ТЕОРИЯ ВЕРОЯТНОСТЕЙ

А.В. Колесников

Содержание

1. Лекция 1 (вводная)

Настоящий курс представляет собой продолжение вводного "элементарного" курса теории вероятностей. Основное содержание курса составляет теория случайных процессов, более точно: построение винеровского процесса (броуновского движения), стохастический анализ и введение в теорию диффузионных процессов. Уделено большое внимание теории слабой сходимости мер, в том числе на бесконечномерных пространствах. Теория конечных марковских цепей представлена выборочно. Теория марковских цепей со счетным числом состояний, а также дискретных цепей с непрерывным временем, в общем виде не рассказывается, представлена в виде отдельных примеров. Некоторые сложные вопросы теории диффузионных процессов (марковские свойства, уравнения Колмогорова, аккуратное доказательство формулы Ито) представлены обзорно.

Предполагается, что слушатель хорошо владеет следующим материалом теории меры и теории вероятностей. Для ознакомления с ним рекомендуются книги [1], [4], [5].

(1) Сигма-алгебры. Определение, примеры, базовые свойства.
(2) Борелевская σ-алгебра на $\mathbb{R}^n$.
(3) Меры. Мера Лебега. Вероятностные пространства.
(4) Измеримые множества. Случайные величины. Интеграл Лебега/математическое ожидание.
(5) Независимость.
(6) Абсолютная непрерывность меры относительно другой меры. Теорема Радона-Никодима.
(7) Характеристические функции (преобразование Фурье).
(8) Закон больших чисел. Центральная предельная теорема.

Вводная лекция данного курса включают в себя материал, необходимый для понимания дальнейшего. Этот материал, как правило, входит в элементарный курс, но часто не усваивается должным образом из-за нехватки времени или предпочтений преподавателя.

1.1. Многомерные распределения. Гауссовские распределения. Преобразования распределений при отображениях.

Всюду далее $(\Omega, \mathcal{F}, P)$ — вероятностное пространство.

Определение 1.1. *(Измеримые отображения) Отображение $F : \Omega \to Y$, где Y — множество, наделенное σ-алгеброй $\mathcal{B}$, называется измеримым ($\mathcal{B}$-измеримым), если $F^{-1}(B) \in \mathcal{F}$, для любого $B \in \mathcal{B}$.*

Как правило, отображения, которые мы будем рассматривать, будут принимать значения в $\mathbb{R}^n$. Пространство $\mathbb{R}^n$ всегда будет наделяться борелевской σ-алгеброй. Следовательно, измеримым отображением $F : \Omega \to \mathbb{R}^n$ будет считаться любое отображение со свойством $F^{-1}(B) \in \mathcal{F}$ для любого борелевского B.

Определение 1.2. *(Образы мер) Каждое измеримое отображением $F : \Omega \to Y$, где Y — множество, наделенное σ-алгеброй $\mathcal{B}$, определяет меру-образ $\mu = P \circ F^{-1}$ на $\mathcal{B}$ по формуле*

$$\mu(A) = P(F^{-1}(A)).$$

Определение 1.3. *Случайным вектором $\xi : \Omega \to \mathbb{R}^n$ называется любое измеримое отображение из Ω в $\mathbb{R}^n$. Мера $\mu_\xi = P \circ \xi^{-1}$ называется его распределением.*

Определение 1.4. *Случайный вектор $\xi : \Omega \to \mathbb{R}^n$ называется непрерывным, если мера $\mu_\xi = P \circ \xi^{-1}$ обладает плотностью относительно меры Лебега, т.е., существует функция $\rho : \mathbb{R}^n \to \mathbb{R}^+$, удовлетворяющая условию $\int \rho \, dx = 1$ и*

$$P(\xi \in B) = \mu_\xi(B) = \int_B \rho \, dx.$$

Определение 1.5. *Маргинальным распределением вектора ξ называется одномерное распределение его i-й компоненты ξ_i*

$$\xi = (\xi_1, \cdots, \xi_i, \cdots, \xi_n)$$

Нетрудно видеть, что i-е маргинальное распределение непрерывного случайного вектора задается формулой

$$P(\xi_i \in B) = \int_{\{x : x_i \in B\}} \rho_\xi \, dx.$$

Существует простой критерий независимости совокупности случайных величин (см. [4]).

Теорема 1.6. *Компоненты $\{\xi_i\}$ случайного вектора ξ являются независимыми в совокупности тогда и только тогда, когда многомерная функция распределения ξ, заданная формулой $F(x) = P(\xi_1 < x_1, \cdots, \xi_i < x_i, \cdots, \xi_n < x_n)$ представляется в виде произведения*

$$F(x) = \prod_{i=1}^n F_i(x_i),$$

где F_i — функция распределения соответствующего маргинального распределения.

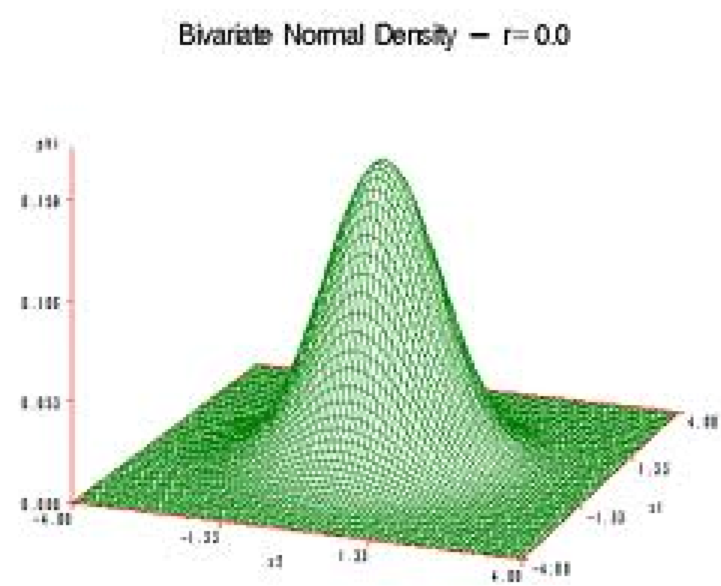
Определение 1.7. *Случайный n-мерный вектор ξ называется гауссовским, если его характеристическая функция*

$$\varphi(t) = \mathbb{E}(e^{i\langle t, \xi \rangle}), \; t \in \mathbb{R}^n$$

имеет вид

$$\varphi(t) = \exp\Big(i\langle t, a\rangle - \frac{1}{2}\langle At, t\rangle\Big),$$

где $a \in \mathbb{R}^n$ и A — симметричная неотрицательная матрица.



В случае невырожденной матрицы A можно явно указать плотность ξ.

Теорема 1.8. *Если A невырождена, то ξ имеет плотность*

$$f(x) = \sqrt{\frac{\det Q}{(2\pi)^n}} \exp\Big(-\frac{\langle Q(x-a), (x-a)\rangle}{2}\Big),$$

где $Q = A^{-1}$.

Доказательство. Переходя к новому ортогональному базису, можно считать, что Q диагональна. Тогда теорема легко следует из соответствующей одномерной теоремы. □

В общем случае надо приблизить матрицу A по норме невырожденными матрицами и воспользоваться соотношением между слабой сходимостью мер и поточечной сходимостью функционалов (теорема Бохнера).

