МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА
Механико-математический факультет

Курс лекций по теории случайных процессов

Лектор — Андрей Михайлович Зубков

III курс, 6 семестр, поток математиков

Москва, 2007 г.

Содержание

Предисловие от редакции

Это полный курс лекций, однако при наборе в текст могли вкрасться опечатки, неточности и ошибки. Пожалуйста, сообщайте о них нам по электронной почте. В текущей версии есть незначительные пробелы, которые хотелось бы заполнить, но на это пока что не хватает времени. Если кто-нибудь вдруг захочет этим заняться, сообщайте, мы вышлем вам исходники.

Александр Харитонов (kalkin@mexmat.net), Степан Кузнецов (skuzn@inbox.ru).

Последняя компиляция: 22 сентября 2007 г.
Обновления документа — на сайте http://rain.aliso.ru/mexmat,
Об опечатках и неточностях пишите на kalkin@mexmat.net, skuzn@inbox.ru.

1. Немного философии

В теории вероятностей рассматривается вероятностное пространство $(\Omega, \mathscr{F}, \mathsf{P})$ и случайные величины $\xi\colon \Omega \longrightarrow \mathbb{R}$ (или в другое множество, вроде $\mathbb{Z}$ или $\mathbb{R}^n$), т.е. функции, измеримые относительно $\mathscr{F}$ и σ-алгебры на образе. В теории случайных процессов рассматривается отображение Ω в некоторое пространство функций, например $\mathbb{R}^\infty$ — множество последовательностей вещественных чисел или $\mathbb{R}^\mathbb{R}$ — множество отображений $\mathbb{R} \longrightarrow \mathbb{R}$.

Сравним две теоремы из курса теории вероятностей:

ЗБЧ. Пусть $\xi_1, \xi_2, \ldots$ — независимые одинаково распределённые случайные величины, $\mathsf{M}\xi_1 = a$, $\mathsf{D}\xi_1 < \infty$, $S_n = \xi_1 + \cdots + \xi_n$. Тогда для всех $\varepsilon > 0$

$$\mathsf{P}\left\{\left|\frac{1}{n}S_n - a\right| > \varepsilon\right\} \xrightarrow[n\to\infty]{} 0.$$

УЗБЧ. При тех же условиях имеем

$$\mathsf{P}\left\{\lim_{n\to\infty} \frac{1}{n}S_n = a\right\} = 1.$$

В ЗБЧ речь идёт о распределениях отдельных сумм, а в УЗБЧ — уже о всей последовательности $\{S_n\}$ — это уже теорема о случайном процессе (случайной последовательности).

Рассмотрим ещё одну теорему из курса теории вероятностей:

ЦПТ. В условиях ЗБЧ имеем:

$$\mathsf{P}\left\{\frac{S_n - na}{\sqrt{n\mathsf{D}\xi_1}} \leqslant x\right\} \xrightarrow[n\to\infty]{} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{u^2}{2}}\,du.$$

Здесь — опять свойство распределений отдельных сумм.

2. Закон повторного логарифма

Пусть $\zeta_1, \zeta_2, \ldots$ — произвольная последовательность случайных величин, $\psi(n)$ — детерминированная функция натурального аргумента. Рассмотрим последовательность событий $\{\zeta_n > \psi(n)\}$. Положим $\nu(\psi) = \sum\limits_{n=1}^{\infty} \chi\{\zeta_n > \psi(n)\}$, где $\chi(A)$ — индикатор события A.

Определение. ψ — верхняя функция максимумов, если $\mathsf{P}\{\nu(\psi) < \infty\} = 1$. ψ — нижняя функция максимумов, если $\mathsf{P}\{\nu(\psi) = \infty\} = 1$. ψ — правильная функция максимумов, если для любого $\varepsilon > 0$ $(1+\varepsilon)\psi$ — верхняя функция максимумов, а $(1-\varepsilon)\psi$ — нижняя функция максимумов. Аналогично определяются функции минимумов.

Утверждение 2.1. *Пусть $\{\zeta_n\}$ — последовательность независимых случайных величин, $\zeta_n \sim \mathcal{N}(0,1)$. Тогда*

$$\mathsf{P}\left\{\limsup_{n\to\infty} \frac{\zeta_n}{\sqrt{2\ln n}} = 1\right\} = 1 \qquad \text{и} \qquad \mathsf{P}\left\{\liminf_{n\to\infty} \frac{\zeta_n}{\sqrt{2\ln n}} = -1\right\} = 1.$$

Лемма 2.2 (лемма Бореля–Кантелли). *Пусть на $(\Omega, \mathscr{F}, \mathsf{P})$ задана последовательность событий $\{A_n\}$. $A^* = \{\omega \in \Omega \mid \sum\limits_{n=1}^{\infty} \chi(A_n) = \infty\} = \{$выполняется бесконечно много событий из этой последовательности$\}$. Тогда*

1. *если $\sum\limits_{n=1}^{\infty} \mathsf{P}(A_n) < \infty$, то $\mathsf{P}(A^*) = 0$;*
2. *если A_n попарно независимы и $\sum\limits_{n=1}^{\infty} \mathsf{P}(A_n) = \infty$, то $\mathsf{P}(A^*) = 1$.*

□ Смотри курс теории вероятностей. ■

Утверждение 2.3. $1 - \Phi(x) = \dfrac{1-\varepsilon_x}{x}\varphi(x)$, *где $\varepsilon_x \longrightarrow 0$ при $x \longrightarrow \infty$, $\varepsilon_x > 0$, $\varphi(x) = \dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$.*

□ Совершаем замену $u = x + v$:

$$1 - \Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{x}^{+\infty} e^{-\frac{u^2}{2}}\,du = \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{+\infty} e^{-\frac{(x+v)^2}{2}}\,dv = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \int\limits_{0}^{+\infty} e^{-xv-\frac{v^2}{2}}\,dv \stackrel{?}{=} \frac{1-\varepsilon_x}{x}\varphi(x).$$

Чтобы обосновать последний переход (доказать, что ε_x такое, какое надо), оценим интеграл. Совершим замену $xv = z$ ($x\,dv = dz$):

$$\int\limits_0^{+\infty} e^{-xv-\frac{v^2}{2}}\,dv = \int\limits_0^{+\infty} e^{-z-\frac{z^2}{2x^2}}\,\frac{dz}{x}.$$

$$\int\limits_0^{+\infty} e^{-z-\frac{z^2}{2x^2}}\,dz < \int\limits_0^{+\infty} e^{-z}\,dz = 1,$$

откуда $\varepsilon_x > 0$. Оценим с другой стороны:

$$\int\limits_0^{+\infty} e^{-z-\frac{z^2}{2x^2}}\,dz > \int\limits_0^{\sqrt{x}} e^{-z}e^{-\frac{z^2}{2x^2}}\,dz > e^{-\frac{1}{2x}}\int\limits_0^{\sqrt{x}} e^{-z}\,dz > \left(1-\frac{1}{2x}\right)\left(1-e^{-\sqrt{x}}\right) \longrightarrow 1\ (x\longrightarrow\infty),$$

откуда $\varepsilon_x \longrightarrow 0$ при $x \longrightarrow \infty$. ■

□ [Доказательство утверждения 2.1] Положим $\psi(n) = \sqrt{2\ln n}$, $A_n(b) = \{\zeta_n > b\psi(n)\}$,

$$A^*(b) = \left\{\omega \mid \sum_{n=1}^{\infty}\chi(A_n(b)) = \infty\right\} = \left\{\omega \mid \limsup_{n\to\infty}\frac{\zeta_n}{\psi(n)} \geqslant b\right\}.$$

В силу утверждения 2.3

$$\sum_{n=1}^{\infty}\mathsf{P}(A_n(b)) = \sum_{n=1}^{\infty} e^{-\frac{(b\psi(n))^2}{2}}\,\frac{1-\varepsilon_n}{b\psi(n)\sqrt{2\pi}} = \sum_{n=1}^{\infty}\frac{1}{n^{b^2}}\,\frac{1-\varepsilon_n}{b\sqrt{2\pi}\sqrt{2\ln n}} \begin{cases} = \infty & \text{при } b \leqslant 1,\\ < \infty & \text{при } b > 1.\end{cases}$$

Отсюда в силу леммы Бореля–Кантелли для всех $b > 1$

$$\mathsf{P}\left\{\limsup_{n\to\infty}\frac{\zeta_n}{\psi(n)} \geqslant b\right\} = 0,$$

а поэтому

$$\mathsf{P}\left\{\limsup_{n\to\infty}\frac{\zeta_n}{\psi(n)} \leqslant 1\right\} = 1.$$

Кроме того, для всех $b < 1$

$$\mathsf{P}\left\{\limsup_{n\to\infty}\frac{\zeta_n}{\psi(n)} \geqslant b\right\} = 1,$$

откуда

$$\mathsf{P}\left\{\limsup_{n\to\infty}\frac{\zeta_n}{\psi(n)} \geqslant 1\right\} = 1.$$

Окончательно,

$$\mathsf{P}\left\{\limsup_{n\to\infty}\frac{\zeta_n}{\psi(n)} = 1\right\} = 1,$$

что и завершает доказательство утверждения. ■

Теорема 2.4 (закон повторного логарифма). *Если $\zeta_1, \zeta_2, \ldots$ — независимые случайные величины, $\zeta_n \sim \mathcal{N}(0,1)$, $S_n = \zeta_1 + \cdots + \zeta_n$, то*

$$\mathsf{P}\left\{\limsup_{n\to\infty}\frac{S_n}{\sqrt{2n\ln\ln n}} = 1\right\} = 1 \qquad \text{и} \qquad \mathsf{P}\left\{\liminf_{n\to\infty}\frac{S_n}{\sqrt{2n\ln\ln n}} = -1\right\} = 1.$$

Заметим, что $S_n \sim \mathcal{N}(0,n)$, поэтому $\dfrac{S_n}{\sqrt{n}} \sim \mathcal{N}(0,1)$. Положим $M(t) = \max\limits_{1\leqslant k\leqslant t} S_k$, $h(t) = \sqrt{2t\ln\ln t}$ $(t > 0)$. Сформулируем две вспомогательные леммы.

Лемма 2.5. *Для всех $r > 1$*

$$\sum_{k=1}^{\infty}\mathsf{P}\left\{\frac{M(r^{k+1})}{h(r^k)} > r\right\} < \infty.$$

Лемма 2.6. *Для всех натуральных v*

$$\sum_{k=1}^{\infty}\mathsf{P}\left\{\frac{S_{v^k} - S_{v^{k-1}}}{h(v^k)} > c(v)\right\} = \infty, \qquad \text{где } c(v) = \sqrt{1-\frac{1}{v}}.$$

Эти леммы мы докажем потом, а сейчас с их помощью докажем теорему.

□ [Доказательство закона повторного логарифма] Пусть $r > 1$. Из леммы 2.5 и леммы Бореля–Кантелли имеем, что

$$\mathsf{P}\left\{\limsup_{k\to\infty}\frac{M(r^{k+1})}{h(r^k)}\leqslant r\right\}=1.$$

Положим $k(n)=[\log_r n]$, т.е. $r^{k(n)}\leqslant n<r^{1+k(n)}$ (квадратные скобки обозначают взятие целой части). $k(n)\longrightarrow\infty$ при $n\longrightarrow\infty$. Имеем:

$$\frac{S_n}{h(n)}\leqslant\frac{M(r^{k(n)+1})}{h(r^{k(n)})}\quad\Rightarrow\quad\limsup_{n\to\infty}\frac{S_n}{h(n)}\leqslant\limsup_{n\to\infty}\frac{M(r^{n+1})}{h(r^n)}.$$

Отсюда

$$\forall\, r>1\quad\mathsf{P}\left\{\limsup_{n\to\infty}\frac{S_n}{h(n)}\leqslant r\right\}=1\qquad\Rightarrow\qquad\mathsf{P}\left\{\limsup_{n\to\infty}\frac{S_n}{h(n)}\leqslant 1\right\}=1.$$

Пусть v — целое число, $v>1$. $S_{v^k}-S_{v^{k-1}}=\sum\limits_{n=v^{k-1}+1}^{v^k}\zeta_n$ — семейство независимых случайных величин. По лемме 2.6 и лемме Бореля–Кантелли

$$\mathsf{P}\left\{\limsup_{k\to\infty}\frac{S_{v^k}-S_{v^{k-1}}}{h(v^k)}>c(v)\right\}=1.$$

Из соотношения $h(v^k)=\sqrt{2v^k\ln\ln v^k}=(1+o(1))h(v^{k-1})\sqrt{v}$ (при $k\longrightarrow\infty$) и ранее доказанного получаем:

$$\mathsf{P}\left\{\limsup_{k\to\infty}\frac{S_{v^{k-1}}}{h(v^k)}=\limsup_{k\to\infty}\frac{S_{v^{k-1}}}{h(v^{k-1})\sqrt{v}}\leqslant\frac{1}{\sqrt{v}}\right\}=1.$$

Так как распределение $\mathcal{N}(0,1)$ симметрично относительно нуля, получаем:

$$\mathsf{P}\left\{\liminf_{k\to\infty}\frac{S_{v^{k-1}}}{h(v^k)}\geqslant-\frac{1}{\sqrt{v}}\right\}=1.$$

Из анализа известно, что если $\limsup\limits_{k\to\infty}a_k>c$ и $\liminf\limits_{k\to\infty}b_k>d$, то $\limsup\limits_{k\to\infty}(a_k+b_k)>c+d$. Отсюда

$$\mathsf{P}\left\{\limsup_{k\to\infty}\frac{S_{v^k}}{h(v^k)}>c(v)-\frac{1}{\sqrt{v}}\right\}=1\quad\forall\, v<\infty\qquad\Rightarrow\qquad\forall\, v<\infty\quad\mathsf{P}\left\{\limsup_{n\to\infty}\frac{S_n}{h(n)}>c(v)-\frac{1}{\sqrt{v}}\right\}=1.$$

$c(v)=\sqrt{1-\frac{1}{v}}$, поэтому $c(v)-\frac{1}{\sqrt{v}}$ близко к 1 при достаточно больших v, откуда получаем, что

$$\mathsf{P}\left\{\limsup_{n\to\infty}\frac{S_n}{h(n)}\geqslant 1\right\}=1.$$

Окончательно,

$$\mathsf{P}\left\{\limsup_{n\to\infty}\frac{S_n}{h(n)}=1\right\}=1.$$

Второе утверждение следует из первого в силу симметрии стандартного нормального распределения. ■

Лемма 2.7 (аналог неравенства Колмогорова). *Если $\xi_1,\dots,\xi_n$ — независимые случайные величины, имеющие симметричное распределение ($\mathsf{P}\{\xi_k\leqslant x\}=\mathsf{P}\{\xi_k\geqslant -x\}$ для всех x), то для всех $a>0$ $\mathsf{P}\{\max\limits_{k=1,\dots,n}S_k>a\}\leqslant 2\mathsf{P}\{S_n>a\}$, где $S_k=\xi_1+\dots+\xi_k$.*

□ Пусть τ — случайная величина, равная номеру первого члена последовательности $\{S_k\}$, который больше a (или $n+1$, если такого нет):

$$\tau=\begin{cases}k, & \text{если } S_1,\dots,S_{k-1}\leqslant a,\ S_k>a;\\ n+1, & \text{если } S_1,\dots,S_n\leqslant a.\end{cases}$$

Случайные величины S_k и S_n-S_k независимы и имеют симметричное распределение. Из симметрии $\mathsf{P}\{S_n-S_k\geqslant 0\}=\mathsf{P}\{S_n-S_k\leqslant 0\}\geqslant\frac{1}{2}$. Далее, $\mathsf{P}\{S_n>a,\tau=k\}\geqslant\mathsf{P}\{S_n\geqslant S_k>a,\tau=k\}=\mathsf{P}\{\tau=k,S_n-S_k\geqslant 0\}=\mathsf{P}\{\tau=k\}\mathsf{P}\{S_n-S_k\geqslant 0\}\geqslant\frac{1}{2}\mathsf{P}\{\tau=k\}$, где предпоследнее равенство имеет место в силу независимости события $\{\tau=k\}$ от S_n-S_k. Суммируя по k, получаем:

$$\mathsf{P}\{S_n>a\}=\sum_{k=1}^{n}\mathsf{P}\{S_n>a,\tau=k\}\geqslant\sum_{k=1}^{n}\frac{1}{2}\mathsf{P}\{\tau=k\}=\frac{1}{2}\mathsf{P}\{\max S_k>a\},$$

что и требовалось. ■

□ [Доказательство леммы 2.5] По утверждению 2.3

$$1-\Phi(u) \leqslant \frac{1}{u\sqrt{2\pi}}e^{-\frac{u^2}{2}} < \frac{1}{2u}e^{-\frac{u^2}{2}}.$$

$S_n/\sqrt{n} \sim \mathcal{N}(0,1)$, поэтому в силу леммы 2.7

$$\mathsf{P}\{M(n) > u\sqrt{n}\} \leqslant 2\mathsf{P}\{S_n > u\sqrt{n}\} = 2\mathsf{P}\left\{\frac{S_n}{\sqrt{n}} > u\right\} = 2(1-\Phi(u)) < \frac{1}{u}e^{-\frac{u^2}{2}}.$$

Полагая $n = r^{k+1}$ и $u = \sqrt{2r\ln\ln r^k}$, получаем:

$$\mathsf{P}\left\{\frac{M(r^{k+1})}{h(r^k)} > r\right\} = \mathsf{P}\left\{M(r^{k+1}) > rh(r^k)\right\} = \mathsf{P}\left\{M(r^{k+1}) > \sqrt{r^{k+1}}\sqrt{2r\ln\ln r^k}\right\} <$$

$$< \frac{e^{-r\ln\ln r^k}}{\sqrt{2r\ln\ln r^k}} = \frac{1}{\sqrt{2r\ln\ln r^k}(k\ln r)^r} = O\left(\frac{1}{k^r}\right),$$

а этот ряд сходится при $r > 1$. ■

□ [Доказательство леммы 2.6] $S_{v^k} - S_{v^{k-1}} \sim \mathcal{N}(0, v^k - v^{k-1})$, $\zeta_1 \sim \mathcal{N}(0,1)$, $\zeta_1\sqrt{v^k - v^{k-1}} \sim \mathcal{N}(0, v^k - v^{k-1})$. Далее (используем совпадение распределений и опять применяем утверждение 2.3),

$$\mathsf{P}\left\{\frac{S_{v^k} - S_{v^{k-1}}}{h(v^k)} > c(v)\right\} = \mathsf{P}\left\{\frac{\zeta_1\sqrt{v^k - v^{k-1}}}{h(v^k)} > c(v)\right\} = \mathsf{P}\left\{\zeta_1 > \frac{c(v)h(v^k)}{\sqrt{v^k - v^{k-1}}}\right\} =$$

$$= \mathsf{P}\left\{\zeta_1 > \sqrt{2\ln\ln v^k}\right\} \sim \frac{1}{2\sqrt{\pi\ln\ln v^k}}e^{-\ln\ln v^k} = \frac{1}{\sqrt{\pi\ln k}\,k\ln v},$$

а этот ряд расходится. ■

3. Цепи Маркова

3.1. Определения

Пусть $\xi_0, \xi_1, \ldots$ — последовательность случайных величин, принимающих значения в конечном множестве. Распределение этой последовательности определяется счётным семейством совместных распределений её конечных отрезков, т.е. вероятностями $P(\xi_0 = i_0, \ldots, \xi_n = i_n)$. Эти вероятности можно записать также как

$$\mathsf{P}\{\xi_0 = i_0, \ldots, \xi_n = i_n\} = \mathsf{P}\{\xi_0 = i_0\}\prod_{k=1}^{n}\mathsf{P}\{\xi_k = i_k \mid \xi_0 = i_0, \ldots, \xi_{k-1} = i_{k-1}\} \qquad (*)$$

Если случайные величины ξ_i независимы в совокупности, то эта формула значительно упрощается и принимает вид

$$\mathsf{P}\{\xi_0 = i_0, \ldots, \xi_n = i_n\} = \prod_{k=0}^{n}\mathsf{P}\{\xi_k = i_k\}$$

Цепи Маркова — это случайные последовательности, занимающие промежуточное положение между полностью зависимыми и полностью независимыми.

Определение. Цепь Маркова с конечным или счётным множеством состояний S и дискретным временем — это такая последовательность случайных величин $\xi_0, \xi_1, \ldots$, принимающих значения в S, что

$$\forall t \geqslant 0\,\forall\, i_0, \ldots, i_{t+1} \in S \quad \mathsf{P}\{\xi_{t+1} = i_{t+1} \mid \xi_0 = i_0, \ldots, \xi_t = i_t\} = \mathsf{P}\{\xi_{t+1} = i_{t+1} \mid \xi_t = i_t\} = p^{(t)}_{i_t i_{t+1}}.$$

$p^{(t)}_{i_t i_{t+1}}$ называются *переходными вероятностями* этой цепи.
Если $p^{(t)}_{i_t i_{t+1}}$ не зависит от t, то цепь называется *однородной по времени.*

Для цепей Маркова формула $(*)$ упрощается и принимает вид

$$\mathsf{P}\{\xi_0 = i_0, \ldots, \xi_t = i_t\} = \mathsf{P}\{\xi_0 = i_0\}\prod_{k=1}^{t}p^{(k)}_{i_{k-1}i_k}$$

Для удобства обозначим $p_{i_0}^{(0)} := \mathsf{P}\left\{\xi_0 = i_0\right\}$.
Распределение однородной цепи Маркова в случае $S = 1, \ldots, n$ задаётся $n + n^2$ параметрами, а именно, начальным вектором $\bar{p}^{(0)} = (p_1^{(0)}, \ldots, p_n^{(0)})$ и матрицей переходных вероятностей $P = (p_{ij})$. В случае неоднородности по времени можно задать счётное семейство матриц переходных вероятностей $P^{(t)} = (p_{ij}^{(t)})$.

Определение. Квадратная матрица $A = (a_{ij})$, такая, что $a_{ij} \geqslant 0$ и $\forall i \; \sum_k a_{ik} = 1$, называется *стохастической*. Если вместо второго условия выполнено только $\forall i \; \sum_k a_{ik} \leqslant 1$, то матрица называется *полустохастической*. Наконец, если как A, так и A^T стохастические, то A называется *дважды стохастической.*

Очевидно, что матрицы переходных вероятностей цепей Маркова являются стохастическими и наоборот, любая стохастическая матрица задаёт некоторую однородную по времени цепь Маркова.

Замечание. Если P, Q — стохастические матрицы и определено их произведение PQ, то PQ также является стохастической.

3.2. Примеры цепей Маркова

1. Пусть ξ_t — независимые целочисленнные случайные величины. Тогда $S_t = \xi_1 + \ldots + \xi_t$ образуют цепь Маркова.

2. $\eta_1, \ldots$ — независимые одинаково распределённые случайные величины со значениями в $\{\pm 1\}$, $\mathsf{P}\left\{\eta_i = 1\right\} = p$, $\mathsf{P}\left\{\eta_i = -1\right\} = q$.
$$\xi_t = \max\left\{\xi_{t-1} + \eta_t, 0\right\}, \quad t \in \mathbb{N}$$
Такой процесс называется *случайным блужданием с отражающим экраном в нуле.* Найдём матрицу его переходов:
$$\mathsf{P}\left\{\xi_{t+1} = i + 1 \mid \xi_t = i\right\} = p,$$
$$\mathsf{P}\left\{\xi_{t+1} = i - 1 \mid \xi_t = i\right\} = \begin{cases} q, & i \geqslant 1 \\ 0, & i = 0 \end{cases}$$
$$\mathsf{P}\left\{\xi_{t+1} = i \mid \xi_t = i\right\} = \begin{cases} 0, & i > 0 \\ q, & i = 0 \end{cases}$$
Матрица (бесконечная) вероятностей переходов будет иметь вид
$$\begin{pmatrix} q & p & 0 & 0 & \ldots \\ q & 0 & p & 0 & \ldots \\ 0 & q & 0 & p & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

3.2.1. Модель Эренфестов для диффузии

Рассмотрим следующую модель распространения газа между двумя сосудами, в одном из которых изначально вакуум, а в другом — n частиц. Каждый момент времени будем брать случайную частицу и перемещать её в другой сосуд. Последовательность $\xi_t = \{$число частиц в 1 сосуде в момент времени $t\}$ является цепью Маркова.
$$\mathsf{P}\left\{\xi_{t+1} = k + 1 \mid \xi_t = k\right\} = \frac{n-k}{n} = 1 - \frac{k}{n}$$
$$\mathsf{P}\left\{\xi_{t+1} = k - 1 \mid \xi_t = k\right\} = \frac{k}{n}$$
Матрица вероятностей переходов:
$$\begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ \frac{1}{n} & 0 & 1 - \frac{1}{n} & 0 & \ldots & 0 \\ 0 & \frac{2}{n} & 0 & 1 - \frac{2}{n} & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 - \frac{1}{n} & 0 & \frac{1}{n} \\ 0 & 0 & \ldots & 0 & 1 & 0 \end{pmatrix}$$

3.3. Свойства матрицы переходных вероятностей

Рассмотрим однородную цепь Маркова ξ_t $(S = 1, \ldots, n, P = (p_{ij}))$ с дискретным временем и зададимся вопросом посчитать вероятности перехода между состояниями за m шагов, т.е. найдём
$$p_{ij}(m) := \mathsf{P}\left\{\xi_{t+m} = j \mid \xi_t = i\right\}$$

Пусть $P(m) = (p_{ij}(m)), p_i(t) = \mathsf{P}\{\xi_t = i\}, \bar{p}(t) = (p_i(t), i \in S)$.

Лемма 3.1. *Для однородной по времени цепи Маркова с матрицей переходных вероятностей P $\forall m \geqslant 1$ имеем*

$$P(m) = P^m, \quad \bar{p}(t+m) = \bar{p}(t)P^m.$$

Если цепь не однородна, то имеют место аналогичные равенства:

$$(p_{ij}^{t,m}) = (\mathsf{P}\{\xi_{t+m} = j \mid \xi_t = i)\}) = P^{(t)} \dots P^{(t+m-1)}, \quad \bar{p}(t+m) = \bar{p}(t)P^{(t)} \dots P^{(t+m-1)}.$$

□ Докажем только первое равенство, остальные доказываются абсолютно аналогично. Рассуждаем по индукции (база, случай $m = 1$, очевидна). Имеем

$$p_{ij}(m+1) = \sum_{k \in S} p_{ik}(m)p_{kj} = \left(P^{(m)}P\right)_{ij},$$

что и требовалось. ■

3.4. Классификация состояний цепей Маркова

Пусть имеется (однородная) цепь Маркова с дискретным временем, S — множество её состояний, $p_{ij}(m)$ — вероятность перехода между состояниями i и j за m шагов.

Определение. Состояние j *следует* за состоянием i ($i \to j$), если $\exists m : p_{ij}(m) > 0$.
Если $i \to j, j \to i$, то состояния i и j называются *сообщающимися* ($i \leftrightarrow j$)). Легко видеть, что $\to$ транзитивно, а $\leftrightarrow$ задаёт на S отношение эквивалентности.
Состояние i называется:

- *поглощающим*, если $p_{ii} = 1$.
- *периодическим* с периодом $d > 1$, если НОД$(m : p_{ii}(m) > 0) = d$.
- *непериодическим*, если оно периодическое с периодом 1.
- *несущественным*, если $\exists j \in S : i \to j, j \not\to i$.
- *существенным*, если оно не является несущественным.
- *возвратным*, если оно существенно и $\mathsf{P}\{\exists t : \xi_t = i \mid \xi_0 = i\} = 1$.
- *возвратным нулевым*, если оно возвратно и $p_{ii}(t) \to 0(t \to \infty)$.
- *возвратным положительным*, если оно возвратно и $\limsup_{t\to\infty} p_{ii}(t) \geqslant 0$.

Рассмотрим теперь множество классов эквивалентности в S, задаваемых отношением $\leftrightarrow$. На них можно задать структуру ориентированного графа, направив ребро из K в L, когда $\exists x \in K, y \in L : x \to y$.

Определение. *Финальными* назовём те классы, из которых в этом графе не выходит рёбер.

Определение. Цепь Маркова, у которой все состояния образуют один класс, называется *неразложимой*.

3.4.1. Примеры

1. **Периодические цепи**

   Рассмотрим цепь Маркова с тремя состояниями и матрицей переходных вероятностей

   $$P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} .....$$

   Очевидно, $P^3 = E$ и $\forall\, a, k : P^{a+3k} = P^a$, то есть процесс циклически переходит между 3 состояниями. Этот пример тривиален, однако по его подобию можно построить более сложный циклический процесс, а именно, пусть P_1,P_2 и P_3 — стохастические матрицы (одинакового размера). Рассмотрим цепь Маркова со следующей блочной матрицей переходов:

   $$P^* = \begin{pmatrix} 0 & 0 & P_1 \\ P_2 & 0 & 0 \\ 0 & P_3 & 0 \end{pmatrix} .....$$

   Если матрицы P_i задают неразложимые цепи, то в новой цепи граф классов эквивалентности будет иметь такой же вид, что и в предыдущем примере, хотя внутри классов процесс перемещаться будет случайно.

