каталана задаются практически так же. Те, кто ходил на курс дискретной математики, помнят, что есть соответствие между числами Каталана и количеством путей, отвечающим свойствам:  $\sum_{i=1}^{2n} \varepsilon_i = 0$  и  $\sum_{i=1}^k \varepsilon_i \geq 0, k \in [2n - 1]$ .

Обозначим количество таких путей через  $C_n$ .

**Лемма.**  $\tilde{C}_{n+1} = C_n$ .

*Доказательство.* Рассмотрим любой путь, соответствующий  $\tilde{C}_{n+1}$ . Заметим, что первые и последние шаги обязательно  $+1$  и  $-1$  соответственно. Поэтому, при «поднятии» оси  $Oi$  на единицу мы получим, что перед нами путь из  $C_n$ , действительно, это так, так как префиксные суммы уменьшились на один и не стали отрицательными, а сумма по-прежнему сохранилась нулевой.

В другую сторону доказывается аналогично. См. иллюстрацию.



□

Получается, что

$$P(S_1 > 0, \dots, S_{2n-1} > 0, S_{2n} = 0) = C_{n-1}(pq)^n$$

И соответственно:

$$P(S_1 \neq 0, \dots, S_{2n-1} \neq 0, S_{2n} = 0) = 2C_{n-1}(pq)^n$$

Внимательный читатель заметит, что чтобы посчитать самую исходную вероятность, надо просто сложить все такие выше по всем  $n \in \mathbb{N}$ . О том, как такие вещи складывать (в частности о характеристических функциях), мы поговорим в следующей лекции.

## 1.2 Лекция от 31.01.2017

### 1.2.1 Числа Каталана. Реккурентное соотношение. Производящая функция

**Лемма.** Обозначим  $C_0 = 1$ , тогда имеет место следующее равенство:

$$C_n = \sum_{k=0}^{n-1} C_k C_{n-1-k}$$

*Доказательство.* Будем опять рассуждать в терминах положительной траектории (см. рис с предыдущей лекции).

Пусть  $2k > 0$  первый момент, когда наша траектория придёт в ноль. Действительно, такой момент найдётся, потому что в момент времени  $2n$  мы придём в ноль.

Тогда от 0 до  $2k$  положительная траектория, от  $2k$  до  $2n$  неотрицательная, поэтому всего таких путей  $\tilde{C}_k C_{n-k}$ . Чтобы получить все траектории, надо сложить все такие вещи по всем  $k = 1, \dots, n$ , поэтому мы получим, что имеет место равенство по последней лемме из предыдущей лекции и тем, что  $C_0 = 1$ :

$$C_n = \sum_{k=1}^n \tilde{C}_k C_{n-k} = \sum_{k=1}^n C_{k-1} C_{n-k} = \text{«замена } j = k - 1\text{»} = \sum_{j=0}^{n-1} C_j C_{n-1-j}$$

□

**Определение 8.** Для последовательности  $\{a_n\}_{n=0}^{+\infty}$  производящей функцией называется  $f(x) = \sum_{n=0}^{+\infty} a_n x^n$ .

Ряд может расходиться на некоторых  $x$  и тогда мы просто рассматриваем ряд формально, на них можно ввести операции сложения и прочие операции.

Заметим, что  $C_n \leq 2^{2n}$ , потому что всего путей не больше  $2^{2n}$ , на самом деле их меньше аж примерно в  $n^{3/2}$  раза, но это нам нужно для того, чтобы производящая функция чисел Каталана была такова, что при  $|x| < \frac{1}{4}$  ряд сходился. Как мы потом увидим, он будет сходиться и при  $|x| = \frac{1}{4}$ .

Сейчас мы будем рассматривать только  $f(x) = \sum_{n=0}^{+\infty} C_n x^n$  при  $|x| < \frac{1}{4}$ .

**Теорема 2.**  $f(x) = \frac{1 - \sqrt{1 - 4x}}{2x}$ .

*Доказательство.*

$$f^2(x) = \left( \sum_{n=0}^{+\infty} C_n x^n \right)^2 = \sum_{n=0}^{+\infty} \left( \sum_{k=0}^n C_k C_{n-k} \right) x^n.$$

Внимательный читатель заметит, что мы в скобках в точности получили рекуррентное соотношение для чисел Каталана для  $C_{n+1}$ . Поэтому это равно:

$$\sum_{n=0}^{+\infty} \left( \sum_{k=0}^n C_k C_{n-k} \right) x^n = \sum_{n=0}^{+\infty} C_{n+1} x^n = \frac{f(x) - 1}{x}.$$

Где последнее равенство возникает из-за того, что  $C_0 = 1$ .

Откуда мы получаем квадратное уравнение относительно  $f(x)$ .

$$x f^2(x) - f(x) + 1 = 0.$$

Решая уравнение, получим, что

$$f(x) = \frac{1 \pm \sqrt{1-4x}}{2x}$$

Но мы знаем, что  $f(0) = 1$ , поэтому с плюсом не подходит, так как предел будет не тот. Откуда единственный подходящий вариант будет

$$f(x) = \frac{1 - \sqrt{1-4x}}{2x}$$

□

Заметим, что и при  $\frac{1}{4}$  мы получим конечное число, поэтому по непрерывности можно сказать, что и в  $x = \frac{1}{4}$  ряд сходится.