Теорема 1.9. *Линейный образ гауссовского распределения является гауссовским распределением.*

Определение 1.10. *Ковариационной матрицей случайного вектора ξ называется матрица C, где*

$$C_{ij} = \mathbb{E}\big(\xi_i - \mathbb{E}(\xi_i)\big)\big(\xi_j - \mathbb{E}(\xi_j)\big).$$

Упражнение 1.11. *Докажите, что C симметрична и неотрицательно определена.*

Теорема 1.12. *Для гауссовского случайного вектора $a = \mathbb{E}\xi$ и*

$$C = A.$$

Следствие 1.13. *У гауссовского случайного вектора компоненты независимы тогда и только тогда, когда они некоррелированы (C диагональна).*

Теорема 1.14. *(Свертка) Пусть ξ, η — независимые непрерывные случайные величины с плотностями распределения f_ξ, f_η. Тогда плотность распределения с.в. $\theta = \xi + \eta$ задается формулой свертки*

$$f_{\xi+\eta}(x) = \int_{-\infty}^{+\infty} f_\xi(t) f_\eta(x-t)\, dt.$$

Пример 1.15. *Распределение χ^2.*

Пусть ξ_1, ξ_2 — независимые стандартные гауссовские (т.е. $\xi_i \sim \mathcal{N}(0,1)$) случайные величины. Найдем распределение с.в.

$$\eta = \xi_1^2 + \xi_2^2.$$

Плотность распределения ξ_1^2 находится из соотношения

$$P(\xi_1^2 \le x) = \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{x}}^{\sqrt{x}} e^{-\frac{t^2}{2}}\, dt.$$

Дифференцируя по x, получаем

$$f_{\xi_1^2}(x) = \begin{cases} 0, \ x \le 0 \\ \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}}, \ x > 0. \end{cases}$$

Следовательно, по формуле свертки, при $x > 0$

$$f_\eta(x) = \frac{1}{2\pi}\int_0^x \frac{e^{-\frac{t}{2}}e^{-\frac{(x-t)}{2}}}{\sqrt{t(x-t)}}\,dt = \frac{e^{-\frac{x}{2}}}{2\pi}\int_0^x \frac{1}{\sqrt{t(x-t)}}\,dt = \frac{e^{-\frac{x}{2}}}{2\pi}\int_0^1 \frac{1}{\sqrt{t(1-t)}}\,dt$$
$$= B\Big(\frac{1}{2},\frac{1}{2}\Big)\frac{1}{2\pi}e^{-\frac{x}{2}} = \frac{1}{2}e^{-\frac{x}{2}}.$$

Применяя индукцию, легко доказать что с.в.

$$\eta = \xi_1^2 + \xi_2^2 + \cdots + \xi_n^2,$$

где $\xi_i \sim \mathcal{N}(0,1)$ и все ξ_i независимы, имееет плотность распределения

$$f_\eta(x) = \begin{cases} 0, \ x \le 0 \\ \frac{1}{2^{n/2}\Gamma(n/2)}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}, \ x > 0. \end{cases}$$

Это распределение называется распределением χ^2 с n степенями свободы.

Для вычисления формулы плотности для другого классического распределения мы применим следующее простое наблюдение, следующее непосредственно из формулы замены переменной

Теорема 1.16. *Пусть непрерывные случайные векторы ξ и η с плотностями распределения f_ξ, f_η связаны соотношением $F(\eta) = \xi$, где $F : \mathbb{R}^n \to \mathbb{R}^n$ — диффеоморфизм. Тогда выполнено соотношение:*

$$f_\eta = f_\xi(F)|\det DF|.$$

Пример 1.17. *Распределение Стьюдента.*

Найдем распределение с.в.

$$\frac{\xi_0}{\sqrt{\frac{1}{n}(\xi_1^2 + \xi_2^2 + \cdots + \xi_n^2)}},$$

где $\xi_i \sim \mathcal{N}(0,1)$ и все ξ_i независимы. Для этого рассмотрим двумерный случайный вектор

$$(X,Y) = \Big(\xi_0, \xi_1^2 + \xi_2^2 + \cdots + \xi_n^2\Big).$$

В силу независимости его плотность распределения задается формулой ($y > 0$)

$$f_{X,Y} = C(n)e^{-\frac{x^2}{2}-\frac{y}{2}}y^{\frac{n}{2}-1}.$$

Пусть (U,V) — случайный вектор, связанный с (X,Y) соотношением

$$U = X, \ V = \frac{X}{\sqrt{Y/n}}.$$

Очевидно, V имеет искомую плотность распределения. Выразим X, Y через U, V:

$$X = U, \ Y = n\frac{U^2}{V^2}$$

Модуль якобиана этого отображения, как легко проверить, равен

$$\frac{2nU^2}{|V|^3}.$$

По предыдущему предложению вектор (U,V) имеет плотность распределения

$$f_{U,V} = C(n)\exp\Big(-\frac{u^2}{2} - \frac{nu^2}{2v^2}\Big)\frac{|u|^{n-2}}{|v|^{n-2}}\frac{u^2}{|v|^3}.$$

Для того, чтобы получить плотность V , необходимо проинтегрировать полученную функцию по u. Получаем

$$f_V = \frac{C(n)}{|v|^{n+1}} \int_{-\infty}^{\infty} |u|^n \exp\Big(-\frac{(n+v^2)u^2}{2v^2}\Big)\, du.$$

Сделав в интеграле замену переменной

$$s = u\frac{\sqrt{n+v^2}}{v},$$

получим, что для некоторой подходящей константы $C(n)$ плотность V равна

$$f_V = \frac{C(n)}{(n+v^2)^{\frac{n+1}{2}}}.$$

Ниже приведена более точная формулировка (с точной константой).

Следствие 1.18. *Случайная величина*

$$\eta = \frac{\xi_0}{\sqrt{\frac{1}{n}(\xi_1^2+\xi_2^2+\cdots+\xi_n^2)}}$$

имеет распределение Стьюдента с n степенями свободы, которое задается плотностью

$$f_\eta(x) = \frac{1}{\sqrt{\pi n}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{1}{(1+\frac{x^2}{n})^{\frac{n+1}{2}}}.$$

1.2. **Различные виды сходимости случайных величин. Равномерная интегрируемость. Лемма Бореля-Кантелли. Закон 0-1.**

Для усвоения материала лекций надо владеть следующими понятиями: сходимость почти наверное (почти всюду), сходимость по вероятности (по мере), в среднем (в $L^p(\mu)$, $p \geq 1$), по распределению (слабая сходимость мер) и знать соотношения между ними (см. [4], глава 2(10)):

сходимость п.н. $\to$ сходимость по вероятности $\to$ сходимость по распределению.

сходимость в среднем $\to$ сходимость по вероятности.

Из сходимости по вероятности не следует сходимости п.н., но верна следующая теорема.

Теорема 1.19. *(Рисс) Из последовательности с.в $\{\xi_n\}$, сходящейся по вероятности, можно выделить почти наверное сходящуюся подпоследовательность.*

Чрезвычайно важными являются следующие теоремы о сходимости под знаком мат. ожидания

Теорема 1.20. *Пусть $\{\xi_n\}$ — последовательность с.в. Тогда*

1) *(Лебег). Если $\xi_n \to \xi$ п.н. и $|\xi_n| < \eta$, $\eta \in L^1(P)$ то*

$$\lim_n \mathbb{E}(\xi_n) = \mathbb{E}(\xi).$$

2) *(Фату). Если $\xi_n \geq \eta$ и $\mathbb{E}(\eta) > -\infty$, то*

$$\underline{\lim}_n \mathbb{E}(\xi_n) \geq \mathbb{E}(\underline{\lim}_n \xi_n).$$

Определение 1.21. *Последовательность с.в. $\{\xi_n\}$ называется равномерно интегрируемой, если*

$$\lim_{N\to+\infty}\sup_i \mathbb{E}(|\xi_i| I_{|\xi_i|\geq N}) = 0.$$

Теорема 1.22. *Пусть $\xi_i \to \xi$ по мере. Если $\{\xi_i\}$ — равномерно интегрируемая последовательность, то имеет место сходимость $\xi_i \to \xi$ в $L^1(P)$.*

Важным техническим средством для доказательства теорем о суммах независимых с.в. является лемма Бореля-Кантелли.

