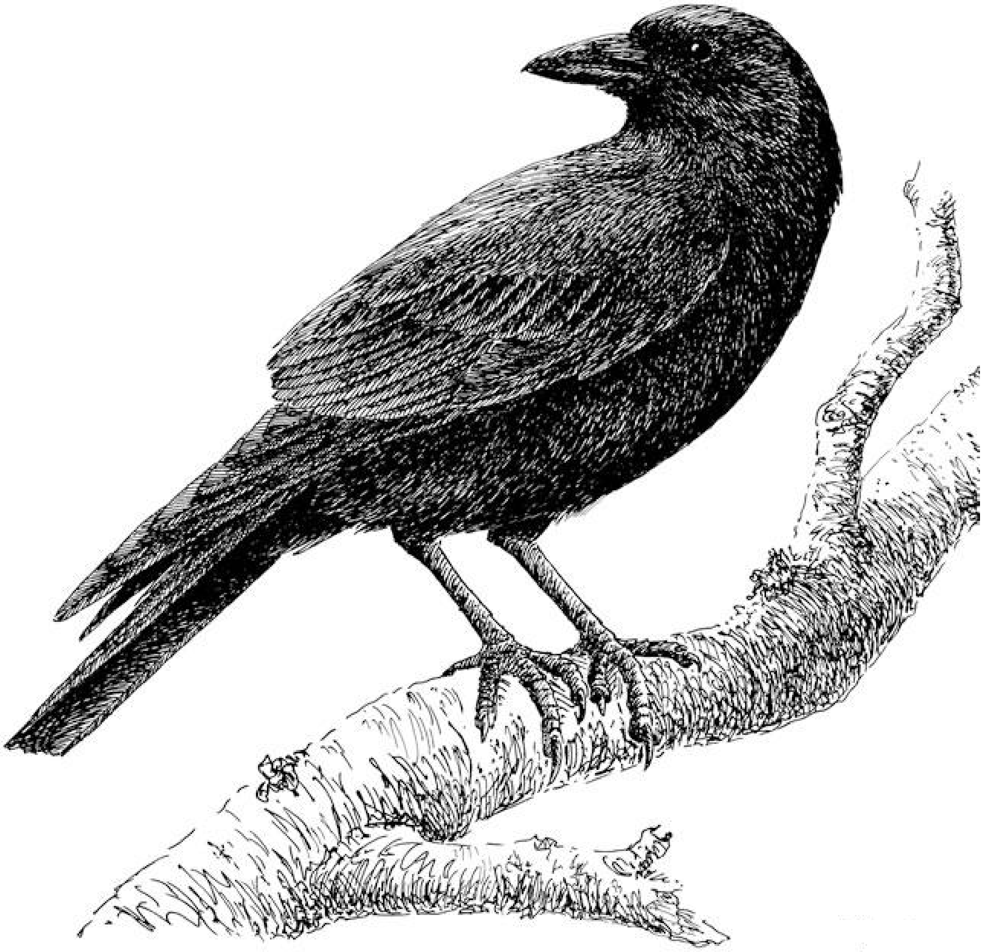


Теория вероятностей

Конспекты лекций и семинаров

Лектор: Д.А. Шабанов

Конспекты вели Денис Беляков, Никита Попов и Алексей Хачиянц

НИУ ВШЭ, 2016-2017

Оглавление

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.

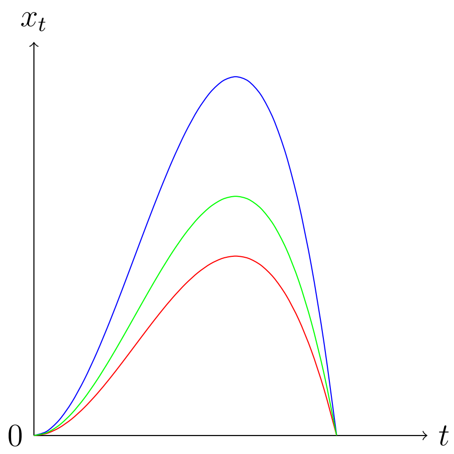
Определение 2. Классифицируем случайные процессы:

- 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 $(\Omega, \mathcal{F}, \mathsf{P})$ и все $x_t : \Omega \to \mathbb{R}$.

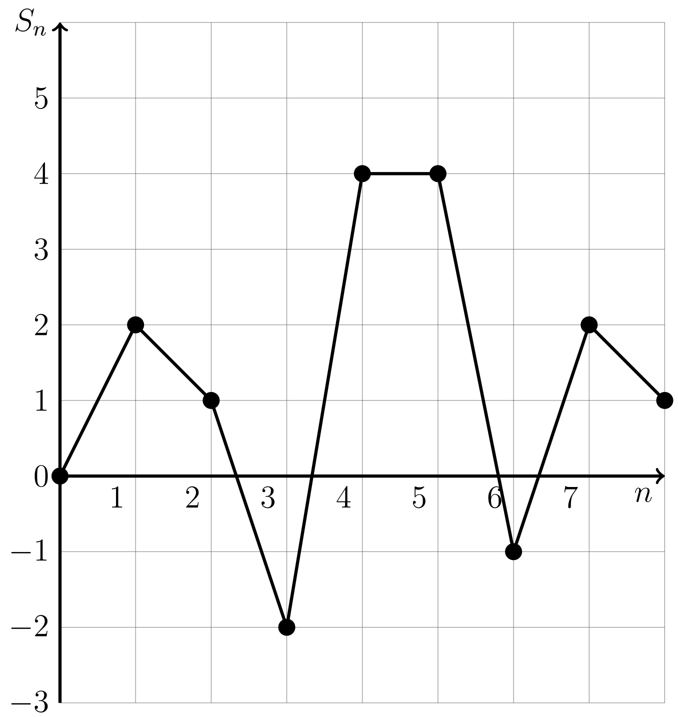
Определение 3. Для фиксированного $\omega_0 \in \Omega$ функция $\tilde{x}_t(t) = x_t(\omega_0), t \in T$ является *траекторией* (или *реализацией*) случайного процесса.

Пример. $x_t = f(t)\xi$, где $f(t)$ — какая-то функция, ξ — случайная величина. Приведем пример траекторий для некоторых ω_0.



1.1.2 Случайные блуждания

Определение 4. Пусть $\{\xi_n, \in \mathbb{N}\}$ — 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.*



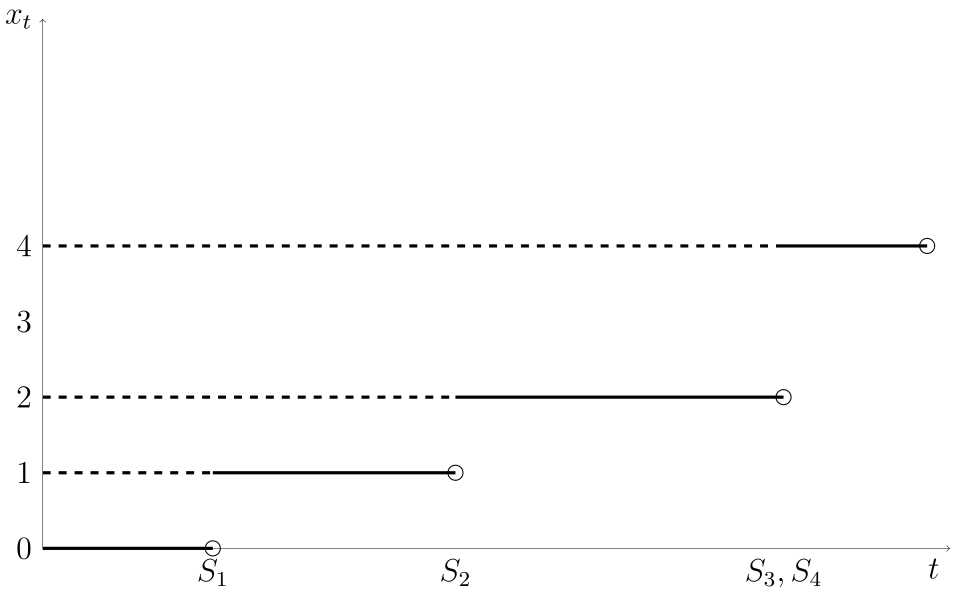
Пример. Физической моделью, соответствующую случайному блужданию, могут являться прыжки кузнечика.

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 $t \geqslant 0$ рассмотрим такие случайные величины: $x_t = \max\{n : S_n \leqslant t\}$ (если максимума не существует, положим $x_t = +\infty$).

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

График можно описать как-то так для какого-то ω_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 ξ_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* $\mathsf{P}(\xi_i = 0) < 1$.

Доказательство. Пусть $t > 0$ фиксировано. $\mathsf{P}(x_t = +\infty) = \mathsf{P}(\forall n : S_n \leqslant t)$.

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

Выпишем тривиальное равенство пределов:

$$\lim_{n\to+\infty} \mathsf{P}(S_n \leqslant t) = \lim_{n\to+\infty} \mathsf{P}\left(\frac{S_n}{n} \leqslant \frac{t}{n}\right)$$

Есть 2 случая:

- $\mathsf{E}[\xi_i]$ конечно и равно $A > 0$ (больше нуля, так как $\mathsf{P}(\xi_i = 0) < 1$). Тогда по закону больших чисел имеем, что $\frac{S_n}{n} \xrightarrow{\mathsf{P}} A$ (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:

$$= \lim_{n\to+\infty} \mathsf{P}\left(\frac{S_n}{n} \leqslant \frac{t}{n}\right) \leqslant \lim_{n\to+\infty} \mathsf{P}\left(\frac{S_n}{n} \leqslant \frac{A}{2}\right)$$

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

$$= \lim_{n \to +\infty} \mathsf{P}\left(A \leqslant \frac{A}{2}\right) = 0.$$

Последнее равенство идёт из-за неотрицательности A.

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

- $\mathsf{E}[\xi_i] = +\infty$.