2. **Несущественные состояния**

Рассмотрим для $\alpha, \beta, \gamma > 0$ цепь с матрицей переходов

$$\begin{pmatrix} \alpha & \beta & \gamma \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .....$$

Видно, что состояние 1 несущественно, а 2 и 3 — поглощающие. При этом

$$p_{11}(t) = \alpha^t \longrightarrow 0 \quad (t \longrightarrow \infty)$$
$$p_{12}(t) = (1-\alpha)^t \frac{\beta}{\beta+\gamma} \longrightarrow \frac{\beta}{\beta+\gamma} \quad (t \longrightarrow \infty)$$
$$p_{13}(t) = (1-\alpha)^t \frac{\gamma}{\beta+\gamma} \longrightarrow \frac{\gamma}{\beta+\gamma} \quad (t \longrightarrow \infty)$$

На основе этого процесса, взяв стохастические матрицы соотвествующих размеров, можно аналогичным образом получить более сложный процесс с тем же графом классов эквивалентности. Он будет иметь матрицу

$$\begin{pmatrix} P_{11} & P_{12} & P_{13} \\ 0 & P_{22} & 0 \\ 0 & 0 & P_{33} \end{pmatrix}$$

3.4.2. Критерий возвратности состояния

Теорема 3.2 (Критерий возвратности состояния). *Состояние j однородной цепи Маркова возвратно тогда и только тогда, когда*

$$\sum_{m=1}^{\infty} p_{jj}(m) = \infty.$$

Обозначим $T_0 = 0, T_k = \min\{T \mid T > T_{k-1}, \xi_T = j\}$ (считаем $\min \varnothing = \infty$).
Для доказательства этой теоремы нам понадобится следующая

Лемма 3.3. *Если цепь Маркова однородна, то при условии $\xi_0 = j$ случайные величины $\Delta_k = T_k - T_{k-1}$ (где k такое, что $T_k < \infty$) независимы и одинаково распределены.*

$\square$ По определению T_k $T_k = t \Rightarrow \xi_t = j$.

$$\mathsf{P}\{\xi_{t+v} = i_{t+v}, v = 1, \dots, m \mid \xi_0 = j, \dots, \xi_t = j\} = \mathsf{P}\{\xi_{t+v} = i_{t+v}, v = 1, \dots, m \mid \xi_t = j\} =$$
$$= \mathsf{P}\{\xi_v = i_{t+v}, v = 1, \dots, m \mid \xi_0 = j\} \quad \forall m, \forall i_1, \dots, i_{t+m} \in S,$$

то есть $\forall t: \ T_k = t$ распределение последовательности $\{\xi_{t+1}, \dots\}$ совпадает с распределением последовательности $\{\xi_1, \dots\}$, а это означает, что $\Delta_k = T_{k+1} - T_k$ распределена так же, как $T_1 - T_0 = T_1$.

Независимость Δ_k в совокупности следует из определения цепи Маркова. Докажем, например, независимость Δ_1 и Δ_2:

$$\mathsf{P}\{\Delta_1 = m, \Delta_2 = n\} = \mathsf{P}\{\xi_0 = j, \xi_1 \neq j, \dots, \xi_m = j, \xi_{m+1} \neq j, \dots, \xi_{m+n} = j\} =$$
$$= \sum_{x_1,\dots,x_{m-1} \neq j} \sum_{y_1,\dots,y_{n-1} \neq j} \mathsf{P}\{\xi_{m+1} = y_1, \dots, \xi_{m+n} = j \mid \xi_0 = j, \xi_1 = x_1, \dots, \xi_m = j\} \mathsf{P}\{\xi_0 = j, \dots, \xi_m = j\} =$$
$$= \mathsf{P}\{\Delta_1 = m\} \sum_{y_1,\dots,y_{n-1} \neq j} \mathsf{P}\{\xi_{m+1} = y_1, \dots, \xi_{m+n} = j \mid \xi_m = j\} = \mathsf{P}\{\Delta_1 = m\} \mathsf{P}\{\Delta_2 = n\}.$$

$\blacksquare$

Докажем теперь нашу теорему.

$\square$ [Доказательство теоремы 3.2] При условии $\xi_0 = j$ $\{\omega : \exists t > 0 : \ \xi_t = j\} = \{\omega : \ \Delta_1 < \infty\}$.
Поэтому j возвратно $\Leftrightarrow \mathsf{P}\{\Delta_1 < \infty\} = 1$.
Положим

$$I_n(\omega) = \begin{cases} 1, & \xi_n = j, \\ 0 & \text{иначе.} \end{cases}$$

$\forall N < \infty$ положим $\nu(N) = \sum_{n=1}^{N} I_n = \max\{k \mid T_k \leqslant N\}$.

$$\mathsf{M}\{I_n \mid \xi_0 = j\} = \mathsf{P}\{\xi_n = j \mid \xi_0 = j\} = p_{jj}(n), \ \mathsf{M}\nu(N) = \sum_{n=1}^{N} p_{jj}(n).$$

Если $\mathsf{P}\{\Delta_1 < \infty\} = 1$, то $\forall k < \infty\ \mathsf{P}\{T_k < \infty\} = 1$, следовательно,

$$\forall k < \infty \exists N(k) : \mathsf{P}\{T_k < N(k)\} = \mathsf{P}\{\nu(N(k)) \geqslant k\} > \frac{1}{2} \Rightarrow M(\nu(N(k))) \geqslant \frac{k}{2}.$$

Так как $N(k) \longrightarrow \infty$ $(k \longrightarrow \infty)$, то

$$\sum_{n=1}^{N} p_{jj}(n) \longrightarrow \infty \quad (N \longrightarrow \infty).$$

Если же $p = \mathsf{P}\{\Delta_1 < \infty\} < 1$, то рассмотрим

$$\nu = \lim_{N \longrightarrow \infty} \nu(N) = \max\{k : T_k < \infty\}$$
$$\mathsf{P}\{\nu = N\} = \mathsf{P}\{\min\{k : \Delta_k = \infty\} = N+1\} = p^n(1-p),$$

так как Δ_i независимы. Но тогда

$$\forall N \sum_{k=1}^{N} p_{jj}(k) = \mathsf{M}\nu(N) \leqslant \mathsf{M}\nu = \frac{p}{1-p} < \infty,$$

что завершает доказательство. ■

Пример 4.1. *Случайное блуждание по целочисленным точкам.* $\xi_1, \xi_2, \ldots$ — независимые случайные величины; $\mathsf{P}\{\xi_k = 1\} = p$, $\mathsf{P}\{\xi_k = -1\} = q = 1-p$. $S_0 = 0$, $S_n = S_{n-1} + \xi_n = \xi_1 + \cdots + \xi_n$ — случайное блуждание. Это цепь Маркова с множеством состояний $S = \mathbb{Z}$. Из симметрии задачи ясно, что либо все состояния будут возвратными, либо все не будут возвратными одновременно. Применим ранее доказанный критерий для определения возвратности состояния 0.

$$p_{00}(n) = \mathsf{P}\{S_n = 0\} = \begin{cases} 0, & \text{если } n \text{ нечётно;} \\ C_n^{\frac{n}{2}} p^{\frac{n}{2}} q^{\frac{n}{2}}, & \text{если } n \text{ чётно.} \end{cases}$$

Нас интересует сходимость ряда $\sum\limits_{n=1}^{\infty} C_{2n}^n p^n q^n$. Из анализа известна формула Стирлинга: $k! = (1+o(1))\sqrt{2\pi k}\left(\frac{k}{e}\right)^k$ при $k \longrightarrow \infty$. Отсюда получаем:

$$C_{2n}^n \sim \frac{\sqrt{2\pi \cdot 2n}\left(\frac{2n}{e}\right)^{2n}}{2\pi n\left(\frac{n}{e}\right)^{2n}} = \frac{2^{2n}}{\sqrt{\pi n}},$$

откуда $C_{2n}^n p^n q^n \sim \dfrac{(4pq)^n}{\sqrt{\pi n}}$, то есть сходимость нашего ряда равносильна сходимости ряда $\sum\limits_{n=1}^{\infty} \dfrac{(4pq)^n}{\sqrt{\pi n}}$.

Если $p \neq q$, то $pq = p(1-p) < \frac{1}{4}$, следовательно $4pq < 1$ и ряд сходится как геометрическая прогрессия. Если же $p = q$, то ряд принимает вид $\sum \frac{1}{\sqrt{\pi n}}$ и расходится. Применяя критерий возвратности, получаем, что любое состояние возвратно тогда и только тогда, когда блуждание симметрично $(p = q)$.

Пример 4.2. *Двумерное целочисленное случайное блуждание.* $\xi_1, \xi_2, \ldots$ независимы; $\mathsf{P}\{\xi_k = 1\} = p$, $\mathsf{P}\{\xi_k = -1\} = q = 1-p$. $S_0 = (0,0)$, $S_n = S_{n-1} + (\xi_{2n-1}, \xi_{2n}) = (\xi_1, \xi_2) + (\xi_3, \xi_4) + \cdots + (\xi_{2n-1}, \xi_{2n})$. Множество состояний $S = \mathbb{Z}^2$.

$S_n = (S_n^{(1)}, S_n^{(2)})$, $S_n^{(1)} = \sum\limits_{k=1}^{n} \xi_{2k-1}$, $S_n^{(2)} = \sum\limits_{k=1}^{n} \xi_{2k}$. $S_n^{(1)}$ и $S_n^{(2)}$ суть одномерные блуждания из предыдущего примера. $p_{(0,0)(0,0)}(2n-1) = 0$. В силу независимости $p_{(0,0)(0,0)}(2n) = \mathsf{P}\{S_n^{(1)} = 0, S_n^{(2)} = 0\} = p_{00}^2(n) \sim \dfrac{(4pq)^{2n}}{\pi n}$. Условия сходимости/расходимости ряда те же, что и в первом случае.

А вот в трёхмерном случае соответствующая вероятность возврата $\sim \dfrac{(4pq)^{3n}}{(\pi n)^{3/2}}$ — ряд сходится при всех p, q, то есть трёхмерное блуждание уже всегда будет невозвратным.

3.5. Предельная теорема для конечных цепей Маркова

Теорема 3.4. *Пусть $\{\xi_k\}$ — цепь Маркова со множеством состояний $S = \{1, \ldots, n\}$ и матрицей переходных вероятностей $P = (p_{ij})$. Если существует такое $v < \infty$, что все элементы матрицы P^v положительны, то существует предел $\lim\limits_{t\to\infty} p_{ij}(t) = \pi_j > 0$, $j \in S$, не зависящий от начального состояния i, причём $(\pi_1, \ldots, \pi_n)$ — единственное решение системы уравнений:*

$$\begin{cases} \sum_{k=1}^{n} x_k p_{kj} = x_j, \quad j \in S \\ \sum_{k=1}^{n} x_k = 1 \end{cases}$$

□ $p(t) = (p_1(t), \ldots, p_n(t)) = p(0)P^t$ для всех натуральных t ($p(0)$ — произвольное начальное состояние). Нужно показать, что $p(0)P^t \longrightarrow \pi = (\pi_1, \ldots, \pi_n)$ при всех возможных $p(0)$.

Поделим t на v с остатком: $t = kv + r$, $0 \leqslant r < v$. $pP^t = pP^{kv+r} = (pP^r)(P^v)^k$. Если мы докажем, что для всех x $x(P^v)^k \longrightarrow \pi$, то это будет верно и для pP^t (разбиваем P^t на подпоследовательности). Поэтому далее все выкладки будут производиться с матрицей P^v.

Рассмотрим оператор $A\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$, действующий так: $Ax = xP^v$. $G_n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_j \geqslant 0, \sum x_j = 1\} \subset \mathbb{R}^n$ — множество вероятностных распределений на n-элементном множестве. Покажем, что G_n инвариантно относительно A.

Утверждение 3.5. $A(G_n) \subseteq G_n$.

□ Пусть $x \in G_n$. $y = Ax = xP^v = (y_1, \ldots, y_n)$. $y_j = \sum\limits_{k=1}^{n} x_k p_{kj}(v) \geqslant 0$. Первое свойство проверено. Далее,

$$\sum_{j=1}^{n} y_j = \sum_{j=1}^{n}\sum_{k=1}^{n} x_k p_{kj}(v) = \sum_k x_k \underbrace{\sum_j p_{kj}(v)}_{=1} = \sum_k x_k = 1.$$

Отсюда $y \in G_n$. ■

Введём в $\mathbb{R}^n$ метрику $\rho(x, y) = \sum\limits_{k=1}^{n} |x_k - y_k|$.

Утверждение 3.6. *В метрике ρ отображение A с матрицей P^v является сжимающим с коэффициентом сжатия $\leqslant 1 - \varepsilon$, где $\varepsilon = \min\limits_{i,j} p_{ij}(v) > 0$.*

□ Вычислим $\rho(xP^v, yP^v)$:

$$\rho(xP^v, yP^v) = \sum_{k=1}^{n}\left|\sum_{i=1}^{n} x_i p_{ik}(v) - \sum_{i=1}^{n} y_i p_{ik}(v)\right| = \sum_{k=1}^{n}\left|\sum_{i=1}^{n}(x_i - y_i)p_{ik}(v)\right|.$$

Далее, т.к. $x, y \in G_n$, $\sum\limits_{i=1}^{n}(x_i - y_i) = 1 - 1 = 0$. Поэтому суммы положительных и отрицательных разностей в этой сумме совпадают:

$$\sum_{i=1}^{n}\max\{x_i - y_i, 0\} = -\sum_{i=1}^{n}\min\{x_i - y_i, 0\} = \frac{1}{2}\sum_{i=1}^{n}|x_i - y_i| = \frac{1}{2}\rho(x, y).$$

Значит, существуют такие индексы r и s, что

$$x_r - y_r \geqslant \frac{1}{2n}\rho(x, y) > 0 > -\frac{1}{2n}\rho(x, y) \geqslant x_s - y_s.$$

Для всех $a, b, c, d > 0$ верны следующие соотношения: $|a - b| = a + b - 2\min\{a, b\}$, $\min\{ac, bd\} \geqslant \min\{a, b\}\min\{c, d\}$. Отсюда $|(x_r - y_r)p_{rk}(v) + (x_s - y_s)p_{sk}(v)| \leqslant (x_r - y_r)p_{rk}(v) + |x_s - y_s|p_{sk}(v) - 2\min\{|x_r - y_r|, |x_s - y_s|\} \cdot \min\{p_{sk}(v), p_{rk}(v)\} \leqslant (x_r - y_r)p_{rk}(v) + |x_s - y_s|p_{sk}(v) - \frac{1}{n}\rho(x, y)\varepsilon$. Итак,

$$\left|\sum_{i=1}^{n}(x_i - y_i)p_{ik}(v)\right| \leqslant \sum_{i=1, i\neq r,s}^{n}|x_i - y_i|p_{ik}(v) + |x_r - y_r|p_{rk}(v) + |x_s - y_s|p_{sk}(v) - \frac{\varepsilon}{n}\rho(x, y) = \sum_{i=1}^{n}|x_i - y_i|p_{ik}(v) - \frac{\varepsilon}{n}\rho(x, y).$$

Далее,

$$\rho(xP^v, yP^v) \leqslant \sum_{k=1}^{n}\left(\sum_{i=1}^{n}|x_i - y_i|p_{ik}(v) - \frac{\varepsilon}{n}\rho(x, y)\right) = (1 - \varepsilon)\rho(x, y),$$

что и требовалось. ■

Сжимающее отображение имеет единственную неподвижную точку, поэтому имеет место предельное соотношение. Осталось разобраться с системой уравнений.

π — неподвижная точка отображения A. Следовательно, $\pi P^v = \pi$ и для всех p $pP^t \longrightarrow \pi$ $(t \longrightarrow \infty)$. В частности, $(\pi P)P^{vt} \longrightarrow \pi$, поэтому πP — неподвижная точка матрицы P^v, а эта точка единственна, следовательно $\pi P = \pi$, то есть π — неподвижная для P. ■

Пример 5.1. *Исправленная модель Эренфестов для диффузии.* ξ_t — число частиц в первом сосуде в момент t (всего n частиц). Случайно выбираем частицу и с вероятностью $\frac{1}{2}$ перекладываем её (с вероятностью $\frac{1}{2}$ ничего не делаем). Матрица имеет вид:

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & \ldots & 0 \\ \frac{1}{2n} & \frac{1}{2} & \frac{1}{2} - \frac{1}{2n} & 0 & \ldots & 0 \\ 0 & \frac{2}{2n} & \frac{1}{2} & \frac{1}{2} - \frac{2}{2n} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix}.$$

(Мы исправили модель, чтобы появилась диагональ и выполнялись условия теоремы.) Вычислим предельные распределения $\pi = (\pi_1, \ldots, \pi_n)$ из системы уравнений

$$\begin{cases} \pi_1 = \frac{1}{2}\pi_0 + \frac{1}{2}\pi_1 + \frac{2}{2n}\pi_2 \\ \ldots \\ \pi_k = \frac{n-k+1}{2n}\pi_{k-1} + \frac{1}{2}\pi_k + \frac{k+1}{2n}\pi_{k+1} \quad (1 \leqslant k \leqslant n-1) \\ \ldots \\ \pi_n = \frac{1}{2n}\pi_{n-1} + \frac{1}{2}\pi_n \\ \sum \pi_k = 1 \end{cases}$$

Преобразуем:

$$\begin{cases} \pi_0 = \frac{1}{n}\pi_1 \\ \pi_1 = \pi_0 + \frac{2}{n}\pi_2 \\ \ldots \\ \pi_k = \frac{n-k+1}{n}\pi_{k-1} + \frac{k+1}{n}\pi_{k+1} \\ \ldots \\ \pi_n = \frac{1}{n}\pi_{n-1} \end{cases}$$

Индукцией показываем, что $\pi_k = C_n^k \pi_0$:

$$\pi_{k+1} = \frac{n}{k+1}\left(C_n^k - \frac{n-k+1}{n}C_n^{k-1}\right)\pi_0 = \frac{n}{k+1}C_{n-1}^{k-1}\left(\frac{n}{k} - 1\right)\pi_0 = C_n^{k+1}\pi_0.$$

Из соотношения $\sum \pi_k = 1$ получаем: $1 = \sum C_n^k \pi_0 = 2^n \pi_0$, откуда $\pi_0 = \frac{1}{2^n}$, $\pi_k = C_n^k \frac{1}{2^n}$, т.е предельное распределение цепи биномиально.

При больших значениях n (числа частиц) вероятность того, что все молекулы соберутся в одном сосуде, (равная $\pi_0 = \frac{1}{2^n}$) исчезающе мала.

4. Ветвящиеся процессы

Имеются частицы, которые делятся на поколения. В каждом поколении каждая частица порождает некоторое (случайное) количество потомков, а сама погибает.

Обозначим через $\mu(t)$ число частиц в t-м поколении, γ — случайное число потомков у одной частицы (γ — неотрицательная целочисленная случайная величина). k-я частица в t-м поколении порождает γ_{tk} частиц; все величины γ_{tk} независимы и распределены так же, как γ. В начальный момент имеется $\mu(0)$ частиц (обычно $\mu(0) = 1$). Каждая частица независимо порождает потомков:

$$\mu(t+1) = \begin{cases} 0, & \text{если } \mu(t) = 0, \\ \gamma_{t1} + \cdots + \gamma_{tk}, & \text{если } \mu(t) = k. \end{cases}$$

Возникают следующие вопросы:

1. Когда процесс вырождается?
2. Насколько быстро растёт $\mu(t)$?

4.1. Производящие функции

Пусть ξ — целочисленная неотрицательная случайная величина. Положим

$$f_\xi(s) = \sum_{k=0}^{\infty} s^k \mathsf{P}\{\xi = k\} = \mathsf{M} s^\xi.$$

Функция f_ξ называется *производящей функцией* случайной величины ξ.

Свойства производящей функции:

1. $f_\xi(0) = \mathsf{P}\{\xi = 0\}$, $f_\xi(1) = 1$;
2. $\frac{d}{ds}f_\xi(s) = \sum k s^{k-1}\mathsf{P}\{\xi = k\}$, $\frac{d}{ds}f_\xi(1) = \mathsf{M}\xi$;
3. $\frac{d^2}{ds^2}f_\xi(1) = \mathsf{M}\xi(\xi - 1)$ — второй факториальный момент;
4. если $\xi_1, \ldots, \xi_n$ независимы, то $f_{\xi_1 + \cdots + \xi_n}(s) = \mathsf{M}s^{\xi_1 + \cdots + \xi_n} = \mathsf{M}s^{\xi_1} \ldots \mathsf{M}s^{\xi_n} = \prod_{k=1}^{n} f_{\xi_k}(s)$.

Лемма 4.1. *Пусть $\nu, \gamma_1, \gamma_2, \ldots$ — независимые целочисленные неотрицательные случайные величины; $\gamma_1, \gamma_2, \ldots$ одинаково распределены; $f(s) = \mathsf{M} s^{\nu}$, $g(s) = \mathsf{M} s^{\gamma_1}$. Тогда если $\mu = \begin{cases} 0, & \nu = 0, \\ \gamma_1 + \cdots + \gamma_k, & \nu = k. \end{cases}$, то $\mathsf{M} s^{\mu} = f(g(s))$.*

□ Применяя формулу полной вероятности и пользуясь независимостью, получаем:

$$\mathsf{M} s^{\mu} = \sum_{k=0}^{\infty} \mathsf{P}\{\nu = k\} \mathsf{M} s^{\gamma_1 + \cdots + \gamma_k} = \sum_{k=0}^{\infty} \mathsf{P}\{\nu = k\} \left(\mathsf{M} s^{\gamma_1}\right)^k = f(g(s)).$$

■

Если ветвящийся процесс начинается с одной частицы, то $\mu(1) = \gamma$, $f(s) = \mathsf{M}\left(s^{\mu(1)} \mid \mu(0) = 1\right)$. Положим $\varphi(t, s) = \mathsf{M} s^{\mu(t)}$.

Теорема 4.2. *Последовательность $\{\varphi(t, s)\}_{t=0}^{\infty}$ удовлетворяет рекуррентному соотношению*

$$\varphi(0, s) = \mathsf{M} s^{\mu(0)}, \quad \varphi(t+1, s) = \varphi(t, f(s)), t > 0$$

Если при этом $\mathsf{P}\{\mu(0) = 1\} = 1$, *то*

$$\varphi(0, s) = s, \quad \varphi(t+1, s) = f(\varphi(t, s)), t > 0$$

□

1. $\varphi(t+1, s) = \mathsf{M} s^{\gamma_{t_1} + \ldots + \gamma_{t_{\mu(t)}}} = \varphi(t, f(s))$ в силу предыдущей леммы.
2. Каждый из непосредственных потомков 1-й частицы порождает свой ветвящийся процесс с теми же свойствами, поэтому $\mu(t+1)$ распределена, как сумма $\mu(1)$ независимых экземпляров $\mu(t)$:

$$\varphi(t+1, s) = \mathsf{M} s^{\mu_1(t) + \ldots + \mu_{\mu(1)}(t)} = f(\varphi(t, s)),$$

т.к. $\gamma = \mu(1)$.

■

Следствие 4.1. *Если $\mu(0) = 1$, то $\varphi(t, s) = f_t(s) = f(f(\ldots f(s) \ldots))$ (t раз)*

Обозначим $\mathsf{M}\gamma = A = f'(1)$. $\mathsf{M}\mu(t) = \frac{d}{ds} \mathsf{M} s^{\mu(t)}\Big|_{s\,=\,1}$. Из рекуррентного соотношения имеем

$$\mathsf{M}\mu(t+1) = \varphi'_s(t+1, s)\Big|_{s\,=\,1} = \frac{d}{ds}\varphi(t, f(s))\Big|_{s\,=\,1} = \left[\varphi'_u(t, u)\Big|_{u\,=\,f(s)} f'(s)\right]\Big|_{s\,=\,1} = \mathsf{M}\mu(t) \cdot A,$$

Определение.

- Если $A < 1$, то процесс называется *докритическим*. В этом случае $\mathsf{M}\mu(t) \longrightarrow 0, \mathsf{P}\{\mu(t) > 0\} \longrightarrow 0 \quad (t \longrightarrow \infty)$.
- В случае $A = 1$ процесс называется *критическим*. $\mathsf{P}\{\mu(t) > 0\} \longrightarrow 0$, если $\mathsf{P}\{\gamma = 1\} \neq 1$ (это мы докажем потом).
- Если, наконец, $A > 1$, то $\mathsf{M}\mu(t) \longrightarrow \infty \quad (t \longrightarrow \infty)$ с экспоненциальной скоростью. Такие процессы называются *надкритическими*.

4.2. Вероятность вырождения ветвящегося процесса

Положим $q = \lim\limits_{t \longrightarrow \infty} \mathsf{P}\{\mu(t) = 0 \mid \mu(0) = 1\}$

Теорема 4.3. *q является наименьшим неотрицательным корнем уравнения $f(s) = s$.*

□ $\mathsf{P}\{\mu(t) = 0 \mid \mu(0) = 1\} = \sum_{k=0}^{\infty} \mathsf{P}\{\mu(t) = k\} s^k\Big|_{s\,=\,0} = \varphi(t, 0) = f_t(0)$.

Лемма 4.4. *Если $f(s) = \sum\limits_{k=0}^{\infty} p_k s^k$ — производящая функция распределения $\{p_k\}$, $f(s) \not\equiv s$, то*

1. *$A = f'(1) \leqslant 1 \Rightarrow \forall s \in [0, 1)\, f(s) > s$.*
2. *$A = f'(1) > 1 \Rightarrow \exists s_0 \in [0, 1)\, f(s_0) = s_0, f(s) > s$ при $s < s_0$, $f(s) < s$ при $s_0 < s < 1$.*

□ Заметим, что $f'(s) = \sum kp_k s^{k-1} \geqslant 0, f''(s) = \sum k(k-1)p_k s^{k-2} > 0$, поэтому f выпукла вниз и не убывает. Отсюда, так как $f(1) = 1$, видно, что в случае $A \leqslant 1$ график $f(s)$ лежит выше диагонали на $[0, 1)$, а в случае $A > 1$ обязательно на $[0, 1)$ её пересекает. (Если кто-нибудь нарисует картинку, будет очень хорошо) ■

Докажем по индукции, что $0 \leqslant f_t(0) \leqslant s_0$. При $t = 0$ это очевидно. Из леммы имеем

$$0 \leqslant f_t(0) \leqslant f(f_t(0)) = f_{t+1}(0) \leqslant f(s_0) = s_0,$$

т.е. $f_{t+1}(0) \geqslant f_t(0) \quad \forall t$, следовательно, существует $\lim_{t\longrightarrow\infty} f_t(0) = q \leqslant s_0$.

$$q = \lim_{t\longrightarrow\infty} f_t(0) = \lim_{t\longrightarrow\infty} f(f_t(0)) = f(\lim_{t\longrightarrow\infty} f_t(0)) = f(q),$$

т.к. f непрерывна. Следовательно, $q = s_0$. ■

Следствие 4.2. *Вероятность вырождения $q < 1 \Leftrightarrow A = f'(1) > 1$ (кроме вырожденного случая $\mathsf{P}\{\gamma = 1\} = 1$).*

Пример 2.1. $\mathsf{P}\{$ребёнок — девочка$\} = 0.486$, $\mathsf{P}\{$девочка доживёт до 18 лет$\} = 0.97$, A — среднее число детей. Если $A \cdot 0.486 \cdot 0.97 > 1$, то есть до 18 лет доживает в среднем хотя бы одна девочка, то $A \geqslant 2.124$. Из демографии известно, что A примерно равно 2.14.

Теорема 4.5. *Все положительные состояния ветвящегося процесса с $f(s) \not\equiv s$ несущественны:*

$$\forall k > 0 \quad \lim_{t\longrightarrow\infty} \mathsf{P}\{\mu(t) = k\} = 0.$$

□ От противного, пусть $\exists k_0 > 0 \limsup_{t\longrightarrow\infty} \mathsf{P}\{\mu(t) = k_0\} = v_{k_0} > 0$.

1. Пусть $\mathsf{P}\{\gamma = 0\} = p > 0$. Тогда вероятность вырождения за один шаг

   $$\mathsf{P}\{\mu(t+1) = 0 \mid \mu(t) = k\} = p^k > 0.$$

   $$\forall t\, \mathsf{P}\{\mu(t+1) = 0\} \geqslant \mathsf{P}\{\mu(t) = 0\} + \mathsf{P}\{\mu(t) = k_0\} p^{k_0} \text{ (оцениваем сумму снизу двумя слагаемыми)}$$

   Таким же образом оценивая снизу $\mathsf{P}\{\mu(t) = 0\}$ и т.д., получаем

   $$\mathsf{P}\{\mu(t+1) = 0\} \geqslant p^{k_0} \sum_{r=0}^{t} \mathsf{P}\{\mu(r) = k_0\} \longrightarrow \infty \quad (t \longrightarrow \infty),$$

   так как $\limsup_{t\longrightarrow\infty} \mathsf{P}\{\mu(t) = k_0\} > 0$. Противоречие.

2. Если $\mathsf{P}\{\gamma = 0\} = 0$, то $\mathsf{P}\{\gamma \geqslant 1\} = 1, \mathsf{P}\{\gamma = 1\} < 1$, т.к. $f(s) \not\equiv s$.

   $$\mathsf{P}\left\{\mu(t+1) = \gamma_{t_1} + \ldots + \gamma_{t_{\mu(t)}} \geqslant \mu(t)\right\} = 1,$$

   $$\mathsf{P}\{\mu(t+1) \geqslant k+1 \mid \mu(t) = k\} \geqslant 1 - (\mathsf{P}\{\gamma = 1\})^k \geqslant 1 - \mathsf{P}\{\gamma = 1\} = r > 0,$$

   $$\mathsf{P}\{\mu(t+1) \geqslant k_0+1\} \geqslant \mathsf{P}\{\mu(t) \geqslant k_0+1\} + \mathsf{P}\{\mu(t) = k_0, \mu(t+1) > k_0\} \geqslant \mathsf{P}\{\mu(t) \geqslant k_0+1\} + r\mathsf{P}\{\mu(t) = k_0\} \quad \forall t$$

   Аналогично пункту 1 получаем

   $$\mathsf{P}\{\mu(t+1) \geqslant k_0+1\} \geqslant r \sum_{v=0}^{t} \mathsf{P}\{\mu(t) = k_0\} \longrightarrow \infty \quad (t \longrightarrow \infty),$$

   что опять приводит к противоречию.