Лемма 1.23. (Борель-Кантелли). *Пусть $\{A_n\}$ — последовательность событий. Событие A состоит в том, что события A_n случаются бесконечно часто*

$$A = \bigcap_{n=1}^{\infty}\bigcup_{m\geq n} A_m.$$

1) *Если $\sum_{m=1}^{\infty} P(A_m) < \infty$, то $P(A) = 0$.*
2) *Если $\sum_{m=1}^{\infty} P(A_m) = \infty$ и события $\{A_m\}$ независимы, то $P(A) = 1$.*

Доказательство. a) $P(A) \leq \lim_n P(\bigcup_{m\geq n} A_m) \leq \lim_n \sum_{m\geq n} P(A_m) = 0$.

b) В силу независимости $P(\cap_{m=n}^{\infty} A_m^c) = \prod_{m=n}^{\infty}(1 - P(A_m))$. Заметим, что

$$\ln P(\cap_{m=n}^{\infty} A_m^c) = \sum_{m=n}^{\infty} \ln(1 - P(A_m)) \leq -\sum_{m=n}^{\infty} P(A_m) = -\infty.$$

Следовательно, $P(\cap_{m=n}^{\infty} A_m^c) = 0$. Тогда $0 = P(\cup_{n=1}^{\infty}\cap_{m=n}^{\infty} A_m^c) = P(A^c) = 1 - P(A)$. □

Определение 1.24. *Пусть $\{\xi_n\}$ — последовательность с.в.. Обозначим через $\mathcal{F}^{(n)}$ σ-алгебру, порожденную с.в. $\{\xi_n, \xi_{n+1}, \cdots\}$. Положим*

$$\mathcal{F}^{\infty} = \cap_{n=1}^{\infty}\mathcal{F}^{(n)}.$$

Система множеств $\mathcal{F}^{\infty}$ является пересечением σ-алгебр, следовательно, сама является σ-алгеброй.

Пример 1.25. *Следующие события принадлежат $\mathcal{F}^{\infty}$.*

$$\Big\{\overline{\lim}_n \xi_n \leq c\Big\}, \quad \Big\{\frac{S_n}{n} \text{ сходится}\Big\}.$$

Теорема 1.26. (Закон нуля и единицы. А.Н. Колмогоров.) *Пусть $\{\xi_i\}$ — независимые с.в. с $\mathbb{E}\xi_i = 0$. Тогда для любого события $A \in \mathcal{F}^{\infty}$ имеем $P(A) = 0$ либо $P(A) = 1$.*

Доказательство. Так как A принадлежит σ-алгебре $\mathcal{F}^{(n)}$, порожденной $\{\xi_{n+1}, \xi_{n+2}, \cdots\}$, то A не зависит от σ-алгебры $\mathcal{F}_n$, порожденной $\{\xi_1, \cdots, \xi_n\}$. Но объединение $\mathcal{F}_n$ порождает всю σ-алгебру цилиндрических множеств. Значит, A не зависит от самого себя. Следовательно $P(A) = P(AA) = P^2(A)$. Отсюда вытекает утверждение теоремы. □

Отсюда следует, что события $\{\sum_{n=1}^{\infty}\xi_n$ сходится$\}$, $\{\lim_n \xi_n$ существует$\}$ имеют вероятность ноль либо единица для независимых с.в. $\{\xi_n\}$.

1.3. **Условное математическое ожидание.** Всюду далее $(\Omega, \mathcal{F}, P)$ — вероятностное пространство, а $\mathcal{B} \subset \mathcal{F}$ — σ-алгебра, содержащаяся в $\mathcal{F}$.

Замечание 1.27. *Мы будем считать, что мера P является полной. То есть, если $A \subset B \in \mathcal{F}$ и $P(B) = 0$, то $P(A) = 0$.*

Любое пространство с мерой можно сделать полным, приписав меру ноль всем подмножествам множеств нулевой меры (проверьте, что операция пополнения корректно определена).

Условное математическое ожидание относительно события

Простейшей версией понятия условного математического ожидания является понятие условного математического ожидания относительно события B ненулевой вероятности:

$$\mathbb{E}(\xi|B) = \frac{\mathbb{E}(\xi \cdot I_B)}{P(B)}.$$

Для $\xi = I_A$ математическое ожидание $\mathbb{E}(\xi|B)$ равно условной вероятности $P(A|B)$.

Несложно проверить следующее свойство :

$$\mathbb{E}\xi = \sum_{i=1}^{n} \mathbb{E}(\xi|B_i)P(B_i),$$

если $\{B_i\}$ — разбиение Ω на непересекающиеся множества B_i.

Условное математическое ожидание относительно разбиений

Пусть дано разбиение $\mathcal{B}$ вероятностного пространства Ω на конечное число непересекающихся множеств положительной меры.

$$\Omega = \cup_{i=1}^{n} B_i,\ B_i \cap B_j = \emptyset,\ i \neq j.$$

Условным математическим ожиданием с.в. ξ относительно этого разбиения называется случайная величина

$$\mathbb{E}(\xi|\mathcal{B}) = \sum_{i=1}^{n} \frac{\mathbb{E}(\xi \cdot I_{B_i})}{P(B_i)} I_{B_i}(\omega) = \sum_{i=1}^{n} \mathbb{E}(\xi|B_i) I_{B_i}(\omega). \tag{1}$$

Замечание 1.28. *Нетрудно проверить следующие свойства условных математических ожиданий*

1)
$$\mathbb{E}(\mathbb{E}(\xi|\mathcal{B})) = \mathbb{E}\xi.$$

2) *Если $\xi = I_A$, то*
$$\mathbb{E}(I_A|\mathcal{B}) = \sum_{i=1}^{n} P(A|B_i) I_{B_i}(\omega)$$

3) *1) + 2) влечет формулу полной вероятности*
$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

4) *Для любого множества B_i*
$$\mathbb{E}(\mathbb{E}(\xi|\mathcal{B}) I_{B_i}) = \mathbb{E}(\xi I_{B_i}).$$

Из свойства 4) следует, что

$$\mathbb{E}(\mathbb{E}(\xi|\mathcal{B})\varphi) = \mathbb{E}(\xi\varphi) \tag{2}$$

для любой функции φ, измеримой относительно разбиения.

Условное математическое ожидание как проекция

Пусть ξ — случайная величина с $\mathbb{E}(\xi^2) < \infty$. Условное математическое ожидание можно характеризовать следующим образом: для любой $\mathcal{B}$-измеримой функции φ

$$\mathbb{E}(\xi - \varphi)^2 \geq \mathbb{E}(\xi - \mathbb{E}(\xi|\mathcal{B}))^2,$$

т.е. минимум функционала $\varphi \to \mathbb{E}(\xi - \varphi)^2$ среди $\mathcal{B}$-измеримых функций достигается на $\mathbb{E}(\xi|\mathcal{B})$.