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

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

  А значит, что $\mathsf{P}(S_n \leqslant t) \leqslant \mathsf{P}\left(\hat{S}_n \leqslant t\right)$ (так как $S_n \geqslant \hat{S}_n$, а значит, что $\hat{S}_n$ принимает меньшие значения).

  Но мы уже всё доказали для конечного матожидания $\hat{\xi}_i$, поэтому получаем, что $\mathsf{P}\left(\hat{S}_n \leqslant t\right) \to 0$, откуда и $\mathsf{P}(S_n \leqslant t) \to 0$, 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 $x_t = \max\{n : \xi_1 + \ldots + \xi_n \leqslant t\}$.

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

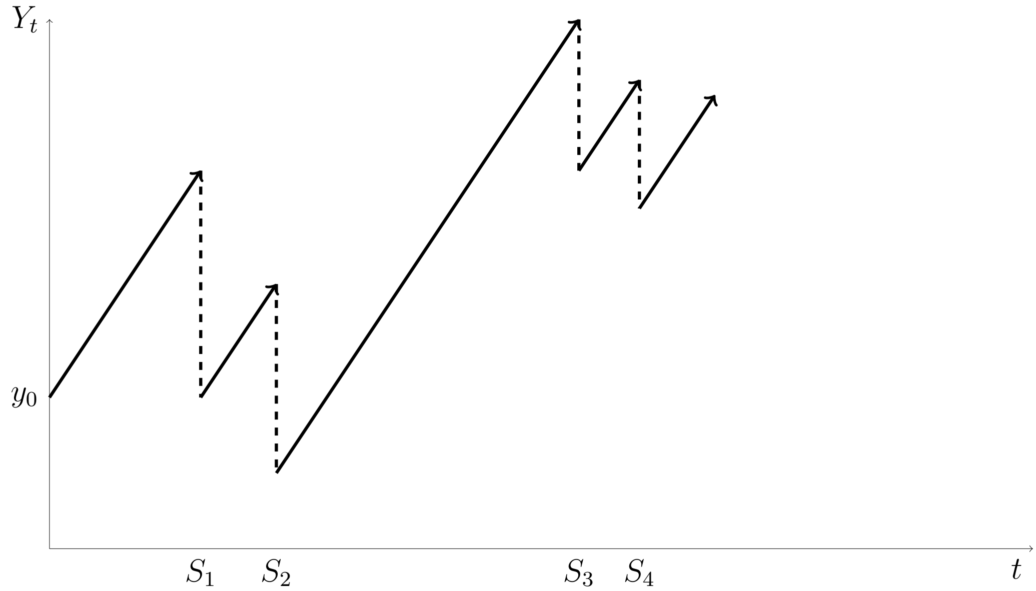
$$Y_t = y_0 + ct - \sum_{k=1}^{x_t} \eta_k$$

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.
- η_k — размер выплаты с номером k.
- ξ_k — время между $k-1$-й и k-й выплатой.
- x_t — количество выплат к моменту времени t.
- $\sum\limits_{k=1}^{x_t} \eta_k$ — 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 $p \in [0,1]$

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

Положим $S_0 = 0, S_n = \xi_1 + \ldots \xi_n, n \in \mathbb{N}$. 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 $\mathsf{E}[S_n] = (p-q)n$. Действительно, ответ можно легко получить, если раскрыть по линейности, и воспользоваться совершенно ясным фактом, который получается из определения, что $\mathsf{E}[\xi_i] = p - q$.

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 $\mathsf{P}(\exists\ n > 0 : S_n = 0)$.

Попробуем найти $\mathsf{P}(S_n = x)$. 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 $n \geqslant |x|$, 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:

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

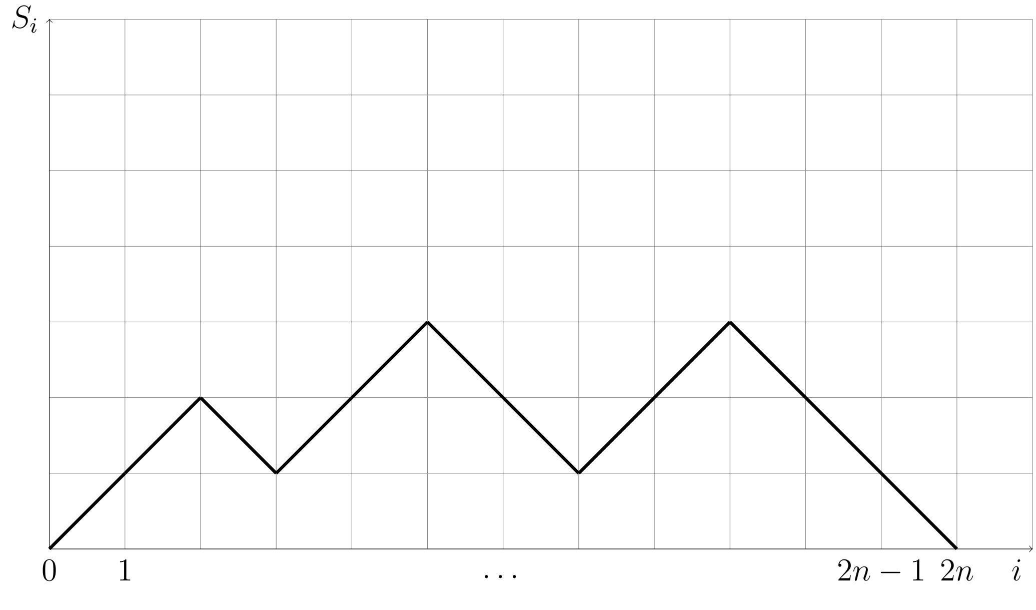
Но этого мало. Мы не сможем как-то легко посчитать вероятность существования n, что $S_n = 0$.

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

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

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

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

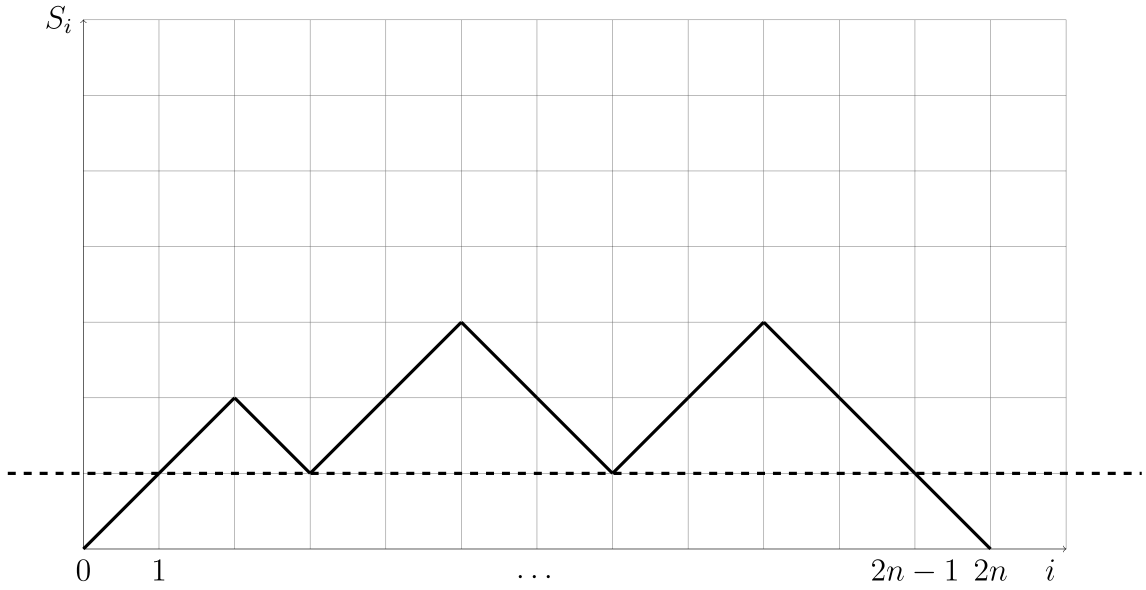
А теперь вспомним, что числа Каталана задаются практически так же. Те, кто ходил на курс дискретной математики, помнят, что есть соответствие между числами Каталана и количеством путей, отвечающим свойствам: $\sum\limits_{i=1}^{2n} \varepsilon_i = 0$ и $\sum\limits_{i=1}^{k} \varepsilon_i \geqslant 0,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

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 нулевая.

В другую сторону доказывается аналогично. См. иллюстрацию.



□

Получается, что

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

И соответственно:

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

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 $C_0 = 1$:

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

$\square$

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

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

Заметим, что $C_n \leqslant 2^{2n}$, 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 сходится и при $|x| = \frac{1}{4}$.

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

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

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 реккурентное 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 $pq \leqslant \frac{1}{4}$.

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

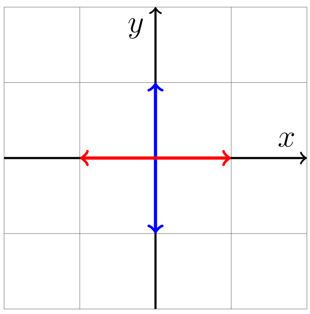
□

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

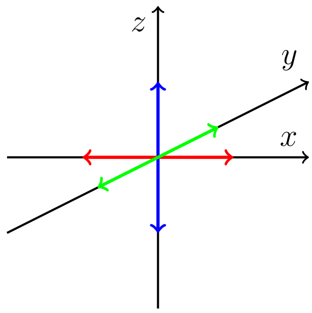
1.2.3 Многомерный случай

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 $\mathsf{P}(S_{2n}=0)$. При $d=1$ и $p=q=\frac{1}{2}$ получим, что $\mathsf{P}(S_{2n}=0)=\binom{2n}{n}4^{-n}\approx\Theta\left(\frac{1}{\sqrt{n}}\right)$, 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 $f(x)=\sum\limits_{n=0}^{+\infty}C_nx^n$.

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

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 $(2n-3)!!$:

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

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. $\mathsf{P}(S_1\neq 0,\dots,S_{2n-1}\neq 0,S_{2n}=0)=2C_{n-1}(pq)^n=\binom{2n}{n}(pq)^n\frac{1}{2n-1}$ — 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. $\mathsf{E}[X] = +\infty$.