■

Теорема 4.6 (предельная теорема для надкритических ветвящихся процессов). *Если $A = \mathsf{M}\gamma > 1$ и $0 < \sigma^2 = \mathsf{D}\gamma < \infty$, то существует случайная величина $X\colon \mathsf{P}\{X > 0\} > 0$ и $\mathsf{P}\left\{\frac{\mu(t)}{A^t} \longrightarrow X, t \longrightarrow \infty\right\} = 1$. При этом $\mathsf{M}X = 1$ и характеристическая функция $\psi(u) = \mathsf{M}e^{iuX}$ удовлетворяет уравнению $\psi(Au) = f(\psi(u))$.*

Замечание. На самом деле эта теорема справедлива при гораздо более слабых условиях. Необходимым и достаточным условием существования такой величины X является $\mathsf{M}(\gamma \ln(1+\gamma)) < \infty$.

□ Определим последовательность дискретных случайных величин $X(t) = \frac{\mu(t)}{A^t}$. Покажем, что $\{X(t)\}$ фундаментальна в $L_2(\Omega)$, то есть

$$\forall \varepsilon > 0\, \exists s_0 : \mathsf{M}(X(s) - X(t))^2 < \varepsilon \quad \forall s, t > s_0.$$

$\mathsf{M}\left(\mu(t+k) \mid \mu(t)=m\right)=A^k m$, следовательно, $\mathsf{M}\left(X(t+k) \mid X(t)=x\right)=\mathsf{M}\left(\frac{\mu(t+k)}{A^{t+k}} \mid \mu(t)=xA^t\right)=\frac{xA^tA^k}{A^{k+t}}=x$, а, значит, $\mathsf{M}\left(X(s)-X(t) \mid X(k)\right)=0 \quad \forall s,t\geqslant k$.

Имеем[1]

$$\mathsf{M}\left((X(t)-X(t-1))(X(s)-X(s-1))\right)=\mathsf{M}\left((X(t)-X(t-1))\,\mathsf{M}\left(X(s)-X(s-1) \mid X(t),X(t-1)\right)\right)= \\ =\mathsf{M}\left(X(t)-X(t-1)\right)\mathsf{M}\left(X(s)-X(s-1) \mid X(t)\right)=0.$$

$$\mathsf{M}\left(X(t)-X(t-1)\right)^2=\mathsf{M}\left(\frac{\mu(t)}{A^t}-\frac{\mu(t-1)}{A^{t-1}}\right)^2=\frac{1}{A^{2t}}\mathsf{M}\mathsf{M}\left\{\left((\gamma_{t-1,1}-A)+\ldots+(\gamma_{t-1,\mu(t-1)}-A)\right)^2 \mid \mu(t-1)\right\}= \\ =\frac{1}{A^{2t}}\mathsf{M}\left(\sigma^2\mu(t-1)\right)=\sigma^2\frac{A^{t-1}}{A^{2t}}=\frac{\sigma^2}{A^{t+1}}.$$

Тогда при $t<s$ справедливо

$$\mathsf{M}\left(X(s)-X(t)\right)^2=\mathsf{M}\left(\sum_{v=t+1}^{s}(X(v)-X(v-1))\right)^2=\sum_{v=t+1}^{s}\mathsf{M}\left(X(v)-X(v-1)\right)^2=\sum_{v=t+1}^{s}\frac{\sigma^2}{A^{v+1}}\leqslant \\ \leqslant\sigma^2\frac{1}{A^{t+2}(1-\frac{1}{A})}\longrightarrow 0 \quad (t\longrightarrow\infty).$$

Фундаментальность доказана, стало быть, $X(t)$ сходится к X в $L_2(\Omega)$. Вообще говоря, это не означает сходимости почти наверное, но покажем, что в нашем случае она всё же имеет место. Если мы найдём такую функцию $f(t)\longrightarrow 0,\ t\longrightarrow\infty$, что $\sum_{t=1}^{\infty}\mathsf{P}\left\{|X(t)-X|>f(t)\right\}<\infty$, то по лемме Бореля–Кантелли с вероятностью 1 будет выполняться бесконечно много событий вида $|X(t)-X|\leqslant f(t), t>t_0$, т.е. $X(t)\xrightarrow{\text{п.н.}}X$. В качестве такой $f(t)$ можно взять, например, $\frac{1}{A^{t/3}}$:

$$\mathsf{P}\left\{|X(t)-X|>\frac{1}{A^{t/3}}\right\}\leqslant\frac{c}{A^tf^2(t)}=\frac{c}{A^{t/3}}$$

для некоторой константы c в силу неравенства Чебышёва для случайной величины $X(t)-X$, а оценку для её второго момента можно получить предельным переходом из выражения для $\mathsf{M}\left(X(s)-X(t)\right)^2$.

Докажем теперь свойство характеристической функции X. Пусть $\psi_t(u)=\mathsf{M}e^{iuX(t)}$. $X(t)\xrightarrow{\text{п.н.}}X\Rightarrow\psi_t(u)\longrightarrow\psi(u)\,\forall u\in\mathbb{R}$.

$$\psi_{t+1}(u)=\mathsf{M}e^{iuX(t+1)}=\mathsf{M}\exp\left\{iu\frac{\mu(t+1)}{A^{t+1}}\right\}=\varphi\left(t+1,\exp\left\{iu\frac{1}{A^{t+1}}\right\}\right)=f\left(\varphi\left(t,\exp\left(i\frac{u}{A^{t+1}}\right)\right)\right)= \\ =f\left(\mathsf{M}\exp\left(iu\frac{\mu(t)}{A^{t+1}}\right)\right)=f\left(\mathsf{M}e^{iuX(t)/A}\right)=f\left(\psi_t\left(\frac{u}{A}\right)\right).$$

Осталось заметить, что левая часть этого равенства сходится к $\psi(u)$, а правая — к $f\left(\psi(\frac{u}{A})\right)$. ■

5. Конечномерные распределения случайных процессов

Определение. Пусть дано вероятностное пространство $(\Omega,\mathscr{B},\mathsf{P})$ и случайный процесс $\xi\colon\Omega\longrightarrow G$ (G — некоторое функциональное пространство), т.е. совокупность случайных величин $\xi(t)$, и конечный набор точек $t_1,\ldots,t_k$. Распределение вектора $(\xi(t_1),\ldots,\xi(t_k))$ называется *конечномерным распределением* процесса ξ.

5.1. Семейства согласованных распределений

Пример 1.1. Заметим, что случайный процесс с заданными конечномерными распределениями может и не существовать. Положим, например, $\mathsf{P}\{\xi(1)=m,\xi(2)=n\}=\frac{1}{4}$ $(n,m=0,1)$ и $\mathsf{P}\{\xi(2)\leqslant x,\xi(3)\leqslant y\}=\Phi(x)\Phi(y)$, где Φ — функция стандартного нормального распределения.

Чтобы такого не было, вводится требование *согласованности.* Интуитивно это означает, что результат вычисления вероятности любого события не зависит от способа его вычисления по конечномерным распределениям. Для формального определения ограничимся случаем действительнозначного случайного процесса.

Определение. Семейство распределений (вероятностных мер) $\mathscr{P}=\{\mathsf{P}_{t_1,\ldots,t_k},t_1<\cdots<t_k,k=1,2,\ldots\}$ ($\mathsf{P}_{t_1,\ldots,t_k}$ — вероятностная мера на $\mathbb{R}^k$) называется *согласованным*, если для любых упорядоченных наборов

[1] Предпоследнее равенство в следующей цепочке верно в силу того, что $\mu(t)$, а вместе с ними и $X(t)$, образуют цепи Маркова, поэтому условные распределения относительно $X(t),X(t-1)$ и просто $X(t)$ совпадают. — *примеч. А.Х.*

$t_1 < \dots < t_n$, $s_1 < \dots < s_m$, $v_1 < \dots < v_k$, для которых $\{v_1, \dots, v_k\} = \{t_1, \dots, t_n\} \cap \{s_1, \dots, s_m\}$ и для любых борелевских множеств $C_1, \dots, C_k \subseteq \mathbb{R}$ имеем

$$\mathsf{P}_{v_1,\dots,v_k}(C_1 \times \dots \times C_k) = \mathsf{P}_{t_1,\dots,t_n}(A_1 \times \dots \times A_n) = \mathsf{P}_{s_1,\dots,s_m}(B_1 \times \dots \times B_m),$$

где

$$A_j = \begin{cases} C_i, & t_j = v_i, \\ \mathbb{R} & \text{в ином случае} \end{cases} \qquad \text{и} \qquad B_j = \begin{cases} C_i, & s_j = v_i, \\ \mathbb{R} & \text{в ином случае.} \end{cases}$$

Доказательство следующей теоремы можно прочитать в книжке Боровкова. В нашем курсе его не будет.

Теорема 5.1 (теорема Колмогорова о согласованных распределениях). *Если на $\mathbb{R}^{\mathbb{R}}$ задано семейство согласованных распределений $\mathscr{P}$, то на пространстве $(\mathbb{R}^{\mathbb{R}}, \mathscr{F})$ (здесь $\mathscr{F}$ — борелевская σ-алгебра на $\mathbb{R}^{\mathbb{R}}$, т.е. σ-алгебра, порождённая цилиндрическими борелевскими множествами) существует единственная мера $\mathbf{P}$ такая, что для любого конечного множества $\{t_1, \dots, t_n\} = T \subset \mathbb{R}$ проекция $\mathbf{P}$ на $\mathbb{R}^T$ совпадает с $\mathsf{P}_{t_1,\dots,t_k}$.*

6. Процессы с независимыми приращениями. Пуассоновский процесс

6.1. Определения

Определение. *Процесс с независимыми приращениями* — это такой случайный процесс $\xi(t)$, что для любого упорядоченного набора $t_1 < \dots < t_k$ вектор приращений $(\xi(t_2) - \xi(t_1), \xi(t_3) - \xi(t_2), \dots, \xi(t_n) - \xi(t_{n-1}))$ имеет независимые координаты.

В частности, любой процесс с независимыми приращениями является марковским. Чтобы задать такой процесс, достаточно сопоставить каждой паре $s < t$ распределение случайной величины $\xi(t) - \xi(s)$. *Условия согласованности* в данном случае имеют вид: для любых $t_1 < \dots < t_n$ распределение $\xi(t_n) - \xi(t_1)$ совпадает с распределением суммы $\delta_{t_1 t_2} + \delta_{t_2 t_3} + \dots + \delta_{t_{n-1} t_n}$, где $\delta_{t_k t_{k+1}}$ — независимые в совокупности случайные величины, распределённые так же, как и $\xi(t_{k+1}) - \xi(t_k)$.

Определение. $\nu(t)$ — *однородный пуассоновский процесс* с интенсивностью λ, если ν — процесс с независимыми приращениями и для всех $t > s$

$$\mathsf{P}\{\nu(t) - \nu(s) = k\} = \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)}.$$

Процесс с независимыми приращениями $\nu(t)$ — *пуассоновский процесс* с *ведущей функцией* Λ (Λ — монотонно неубывающая функция), если

$$\mathsf{P}\{\nu(t) - \nu(s) = k\} = \frac{(\Lambda(t) - \Lambda(s))^k}{k!} e^{-(\Lambda(s) - \Lambda(t))}.$$

Однородный пуассоновский процесс является частным случаем пуассоновского процесса вообще: положим $\Lambda(t) = \lambda t$. Обратно, если $\nu(t)$ — однородный пуассоновский процесс с интенсивностью $\lambda = 1$, то процесс $\nu_\Lambda(t) = \nu(\Lambda(t))$ — пуассоновский процесс с ведущей функцией Λ, т.е. произвольный пуассоновский процесс отличается от однородного заменой времени.

6.2. Свойства пуассоновского процесса

Утверждение 6.1. *Если $\nu(t)$ — однородный пуассоновский процесс с интенсивностью λ, то при $h \longrightarrow 0+$ имеем:*

$$\mathsf{P}\{\nu(t+h) - \nu(t) = 0\} = 1 - \lambda h + o(h),$$
$$\mathsf{P}\{\nu(t+h) - \nu(t) = 1\} = \lambda h + o(h),$$
$$\mathsf{P}\{\nu(t+h) - \nu(t) > 1\} = o(h).$$

Обратно, если целочисленный однородный по времени случайный процесс с независимыми приращениями $\nu(t)$ удовлетворяет этим соотношениям, то $\nu(t)$ — пуассоновский с интенсивностью λ.

□ В прямую сторону утверждение почти очевидно. Действительно, по определению $\mathsf{P}\{\nu(t+h) - \nu(t) = k\} = \dfrac{(\lambda h)^k}{k!} e^{-\lambda h}$. При $k = 0$ это превращается в $e^{-\lambda h} = 1 - \lambda h + o(h)$ по формуле Тейлора. По той же формуле при $k = 1$ получаем: $\lambda h e^{-\lambda h} = \lambda h + o(h)$. А для $k > 1$ остаётся $1 - (1 - \lambda h + o(h)) - (\lambda h + o(h)) = o(h)$.

Докажем обратное. Пусть $[t, t+h)$ — произвольный интервал. Разобьём его на n кусочков $[t+k\frac{h}{n}, t+(k+1)\frac{h}{n})$.

$$\nu(t+h) - \nu(t) = \sum_{k=0}^{n-1} \nu\left(t+(k+1)\frac{h}{n}\right) - \nu\left(t+k\frac{h}{n}\right).$$

В этой сумме все слагаемые независимы. Рассмотрим производящую функцию отдельного слагаемого (при $0 < s < 1$ геометрическая прогрессия сходится, и все члены, начиная с s^2, уходят в $o\left(\frac{h}{n}\right)$:

$$\mathsf{M}s^{\nu\left(t+(k+1)\frac{h}{n}\right)-\nu\left(t+k\frac{h}{n}\right)} = 1 - \lambda\frac{h}{n} + o\left(\frac{h}{n}\right) + s\left(\lambda\frac{h}{n} + o\left(\frac{h}{n}\right)\right) + o\left(\frac{h}{n}\right) = e^{-\lambda\frac{h}{n}(1-s)+o\left(\frac{h}{n}\right)}.$$

(Для последнего перехода нужна равномерность о-маленьких по k, которая достигается однородностью процесса по времени.) В силу независимости слагаемых

$$\mathsf{M}s^{\nu(t+h)-\nu(t)} = \prod_{k=0}^{n-1} e^{-\lambda\frac{h}{n}(1-s)+o\left(\frac{h}{n}\right)} = e^{-\lambda h(1-s)+o(1)} \xrightarrow[n\to\infty]{} e^{-\lambda h(1-s)},$$

а это производящая функция распределения Пуассона с параметром λh. ■

Траектория процесса Пуассона монотонно неубывает и локально постоянна (имеет скачки на дискретном множестве). Далее мы покажем, что она скачет не больше, чем на 1.

Определение. Последовательность точек скачков пуассоновского процесса называется *пуассоновским потоком.*

Утверждение 6.2. *С вероятностью 1 все скачки пуассоновского процесса с интенсивностью $\lambda < \infty$ равны 1.*

□ Для пуассоновского процесса $\nu(t)$ с интенсивностью λ положим $\nu_2(t,h) \rightleftharpoons$ количество скачков $\nu(t)$ величины больше 1 на $[t, t+h)$.

$$\sigma_n(t,h) \rightleftharpoons \sum_{k=1}^{n} I_{\left\{\nu(t+\frac{kh}{n})-\nu(t+(k-1)\frac{h}{n})>1\right\}}$$

Как видно, $\forall n \geqslant 1$ $\{\nu_2(t,h) \geqslant 1\} \subseteq \{\sigma_n(t,h) \geqslant 1\}$.

$$\mathsf{M}\sigma_n(t,h) = \sum_{k=1}^{n} \mathsf{P}\left\{\nu(t+\frac{kh}{n}) - \nu(t+(k-1)\frac{h}{n}) > 1\right\} = \sum_{k=1}^{n} \mathsf{P}\{\Delta_k > 1\},$$

где $\{\Delta_k\}$ — независимые в совокупности пуассоновские случайные величины с параметром $\lambda\frac{h}{n}$.

$$\mathsf{P}\{\Delta_k > 1\} = 1 - e^{-\lambda\frac{h}{n}} - \lambda\frac{h}{n}e^{-\lambda\frac{h}{n}} = \lambda\frac{h}{n} + O\left(\frac{1}{n^2}\right) - \lambda\frac{h}{n}e^{-\lambda\frac{h}{n}} = O\left(\frac{1}{n^2}\right) + \lambda\frac{h}{n}\left(1 - e^{-\lambda\frac{h}{n}}\right) = O\left(\frac{1}{n^2}\right).$$

Поэтому

$$\mathsf{M}\sigma_n(t,h) = O\left(\frac{1}{n}\right) \longrightarrow 0, \quad (n \longrightarrow \infty).$$

$$\mathsf{P}\{\nu_2(t,h) \geqslant 1\} \leqslant \lim_{n\longrightarrow\infty} \mathsf{P}\{\sigma_n(t,h) > 1\} \leqslant \lim_{n\longrightarrow\infty} \mathsf{M}\sigma_n(t,h) = 0.$$

■

Пуассоновский поток можно получить как предельное распределение количества успехов в некоторой схеме Бернулли. Например, пусть $\tau_1(N), \ldots, \tau_N(N)$ — независимые случайные величины, $\lambda < N$ и $\mathsf{P}\{\tau_k(N) < x\} = \frac{\lambda x}{N}$, $0 < x < 1$, то есть условное распределение $\tau_k(N)$ на $[0,1]$ равномерно. Обозначим $\eta_N \rightleftharpoons \sum_1^N I_{\tau_i(N)\leqslant 1}$.

Теорема 6.3 (Пуассон). $\eta_N \xrightarrow{\mathscr{D}} Poiss(\lambda)$*, то есть количество точек $\tau_i(N)$, попавших в $[0,1]$, стремится при увеличении N к количеству на этом отрезке точек пуассоновского потока интенсивности λ.*

□ η_N — это количество успехов в схеме Бернулли с N испытаниями и вероятностью успеха $\frac{\lambda}{N}$.

$$\mathsf{P}\{\eta_N = k\} = C_N^k\left(\frac{\lambda}{N}\right)^k\left(1-\frac{\lambda}{N}\right)^{N-k} = C_N^k\frac{\lambda^k}{(N-\lambda)^k}\left(1-\frac{\lambda}{N}\right)^N =$$
$$= \frac{N(N-1)\ldots(N-k+1)}{(N-\lambda)^k}\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{N}\right)^N \longrightarrow \frac{\lambda^k}{k!}e^{-\lambda}, \quad N \longrightarrow \infty.$$

■

Пусть $\nu(t)$ — пуассоновский процесс с интенсивностью λ. Положим $\tau_+(t) = \inf\{u > 0 \mid \nu(t+u) > \nu(t)\}$ — расстояние от t до следующего скачка и $\tau_-(t) = \inf\{u > 0 \mid \nu(t-u) < \nu(t)\}$ — до предыдущего.

Утверждение 6.4. *Для любого $t \in \mathbb{R}$ $\tau_+(t)$ не зависит от $\{\nu(x)\}_{x \leqslant t}$, $\tau_-(t)$ не зависит от $\{\nu(x)\}_{x \geqslant t}$, и* $\mathsf{P}\{\tau_+(t) > u\} = \mathsf{P}\{\tau_-(t) > u\} = e^{-\lambda u}$ *при $u > 0$ (т.е. имеют показательное распределение с параметром λ).*

□ $\mathsf{P}\{\tau_+(t) > u\} = \mathsf{P}\{\nu(t+u) = \nu(t)\} = \mathsf{P}\{\nu(t+u) - \nu(t) = 0\} = e^{-\lambda u}$, т.к. наш процесс пуассоновский. Для τ_- аналогично.

Независимость следует из независимости приращений. ■

При условии наличия скачка в малой окрестности: $\mathsf{P}\{\tau_+(t) > u \mid \nu(t-\varepsilon) < \nu(t)\} = e^{-\lambda u}$ для всех $\varepsilon > 0$. Отсюда $\mathsf{P}\{\zeta_{k+1} - \zeta_k > u\} = e^{-\lambda u}$ (ζ_k — точки скачков). $\tau = \zeta_{k+1} - \zeta_k$ независимы и имеют показательное распределение[2] с параметром λ.

Следствие 6.1 («Парадокс времени ожидания»). *Если $\nu(t)$ — однородный пуассоновский процесс с интенсивностью λ и $\{\zeta_k\}$ — моменты его скачков, то* $\mathsf{M}\{\zeta_{k+1} - \zeta_k\} = \mathsf{M}\tau_+(t) = \mathsf{M}\tau_-(t) = \frac{1}{\lambda}$, *но*
$\mathsf{M}\{\zeta_{k+1} - \zeta_k \mid \zeta_k \leqslant t < \zeta_{k+1}\} = \frac{2}{\lambda}$.

□ Первое равенство верно в силу определения $\tau_\pm(t)$ и ζ_i. При условии $\zeta_k \leqslant t < \zeta_{k+1}$ имеем $\zeta_{k+1} - \zeta_k = \tau_+(t) + \tau_-(t)$, откуда получаем второе равенство. ■

Пример 2.1. *Системы массового обслуживания.* Обозначим через $M \mid G \mid 1$ следующую систему. На входе системы имеется пуассоновский поток «заявок» интенсивности λ, который обрабатывается последовательно (т.е. в один поток; отсюда 1 в $M \mid G \mid 1$), причём длительности рассмотрения заявок независимы и одинаково распределены с функцией распределения G. При обработке заявки все заявки, поступившие после неё, становятся в очередь. *Периодом занятости* системы массового обслуживания назовём случайную величину, равную времени от поступления первой заявки до первого опустошения очереди, *периодом простоя* — время от первого опустошения очереди до прибытия новой заявки. (Докажите, кстати, что это действительно случайные величины).

Теорема 6.5. *Период занятости в системе $M \mid G \mid 1$ конечен с вероятностью* $1 \Leftrightarrow \lambda\mathsf{M}\gamma \leqslant 1 \quad (\gamma \sim G)$.

□ С обслуживанием пуассоновского потока заявок можно ассоциировать ветвящийся процесс μ следующим образом:

- $\mu(0) \rightleftharpoons 1$.
- Считаем потомками точки поколения k те заявки, которые пришли за время её обработки. Таким образом $\mu(1) = |\{$заявки, пришедшие за время обработки первой заявки$\}| = |M(1)|$, $\mu(2) =$ $= |\{$заявки, пришедшие за время обработки $M(1)\}|$, и так далее. Количества потомков у разных точек независимы (это свойство пуассоновского потока).

Конечность периода занятости равносильна тому, что процесс μ выродится.

$$\mathsf{P}\{\text{за время обслуживания заявки придёт k новых}\} = \int\limits_0^\infty \frac{(\lambda x)^k}{k!} e^{-\lambda x} dG(x) \eqqcolon r_k$$

$$\sum_{k=1}^\infty k r_k = \sum_{k=1}^\infty k \int\limits_0^\infty \frac{(\lambda x)^k}{k!} e^{-\lambda x} dG(x) = \int\limits_0^\infty \sum_{k=1}^\infty k \frac{(\lambda x)^k}{k!} e^{-\lambda x} dG(x) = \int\limits_0^\infty \lambda x \sum_{k=1}^\infty \frac{(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x} dG(x) = \int\limits_0^\infty \lambda x dG(x) = \lambda\mathsf{M}\gamma.$$

Здесь мы воспользовались тем, что условное распределение числа точек пуассоновского потока в случайном интервале независимой с точками потока длины γ является пуассоновским с параметром $\lambda\gamma$ (докажите это!). По следствию 4.2 это означает, что μ вырождается почти наверное тогда и только тогда, когда $\lambda\mathsf{M}\gamma \leqslant 1$. ■

Теорема 6.6. *Для системы $M \mid G \mid 1$ математическое ожидание числа заявок, обслуженных за период занятости, равно $\frac{1}{1-\lambda\mathsf{M}\gamma}$, если* $\lambda\mathsf{M}\gamma < 1$

□ $\mu(0) = 1, \mu(t)$ — число точек t-го поколения. Общее число обслуженных заявок равно $\sum_0^\infty \mu(t)$.

$$\mathsf{M}\{\mu(t) \mid \mu(0) = 1\} = A^t, \text{где } A = \mathsf{M}\{\mu(1) \mid \mu(0) = 1\} = \lambda\mathsf{M}\gamma,$$

$$\mathsf{M}\sum_0^\infty \mu(t) = \sum_0^\infty A^t = \frac{1}{1-A}.$$

■

Пример 2.2. *Пуассоновское случайное поле в $\mathbb{R}^d$ (интенсивности λ)* — это случайная совокупность точек, удовлетворяющая следующим условиям:

[2] Заметим, что отсюда очевидно следует утверждение 6.2: если считать скачок длины 2 за два единичных, получим $\tau = 0$, что выполняется (в силу вышесказанного) с вероятностью 0. — *примеч. С.К.*

Пуассоновский поток можно, таким образом, задать ещё и как последовательность случайных точек, приращения которой независимы и распределены экспоненциально. — *примеч. А.Х.*

1. Число точек в открытом множестве A имеет распределение Пуассона с параметром $\lambda \operatorname{mes} A$.
2. Числа точек в непересекающихся множествах независимы.

Пример 2.3. Зададимся вопросом оценить λ, равное количеству микробов на литр воды. Пусть имеется N пробирок объёма z. Допустим, что число микробов в пробирке имеет распределение Пуассона с параметром λz.

$$\mathsf{P}\left\{\text{в пробирке нет микробов}\right\} = e^{-\lambda z} = p. \quad \text{(этот факт можно проверить)}$$

Пусть ν — число пробирок без микробов. Найдём $\widehat{\lambda}$ — оценку максимального правдоподобия для λ.

$$\mathsf{P}\left\{\nu = k\right\} = C_N^k p^k (1-p)^{N-k}.$$

$$k(\widehat{\lambda}) \rightleftharpoons \frac{\partial}{\partial \widehat{\lambda}} \left(e^{-\widehat{\lambda} z k}(1-e^{-\widehat{\lambda} z})^{N-k}\right) = e^{-\widehat{\lambda} k z}(1-e^{-\widehat{\lambda} z})^{N-k-1}\left(-kz(1-e^{-\widehat{\lambda} z}) + (N-k)ze^{-\widehat{\lambda} z}\right) = 0 \Leftrightarrow$$

$$\Leftrightarrow -kz + Nze^{-\widehat{\lambda} z} = 0 \Leftrightarrow \widehat{\lambda} = \frac{1}{z}\ln\frac{N}{\nu}.$$

7. Цепи Маркова с непрерывным временем

Определение. $\left\{\xi(t)\right\}_{t \geqslant 0}$ называется (однородной по времени) цепью Маркова с непрерывным временем и не более чем счётным множеством состояний S, если

$$\forall n \geqslant 1\ \forall 0 \leqslant t_0 < \ldots < t_n\ \forall i_0, \ldots, i_n \in S\ \mathsf{P}\left\{\xi(t_n) = i_n \mid \xi(t_{n-1}) = i_{n-1}, \ldots, \xi(t_0) = i_0\right\} = \mathsf{P}\left\{\xi(t_n) = i_n \mid \xi(t_{n-1})\right\}.$$

Пример 0.1.

1. Пуассоновский процесс (да и любой процесс с независимыми приращениями, на самом деле).
2. $\eta(t) = \nu(t) \mod 2$, где $\nu(t)$ — пуассоновский.
3. $\left\{\zeta(t)\right\}_{t \in \mathbb{R}}$ — независимые одинаково распределённые случайные величины, $\mathsf{P}\left\{\zeta(t) = 1\right\} = p$, $\mathsf{P}\left\{\zeta(t) = 0\right\} = q = 1 - p$.

Определение. Случайный процесс $\xi(t)$ стохастически непрерывен в точке t_0, если $\xi(t_0 + s) \xrightarrow{\mathsf{P}} \xi(t_0)$, $s \longrightarrow 0$.

Сопоставим каждому $i \in S \quad \lambda_i \in (0, \infty)$. Покажем, что цепь Маркова с дискретным временем можно рассматривать как цепь с непрерывным временем, которая «сидит» в состоянии i случайное время, распределённое показательно с параметром λ_i.

Замечание.[О свойствах показательного распределения]

1. $\tau \sim \exp(1) \Rightarrow \frac{\tau}{\lambda} \sim \exp(\lambda)$.
2. $\tau \sim \exp(\lambda) \Rightarrow \mathsf{P}\left\{\tau > x + y \mid \tau > x\right\} = e^{-\lambda y}$, $y \geqslant 0$ (свойство отсутствия памяти у показательного распределения).