Действительно,

$$\mathbb{E}(\xi - \varphi)^2 = \mathbb{E}(\xi - \mathbb{E}(\xi|\mathcal{B}))^2 + 2\mathbb{E}((\xi - \mathbb{E}(\xi|\mathcal{B}))(\mathbb{E}(\xi|\mathcal{B}) - \varphi)) + \mathbb{E}(\mathbb{E}(\xi|\mathcal{B}) - \varphi)^2$$

В силу свойства (2) среднее слагаемое равно нулю, поэтому

$$\mathbb{E}(\xi - \varphi)^2 = \mathbb{E}(\xi - \mathbb{E}(\xi|\mathcal{B}))^2 + \mathbb{E}(\mathbb{E}(\xi|\mathcal{B}) - \varphi)^2 \geq \mathbb{E}(\xi - \mathbb{E}(\xi|\mathcal{B}))^2.$$

Вывод: условное математическое ожидание относительно σ-алгебры $\mathcal{B}$ является ортогональной (в смысле $L^2(P)$) проекцией на подпространство σ-измеримых случайных величин.

Условное математическое ожидание относительно σ-алгебры.

Пусть ξ — неотрицательная с.в. со свойством $\mathbb{E}|\xi| < \infty$, а $\mathcal{B} \subset \mathcal{F}$ — некоторая σ-алгебра, содержащаяся в $\mathcal{F}$.

Определение 1.29. *Случайная величина $\mathbb{E}(\xi|\mathcal{B})$ называется условным математическим ожиданием ξ относительно $\mathcal{B}$, если*

1) $\mathbb{E}(\xi|\mathcal{B})$ *измерима относительно σ-алгебры $\mathcal{B}$ и* $\mathbb{E}|\mathbb{E}(\xi|\mathcal{B}))| < \infty$
2) *Для любой ограниченной функции φ, измеримой относительно $\mathcal{B}$, выполнено равенство*

$$\mathbb{E}(\mathbb{E}(\xi|\mathcal{B})\varphi) = \mathbb{E}(\xi\varphi).$$

Пример 1.30. *Для σ-алгебры $\mathcal{B}$, порожденной конечным (или даже счетным) набором множеств, формула (1) задает условное математическое ожидание относительно $\mathcal{B}$. Это немедленно вытекает из пункта 4) Замечания 1.28 и измеримости функции, заданной формулой (1), относительно $\mathcal{B}$.*

Еще один важный пример вытекает из теоремы Фубини.

Пример 1.31. *Пусть $\Omega = \mathbb{R}^n$ и $P = \prod_{i=1}^n \mu_i$ — прямое произведение вероятностных мер. Обозначим через $\mathcal{F}_m$ σ-алгебру, порожденную первыми m координатами. Функции, измеримые относительно $\mathcal{F}_m$ — это борелевские функции вида $g(x_1, \cdots, x_m)$. Тогда для любой интегрируемой борелевской функции f*

$$\mathbb{E}(f|\mathcal{F}_m) = \int f(x_1, \cdots, x_m, x_{m+1}, \cdots, x_n)\ d\mu_{m+1}(x_{m+1}) \cdots d\mu_n(x_n).$$

Существование условного математического ожидания

Пусть теперь $\mathcal{B}$ — произвольная σ-алгебра, содержащаяся в $\mathcal{F}$ и ξ — неотрицательная интегрируемая с.в. Рассмотрим меру $\xi \cdot P$. Эта мера абсолютно непрерывна относительно ограничения μ на σ-алгебру $\mathcal{B}$. По теореме Радона-Никодима существует такая интегрируемая функция g, измеримая относительно пополнения $\mathcal{B}$ относительно P,

$$\mathbb{E}(I_A\xi) = \int I_A \cdot \xi\ dP = \int I_A g\ dP = \mathbb{E}(I_A g)$$

для любого множества $A \subset \mathcal{B}$. Выбрав $\mathcal{B}$-измеримую модификацию, мы получим условное математическое ожидание ξ. Произвольную интегрируемую с.в. ξ надо представить в виде $\xi = \xi_+ - \xi_-$ и положить $\mathbb{E}(\xi|\mathcal{B}) = \mathbb{E}(\xi_+|\mathcal{B}) - \mathbb{E}(\xi_-|\mathcal{B})$.

Теорема 1.32. *Выполнены следующие свойства*

1) *Условное математическое ожидание единственно с точностью до множества меры нуль (относительно сужения P на $\mathcal{B}$!).*
2) $\mathbb{E}(\xi|\mathcal{B}) = \xi$ *почти наверное (т.е. P-п.в.) для всякой с.в. ξ, измеримой относительно $\mathcal{B}$.*
3) $\mathbb{E}(\xi|\mathcal{B}) \geq 0$ *п.н. для всякой с.в. ξ, т.ч. $\xi \geq 0$ п.н.*
4) *Для всякой ограниченной $\mathcal{B}$-измеримой g*

$$\mathbb{E}(\xi \cdot g|\mathcal{B}) = g\mathbb{E}(\xi|\mathcal{B}).$$

Условные меры

В общем случае условное математическое ожидание не обладает явным представлением. Тем не менее, мы разберем важный пример, когда его можно вычислить явно.

Пусть $P = \rho(x, y)\ dxdy$ — вероятностная мера на $\mathbb{R}^2$ с положительной борелевской плотностью. Рассмотрим σ-алгебру $\mathcal{B}_x$, порожденную координатной функцией x. Пусть P_x — проекция (маргинальное распределение) P на ось Ox:

$$P_x(A) = \int_{\mathbb{R}^2} I_A(x)\rho(x, y)\ dxdy.$$

Она имеет плотность $\rho_x(x)$:

$$\rho_x(x) = \int_{\mathbb{R}} \rho(x, y)\ dy.$$

Условной мерой P^x для каждого фиксированного x является вероятностная мера на прямой l_x, проходящей ортогонально Ox через точку $(x, 0)$, заданная плотностью

$$P^x = \rho^x(y)dy = \frac{\rho(x, y)}{\rho_x(y)} \cdot dy = \frac{\rho(x, y)}{\int_{\mathbb{R}} \rho(x, y)dy} \cdot dy.$$

Очевидно, P^x — вероятностная мера.