Доказательство. Если $p \neq q$, тогда $\mathsf{P}(x = +\infty) > 0$, поэтому матожидание уже точно бесконечность.

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

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 находится 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 $L_n = \left|k : k \in \overline{0, \ldots, n}, S_k = 0\right|$.

Хотим понять, какая асимптотика у $\mathsf{E}[L_n]$.

Лемма. $\mathsf{E}[L_n] = \mathsf{E}[|S_{n+1}|]$

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:

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

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

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

Будем раскрывать рекурсивно, откуда получим:

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

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

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

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

Ну и финальный аккорд состоит в том, что $\mathsf{E}[\xi_{k+1}] = 0$, 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 ξ_i равно нулю и дисперсия равна единице мы имеем следующее:

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

По теореме о наследовании сходимости можно заключить, что

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

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 поведений случайных величин.

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

$$\lim_{c \to +\infty} \sup_{n \in \mathbb{N}} \mathsf{E}[|\xi_n|\, \mathbf{I}\{|\xi_n| \geqslant c\}] = 0$$

Поясним определение.

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

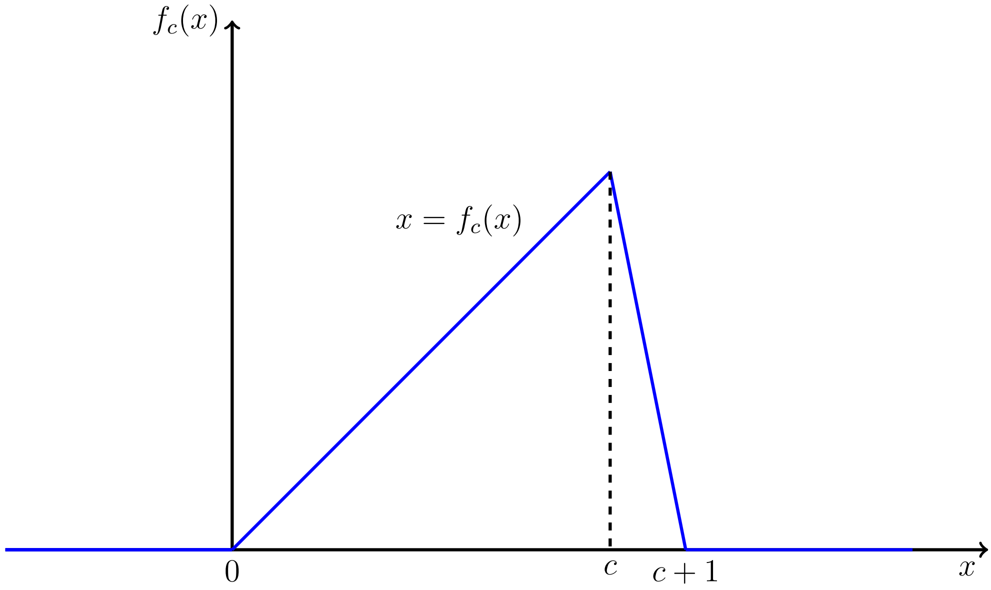
$$\sup_{n \in \mathbb{N}} \int\limits_{|x| \geqslant c} |x| p_{\xi_n}(x) dx \to 0$$

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

Теорема 5. *Пусть $\xi_n \xrightarrow{\mathrm{d}} \xi, \xi_n \geqslant 0$ — все величины с конечным матожиданием, тогда $\mathsf{E}[\xi_n] \to \mathsf{E}[\xi]$ 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 $|\,\mathsf{E}[\xi_n] - \mathsf{E}[\xi]|$. Добавим и вычтем каждое из двух выражений под модулем $\mathsf{E}[f_c(\xi_n)], \mathsf{E}[f_c(\xi)]$ и воспользуемся неравенством треугольника:

$$|\,\mathsf{E}[\xi_n] - \mathsf{E}[\xi]| \leqslant |\,\mathsf{E}[\xi_n] - \mathsf{E}[f_c(\xi_n)]| + |\,\mathsf{E}[\xi] - \mathsf{E}[f_c(\xi)]| + |\,\mathsf{E}[f_c(\xi_n)] - \mathsf{E}[f_c(\xi)]|$$

Зафиксируем произвольное $\varepsilon > 0$. Первое слагаемое не больше $\mathsf{E}[\xi_n\,\mathbf{I}\{\xi_n \geqslant c\}]$, так как при $\xi_n \leqslant c$, величины просто совпадают, от c до $c+1$ будет какой-то остаток, не больший, чем $\mathsf{E}[\xi_n\,\mathbf{I}\{c \leqslant \xi_n \leqslant c+1\}]$, а дальше уже просто $\mathsf{E}[f_c(\xi_n)]$ не даёт никакого вклада. Откуда $\mathsf{E}[\xi_n\,\mathbf{I}\{\xi_n \geqslant c\}] \leqslant \frac{\varepsilon}{2}$ 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 $c \to +\infty$ $f_c(x) \to x$) 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

$$\forall \varepsilon > 0 \varlimsup_{n \to +\infty} |\,\mathsf{E}[\xi_n] - \mathsf{E}[\xi]| \leqslant \varepsilon$$

Откуда есть предел и равен он нулю, что и требовалось доказать. □

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* $\mathsf{E}[\xi_n^2] \leqslant C < +\infty$*, 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 $n \in \mathbb{N}$

$$\mathsf{E}[|\xi_n|\,\mathbf{I}\{|\xi_n| \geqslant t\}] \leqslant \mathsf{E}\left[\frac{\xi_n^2}{t}\,\mathbf{I}\{|\xi_n| \geqslant t\}\right] \leqslant \frac{C}{t} \leqslant \frac{C}{t_0} < \varepsilon$$

Значит предел есть и равен нулю. □

Теперь сформулируем основную теорему:

Теорема 6.

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

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

А также мы выяснили, что

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

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

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

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

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

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

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

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

$$\mathsf{P}\left(\varlimsup_{n\to+\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} = 1\right) = 1$$

Выведем маленькое следствие из этого:

Примечание. Докажем, что из ЗПЛ следует следующий факт:

$$\mathsf{P}\left(\varliminf_{n\to+\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} = -1\right) = 1$$

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

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

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 $\{A_n$ беск. число$\}$ (или $\{A_n$ б. ч.$\}$) называют $\bigcap\limits_{n=1}^{+\infty}\left(\bigcup\limits_{m\geqslant n} A_m\right)$. Событие состоит в том, что произошло бесконечное число событий.

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

1) Если $\sum\limits_{n=1}^{+\infty} \mathsf{P}(A_n)$ *сходится, тогда* $\mathsf{P}(\{A_n$ *б. ч.*$\}) = 0$.

2) Если $\sum\limits_{n=1}^{+\infty} \mathsf{P}(A_n)$ *расходится и* A_n *независимые величины, тогда* $\mathsf{P}(\{A_n$ *б. ч.*$\}) = 1$.

Доказательство. 1)

$$\mathsf{P}(\{A_n \text{ б. ч.}\}) = \mathsf{P}\left(\bigcap_{n=1}^{+\infty}\left(\bigcup_{m\geqslant n} A_m\right)\right) = \lim_{n\to+\infty} \mathsf{P}\left(\bigcup_{m\geqslant n} A_m\right) \leqslant \lim_{n\to+\infty} \sum_{m\geqslant n} \mathsf{P}(A_m) = 0$$

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

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

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

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

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

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

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

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

$$1 - \lim_{n\to+\infty} \lim_{N\to+\infty} \prod_{m=n}^{N} (1 - \mathsf{P}(A_m)) \geqslant 1 - \lim_{n\to+\infty} \lim_{N\to+\infty} e^{-\sum\limits_{m=n}^{N} \mathsf{P}(A_m)}$$

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

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

$\square$

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 $\varepsilon > 0$:

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

Чтобы доказать утверждение теоремы, надо показать, что

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

1.4.2 Следствие из ЦПТ и теоремы Берри-Эссеена

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

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

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*

$$\frac{1}{\sqrt{2\pi}} e^{\frac{-(x+1)^2}{2}} \leqslant 1 - \Phi(x) = \int\limits_{x}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy \leqslant \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

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

$$\int\limits_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy \geqslant \int\limits_x^{x+1} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy \geqslant \frac{1}{\sqrt{2\pi}} e^{\frac{-(x+1)^2}{2}}$$

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:

$$\int\limits_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{y-\frac{y^2}{2}-y}\, dy \leqslant e^{x-\frac{x^2}{2}} \int\limits_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-y}\, dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

□

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

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

Воспользуемся доказанным неравенством, получим, что:

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

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 $o(1)$ запихали всё ненужное.

Теперь подставим наше t

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

Но если мы хотим воспользоваться леммой Бореля-Кантелли, нам надо или чтобы события были независимы, или чтобы ряд сходился. 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 всеми предыдущими значениями блуждания и последнего.

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

$$\mathsf{P}\left(\max_{k \leqslant n} S_k \geqslant a\right) \leqslant 2\,\mathsf{P}(S_n \leqslant a)$$

Доказательство. Давайте введем все нужные события:

$$A = \{\max_{k \leqslant n} S_k \geqslant a\}$$

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

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

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

$$A_k = \{S_1 < a, \ldots, S_{k-1} < a, S_k \geqslant a\}$$

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

Также давайте поймём следующее включение:

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

Действительно, если уж $S_k \geqslant a$, то если $S_n \geqslant S_k$, тогда и $S_n \geqslant a$, 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 $\xi_{k+1}, \ldots, \xi_n$. Это нам пригодится.

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

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

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

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

Из-за независимости событий получаем равенство:

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

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

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

Где последнее равенство из-за дизъюнктности объедения A_k. □

Сейчас мы имеем весь арсенал, чтобы доказать ЗПЛ.