Итак, пусть $\left\{\xi(n)\right\}_{n \in \mathbb{N}}$ — цепь Маркова с дискретным временем, множеством состояний S, матрицей переходных вероятностей P и $\Lambda = \left\{\lambda_i\right\}_{i \in S}$ — множество *интенсивностей выходов.*

7.1. Вложенная цепь Маркова

Пусть $\tau_0, \tau_1, \ldots$ — независимые случайные величины, распределённые показательно с параметром 1. Тогда при *фиксированных* $\left\{\xi_k\right\}$ $\frac{\tau_0}{\lambda_{\xi_0}}, \frac{\tau_1}{\lambda_{\xi_1}}, \ldots$ — последовательность независимых случайных величин, распределённых показательно с параметрами λ_{ξ_k}. Положим

$$T_0 = 0, \quad T_n = \sum_{k=0}^{n-1} \frac{\tau_k}{\lambda_{\xi_k}}, \quad n \in \mathbb{N} \setminus \{0\}$$

Тогда $\xi(t) \rightleftharpoons \xi_n, T_n \leqslant t < T_{n+1}$, будет цепью Маркова с непрерывным временем (проверьте это). Дискретная цепь Маркова ξ_i называется *вложенной цепью Маркова.*

Марковость получившегося процесса с непрерывным временем не очевидна. Докажем это:

Утверждение 7.1. *Если $\lambda^* = \sup\limits_{i \in S} \lambda_i < \infty$, то формула $\xi(t) = \xi_n$ при $T_n \leqslant t < T_{n+1}$ определяет случайный процесс на $[0, +\infty)$, который является цепью Маркова с непрерывным временем.*

□ Докажем сначала, что процесс определён при всех $t \geqslant 0$ (т.е. что $T_n \longrightarrow \infty$ $(n \longrightarrow \infty)$ почти наверное). Имеем:

$$T_n = \sum_{k=0}^{n-1} \frac{\tau_k}{\lambda_{\xi_k}} \geqslant \sum_{k=0}^{n-1} \frac{\tau_k}{\lambda^*} = \frac{1}{\lambda^*} \sum_{k=0}^{n-1} \tau_k.$$

τ_k — последовательность независимых одинаково распределённых случайных величин, $\mathsf{M}\tau_k = 1$. В силу УЗБЧ $\mathsf{P}\{\frac{1}{n} \sum\limits_{k=0}^{n-1} \tau_k \longrightarrow \mathsf{M}\tau_1\} = 1$, откуда почти наверное $\sum\limits_{k=0}^{n-1} \tau_k \longrightarrow \infty$.

Теперь проверим марковость.

Лемма 7.2. *Если события $B_1, B_2, \ldots$ попарно не пересекаются, и A — такое событие, что $\mathsf{P}(A \mid B_k) = p$ (не зависит от k), то $\mathsf{P}(A \mid \bigcup B_k) = p$.*

□ По определению условной вероятности $\mathsf{P}(A \cap B_k) = p \cdot \mathsf{P}(B_k)$. Применяем формулу Байеса и вспоминая, что B_k (а значит, и $A \cap B_k$) попарно не пересекаются, получаем:

$$\mathsf{P}(A \mid \bigcup B_k) = \frac{\mathsf{P}(A \cap (\bigcup B_k))}{\mathsf{P}(\bigcup B_k)} = \frac{\sum \mathsf{P}(A \cap B_k)}{\sum \mathsf{P}(B_k)} = \frac{\sum p \cdot \mathsf{P}(B_k)}{\sum \mathsf{P}(B_k)} = p,$$

что и требовалось. ■

Пример 1.1. Покажем, что если B_k в условии леммы пересекаются, то условная вероятность может отклоняться в обе стороны. Пусть B_1 и B_2 пересекаются и $\mathsf{P}(B_1) = \mathsf{P}(B_2)$.

1. $A = B_1 \cap B_2$, $\mathsf{P}(A) > 0$. $\mathsf{P}(A \mid B_1) = \mathsf{P}(A \mid B_2) = p$, но

$$\mathsf{P}(A \mid B_1 \cup B_2) = \frac{\mathsf{P}(A)}{\mathsf{P}(B_1 \cup B_2)} = \frac{\mathsf{P}(A)}{\mathsf{P}(B_1) + \underbrace{\mathsf{P}(B_2) - \mathsf{P}(A)}_{>0}} > \frac{\mathsf{P}(A)}{\mathsf{P}(B_1)} = \mathsf{P}(A \mid B_1) = p.$$

2. С другой стороны, если взять $A' = B_1 \triangle B_2$ (симметрическая разность), получим: $\mathsf{P}(A' \mid B_1 \cup B_2) = \mathsf{P}(B_1 \triangle B_2 \mid B_1 \cup B_2) = 1 - \mathsf{P}(B_1 \cup B_2 \mid B_1 \cup B_2) = 1 - \mathsf{P}(A \mid B_1 \cup B_2) < 1 - \mathsf{P}(A \mid B_1) = \mathsf{P}(A' \mid B_1)$.

Для доказательства марковости $\xi(t)$ нужно показать, что для любого набора $t_0 < t_1 < \cdots < t_n$ имеем $\mathsf{P}(\xi(t_n) = i_n \mid \xi(t_{n-1}) = i_{n-1}, \ldots, \xi(t_0) = i_0) = \mathsf{P}(\xi(t_n) = i_n \mid \xi(t_{n-1} = i_{n-1})$. Введём вспомогательные события $A = \{\xi(t_n) = i_n\}$, $B_k = \{\xi(t_j) = i_j, j = 0, \ldots, n-1; T_k \leqslant t_{n-1} < T_{k+1}\}$ (из второго условия следует, что $i_{n-1} = \xi_k$), $B'_k = \{\xi(t_{n-1} = i_{n-1}; T_k \leqslant t_{n-1} < T_{k+1}\}$, События A и B_k, а также A, B'_k удовлетворяют условиям леммы, поэтому достаточно показать, что $\mathsf{P}(A \mid B_k) = \mathsf{P}(A \mid B'_k)$ и что $\mathsf{P}(A \mid B'_k) = \mathsf{P}(A \mid B'_l) \quad \forall k, l$.

Заметим, что для любого $D_k\colon D_k \supset B_k, B'_k$ верно $\mathsf{P}(A \mid B_k) = \mathsf{P}(AD_k \min B_k), \mathsf{P}(A \mid B'_k) = \mathsf{P}(AD_k \mid B'_k)$. Возьмём таким D_k $\{T_k \leqslant t_{n-1} < T_{k+1}\}$. $AD_k = \sqcup_{l \geqslant k} AD_k \cap \{T_l \leqslant t_n < t_{l+1}\} = \sqcup_{l \geqslant k} A_{k,l}$. Достаточно показать, что $\forall l \geqslant k \mathsf{P}(A_{k,l} \mid B_k) = \mathsf{P}(A_{k,l} \mid B'_k)$. К сожалению, даже такого дробления событий нам не хватит, поэтому рассмотрим семейства событий

$$B_k = \sqcup_{\alpha \in I} B_\alpha =$$
$$= \sqcup_{\alpha \in I} \{\xi(t_j) = i_j, 0 \leqslant j \leqslant n-1; \xi_k = i_{n-1}, T_k \leqslant t_{n-1} < T_{k+1}; \xi_m = \alpha_m, 0 \leqslant m \leqslant k-1, T_{\alpha(j)} \leqslant t_j < T_{\alpha(j+1)}, 0 \leqslant j \leqslant n-2\},$$
$$B'_k = \sqcup_{\alpha \in I'} B'_\alpha =$$
$$= \sqcup_{\alpha \in I'} \{\xi(t_{n-1}) = i_{n-1}, T_k \leqslant t_{n-1} < T_{k+1}, \xi_k = i_{n-1}; \xi_m = \alpha_m, 0 \leqslant m \leqslant k-1\},$$

где индексами α пробегаются все возможные сочетания значений вложенной цепи и варианты попадания t_j в интервалы между T_i. Заметим, что если вероятность B_α равна нулю, но условную вероятность $P(A_{k,l} \mid B_\alpha)$ можно положить равной чему угодно, в частности, тому, что нам нужно (это не влияет на условную вероятность по условию $\sqcup_\alpha B_\alpha$). Если же вероятность B_α не равна нулю, то из B_α можно выкинуть все ограничения типа $\xi(t_k) = i_k$, так как все $\xi(t_k)$ определены тем, в какой интервал $[T_j, T_{j+1})$ попадает t_k и чему равно ξ_j. Таким образом, без ограничения общности можно считать, что $\{\xi(t_j) = i_j, 0 \leqslant j \leqslant n-1; \xi_k = i_{n-1}, T_k \leqslant t_{n-1} < T_{k+1}; \xi_m = \alpha_m, 0 \leqslant m \leqslant k-1, T_{\alpha(j)} \leqslant t_j < T_{\alpha(j+1)}, 0 \leqslant j \leqslant n-2\} = \{\xi_k = i_{n-1}, \xi_m = \alpha_m, 0 \leqslant m \leqslant k-1, T_k \leqslant t_{n-1} < T_{k+1}, T_{\alpha(j)} \leqslant t_j < T_{\alpha(j+1)}, 0 \leqslant j \leqslant n-2\}$. Заметим также, что на каждом из событий B_α вектор $T = (T_1, \ldots, T_k)$ не зависит от $(\xi_1, \ldots, \xi_k)$, т.к. на B_α $T_i = \sum_0^{i-1} \tau_j / \alpha_j$, а величины $\{\tau_i\} u \{\xi_i\}$ независимы в совокупности. То же относится и к B'_α.

Теперь попытаемся посчитать $\mathsf{P}(A_{k,l} \mid B_\alpha)$. Для этого возьмём наше многострадальное событие $A_{k,l}$ и разобьём его ещё мельче.

$$A_{k,l} = \cup_{(b_{k+1}, \ldots, b_l) \in S^n} A_{k,l} \cap \{\xi_{k+1} = b_{k+1}, \ldots, \xi_l = b_l\} = \cup_\beta A_{k,l,\beta}.$$

Учитывая сказанное выше о независимости ξ и τ, $\mathsf{P}(A_{k,l,\beta})$ распадётся в произведение $\mathsf{P}(\xi_l = i_n, \ldots \xi_{k+1} = b_{k+1} \mid \xi_k = i_{n-1}, \ldots, \xi_0 = \alpha_0) \times \mathsf{P}(T_l \leqslant t_n < T_{l+1} \mid T \in V_\alpha)$, где V_α — некоторое борелевское множество, соответствующее тому, что все компоненты T попадут куда надо ($T_k \leqslant t_{n-1} < T_{k+1}, T_{\alpha(j)} \leqslant t_j < T_{\alpha(j+1)}, 0 \leqslant j \leqslant n-2$). По марковскому свойству эта вероятность равна $\mathsf{P}(\xi_l = i_n, \ldots, \xi_{k+1} = b_{k+1} \mid \xi_k = i_{n-1}) \times \mathsf{P}(T_l \leqslant t_n < T_{l+1} \mid T \in V_\alpha)$. Чтобы увидеть, что $\mathsf{P}(A_{k,l,\beta} \mid B_\alpha) = \mathsf{P}(A_{k,l,\beta} \mid B'_\alpha) = P_\xi \times P_T$, применим нашу лемму, разбив B'_α в объединение событий, проиндексированных всевозможными вариантами попадания точек $t_0, \ldots, t_{n-2}$ в интервалы $[T_j, T_j + 1)$ и значениями элементов $\xi_0, \ldots, \xi_{k-1}$ вложенной цепи.

Докажем, наконец, что $\mathsf{P}(A \mid B'_k) = \mathsf{P}(A \mid B'_l)$.

Лемма 7.3. *Пусть A, $e_1, \ldots$ – независимые неотрицательные случайные величины, причём e_i распределены элкспоненциально (с параметрами λ_i). Тогда при условии $A \leqslant t < A + e_1$ случайные величины $E_1 = A + e_1 - t, e_2, \ldots$ тоже независимы и распределены экспоненциально с теми же параметрами ($A+e_1-t$ — с параметром λ_1). Это небольшое обобщение свойства отсутствия памяти у экспоненциального распределения.*

□ $\mathsf{P}(E_1 \geqslant t_1, \ldots, e_n \geqslant t_n \mid A \leqslant t < A+e_1) = \frac{\mathsf{P}(E_1 \geqslant t_1, A \leqslant t < A+e_1)}{\mathsf{P}(A \leqslant t < A+e_1}\mathsf{P}(e_2 \geqslant t_2) \ldots \mathsf{P}(e_n \geqslant t_n)$, в силу независимости (A, e_1) и $(e_2, \ldots, e_n)$. Вычислим оставшуюся условную вероятность.

$$\mathsf{P}(A + e_1 - t \geqslant t_1, A \leqslant t < A + e_1) = \int_0^t dP_A(x) \int_{t_1+t-x}^{\infty} \lambda_1 e^{-\lambda_1 y} dy = e^{-\lambda_1(t_1+t)} \int_0^t e^x dP_A(x),$$

$$\mathsf{P}(A \geqslant t < A + e_1) = \int_0^t dP_A(x) \int_{t-x}^{\infty} \lambda_1 e^{-\lambda_1 y} dy = e^{-\lambda_1 t} \int_0^t e^x dP_A(x).$$

Следовательно, искомая вероятность равна $e^{-\lambda_1 t_1}$, что и требовалось доказать. ■

Для того, чтобы доказать, что $\mathsf{P}(A \mid B'_k) = \mathsf{P}(A \mid B'_l)$, покажем, что эти вероятности можно собрать из одинаковых слагаемых. А именно, $\mathsf{P}(A \mid B'_k) = \sum_{l \geqslant k} \mathsf{P}(A, T_l \leqslant t_n < T_{l+1} \mid B'_k)$. Зафиксируем l и разобьём $\Pi_{l-k} = A \cap \{T_l \leqslant t_n < T_{l+1}\} \cap B'_k$ по всем возможным значениям промежуточных элементов цепи $Ъ = (\xi_{k+1}, \ldots, \xi_l)$, а условие — по значениям $\ddot{E} = (\xi_0, \ldots, \xi_{k-1})$, а далее, конечно же, докажем, что условная вероятность любого из $\Pi_{l-k}(Ъ)$ при условии $B'_k(\ddot{E})$ не зависит от $\ddot{E}$. Действительно, зафиксировав $Ъ$ и $\ddot{E}$, получаем, что наша условная вероятность распадается в произведение марковской переходной вероятности (по $\ddot{E}$) и условной вероятности попадания вектора $(T_k, \ldots, T_l) = T(\tau_k, \ldots, \tau_l; i_{n-1}, Ъ) = T(\tau_k + T_k - t_{n-1}, \ldots, \tau_i; i_{n-1}; Ъ)$ в некоторое множество при условии $T_k \leqslant t_{n-1} < T_k + \tau_k$. T_k есть функция от $(\tau_1, \ldots, \tau_{k-1}; \ddot{E})$ и поэтому независим с $\tau_k, \ldots$. Применяем лемму, которая говорит, что эта вероятность не зависит от распределения T_k, а, следовательно, и от k. ■

Пример 1.2. *Нерегулярная цепь Маркова*: λ_k не ограничены. $S = \mathbb{N}$, вложенная цепь Маркова имеет вид $\xi_k = k$ (детерминированный процесс), $\lambda_k = (k+1)^2$, $k \geqslant 0$. Матрица вероятностей переходов:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots \\ 0 & 0 & 1 & 0 & \ldots \\ 0 & 0 & 0 & 1 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$\mathsf{P}\left(\frac{\tau_k}{\lambda_k} > x\right) = e^{-(k+1)^2 x}$, $\mathsf{M}\frac{\tau_k}{\lambda_k} = \frac{1}{(k+1)^2}$ для всех $k \geqslant 0$. Отсюда

$$\mathsf{M}T_n = \sum_{k=0}^{n-1} \mathsf{M}\frac{\tau_k}{\lambda_{\xi_k}} = \sum_{k=0}^{n-1} \frac{1}{(k+1)^2} < 3$$

для всех n (ряд сходится). Последовательность $\{T_n\}$ ограничена и монотонно возрастает $\Rightarrow$ существует предел $T = \lim\limits_{n\to\infty} T_n$, $\mathsf{M}T < 3 < \infty$. К моменту T цепь совершит бесконечное число скачков и уйдёт в бесконечность. Траектория строится только на $[0; T)$.

Пусть $\xi(t)$ — однородная цепь Маркова с непрерывным временем. Величина $p_{ij}(t) = \mathsf{P}(\xi(s+t) = j \mid \xi(s) = i)$ называется вероятностью перехода из состояния i в состояние j за время t (в силу однородности она не зависит от s. $P(t) = (p_{ij}(t))_{i,j \in S}$ — матрица переходных вероятностей за время t.

Лемма 7.4. *Если $\xi(t)$ — однородная цепь Маркова с непрерывным временем, то для любых $t, s \geqslant 0$ $P(t+s) = P(t)P(s)$.*

□ Как и в дискретном случае, воспользуемся формулой полной вероятности, марковским свойством и однородностью:

$$\begin{aligned} &p_{ij}(t+s) = \mathsf{P}(\xi(s+t) = j \mid \xi(0) = i) = \\ &= \sum_{k \in S} \mathsf{P}(\xi(t) = k \mid \xi(0) = i) \cdot \mathsf{P}(\xi(t+s) = j \mid \xi(t) = k, \xi(0) = i) = \sum_{k \in S} \mathsf{P}(\xi(t) = k \mid \xi(0) = i) \cdot \mathsf{P}(\xi(t+s) = j \mid \xi(t) = k) = \\ &= \sum_{k \in S} \mathsf{P}(\xi(t) = k \mid \xi(0) = i) \cdot \mathsf{P}(\xi(s) = j \mid \xi(0) = k) = \sum_{k \in S} p_{ik}(t) p_{kj}(s), \end{aligned}$$

что и требовалось. ■

Для однородных цепей Маркова множество $\{P(s) \mid s \geqslant 0\}$ с операцией умножения матриц образует коммутативную полугруппу. В неоднородном случае матрица переходных вероятностей зависит ещё и от s, и формула приобретает вид $P(s, s+t+u) = P(s, s+t) \cdot P(s+t, s+t+u)$.

В силу предыдущей леммы для однородной цепи Маркова $P(k) = (P(1))^k$, $P(1) = (P(\frac{1}{n}))^n$.

Рассмотрим последовательность состояний $\xi(t)$ в целые моменты времени: $\xi(0), \xi(1), \xi(2), \ldots$. Это будет цепью Маркова с дискретным временем. Она называется *вложенной по-другому* цепью Маркова. Таким образом, имеются два различных понятия вложения дискретной цепи Маркова в непрерывную.

Лемма 7.5. *Если $\xi(t)$ — однородная цепь Маркова с непрерывным временем и конечным или счётным множеством состояний S с вложенной (в первом смысле) цепью Маркова ξ_n и матрицей вероятностей переходов в моменты скачков $P = (p_{ij})$, $\lambda_i < \infty$, и если состояние j вложенной цепи следует за состоянием i, то для любого $t > 0$ $p_{ij}(t) > 0$.*

□ Если $i \longrightarrow j$, то существует $k < \infty$ и существуют $i_1, \ldots, i_k \in S$ такие, что $p_{ii_1}p_{i_1i_2}\ldots p_{i_{k-1}i_k}p_{i_kj} > 0$ (из i можно перейти в j с ненулевой вероятностью за конечное число шагов). Запишем время перехода из i в j:

$$\sigma = \frac{\tau_0}{\lambda_i} + \frac{\tau_1}{\lambda_{i_1}} + \cdots + \frac{\tau_k}{\lambda_{i_k}}.$$

$p_{ij}(t) = \mathsf{P}(\xi(t) = j \mid \xi(0) = i) = p_{ii_1}\ldots p_{i_kj} \cdot \mathsf{P}(\sigma < t < \sigma + \frac{\tau_{k+1}}{\lambda_j})$ (нужно, чтобы мы за время t попали в состояние j и не успели из него выйти). Мы хотим показать, что $\mathsf{P}(\sigma < t < \sigma + \frac{\tau_{k+1}}{\lambda_j}) > 0$.

σ есть сумма независимых показательно распределённых случайных величин $\Rightarrow$ распределение σ есть свёртка распределений этих величин $\Rightarrow$ плотность f_σ распределения σ положительна на всём луче $[0, +\infty)$: $f_\sigma(x) > 0$ при $x \geqslant 0$. Отсюда

$$\mathsf{P}\left(\sigma < t < \sigma + \frac{\tau_{k+1}}{\lambda_j}\right) = \int\limits_0^t \underbrace{f_\sigma(x)}_{>0} \cdot \underbrace{\mathsf{P}\left(\frac{\tau_{k+1}}{\lambda_j} > t - x\right)}_{=e^{-(t-x)\lambda_j}>0} dx > 0,$$

что и требовалось. ■

7.2. Предельная теорема

Теорема 7.6. *Если $\xi(t)$ — однородная цепь Маркова с непрерывным временем и конечным множеством состояний S с интенсивностями выходов $\lambda_i < \infty$ и если все её состояния сообщаются, то для любых $i, j \in S$ существует предел $\lim\limits_{t\to\infty} p_{ij}(t) = p_j > 0$, не зависящий от i.*

□ Рассмотрим вложенную по-другому цепь Маркова с дискретным временем: $\xi_n = \xi(n)$. Все элементы матрицы вероятностей переходов ($P(1)$) положительны (по лемме 7.5), поэтому применима предельная теорема для цепей Маркова с дискретным временем (теорема 3.4): существует предел $\lim\limits_{n\to\infty} \mathsf{P}(\xi(n) = j \mid \xi(0) = i) = p_j$, не зависящий от i ($j = 1, \ldots, N$). Получили сходимость по подпоследовательности: $\forall\, x \geqslant 0$ $\lim\limits_{n\to\infty} \mathsf{P}(\xi(n+x) = j \mid \xi(x) = i) = p_j$ ($\{\xi(n+x)\}$ — тоже цепь Маркова с дискретным временем с матрицей $P(1)$). Докажем, что для любой последовательности $t_k \longrightarrow \infty$ $\lim_{k\longrightarrow\infty} \mathsf{P}(\xi(t_k) = j \mid \xi(0) = i) = p_j$. Очевидно, имеет место представление $t_k = n_k - d_k$, где $n_k \in \mathbb{N}, n_k \longrightarrow \infty, d_k \in (0,1)$. $\mathsf{P}(\xi(n_k - d_k) = j \mid \xi(0) = i) = \mathsf{P}(\xi(n_k) = j \mid \xi(d_k) = i) = \sum_l p_{lj}(n_k - 1)p_{il}(1 - d_k)$. Так как для всех l $p_{lj}(n_k - 1) \longrightarrow p_j \quad (k \longrightarrow \infty)$ и $\sum_l p_{il}(1 - d_k) \equiv 1$ (и $p_{il}(1 - d_k) \geqslant 0$!), то и вся сумма сходится к p_j. ■

Теорема 7.7. *Пусть $\xi(t)$ — однородная цепь Маркова с конечным или счётным множеством состояний S, $P = (p_{ij})$ — матрица вероятностей переходов в моменты скачков, $p_{ii} = 0$ для всех i, $\lambda^* = \sup\limits_{i\in S} \lambda_i < \infty$. Тогда существуют пределы*

$$q_i = \lim_{t\to 0+} \frac{1 - p_{ii}(t)}{t} = \lambda_i \qquad \text{и} \qquad q_{ij} = \lim_{t\to 0+} \frac{p_{ij}(t)}{t} = \lambda_i p_{ij}.$$

□ Пусть $\xi(0) = i$, $\tau_i = \inf\{t > 0 : \xi(t) \neq i\}$ — момент первого выхода из i. $\mathsf{P}\{\tau_i > x\} = e^{-\lambda_i x}$ при $x \geqslant 0$. По формуле полной вероятности:

$$\mathsf{P}\{\tau_i > t\} \leqslant p_{ii}(t) = \mathsf{P}(\tau_i > t) + \int\limits_0^t \sum_{j\in S\setminus\{i\}} p_{ij}p_{ji}(t-u)\, d\mathsf{P}\{\tau_i \leqslant u\}.$$

Далее, $p_{ji}(t-u) \leqslant \mathsf{P}\{\tau_j \leqslant t-u\} = 1 - e^{-\lambda_j(t-u)}$, $d\mathsf{P}\{\tau_i \leqslant u\} = \lambda_i e^{-\lambda_i u}\, du$, откуда

$$\int\limits_0^t \sum_{j\in S\setminus\{i\}} p_{ij} p_{ji}(t-u)\, d\mathsf{P}\{\tau_i \leqslant u\} \leqslant \int\limits_0^t \sum_{j\in S\setminus\{i\}} p_{ij} \underbrace{\left(1 - e^{-\lambda_j(t-u)}\right)}_{\leqslant \lambda_j(t-u) \leqslant \lambda^*(t-u)} \lambda_i \underbrace{e^{-\lambda_i u}}_{\leqslant 1}\, du \leqslant$$

$$\leqslant \int\limits_0^t \sum_{j\in S\setminus\{i\}} p_{ij} \lambda^*(t-u)\lambda_i\, du \leqslant \lambda^*\lambda_i \frac{t^2}{2} = O(t^2) \quad (t \longrightarrow 0+),$$

где предпоследний переход верен в силу того, что $\sum\limits_j p_{ij} \leqslant 1$, $\int\limits_0^t (t-u)\, du = t^2/2$.

Итак, $e^{-\lambda_i t} = \mathsf{P}\{\tau_i > t\} \leqslant p_{ii}(t) \leqslant e^{-\lambda_i t} + O(t^2) \Rightarrow 1 - e^{-\lambda_i t} \leqslant 1 - p_{ii}(t) \leqslant 1 - e^{-\lambda_i t} + O(t^2)$, откуда (т.к. $1 - e^{-\lambda_i t} = \lambda_i t + O(t^2)$) получаем, что $1 - p_{ii}(t) = \lambda_i t + O(t^2)$, что и доказывает первое предельное соотношение.

Пусть теперь $i \neq j$. Оценим $p_{ij}(t)$ снизу по формуле полной вероятности вероятностью того, что переход из i в j произошёл сразу (без промежуточных состояний) и оценим получившийся интеграл:

$$p_{ij}(t) \geqslant \int\limits_0^t p_{ij} \mathsf{P}\{\tau_j > t-u\}\, d\mathsf{P}\{\tau_i \leqslant u\} = p_{ij} \int\limits_t^0 e^{-\lambda_j(t-u)} \lambda_i e^{-\lambda_i u}\, du = p_{ij}\lambda_i t + o(t) \quad (t \longrightarrow 0+)$$

(экспоненты равномерно стремятся к 1 при $t \longrightarrow 0+$). Отсюда получаем, что

$$\limsup_{t\to 0+} \frac{p_{ij}(t)}{t} \geqslant \liminf_{t\to 0+} \frac{p_{ij}(t)}{t} \geqslant \lambda_i p_{ij}$$

Суммируем по j:

$$\sum_{j\neq i} \liminf_{t\to 0+} \frac{p_{ij}(t)}{t} \geqslant \sum_{j\neq i} \lambda_i p_{ij} = \lambda_i,$$

т.к. $\sum\limits_{j\neq i} p_{ij} = 1$ ($p_{ii} = 0$ по условию).

$\lambda^* = \sup \lambda_i < \infty$, следовательно цепь Маркова почти наверное не уходит в бесконечность. Поэтому $\sum\limits_{j\in S} p_{ij}(t) = 1$, откуда $\sum\limits_{j\neq i} p_{ij}(t) = 1 - p_{ii}(t)$. Отсюда, применяя первое (уже доказанное) соотношение, получаем:

$$\limsup_{t\to 0+} \sum_{j\neq i} \frac{p_{ij}(t)}{t} = \limsup_{t\to 0+} \frac{1 - p_{ii}(t)}{t} = \lambda_i.$$

Далее,

$$\lambda_i = \sum_{j\neq i} \lambda_i p_{ij} \leqslant \sum_{j\neq i} \liminf_{t\to 0+} \frac{p_{ij}(t)}{t} \leqslant \limsup_{t\to 0+} \sum_{j\neq i} \frac{p_{ij}(t)}{t} = \lambda_i.$$

Но начало и конец цепочки совпадают, значит все неравенства обращаются в равенства. Отсюда для любого j верхний и нижний пределы $p_{ij}(t)/t$ совпадают и равны $\lambda_i p_{ij}$, откуда получается второе соотношение. ■

Числами q_i и q_{ij} можно задавать цепь Маркова. Они образуют *матрицу интенсивностей переходов* $Q = (q_{ij})_{i,j\in S}$ (q_{ii} считается равным $-q_i$). Сумма элементов в каждой строке этой матрицы равна нулю.

Утверждение теоремы 7.7 можно переписать в матричной форме:

$$\lim_{t\to 0+} \frac{P(t) - P(0)}{t} = Q$$

(считается, что $P(0)$ есть единичная матрица).

Теорема 7.8 (уравнения Колмогорова). *В условиях предыдущей теоремы справедливы следующие равенства:*

1. $P'(t) = P(t)Q$ *(обратное уравнение Колмогорова);*
2. $P'(t) = QP(t)$ *(прямое уравнение Колмогорова).*

Эта система о.д.у. имеет решение, записываемое в виде $P(t) = e^{tQ}$, *где* $e^A = E + \sum\limits_{k=1}^{\infty} \frac{1}{k!} A^k$ *— экспонента матрицы* A.

☐ По лемме 7.4 $P(t+s) = P(t)P(s)$. Отсюда ($P(0) = E$): $P(s+t) - P(t) = P(t)(P(s) - P(0)) = (P(s) - P(0))P(t)$. По теореме 7.7 получаем[3]:

$$P'(t) = \lim_{s\to 0+} \frac{P(t+s) - P(t)}{s} = \lim_{s\to 0+} \frac{P(t)(P(s) - P(0))}{s} = P(t) \lim_{s\to 0+} \frac{P(s) - P(0)}{s} = P(t)Q.$$

Аналогично получаем, что $P'(t) = QP(t)$. Далее, $P''(t) = (P'(t))' = P'(t)Q = P(t)Q^2$, …, $P^{(k)}(t) = P(t)Q^k$. Разлагаем $P(t)$ в ряд Тейлора в нуле:

$$P(t) = P(0) + \sum_{k=1}^{\infty} \frac{t^k}{k!} P^{(k)}(0) = E + \sum_{k=1}^{\infty} \frac{(tQ)^k}{k!} = e^{tQ},$$

что и требовалось. ■

8. Мартингалы

8.1. Напоминание про условные математические ожидания (в дискретном случае)

Из курса теории вероятностей мы знаем про условную вероятность. Пусть есть вероятностное пространство $(\Omega, \mathscr{F}, \mathsf{P})$, $B, C \in \mathscr{F}$. *Условная вероятность* события C при условии события B:

$$\mathsf{P}(C \mid B) := \frac{\mathsf{P}(CB)}{\mathsf{P}(B)}.$$

Будем теперь изменять событие C при фиксированном B. Получим функцию на $\mathscr{F}$ — новую (условную) вероятностную меру на $(\Omega, \mathscr{F})$, сосредоточенную на множестве B. По этой новой мере можно брать математическое ожидание — *условное математическое ожидание при условии события* B.