Мера P восстанавливается по проекции и набору условных мер

$$P(dx\ dy) = \int P^x(dy) \oplus P_x(dx).$$

Последнее равенство означает, что для любой интегрируемой функции φ выполнено соотношение

$$\int_{\mathbb{R}^2} \varphi(x, y)P(dx\ dy) = \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} \varphi(x, y)P^x(dy) \Big)\ P_x(dx).$$

Условное математическое ожидание функции ξ относительно $\mathcal{B}_x$ оказывается равным интегралу по условной мере и задается формулой

$$g(x) = \mathbb{E}(\xi|\mathcal{B}_x) = \int \xi(x, y)\rho^x(y)dy = \frac{\int \xi(x, y)\rho(x, y)\ dy}{\int \rho(x, y)\ dy}.$$

Действительно, эта функция $\mathcal{B}_x$-измерима (точнее, измерима относительно пополнения $\mathcal{B}_x$ по мере Лебега). Для функции $\varphi(x)$, зависящей только от x, имеем (по теореме Фубини)

$$\mathbb{E}(g\varphi) =$$

$$= \int_{\mathbb{R}^2} \varphi(x) \frac{\int_{\mathbb{R}} \xi(x,y)\rho(x,y)\ dy}{\int_{\mathbb{R}} \rho(x,y)\ dy} \rho(x,y)\ dxdy\ = \int_{\mathbb{R}} \Bigl(\int_{\mathbb{R}} \varphi(x) \frac{\int_{\mathbb{R}} \xi(x,y)\rho(x,y)\ dy}{\int_{\mathbb{R}} \rho(x,y)\ dy}\ \rho(x,y) dy \Bigr) dx$$

$$= \int_{\mathbb{R}} \Bigl(\varphi(x) \frac{\int_{\mathbb{R}} \xi(x,y)\rho(x,y)\ dy}{\int_{\mathbb{R}} \rho(x,y)\ dy} \int_{\mathbb{R}} \rho(x,y) dy \Bigr) dx = \int_{\mathbb{R}} \varphi(x) \int_{\mathbb{R}} \xi(x,y)\rho(x,y)\ dy\ dx$$

$$= \int_{\mathbb{R}^2} \varphi(x)\xi(x,y)\rho(x,y)\ dxdy = \mathbb{E}(\xi\varphi).$$

2. Лекция 2. Марковские цепи

2.1. **Базовые сведения.** Настоящая лекция имеет также обзорный характер, поскольку в ней представлены базовые сведения о марковских цепях, с которыми большинство слушателей, как правило, знакомо. Речь будет идти в основном об однородных конечных цепях.

Определение 2.1. *Последовательность случайных величин $\{X_n\}$, $n \ge 0$ со значениями в некотором не более чем счетном множестве S (множестве состояний) называется (однородной) марковской цепью, если для любых значений $k_i \in S$ выполнено соотношение*

$$P(X_{n+1} = k | X_0 = k_0, \cdots, X_n = k_n) = P(X_{n+1} = k | X_n = k_n) = P(X_1 = k | X_0 = k_n)$$

(будущее зависит от прошлого через настоящее).

Пример 2.2. *Пусть $X_n = \sum_{i=1}^n \xi_i$, где $\{\xi_n\}$ — последовательность независимых бернуллиевских с.в. Тогда $\{X_n\}$ — марковская цепь. Действительно,*

$$\begin{aligned} P(X_{n+1} = k | X_0 = k_0, \cdots, X_n = k_n) &= P(X_{n+1} - X_n = k - k_n | X_0 = k_0, \cdots, X_n = k_n) \\ &= P(X_{n+1} - X_n = k - k_n) = P(\xi_i = k - k_n) = P(X_1 = k | X_0 = k_n). \end{aligned}$$

Замечание 2.3. *Почти везде ниже мы будем считать, что S конечно : $|S| = d < \infty$.*

Величина

$$p_{ik} = P(X_1 = k | X_0 = i)$$

называется переходной вероятностью. Ее можно интерпретировать как вероятность перехода цепи из состояния i в состояние k за единицу времени. Матрица P: $(P)_{ik} = p_{ik}$ называется матрицей перехода. Очевидно следующее свойство:

$$\sum_{k \in S} p_{ik} = 1.$$

Из определения марковской цепи видно, что распределение X_n полностью определяется X_0 и матрицей P.

Пример 2.4. *Простейший пример: марковская цепь с двумя состояниями, матрица перехода которой задается так:*

$$\begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

где $0 < \alpha, \beta < 1$.

Часто марковскую цепь удобно изображать в виде графа, где состояния соединены направленными стрелками (направление стрелки от состояния i к состоянию k соответствуют вероятности p_{ik}).

Упражнение 2.5. *([3], пример 1.1.2). Изобразите граф, соответствующий матрице перехода*

$$\begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Упражнение 2.6. *Докажите, что*

$$P(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n) = P(X_0 = i_0) p_{i_0 i_1} \cdots p_{i_{n-1} i_n}.$$

Предположим теперь, что распределение с.в. X_0 задано вектором (строкой) λ:

$$P(X_0 = i) = \lambda_i, \ \lambda = (\lambda_1, \cdots, \lambda_d)$$

Тогда, как нетрудно видеть

$$P(X_1 = k) = \sum_{i \in S} P(X_1 = k | X_0 = i) P(X_0 = i) = \sum_{i \in S} \lambda_i p_{ik}.$$

Т.е., распределение вероятностей с.в. X_1 задается вектором

$$\lambda P$$

(умножение происходит слева!). Рассуждая по индукции, мы приходим к выводу, что распределение X_n задается вектором

$$\lambda P^n.$$

Пример 2.7. *([3], пример 1.16.18) Блоха прыгает по вершинам треугольника, перепрыгивая на одну из свободных вершин с вероятностью 1/2. Найти вероятность, что через n прыжков она окажется на месте старта.*

Матрица переходных вероятностей равна

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

Искомая вероятность равна диагональному элементу матрицы P^n. С помощью системы собственных векторов матрицы P найдите P^n и докажите, что искомая вероятность равна $\frac{1}{3} + \frac{2}{3}\left(-\frac{1}{2}\right)^n$.

Упражнение 2.6 обобщается следующим образом.

Теорема 2.8.

$$P(X_{k_1} = i_1, \cdots, X_{k_n} = i_n) = (\lambda P^{k_1})_{i_1} (P^{k_2 - k_1})_{i_1 i_2} \cdots (P^{k_n - k_{n-1}})_{i_{n-1} i_n}.$$

2.2. **Регулярность.**

Определение 2.9. *Марковская цепь называется регулярной, если существует такое n_0, что $(P^{n_0})_{ij} > 0$ для любых $i, j \in S$.*

Определение 2.10. *Марковская цепь называется неприводимой, если для любых $i, j \in S$ существует такое n_0, что $(P^{n_0})_{ij} > 0$.*

Пример 2.11. $P = \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}$, $Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, *цепь регулярна для* P *и приводима и нерегулярна для* Q.

Определение 2.12. *Для состояний* $i \neq k$ *обозначим через* T_{ik} *первый момент достижения* k *для цепи, стартовавшей из* i*:* $T_{ik} = \min\{n \geq 0 : X_n = k | X_0 = i\}$, *а через* $\mu_{ik} = \mathbb{E} T_{ik}$ *— среднее время достижения* k.

Для состояния i *обозначим через* T_i *первый момент возвращения в* i *для цепи, стартовавшей из* i*:* $T_i = \min\{n > 0 : X_n = i | X_0 = i\}$, *а через* $\mu_i = \mathbb{E} T_i$ *— среднее время возвращения в* i.

Нетрудно придумать примеры, когда $P(T_{ik} = \infty) > 0$ ([8], пример 9.3.6). Такая ситуация осуществима, например, благодаря так называемым поглощающим состояниям. Тем не менее, имеет место следующая теорема.

Теорема 2.13. *([8], теорема 9.3.15). Если цепь регулярна, то* T_{ik} *имеет конечное среднее. Более того,* $P(T_{ik} > n) \leq c\lambda^n$ *для некоторых констант* $c > 0, \lambda < 1$.

Упражнение 2.14. *([3], пример 1.3.6.) Лягушка прыгает вверх по лестнице из* N *ступеней. При прыжке с подножия лестницы (нулевая ступень) она с вероятностью* $\beta >$ *оказывается на первой ступени, а с вероятностью* $1 - \beta$ *остается на месте. После прыжка на других ступенях она с вероятностью* $\alpha > 0$ *оказывается на ступень выше, с вероятностью* α *на ступень ниже, с вероятностью* $1 - 2\alpha$ *остается на месте. Докажите, что среднее количество прыжков, которое она совершит, прежде чем достигнет лестницы, равно* $\frac{N(N-1)}{2\alpha} + \frac{N}{\beta}$.