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

$$\begin{cases} \varepsilon \in (0,1] \\ \lambda = 1 + \varepsilon \\ n_k = \lambda^k \\ k_0 : k \geqslant k_0, \text{ что } \ln\ln k > 1 \\ C_k = \bigcup\limits_{\substack{n > n_{k-1}, \\ n \leqslant n_k}} A_k \end{cases}$$

Несложно понять, что $\{A_n \text{ б.ч.}\}$ совпадает с $\{C_k \text{ б.ч.}\}$.
Будем считать $\mathsf{P}(C_k)$:

$$\mathsf{P}(C_k) \leqslant \mathsf{P}\left(\max_{n \leqslant n_k} S_n \geqslant \lambda\sqrt{2n_{k-1}\ln\ln n_{k-1}}\right)$$

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

$$\mathsf{P}(C_k) \leqslant 2\,\mathsf{P}\left(S_{\lfloor n_k \rfloor} \geqslant \lambda\sqrt{2n_{k-1}\ln\ln n_{k-1}}\right)$$

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

$$2\,\mathsf{P}\left(S_{\lfloor n_k \rfloor} \geqslant \lambda\sqrt{2n_{k-1}\ln\ln n_{k-1}}\right) \leqslant 2\exp\left(\frac{-\lambda^2 2n_{k-1}\ln\ln n_{k-1}}{2\lfloor n_k \rfloor}(1+o(1))\right) + \mathcal{O}\left(\frac{1}{\sqrt{n_k}}\right)$$

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

Поэтому это всё эквивалентно:

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

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

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

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

Рассмотрим последовательность $\{-S_n, n \in \mathbb{N}\}$. По пункту один мы выяснили, что $D_k = \{-S_{n_k} \leqslant 2\sqrt{2n_k \ln\ln n_k}\}$ 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.

$$Q_k = \{S_{n_k} \geqslant \lambda\sqrt{2n_k \ln\ln n_k}\} = \{S_{n_k} - S_{n_{k-1}} \geqslant \lambda\sqrt{2n_k \ln\ln n_k} - S_{n_{k-1}}\} \supseteq$$
$$\supseteq \{S_{n_k} - S_{n_{k-1}} \geqslant \lambda\sqrt{2n_k \ln\ln n_k} + 2\sqrt{2n_{k-1} \ln\ln n_{k-1}}\}$$

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

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

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

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

Заметим, что все Q_k независимы, так как не пересекаются по ξ_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

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

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 ε).

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

А теперь давайте поймём, что $\frac{\lambda'^2}{1-\frac{1}{N}} < 1$ можно сделать. Действительно, при $N \to +\infty$ 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 рассходящегося рядов. События независимы, поэтому по лемме Бореля-Кантелли вероятность б.ч. равна единице. □

Закон повторного логарифма, конечно, работает и при других ограничениях на ξ_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 σ^2. Тогда*

$$\mathsf{P}\left(\overline{\lim_{n\to+\infty}} \frac{S_n}{\sqrt{2n\sigma^2 \ln\ln n}} = 1\right) = 1$$

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

Тогда введём X_n — число частиц в n-ом поколении. Тогда $X_0 = 1$, а $X_n = \sum\limits_{k=1}^{X_{n-1}} \xi_k^{(n)}$. Фактически, $\xi_k^{(n)}$ отвечает за число потомков k-ой частицы в $n-1$-ом поколении.

Такой процесс называется *процессом Гальтона-Ватсона с законом размножения частиц ξ*.

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

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

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

1.5 Лекция от 21.02.2017

1.5.1 Производящая функция случайной величины

Несколько лекций назад мы познакомились с производящей функцией последовательности. А теперь давайте введём понятия для произвольной случайной величины:

Определение 14. *Производящей функцией* случайной величины ξ называют функцию $\varphi_\xi(z) = \mathsf{E}\left[z^\xi\right]$ при $z \geqslant 0$.

Начнём сразу говорить о каких-то не очень сложных и одновременно важных свойствах.

1. $\varphi_\xi(1) = 1$. Тривиально.

2. $\varphi'_\xi(1) = \mathsf{E}[\xi]$. Оно нам только понадобиться для дискретного случая, где это свойство легко выводится (см. ниже).

3. Если ξ, η независимы, тогда
$$\varphi_{\xi+\eta}(z) = \varphi_\xi(z) \cdot \varphi_\eta(z).$$
Тривиально следует из свойств матожидания.

4. Если $\xi \in \mathbb{Z}_+$, тогда
$$\varphi_\xi(z) = \sum_{k=0}^{+\infty} z^k \, \mathsf{P}(\xi = k).$$
Очень простое и понятное свойство. Этот ряд сходится абсолютно при $|z| \leqslant 1$, так как сумма ограничена суммой вероятностей. Можно заметить, что для дискретной величины свойство 2 доказывается очень легко.

5. При $|z| < 1$ дискретную характеристическую функцию можно дифференцировать бесконечное число раз, так как по формуле Коши-Адамара радиус сходимости останется тот же. При $|z| = 1$ непонятно что будет происходить, так как ряд $k \, \mathsf{P}(\xi = k)$ может уже расходиться.

6. В дискретной случайной величине $\varphi_\xi(0) = \mathsf{P}(\xi = 0)$, если k раз продифференцировать и подставить $z = 0$, понятно, что получим
$$\mathsf{P}(\xi = k) = \frac{1}{k!} \frac{\partial^k}{\partial z^k} \varphi_\xi(z)$$

Как ни странно, именно производящие функции помогают получить результаты в случайных дискретных процессах.

1.5.2 Вероятность вырождения

Перед тем, как доказывать сразу окончательный результат, докажем несколько лемм.

Лемма (О зависимости X_n и X_{n-1}). *Пусть X_n — ветвящийся случайный процесс с законом размножения частиц ξ.*

Тогда $\varphi_{X_n}(z) = \varphi_{X_{n-1}}(\varphi_\xi(z))$.

Доказательство.

$$\varphi_{X_n}(z) = \mathsf{E}\big[z^{X_n}\big] = \mathsf{E}\left[z^{\sum\limits_{k=1}^{X_{n-1}} \xi_k^{(n)}}\right]$$

Заметим, что X_{n-1} принимает какое-то значение дискретное значение, поэтому если ввести случайную величину $\eta = \sum\limits_{m=0}^{+\infty} \mathbf{I}\{X_{n-1} = m\}$, то $\eta = 1$ всегда. Поэтому продолжим равенство:

$$\mathsf{E}\left[z^{\sum\limits_{k=1}^{X_{n-1}} \xi_k^{(n)}}\right] = \mathsf{E}\left[z^{\sum\limits_{k=1}^{X_{n-1}} \xi_k^{(n)}} \sum_{m=0}^{+\infty} \mathbf{I}\{X_{n-1} = m\}\right] = \sum_{m=0}^{+\infty} \mathsf{E}\left[z^{\sum\limits_{k=1}^{X_{n-1}} \xi_k^{(n)}} \mathbf{I}\{X_{n-1} = m\}\right]$$

Заметим, что X_{n-1} зависит только от $\xi_k^{(\ell)}$, где $\ell \leqslant n-1$, а k — любое, поэтому можно продолжить равенство с использованием независимости случайных величин (и ещё вместо X_{n-1} подставим m):

$$\sum_{m=0}^{+\infty} \mathsf{E}\left[z^{\sum\limits_{k=1}^{m} \xi_k^{(n)}} \mathbf{I}\{X_{n-1} = m\}\right] = \sum_{m=0}^{+\infty} \mathsf{E}\left[z^{\sum\limits_{k=1}^{m} \xi_k^{(n)}}\right] \mathsf{P}(X_{n-1} = m) =$$

Заметим, что $\mathsf{E}\left[z^{\sum\limits_{k=1}^{m} \xi_k^{(n)}}\right] = (\varphi_\xi(z))^m$ из-за свойства 3 производящих функций и независимости величин $\xi_k^{(n)}$. Откуда получаем:

$$= \sum_{m=0}^{+\infty} (\varphi_\xi(z))^m \, \mathsf{P}(X_{n-1} = m) = \varphi_{X_{n-1}}(\varphi_\xi(z))$$

□

Следствие. *Воспользуемся тем, что $X_0 = 1$, тогда рекурсивно получим:*

$$\varphi_{X_n}(z) = \underbrace{\varphi_\xi(\ldots(\varphi_\xi(z))\ldots)}_{n \text{ раз}}$$

Откуда получаем (если внимательно приглядеться к результату леммы), что $\varphi_{X_n}(z) = \varphi_\xi(\varphi_{X_{n-1}}(z))$.

Введем обозначение $q = \mathsf{P}(X_n = 0)$. Тогда сформулируем следующую лемму:

Лемма. *$q_n \leqslant q_{n+1}$ и $q = \lim\limits_{n\to+\infty} q_n$, где q — вероятность вырождения процесса.*

Доказательство. Первое неравенство совершенно тривиально, так как $\{X_n = 0\} \subseteq \{X_{n+1} = 0\}$, так как если $X_n = 0$, то уж и $X_{n+1} = 0$.

$\{\exists n : X_n = 0\} = \bigcup\limits_{n=1}^{+\infty} X_n$. В силу непрерывности вероятности получаем, что

$$q = \lim_{n \to +\infty} \mathsf{P}(X_n = 0) = \lim_{n \to +\infty} q_n$$

□

Теорема 11. *Если q — вероятность вырождения, тогда $q = \varphi_\xi(q)$.*

Доказательство. $q_n = \mathsf{P}(X_n = 0) = \varphi_{X_n}(0)$ (см. свойство 6). По следствию из леммы о зависимости X_n и X_{n-1} получаем, что

$$q_n = \mathsf{P}(X_n = 0) = \varphi_{X_n}(0) = \varphi_\xi(\varphi_{X_{n-1}}(0)).$$

Заметим, что у $\varphi_{X_n}(0)$ существует предел, так как по лемме выше у q_n существует предел, поэтому устремим $n \to +\infty$, откуда

$$q = \varphi_\xi(q), \text{ что и требовалось доказать.}$$

□

Мы уже получили хороший аппарат для нахождения q. Но заметим, что $q = 1$ всегда подходит, но иногда вероятность вырождения меньше единицы. Так вот, следующая теорема утверждает, что таких q не больше двух и если их два, то надо брать наименьшее.