Теперь будем менять событие B. Пусть дано разбиение $\Omega = \bigsqcup\limits_{k\geqslant 1} B_k$. С этим распределением связана σ-алгебра $\mathscr{B} = \sigma\{B_k\}$, состоящая из всевозможных объединений событий B_k. Для фиксированного события C получаем набор чисел (условных вероятностей), заиндексированный индексом k.

Условная вероятность относительно $\mathscr{B}$ — ступенчатая функция на Ω (это случайная величина), которая на каждом B_k равна $\mathsf{P}(C \mid B_k)$ (заметим, что эта функция $\mathscr{B}$-измерима).

Пусть теперь дана дискретная случайная величина ξ ($\{\omega : \xi = x_k\} = A_k$). *Условное математическое ожидание* ξ относительно $\mathscr{B}$ есть $\mathsf{M}(\xi \mid \mathscr{B})(\omega) = \sum x_k \mathsf{P}(A \mid \mathscr{B})(\omega)$ (это — тоже случайная величина, измеримая относительно $\mathscr{B}$, т.е. ступенчатая функция).

Дискретная случайная величина ξ порождает разбиение $\Omega = \bigsqcup\limits_{k} A_k$. Соответствующая σ-алгебра называется $\sigma(\xi)$. Пишем $\mathsf{M}(\eta \mid \xi) := \mathsf{M}(\eta \mid \sigma(\xi))$.

Разумеется, всю эту теорию можно развивать для произвольных σ-алгебр. Такие вещи обычно рассказывают в курсе математической статистики. Здесь мы просто напоминаем про у.м.о. в простейшем случае...

Свойства у.м.о.:

1. Если $C \in \mathscr{B}$, то $\mathsf{M}\chi_C\xi = \mathsf{M}(\chi_C \mathsf{M}(\xi \mid \mathscr{B}))$ (χ_C — индикатор C). В частности (если $C = \Omega$): $\mathsf{M}\xi = \mathsf{M}\mathsf{M}(\xi \mid \mathscr{B})$.
2. Линейность: $\mathsf{M}(a\xi + b\eta \mid \mathscr{B}) = a\mathsf{M}(\xi \mid \mathscr{B}) + b\mathsf{M}(\eta \mid \mathscr{B})$.
3. $\mathsf{M}(\chi_A \mid \mathscr{B}) = \mathsf{P}(A \mid \mathscr{B})$.
4. Если $\mathscr{B} = \sigma(\xi)$, то $\mathsf{M}(\xi \mid \mathscr{B}) = \xi$.
5. Если $\mathscr{B}_1 \subseteq \mathscr{B}_2$ ($\mathscr{B}_2$ соответствует подразбиению $\mathscr{B}_1$ — более грубой σ-алгебры), то $\mathsf{M}(\mathsf{M}(\xi \mid \mathscr{B}_2) \mid \mathscr{B}_1) = \mathsf{M}(\xi \mid \mathscr{B}_1)$.
6. Если $\mathscr{B} = \sigma(\eta)$, то $\mathsf{M}(\xi\eta \mid \mathscr{B}) = \eta\mathsf{M}(\xi \mid \mathscr{B})$ (множитель, измеримый относительно условия, можно выносить из-под знака у.м.о.).

8.2. Мартингалы с дискретным временем: определение и простые свойства

В нашем курсе других мартингалов и не будет...

Определение. Пусть $(\Omega, \mathscr{F}, \mathsf{P})$ — вероятностное пространство, на котором заданы последовательность случайных величин $X_0, X_1, X_2, \ldots$ и расширяющаяся последовательность (*поток*) σ-алгебр $\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \mathscr{F}_2 \subseteq \cdots \subseteq \mathscr{F}$. Последовательность пар $\{(X_n, \mathscr{F}_n)\}_{n\in\mathbb{N}}$ называется *мартингалом* (говорят также, что $\{X_n\}$ — мартингал относительно $\{\mathscr{F}_n\}$), если:

[3] На самом деле это всего лишь правая производная (так что в формулировке есть лукавство). ~~Но для разложения Тейлора этого хватит.~~ Ага, щ-щас. Этого для разложения Тейлора, конечно, не хватит, но если вспомнить, что $P(s-t)P(t) = P(s)$ и устремить $t \to 0+$, получим, что $P(t)$ непрерывна слева. Дифференцируемость слева тогда доказывается элементарно заменой t на $-t$.

1. $\mathsf{M}|X_n| < \infty$ для всех n;
2. X_n измерима относительно $\mathscr{F}_n$ для всех n;
3. $\mathsf{M}(X_m \mid \mathscr{F}_n) = X_n$ почти наверное, если $m > n$ (характеристическое свойство мартингала).

Если вместо свойства 3 написать $\mathsf{M}(X_m \mid \mathscr{F}_n) \geqslant X_n$, получится *субмартингал*, а если $\mathsf{M}(X_m \mid \mathscr{F}_n) \leqslant X_n$ — *супермартингал*.

Изменение знака переводит субмартингал в супермартингал (и наоборот). Поэтому иногда эти два понятия объединяют и говорят о *семи-, или полумартингалах*.

Утверждение 8.1. *Для дискретного времени характеристическое свойство мартингала эквивалентно своему частному случаю: для любого* n $\mathsf{M}(X_{n+1} \mid \mathscr{F}_n) = X_n$.

□ $\mathsf{M}(X_m \mid \mathscr{F}_n) = \mathsf{M}(\mathsf{M}(X_m \mid \mathscr{F}_{m-1}) \mid \mathscr{F}_m) = \mathsf{M}(X_{m-1} \mid \mathscr{F}_n)$. Далее по индукции. ■

Утверждение 8.2. *Для мартингала* $\mathsf{M}(X_{n+1} \mid X_n) = X_n$.

□ $\mathsf{M}(X_{n+1} \mid X_n) = \mathsf{M}(\mathsf{M}(X_{n+1} \mid \mathscr{F}_n) \mid \sigma(X_n)) = \mathsf{M}(X_n \mid \sigma(X_n)) = X_n$. ■

8.3. Примеры мартингалов

1. Пусть $\xi_1, \xi_2, \ldots$ — независимые случайные величины. $X_n = \xi_1 + \cdots + \xi_n$, $X_0 = 0$ — случайное блуждание. $\mathscr{F}_n = \sigma(\xi_1, \ldots, \xi_n)$. Если $\mathsf{M}\xi_n = 0$ для всех n, то $\{(X_n, \mathscr{F}_n)\}$ — мартингал, если $\leqslant 0$ — субмартингал, если $\geqslant 0$ — супермартингал.

2. $\xi_1, \xi_2, \ldots$ — независимые случайные величины, $\mathscr{F}_n = \sigma(\xi_1, \ldots, \xi_n)$, $X_0 = 1$, $X_n = \prod\limits_{i=1}^{n} \xi_i$. Если для всех n $\mathsf{M}\xi_n = 1$, то $\{(X_n, \mathscr{F}_n)\}$ — мартингал. Если же $\mathsf{M}\xi_n >= 1$, то субмартингалом такая последовательность, вообще говоря, не будет. Однако если потребовать, например, $\xi_i \geqslant 0$ п.н., то X_i всё же будут образовывать супермартингал.

3. $\xi_1, \xi_2, \ldots$ — независимые случайные величины, $\mathscr{F}_n = \sigma(\xi_1, \ldots, \xi_n)$, $X_1 = X_2 = 0$, $X_n = \xi_1\xi_2 + \xi_2\xi_3 + \ldots + \xi_{n-1}\xi_n$. Если $\forall n$ $\mathsf{M}\xi_n = 0$, то $(X_n, \mathscr{F}_n)$ — мартингал. Если $\forall n$ $\mathsf{M}\xi_{2n} = 0$, то $\forall i$ $\mathsf{M}\xi_i\xi_{i+1} = 0$, но $\{X_n, \mathscr{F}_n\}$ могут и не образовывать мартингал. Рассмотрим, например, $\mathsf{P}\{\xi_{2k} = -1\} = \mathsf{P}\{\xi_{2k} = 1\} = \frac{1}{2}, \mathsf{M}\xi_{2k} = 0$, и $\mathsf{P}\{\xi_{2k-1} = 1\} = \mathsf{P}\{\xi_{2k-1} = 3\} = \frac{1}{2}, \mathsf{M}\xi_{2k-1} = 2 \neq 0$. Тогда
$$\mathsf{M}(X_3 \mid X_2 = 3) = \mathsf{M}(\xi_1\xi_2 + \xi_2\xi_3 \mid \xi_1 = 3, \xi_2 = 1) = 3 + \mathsf{M}(\xi_3) = 5 \neq 3.$$
Таким образом, нарастающая сумма случайных величин с нулевыми математическими ожиданиями не всегда образует мартингал.

4. Пусть $(\Omega, \mathscr{F}, \mathsf{P})$ — вероятностное пространство, $\mathscr{F}_1 \subseteq \mathscr{F}_2 \subseteq \ldots \subseteq \mathscr{F}$ — поток σ-алгебр, ξ — фиксированная случайная величина с $\mathsf{M}\xi < \infty$. Тогда $X_n \rightleftharpoons \mathsf{M}(\xi \mid \mathscr{F}_n)$ — мартингал относительно $\{\mathscr{F}_n\}$.

   □ Измеримость X_n относительно $\mathscr{F}_n$ очевидна, так как X_n — это у.м.о. относительно $\mathscr{F}_n$. Для любых $m > n$ имеем
$$\mathsf{M}(X_m \mid \mathscr{F}_n) = \mathsf{M}(\mathsf{M}(\xi \mid \mathscr{F}_m) \mid \mathscr{F}_n) = \mathsf{M}(\xi \mid \mathscr{F}_n) = X_n, \text{ т.к. } \mathscr{F}_n \subseteq \mathscr{F}_m.$$
   ■

   В качестве примера такого мартингала рассмотрим $\Omega = [0, 1)$ с линейной мерой Лебега, $\xi(\omega) \equiv \omega$, а последовательность $\mathscr{F}_n$ — последовательные дробления $[0, 1]$ на 2^k частей, т.е. $\mathscr{F}_n = \left\{[\frac{k}{2^n}, \frac{k+1}{2^n}) \mid 0 \leqslant k \leqslant 2^n - 1\right\}$. Тогда $X_n(\omega) = \frac{2k+1}{2^{2n+1}}$ при $\frac{k}{2^n} \leqslant \omega < \frac{k+1}{2^n}$.

5. Пусть $X_1, X_2, \ldots$ — случайные величины со значениями в конечном множестве $\{1, \ldots, N\}$. Если $(X_n, \sigma(X_1, \ldots, X_n))$ — мартингал, то состояния 0 и N — поглощающие, т.е. $\mathsf{P}\{X_{m+1} = 0 \mid X_m = 0\} = 1 = \mathsf{P}\{X_{m+1} = N \mid X_m = N\}$.

   □ От противного, если $\mathsf{P}\{X_{m+1} > 0 \mid X_m = 0\} > 0$, то
$$0 = X_m = \mathsf{M}(X_{m+1} \mid X_m = 0) = \sum_k k\mathsf{P}\{X_{m+1} = k \mid X_m = 0\} > 0$$
   . Случай $X_m = N$ аналогичен. ■

6. Пусть $\xi_t, t \in \mathbb{Z}$ — независимые случайные величины, $\mathsf{P}\{\xi_t = 1\} = p, \mathsf{P}\{\xi_t = 0\} = 1 - p$,
$$Y_n \rightleftharpoons \min\{k > 0 \mid \xi_{n-k} + \xi_{n+k} > 0\}.$$
   Y_n не образуют мартингала относительно $\{\sigma(\xi_t, t \leqslant n)\}$ в силу неизмеримости Y_n относительно «будущего»: Y_n не обязательно постоянна на атомах σ-алгебры $\sigma(\xi_{n-1}, \ldots)$, т.к. зависит ещё и от $\xi_{n+1}, \ldots$.

8.4. Предсказуемые последовательности случайных величин. Разложение Дуба

Определение. Последовательность $\{\xi_n\}$ называется *предсказуемой* относительно потока σ-алгебр $\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \dots \subseteq \mathscr{F}$, если ξ_n измерима относительно $\mathscr{F}_{n-1}$.

Неформально говоря, σ-алгебра $\mathscr{F}_n$ описывает всё, что было до момента n.

Пример 4.1. $\zeta_1, \zeta_2, \dots$ — последовательность случайных величин. $\mathscr{F}_n = \sigma(\zeta_1, \dots, \zeta_n)$, $n \geqslant 1$. $f_n : \mathbb{R}^n \to \mathbb{R}$. $\xi_n = f_{n-1}(\zeta_1, \dots, \zeta_{n-1})$ — предсказуема относительно $\{\mathscr{F}_n\}$.

Теорема 8.3 (разложение Дуба[4]). *Пусть $\{(X_n, \mathscr{F}_n)\}$ — субмартингал. Тогда существуют такие мартингал $\{(Z_n, \mathscr{F}_n)\}$ и последовательность случайных величин $\{A_n\}$ такие, что: $A_0 = 0$, $\{A_n\}$ предсказуема относительно $\{\mathscr{F}_n\}$, A_n неубывает с вероятностью 1; $X_n = Z_n + A_n$ почти наверное для всех n. С точностью до множества меры нуль последовательности $\{A_n\}$ и $\{Z_n\}$ определяются однозначно.*

Последовательность $\{A_n\}$ называется *компенсатором.*

□ $A_n = 0 \Rightarrow Z_0 = X_0$ — определяется однозначно. Далее применяем несколько тривиальных тождеств:

$$X_n = X_0 + \sum_{j=0}^{n-1}(X_{j+1} - X_j) = X_0 + \sum_{j=0}^{n-1}((X_{j+1} - \mathsf{M}(X_{j+1} \mid \mathscr{F}_j)) + (\mathsf{M}(X_{j+1} \mid \mathscr{F}_j) - X_j)).$$

Положим ($Z_0 = X_0$)

$$Z_n := Z_0 + \sum_{j=0}^{n-1}(X_{j+1} - \mathsf{M}(X_{j+1} \mid \mathscr{F}_j)) = Z_{n-1} + (X_n - \mathsf{M}(X_n \mid \mathscr{F}_{n-1}) \qquad \text{и} \qquad A_n = \sum_{j=0}^{n-1}(\mathsf{M}(X_{j+1} \mid \mathscr{F}_j) - X_j).$$

$\{X_n\}$ — субмартингал, поэтому все слагаемые в A_n почти наверное неотрицательны $\Rightarrow$ A_n почти наверное неубывает (как нарастающая сумма неотрицательных случайных величин). Каждая скобка в A_n измерима относительно $\mathscr{F}_j \Rightarrow \{A_n\}$ предсказуема относительно $\{\mathscr{F}_n\}$ (A_n $\mathscr{F}_{n-1}$-измерима).

Проверим, что Z_n — мартингал:

$$\begin{aligned}\mathsf{M}(Z_n \mid \mathscr{F}_{n-1}) = \mathsf{M}(Z_{n-1} + (X_n - \mathsf{M}(X_n \mid \mathscr{F}_{n-1})) \mid \mathscr{F}_{n-1}) = \mathsf{M}(Z_{n-1} \mid \mathscr{F}_{n-1}) + \mathsf{M}(X_n \mid \mathscr{F}_{n-1}) - \mathsf{M}(X_n \mid \mathscr{F}_{n-1}) = \\ = Z_{n-1}.\end{aligned}$$

Единственность. От противного. Пусть $X_n = Z_n^* + A_n^*$ почти наверное, Z_n^* — мартингал, A_n^* — неубывающая предсказуемая последовательность, $A_0^* = 0$. Рассмотрим разность $X_{n+1} - X_n$:

$$X_{n+1} - X_n = (Z_{n+1} - Z_n) + (A_{n+1} - A_n) = (Z_{n+1}^* - Z_n^*) + (A_{n+1}^* - A_n^*) \quad \text{п.н.}$$

Вычислим у.м.о. относительно $\mathscr{F}_n$:

$$\mathsf{M}((Z_{n+1} - Z_n) + (A_{n+1} - A_n) \mid \mathscr{F}_n) = \mathsf{M}((Z_{n+1}^* - Z_n^*) + (A_{n+1}^* - A_n^*) \mid \mathscr{F}_n).$$

Далее, $\{Z_n\}$ и $\{Z_n^*\}$ — мартингалы, поэтому $\mathsf{M}(Z_{n+1} - Z_n \mid \mathscr{F}_n) = 0 = \mathsf{M}(Z_{n+1}^* - Z_n^* \mid \mathscr{F}_n)$. $\{A_n\}$ и $\{A_n^*\}$ предсказуемы, откуда $\mathsf{M}(A_{n+1} - A_n \mid \mathscr{F}_n) = A_{n+1} - A_n$, $\mathsf{M}(A_{n+1}^* - A_n^* \mid \mathscr{F}_n) = A_{n+1}^* - A_n^*$. Итак, $A_{n+1} - A_n = A_{n+1}^* - A_n^*$, $A_0 = 0 = A_0^*$. Отсюда по индукции получаем $A_n^* = A_n$. ■

Лемма 8.4. *Если $\{(X_n, \mathscr{F}_n)\}$ — мартингал, $g(x)$ — выпуклая функция и для всех n $M|g(X_n)| < \infty$, то последовательность $\{(g(X_n), \mathscr{F}_n)\}$ — субмартингал.*

Если $\{(X_n, \mathscr{F}_n)\}$ — субмартингал, $g(x)$ выпукла и неубывает, $\mathsf{M}|g(X_n)| < \infty$ для всех n, то $\{(g(X_n), \mathscr{F}_n)\}$ — субмартингал.

□ Нужно доказать[5], что $\mathsf{M}(g(X_{n+1}) \mid \mathscr{F}_n) \geqslant g(X_n)$.

g выпукла $\Rightarrow$ существует неубывающая функция g^*: для всех x, y имеем $g(y) \geqslant g(x) + g^*(x)(y - x)$. (Чтобы доказать это, рассматриваем в данной любую опорную прямую, т.е. такую прямую, что график g лежит целиком выше её. В случае дифференцируемой g всё вообще очевидно.)

Далее,

$$\mathsf{M}(g(X_{n+1}) \mid \mathscr{F}_n) \geqslant \mathsf{M}(g(X_n) + g^*(X_n)(X_{n+1} - X_n) \mid \mathscr{F}_n) = g(X_n) + g^*(X_n)\mathsf{M}(X_{n+1} - X_n \mid \mathscr{F}_n).$$

В первом случае ($\{X_n\}$ — мартингал) второе слагаемое равно нулю, и мы сразу получаем требуемое.

Во втором случае $\mathsf{M}(X_{n+1} - X_n \mid \mathscr{F}_n) \geqslant 0$, $g^*(X_n) \geqslant 0$ (т.к. g неубывает), откуда второе слагаемое неотрицательно, что и даёт субмартингальное свойство. ■

[4] Это не дерево, а американский математик по фамилии Doob.

[5] На самом деле всё это делается в одно касание применением неравенства Йенсена. Здесь оно по сути ещё раз доказано — примеч. С.К.

8.5. Моменты остановки

Определение. Пусть $(\Omega, \mathscr{F}, \mathsf{P})$ — вероятностное пространство, $\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \dots \subseteq \mathscr{F}$ — поток σ-алгебр, ν — целочисленная случайная величина. ν — *момент остановки* относительно $\{\mathscr{F}_n\}$, если для любого n $\{\omega : \nu(\omega) \leqslant n\} \in \mathscr{F}_n$. Говорят также: ν — *марковский момент*, ν — случайная величина, *не зависящая от будущего.*

Пример 5.1. Пусть $\{X_n\}$ — последовательность случайных величин, X_n измерима относительно $\mathscr{F}_n$ для любого n, $B_n \in \mathscr{F}_n$ — последовательность множеств. $\nu = \min\{n : X_n \in B_n\}$ — момент остановки.

Частный случай: $\nu = \min\{n : X_n > f(n)\}$ — первый момент, когда X_n превосходит $f(n)$.

Свойства моментов остановки:

1. Если ν — момент остановки, m — фиксированное натуральное число, то $\nu^* = \min\{m, \nu\}$ — тоже момент остановки.

   $\square$ $m \geqslant n \Rightarrow \{\nu^* \leqslant n\} = \Omega \in \mathscr{F}_n$. $m < n \Rightarrow \{\nu^* \leqslant n\} = \{\nu \leqslant n\} \in \mathscr{F}_n$. $\blacksquare$

2. Пусть ν — момент остановки. Тогда $\{\nu = n\} = \{\nu \leqslant n\} \setminus \{\nu \leqslant n-1\} \in \mathscr{F}_n$; $\{\nu \geqslant n\} = \Omega \setminus \{\nu \leqslant n-1\} \in \mathscr{F}_{n-1}$.

Пусть ν — момент остановки относительно $\{\mathscr{F}_n\}$, $\{(X_n, \mathscr{F}_n)\}$ — (полу)мартингал. Определим значение X_n в случайный момент времени: $X_\nu = \sum\limits_{n \geqslant 0} X_n \mathrm{I}_{\{\nu = n\}}$.

Лемма 8.5. *Если $\{(X_n, \mathscr{F}_n)\}$ — (полу)мартингал и ν — момент остановки относительно $\{\mathscr{F}_n\}$, то остановленная последовательность $\{X_{\min\{\nu,n\}}\}$ — (полу)мартингал относительно $\{\mathscr{F}_n\}$.*

$\square$ $X_{\min\{\nu,n\}} = \sum\limits_{m=0}^{n} X_m \mathrm{I}_{\{\nu=m\}} + X_n \mathrm{I}_{\{\nu>n\}}$, откуда $X_{\min\{\nu,n\}}$ измерима относительно $\mathscr{F}_n$. Далее, $X_{\min\{\nu,n+1\}} = \sum\limits_{m=0}^{n+1} X_m \mathrm{I}_{\{\nu=m\}} + X_{n+1}\mathrm{I}_{\nu>n+1} = \sum\limits_{m=0}^{n} X_m \mathrm{I}_{\{\nu=m\}} + X_{n+1}\mathrm{I}_{\nu>n}$. Отсюда $X_{\min\{\nu,n+1\}} - X_{\min\{\nu,n\}} = \mathrm{I}_{\{\nu>n\}}(X_{n+1} - X_n)$. Поэтому $\mathsf{M}(X_{\min\{\nu,n+1\}} - X_{\min\{\nu,n\}} \mid \mathscr{F}_n) = \mathsf{M}(\mathrm{I}_{\{\nu>n\}}(X_{n+1} - X_n) \mid \mathscr{F}_n) = \mathrm{I}_{\{\nu>n\}}\mathsf{M}(X_{n+1} - X_n \mid \mathscr{F}_n)$ — знак совпадает со знаком $\mathsf{M}(X_{n+1} - X_n \mid \mathscr{F}_n)$ (0 для мартингала, „+" для субмартингала, „−" для супермартингала), что и требовалось. $\blacksquare$

Пусть $\{X_n\}$ — мартингал. Тогда (из мартингального свойства) $\mathsf{M}X_n = \mathsf{M}X_0$. Верно ли это для X_ν? Казалось бы, так как $\min\{n, \nu\} \longrightarrow \nu$ $(n \longrightarrow \infty)$, получаем $\mathsf{M}X_0 = \mathsf{M}X_{\min\{n,\nu\}} \longrightarrow \mathsf{M}X_\nu$. Но последний предельный переход (по сути, переход к пределу под знаком интеграла) не всегда законен.

Пример 5.2. $\xi_n \sim \mathcal{N}(0,1)$ — независимые; $X_0 = 0$, $X_n = \xi_1 + \dots + \xi_n$ — случайное блуждание ($\{X_n\}$ — мартингал относительно $\{\sigma(X_1, \dots, X_n)\}$); $\nu = \min\{n : X_n < 0\}$ — момент остановки.

По закону повторного логарифма:

$$\limsup_{n\to\infty} \frac{X_n}{\sqrt{2n\ln\ln n}} = +1 \qquad \text{и} \qquad \liminf_{n\to\infty} \frac{X_n}{\sqrt{2n\ln\ln n}} = -1,$$

поэтому почти наверное X_n меняет знак бесконечное число раз $\Rightarrow \mathsf{P}\{\nu < \infty\} = 1 \Rightarrow \mathsf{M}X_\nu < 0 = \mathsf{M}X_0$.

Пусть $\{\mathscr{F}_n\}$ — поток σ-алгебр, ν — момент остановки. Определим σ-алгебру $\mathscr{F}_\nu$ следующими эквивалентными способами:

$$\mathscr{F}_\nu = \left\{A \in \mathscr{F} : A = \bigcup_{m=0}^{\infty} (\{\nu = m\} \cap B_m), B_m \in \mathscr{F}_m \;\forall m\right\} = \{A \in \mathscr{F} : A \cap \{\nu = m\} \in \mathscr{F}_m \;\forall m\} =$$
$$= \sigma\left\{\{\nu \leqslant n\} \cap B_n, B_n \in \mathscr{F}_n\right\}$$

σ-алгебра $\mathscr{F}_\nu$, таким образом, порождается совокупностью событий вида $\{\nu \leqslant n\} \cap B_n, B_n \in \mathscr{F}_n$. ν измерима относительно $\mathscr{F}_\nu$ (достаточно взять $B_n = \Omega \;\forall n$). Если $\{X_n, \mathscr{F}_n\}$ — мартингал, то X_ν измерима относительно $\mathscr{F}_\nu$:

$$\{X_\nu \in B\} = \cup_{n=0}^{\infty} \{\nu = n, X_n \in B\} \in \mathscr{F}_\nu, \text{т.к. } X_n \; \mathscr{F}_n\text{-измерима.}$$

Вообще говоря, $\mathscr{F}_\nu \neq \mathscr{F}_n \forall n$.

Утверждение 8.6. *Пусть ν_1, ν_2 – два момента остановки относительно $\{\mathscr{F}_n\}$. Тогда $\{\nu_2 \geqslant \nu_1\} \in \mathscr{F}_{\nu_1} \cap \mathscr{F}_{\nu_2}$.*

$\square$

$$\{\nu_2 \geqslant \nu_1\} = \begin{cases} \cup_{n=0}^{\infty} (\{\nu_1 = n\} \cap \{\nu_2 \geqslant n\}) \in \mathscr{F}_{\nu_1}, & \text{т.к. } \{\nu_2 \geqslant n\} \in \mathscr{F}_n, \\ \cup_{n=0}^{\infty} (\{\nu_2 = n\} \cap \{\nu_1 \leqslant n\}) \in \mathscr{F}_{\nu_2}, & \text{т.к. } \{\nu_1 \leqslant n\} \in \mathscr{F}_n. \end{cases}$$

$\blacksquare$

Теорема 8.7 (Сохранение мартингальности для одного момента остановки). *Пусть* $\{(X_n, \mathscr{F}_n)\}$ *— мартингал,* ν *— момент остановки относительно* $\{\mathscr{F}_n\}$, $\mathsf{M}|X_\nu| < \infty, \liminf_{n\longrightarrow\infty} \mathsf{M}\left(|X_n| \mathrm{I}\{\nu > n\}\right) = 0$. *Тогда с вероятностью* 1

$$\mathsf{M}\left(X_\nu \mid \mathscr{F}_0\right) = X_0.$$

Аналогичное свойство выполняется для полумартингалов (с заменой равенства на соответствующее неравенство).

Перед тем, как доказать эту теорему, сделаем замечание общего характера.

Замечание. Пусть f, g — измеримые функции на $(\Omega, \mathscr{F}), \mathsf{P}\{f \neq g\} > 0$. Тогда $\mathsf{P}\{f > g\} > 0$ или $\mathsf{P}\{f < g\} > 0$. Пусть $\mathsf{P}\{f > g\} > 0$. В этом случае

$$\int_{\{\omega\colon f > g\}} f\mathsf{P}(d\omega) > \int_{\{\omega\colon f>g\}} g\mathsf{P}(d\omega).$$

Следовательно, если $\forall A \in \mathscr{F}\colon \int_A f\mathsf{P}(d\omega) = \int_A g\mathsf{P}(d\omega)$, то $f = g$ п.н. В нашем случае достаточно будет рассматривать $A \in \mathscr{F}_0$, т.е. доказать, что X_0 является вариантом $\mathsf{M}(X_\nu \mid \mathscr{F}_0)$; по свойству у.м.о. все его варианты совпадают почти наверное.

□ [Доказательство теоремы 8.7] Рассмотрим $f(\omega) = \mathsf{M}\left(X_\nu \mid \mathscr{F}_0\right), g(\omega) = X_0$. Согласно замечанию, достаточно установить, т.к. f, g — $\mathscr{F}_0$-измеримые функции, что $\forall A \in \mathscr{F}_0$

$$\mathsf{M}\left(X_\nu \mathrm{I}_A\right) = \mathsf{M}\mathsf{M}\left(\mathrm{I}_A X_\nu \mid \mathscr{F}_0\right) = \mathsf{M}\left(\mathsf{M}(X_\nu \mid \mathscr{F}_0)\mathrm{I}_A\right) = \mathsf{M}\left(X_0 \mathrm{I}_A\right),$$

то есть доказать, что $\mathsf{M}\left(X_\nu \mathrm{I}_A\right) = \mathsf{M}\left(X_0 \mathrm{I}_A\right)$.