### 1.2.2 Вероятность возвращения

**Теорема 3.** *Случайное блуждание возвратно с вероятностью  $1 - |p - q|$ .*

*Доказательство.* Заметим, что подставить  $pq$  в производящую функцию можно, так как  $pq \leq \frac{1}{4}$ .

$$\begin{aligned} P(\exists n : S_n = 0) &= \sum_{n=1}^{+\infty} P(S_1 \neq 0, \dots, S_{2n-1} \neq 0, S_{2n} = 0) = \\ &= \sum_{n=1}^{+\infty} 2C_{n-1}(pq)^n = 2pq \sum_{n=0}^{+\infty} C_n(pq)^n = 2pq \frac{1}{2pq} (1 - \sqrt{1-4pq}) = \text{«так как } 1 = (p+q)^2\text{»} = \\ &= 1 - \sqrt{(p-q)^2} = 1 - |p - q| \end{aligned}$$

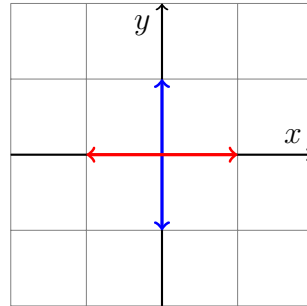
□

*Примечание.* Если  $p = q = \frac{1}{2}$ , то случайное блуждание возвратно с вероятностью один.

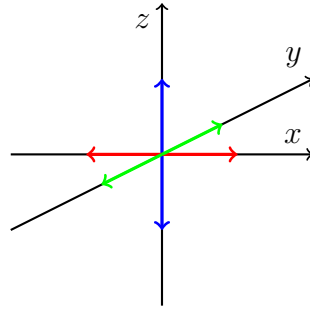
### 1.2.3 Многомерный случай

Абстрактно поговорим о многомерном случае. То есть мы находимся в  $\mathbb{Z}^d, d > 1$ .

На плоскости мы можем идти в 4 разные стороны с равными вероятностями. Как ни странно, вероятность возвращения в ноль в данном случае будет тоже равна один.



В трёхмерном случае не всё так однозначно. У нас есть шесть направлений. И в данном случае вероятность возвращения будет строго меньше единицы. С этим разобраться мы предложим читателю в будущих задачах.



На самом деле всё это зависит от  $P(S_{2n} = 0)$ . При  $d = 1$  и  $p = q = \frac{1}{2}$  получим, что  $P(S_{2n} = 0) = \binom{2n}{n} 4^{-n} \approx \Theta\left(\frac{1}{\sqrt{n}}\right)$ , ряд расходится.

При  $d = 2$  будет  $\Theta\left(\frac{1}{n}\right)$ , что тоже расходится, с каждой новой размерностью добавляется этот самый  $\sqrt{n}$  в знаменателе, поэтому с  $d = 3$  ряд будет сходящимся. Это была подсказка на будущие задачи, ничего тут мы пока не доказывали.

### 1.2.4 Числа Каталана через биномиальные коэффициенты

Вспомним некоторые факты про ряд Тейлора, а именно, что

$$(1+x)^\alpha = \sum_{n=0}^{+\infty} \binom{\alpha}{n} x^n,$$

где  $\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}$ .

Поэтому давайте преобразуем ряд для  $f(x) = \sum_{n=0}^{+\infty} C_n x^n$ .

$$\sqrt{1-4x} = \sum_{n=0}^{+\infty} \binom{1/2}{n} x^n = 1 + \sum_{n=1}^{+\infty} \frac{\frac{1}{2}(\frac{1}{2}-1)\dots(\frac{1}{2}-n+1)}{n!} (-4x)^n =$$

Вынесем  $1/2$  из каждой дроби, везде поменяем знак, получим, что это домножится на  $(-1)^{n-1}$ , что совместно с  $(-4x)^n$  даст знак минус, в скобках останется  $(2n-3)!!$ :

$$= 1 - \sum_{n=1}^{+\infty} \frac{4^n x^n (2n-3)!!}{2^n n!} = 1 - \sum_{n=1}^{+\infty} \frac{2^n x^n (2n-3)!!}{n!} =$$

Домножим на  $n!$  и разделим на него, а также домножим на  $2n-1$  и опять же разделим. Воспользуемся тем, что  $(2n)! = (2n-1)!! n! \cdot 2^n$ :

$$= 1 - \sum_{n=1}^{+\infty} \frac{(2n)! x^n}{n! n!} \frac{1}{2n-1} = 1 - \sum_{n=1}^{+\infty} \binom{2n}{n} \frac{x^n}{2n-1}$$

Откуда:

$$f(x) = \frac{1}{2x} (1 - \sqrt{1-4x}) = \frac{1}{2} \sum_{n=1}^{+\infty} \binom{2n}{n} \frac{1}{2n-1} x^{n-1} = \frac{1}{2} \sum_{n=0}^{+\infty} \binom{2n+2}{n+1} \frac{1}{2n+1} x^n = \sum_{n=0}^{+\infty} \binom{2n}{n} \frac{1}{n+1} x^n$$

В последнем равенстве можно убедиться непосредственно.

Откуда получаем  $C_k = \frac{1}{k+1} \binom{2k}{k}$ .

*Примечание.*  $P(S_1 \neq 0, \dots, S_{2n-1} \neq 0, S_{2n} = 0) = 2C_{n-1}(pq)^n = \binom{2n}{n} (pq)^n \frac{1}{2n-1}$  — распределение первого момента в нуле.

### 1.2.5 Математическое ожидание первого момента возвращения в ноль

**Определение 9.**  $X = \min(2n : S_{2n} = 0)$ .

**Теорема 4.**  $E[X] = +\infty$ .

*Доказательство.* Если  $p \neq q$ , тогда  $P(x = +\infty) > 0$ , поэтому математическое ожидание уже точно бесконечность.

Если  $p = q = \frac{1}{2}$ , тогда  $E[X] = \sum_{n=1}^{+\infty} P(X = 2n) \cdot (2n) = \sum_{n=1}^{+\infty} \binom{2n}{n} \frac{2n}{2^{n-1} 4^n} = \Theta\left(\frac{1}{\sqrt{n}}\right)$ , что расходится  $\square$

Получается, что да, в ноль мы вернемся, но очень не скоро и бесконечно долго будем ходить вне. Вот такой вот парадокс. Перейдём к следующему пункту.