Теорема 12 (О вероятности вырождения)**.** *Пусть $\mathsf{P}(\xi = 1) < 1$ и $\mu = \mathsf{E}[\xi]$ (μ необязательно конечно).*

1) Если $\mu \leqslant 1$, тогда $z = \varphi_\xi(z)$ имеет ровно одно решение — единица;

2) Если $\mu > 1$, тогда $z = \varphi_\xi(z)$ имеет один корень z_0 на $[0, 1)$ и $q = z_0$.

Доказательство. Рассмотрим два случая в порядке их очереди.

1) Если $\mathsf{P}(\xi = 0) = 1$, тогда утверждение очевидно. Пусть $\mathsf{P}(\xi = 0) < 1$. Рассмотрим производную $\varphi'_\xi(z)$.

$$\varphi'_\xi(z) = \sum_{k=1}^{+\infty} k z^{k-1} \mathsf{P}(\xi = k)$$

Поэтому из формулы можно заключить, что $\varphi'_\xi(z) \nearrow$ с ростом $z > 0$, так как $\mathsf{P}(\xi = 0) < 1$.

Поэтому имеем по свойству 1 и формуле конечных приращений, что $(z < 1)$:

$$1 - \varphi_\xi(z) = \varphi_\xi(1) - \varphi_\xi(z) = \varphi'_\xi(\theta)(1 - z)$$

Где $\theta \in (0, z)$. Но производная строго возрастающая функция, как мы показали выше, поэтому:

$$\varphi'_\xi(\theta)(1 - z) < \varphi'_\xi(1)(1 - z) = \mu(1 - z) \leqslant (1 - z)$$

Откуда $1 - \varphi_\xi(z) < 1 - z$, откуда $z < \varphi_\xi(z)$ для всех $z \in [0, 1)$.

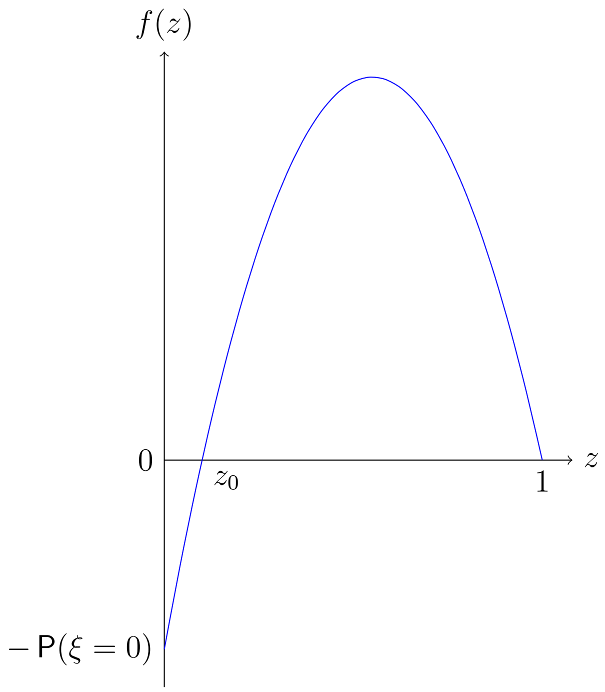
2) Заметим, что существует $k \geqslant 2$, что $\mathsf{P}(\xi = k) > 0$, иначе матожидание было бы не больше единицы. Поэтому $\varphi''_\xi(z) = \sum\limits_{k=2}^{+\infty} k(k-1)z^{k-2}\,\mathsf{P}(\xi = k)$ строго положительная функция.

Введем функцию $f(z) = z - \varphi_\xi(z)$. Тогда $f'(z) = 1 - \varphi'_\xi(z)$.

Заметим, что $f'(0) = 1 - \varphi'_\xi(0) = 1 - \mathsf{P}(\xi = 1) > 0$ (по условию). А $f'(1) = 1 - \mu < 0$.

Получаем, что производная у f сначала возрастает, потом убывает, так как $\varphi'_\xi(0)$ строго возрастающая функция. $f(1) = 0$ по свойству 1 производящих функций.

$f(0) = -\varphi_\xi(0) = -\mathsf{P}(\xi = 0)$, откуда давайте поймём, что мы начинаем из неположительного значения, возрастаем, а потом в единице приходим в ноль — тут будет строго один корень на $[0, 1)$ из-за непрерывности и поведения функции.



Рассмотрим 2 случая:
1) $\mathsf{P}(\xi = 0) = 0$. Очевидно, что тогда вероятность вырождения равна нулю, так как каждый представитель будет иметь хотя бы одного потомка.
2) $\mathsf{P}(\xi = 0) > 0$, тогда обозначим этот корень за z_0.
Несложно заметить, что $z - \varphi_\xi(z) < 0$ только при $0 \leqslant z < z_0$. Остался последний вопрос про то, какое q надо всё же выбрать.
Вообще, как мы выше показали, $q_n - q_{n+1} \leqslant 0$. Но давайте в этом случае покажем строгое неравенство.
Поймём, что

$$q_{n+1} = q_n + \sum_{m=1}^{+\infty} \mathsf{P}(X_n = m)(\mathsf{P}(\xi = 0))^m$$

То есть мы хотим, чтобы X_n приняло какое-то значение и чтобы на следующем шаге все представители остались без потомков. Хотя бы при одном $m > 0$ мы точно знаем, что $\mathsf{P}(X_n = m) > 0$ (так как $\mathsf{P}(\xi > 0) > 0$, так как иначе матожидание было бы меньше единицы) и $\mathsf{P}(\xi = 0) > 0$. Откуда действительно строгое неравенство.

Вспомним, что $\varphi_\xi(q_n) = q_{n+1}$, поэтому для всех n выполняется, что $q_n - \varphi_\xi(q_n) < 0$, значит все $q_n \in [0, z_0)$. Устремляем n к бесконечности и получаем, что $q \in [0, z_0]$, а только в z_0 функция обращается в ноль. □

Но в большинстве случаев это q найти невозможно. Следующий пример покажет, что даже в пуассоновском распределении будет трансцендентное уравнение.

Пример. Пусть $\xi \sim \mathrm{Pois}(c)$. Тогда

$$\varphi_\xi(z) = \sum_{k=0}^{+\infty} z^k\,\mathsf{P}(\xi = k) = \sum_{k=0}^{+\infty} \frac{(zc)^k}{k!} e^{-c} = e^{c(z-1)}$$

Откуда надо найти $q = e^{c(q-1)}$. Введем $p = 1 - q$ — вероятность невырождения:

$$e^{-pc} + p = 1$$

Если $c < 1$, тогда будет ровно один корень, а при $c \geqslant 1$ будет уже два корня (оставим это утверждение для доказательства исходя только из функции, как упражнение читателю).

Немного обозначим терминологию.

Определение 15. Будем классифицировать случайные процессы в зависимости от среднего значения закона размножения.

- Если $\mu < 1$, тогда случайный процесс называют *докритическим.*
- Если $\mu = 1$, тогда случайный процесс называют *критическим.*
- Если $\mu > 1$, тогда случайный процесс называют *надкритическим.*

В первых двух случаях мы доказали, что $\mathsf{P}(X_n = 0) \to 1$, то есть фактически $X_n \xrightarrow{\mathsf{P}} 0$. Следующую теорему мы оставим без доказательства, но она показывает, как растет количество потомков на уровне n.

Теорема 13 (О надкритическом случае)**.** *Пусть $\mu > 1$, $\mathsf{D}[\xi] = \sigma^2 < +\infty$ и $W_n = \frac{X_n}{\mu^n}$, тогда существует такая случайная величина W, что выполняются четыре утверждения:*

$$1)\ W_n \xrightarrow{п.н.} W;$$

$$2)\ W_n \xrightarrow{L^2} W;$$

$$3)\ \mathsf{P}(W = 0) = q;$$

$$4)\ \mathsf{E}[W] = 1, \mathsf{D}[W] = \frac{\sigma^2}{\mu^2 - \mu}.$$

1.5.3 Общее число частиц после n-го хода

Теперь будем оценивать количество частиц/экземпляров до момента времени n. Введем случайную величину $Y_n = 1 + X_1 + \ldots + X_n$. Ясно, что Y_n как раз и отвечает за это самое количество.

Лемма (О зависимости Y_n и Y_{n-1})**.**

$$\varphi_{Y_n}(z) = z \cdot \varphi_\xi(\varphi_{Y_{n-1}}(z))$$

Доказательство. Аналогично распишем, как и в лемме о зависимости X_n и X_{n-1}.

$$\varphi_{Y_n}(z) = \mathsf{E}\left[z^{Y_n}\right] = \sum_{m=0}^{+\infty} \mathsf{E}\left[z^{Y_n}\, \mathbf{I}\{X_1 = m\}\right]$$

То есть посмотрим, на сколько мы в первый раз разбились. Пусть A_i — количество всего потомков у i-й частицы после первого хода. Заметим, что $A_i \stackrel{d}{=} Y_{n-1}$, так как мы убрали первый уровень.

Это разбиение нам на руку, так как A_i независимы в совокупности, а также независимы с X_1.

$$Y_n = 1 + \sum_{k=1}^{n} A_k\, \mathbf{I}\{X_1 = m\}$$

Откуда мы получаем, что

$$\sum_{m=0}^{+\infty} \mathsf{E}\big[z^{Y_n}\,\mathbf{I}\{X_1 = m\}\big] = z\sum_{m=0}^{+\infty} \mathsf{E}\left[z^{\sum\limits_{k=1}^{m} A_k}\right]\mathsf{P}(X_1 = m)$$

Легко видеть, что $\mathsf{E}\big[z^{A_k}\big] = \varphi_{Y_{n-1}}(z)$, так как мы пользовались только распределением. Поэтому подставим по свойству 3 производящих функций:

$$z\sum_{m=0}^{+\infty} \mathsf{E}\left[z^{\sum\limits_{k=1}^{m} A_k}\right]\mathsf{P}(X_1 = m) = z\sum_{m=0}^{+\infty}(\varphi_{Y_{n-1}}(z))^m\,\mathsf{P}(X_1 = m) = z\varphi_\xi(\varphi_{Y_{n-1}}(z))$$

Последнее равенство следует из того, что $\xi \overset{d}{=} X_1$. □

А теперь давайте попробуем установить, будет ли какой-то предел у характеристических функции Y_n. На самом деле, предыдущее равенство указывает, что при $|z| < 1$ будет «убывание», давайте строго сформулируем утверждение.