Лемма 8.8. *Если* $\{(X_n, \mathscr{F}_n)\}$ *— мартингал,* $A \in \mathscr{F}_0$, ν *— момент остановки относительно* $\{\mathscr{F}_n\}$ *и выполнены условия теоремы 8.7, то*

$$\forall n \geqslant 0\colon \mathsf{M}\left(X_0 \mathrm{I}_A\right) = \mathsf{M}\left(X_{\min\{n,\nu\}} \mathrm{I}_A\right) = \mathsf{M} X_\nu \mathrm{I}_{A\cap\{\nu\leqslant n\}} + \mathsf{M} X_n \mathrm{I}_{A\cap\{\nu>n\}}.$$

□ Первое равенство выполнено в силу того, что $X_{\min\{n,\nu\}}$ — мартингал, второе получается разложением I_A. ■

По условию $\liminf_{n\longrightarrow\infty} \mathsf{M} X_n \mathrm{I}_{A\cap\{\nu>n\}} = 0$ существует последовательность $\{n_k\}\colon \mathsf{M} X_{n_k} \mathrm{I}_{A\cap\{\nu>n_k\}} \longrightarrow 0$, $(k \longrightarrow \infty)$. $X_\nu \mathrm{I}_{A\cap\{\nu\leqslant n\}} \uparrow X_\nu \mathrm{I}_A$ как функции на Ω, $|X_\nu \mathrm{I}_{A\cap\{\nu\leqslant n\}}| \leqslant |X_\nu \mathrm{I}_A|$, $\mathsf{M}|X_\nu \mathrm{I}_A| \leqslant \mathsf{M}|X_\nu| < \infty$. Поэтому по теореме Лебега о мажорированной сходимости

$$\mathsf{M} X_\nu \mathrm{I}_{A\cap\{\nu\leqslant n\}} \longrightarrow \mathsf{M} X_\nu I_A,$$

откуда немедленно получаем $\mathsf{M} X_0 \mathrm{I}_A = \mathsf{M} X_\nu \mathrm{I}_A$. ■

Теорема 8.9 (Сохранение мартингальности для двух моментов остановки). *Пусть* $\{(X_n, \mathscr{F}_n)\}$ *— мартингал,* ν_1, ν_2 *— моменты остановки относительно* $\{\mathscr{F}_n\}$, $\mathsf{M}|X_{\nu_1}| < \infty, \mathsf{M}|X_{\nu_2}| < \infty$,
$\liminf_{n\longrightarrow\infty} \mathsf{M}|X_n| \mathrm{I}\{\nu_2 > n\} = 0$. *Тогда* $\mathsf{M}\left(X_{\nu_2} \mid \mathscr{F}_{\nu_1}\right)(\omega) = X_{\nu_1}(\omega)$ *для п.в.* ω, *таких что* $\nu_2(\omega) \geqslant \nu_1(\omega)$. *Аналогичное утверждение верно для семимартингалов с заменой равенства на неравенство соответствующего знака.*

□ $\{(X_n, \mathscr{F}_n)\}$ — мартингал на $(\Omega, \mathscr{F}, \mathsf{P})$. Так как время дискретно, то $\nu_1 \in \mathbb{N} \cup \{0\}$. По свойству моментов остановки

$$\{\nu_2 \geqslant \nu_1\} = \bigcup_{m=0}^{\infty} N_m, \quad N_m = \{\nu_1 = m, \nu_2 \geqslant m\} \in \mathscr{F}_m.$$

При $m_1 \neq m_2$ $N_{m_1} \cap N_{m_2} = \varnothing$. Зафиксируем m. Нам достаточно показать, что для данного m для почти всех $\omega \in N_m$ выполняется

$$\mathsf{M}\left(X_{\nu_2} \mid \mathscr{F}_{\nu_1}\right)(\omega) = X_{\nu_1}(\omega).$$

Рассмотрим ограничение вероятностного пространства на множество N_m с условной вероятностной мерой (в случае $\mathsf{P}(N_m) > 0$; остальные N_m отбрасываем) и применим теорему 8.7. Новое вероятностное пространство — $(N_m, \mathscr{F}^{(m)}, \mathsf{P}^{(m)})$, где $\mathscr{F}^{(m)} = \{A \cap N_m \mid A \in \mathscr{F}\}$, $\mathsf{P}^{(m)}\{B\} = \mathsf{P}\{B \mid N_m\}$. Через $\mathsf{M}^{(m)}$ обозначим мат. ожидание по мере $\mathsf{P}^{(m)}$. Новым потоком σ-алгебр объявим $\mathscr{F}_n^{(m)} = \{A \in \mathscr{F}_n\colon A \subset N_m\}, n \geqslant m$. Очевидно, $\mathscr{F}_m^{(m)} \subseteq \mathscr{F}_{m+1}^{(m)} \subseteq \ldots \subseteq \mathscr{F}^{(m)}$.

Теперь проверим, что $\left\{(X_n\Big|_{N_m}, \mathscr{F}_n^{(m)})\right\}$ — мартингал на новом вероятностном пространстве.

$$\mathsf{M}^{(m)}\left(X_{n+1}\Big|_{N_m} \mid \mathscr{F}_n^{(m)}\right)(\omega) = \mathsf{M}\left(X_{n+1} \mid \mathscr{F}_n\right)(\omega) = X_n(\omega) = X_n\Big|_{N_m}, \text{ т.к. } \omega \in N_m \text{ и } \{(X_n, \mathscr{F}_n)\} \text{ — мартингал.}$$

На N_m $\nu_1 = \mathrm{const}$, а ν_2 — момент остановки относительно нового потока σ-алгебр:

$$\Xi_n = \{\nu_2 \leqslant n, \nu_1 = m\} \in \mathscr{F}_n,\ \Xi_n \subseteq N_m \Rightarrow \Xi_n \in \mathscr{F}_n^{(m)}.$$

Так как мы доказываем свойство, выполняющееся почти наверное, то имеет смысл рассматривать только N_m с $\mathsf{P}\{N_m\} > 0$. Отметим, что так как момент остановки может принимать и бесконечное значение, то, вообще говоря, $\cup_m N_m \neq \Omega$. Именно поэтому в следующей строчке мы имеем не равенство, а оценку снизу.

$$\mathsf{M}|X_n|\mathrm{I}\{\nu_2 > n\} \geqslant \sum_{m=0}^{\infty} \mathsf{M}|X_n|\mathrm{I}\{\nu_2 > n\}\,\mathrm{I}\{N_m\} = \sum_{m=0}^{\infty} \mathsf{M}\left(|X_n|\mathrm{I}\{\nu_2 > n\} \mid N_m\right)\mathsf{P}\{N_m\} =$$
$$= \sum_{m=0}^{\infty} \mathsf{M}\left(|X_n|\Big|_{N_m} \mathrm{I}\{\nu_2 > n\} \mid N_m\right)\mathsf{P}\{N_m\}.$$

Покажем теперь, что для каждого m $\liminf_{n\longrightarrow\infty} \mathsf{M}\left(|X_n|\Big|_{N_m} \mathrm{I}\{\nu_2 > n\} \mid N_m\right) = 0$. Действительно, в противном случае, так как все слагаемые в сумме выше (мы не учитываем те m, для которых $\mathsf{P}\{N_m\} = 0$) неотрицательны, мы бы имели $\liminf_{n\longrightarrow\infty} \mathsf{M}|X_n|\mathrm{I}\{\nu_2 > n\} > 0$, что противоречит условию.

Таким образом, для мартингала и момента остановки на новом вероятностном пространстве выполнены условия теоремы 8.7 и тем самым нужное равенство выполняется на каждом из N_m почти наверное. ■

Это рассуждение весьма мудро. Приведём более простое доказательство, которое к тому же не требует теоремы для одного момента остановки.

□ Нам нужно доказать, что для любого $A \in \mathscr{F}_{\nu_1} \cap \{\nu_2 \geqslant \nu_1\}$ $\mathsf{M}X_{\nu_2}\mathrm{I}_A = \mathsf{M}X_{\nu_1}\mathrm{I}_A$. Заметим, что если мы это докажем для событий $A_n \in \mathscr{F}_{\nu_1} \cap \{\nu_2 \geqslant \nu_1\} \cap \{\nu_1 = n\}$, то по теореме Лебега получим требуемое и для A ($|X_{\nu_i}\mathrm{I}_{\cup_1^N A_n}| \leqslant |X_{\nu_i}\mathrm{I}_A| \leqslant |X_{\nu_i}| \in L_1(\Omega)$). По определению σ-алгебры $\mathscr{F}_{\nu_1}$, учитывая, что $\{\nu_2 \geqslant \nu_1\}$ ей принадлежит, имеем, что $A_n \in \mathscr{F}_n$. Тогда

$$\mathsf{M}X_{\nu_1}\mathrm{I}_{A_n} = \mathsf{M}X_n\mathrm{I}_{A_n} = \mathsf{M}X_{\nu_2}\mathrm{I}_{A_n\cap\{\nu_2=n\}} + \mathsf{M}X_n\mathrm{I}_{A_n\cap\{\nu_2>n\}}.$$

Заметим, что $A_n \cap \{\nu_2 > n\} \in \mathscr{F}_n$, поэтому $\mathsf{M}X_n\mathrm{I}_{A_n\cap\{\nu_2>n\}} = \mathsf{M}X_{n+1}\mathrm{I}_{A_n\cap\{\nu_2\geqslant n+1\}}$. Продолжая эту процедуру по индукции, получим, что

$$\mathsf{M}X_n\mathrm{I}_{A_n} = \mathsf{M}X_{\nu_2}\mathrm{I}_{A_n\cap\{n\leqslant\nu_2\leqslant m\}} + \mathsf{M}X_m\mathrm{I}_{A_n\cap\{\nu_2>m\}}.$$

Вот и всё: по теореме Лебега первое слагаемое сходится к $\mathsf{M}X_{\nu_2}\mathrm{I}_{A_n}$ ($A_n \subset \{\nu_2 \geqslant n\}$) и существует подпоследовательность m_k, по которой второе слагаемое стремится к нулю. Следовательно, $\mathsf{M}X_n\mathrm{I}_{A_n} = \mathsf{M}X_{\nu_2}\mathrm{I}_{A_n}$. ■

Пример 5.3.(«Петербургская игра») Пусть $\xi_1, \xi_2, \ldots$ — независимые равновероятные испытания Бернулли, $\mathsf{P}\{\xi_i = \pm 1\} = 1/2$. Положим

$$X_0 = 0, X_1 = \xi_1, \ldots, X_n = \xi_1 + \sum_{k=2}^{n} \xi_k 2^{k-1}\mathrm{I}\{\xi_1 = \ldots = \xi_{k-1} = -1\}.$$

Проверим, что $\{(X_n, \sigma(\xi_1, \ldots, \xi_n))\}$ — мартингал.

$$\mathsf{M}\left(X_{n+1} \mid \sigma(\xi_1, \ldots, \xi_n)\right) =$$
$$= \xi_1 + \sum_{k=2}^{n} 2^{k-1}\xi_k\mathrm{I}\{\xi_1 = \ldots = \xi_{k-1} = -1\} + 2^n\mathsf{M}\left(\xi_{n+1}\mathrm{I}\{\xi_1 = \ldots = \xi_n = -1\} \mid \sigma(\xi_1, \ldots, \xi_n)\right) =$$
$$= X_n + 2^n\mathrm{I}\{\xi_1 = \ldots = \xi_n = -1\}\mathsf{M}\xi_{n+1} = X_n.$$

Здесь мы пользовались тем, что X_n и $\mathrm{I}\{\xi_1 = \ldots = \xi_n = -1\}$ измеримы относительно $\sigma(\xi_1, \ldots, \xi_n)$, а X_{n+1} от неё не зависит. Проверим выполнение условия теоремы.

$$\{\nu > n\} = \{\xi_1 = \ldots = \xi_n = -1\}.$$

На $\{\nu > n\}$ $X_n = -1 - 2 - 4 - \ldots - 2^{n-1} < 0$.

$$X_\nu = -1 - 2 - \ldots - 2^{\nu-2} + 2^{\nu-1} = 1 \Rightarrow \mathsf{M}|X_\nu| < \infty.$$

$$\mathsf{M}|X_n|\mathrm{I}\{\nu > n\} = (1 + 2 + \ldots + 2^{n-1})\mathsf{P}\{\nu > n\} = \frac{2^n - 1}{2^n} \longrightarrow 1 \neq 0.$$

Условие теоремы, таким образом, не выполнено, равно как и заключение: $\mathsf{M}X_\nu = 1 \neq 0 = \mathsf{M}X_0$. Теперь определим собственно игру. Она будет проводиться по турам, каждый тур состоит из этапов, на которых подбрасывается монетка. Ставка начинается с одного рубля и от тура к туру каждый раз увеличивается в два раза. При выпадении решки текущую ставку выплачивает второй игрок, при выпадении орла ставку выплачивает первый игрок и тур заканчивается. Казалось бы, всё честно, но к моменту остановки (выпадению орла) первый игрок теряет в среднем, как мы показали, один рубль.

Модифицируем игру, сделав ставки во всех раундах, начиная с 1001-го, нулевыми. Тогда условия теоремы будут выполнены и $\mathsf{M}X_\nu = 0$, то есть игра вроде бы справедлива. Но от исходной игры она отличается только в случае выпадения решётки 1000 раз подряд, что имеет вероятность $1/2^{1000}$. Справедливости, как видим, придётся ждать долго...

Считаем $X_{-1} = 0, \Delta X_n \rightleftharpoons X_n - X_{n-1}, (n \leqslant 0)$.

Лемма 8.10. *Пусть $\{X_n\}$ измеримы относительно $\{\mathscr{F}_n\}$, ν — момент остановки относительно потока $\{\mathscr{F}_n\}$, $\mathsf{P}\{\nu < \infty\} = 1, \mathsf{M}\sum_{k=0}^{\nu} |\Delta X_k| < \infty$. Тогда $\mathsf{M}|X_\nu| < \infty$ и $M|X_n|\mathrm{I}\{\nu > n\} \longrightarrow 0 \quad (n \longrightarrow \infty)$ (то есть, на самом деле, ещё и выполнены условия теоремы 8.7).*

□

$$Y_\nu \rightleftharpoons \sum_{k=0}^{\nu} |\Delta X_k| \geqslant |\sum_{k=0}^{\nu}(X_k - X_{k-1})| = |X_\nu|.$$

$\mathsf{M}|X_\nu| \leqslant \mathsf{M}Y_\nu < \infty$, поэтому

$$\mathsf{M}|X_n|\mathrm{I}\{\nu > n\} \leqslant \mathsf{M}\sum_{k=0}^{n} |\Delta X_k| \mathrm{I}\{\nu > n\} \leqslant \mathsf{M}Y_\nu \mathrm{I}\{\nu > n\}.$$

Так как $Y_\nu \geqslant 0$ и $Y_\nu \in L_1(\Omega)$, а также $\nu < \infty$ п.н. $\Rightarrow \mathsf{P}\{\nu > n\} \longrightarrow 0\,(n \longrightarrow \infty)$, то $\mathsf{M}Y_\nu \mathrm{I}\{\nu > n\} \longrightarrow 0$, $(n \longrightarrow \infty)$. ■

Теорема 8.11 (достаточные условия сохранения мартингальности). *Пусть $\{(X_n, \mathscr{F}_n)\}$ — мартингал, ν — момент остановки относительно $\{\mathscr{F}_n\}$, $\mathsf{M}\nu < \infty$ и*

$$\exists C < \infty\, \forall n \geqslant 0\ \ \{\omega\colon \nu \geqslant n\} \subseteq \{\omega\colon \mathsf{M}\left(|\Delta X_n| \mid \mathscr{F}_{n-1}\right) \leqslant C\}.$$

Тогда $\mathsf{M}|X_\nu| < \infty$ и $\mathsf{M}X_\nu = \mathsf{M}X_0$.

□ Нужно проверить условие предыдущей леммы, т.е. доказать, что $\mathsf{M}\sum_{k=0}^{\nu} |\Delta X_k| < \infty$.

$$\mathsf{M}\sum_{k=0}^{\nu} |\Delta X_k| = \mathsf{M}\sum_{n\geqslant 0} \mathrm{I}\{\nu = n\} \sum_{k=0}^{n} |\Delta X_k| = \sum_{n\geqslant 0}\sum_{k=0}^{n} \mathsf{M}|\Delta X_k| \mathrm{I}\{\nu = n\} = \sum_{k=0}^{\infty}\sum_{n=k}^{\infty} \mathsf{M}|\Delta X_k|\mathrm{I}\{\nu = n\} =$$
$$= \sum_{k=0}^{\infty} \mathsf{M}|\Delta X_k| \mathrm{I}\{\nu \geqslant k\}.$$

Заметим, что $\{\nu \geqslant k\} = \Omega \setminus \{\nu < k\} \in \mathscr{F}_{k-1}$.

$$\mathsf{M}|\Delta X_k|\mathrm{I}\{\nu \geqslant k\} = \mathsf{M}\mathsf{M}\left(|\Delta X_k|\mathrm{I}\{\nu \geqslant k\} \mid \mathscr{F}_{k-1}\right) = \mathsf{M}\mathrm{I}\{\nu \geqslant k\}\, \mathsf{M}\left(|\Delta X_k| \mid \mathscr{F}_{k-1}\right) \leqslant C\mathsf{P}\{\nu \geqslant k\},$$

$$\mathsf{M}\sum_{k=0}^{\nu} |\Delta X_k| \leqslant \sum_{k=0}^{i} nftyC\mathsf{P}\{\nu \geqslant k\} = C\sum_{k=0}^{\infty}\sum_{n=k}^{\infty} \mathsf{P}\{\nu = n\} = C\sum_{n=0}^{\infty}\sum_{k=0}^{n} \mathsf{P}\{\nu = n\} = C\sum_{n=0}^{\infty}(1+n)\mathsf{P}\{\nu = n\} =$$
$$= C(\mathsf{M}\nu + 1) < \infty.$$

■

8.6. Тождество Вальда

Теорема 8.12 (Тождество Вальда). *Пусть $\xi_1, \ldots, \xi_n$ — независимые случайные величины и $\forall k \quad \mathsf{M}\xi_k = a$, $\mathsf{M}|\xi_k| \leqslant C < \infty$. Положим $S_0 = 0$, $S_n = \xi_1 + \ldots + \xi_n, n \in \mathbb{N}$. Пусть $\nu \geqslant 1$ — момент остановки относительно потока σ-алгебр $\{\sigma(\xi_1, \ldots, \xi_n)\}$ и $\mathsf{M}\nu < \infty$. Тогда*

$$\mathsf{M}S_\nu = a\mathsf{M}\nu.$$

□ Положим $X_n \rightleftharpoons S_n - na = (\xi_1 - a) + \ldots + (\xi_n - a), \mathscr{F}_n \rightleftharpoons \sigma(\xi_1, \ldots, \xi_n)$. Тогда $\{(X_n, \mathscr{F}_n)\}$ — мартингал.

$$\mathsf{M}\left(|X_n - X_{n-1}| \mid \mathscr{F}_{n-1}\right) = \mathsf{M}\left(|S_n - na - (S_{n-1} - (n-1)a)| \mid \mathscr{F}_{n-1}\right) = \mathsf{M}\left(|\xi_n - a| \mid \mathscr{F}_{n-1}\right) =$$
$$= \mathsf{M}|\xi_n - a| \leqslant |a| + \mathsf{M}|\xi_n| \leqslant C + a < \infty.$$

Тем самым условия предыдущей теоремы выполнены и $\mathsf{M}X_\nu = 0 \Leftrightarrow \mathsf{M}\left(S_\nu - \nu a\right) = 0 \Leftrightarrow \mathsf{M}S_\nu = a\mathsf{M}\nu$. ■

Приведём ещё одно доказательство тождества Вальда, следуя Колмогорову и Прохорову.

□ Положим $\chi_i \eqqcolon \mathrm{I}\{\nu \geqslant i\}$, $\{\nu \geqslant i\} \in \mathscr{F}_{i-1} \Rightarrow \xi_i$ и χ_i независимы $\forall i$. Покажем, что ряд

$$S_\nu = \sum_{i=1}^{\nu} \xi_i = \sum_{i=1}^{\infty} \xi_i \chi_i$$

сходится абсолютно.

$$\mathsf{M}|\chi_i \xi_i| = \mathsf{M}\chi_i \mathsf{M}|\xi_i| \leqslant C\mathsf{P}\{\nu \geqslant i\},$$
$$\sum_{i=1}^{\infty} \mathsf{M}|\chi_i \xi_i| \leqslant C \sum_{i=1}^{\infty} \mathsf{P}\{nu \geqslant i\} < \infty,$$

т.к. сумму последнего ряда мы уже считали. Вследствие абсолютной сходимости ряд можно почленно проинтегрировать, получив

$$\mathsf{M}S\nu = \sum_{i=1}^{\infty} \mathsf{M}\xi_i \chi_i = \sum_{i=1}^{\infty} \mathsf{M}\xi_i \mathsf{M}\chi_i = a \sum_{i=1}^{\infty} \mathsf{P}\{\nu \geqslant i\} = a\mathsf{M}\nu.$$

■

Без абсолютной сходимости, вообще говоря, при почленном интегрировании может получиться бред.

Пример 6.1. $\xi_1, \xi_2, \ldots \sim \mathcal{N}(0,1), S_n \eqqcolon \xi_1 + \ldots + \xi_n, \nu = \min\{n \mid S_n < 0\}, S_\nu = \sum_{i=1}^{\infty} \xi_i \chi_i$. $\mathsf{P}\{S_\nu < 0\} = 1 \Rightarrow \mathsf{M}S_\nu < 0$, но $\mathsf{M}\xi_i \chi_i = \mathsf{M}\xi_i \mathsf{M}\chi_i = 0$, поэтому при формальном интегрировании получили бы $\mathsf{M}S_\nu = 0$.

$$0 > \mathsf{M}S_\nu \neq \sum_{i=1}^{\infty} \mathsf{M}\xi_i \chi_i, \mathsf{M}S_\nu = \mathsf{M}\sum_{i=1}^{\infty} \xi_i \chi_i.$$

Это показывает, что математическое ожидание не всегда σ-аддитивно.

8.6.1. Преобразование Лапласа. Фундаментальное тождество Вальда

Определение. Пусть X — неотрицательная случайная величина. Преобразованием Лапласа от X называется функция $\psi_X(\lambda)$, определённая для $\lambda \in \{z \mid Rez > 0\}$ следующим образом:

$$\psi_X \lambda \eqqcolon \mathsf{M}e^{-\lambda X}.$$

Перечислим некоторые его свойства.

1. Если $X_1, \ldots, X_n$ независимы, то $\psi_{X_1+\ldots+X_n}(\lambda) = \prod_{i=1}^n \psi_{X_i}(\lambda)$.
2. $\psi_X(0) = 1$.
3. $|\psi_X(\lambda)| \leqslant \psi_X(0) = 1$ в правой полуплоскости.
4. $\psi_X^{(n)}(0) = (-1)^n \mathsf{M}X^n$, если $\mathsf{M}|X^n| < \infty$.
5. Для случайного блуждания ($S_n = \xi_1 + \ldots + \xi_n, \xi_j$ независимы и одинаково распределены) $\psi_{S_n}(\lambda) = (\psi_{\xi_1}(\lambda))^n$.
6. $\psi_\xi''(z) = \mathsf{M}\xi^2 e^{-z\xi} \geqslant 0$ при $z \in \mathbb{R}$, т.е. преобразование Лапласа выпукло вниз.

Теорема 8.13 (Фундаментальное тождество Вальда). *Пусть $\xi, \xi_1, \ldots$ — независимые одинаково распределённые случайные величины, $S_0 = 0$, $S_n = \xi_1 + \ldots + \xi_n, n \geqslant 1$, $\psi(\lambda) = \mathsf{M}e^{-\lambda\xi} < \infty$, ν — момент остановки относительно $\mathscr{F}_n \eqqcolon \{\sigma(\xi_1, \ldots, \xi_n)\}$, $\mathsf{M}\nu < \infty$ и $\exists C < \infty\, \forall n \geqslant 0$ $\{\nu \geqslant n\} \subseteq \{|S_n| \leqslant C\}$, то есть момент остановки не наступает, пока случайное блуждание S_n не выйдет из полосы $\{|x| \leqslant C\}$. Тогда*

$$\forall \lambda\colon \psi(\lambda) \geqslant 1 \Rightarrow \mathsf{M}\left(\frac{e^{-\lambda S_\nu}}{\psi^\nu(\lambda)}\right) = 1.$$

□ Введём вспомогательный мартингал X_n: положим $X_0 \eqqcolon 1$, $X_n \eqqcolon \dfrac{e^{-zS_n}}{\psi^n(z)}, n \geqslant 1$. Очевидно, что X_n измеримо относительно $\mathscr{F}_n$, и что $\mathsf{P}\{X_n \geqslant 0\} = 1$[6].

$$\mathsf{M}X_n = \frac{\mathsf{M}e^{-zS_n}}{\psi^n(z)} = \frac{\psi^n(z)}{\psi^n(z)} = 1,$$

[6] Мы здесь таки считаем z и ξ_i действительными, а z ещё и неотрицательным

$$X_{n+1} = \frac{e^{-zS_{n+1}}}{\psi^{n+1}(z)} = X_n \frac{e^{-z\xi_{n+1}}}{\psi(z)},$$

$$\mathsf{M}\left(X_{n+1} \mid \mathscr{F}_n\right) = \mathsf{M}\left(X_n \frac{e^{-z\xi_{n+1}}}{\psi(z)} \mid \mathscr{F}_n\right),$$

что в силу того, что X_n является $\mathscr{F}_n$-измеримой, равно

$$X_n \frac{\mathsf{M}e^{-z\xi_{n+1}}}{\psi(z)} = X_n,$$

то есть $\{(X_n, \mathscr{F}_n)\}$ — действительно мартингал.

Покажем, что существует C': $\mathsf{M}\left(|X_n - X_{n-1}| \mid \mathscr{F}_{n-1}\right) \leqslant C'$ на $\{\nu \geqslant n\}$ для всех n (тогда можно будет применить теорему 8.11 о сохранении мартингального свойства).

$$\mathsf{M}\left(|X_n - X_{n-1}| \mid \mathscr{F}_{n-1}\right) = \mathsf{M}\left(X_{n-1} |\frac{e^{-z\xi_n}}{\psi(z)} - 1| \mid \mathscr{F}_{n-1}\right) = X_{n-1}\mathsf{M}|\frac{e^{-z\xi_n}}{\psi(z)} - 1| \leqslant X_{n-1}\mathsf{M}\left(1 + \frac{e^{-z\xi_n}}{\psi(z)}\right) \leqslant$$
$$\leqslant 2X_{n-1},$$

т.к. по условию $\psi(z) \leqslant 1$, а $X_{n-1} \geqslant 0$. На множестве $\{\omega \colon \nu \geqslant n\} \in \mathscr{F}_{n-1}$ имеем по условию нашей теоремы

$$\mathsf{M}\left(|X_n - X_{n-1}| \mid \mathscr{F}_{n-1}\right) \leqslant 2X_{n-1} \leqslant 2\frac{e^{zC}}{\psi^{n-1}(z)} \leqslant 2e^{zC} \eqqcolon C' < \infty.$$

Применяя к мартингалу X_n теорему 8.11, получаем $\mathsf{M}X_\nu = \mathsf{M}X_0 = 1$, что и требовалось. ■

8.7. Применения тождества Вальда

8.7.1. Теорема восстановления

Пусть $\xi_1, \xi_2, \ldots$ — независимые одинаково распределённые неотрицательные случайные величины, $\mathsf{M}\xi_1 = a < \infty$. Определим по ним случайное блуждание S_n, положив $S_0 = 0$, $S_n = \xi_1 + \ldots + \xi_n$. $\nu(t) = \min\{n \colon S_n > t\} = \sum_{k=0}^{\infty} \mathrm{I}\{S_k \leqslant t\}$ — момент остановки (относительно $\{\sigma(\xi_1, \ldots, \xi_n)\}$).

Определение. *Процессом восстановления* называется процесс $\{\nu(t)\}_t$. *Функцией восстановления* называется функция $U(t) = \mathsf{M}\nu(t)$.

$U(t)$ не убывает, следовательно, можно считать, что она задаёт меру на $[0, +\infty)$: $U([a,b]) \eqqcolon U(b) - U(a)$ (доопределим в точках разрыва U до непрерывной слева, так всегда можно сделать).