### 1.2.6 Среднее время в нуле

Оказывается, что в нуле мы будем находиться не так мало. За  $n$  шагов примерно  $\sqrt{n}$  раз мы будем в нуле. Это по большей мере связано с тем, что нули расположены рядом, то есть есть большая вероятность, что если мы пришли в ноль, то через малое количество шагов окажемся опять там.

**Определение 10.** Будем рассматривать симметричное простое случайное блуждание. Тогда  $L_n = |k : k \in \overline{0, \dots, n}, S_k = 0|$ .

Хотим понять, какая асимптотика у  $E[L_n]$ .

**Лемма.**  $E[L_n] = E[|S_{n+1}|]$

*Доказательство.* Распишем  $|S_{n+1}|$ :

$$|S_{n+1}| = \begin{cases} 1, & \text{если } |S_n| = 0 \\ S_n + \xi_{n+1}, & \text{если } S_n > 0 \\ -S_n - \xi_{n+1}, & \text{если } S_n < 0 \end{cases}$$

Действительно, при положительном  $S_n$  мы не поменяем знак, поэтому надо лишь добавить то, что мы выбирали на следующем ходу, при отрицательном  $S_n$  аналогично, а если  $S_n = 0$ , то в любом случае модуль равен единице.

Запишем  $|S_{n+1}|$  через индикаторы:

$$|S_{n+1}| = I\{|S_n| = 0\} + (S_n + \xi_{n+1})I\{S_n > 0\} - (S_n + \xi_{n+1})I\{S_n < 0\}$$

Заметим, что  $S_n I\{S_n > 0\} - S_n I\{S_n < 0\} = |S_n|$  (в этом легко убедиться, проверив несколько случаев). Поэтому, введя функцию знака, можно утверждать, что:

$$|S_{n+1}| = I\{|S_n| = 0\} + |S_n| + \xi_{n+1} \text{sgn}(S_n)$$

Будем раскрывать рекурсивно, откуда получим:

$$|S_{n+1}| = \sum_{k=0}^n I\{|S_k| = 0\} + |S_0| + \sum_{k=0}^n \xi_{k+1} \text{sgn}(S_k)$$

Заметим, что  $L_n = \sum_{k=0}^n I\{|S_k| = 0\}$ , поэтому давайте запишем под знаком матожидания с использованием, что  $E[|S_0|] = 0$ .

$$E[|S_{n+1}|] = E[L_n] + \sum_{k=0}^n E[\xi_{k+1} \operatorname{sgn}(S_k)]$$

Но величины  $\xi_{k+1}$  и  $\operatorname{sgn}(S_k)$  независимы, так как нет пересечений по ходам, поэтому это распадается в произведение матожиданий.

Ну и финальный аккорд состоит в том, что  $E[\xi_{k+1}] = 0$ , поэтому мы доказали равенство.  $\square$

Про асимптотику поговорим в следующей лекции.

## 1.3 Лекция от 07.02.2017

### 1.3.1 Среднее время в нуле

Вспомним, что по ЦПТ и тем фактом, что среднее  $\xi_i$  равно нулю и дисперсия равна единице мы имеем следующее:

$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1)$$

По теореме о наследовании сходимости можно заключить, что

$$\frac{|S_n|}{\sqrt{n}} \xrightarrow{d} |\mathcal{N}(0, 1)|$$

А теперь мы хотим осознать, а как ведет себя среднее левой части. Заметим, что просто матожидание мы не можем взять, так как функция  $f(x) = x$  не ограничена и эквивалентным определением сходимости по распределению напрямую воспользоваться нельзя. Поэтому нужны более сильные знания о поведении случайных величин.

**Определение 11.** Последовательность  $\{\xi_n, n \in \mathbb{N}\}$  называется равномерно интегрируемой, если

$$\lim_{c \rightarrow +\infty} \sup_{n \in \mathbb{N}} E[|\xi_n| I\{|\xi_n| \geq c\}] = 0$$

Поясним определение.

Если у случайной величины есть плотность, то фактически мы говорим, что

$$\sup_{n \in \mathbb{N}} \int_{|x| \geq c} |x| p_{\xi_n}(x) dx \rightarrow 0$$

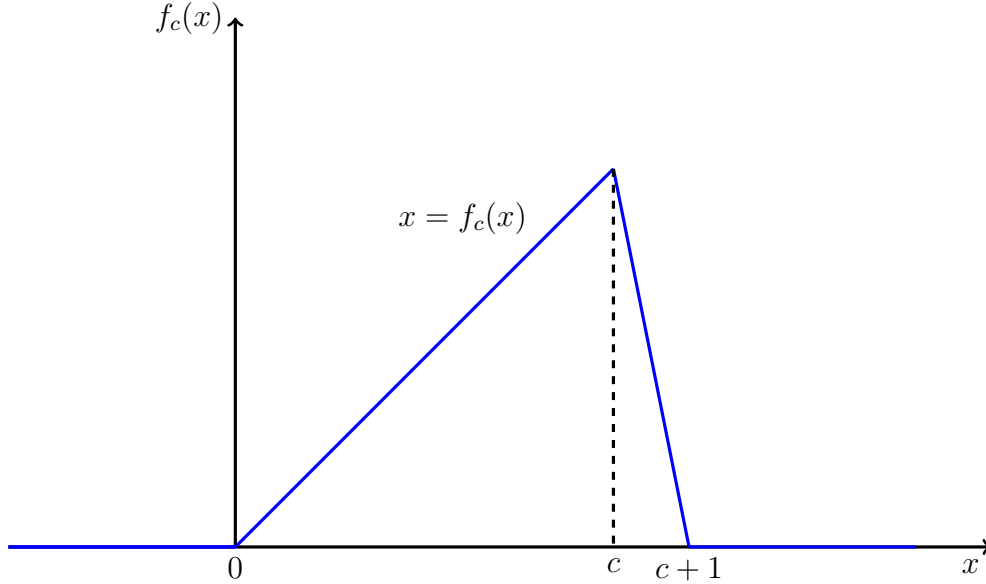
То есть модуль среднего на бесконечности в обе стороны близок к нулю.