В дальнейшем нам понадобится следующая версия *сильного марковского свойства*

Теорема 2.15. *(Сильное марковское свойство) Пусть* T *— первый момент достижения цепью состояния* d*. Выполнено следующее соотношение:*

$$P(X_{T+m} = k | X_r = x_r,\ 1 \leq r \leq T,\ X_T = d) = (P^m)_{dk}$$

для $x_r \neq d$.

Доказательство. Обозначим событие $\{X_r = x_r,\ 0 \leq r < T\}$ через $A(T)$. Тогда

$$P(X_{T+m} = k | X_r = x_r,\ 0 \leq r \leq T,\ X_T = d) = \frac{P(X_{T+m} = k, A(T), X_T = d)}{P(A(T), X_T = d)}.$$

$$P(X_{T+m} = k, A(T), X_T = d) = \sum_{t=1}^{\infty} P(X_{t+m} = k, A(t), T = t, X_t = d)$$
$$= \sum_{t=1}^{\infty} P(X_{t+m} = k | A(t), T = t, X_t = d) P(A(t), T = t, X_t = d) = \sum_{t=1}^{\infty} (P^m)_{dk} P(A(t), T = t, X_t = d)$$
$$= (P^m)_{dk} P(A(T), X_T = d).$$

□

2.3. **Существование и единственность инвариантных распределений.**

Определение 2.16. *Вероятностное распределение* $\pi = (\pi_1, \cdots, \pi_d))$ *на* S *называется* **инвариантным** *или* **стационарным** *относительно цепи* $\{X_n\}$, *если* $\pi P = \pi$, *где* P *— соответствующая матрица перехода.*

Упражнение 2.17. *Найдите инвариантные распределения для* $P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$.

Теорема 2.18. *Конечная марковская цепь обладает инвариантным распределением* π *(вообще говоря, неединственным).*

Теорема вытекает из теоремы о неподвижной точке. Доказательство остается в качестве упражнения.

Для конечной регулярной марковской цепи рассмотрим произвольное состояние $s \in S$. Предположим, что $X_0 = s$. Пусть $\rho_k(s)$ — среднее число посещений состояния k между возвращением в состояние s (считаем, что $\rho_s(s) = 1$).

Теорема 2.19. *Мера* $\pi_k = \frac{\rho_k(s)}{\mu_s}$ *— инвариантная*

Доказательство. Здесь $\mu_s = \mathbb{E}T_s$ — среднее время возвращения в состояние s.

Зафиксируем $k \in S, k \neq s$. Пусть $I_n = \{X_n = k, X_i \neq s, 0 < i \leq n\} = \{X_n = k, T_s \geq n\}$. Общее число посещений k до возвращения в s равно

$$R_k = \sum_{n=1}^{\infty} I_n.$$

Очевидно, время возвращения в s равно

$$T_s = 1 + \sum_{k \neq s} R_k.$$

Возьмем математическое ожидание от обеих частей.

$$\mu_s = \sum_{k \in S} \rho_k(s).$$

По определению $\rho_k(s)$:

$$\rho_k(s) = \mathbb{E}R_k = \sum_{n=1}^{\infty} \mathbb{E}I_n = \sum_{n=1}^{\infty} P(X_n = k, T_s \geq n).$$

Очевидно, для $n = 1$ имеем: $P(X_1 = k, T_s \geq 1) = p_{sk}$. Для $n \geq 2$

$$P(X_n = k, T_s \geq n) = \sum_{j \neq s} P(X_n = k, X_{n-1} = j, T_s \geq n-1) =$$

$$= \sum_{j \neq s} P(X_n = k | X_{n-1} = j, T_s \geq n-1) P(X_{n-1} = j, T_s \geq n-1).$$

В силу сильного марковского свойства последнее выражение равно $\sum_{j \neq s} p_{jk} P(X_{n-1} = j, T_s \geq n-1)$. Отсюда следует, что

$$\rho_k(s) = p_{sk} + \sum_{j \neq s} p_{jk}\rho_j(s) = \sum_j p_{jk}\rho_j(s).$$

Разделив выражение на μ_s, получаем инвариантность π. □

Теорема 2.20. *Для регулярной марковской цепи инвариантное распределение* π *единственно, причем*

$$\pi_s = \frac{1}{\mu_s},$$

Доказательство. Напомним, что $T_{ik} = \min\{n \geq 0, X_n = k | X_0 = i\}$ $(T_{kk} = 0, \mu_{kk} = 0)$, $\mu_{ik} = \mathbb{E} T_{ik}$. Для $i \neq k$ в силу марковского свойства

$$\mu_{ik} = \sum_j p_{ij}(1 + \mu_{jk}) = 1 + \sum_j p_{ij}\mu_{jk}.$$

Аналогично, для времени возвращения

$$\mu_k = 1 + \sum_j p_{kj}\mu_{jk}.$$

Следовательно

$$\mu_{ik} + \delta_{ik}\mu_k = 1 + \sum_j p_{ij}\mu_{jk}$$

Тогда

$$\sum_i \pi_i \mu_{ik} + \sum_i \pi_i \delta_{ik}\mu_k = 1 + \sum_i \sum_j \pi_i p_{ij}\mu_{jk}.$$

В силу инвариантности последнее слагаемое равно $\sum_i \pi_i \mu_{ik}$. Таким образом $1 = \sum_i \pi_i \delta_{ik}\mu_k = \pi_k \mu_k$. □

2.4. **Эргодичность.** Замечательным обстоятельством является тот факт, что при весьма общих обстоятельствах при больших значениях n наблюдается сходимость процесса к инвариантному распределению (эргодичность).

Теорема 2.21. *Предположим, что существует предел* $\Pi_{ij} = \lim_{n\to\infty} p_{ij}(n)$. *Тогда каждая строка* $\pi^{(i)}$ *матрицы* Π *задает инвариантное распределение.*

Доказательство.

$$(\pi^{(i)}P)_j = \sum_l \pi_{il}p_{lj} = \lim_n \sum_l p^{(n)}_{il}p_{lj} = \lim_n p^{(n+1)}_{ij} = (\pi^{(i)})_j.$$

□

Теорема 2.22. (Эргодическая теорема). *Пусть* $\{X_n\}$ — *регулярная марковская цепь. Для любых* i, k *существует предел* $\pi_k = \lim_{n\to\infty} p_{ik}(n)$. *При этом* $\pi = (\pi_0, \cdots, \pi_d)$ *является инвариантным распределением.*

Доказательство. ([4], 1.12). Положим

$$m^{(n)}_j = \min_i (P^n)_{ij}, \;\; M^{(n)}_j = \min_i (P^n)_{ij}.$$

Заметим, что $m^{(n)}_j$ — неубывающая по n последовательность. Действительно

$$m^{(n+1)}_j = \min_i (P^{(n+1)})_{ij} = \min_i \sum_k p_{ik}(P^n)_{kj} \geq \min_i \sum_k p_{ik} \min_k (P^n)_{kj} = \min_i \sum_k p_{ik} m^{(n)}_j = m^{(n)}_j.$$

Аналогично $M^{(n)}_j$ — невозрастающая по n последовательность.