Теорема 8.14. *Если $U(t)$ — функция восстановления, построенная по неотрицательным независимым одинаково распределённым случайным величинам $\xi_1, \xi_2, \ldots$, $\psi_{\xi_1}(z) = \psi(z) = \mathsf{M}e^{-z\xi_1} = \int_{\mathbb{R}} e^{-zx} dF(x)$, где F — функция распределения ξ_1, то преобразование Лапласа меры U, равное по определению $\int\limits_0^\infty e^{-zt} dU(t)$, равно $\frac{1}{1-\psi(z)}$ в тех точках, где оно определено. Если при этом $\xi_1, \ldots$ — целочисленные с производящей функцией $f(s) = \mathsf{M}s^{\xi_1}$, $u_n = U(n) - U(n-1)$, то производящая функция последовательности u_n $\sum_{n=0}^\infty s^n u_n = \frac{1}{1-f(s)}$.*

□ $\nu(t) = \sum_{n=0}^\infty \mathrm{I}\{S_n \leqslant t\} \Rightarrow U(t) = \sum_{n=0}^\infty \mathsf{P}\{S_n \leqslant t\}$. Перейдём к преобразованию Лапласа:

$$\int_0^\infty e^{-zt} dU(t) = \sum_{n=0}^\infty \int_0^\infty e^{-zt} dP\{S_n \leqslant t\} = \sum_{n=0}^\infty \mathsf{M}e^{-zS_n} = \sum_{n=0}^\infty \psi^n(z) = \frac{1}{1-\psi(z)}.$$

Второе утверждение доказывается аналогично. ■

Теорема 8.15. *При тех же условиях, $\mathsf{M}\xi_1 = a > 0, a < \infty$, существует*

$$\lim_{t \longrightarrow \infty} \frac{U(t)}{t} = \frac{1}{a}.$$

□ $\nu(t)$ — момент остановки, $\mathsf{M}\left(|S_n - S_{n-1}| \mid \mathscr{F}_{n-1}\right) = \mathsf{M}|\xi_n| = \mathsf{M}\xi_n = a < \infty$. $\psi(z) \eqqcolon \mathsf{M}e^{-z\xi_1}$. Покажем, что $\mathsf{M}\nu(t) < \infty$: для каждого t $\mathrm{I}\{x \leqslant t\} \leqslant e^{-z(x-t)}, z \geqslant 0$.

$$\mathsf{P}\{\nu(t) > k\} = \mathsf{P}\{S_k \leqslant t\} = \mathsf{M}\mathrm{I}\{S_k \leqslant t\} \leqslant \mathsf{M}e^{-z(S_k - t)} = \psi^k(z)e^{zt},$$

следовательно,

$$\mathsf{M}\nu(t) = \sum_{k=0}^\infty \mathsf{P}\{\nu(t) > k\} \leqslant \sum_{k=0}^\infty e^{zt}\psi^k(z) = \frac{e^{zt}}{1-\psi(z)} < \infty, \quad \text{т.к. при } z > 0\, |\psi(z)| < 1.$$

Поэтому применимо тождество Вальда, из которого получаем

$$\mathsf{M}S_{\nu(t)} = a\mathsf{M}\nu(t) = aU(t).$$

По определению $S_{\nu(t)} > t$, следовательно, $aU(t) > t$, откуда

$$\liminf_{t\longrightarrow\infty} \frac{U(t)}{t} \geqslant \frac{1}{a}.$$

$S_{\nu(t)-1} \leqslant t$, но $\nu(t)-1$, вообще говоря, не является моментом остановки относительно $\{\sigma(\xi_1,\ldots,\xi_n)\}$. Если бы наши случайные величины ξ_i были бы ограничены, т.е. существовало бы w: $\mathsf{P}\{\xi_1 \leqslant w\} = 1$, то было бы верно $\mathsf{P}\left\{S_{\nu(t)} = S_{\nu(t)-1} + \xi_{\nu(t)} \leqslant t + w\right\} = 1$. Почему бы так, собственно, и не сделать?

Введём $\forall w$ случайную величину $\xi_k(w) \rightleftharpoons \min\{\xi_k, w\}$[7], разумеется, $\xi_k(w) \leqslant w$ п.н., $S_n(w) \rightleftharpoons \xi_1(w) + \ldots + \xi_k(w) \leqslant S_n, \nu(t,w) \rightleftharpoons \min\{n\colon S_n(w) > t\} \geqslant \nu(t), U(t,w) \rightleftharpoons \mathsf{M}\nu(t,w) \geqslant U(t)$.

Применим к $S_n(w)$ и $\nu(t,w)$ тождество Вальда[8]:

$$\mathsf{M}S_{\nu(t,w)}(w) = \mathsf{M}\xi_1(w)U(t,w),$$

$S_{\nu(t,w)}(w) = S_{\nu(t,w)-1}(w) + \xi_{\nu(t,w)}(w) \leqslant t + w$ п.н., следовательно, $U(t) \leqslant U(t,w) = \mathsf{M}\left(\frac{S_{\nu(t,w)}(w)}{\mathsf{M}\xi_1(w)}\right) \leqslant \frac{t+w}{\mathsf{M}\xi_1(w)} \quad \forall t, w$. Поэтому

$$\limsup_{t\longrightarrow\infty} \frac{U(t)}{t} \leqslant \frac{1}{\mathsf{M}\xi_1(w)} \lim_{t\longrightarrow\infty} \frac{t+w}{t} = \frac{1}{\mathsf{M}\min\{\xi_1, w\}} \quad \forall w.$$

Но так как $\mathsf{M}\min\{\xi_1, w\} \longrightarrow \mathsf{M}\xi_1 = a > 0 \quad (w \longrightarrow \infty)$, то $\frac{1}{\mathsf{M}\xi_1(w)} \longrightarrow \frac{1}{\mathsf{M}\xi_1} = \frac{1}{a}$, следовательно, $\lim_{t\longrightarrow\infty} \frac{U(t)}{t} = \frac{1}{a}$. ■

8.7.2. Неравенства Дуба и Колмогорова

Пусть $\{X_n, \mathscr{F}_n\}$ — субмартингал, т.е. X_n измерима относительно $\mathscr{F}_n$, $\mathsf{M}|X_n| < \infty$, $\mathsf{M}(X_{n+1} \mid \mathscr{F}_n) \geqslant X_n \forall n$.

Лемма 8.16 (Неравенство Маркова). *Пусть $\xi \geqslant 0, \mathsf{M}\xi < \infty$. Тогда $\forall x > 0$ $\mathsf{P}\{\xi \geqslant x\} \leqslant \frac{\mathsf{M}\xi}{x}$.*

□ $\mathsf{P}\{\xi \geqslant x\} = \mathsf{M}\mathrm{I}\{\xi \geqslant x\} \leqslant \mathsf{M}\mathrm{I}\{\xi \geqslant x\}\frac{\xi}{x} = \frac{1}{x}\mathsf{M}\xi\mathrm{I}\{\xi \geqslant x\} \leqslant \frac{\mathsf{M}\xi}{x}$. ■

Теорема 8.17 (Doob). *Пусть $\{X_n, \mathscr{F}_n\}$ — неотрицательный субмартингал. Тогда* [9]

$$\forall x > 0, n > 0\ \mathsf{P}\left\{\max_{1\leqslant k\leqslant n} X_k \geqslant x\right\} \leqslant \frac{1}{x}\mathsf{M}X_n.$$

□ Введём моменты остановки

$$\nu \rightleftharpoons \min\{k \geqslant 0\colon X_k \geqslant x\}, \nu(n) \rightleftharpoons \min\{\nu, n\} \leqslant n \Rightarrow \liminf_{m\longrightarrow\infty} \mathsf{M}|X_m|\mathrm{I}\{\nu(n) \geqslant m\} = 0.$$

Положим $\nu_1 = \nu(n), \nu_2 \equiv n$. По теореме 8.9 имеем

$$\mathsf{M}X_{\nu_2} = \mathsf{M}X_n \geqslant \mathsf{M}X_{\nu_1} = \mathsf{M}X_{\nu(n)}.$$

Заметим, что

$$\left\{\max_{1\leqslant k\leqslant n} X_k \geqslant x\right\} = \left\{X_{\nu(n)} \geqslant x\right\},$$

следовательно,

$$\mathsf{P}\left\{\max_{1\leqslant k\leqslant n} X_k \geqslant x\right\} = \mathsf{P}\left\{X_{\nu(n)} \geqslant x\right\} \leqslant \frac{\mathsf{M}X_{\nu(n)}}{x} \leqslant \frac{\mathsf{M}X_n}{x},$$

где предпоследнее неравенство выполнено в силу неравенства Маркова. ■

Следствие 8.1 (Неравенство Колмогорова). *Пусть $\{X_n, \mathscr{F}_n\}$ — мартингал, $\mathsf{M}X_n^2 < \infty$. Тогда $\{X_n^2, \mathscr{F}_n\}$ — субмартингал и*

$$\mathsf{P}\left\{\max_{1\leqslant k\leqslant n} |X_k| \geqslant x\right\} \leqslant \frac{\mathsf{M}X_n^2}{x^2}.$$

□ Выпуклая вниз измеримая функция от мартингала есть субмартингал. ■

[7] w называется *уровнем усечения*.

[8] Контрольный вопрос: а почему это $\mathsf{M}\nu(t,w) < \infty$? А вдруг равно?

[9] Это называется неравенством Дуба.

8.7.3. Теорема Дуба о среднем числе пересечений полосы мартингалом

Пусть есть случайный процесс X_t и два уровня $-\infty < a < b < +\infty$, и процесс, для простоты, стартует «между» ними, т.е. $X_0 \in (a, b)$. Положим $\nu_0 \rightleftharpoons 0$, $\nu_{2k-1} \rightleftharpoons \min\{t > \nu_{2k-2} \colon X_t \leqslant a\}$, $\nu_{2k} \rightleftharpoons \min\{t > \nu_{2k-1} \colon X_t \geqslant b\}$. Определим

$$\beta_n(a,b) \rightleftharpoons \begin{cases} 0, & \nu_2 > n, \\ \max\{k \colon \nu_{2k} \leqslant n\}, & \nu_2 \leqslant n. \end{cases}$$

Говоря неформально, $\beta_n(a,b)$ — это количество полных пересечений процессом X_t полосы (a,b) «снизу вверх» на промежутке $(0,n)$.

Обозначим через x_+ $\max\{0,x\}$.

Теорема 8.18 (Неравенство Дуба). *Если $(X_n, \mathscr{F}_n)$ — субмартингал, то*

$$\forall n, \forall -\infty < a < b < +\infty \colon \mathsf{M}\beta_n(a,b) \leqslant \frac{\mathsf{M}(X_n - a)_+}{b-a}.$$

□ $f(x) = (x-a)_+$ — выпуклая вниз неубывающая функция, следовательно, $((X_n - a)_+, \mathscr{F}_n)$ — субмартингал.

Далее считаем $X_n \geqslant 0$ (заменяем $(X_n - a)_+$ на X_n, вместо $[a,b]$ берём $[0, b-a] = [0,c]$). Нужно доказать, что $\mathsf{M}\beta_n(0,c) \leqslant \frac{\mathsf{M}X_n}{c}$. Введём случайные величины η_t следующим образом:

$$\eta_t = \begin{cases} 1, & t \in \bigcup_{k\geqslant 1}(\nu_{2k-1}, \nu_{2k}] \\ 0, & t \in \bigcup_{k\geqslant 0}(\nu_{2k}, \nu_{2k+1}], \end{cases}$$

т.е. $\eta_t = 1$ тогда и только тогда, когда мы идём от нечётного ν к чётному (включая чётный конец интервала). η_t определяется чётностью $\max\{k \colon \nu_k < t\}$, следовательно, выражается через $X_1, \ldots, X_{t-1}$, а поэтому измерима относительно $\mathscr{F}_{t-1}$.

$$\eta_t - \eta_{t+1} \neq 0 \Leftrightarrow t \in \{\nu_1, \nu_2, \ldots\},$$

$$\eta_t - \eta_{t+1} = \begin{cases} -1, & t \in \{\nu_{2k-1} \colon k > 0\} = \{X_t = 0\} \\ 1, & t \in \{\nu_{2k} \colon k \geqslant 1\} = \{X_t \geqslant c\}. \end{cases}$$

$$c\mathrm{I}\{\exists k \colon t = \nu_{2k}\} \leqslant (\eta_t - \eta_{t+1})X_t,$$

т.к. на тех ω, где индикатор слева равен 1, $\eta_t - \eta_{t+1} = 1$, а $X_t \geqslant c$. Поэтому

$$c\beta_n(0,c) = c\max\{k \colon \nu_{2k} \leqslant n\} = \sum_{t=0}^{n} c\mathrm{I}\{\exists k \colon t = \nu_{2k}\} \leqslant \sum_{n=0}^{n}(\eta_t - \eta_{t+1})X_t = \sum_{t=1}^{n}\eta_t(X_t - X_{t-1}) - \eta_{n+1}X_n.$$

Учитывая, что $\eta_{n+1}, X_n \geqslant 0$, получаем

$$c\mathsf{M}\beta_n(0,c) \leqslant \mathsf{M}\sum_{t=1}^{n}\eta_t(X_t - X_{t-1}) = \sum_{t=1}^{n}\mathsf{M}\left(\eta_t \mathsf{M}\left(X_t - X_{t-1} \mid \mathscr{F}_{t-1}\right)\right) = \sum_{t=1}^{n}\mathsf{M}\left(\eta_t \mathsf{M}\left(X_t \mid \mathscr{F}_{t-1}\right) - X_{t-1}\right),$$

что в силу того, что $\eta_t \in \{0,1\}$ и $\mathsf{M}(X_t \mid \mathscr{F}_{t-1}) - X_{t-1} \geqslant 0$ даёт

$$c\mathsf{M}\beta_n(0,c) \leqslant \sum_{t=1}^{n}\mathsf{M}\left(\mathsf{M}(X_t \mid \mathscr{F}_{t-1}) - X_{t-1}\right) = \sum_{t=1}^{n}(\mathsf{M}X_t - \mathsf{M}X_{t-1}) = \mathsf{M}X_n - \mathsf{M}X_0 \leqslant \mathsf{M}X_n.$$

■

8.7.4. Теорема Дуба о сходимости субмартингалов

Теорема 8.19 (Дуба о сходимости субмартингалов). *Пусть $(X_n, \mathscr{F}_n)$ — субмартингал, $\sup_n \mathsf{M}|X_n| < \infty$. Тогда существует такая случайная величина X, что $\mathsf{P}\{\lim_{n\longrightarrow\infty} X_n = X\} = 1, \mathsf{M}|X| < \infty$.*

□ Рассуждаем от противного. Пусть с положительной вероятностью предела не существует, т.е. для случайных величин $\xi_* \rightleftharpoons \liminf_{n\longrightarrow\infty} X_n$, $\xi^* \rightleftharpoons \limsup_{n\longrightarrow\infty} X_n$ выполнено

$$\mathsf{P}\{\xi^* > \xi_*\} > 0.$$

Заметим, что $\{\omega \colon \xi^* > \xi_*\} = \cup_{a,b\in\mathbb{Q}}\{\omega \colon \xi_* < a < b < \xi^*\}$, а посему в силу σ-аддитивности вероятностной меры

$$0 < \mathsf{P}\{\xi^* > \xi_*\} \leqslant \sum_{a,b\in\mathbb{Q}} \mathsf{P}\{\xi_* < a < b < \xi^*\},$$

следовательно, $\exists a, b \in \mathbb{Q}\colon \mathsf{P}\{\xi_* < a < b < \xi^*\} > 0$, то есть X_n с положительной вероятностью бесконечное число раз пересекает полосу (a, b). Обозначим $\beta^*(a,b) \rightleftharpoons \sup_n \beta_n(a,b)$. Ясно, что $\beta_n(a,b) \uparrow \beta^*(a,b), n \longrightarrow \infty$. Имеем

$$\forall k < \infty\colon\ \mathsf{P}\{\beta_n(a,b) \geqslant k \mid \xi^* > b > a > \xi_*\} \longrightarrow 1 \quad (n \longrightarrow \infty).$$

Поэтому $\mathsf{P}\{\beta^*(a,b) = \infty\} = \mathsf{P}\{\xi^* > b > a > \xi_*\} > 0$, значит, $\mathsf{M}\beta^*(a,b) = +\infty$.

С другой стороны, по неравенству Дуба

$$\mathsf{M}\beta_n(a,b) \leqslant \frac{\mathsf{M}\,(X_n - a)_+}{b-a} = \frac{\mathsf{M}\,(X_n) + |a|}{b-a} \leqslant \frac{\sup_n \mathsf{M}|X_n| + |a|}{b-a} = c < \infty,$$

откуда по теореме Б. Леви $\mathsf{M}\beta^*(a,b) \leqslant c < \infty$, что противоречит полученному выше. Таким образом, $\xi^* \overset{\text{п.н.}}{=} \xi_*$. Осталось заметить, что $\mathsf{M}|X| < \infty$ по теореме Фату. ■

Следствие 8.2. *Если $X_n \leqslant 0$ — субмартингал, то* $\mathsf{P}\{\exists \lim_{n\longrightarrow\infty} X_n = X\} = 1$

□ $\mathsf{M}|X_n| \downarrow, n \longrightarrow \infty \Rightarrow \sup_n \mathsf{M}|X_n| < \infty$. ■

Следствие 8.3. *Если $(X_n, \mathscr{F}_n)$ — неотрицательный мартингал, то* $\mathsf{P}\{\exists \lim_{n\longrightarrow\infty} X_n = X\} = 1$.

□ $\mathsf{M}|X_n| = \mathsf{M}X_n = \mathsf{M}X_0 < \infty$. ■

Пример 7.1. Пусть $\mu(t)$ — ветвящийся процесс, A — мат. ожидание числа потомков, тогда $X_n = \frac{\mu(n)}{A^n}$ образуют мартингал, $X_n \geqslant 0$, причём сходимость есть даже при $A < 1$!

Следствие 8.4 (Теорема Колмогорова о сходимости рядов). *Пусть $\xi_1, \ldots$ — независимые случайные величины, $\mathsf{M}\xi_k = 0$, $\mathsf{D}\xi_k = \sigma_k^2 < \infty$, $\sum_{k=1}^{\infty} \sigma_k^2 = \sigma^2 < \infty$. Тогда существует случайная величина S, к которой частичные суммы $S_n = \xi_1 + \ldots + \xi_n$ сходятся почти наверное.*

□ S_n — мартингал. По неравенству Йенсена $\mathsf{M}|S_n| \leqslant \sqrt{\mathsf{M}S_n^2} = \sqrt{\mathsf{D}S_n} = \sqrt{\sum_1^n \sigma_i^2} \leqslant \sigma < \infty$. Применяем теорему Дуба. ■

9. Процесс броуновского движения

Определение. Процессом броуновского движения (стандартным винеровским процессом) называется случайный процесс $W(t), t \geqslant 0$, обладающий следующими свойствами:

1. $\mathsf{P}\{W(0) = 0\} = 1$.
2. $W(t)$ — процесс с независимыми приращениями.
3. $\forall 0 \leqslant s \leqslant t\colon\ W(t) - W(s) \sim \mathcal{N}(0, t-s)$.

Лемма 9.1. *Совместное распределение значений $(W(t_1), \ldots, W(t_n))$, $0 = t_0 < \ldots < t_n$ является многомерным нормальным с ковариационной матрицей $\Sigma = (\min(t_i, t_j))$.*

□ По определению процесса $W(t)$ плотность распределения вектора $(W(t_1) - W(t_0), \ldots, W(t_n) - W(t_{n-1}))$ равна

$$p_{t_1,\ldots,t_n}(y_1, \ldots, y_n) = \frac{1}{\prod_{k=1}^n \sqrt{2\pi(t_k - t_{k-1})}} \prod_{k=1}^n e^{-\frac{y_k^2}{2(t_k - t_{k-1})}}.$$

Из теории меры и интеграла хорошо известен следующий

Факт 1. *Если вектор $\xi = (\xi_1, \ldots, \xi_n)$ имеет плотность распределения $\rho(x)$, а $H\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ — невырожденный в каждой точке диффеоморфизм, то вектор $\zeta = H(\xi)$ имеет плотность распределения $q(x) = \frac{\rho(H^{-1}(x))}{|J_H(H^{-1}(x))|}$.*

Рассмотрим линейное преобразование $\mathbb{R}^n$, переводящее $(y_1, \ldots, y_n)$ в $(y_1, y_1 + y_2, \ldots, y_1 + \ldots + y_n)$. Как видно, его якобиан равен 1, и поэтому, применяя наш факт, получаем то, что нужно (проведите выкладку с подстановкой H^{-1} самостоятельно).

Пусть $s < t$. Найдём $\operatorname{cov}(W(s), W(t)) = \operatorname{cov}(W(s), W(s) + W(t) - W(s)) = \operatorname{cov}(W(s), W(s)) + \operatorname{cov}(W(s), W(t) - W(s)) = \mathsf{D}W(s) = s = \min(s, t)$. ■

9.1. Теорема Колмогорова о непрерывной модификации

Определение. Случайные процессы $\xi(t)$ и $\eta(t)$ называются *стохастически эквивалентными*, если

$$\forall t\colon \mathsf{P}\{\xi(t) = \eta(t)\} = 1.$$

$\xi(t)$ называется модификацией $\eta(t)$ (и наоборот).

Из определения видно, что если $\xi(t)$ и $\eta(t)$ стохастически эквивалентны, то $\mathsf{P}\{\xi(t_k) = \eta(t_k), k = 1, \ldots, n\} = 1$.

Пример 1.1. Пусть $\xi(t) \equiv 0, t \in [0,1], r(t) = \mathrm{I}_{\mathbb{Q}}(t), \gamma \sim \mathcal{R}[0,1]$. Положим $\eta(t) = r(t+\gamma)$. Тогда $\mathsf{P}\{\xi(t) \neq \eta(t)\} = \mathsf{P}\{\gamma + t \in Q\} = 0$. Таким образом, $\xi(t)$ и $\eta(t)$ стохастически эквивалентны, но при этом $\mathsf{P}\left\{\sup_{[0,1]} \xi(t) = 0\right\} = 1, \mathsf{P}\left\{\sup_{[0,1]} \eta(t) = 1\right\} = 1$.

Утверждение 9.2. *Пусть $\xi(t)$, $\eta(t), t \in [0,1]$ — два случайных процесса на $(\Omega, \mathscr{F}, \mathsf{P})$. Если существует счетное детерминированное $A \subset [0,1]$ и функционал $F\colon \mathbb{R}^A \longrightarrow \mathbb{R}^{[0,1]}$, такой, что $\mathsf{P}\{\xi(t) = \eta(t)\} = 1 \quad \forall t \in A$ и $\mathsf{P}\{\xi(t) = F(\xi(x), x \in A)(t)\ \forall x \in [0,1]\} = \mathsf{P}\{\eta(t) = F(\eta(x), x \in A)(t)\ \forall x \in [0,1]\} = 1$, то $\mathsf{P}\{\xi(t) = \eta(t)\ \forall t \in [0,1]\} = 1$.*

□ Очевидно. ■

Теорема 9.3 (Колмогоров). *Пусть $\xi(t)$ — случайный процесс, $t \in [0,1]$. Если существуют такие $b > a > 0, c < \infty$, что*

$$\forall t, t+h \in [0,1] \quad \mathsf{M}|\xi(t+h) - \xi(t)|^a < \frac{c|h|}{|\log|h||^{1+b}},$$

то $\xi(t)$ имеет на $[0,1]$ непрерывную модификацию.

□ Обозначим $\Delta_h \xi(t) \rightleftharpoons \xi(t+h) - \xi(t), \varepsilon(h) \rightleftharpoons \frac{1}{|\log|h||^\beta}, 1 < \beta < \frac{b}{a}$. Имеем по неравенству Маркова

$$\mathsf{P}\{|\Delta_h \xi(t)| > \varepsilon(h)\} \leqslant \frac{\mathsf{M}|\Delta_h \xi(t)|^a}{\varepsilon^a(h)} \leqslant \frac{c|h|}{|\log|h||^{1+b-a\beta}} = \frac{c|h|}{|\log|h||^{1+\delta}} \eqqcolon q(h),$$

где $\delta = b - a\beta > 0$. Как видно, $\varepsilon(h) \downarrow 0, q(h) \downarrow 0, (h \longrightarrow 0)$.

$$\sum_{n=1}^{\infty} \varepsilon(2^{-n}) = \sum_{n=1}^{\infty} \frac{c^*}{n^\beta} < \infty, \sum_{n=1}^{\infty} 2^n q(2^{-n}) = \sum_{n=1}^{\infty} \frac{c^{**}}{n^{1+\delta}} < \infty.$$

Положим $\xi_n(t) = \xi\left(\frac{r}{2^n}\right) + 2^n\left(t - \frac{r}{2^n}\right)\Delta_{\frac{1}{2^n}}\xi\left(\frac{r}{2^n}\right)$ для $\frac{r}{2^n} \leqslant t \leqslant \frac{r+1}{2^n}$. Докажем несколько вспомогательных лемм.

Лемма 9.4. *Последовательность $\xi_n(t)$ при $n \longrightarrow \infty$ почти наверное сходится к случайному процессу $\eta(t)$, траектории которого непрерывны почти наверное.*

□

$$\begin{aligned}
\Big|\xi_{n+1}(t) - \xi_n(t)\Big| &\leqslant \left|\xi\left(\frac{2r+1}{2^{n+1}}\right) - \frac{1}{2}\left(\xi\left(\left(\frac{r}{2^n}\right)\right) + \xi\left(\frac{r+1}{2^n}\right)\right)\right| = \\
&= \frac{1}{2}\left|\left(\xi\left(\frac{2r+1}{2^{n+1}}\right) - \xi\left(\frac{r}{2^n}\right)\right) + \left(\xi\left(\frac{2r+1}{2^{n+1}}\right) - \xi\left(\frac{r+1}{2^n}\right)\right)\right| \leqslant \frac{1}{2}\left(\left|\Delta_{\frac{1}{2^{n+1}}}\xi\left(\frac{r}{2^n}\right)\right| + \left|\Delta_{\frac{1}{2^{n+1}}}\xi\left(\frac{2r+1}{2^{n+1}}\right)\right|\right).
\end{aligned}$$

$\mathsf{P}_{n,r} \rightleftharpoons \mathsf{P}\left\{\sup_{\frac{r}{2^n} \leqslant t \leqslant \frac{2r+1}{2^{n+1}}} |\xi_{n+1}(t) - \xi_n(t)| > \varepsilon\left(\frac{1}{2^{n+1}}\right)\right\} \leqslant \mathsf{P}\left\{\frac{1}{2}\left(\left|\Delta_{\frac{1}{2^{n+1}}}\xi\left(\frac{r}{2^n}\right)\right| + \left|\Delta_{\frac{1}{2^{n+1}}}\xi\left(\frac{2r+1}{2^{n+1}}\right)\right|\right) > \varepsilon\left(\frac{1}{2^{n+1}}\right)\right\} \leqslant 2q\left(\frac{1}{2^{n+1}}\right)$. Поэтому

$$\mathsf{P}\left\{\sup_{[0,1]} |\xi_{n+1}(t) - \xi_n(t)| > \varepsilon\left(\frac{1}{2^{n+1}}\right)\right\} \leqslant \sum_{r=0}^{2^n-1} \mathsf{P}_{n,r} \leqslant 2^{n+1} q\left(\frac{1}{2^{n+1}}\right).$$

Т.к. $\sum_n 2^{n+1} q\left(\frac{1}{2^{n+1}}\right) < \infty$, то по лемме Бореля–Кантелли для почти всех $\omega \in \Omega$ найдётся

$$n(\omega)\colon \forall n > n(\omega) \sup_{[0,1]} |\xi_{n+1}(t) - \xi_n(t)| \leqslant \varepsilon\left(\frac{1}{2^{n+1}}\right).$$

С другой стороны, т.к. $\sum_n \varepsilon\left(\frac{1}{2^{n+1}}\right) < \infty$, то для $m > n$ $\sup_{[0,1]} |\xi_n(t) - \xi_m(t)| \leqslant \varepsilon_n = \sum_n^\infty \varepsilon(2^{-k}) \longrightarrow 0 \quad (n \longrightarrow \infty)$. Поэтому $\{\xi_n(t)\}$ почти наверное фундаментальна, стало быть, почти наверное $\exists \lim_{n\longrightarrow\infty} \xi_n(t) = \eta(t)$, причём эта сходимость почти наверное равномерная ($\sup_{[0,1]} |\xi_n(t) - \eta(t)| \leqslant \varepsilon_n \forall n > n(\omega)$), поэтому траектории $\eta(t)$ почти наверное непрерывны, как равномерные пределы непрерывных (на самом деле кусочно-линейных) траекторий $\xi_n(t)$. ■

Лемма 9.5. *Процессы $\eta(t)$ и $\xi(t)$ стохастически эквивалентны.*

□ Если t — двоично-рациональная точка, то при достаточно больших m $\eta(t) = \xi_m(t) = \xi(t)$ (по построению). В противном случае положим $r_n = [t2^n]$ — номер интервала n-го разбиения, в который попала точка t. Очевидно, что $0 < t - \dfrac{r_n}{2^n} < \dfrac{1}{2^n}$, откуда $\dfrac{r_n}{2^n} \longrightarrow t$ $(n \longrightarrow \infty)$. Далее (используя монотонность ε и q),

$$\mathsf{P}\left\{\left|\xi\left(\frac{r_n}{2^n}\right) - \xi(t)\right| > \varepsilon\left(\frac{1}{2^n}\right)\right\} \leqslant \mathsf{P}\left\{\left|\xi\left(\frac{r_n}{2^n}\right) - \xi(t)\right| > \varepsilon\left(t - \frac{r_n}{2^n}\right)\right\} \leqslant q\left(t - \frac{r_n}{2^n}\right) \leqslant q\left(\frac{1}{2^n}\right).$$

Ряд $\sum q\left(\frac{1}{2^n}\right)$ сходится. Применяем лемму Бореля–Кантелли: выполняется лишь конечное число из этих событий, поэтому

$$\mathsf{P}\left\{\xi\left(\frac{r_n}{2^n}\right) \xrightarrow[n\to\infty]{} \xi(t)\right\} = 1.$$

Но для η аналогичное соотношение выполняется в силу непрерывности η п.н.; в двоично-рациональных точках ξ и η совпадают. Отсюда $\xi(t) = \eta(t)$ почти наверное. ■

■

Пример 1.2. Пуассоновский процесс с интенсивностью $\lambda \in (0;\infty)$. Траектории разрывны.

$$\mathsf{M}\left|\xi(t+h) - \xi(t)\right|^a \geqslant \mathsf{P}\left\{\xi(t+h) > \xi(t)\right\} = 1 - e^{-\lambda h} = \lambda h(1 + o(h)) \quad \text{при } h \longrightarrow 0.$$

Нет логарифмического множителя. Теорема Колмогорова неприменима.