**Теорема 5.** Пусть  $\xi_n \xrightarrow{d} \xi, \xi_n \geq 0$  — все величины с конечным матожиданием, тогда  $E[\xi_n] \rightarrow E[\xi]$  тогда и только тогда, когда последовательность  $\{\xi_n, n \in \mathbb{N}\}$  равномерно интегрируема.



*Доказательство.* Докажем в одну сторону, в другую оставим в качестве домашней задачи.

Пусть последовательность равномерно интегрируема. Тогда для  $c > 0$  рассмотрим функцию  $f_c(x)$ :



То есть до  $0$  эта функция тождественно ноль, потом ведет себя так же, как и аргумент до  $c$ , потом резко убывает и снова становится нулем. Мы вводим такую функцию, чтобы приблизить  $f(x) = x$ , но ограниченным образом.

Заметим, что  $f_c(x)$  непрерывна и ограничена. Оценим  $|E[\xi_n] - E[\xi]|$ . Добавим и вычтем каждое из двух выражений под модулем  $E[f_c(\xi_n)]$ ,  $E[f_c(\xi)]$  и воспользуемся неравенством треугольника:

$$|E[\xi_n] - E[\xi]| \leq |E[\xi_n] - E[f_c(\xi_n)]| + |E[\xi] - E[f_c(\xi)]| + |E[f_c(\xi_n)] - E[f_c(\xi)]|$$

Зафиксируем произвольное  $\varepsilon > 0$ . Первое слагаемое не больше  $E[\xi_n I\{\xi_n \geq c\}]$ , так как при  $\xi_n \leq c$ , величины просто совпадают, от  $c$  до  $c+1$  будет какой-то остаток, не больший, чем  $E[\xi_n I\{c \leq \xi_n \leq c+1\}]$ , а дальше уже просто  $E[f_c(\xi_n)]$  не даёт никакого вклада. Откуда  $E[\xi_n I\{\xi_n \geq c\}] \leq \frac{\varepsilon}{2}$  при всех  $c > c_0(\varepsilon)$ . Эта оценка работает по условию теоремы при всех  $n$ .

Второе слагаемое для всех  $c > c_1(\varepsilon)$  тоже не больше, чем  $\frac{\varepsilon}{2}$ , так как мы приближаем  $f_c(x)$  положительную функцию (и при  $c \rightarrow +\infty$   $f_c(x) \rightarrow x$ ) и матожидание конечно, то с какого-то момента разность станет очень маленькой.

Третье слагаемое стремится к нулю с ростом  $n$  из-за эквивалентного определения сходимости по распределению.

Поэтому получаем, что

$$\forall \varepsilon > 0 \quad \overline{\lim}_{n \rightarrow +\infty} |E[\xi_n] - E[\xi]| \leq \varepsilon$$

Откуда есть предел и равен он нулю, что и требовалось доказать.  $\square$

Но теперь, чтобы наконец-то понять асимптотику среднего в нуле, давайте поймём, какие условия «полегче» нужно наложить, чтобы последовательность была равномерно интегрируемой. Это можно сделать из следующей леммы:

**Лемма.** Пусть  $\{\xi_n, n \in \mathbb{N}\}$  — последовательность случайных величин. Если для всех  $n \in \mathbb{N}$  выполнено, что  $E[\xi_n^2] \leq C < +\infty$ , тогда последовательность равномерно интегрируема.

*Доказательство.* Для любого  $\varepsilon > 0$  выберем  $t_0$  такое, что  $t_0 > \frac{C}{\varepsilon}$ . Тогда для любого  $t > t_0$  и любого  $n \in \mathbb{N}$

$$E[|\xi_n| I\{|\xi_n| \geq t\}] \leq E\left[\frac{\xi_n^2}{t} I\{|\xi_n| \geq t\}\right] \leq \frac{C}{t} \leq \frac{C}{t_0} < \varepsilon$$

Значит предел есть и равен нулю. □

Теперь сформулируем основную теорему:

**Теорема 6.**

$$E[L_n] \sim \sqrt{\frac{2n}{\pi}}$$

*Доказательство.* По последней лемме из прошлой лекции имеем  $E[L_n] = E[|S_{n+1}|]$ .

А также мы выяснили, что

$$\frac{|S_n|}{\sqrt{n}} \xrightarrow{d} |\mathcal{N}(0, 1)|$$

Посмотрим на второй момент левой части и вспомним, что среднее у  $S_n$  равно нулю:

$$E\left[\frac{S_n^2}{n}\right] = \frac{D[S_n]}{n} = 1$$

Последнее равенство следует из того, что дисперсия независимых величин равна сумме дисперсий.

Получаем, что второй момент всегда ограничен единицей, значит по теореме 5 мы получаем, что

$$E\left[\frac{|S_n|}{\sqrt{n}}\right] \rightarrow E[|\mathcal{N}(0, 1)|] = \sqrt{\frac{2}{\pi}}$$

Где последнее равенство легко проверяется интегрированием (см одно из ДЗ обычного курса).

Получаем, что  $E[L_n] \sim \sqrt{\frac{2n}{\pi}}$  □

### 1.3.2 Геометрия траекторий. Закон повторного логарифма

Начнём сразу с теоремы, потом будем пояснять её смысл и постепенно доказывать.