Лемма (О пределе характеристических функций Y_n)**.** *Для любого $z \in [0, 1)$ существует предел $\lim\limits_{n\to+\infty} \varphi_{Y_n}(z) = \rho(z)$, притом $\rho(z) = z\varphi_\xi(\rho(z))$.*

Доказательство. Докажем, что для любого $z \in [0, 1)$ верно, что $\varphi_{Y_{n+1}}(z) \leqslant \varphi_{Y_n}(z)$, тогда мы докажем утверждение, так как снизу эти выражения ограничены нулём.

Будем вести индукцию по n. Удивительным образом база доказывается дольше перехода.

База $n = 0$. Выпишем очевидные равенства из свойств производящих функций:

$$\varphi_{Y_1}(z) = \varphi_{1+X_1}(z) = \varphi_1(z)\varphi_{X_1}(z) = \varphi_1(z)\varphi_\xi(z)$$

Теперь у случайной величины «один» будет в точке z значение z по определению характеристической функции. $X_1 \overset{d}{=} \xi$, так как это количество первых потомков от первого. Поэтому верно последнее равенство.

Но заметим, что $\varphi_\xi(z) < 1$, так как это меньше суммы всех вероятностей.

Поэтому мы получили, что $\varphi_{Y_1}(z) < z = \varphi_{Y_0}(z)$. Последнее равенство следует из того, что $Y_0 \overset{d}{=} 1$.

Переход $n - 1 \to n$. По лемме о зависимости мы получаем:

$$\varphi_{Y_{n+1}}(z) = z\varphi_\xi(\varphi_{Y_n}(z))$$

Воспользуемся предположением индукции и тем, что характеристическая функция ξ не убывает с ростом z, откуда:

$$z\varphi_\xi(\varphi_{Y_n}(z)) \leqslant z\varphi_\xi(\varphi_{Y_{n-1}}(z)) = \varphi_{Y_n}(z)$$

Первое утверждение, о котором мы дополнительно заявляли, сразу следует из предела, который мы доказали.

□

Также давайте доопределим $\rho(1) = q$.

1.6 Лекция от 28.02.2017

1.6.1 Общее число частиц после n-го хода

Будем считать, что $Y_n \to Y$ (Y может быть бесконечностью, сходимость имеется в виду предельная, то есть смотрят на предел Y_n и вероятность каждого значения (в том числе и бесконечность) — это вероятность принять это значение Y_n-ыми). Давайте поймём, что $\rho(z)$ — производящая функция величины $Y\,\mathbf{I}\{Y < +\infty\}$. Действительно, если ограничить Y конечными значениями, то получим, что

$$\varphi_{Y\,\mathbf{I}\{Y<+\infty\}}(z) = \sum_{k=0}^{+\infty} z^k\,\mathsf{P}(Y = k)$$

А теперь давайте поймём, что это в точности $\rho(z)$ при $z \in [0,1]$. В единице как раз вероятность вырождения, тут всё отлично. В остальных точках это предел производящих функций, то есть предел $\mathsf{E}\left[z^{Y_n}\right] \to \mathsf{E}\left[z^{Y\,\mathbf{I}\{Y<+\infty\}}\right]$, действительно, у величин Y_n есть сходимость по распределению к $Y\,\mathbf{I}\{Y < +\infty\}$, а z^k при $z \in [0,1)$ ограниченная функция. Поэтому в пределе как раз и получим $\rho(z)$.

Давайте рассмотрим пример после не самых тривиальных следствий.

Пример. $\xi \sim \mathrm{Geom}(p), p \in (0,1)$, причём сделаем его с нуля ($\mathsf{P}(\xi = k) = p(1-p)^k, k \in \mathbb{Z}_+$). Пусть $\{X_n, n \in \mathbb{Z}_+\}$ — ветвящийся процесс с законом размножения ξ. Интересно было бы найти $\rho, \mathsf{P}(Y = k)$. Давайте этим и займёмся.

$$\varphi_\xi(z) = \mathsf{E}\left[z^\xi\right] = \sum_{k=0}^{+\infty} z^k (1-p)^k p = \frac{p}{1 - z(1-p)}$$

Откуда из леммы о пределе характеристических функций получаем:

$$\rho = \frac{zp}{1 - \rho(1-p)}$$

После недолгих вычислений с учётом, что $p \neq 1$ получим:

$$\rho(z) = \frac{1 \pm \sqrt{1 - 4zp(1-p)}}{2(1-p)}$$

Надо теперь выбрать, какой знак нам подойдет. Если подставить $z = 0$, то по лемме предел должен быть равен нулю ($Y_n > 0$ всегда из-за начальной частицы).

Поэтому надо выбрать знак «минус», чтобы сошлось. Отлично, мы знаем производящую функцию, давайте теперь вычислять коэффициенты через ряд Тейлора и производные.

$$\rho(z) = -\sum_{k=1}^{+\infty} \binom{1/2}{k} (-4p(1-p))^k \frac{1}{2(1-p)} z^k = -\sum_{k=1}^{+\infty} \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\cdot\ldots\cdot\left(\frac{1}{2}-k+1\right)}{2k!(1-p)} 2^k (-4p(1-p))^k z^k =$$
$$= \sum_{k=1}^{+\infty} \frac{(2k-3)!!2^k}{k!} (p(1-p))^k \frac{1}{2(1-p)} z^k = \sum_{k=1}^{+\infty} \frac{1}{2}\binom{2k}{k} \frac{1}{2k-1} p^k (1-p)^{k-1} z^k$$

Откуда $\mathsf{P}(Y = k) = \frac{1}{2}\binom{2k}{k}\frac{1}{2k-1} p^k (1-p)^{k-1}$.

На самый конец давайте ещё вычислим математическое ожидание частиц при условии, что процесс конечен. Это даёт хорошее приближение Y_n, но так изворачиваемся мы потому, что через производящие функции легче посчитать это значение у Y, нежели возиться с Y_n и пределами

$$\mathsf{E}[Y \mid Y < +\infty] = \sum_{k=0}^{+\infty} k\,\mathsf{P}(Y = k \mid Y < +\infty) = \sum_{k=0}^{+\infty} \frac{k\,\mathsf{P}(Y = k)}{q} = \frac{\rho'(1)}{q}$$

Теперь вычислим производную слева и справа у давно нам известного равенства $\rho(z) = z\varphi_\xi(\rho(z))$ и сразу подставим единицу:

$$\rho'(1) = \varphi_\xi(\rho(1)) + \rho'(1)\varphi'_\xi(\rho(1)) \implies \rho'(1) = \frac{q}{1 - \varphi'_\xi(\rho(1))}$$

Откуда наконец-то вычисляется наш финальный результат:

$$\mathsf{E}[Y \mid Y < +\infty] = \frac{1}{1 - \varphi'_\xi(q)}$$

На этом случайные процессы Гальтона-Ватсона заканчиваются.

1.6.2 Случайные графы

Случайные графы можно описать различными способами. Существует несколько моделей создавать графы, о которых мы поговорим немного позже. Случайные графы нашли практическое применение во всех областях, где нужно смоделировать сложные сети — известно большое число случайных моделей графов, отражающих разнообразные типы сложных сетей в различных областях. Но, к сожалению, граф, задающий интернет, не описывается какой-то простой моделью. Об этом мы тоже поговорим в этой лекции.

1.6.3 Классические модели

Существуют две классические модели, задающие случайный граф. Вообще, мы всегда будем работать с графом на n вершинах и каким-то образом выбирать остовный подграф на этих вершинах. Напомним, что остовный подграф — подграф на этих же вершинах, но с каким-то (возможно пустым) подмножеством ребер.

1. **Биномиальная**

   Определение 16. $G(n,p)$ — случайный граф, который имеет следующее распределение для любого остовного подграфа G:

   $$\mathsf{P}(G(n,p) = G) = p^{|E(G)|}(1-p)^{\binom{n}{2} - |E(G)|}$$

   Другими словами, любое ребро включается в случайный граф с вероятностью p.

2. **Равномерная**

   Определение 17. Пусть $\mathfrak{G}_m$ — все остовные подграфы с количеством ребер m. Тогда $G(n,m)$ — случайный граф, который имеет равномерное распределение для любого $G \in \mathfrak{G}_m$

   $$\mathsf{P}(G(n,m) = G) = \frac{1}{\binom{\binom{n}{2}}{m}}$$

Несложно заметить, что если $p(n)$ и $m(n)$ таковы, что $\binom{n}{2}p \sim m$, то вероятности будут не сильно отличаться, т.е. $\mathsf{P}(G(n,p)$ обладает свойством $Q) \sim \mathsf{P}(G(n,m)$ обладает свойством $Q)$, потому что мы выбираем в среднем около m ребер. Конечно, это нестрогое рассуждение, но эквивалентность этих моделей и вправду (как мы увидим позднее) почти всегда выполняется для всех адекватных свойств Q.

Приведем ещё некоторые (менее популярные) модели случайных графов.

3. **Случайный d-регулярный граф**

   Просто равномерно строится случайный d-регулярный граф. Один из хороших способов строить такой граф — взять nd точек, после этого разбить по n групп по d точек в каждую группу. После этого надо взять произвольное совершенное паросочетание так, чтобы в каждой группе не было ребер между собой. После этого надо сжать каждую группу в одну вершину и сказать, что мы построили граф. Оставим упражнение читателю для доказательства, что так равновероятно будет получен любой регулярный граф.