9.2. Свойства винеровского процесса

Лемма 9.6. *Стандартный винеровский процесс имеет модификацию с п.н. непрерывными траекториями.*

□ Проверим условия теоремы Колмогорова. Положим $a = 4$. $w(t+h) - w(t) \sim \mathcal{N}(0, h)$ $(h > 0)$. $\mathsf{M}(w(t+h) - w(t))^4 = 3h^2$ (интеграл берётся по частям), а $3h^2 = o\left(\frac{|h|}{\ln^6|h|}\right)$ (например) — теорема Колмогорова работает. ■

Лемма 9.7. *Если $w(t)$ — стандартный винеровский процесс, то процессы $w_a(t) = \dfrac{1}{\sqrt{a}}w(ta)$ (для любого $a > 0$) и $w^*(t) = tw\left(\dfrac{1}{t}\right)$, $w^*(0) = 0$ имеют такие же распределения, как и $w(t)$.*

□ Независимость приращений очевидна (в силу монотонности преобразований). Очевидно, что $w_a(0) = 0$. Распределение приращений:

$$w_a(t) - w_a(s) = \frac{w(at) - w(as)}{\sqrt{a}} \sim \frac{\mathcal{N}(0, a(t-s))}{\sqrt{a}} = \mathcal{N}(0, t-s);$$

$$w^*(t) - w^*(s) = tw\left(\frac{1}{t}\right) - sw\left(\frac{1}{t}\right) = (t-s)w\left(\frac{1}{t}\right) - s\left(w\left(\frac{1}{s}\right) - w\left(\frac{1}{t}\right)\right)$$

Уменьшаемое и вычитаемое суть приращения на непересекающихся отрезках и потому независимы.

$(t-s)w(1/t) \sim \mathcal{N}(0, (t-s)^2/t)$, $s(w(1/s) - w(1/t)) \sim \mathcal{N}(0, s^2(1/s - 1/t))$, откуда

$$w^*(t) - w^*(s) \sim \mathcal{N}\left(0, \frac{(t-s)^2}{t} + s^2\left(\frac{1}{s} - \frac{1}{t}\right)\right) = \mathcal{N}(0, t-s).$$

В частности, $\mathsf{M}(w^*(0))^2 = 0$, то есть $w^*(0) = 0$ п.н. ■

Лемма 9.8.

$$\mathsf{P}\left\{\sup_{0\leqslant u\leqslant t} w(u) \geqslant x\right\} = 2\mathsf{P}\{w(t) \geqslant x\} = \frac{2}{\sqrt{2\pi t}}\int\limits_x^{+\infty} e^{-\frac{v^2}{2t}}\,dv$$

□ Положим $\tau_x = \inf\{u : w(u) \geqslant x\}$ — первое пересечение уровня x. $w(\tau_x) = x$. Имеем:

$$\{\tau_x > v\} = \left\{\sup_{0\leqslant u\leqslant v} w(u) < x\right\} \in \mathscr{F}_v := \sigma(w(u), u \leqslant v);$$

$$\{\tau_x = v\} = \bigcap_{n=1}^{\infty}\left\{\sup_{0\leqslant u\leqslant v-\frac{1}{n}} w(u) < x, \sup_{0\leqslant u\leqslant v} w(u) = x\right\} \in \mathscr{F}_v,$$

поэтому τ_x — момент остановки относительно $\{\mathscr{F}_v\}$. Если $\tau_x = v$, то $w(t) - w(v) = w(t) - w(\tau_x)$ — не зависит от $\{w(y), y \leqslant v\}$ и имеет распределение $\mathcal{N}(0, t-v)$. Отсюда (при $v < t$) $\mathsf{P}(w(t) \geqslant x \mid \tau_x = v) = \mathsf{P}(w(t) \leqslant x \mid \tau_x = v) = \frac{1}{2}$ (симметричность нормального распределения). Далее, $\mathsf{P}(w(t) \geqslant x \mid \tau_x) = \frac{1}{2}$ на $\{\tau_x < t\}$. Поэтому $\mathsf{P}\{w(t) \geqslant x\} = \mathsf{M}\mathsf{P}(w(t) \geqslant x \mid \tau_x)\mathrm{I}_{\{\tau_x < t\}} = \frac{1}{2}\mathsf{P}\{\tau_x < t\} = \frac{1}{2}\mathsf{P}\left\{\sup_{0\leqslant u\leqslant t} w(u) \geqslant x\right\}$. ■

Приведённое выше доказательство не является вполне строгим, поскольку мы на самом деле не обосновали то, что $\mathsf{P}(w(t) - w(v) \mid \tau_x)(\omega \in \{\tau_x = v\}) = \mathsf{P}(w(t) - w(\tau_x(\omega)) \mid \tau_x)(\omega)$ хотя бы почти всюду. На множествах $\{\tau_x = v\}$ это действительно так, но их несчётное число, поэтому про поведение этой условной вероятности на их объединении ничего сказать нельзя. Для того, чтобы найти распределение $\sup_{[0,t]} w(u)$, нам потребуется несколько лемм.

Лемма 9.9 (Марковское свойство винеровского процесса). *Для любого $a \geqslant 0$ процесс $Y(t,\omega) = W(t+a,\omega) - W(a,\omega)$ является стандартным винеровским и не зависит от σ-алгебры $\mathscr{F}_a = \sigma\{X(t), t \leqslant a\}$.*

$\square$ То, что $Y(0) = 0$ и что конечномерные распределения $\{Y(t)\}$ совпадают с нужными, очевидно. Для того, чтобы доказать независимость σ-алгебр $\sigma\{Y(t)\}$ и $\mathscr{F}_a$, достаточно для любых конечных наборов $0 \leqslant t_1 < \ldots < t_n \leqslant a$, $0 \leqslant s_1 < \ldots < s_m$ доказать независимость событий $\{\omega\colon X(t_1) \in B_1, \ldots, X(t_n) \in B_n\}$ и $\{\omega\colon Y(s_1) \in C_1, \ldots, Y(s_m) \in C_m\} = \{\omega\colon X(s_1+a) - X(a) \in C_1, \ldots, X(s_m+a) - X(a) \in C_m\}$, что очевидно, так как эти события выражаются через независимые группы приращений процесса X, а именно,
$(X(t_1), \ldots, X(t_n) - X(t_{n-1}))$ и $(X(s_1+a) - X(a), \ldots, X(s_m+a) - X(s_{m-1}+a))$. $\blacksquare$

На самом деле верно ещё более сильное утверждение.

Лемма 9.10 (Строго марковское свойство винеровского процесса). *Пусть $\{W(t), t \geqslant 0\}$ — винеровский процесс, $\mathscr{F}_t = \sigma\{X(s), s \leqslant t\}$ — его естественная фильтрация и τ — момент остановки относительно потока $\mathscr{F}_t$. Тогда процесс $\{Y(t,\omega) = X(t+\tau(\omega)) - X(\tau(\omega)), t \geqslant 0\}$ также является винеровским, причём не зависящим от σ-алгебры $\mathscr{F}_\tau$.*

$\square$ Доопределим на множестве $\{\tau = \infty\}$ (оно имеет меру ноль, так как τ – момент остановки) процесс Y тождественно нулевыми траекториями. Покажем, что мы получили действительно случайный процесс, т.е., что $Y(t)$ есть случайная величина для любого t. Определим последовательность дискретных марковских моментов τ_n, сходящуюся к τ, следующим образом:

$$\tau_n(\omega) = \sum_{k=1}^{\infty} k2^{-n} \mathrm{I}_{(k-1)2^{-n} < \tau(\omega) \leqslant k2^{-n}}$$

Очевидно, что $\tau_n \xrightarrow{\text{п.н.}} \tau$ и что это марковские моменты относительно $\mathscr{F}_n$, т.к. $\{\tau_n \leqslant t\} = \{\tau \leqslant k2^{-n}\} \in \mathscr{F}_{k2^{-n}} \subset \mathscr{F}_t$, где $k = \max\{l\colon l2^{-n} \leqslant t\}$. Так как траектории $W(t)$ непрерывны п.н., то $W(t+\tau_n) \xrightarrow{\text{п.н.}} W(t+\tau), n \longrightarrow \infty$. Для любых $n \in \mathbb{N}, z \in \mathbb{R}, t \geqslant 0$ имеем

$$\{\omega\colon W(t+\tau_n(\omega),\omega) \leqslant z\} = \cup_{k=1}^{\infty} \left\{\omega\colon W(t+k2^{-n},\omega), \tau_n(\omega) = k2^{-n}\right\} \in \mathscr{F}.$$

Следовательно, $W(t+\tau)$, а заодно и $Y(t)$ являются случайными величинами как пределы сходящихся почти наверное случайных величин. Докажем, что $Y(t)$ не зависит от $\mathscr{F}_\tau$ и заодно проверим, что это винеровский процесс. Для этого, как и в предыдущей лемме, достаточно проверить, что для любого $A \in \mathscr{F}_\tau$, любых точек $0 \leqslant t_1 < \ldots < t_m$ и любого $B \in \mathcal{B}(\mathbb{R}^m)$ $\mathsf{P}\{A \cap \{\xi \in B\}\} = \mathsf{P}\{A\}\mathsf{P}\{\xi \in B\}$, где $\xi = (Y(t_1), \ldots, Y(t_m)) \in \mathbb{R}^m$. Достаточно показать это для замкнутого B, поскольку для любой меры P в R^m и для любого борелевского B найдётся такое замкнутое подмножество $F_\varepsilon \subset B$, что $\mathsf{P}(B \setminus F_\varepsilon) < \varepsilon$. Итак, нам нужно показать, что $\mathsf{MI}_A \mathrm{I}_{\xi \in B} = \mathsf{MI}_A \mathsf{MI}_{\xi \in B}$. Мы это докажем, если покажем, что для любой непрерывной и ограниченной функции выполнено $\mathsf{MI}_A f(\xi) = \mathsf{MI}_A \mathsf{M} f(\xi)$. Действительно, тогда, взяв $f_k(x) = \max(0, 1 - k\rho(x,B))$, получим требуемое из теоремы Лебега ($f_k(x) \longrightarrow \mathrm{I}_B(x), k \longrightarrow \infty$). Опять же по теореме Лебега имеем

$$\mathsf{MI}_A f(\xi) = \lim_n \mathsf{MI}_a f(\xi_n),$$

где $\xi_n = (W(t_1+\tau_n) - W(\tau_n), \ldots, W(t_m+\tau_n) - W(\tau_n)) \xrightarrow{\text{п.н.}} \xi, n \longrightarrow \infty$. Воспользовавшись счётной аддитивностью интеграла Лебега, получаем

$$\mathsf{MI}_A f(\xi_n) = \sum_{k=1}^{\infty} \mathsf{MI}_A f(\xi_n) \mathrm{I}_{\tau_n = k2^{-n}} = \sum_{k=1}^{\infty} \mathsf{MI}_{A \cap \{\tau_n = k2^{-n}\}} f(\xi_{n,k}),$$

где $\xi_{n,k} = (W(t_1+k2^{-n}) - W(k2^{-n}), \ldots, W(t_m+k2^{-n}) - W(k2^{-n}))$. Т.к. $A \in \mathscr{F}_\tau$, то $A \cap \{\tau_n = k2^{-n}\} = A \cap \{(k-1)2^{-n} < \tau \leqslant k2^{-n}\} \in \mathscr{F}_{k2^{-n}}$. По предыдущей лемме $\xi_{n,k}$ не зависит от $\mathscr{F}_{k2^{-n}}$. При этом распределение $\xi_{n,k}$ совпадает с распределением $(W(t_1), \ldots, W(t_m))$. Поэтому

$$\mathsf{MI}_A f(\xi_n) = \mathsf{M} f(W(t_1), \ldots, W(t_m)) \sum_{k=1}^{\infty} \mathsf{MI}_{A \cap \{\tau_n = k2^{-n}\}} = \mathsf{M} f(W(t_1), \ldots, W(t_m)) \mathsf{MI}_A.$$

Осталось показать, что $\mathsf{M} f(Y(t_1), \ldots, Y(t_m)) = \mathsf{M} f(W(t_1), \ldots, W(t_m))$, что немедленно следует из доказанного при $A = \Omega$. Взяв ограниченные непрерывные f_k, сходящиеся к индикаторным функциям, по теореме Лебега получим, что конечномерные распределения W и Y совпадают, следовательно, Y — винеровский процесс. $\blacksquare$

Ключевым инструментом для нахождения распределения $\sup W(t)$ на отрезках (и много чего другого, на самом деле) является следующая

Теорема 9.11 (Принцип отражения). *Пусть $\{W(t), t \geqslant 0\}$ — винеровский процесс, τ — момент остановки относительно $\mathscr{F}_t = \sigma\{X(s), s \leqslant t\}$ Положим*

$$Z(t,\omega) = \begin{cases} W(t,\omega), & 0 \leqslant t \leqslant \tau(\omega), \\ 2W(\tau(\omega),\omega) - W(t,\omega), & t > \tau(\omega), \\ W(t,\omega), & \tau(\omega) = \infty. \end{cases}$$

Каждая траектория Z получается отражением траектории X относительно линии уровня $X(\tau)$ при $x > \tau$. Оказывается, $Z(t,\omega)$ – тоже винеровский процесс.

$\square$ По определению $Z(t,\omega) = W(t,\omega)\mathrm{I}_{\{\tau \geqslant t\}} + (2W(\tau(\omega),\omega) - W(t,\omega))\mathrm{I}_{\{\tau < t\}}$, следовательно, это случайная величина при каждом $t \geqslant 0$. Траектории Z непрерывны п.н., поэтому нетрудно показать, что $Z(\cdot,\omega)$ есть случайный элемент со значениями в метрическом пространстве $C_0[0,\infty) = \{f \in C[0,\infty), f(0) = 0\}$, метрика на котором задаётся равномерной сходимостью на компактах, т.е.

$$\rho(f,g) = \sum_{n=1}^{\infty} 2^{-n} \frac{\sup_{[0,n]} |f(t) - g(t)|}{1 + \sup_{[0,n]} |f(t) - g(t)|}.$$

Это доказывается так: во-первых, $C[0,\infty)$ сепарабельно (на любом компакте функция равномерно непрерывна и к ней сходится последовательность линейных функций с изломами в рациональных точках и рациональными значениями в точках излома, а потом берём диагональную последовательность таких функций), следовательно, его борелевская σ-алгебра порождается всеми открытыми шарами, а прообраз каждого такого шара, как нетрудно показать, измерим. А именно, $B_r(x) = \cap_N \left\{y \colon \sum_{k=1}^{N} 2^{-k} \frac{\sup_{\mathbb{Q}\cap[0,k]} |x(t)-y(t)|}{1+\sup_{\mathbb{Q}\cap[0,k]}} \leqslant r\right\}$. Это пересечение замкнутых множеств в $\mathbb{R}^N$ (вектора в $\mathbb{R}^N$ соответствуют наборам из $\sup_{\mathbb{Q}\cap[0,k]} |x(t) - y(t)|$), оно замкнуто, следовательно, его дополнение открыто и представляется в виде счётного объединения кубов, а прообразы кубов измеримы.

Положим $Y(t,\omega) = W(t+\tau(\omega),\omega) - W(\tau(\omega),\omega)$. Как мы знаем, это винеровский процесс. Положим $X(t,\omega) = W(\min(t,\omega),\omega)$ — это процесс, «остановленный» в точке τ. Аналогично доказанному получаем, что X, Y — случайные элементы со значениями в $C_0[0,\infty)$. Рассмотрим метрическое пространство $V = [0,\infty) \times C_0[0,\infty) \times C_0[0,\infty)$ с метрикой, равной максимуму из покоординатных расстояний и рассмотрим отображение $h : V \longrightarrow C_0[0,\infty)$, задаваемое следующим образом:

$$h(b,f,g)(t) = f(t)\mathrm{I}_{[0,b]}(t) + (f(b) + g(t-b))\mathrm{I}_{(b,\infty)}(t).$$

Как легко видеть, оно непрерывно и для почти всех ω (кроме тех, для которых τ бесконечен или траектория одного из процессов разрывна)

$$h(\tau(\omega), X(\cdot,\omega), Y(\cdot,\omega)) = W(\cdot,\omega), h(\tau(\omega), X(\cdot,\omega), -Y(\cdot,\omega)) = Z(\cdot,\omega).$$

Если мы докажем, что на $(V, \mathcal{B}(V))$ случайные вектора (τ, X, Y) и $(\tau, X, -Y)$ имеют одинаковое распределение, то в силу измеримости h получим утверждение теоремы. Докажем, что (τ, X) измерим относительно $\mathscr{F}_\tau$, тогда в силу независимости Y от $\mathscr{F}_\tau$ распределение распадётся в произведение распределений $(\tau, X) \otimes (\pm Y)$, распределения же Y и $-Y$ совпадают в силу симметричности винеровского процесса. Для измеримости вектора относительно $\mathscr{F}_\tau$ необходима и достаточна измеримость каждой его компоненты. $\mathscr{F}_\tau$-измеримость τ следует из определения $\mathscr{F}_\tau$ (проверьте), для того, чтобы доказать $\mathscr{F}_\tau$-измеримость X, введём $\mathscr{F}_\tau$-измеримую последовательность X_n, сходящуюся к X п.н. Положим

$$\alpha_n(\omega) = \sum_{k=0}^{\infty} k 2^{-n} \mathrm{I}_{k2^{-n} \leqslant \tau(\omega) < (k+1)2^{-n}} \uparrow \tau(\omega), n \longrightarrow \infty,$$

и определим $X_n(t,\omega) = W(\min(t, \alpha_n(\omega)), \omega)$. Проверка $\mathscr{F}_\tau$-измеримости X_n тривиальна:

$$\{\omega \colon X_n(t,\omega) \leqslant x, \tau \leqslant s\} = \cup_k \left\{X(\min(t, k2^{-n})) \leqslant x, \tau \leqslant s, \tau \in [k2^{-n}, (k+1)2^{-n})\right\} \in \mathscr{F}_s.$$

$\blacksquare$

Докажем, наконец, то, ради чего мы это всё вводили.

Теорема 9.12. *Пусть $M(t) = \sup_{[0,t]} W(s)$ (это случайная величина, т.к. $f \mapsto sup_{[0,t]} f$ есть непрерывное отображение $C_0[0,\infty) \longrightarrow \mathbb{R}$). Тогда $\forall x, y, t \geqslant 0$*

$$\mathsf{P}(W(t) < y - x, M(t) \geqslant y) = P(W(t) > y + x).$$

$\square$ Если $y = 0$, то утверждение теоремы тривиально. Пусть $y > 0$. Тогда $\tau_y = \inf\{s \in [0,\infty) \colon W(s) \geqslant y\}$ есть момент остановки относительно $\{\mathscr{F}_t = \sigma\{W(t)\}\}$. Он конечен п.н. в силу закона повторного логарифма для случайного блуждания $W(n)$ (почти наверное существует целая точка, в которой $W(n)$ сколь угодно велико). Рассмотрим «отражённый» относительно τ_y процесс Z. Тогда $\sigma_y = \inf\{s \geqslant 0 \colon Z(s) = y\}$ — тоже момент остановки относительно $\{\sigma\{Z(t)\}\}$, причём при всех $y \geqslant 0$, $\omega \in \{\tau_y < \infty\}$ имеем $\sigma_y(\omega) = \tau_y(\omega)$, потому что до момента $\tau_y(\omega)$ траектории X и Z совпадают. Заметим, что $\{\tau_y \leqslant t\} = \{M(t) \geqslant y\}$ Поэтому для любого $B \in \mathcal{B}(C_0[0,\infty))$, $t \geqslant 0$ имеем

$$\mathsf{P}(\tau_y \leqslant t, W \in B) = P(W \in \hat{B} \cap B),$$

где $\hat{B} = \{f\colon \sup_{[0,t]} \geqslant t\} \in \mathcal{B}(C_0[0,\infty))$ в силу непрерывности взятия sup. Одновременно и

$$\mathsf{P}(\sigma_y \leqslant t, Z \in B) = \mathsf{P}(Z \in \hat{B} \cap B) = \mathsf{P}(W \in \hat{B} \cap B),$$

так как Z — тоже винеровский процесс. Получили, что (τ_y, W) и (σ_y, Z) распределены одинаково. В силу непрерывности траекторий W имеем $W(\tau_y(\omega), \omega) = y$ п.н., поэтому при $t \geqslant \sigma_y(\omega)$ получаем $Z(t,\omega) = 2W(\tau_y(\omega), \omega) - W(t,\omega) = 2y - W(t,\omega)$. Следовательно,

$$\mathsf{P}(M(t) \geqslant y, W(t) < y - x) = \mathsf{P}(\sigma_y \leqslant t, Z(t) < y - x) =$$
$$\mathsf{P}(\sigma_y \leqslant t, W(t) > y + x) = \mathsf{P}(\tau_y \leqslant t, W(t) > y + x) = \mathsf{P}(M(t) \geqslant y, W(t) > y + x) = \mathsf{P}(W(t) > y + x), x \geqslant 0.$$

■

Следствие 9.1.

$$\mathsf{P}(M(t) \geqslant y) = 2\mathsf{P}(W(t) \geqslant y).$$

□ Возьмём в предыдущей теореме $x = 0$. Получим

$$\mathsf{P}(W(t) < y, M(t) \geqslant y) = \mathsf{P}(W(t) > y).$$

Имеем $\mathsf{P}(M(t) \geqslant y) = \mathsf{P}(M(t) \geqslant y, W(t) < y) + \mathsf{P}(M(t) \geqslant y, W(t) \geqslant y) = \mathsf{P}(W(t) > y) + \mathsf{P}(W(t) \geqslant y) = 2\mathsf{P}(W(t) \geqslant y)$ (в силу непрерывности нормального распределения $\mathsf{P}(W(t) = y) = 0$. ■

9.3. Закон повторного логарифма для винеровского процесса

Теорема 9.13 (закон повторного логарифма для винеровского процесса).

$$\mathsf{P}\left\{\limsup_{t\longrightarrow\infty} \frac{w(t)}{\sqrt{2t\ln\ln t}} = 1\right\} = \mathsf{P}\left\{\liminf_{t\longrightarrow\infty} \frac{w(t)}{\sqrt{2t\ln\ln t}} = -1\right\} = 1;$$
$$\mathsf{P}\left\{\limsup_{t\longrightarrow 0+} \frac{w(t)}{\sqrt{2t\ln\ln\frac{1}{t}}} = 1\right\} = \mathsf{P}\left\{\liminf_{t\longrightarrow 0+} \frac{w(t)}{\sqrt{2t\ln\ln\frac{1}{t}}} = -1\right\} = 1.$$

□ Вторая строчка следует из первой (рассмотрим w^*). Значения винеровского процесса в целых точках $\{w(n)\}$ образуют случайное блуждание с распределением $\mathcal{N}(0,1)$ — применяем закон повторного логарифма для случайного блуждания и следующую лемму:

Лемма 9.14.

$$\mathsf{P}\left\{\lim_{t\to\infty}\left|\frac{w(t)}{\sqrt{2t\ln\ln t}} - \frac{w([t])}{\sqrt{2t\ln\ln t}}\right| = 0\right\} = 1.$$

□ Из однородности винеровского процесса по времени и симметричности нормального распределения получаем:

$$\mathsf{P}\left\{\sup_{n\leqslant t\leqslant n+1} |w(t) - w(n)| > x\right\} \leqslant \mathsf{P}\left\{\sup_{0\leqslant t\leqslant 1} w(t) > x\right\} + \mathsf{P}\left\{\inf_{0\leqslant t\leqslant 1} w(t) < -x\right\} =$$
$$= 2\mathsf{P}\left\{\sup_{0\leqslant t\leqslant 1} w(t) > x\right\} = \frac{4}{\sqrt{2\pi}}\int\limits_x^{+\infty} e^{-\frac{v^2}{2}}\,dv = O\left(e^{-\frac{x^2}{2}}\right).$$

Положим $A_n = \left\{\sup_{n\leqslant t\leqslant n+1} |w(t) - w(n)| > \sqrt{2n}\right\}$. $\mathsf{P}(A_n) = O(e^{-n})$ — ряд по n сходится.

$$\mathsf{P}\left\{\sup_{t\geqslant N} \frac{|w(t) - w([t])|}{\sqrt{2t}} > 1\right\} = \mathsf{P}\left(\bigcup_{n=N}^{\infty} A_n\right) \leqslant \sum_{n=N}^{\infty} \mathsf{P}(A_n) \longrightarrow 0 \quad (N \longrightarrow \infty).$$

■

■

9.4. Неограниченность вариации траекторий винеровского процесса

Из закона повторного логарифма следует, что траектории винеровского процесса $w(t)$ при $t \longrightarrow 0$ осциллируют между кривыми $w = \sqrt{2t\ln\ln\frac{1}{t}}$ и $w = -\sqrt{2t\ln\ln\frac{1}{t}}$.

Напомним некоторые факты из действительного анализа. *Вариацией* f на отрезке $[a,b]$ называется $\mathrm{Var}_{[a,b]} f = \sup_T \sum_{j=1}^N |f(t_j) - f(t_{j-1}|$, где T пробегает разбиения $a = t_0 < \ldots < t_N = b$. Если f дифференцируема, то $\mathrm{Var}_{[a,b]} f = \int\limits_a^b |f'(t)| dt$.

Лемма 9.15. *Обозначим* $\Delta t_j \rightleftharpoons t_j - t_{j-1}$. *Если* $w(t)$ *— процесс броуновского движения, то*

$$\mathsf{P}\left\{\mathrm{Var}_{[0,1]} w = \infty\right\} = 1$$

и для любых $t > 0, \varepsilon > 0$ *и для любой последовательности* $0 = t_0 < t_1 < \ldots$ *имеем*

$$\lim_{\max \Delta t_j \longrightarrow 0} \mathsf{P}\left\{|\sum_{j:\, t_j \leqslant t} (w(t_j) - w(t_{j-1}))^2 - t| > \varepsilon\right\} = 0.$$

$\square$ $\mathsf{M}|w(t_j) - w(t_{j-1})| = \mathsf{M}|w(\Delta t_j)|$. Так как $w(\Delta t_j) \sim \mathcal{N}(0, \Delta t_j)$, то

$$\mathsf{M}|w(\Delta t_j)| = \frac{1}{\sqrt{2\pi \Delta t_j}} \int\limits_{-\infty}^{+\infty} |x| e^{-\frac{x^2}{2\Delta t_j}} dx = \frac{2}{\sqrt{2\pi \Delta t_j}} \int\limits_0^{\infty} x e^{-\frac{x^2}{2\Delta t_j}} dx = \sqrt{\frac{2\Delta t_j}{\pi}} > \Delta t_j \sqrt{\frac{2}{\pi \max \Delta t_j}}.$$

$$\mathsf{D}|w(\Delta t_j)| = \mathsf{M} w^2(\Delta t_j) - \mathsf{M}^2 |w(\Delta t_j)| = \Delta t_j - \frac{2\Delta t_j}{\pi} = \Delta t_j \left(1 - \frac{2}{\pi}\right).$$

$$m(\{t_j\}) \rightleftharpoons \mathsf{M} \sum_{j:\, t_j \leqslant 1} |w(t_j) - w(t_{j-1})| \leqslant \sqrt{\frac{2}{\pi \max \Delta t_j}} \sum_{j:\, t_j \leqslant 1} \Delta t_j.$$

$$d(\{t_j\}) \rightleftharpoons \mathsf{D} \sum_{j:\, t_j \leqslant 1} |w(t_j) - w(t_{j-1})| = \left(1 - \frac{2}{\pi}\right) \sum_{j:\, t_j \leqslant 1} \Delta t_j = 1 - \frac{2}{\pi} + O(1).$$

$$\mathsf{P}\left\{\sum_{j:\, t_j \leqslant 1} |w(t_j) - w(t_{j-1})| < x\right\} \leqslant \mathsf{P}\left\{\left|\sum_{j:\, t_j \leqslant 1} |w(t_j) - w(t_{j-1})| - m(\{t_j\})\right| \geqslant m(\{t_j\}) - x\right\} \leqslant$$
$$\leqslant \frac{d(\{t_j\})}{m(\{t_j\}) - x)^2} \longrightarrow 0 \quad (\max \Delta t_j \longrightarrow 0).$$

Вторая часть утверждения доказывается аналогично:

$$\mathsf{M}\left(w(t_j) - w(t_{j-1})\right)^2 = \mathsf{M} w^2(\Delta t_j) = \Delta t_j, \mathsf{M}\left(w(t_j) - w(t_{j-1})\right)^4 = \mathsf{M} w^4(\Delta t_j) = 3(\Delta t_j)^2,$$
$$\mathsf{D}\left(w(t_j) - w(t_{j-1})\right)^2 = \mathsf{M} w^4(\Delta t_j) - (\mathsf{M} w^2(\Delta t_j))^2 = 2(\Delta t_j)^2.$$

$$\mathsf{M} \sum_{j:\, t_j \leqslant t} \left(w(t_j) - w(t_{j-1})\right)^2 = \sum_{j:\, t_j \leqslant t} \Delta t_j \longrightarrow t \quad (\max \Delta t_j \longrightarrow 0).$$

$$\mathsf{D} \sum_{j:\, t_j \leqslant t} \left(w(t_j) - w(t_{j-1})\right)^2 = 2 \sum_{j:\, t_j \leqslant t} (\Delta t_j)^2 \longrightarrow 0 \quad (\max \Delta t_j \longrightarrow 0).$$

$\blacksquare$