**Теорема 7** (Закон повторного логарифма). Пусть  $\{S_n, n \in \mathbb{Z}_+\}$  — простое симметричное случайное блуждание.

Тогда

$$P\left(\overline{\lim}_{n \rightarrow +\infty} \frac{S_n}{\sqrt{2n \ln \ln n}} = 1\right) = 1$$

Выведем маленькое следствие из этого:

*Примечание.* Докажем, что из ЗПЛ следует следующий факт:

$$\mathbb{P}\left(\lim_{n \rightarrow +\infty} \frac{S_n}{\sqrt{2n \ln \ln n}} = -1\right) = 1$$

Ну это легко понять, если заменить  $S_n = -X_n$  и надо лишь осознать, что  $\{X_n, n \in \mathbb{Z}_+\}$  тоже случайное блуждание, но это совершенно ясно из определения.

Давайте будем пояснять смысл.

Закон повторного логарифма занимает промежуточное положение между законом больших чисел и центральной предельной теоремой. Мы знаем, что

$$\frac{S_n}{n} \xrightarrow{\mathbb{P}} 0, \frac{S_n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1)$$

Центральная предельная теорема утверждает, что суммы  $S_n$  с делителем  $\sqrt{n}$  сходятся к стандартному нормальному распределению, и эта последовательность сумм не сходится к какой-либо конкретной величине ни по вероятности, ни почти наверное, а бесконечно блуждает.

Таким образом величина  $S_n/\sqrt{2n \ln \ln n}$  будет к любой точке отрезка  $[-1, 1]$  бесконечное число раз приближаться сколь угодно близко почти наверное.

Так же ЗПЛ означает, что с вероятностью один график блуждания лежит между  $(1 + \varepsilon)\sqrt{2n \ln \ln n}$  и  $-(1 + \varepsilon)\sqrt{2n \ln \ln n}$  для любого  $\varepsilon > 0$  и бесконечное число раз выходит за пределы  $(1 - \varepsilon)\sqrt{2n \ln \ln n}$  и  $-(1 - \varepsilon)\sqrt{2n \ln \ln n}$ .

Вспомним лемму Бореля-Кантеля, а сам ЗПЛ докажем на следующей лекции.

**Определение 12.** Пусть  $\{A_n, n \in \mathbb{N}\}$  – события. Тогда событием  $\{A_n \text{ беск. число}\}$  (или  $\{A_n \text{ б. ч.}\}$ ) называют  $\bigcap_{n=1}^{+\infty} \left( \bigcup_{m \geq n} A_m \right)$ . Событие состоит в том, что произошло бесконечное число событий.

**Теорема 8** (Лемма Бореля-Кантелли). Пусть  $\{A_n, n \in \mathbb{N}\}$ .

- 1) Если  $\sum_{n=1}^{+\infty} \mathbb{P}(A_n)$  сходится, тогда  $\mathbb{P}(\{A_n \text{ б. ч.}\}) = 0$ .
- 2) Если  $\sum_{n=1}^{+\infty} \mathbb{P}(A_n)$  расходится и  $A_n$  независимые величины, тогда  $\mathbb{P}(\{A_n \text{ б. ч.}\}) = 1$ .

*Доказательство.* 1)

$$\mathbb{P}(\{A_n \text{ б. ч.}\}) = \mathbb{P}\left(\bigcap_{n=1}^{+\infty} \left( \bigcup_{m \geq n} A_m \right)\right) = \lim_{n \rightarrow +\infty} \mathbb{P}\left(\bigcup_{m \geq n} A_m\right) \leq \lim_{n \rightarrow +\infty} \sum_{m \geq n} \mathbb{P}(A_m) = 0$$

Последнее равенство верно, так как остаток сходящегося ряда стремится к нулю.

2) Тут уже напрямую не получится, надо провести более тонкий анализ:

$$\mathbb{P}(\{A_n \text{ б. ч.}\}) = \lim_{n \rightarrow +\infty} \mathbb{P}\left(\bigcup_{m \geq n} A_m\right) = 1 - \lim_{n \rightarrow +\infty} \mathbb{P}\left(\bigcap_{m \geq n} \overline{A_m}\right) =$$

Хочется воспользоваться независимостью, но в данном случае надо применять один трюк, чтобы это стало возможным. Поставим повторный предел:

$$= 1 - \lim_{n \rightarrow +\infty} \lim_{N \rightarrow +\infty} \mathbf{P} \left( \bigcap_{m=n}^N \overline{A_m} \right) =$$

Вот теперь уже можно пользоваться независимостью событий (а значит и их дополнений).

$$1 - \lim_{n \rightarrow +\infty} \lim_{N \rightarrow +\infty} \prod_{m=n}^N (1 - \mathbf{P}(A_m))$$

Применим неравенство, что  $1 - x \leq e^{-x}$  и поэтому можно написать, что

$$1 - \lim_{n \rightarrow +\infty} \lim_{N \rightarrow +\infty} \prod_{m=n}^N (1 - \mathbf{P}(A_m)) \geq 1 - \lim_{n \rightarrow +\infty} \lim_{N \rightarrow +\infty} e^{-\sum_{m=n}^N \mathbf{P}(A_m)}$$

Но остаток расходящегося ряда стремится к  $+\infty$ , поэтому имеем равенство

$$= 1 - \lim_{n \rightarrow +\infty} e^{-\infty} = 1$$

□

## 1.4 Лекция от 14.02.2017

### 1.4.1 Закон повторного логарифма

В этой лекции докажем этот закон и поймём, что обычной техникой оценивания через ЦПТ и некоторые другие теоремы вообще не работают.