Достаточно доказать, что $M^{(n)}_j - m^{(n)}_j$ сходится к нулю. Пусть $\varepsilon = \min_{i,j}(P^{n_0})_{ij} > 0$.

$$\begin{aligned}(P^{n_0+n})_{ij} &= \sum_k (P^{n_0})_{ik}(P^n)_{kj} = \sum_k [(P^{n_0})_{ik} - \varepsilon (P^n)_{jk}](P^n)_{kj} + \varepsilon \sum_k (P^n)_{kj}(P^n)_{jk} \\ &= \sum_k [(P^{n_0})_{ik} - \varepsilon (P^n)_{jk}](P^n)_{kj} + \varepsilon P^{(2n)}_{jj}.\end{aligned}$$

В силу того, что $(P^{n_0})_{ik} - \varepsilon(P^n)_{jk} \geq 0$, получаем

$$(P^{n_0+n})_{ij} \geq \min_k (P^n)_{kj} \sum_k [(P^{n_0})_{ik} - \varepsilon(P^n)_{jk}] + \varepsilon P_{jj}^{(2n)} = (1-\varepsilon) m_j^{(n)} + \varepsilon P_{jj}^{(2n)}.$$

Следовательно

$$m_j^{(n+n_0)} \geq (1-\varepsilon) m_j^{(n)} + \varepsilon P_{jj}^{(2n)}.$$

Аналогично $M_j^{(n+n_0)} \leq (1-\varepsilon) M_j^{(n)} + \varepsilon P_{jj}^{(2n)}$. Следовательно

$$M_j^{(n+n_0)} - m_j^{(n+n_0)} \leq (1-\varepsilon)(M_j^{(n)} - m_j^{(n)}).$$

Отсюда следует утверждение теоремы (почему?). □

3. Лекция 3. Случайное блуждание. Комбинаторные задачи. Возвратность.

3.1. **Подсчет числа путей. Принцип отражения.** Материал: [7] (3.10).

(Несимметричное) случайное блуждание, стартующее их точки a.

$$S_n = a + X_1 + \ldots + X_n,$$

X_i — последовательность независимых случайных величин, принимающих значение -1 с вероятностью $0 < q < 1$ и значение 1 с вероятностью $p = 1 - q$.

Пусть $N_n(a,b)$ — число всевозможных путей со свойством $S_0 = a, S_n = b$, а $N_n^0(a,b)$ — число тех из них, которые содержат точку $(k,0)$ на оси x .

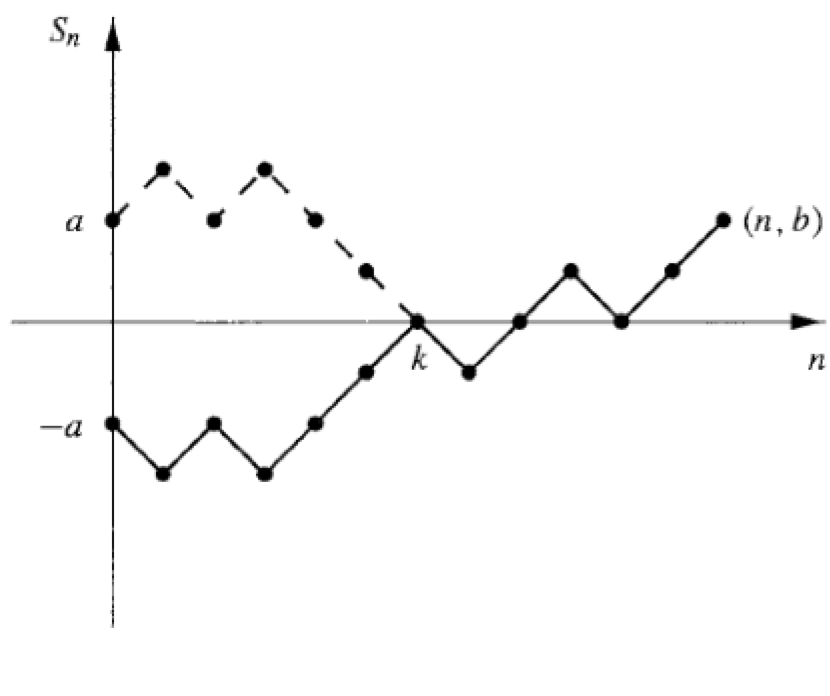
Упражнение 3.1. *Докажите, что*

$$N_n(a,b) = C_n^{\frac{1}{2}(n+b-a))},$$

если множество путей со свойством $S_0 = a, S_n = b$ непусто.

Теорема 3.2. (Принцип отражения) *Пусть $a, b > 0$. Тогда $N_n^0(a,b) = N_n(-a,b)$.*

Теорема доказывается путем установления взаимно однозначного соответствия между двумя множествами путей (см. рисунок).



3.2. **Распределение максимума случайного блуждания. Вероятность возвращения для одномерного (несимметричного) случайного блуждания.**

Теорема 3.3. *Пусть* $S_0 = 0$, $M_n = \max_{0\le r\le n} S_r$. *Тогда для любого* $r \ge 1$
$$P(M_n \ge r, S_n = b) = P(S_n = b),$$
если $b \ge r$ *и*
$$P(M_n \ge r, S_n = b) = (q/p)^{r-b} P(S_n = 2r - b),$$
если $b < r$.

Доказательство. Случай $b \ge r$ очевиден. Пусть $b \le r$. Воспользуемся принципом отражения. Найдем первую точку τ_r, в которой S достигает уровня $r : S_{\tau_r} = r$. Сопоставим пути S_n отраженный путь $\tilde{S}_n$: $S_n = \tilde{S}_n$, если $n \le \tau_r$ и $\tilde{S}_n = 2r - S_n$, если $n \ge \tau_r$. Нетрудно видеть, что таким способом задается взаимно однозначное соотношение между путями, оканчивающими в $2r - b$ и путями, оканчивающимися в b и максимумом $\ge r$. Убедитесь, что вероятность задается искомой формулой. □

Упражнение 3.4. *Докажите, что в симметричном случае* $p = q = 1/2$
$$P(M_n \ge r) = 2P(S_n \ge r + 1) + P(S_n = r).$$

Теорема 3.5. *Вероятность* $P(S_n = 0, \text{для некоторого } n > 1 \mid S_0 = 0)$ *того, что случайное блуждание вернется в точку старта равно* $1 - |p - q]$. *В частности, одномерное симметричное случайное блуждание возвратно.*