4. **Случайный двудольный граф**

   Тут всё просто, если есть полный граф $K_{n,m}$, то с вероятностью p мы берем каждое ребро.

1.6.4 Случайный процесс на графах

Генерировать графы можно и как случайный процесс. То есть добавлять рёбра (или даже несколько ребер) в какие-то моменты времени. Существуют два вида времени, для которых хоть сколько-нибудь осмысленно рассматривать случайный процесс на графах.

1. **Дискретное время** Тут достаточно всё интуитивно. Каждый момент времени мы добавляем какое-то ребро, которое ещё не было добавлено до этого.

   Определение 18. $\tilde{G} = (\tilde{G}(n,m), m = 0, \ldots, m = \binom{n}{2})$ таков, что $\tilde{G}(n,0)$ пуст и $\tilde{G}(n,m) = \tilde{G}(n,m-1)$ плюс какое-то ребро из тех, кто не вошёл в $\tilde{G}(n,m)$.

   Практически очевидно, что $\tilde{G}(n,m) \stackrel{d}{=} G(n,m)$ (можно воспользоваться идеей перенумерации ребер).

   Но такой процесс позволяет нам позволяет смотреть динамику получившихся графов. Давайте анонсируем некоторые результаты, один из которых мы позже докажем.

   Пусть τ_1 — первый момент, когда нет изолированных вершин, τ_2 — первый момент, когда граф становится связным, τ_3 — когда степень каждой вершины становится хотя бы двойка, а τ_4 — первый момент, когда граф становится гамильтоновым.

   Ясно, что между этими величинами можно поставить следующие соотношения:

   - $\tau_1 \leqslant \tau_2$, т.к в связных графах нет изолированных вершин;
   - $\tau_1 \leqslant \tau_3$, т.к если у каждой вершины хотя бы степень 2, то нет ни одной изолированной;
   - $\tau_3 \leqslant \tau_4$, т.к гамильтонов цикл уж точно нам даёт, что из каждой вершины выходит хотя бы 2 ребра.

Но связь между этими случайными величинами более тонкая.

Теорема 14.

$$\mathsf{P}(\tau_1 = \tau_2) = 1, n \to +\infty$$
$$\mathsf{P}(\tau_3 = \tau_4) = 1, n \to +\infty$$

Пока оставим без доказательства, но первый пункт мы точно докажем в курсе, второй доказывается тяжело.

Если так вдуматься, то получается, что как только перестают существовать изолированные вершины, так сразу граф становится связен. И как только степень вершин хотя бы два, то граф уже гамильтонов. Возможно это противоречит интуиции, но так и есть и вскоре мы с этим разберемся.

2. **Непрерывное время**

   Определение 19. Занумеруем как-то ребра полного графа K_n, тогда $\tilde{G} = (G(n,t), t \geqslant 0)$ называется случайным процессом на графе с непрерывным временем и $\tilde{G}(n,t)$ состоит из тех ребер i для которых $t(i) \leqslant t$.

   $t(i)$ можно выбирать как угодно, возможно из какого-то распределения случайно или только некоторые из них случайно. Это понятие достаточно хорошо коррелирует с биномиальной моделью, а именно $\tilde{G}(n,t) \overset{d}{=} G(n,p)$ при $p = \mathsf{P}(t(i) \leqslant t)$ — если каждое $t(i)$ одинаково выбирается случайно каким-либо образом.

3. **Triangle free process**

   То же, что и в дискретном процессе, только мы не берем на шаге ребро, если оно образовывает треугольник. И останавливаемся, когда такой ход сделать нельзя.

   Данный процесс получил широкое распространение из-за теории Рамсея, так как размер независимого множества получается не самым маленьким (порядка $\Omega(\sqrt{n \ln n})$) и одновременно получилась точная оценка ($\mathcal{O}(\sqrt{n \ln n})$), что дало огромное количество значений для чисел Рамсея $R(3,n)$ вплоть до $n = 9$.

1.6.5 Граф интернета

Одним из самых важных графов для поисковых компаний на данное время — является граф интернета. Но, к сожалению, его не удалось как-то просто моделировать по многим причинам. Опишем какие свойства у графа интернета есть (мы рассматриваем модель, что вершины это ссылки, а нахождение каких-то ссылок на другие сайты — это рёбра).

- Эмпирически было выяснено, что ребер в таком графе равно $\mathcal{O}(n)$. То есть граф является разреженным;

- Степенной закон. Вершин степени d в таком равно примерно $\frac{cn}{d^{2.3}}$. Гораздо удивительнее здесь константа 2.3, которая не меняется с течением времени.

- Константный малый диаметр. Как и в мире существует гипотеза о шести рукопожатиях, так и в веб-графе существует закон шести кликов — это факт, состоящий в том, что диаметр веб-графа равен шести.

- В 1999 году двое исследователей — А. Л. Барабаши и Р. Альберт — предложили крайне простую идею, которая, однако ж, оказалась весьма продуктивной. Идея заключалась в том, что, когда новый сайт появляется на свет, он, скорее всего, «предпочитает» сослаться на те сайты, которые и без того уже многими цитированы. То есть с каждым новым сайтом добавляются ссылки, пропорциональному степени входящих и исходящих ребер того или иного сайта.

- Будем называть два узла соседями, если существует связь между ними. Для веб-графа характерно, что два узла, соседних к какому-либо узлу, часто также являются соседями между собой. Чтобы охарактеризовать это явление и был предложен кластерный коэффициент C_i узла i. Предположим, что узел имеет степень k_i , это значит, что у него k_i соседей и между ними может быть максимум $k_i(k_i-1)/2$ связей. Тогда

  $$C_i = \frac{2n_i}{k_i(k_i-1)},$$

  где n_i — число связей между соседями узла i.

  У веб-графа, тем не менее, наблюдается постоянный и немаленький кластерный коэффициент.

Сложность заключается в том, что простыми моделями такой граф не описать, так как у каждой из них существуют те или иные отклонения от пунктов, которые написаны выше. Кроме того, ясно, что в гипотетических моделях веб-графа у более старых вершин гораздо больше шансов иметь большую входящую степень, нежели у более новых вершин. Это заведомо плохо согласуется с новостными «взрывами», которые ежедневно случаются в интернете: едва появляется страница с важной новостью, как на нее приходят тысячи ссылок. Мы уж не говорим о том, что сайты, страницы и ссылки на них зачастую умирают, и это тоже сложность для отражения в моделях.

1.6.6 Монотонные свойства

Но мы всё же остановим свой взгляд на классических моделях графов. Чтобы что-то доказывать, нужно как-то классифицировать свойства графов, так как у нас уже есть различные процессы, которые добавляют рёбра или, наоборот, убирают их.

Напомним, что свойством мы называем любое подмножество графов на n вершинах.

Определение 20. Свойство Q называется *возрастающим*, если для любого G_1 обладающим этим свойством и тем, что $G_1 \subseteq G_2$ следует, что G_2 тоже обладает этим свойством.

Пример. Примеров достаточно много, таких как:

- Связность;

- Отсутствие изолированных вершин;

- Наличие любой клики.

Определение 21. Свойство Q называется *убывающим*, если для любого G_1 обладающим этим свойством и тем, что $G_1 \supseteq G_2$ следует, что G_2 тоже обладает этим свойством.

Пример. Здесь немного другого типа примеры, которые выполняются для пустых графов.

- Несвязность;
- Планарность;
- Ацикличность;
- Двудольность.

Определение 22. Свойство Q называется *выпуклым*, если для любого G_1 и G_2 обладающими этим свойством и G_3 таким, что $G_1 \subseteq G_3 \subseteq G_2$ следует, что G_3 обладает этим свойством.

Пример. Число изолированных вершин равно k (оно не монотонное).

Для разнообразия приведем пример не выпуклого и не монотонного свойства — «максимальная компонента связности является деревом». Желающие могут самостоятельно придумать примеры, которые показывают, что это свойство не монотонное и не выпуклое.

1.7 Лекция от 07.03.2017

Первые 25 минут лекции были посвящены холивару по поводу домашнего задания по случайным блужданием. В частности по девятой задаче и корректности перехода к пределу в функциональных рядах.

1.7.1 Пороговые функции

Лемма (О монотонности вероятности)**.** *Пусть Q возрастающее свойство на n вершинах, тогда если $0 \leqslant p_1 < p_2 \leqslant 1$ и $0 \leqslant m_1 < m_2 \leqslant \binom{n}{2}$, то*

$$\mathsf{P}(G(n,p_1) \in Q) \leqslant \mathsf{P}(G(n,p_2) \in Q), \mathsf{P}(G(n,m_1) \in Q) \leqslant \mathsf{P}(G(n,m_2) \in Q)$$

Заметим, что при равенствах вероятностей и так понятно, что будет происходить, поэтому поставим строгие неравенства.

Доказательство. Легко упражнение заключается в том, что если $G(n,p')$ и $G(n,p'')$ независимые два случайных графа, то $G(n,p') \cup G(n,p'') \overset{d}{=} G(n,p'+p''-p'p'')$.