Вспомним, что мы хотим глобально. Введем 2 события для произвольного  $\varepsilon > 0$ :

$$A_n = \{S_n \geq (1 + \varepsilon)\sqrt{2n \ln \ln n}\}, B_n = \{S_n \leq (1 - \varepsilon)\sqrt{2n \ln \ln n}\}$$

Чтобы доказать утверждение теоремы, надо показать, что

$$\mathbf{P}(\{A_n \text{ б.ч.}\}) = 0; \mathbf{P}(\{B_n \text{ б.ч.}\}) = 1$$

### 1.4.2 Следствие из ЦПТ и теоремы Берри-Эссеена

Если вспомнить ЦПТ и теорему Берри-Эссеена, тогда (учитывая, что третий момент конечен и  $\mathbf{E}[\xi_1] = 0, \mathbf{D}[\xi_1] = 1$ ) получаем, что

$$\sup_{x \in \mathbb{R}} \left| \mathbf{P} \left( \frac{S_n - n \mathbf{E}[\xi_1]}{\sqrt{n \mathbf{D}[\xi_1]}} < x \right) - \Phi_{\mathcal{N}(0,1)}(x) \right| \leq \frac{1}{2} \frac{\mathbf{E}[|\xi_1 - \mathbf{E}[\xi_1]|^3]}{\sqrt{n \mathbf{D}[\xi_1]}^{3/2}} = \mathcal{O} \left( \frac{1}{\sqrt{n}} \right)$$

Докажем одну лемму, которая нам в будущем пригодится, а именно она относительно неплохо оценивает функцию распределения нормального стандартного распределения:

**Лемма.** При всех достаточно больших  $x$  (скажем,  $x > 1$ ) выполняется

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{(x+1)^2}{2}} \leq 1 - \Phi(x) = \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \leq \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

*Доказательство.* Оценить снизу совсем просто. Действительно, интеграл убывает экспоненциально, поэтому основная его часть концентрируется около  $x$ :

$$\int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \geq \int_x^{x+1} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \geq \frac{1}{\sqrt{2\pi}} e^{-\frac{(x+1)^2}{2}}$$

Где последнее неравенство следует из того, что функция под интегралом не меньше, чем написанная в выражении. По-другому можно воспользоваться теоремой о среднем и показать оценку уже напрямую.

Оценить сверху немного сложнее. Заметим, что при  $x > 1$  будет выполнено следующее неравенство для всех  $y > x$ :

$$y - \frac{y^2}{2} < x - \frac{x^2}{2}$$

В этом легко убедиться, так как у этого квадратного сравнения будут корни  $y_1 = x, y_2 = 2 - x$ , но при  $x > 1$  будет выполнено  $y_1 > y_2$ , а мы знаем, что  $y > x$ , поэтому действительно парабола будет принимать положительное значение.

А теперь давайте оценивать интеграл с помощью сравнений функций и обычного интегрирования:

$$\int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{y - \frac{y^2}{2} - y} dy \leq e^{x - \frac{x^2}{2}} \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-y} dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

□

Из теоремы Берри-Эссеена можно даже написать равенство (записать остаток в О-большое) для  $P(S_n \geq t)$  при подстановке  $x = \frac{t}{\sqrt{n}}$ :

$$P(S_n \geq t) = 1 - \Phi\left(\frac{t}{\sqrt{n}}\right) + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$$

Воспользуемся доказанным неравенством, получим, что:

$$P(S_n \geq t) = e^{-\frac{t^2}{2n}(1+o(1))} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$$

Мы знаем, что при  $t = \sqrt{2n \ln \ln n}$  следует, что  $x = \frac{t}{\sqrt{n}} > 1$ , поэтому неравенством мы корректно воспользовались, а в  $o(1)$  записали всё ненужное.

Теперь подставим наше  $t$

$$P(S_n \geq t) \sim \frac{1}{(\ln n)^{1+o(1)}} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$$

Но если мы хотим воспользоваться леммой Бореля-Кантелли, нам надо или чтобы события были независимы или ряд сходил. Но у нас тут ряд логарифмов! А события очевидно все зависимы. Ряд расходится, какую там степень бы не написать. Да даже больше — остаток расходится! Плохо, нужна другая техника, чтобы доказать ЗПЛ.

### 1.4.3 Доказательство ЗПЛ

Следующая лемма показывает, насколько максимально мы можем уйти. Точнее связь между

**Теорема 9.** Пусть  $\{\xi_n, n \in \mathbb{N}\}$  независимые одинаково распределенные случайные величины с симметричным распределением (то есть  $\xi_k \stackrel{d}{=} -\xi_k$ ), а  $S_n = \xi_1 + \dots + \xi_n$ . Тогда  $\forall a > 0$ :

$$P\left(\max_{k \leq n} S_k \geq a\right) \leq 2P(S_n \leq a)$$

*Доказательство.* Давайте введем все нужные события:

$$A = \{\max_{k \leq n} S_k \geq a\}$$

Понятное дело, что без  $A$  не обойтись, если мы хотим доказать теорему. Аналогично не обойтись без  $B$ :

$$B = \{S_n \geq a\}$$

А теперь давайте попытаемся представить  $A$  в виде дизъюнктивного объединения каких-то событий. Для этого часто в теории вероятностей вводят события первых моментов:

$$A_k = \{S_1 < a, \dots, S_{k-1} < a, S_k \geq a\}$$

Действительно,  $A = \bigsqcup_{k=1}^n A_k$ , так как  $A_k$  не пересекаются и образуют всё  $A$ .