Действительно, посмотрим на вероятность получения одного графа с количеством ребер $|E(G)|$ с вероятностью $p'+p''-p'p''$. По определению вероятность равна $(p'+p''-p'p'')^{|E(G)|}(1-p'-p''+p'p'')^{\binom{n}{2}-|E(G)|}$. А теперь посмотрим на левую часть. Оба подграфа должны быть подграфом фиксированного G. Вероятность того, что $G(n,p')$ является подграфом G равна

$$\sum_{m=0}^{|E(G)|} \binom{|E(G)|}{m} p'^m (1-p')^{\binom{n}{2}-m}$$

А теперь надо, чтобы второй процесс выбрал все недостающие ребра (это происходит с вероятностью $p''^{|E(G)|-m}$), а также можно выбирать выбранные уже первым процессом ребра. Вероятность такого равна $(1-p'')^{\binom{n}{2}-|E(G)|}$, так как запрещено брать оставшиеся ребра. В итоге вероятность объединения равна:

$$\sum_{m=0}^{|E(G)|}\binom{|E(G)|}{m}p'^m p''^{|E(G)|-m}(1-p')^{\binom{n}{2}-m}(1-p'')^{\binom{n}{2}-|E(G)|}$$

Вынесем $(1-p'')^{\binom{n}{2}-|E(G)|}(1-p')^{\binom{n}{2}-|E(G)|}=(1-p'-p''+p'p'')^{\binom{n}{2}-|E(G)|}$, что даёт один множитель. А теперь давайте посчитаем оставшийся множитель:

$$\sum_{m=0}^{|E(G)|}\binom{|E(G)|}{m}p'^m p''^{|E(G)|-m}(1-p')^{|E(G)|-m}=(p'+p''(1-p'))^{|E(G)|}=(p'+p''-p'p'')^{|E(G)|}.$$

Отлично, с упражнением разобрались, давайте теперь доказывать теорему.

Введем $p'=\frac{p_2-p_1}{1-p_1}$ (тут как раз потребовалось, что $1-p_1>0$) и возьмём $G(n,p')$ независимый с $G(n,p_1)$, тогда (из-за того, что свойство возрастающее):

$$\mathsf{P}(G(n,p_1)\in Q)\leqslant\mathsf{P}(G(n,p_1)\cup G(n,p')\in Q)=\mathsf{P}(G(n,p'+p_1-p'p_1)\in Q)=\mathsf{P}(G(n,p_2)\in Q).$$

С равномерной моделью всё ещё проще. Вспомним, что $\tilde{G}(n,m_1)\overset{d}{=}G(n,m_1)$, откуда

$$\mathsf{P}(G(n,m_1)\in Q)=\mathsf{P}\Big(\tilde{G}(n,m_1)\in Q\Big)\leqslant\mathsf{P}\Big(\tilde{G}(n,m_2)\in Q\Big)=\mathsf{P}(G(n,m_2)\in Q).$$

Основное неравенство появляется из-за того, что в случайном процессе мы добавляем ребра и свойство возрастающее. □

Определение 23. Пусть $Q(n)$ — возрастающее свойство графов, тогда функция $\hat{p}(n)$ называется пороговой для $Q(n)$, если выполнено следующее

$$\lim_{n\to+\infty}\mathsf{P}(G(n,p(n))\in Q)=\begin{cases}0, p(n)=o(\hat{p}(n))\\1, \hat{p}(n)=o(p(n))\end{cases}$$

Неформально это свойство означает, что пороговая функция это эдакий разделитель для свойства, когда оно никогда не выполняется и когда всегда. При $p(n)=\Theta\left(\hat{p}(n)\right)$ часто возникают интересные результаты. Давайте приведем примеры:

- **Связность**

  Пороговая вероятность в данном случае равна $\hat{p}(n)=\frac{\ln n}{n}$. На пороге верно следующее утверждение, которое мы докажем позднее: если $p(n)=\frac{\ln n+\lambda+o(1)}{n}$, тогда $\mathsf{P}(G \text{ — связен})\to e^{-e^{-\lambda}}$ (да, этот результат производит некоторое впечатление катарсиса). Это распределение называется распределением Гумбеля.

- **Наличие треугольников**

  Пороговая вероятность $\hat{p}(n)=\frac{1}{n}$.

- $\min\deg\geqslant k$

  Пороговая вероятность в данном случае совершенно не зависит от k (что понятно, так как при достаточно больших n уже становится маловероятно, что k как-то сильно играет роль) и равна $\frac{\ln n}{n}$. Также тут известен более точный результат: если $p=\frac{\ln n+(k-1)\ln\ln n+w(n)}{n}$ и $w(n)$ растет, тогда вероятность стремится к единице, если убывает, тогда стремится к нулю.

Наша основная задача до конца лекции — доказать, что пороговая функция всегда существует для всех возрастающих свойств для биномиальной модели (для равномерной надо чуть больше париться, нам оно не потребуется тем более).

Определение 24. Если Q возрастающее свойство, тогда введем функцию

$$f(p) = \mathsf{P}(G(n,p) \in Q)$$

Лемма (Лемма о функции f)**.** *Если Q нетривиальное свойство, тогда $f(p)$ непрерывно строго возрастающая функция.*

Доказательство. То, что $f(p)$ возрастающая напрямую следует из леммы о монотонности вероятности в возрастающем свойстве. Так же отметим, что раз свойство нетривиальное, то $f(0) = 0, f(1) = 1$. Выпишем $f(p)$ в явном виде:

$$f(p) = \sum_G p^{|E(G)|}(1-p)^{\binom{n}{2}-|E(G)|}$$

Это многочлен, притом он не константа из-за того, что на концах отрезка принимает различные значения. И если у нас есть промежуток постоянства со значением C, то у $f(p) - C$ бесконечно много корней, что не может быть у неконстантного многочлена. Значит эта функция непрерывна и строго возрастает. $\square$

Определение 25. Для любого $a \in (0,1)$ введем $p(a,n) = f^{-1}(a)$ — вероятность $p(a,n)$, что $\mathsf{P}(G(n,p(a,n)) \in Q) = a$. Такое значение всегда найдётся из-за строгой монотонности и непрерывности.

1.7.2 Существование пороговой вероятности

Теорема 15 (Критерий пороговости)**.** *Функция $\hat{p}(n)$ является пороговой для возрастающего нетривиального свойства $Q \iff \forall a \in (0,1) \implies p(a,n) = \Theta(\hat{p}(n))$.*

Доказательство. $\Rightarrow$ Пусть $\exists a \in (0,1)$, что $p(a,n) \neq \Theta(\hat{p}(n))$, откуда существует подпоследовательность $\{n_k\}_{k=0}^{+\infty}$ такая, что

$$\left[\begin{aligned} & p(a,n_k) = o(\hat{p}(n_k)) \\ & \hat{p}(n_k) = o(p(a,n_k)) \end{aligned}\right.$$

В любом случае получаем, что или $a = 0$ или $a = 1$, так как $\mathsf{P}(G(n,p(a,n_k)) \in Q) = a$, а те условия дают именно определение пороговой функции. Но мы так же знаем, что $a \in (0,1)$, поэтому получаем противоречие.

$\Leftarrow$ Пусть $\hat{p}$ удовлетворяет данному свойству. Тогда рассмотрим случай, когда $p = o(\hat{p})$, тогда $p = o(p(a,n)) \forall a \in (0,1)$, тогда $\mathsf{P}(G(n,p) \in Q) \leqslant \mathsf{P}(G(n,p(a,n)) \in Q)$ для всех достаточно больших (можно считать, что для всех) n. Откуда

$$\varlimsup_{n\to+\infty} \mathsf{P}(G(n,p) \in Q) \leqslant a,\ \forall a \in (0,1)$$

Откуда и сам предел равен нулю. Другой случай совершенно аналогичен с точностью до замены знаков и нижнего предела. $\square$

Теперь мы наконец-то дошли до самого существования. Также заметим, что для тривиальных свойств пороговые функции не определены (мы не знаем, что такое $o(0)$), но с этими свойствами и так всё ясно. Поэтому будем доказывать теорему для нетривиальных свойств.

Теорема 16 (О пороговой вероятности)**.** *Любое нетривиальное монотонное свойство имеет пороговую вероятность.*

Доказательство. Будем доказывать только для возрастающих свойств. Для убывающих аналогично.

Используя предыдущую лемму, покажем, что $\forall a \in (0,1) \implies p(a,n) = \Theta\left(p\left(\frac{1}{2}, n\right)\right)$.

Зафиксируем $\varepsilon > 0$ и возьмём натуральное m такое, что $(1-\varepsilon)^m < \varepsilon$. Рассмотрим случайные независимые графы $G^{(1)}(n, p(\varepsilon, n)), \ldots, G^{(m)}(n, p(\varepsilon, n))$. Рассмотрим $\tilde{G} = G^{(1)} \cup \ldots \cup G^{(m)}$. Посмотрим на вероятность, что $\tilde{G}$ не обладает этим свойством. Это означает, что каждое из $G^{(i)}$ этим свойством не обладал, поэтому из-за независимости событий получаем, что $\mathsf{P}\Big(\tilde{G} \notin Q\Big) \leqslant (1-\varepsilon)^m < \varepsilon$. Откуда получаем, что $\mathsf{P}\Big(\tilde{G} \in Q\Big) \geqslant 1 - \varepsilon$.

По индукции легко получить, что $\tilde{G} \overset{d}{=} G(n, p')$, где $p' = 1-(1-p(\varepsilon,n))^m$. Из неравенства Бернулли получим, что $p' \leqslant mp(\varepsilon, n)$.

Так как $\mathsf{P}\Big(\tilde{G} \in Q\Big) \geqslant 1 - \varepsilon$, то $p' \geqslant p(1-\varepsilon, n)$ из-за монотонности вероятности. Откуда получаем двойное неравенство:

$$\forall \varepsilon \in \left(0, \frac{1}{2}\right) \implies \frac{1}{m} p(1-\varepsilon, n) \leqslant p(\varepsilon, n) \leqslant p(1-\varepsilon, n)$$

В итоге получаем цепочку неравенств для всех $\varepsilon \in \left(0, \frac{1}{2}\right)$:

$$\frac{1}{m} p(1-\varepsilon, n) \leqslant p(\varepsilon, n) \leqslant p\left(\frac{1}{2}, n\right) \leqslant p(1-\varepsilon, n) \leqslant mp(\varepsilon, n)$$

Откуда видно, что $p(\varepsilon, n) = \Theta\left(p\left(\frac{1}{2}, n\right)\right)$ и $p(1-\varepsilon, n) = \Theta\left(p\left(\frac{1}{2}, n\right)\right)$, так как m не зависит от n. Значит мы доказали для всех $\varepsilon \in (0,1)$ (так как $1 - \varepsilon \in (1/2, 1)$). $\square$