Также давайте поймём следующее включение:

$$A_k \cap \{S_n - S_k \geq 0\} \subseteq A_k \cap B$$

Действительно, если уж  $S_k \geq a$ , то если  $S_n \geq S_k$ , тогда и  $S_n \geq a$ , то есть все события слева включены в правое. Но чем же хорошо это включение? Да тем, что слева независимые события, так как  $A_k$  никак не зависит от  $\xi_{k+1}, \dots, \xi_n$ . Это нам пригодится.

Ура, у нас уже есть какие-то включения, давайте уже что-то оценивать. Будем аккуратно расписывать наше неравенство:

$$P(B) \geq \sum_{k=1}^n P(B \cap A_k)$$

Действительно, мы просто пересекаем событие  $B$  с непересекающимися между собой  $A_k$ . По включению выше мы получаем, что

$$\sum_{k=1}^n P(B \cap A_k) \geq \sum_{k=1}^n P(A_k \cap \{S_n - S_k \geq 0\})$$

Из-за независимости событий получаем, равенство:

$$\sum_{k=1}^n P(A_k \cap \{S_n - S_k \geq 0\}) = \sum_{k=1}^n P(A_k) P(\{S_n - S_k \geq 0\})$$

Но вторая вероятность не меньше  $\frac{1}{2}$  из-за симметричности распределения и тем, что ещё может достигаться равенство. Поэтому последнее неравенство:

$$\sum_{k=1}^n \mathbf{P}(A_k) \mathbf{P}(\{S_n - S_k \geq 0\}) \geq \frac{1}{2} \sum_{k=1}^n \mathbf{P}(A_k) = \frac{1}{2} \mathbf{P}(A)$$

Где последнее равенство из-за дизъюнктивности объединения  $A_k$ .  $\square$

Сейчас мы имеем весь арсенал, чтобы доказать ЗПЛ.

*Доказательство.* Сначала докажем, что  $\mathbf{P}(\{A_n \text{ б.ч.}\}) = 0$ . Для этого введем некоторые обозначения при фиксированном  $\varepsilon > 0$ :

$$\left\{ \begin{array}{l} \varepsilon \in (0, 1] \\ \lambda = 1 + \varepsilon \\ n_k = \lambda^k \\ k_0 : k \geq k_0, \text{ что } \ln \ln k > 0 \\ C_k = \bigcup_{\substack{n > n_{k-1}, \\ n \leq n_k}} A_k \end{array} \right.$$

Несложно понять, что  $\{A_n \text{ б.ч.}\}$  совпадает с  $\{C_k \text{ б.ч.}\}$ .

Будем считать  $\mathbf{P}(C_k)$ :

$$\mathbf{P}(C_k) \leq \mathbf{P}\left(\max_{n \leq n_k} S_n \geq \lambda \sqrt{2n_{k-1} \ln \ln n_{k-1}}\right)$$

Действительно, мы смотрим, что хотя бы один до  $n_k$  (знак  $\leq$  из-за того, что есть ещё целые числа от  $1, \dots, n_{k-1}$ ) больше нужного значения. По теореме 9 мы получаем, что

$$\mathbf{P}(C_k) \leq 2 \mathbf{P}\left(S_{[n_k]} \geq \lambda \sqrt{2n_{k-1} \ln \ln n_{k-1}}\right)$$

А теперь вспомним оценку нашего интеграла (да, тут она уже работает):

$$2 \mathbf{P}\left(S_{[n_k]} \geq \lambda \sqrt{2n_{k-1} \ln \ln n_{k-1}}\right) \leq 2 \exp\left(\frac{-\lambda^2 2n_{k-1} \ln \ln n_{k-1}}{2[n_k]}(1 + o(1))\right) + \mathcal{O}\left(\frac{1}{\sqrt{n_k}}\right)$$

Также заметим, что  $\frac{n_{k-1}}{[n_k]} = \frac{1}{\lambda}(1 + o(1))$ , поэтому это можно внести под то о-малое. А  $\ln \ln n_{k-1} = \ln k + o(1)$ .

Поэтому это всё эквивалентно:

$$= 2 \exp((- \lambda \ln k)(1 + o(1))) + \mathcal{O}\left(\frac{1}{\sqrt{\lambda^k}}\right) = 2k^{-(1+\varepsilon)(1+o(1))} + \mathcal{O}\left(\frac{1}{\sqrt{\lambda^k}}\right)$$

Что несомненно сходится. По лемме Бореля-Кантелли получаем, что как раз вероятность  $\mathbf{P}(C_k) \leq 2k^{-(1+\varepsilon)(1+o(1))} + \mathcal{O}\left(\frac{1}{\sqrt{\lambda^k}}\right)$ , поэтому ряд из  $\mathbf{P}(C_k)$  будет сходиться, поэтому и  $\mathbf{P}(\{C_k \text{ б.ч.}\}) = 0$ , и  $\mathbf{P}(\{A_n \text{ б.ч.}\}) = 0$ .

Для второго подпункта введём свои обозначения при фиксированном  $\varepsilon > 0$ :

$$\begin{cases} n_k = N^k, \text{ } N \text{ мы определим чуть позднее} \\ \lambda = 1 - \varepsilon \\ \varepsilon \in (0, 1) \end{cases}$$

Рассмотрим последовательность  $\{-S_n, n \in \mathbb{N}\}$ . По пункту один мы выяснили, что  $D_k = \{-S_{n_k} \leq 2\sqrt{2n_k \ln \ln n_k}\}$  происходит с вероятностью один конечное число раз.

Поэтому давайте выпишем событие  $B_{n_k}$ . Если мы докажем для какой-то подпоследовательности, что вероятность бесконечного числа равна единице, то тогда понятно, что и для всей последовательности оно будет с вероятностью один.

$$\begin{aligned} Q_k = \{S_{n_k} \geq \lambda\sqrt{2n_k \ln \ln n_k}\} &= \{S_{n_k} - S_{n_{k-1}} \geq \lambda\sqrt{2n_k \ln \ln n_k} - S_{n_{k-1}}\} \supseteq \\ &\supseteq \{S_{n_k} - S_{n_{k-1}} \geq \lambda\sqrt{2n_k \ln \ln n_k} + 2\sqrt{2n_{k-1} \ln \ln n_{k-1}}\} \end{aligned}$$

Где последнее включение следует как раз из-за того, что с вероятностью один мы с какого-то момента имеем корректное неравенство. И если мы докажем для последнего события, что оно бесконечное число раз выполняется с вероятностью один, то автоматически докажем теорему.

Заметим, что

$$\lambda\sqrt{2n_k \ln \ln n_k} + 2\sqrt{2n_{k-1} \ln \ln n_{k-1}} \leq \lambda\sqrt{2n_k \ln \ln n_k} \left(1 + 2\sqrt{\frac{1}{N\lambda^2}}\right) = \lambda'\sqrt{2n_k \ln \ln n_k}$$

Где  $\lambda' \in (\lambda, 1)$ . Такое  $\lambda'$  можно подобрать, если взять  $N$  очень большое. Поэтому усилим оценку и будем доказывать, что будет вероятность один у б.ч. у события

$$\{S_{n_k} - S_{n_{k-1}} \geq \lambda'\sqrt{2n_k \ln \ln n_k}\}$$

Заметим, что все  $Q_k$  независимы, так как не пересекаются по  $\xi_i$ , значит уже есть вера в то, что можно воспользоваться леммой Бореля-Кантелли.

Обозначим за  $Y_k = S_{n_k} - S_{n_{k-1}}$ . Тогда

$$P(Q_k) \geq P(Y_k \geq \lambda'\sqrt{2n_k \ln \ln n_k}) = \exp\left(\frac{-\lambda'^2 2n_k \ln \ln n_k}{2(n_k - n_{k-1})}(1 + o(1))\right) + \mathcal{O}\left(\frac{1}{\sqrt{N^k}}\right) =$$

А теперь давайте всё подряд оценивать.  $\ln \ln n_k = \ln k + \ln \ln N = \ln k + o(1)$  при достаточно больших  $k$  ( $N$  зависит только от  $\varepsilon$ ).

$$= \exp\left(\frac{-\lambda'^2 \ln k}{1 - \frac{1}{N}}(1 + o(1))\right) + \mathcal{O}\left(\frac{1}{\sqrt{N^k}}\right) =$$

А теперь давайте поймём, что  $\frac{\lambda'^2}{1 - \frac{1}{N}} < 1$  можно сделать. Действительно, при  $N \rightarrow +\infty$  следует, что  $\lambda' \downarrow \lambda$ , а значит при увеличении  $N$  выражение стремится к числу, меньшему единице, значит с какого-то момента будет меньше единицы. Отлично! Пусть  $\frac{\lambda'^2}{1 - \frac{1}{N}} = \delta < 1$ . Тогда получаем:



$$= k^{-(1-\delta)(1+o(1))} + \mathcal{O}\left(\frac{1}{\sqrt{N^k}}\right)$$

Что расходится, как сумма сходящегося и расходящегося рядов. События независимы, поэтому по лемме Бореля-Кантелли вероятность б.ч. равна единице.  $\square$

Закон повторного логарифма, конечно, работает и при других ограничениях на  $\xi_i$ . Но нам хватило, как мы считаем, и для случайного блуждания. Оставим общий ЗПЛ без доказательства.

**Теорема 10.** Пусть  $\{\xi_n, n \in \mathbb{N}\}$  независимые одинаково распределенные случайные величины со средним ноль и положительной дисперсией  $\sigma^2$ . Тогда

$$\mathbb{P}\left(\overline{\lim}_{n \rightarrow +\infty} \frac{S_n}{\sqrt{2n\sigma^2 \ln \ln n}} = 1\right) = 1$$

На этом со случайным блужданием мы заканчиваем.

#### 1.4.4 Ветвящиеся случайные процессы

#### 1.4.5 Физическая модель

Физическая модель у этих процессов достаточно естественная. Сначала есть один человек. С какой-то вероятностью он порождает потомков. Потом каждый потомок независимо от остальных с тем же распределением порождает ещё несколько. Все предыдущие поколения умирают. Так повторяется либо пока не останется ни одного потомка, либо бесконечно. Можно считать, что так мы смотрим, вымрет ли род когда-нибудь, насколько долго он будет жить и насколько он будет широким.

#### 1.4.6 Математическая модель

**Определение 13.** Пусть  $\xi$  — случайная величина со значениями в  $\mathbb{Z}_+$  (называемый *законом размножения*). Пусть  $\xi_k^{(n)}$  одинаково распределенные случайные величины с распределением, как у  $\xi$ .

Тогда введём  $X_n$  — число частиц в  $n$ -ом поколении. Тогда  $X_0 = 1$ , а  $X_n = \sum_{k=1}^{X_{n-1}} \xi_k^{(n)}$ .

Фактически,  $\xi_k^{(n)}$  отвечает за число потомков  $k$ -ой частицы в  $n - 1$ -ом поколении.

Такой процесс называется *процессом Гамильтона-Ватсона с законом размножения частиц*  $\xi$ .

Ясно, что нас интересуют изначально вопрос о вырождении, то есть  $\mathbb{P}(\exists n : X_n = 0)$ . Анонсируем сразу ответ, на следующей лекции докажем

Если  $\mathbb{E}[\xi] \leq 1$ , тогда  $\mathbb{P}(\exists n : X_n = 0) = 1$ , кроме случая  $\xi = 1$ .

Если  $\mathbb{E}[\xi] > 1$ , тогда  $\mathbb{P}(\exists n : X_n = 0) < 1$ .