Доказательство. Найдем $P(S_{2n=0}, S_i \ne 0, 0 < i < 2n)$. Эта вероятность равна
$$P(S_1 = 1, S_i \ne 0, S_{2n-1} = 1, 2 \le i \le 2n-2) + P(S_1 = -1, S_i \ne 0, S_{2n-1} = -1, 2 \le i \le 2n-2).$$
Применяя принцип отражения, получаем:
$$P(S_1 = 1, S_i \ne 0, S_{2n-1} = 1, 2 \le i \le 2n-2) = pqP(S_{2n-2} = 1, S_i \ne 0, 1 \le i \le 2n-2 | S_0 = 1)$$
$$= pq\Big(P(S_{2n-2} = 1 | S_0 = 1) - P(S_{2n-2} = 1, \exists i : 1 \le i \le 2n-2, S_i = 0 | S_0 = 1)\Big)$$
$$= pq\Big((pq)^{n-1} C^{n-1}_{2n-2} - (pq)^{n-1} N^0_{2n-2}(1,1)\Big) = (pq)^n \Big(C^{n-1}_{2n-2} - N_{2n-2}(1,-1)\Big)$$
$$= (pq)^n \Big(C^{n-1}_{2n-2} - C^{n-2}_{2n-2}\Big)$$
$$= (pq)^n (2n-2)! \Big(\frac{1}{((n-1)!)^2} - \frac{1}{n!(n-2)!}\Big) = (pq)^n \frac{(2n-2)!}{(n-1)!(n-2)!}\Big(\frac{1}{n-1} - \frac{1}{n}\Big)$$
$$= (pq)^n \frac{(2n-2)!}{n!(n-1)!} = \frac{p^n q^n}{2n-1} C^n_{2n-1}.$$
Аналогичная формула имеет место для $P(S_1 = -1, S_i \ne 0, S_{2n-1} = -1, 2 \le i \le 2n-2)$. Итак
$$P(S_{2n=0}, S_i \ne 0, 0 < i < 2n) = \frac{2p^n q^n}{2n-1} C^n_{2n-1}.$$
Таким образом
$$P(S_i \ne 0, i > 0 | S_0 = 0) = \sum_{i=1}^{\infty} \frac{2p^n q^n}{2n-1} C^n_{2n-1}.$$
Упражнение: проверьте, что если $|x| < 1$, то
$$1 - \sqrt{1-x} = \sum_{n=1}^{\infty} 2\Big(\frac{x}{4}\Big)^n \frac{(2n-2)!}{n!(n-1)!}.$$

Отсюда вытекает, что $P(S_{2n=0}, S_i \neq 0, 0 < i < 2n) = 1 - \sqrt{1-4qp} = 1 - |p-q|$. $\square$

3.3. **Возвратность двумерного случайного блуждания. Невозвратность трехмерного случайного блуждания.** Материал: [3].

Мы обсудим общий способ доказательства возвратности/невозвратности состояния марковской цепи со счетным числом состояний.

Пусть цепь X_n стартует из состояния x_0. Положим

$$I_k = \{\omega : X_n(\omega) = x_0, n > 1 \text{ хотя бы для } k \text{ значений } n\}.$$

Из сильного марковского свойства вытекает

$$P(I_k) = (P(I_1))^k$$

(проверьте!). Отсюда сразу следует, что вероятность возвращения бесконечное число раз равна $\lim_k (P(I_1))^k$, т.е. 1, если $P(I_1) = 1$ (состояние возвратно) или 0, если $P(I_1) < 1$ (состояние невозвратно).

Пусть $J_n = I_{X_n = x_0}$. Тогда случайная величина V, равная общему числу возвращений равна

$$V = \sum_{n=1}^{\infty} J_n.$$

Поэтому

$$\mathbb{E}V = \sum_{n=1}^{\infty} P(X_n = x_0) = \sum_{n=1}^{\infty} P^n_{x_0 x_0}.$$

С другой стороны

$$\mathbb{E}V = \sum_{n=0}^{\infty} P(V > n) = \sum_{n=0} (P(I_1))^{n+1}.$$

Таким образом невозвратность эквивалентна сходимости ряда $\sum_{n=0} (P(I_1))^{n+1}$, что в свою очередь эквивалентно сходимости ряда $\sum_{n=1}^{\infty} P^n_{x_0 x_0}$.

Упражнение 3.6. *Пользуясь этим признаком, докажите возвратность симметричного одномерного случайного блуждания и невозвратность несимметричного.*

Теорема 3.7. *Двумерное симметричное случайное блуждание возвратно, а трехмерное симметричное случайное блуждание невозвратно.*

4. Лекция 4. Слабая сходимость мер. Теорема Прохорова.

4.1. **Определения и свойства слабой сходимости.** Приложения слабой сходимости мер не исчерпываются вероятностными задачами. По своей природе это понятие функционально-аналитическое, поскольку может рассматриваться как *-слабая сходимость на пространстве мер, но оно также широко используется в вариационных задачах и уравнениях в частных производных. Фундаментальные результаты о слабой сходимости мер были получены в работах выдающегося геометра А.Д. Александрова, при этом они были мотивированы его работами по теории выпуклых поверхностей.

Пример 4.1. *Рассмотрим δ-образную последовательность функций $f_\delta(x) = \frac{1}{\sqrt{2\pi\delta^2}} e^{-\frac{x^2}{2\delta^2}}$ на прямой. Каждая функция удовлетворяет условию $\int_{\mathbb{R}} f_\delta(x)\, dx = 1$. Очевидно, $\lim_{\delta\to 0} f_\delta(x) = 0$ если $x \neq 0$ и $\lim_{\delta\to 0} f_\delta(0) = +\infty$. Несмотря на то, что $\lim_{\delta\to 0} f_\delta(x) = 0$ почти всюду, функция, тождественно равная нулю, не является разумным пределом $\{f_\delta\}$. Действительно, как хорошо известно из анализа (и нетрудно проверить самостоятельно), $\lim_{\delta\to 0} \int \varphi(x) f_\delta(x)\, dx = \varphi(0)$ для непрерывной ограниченной* *функции φ. Очевидно, в этом примере не выполняется свойство о предельном переходе под знаком интеграла $\lim_{\delta\to 0}\int \varphi f_\delta(x)\,dx \neq 0$. Естественным решением данной проблемы является следующее: рассматривать $\{f_\delta\}$ как последовательность* **мер** *$\{f_\delta(x)\,dx\}$. Пределом этой последовательности является мера Дирака δ_0, сосредоточенная в точке 0, т.е. мера, определенная условием $\delta(A) = 0$, если $0 \in A$ и $\delta(A) = 1$, если $0 \in A$. Очевидно, $\int f(x)\,d\delta = f(0)$.*

Естественным классом пространств, на которых удобно рассматривать слабую сходимость, является класс так называемых польских пространств.

Определение 4.2. *Польскими пространствами называются полные сепарабельные метрические пространства.*

Мы будем рассматривать борелевские вероятностные меры на польском пространстве X, то есть, меры, определенные на борелевской σ-алгебре $\mathcal{B}(X)$. Метрику X будем обозначать через ρ.

Определение 4.3. *Пусть $\{\mu_n\}$ — последовательность вероятностных мер. Будем говорить, что $\{\mu_n\}$ сходится слабо к вероятностной мере $\{\mu\}$, если для любой непрерывной ограниченной функции $f : X \to \mathbb{R}$ выполнено соотношение*

$$\lim_{n\to\infty}\int_X f\ d\mu_n = \int_X f\ d\mu.$$

Теорема 4.4. *(А.Д. Александров) Последовательность вероятностных мер $\{\mu_n\}$ слабо сходится к вероятностной мере μ тогда и только тогда, когда выполнено одно из условий*

1)
$$\mu(G) \geq \overline{\lim}_n \mu_n(G)$$
для любого замкнутого G

2)
$$\mu(O) \leq \underline{\lim}_n \mu_n(O)$$
для любого открытого O

3) *Для любого множества B, удовлетворяющего условию $\mu(\partial B) = 0$*
$$\mu(B) = \lim_n \mu_n(B).$$

Доказательство. Предположим, что $\mu_n \to \mu$ слабо. Пусть G — замкнутое множество. Рассмотрим функцию $x \to \rho(x, G) = \inf_{y\in G}\rho(x, y)$. Заметим, что эта функция непрерывна (почему?). Положим $f_m(x) = (1 + m\rho(x, G))^{-1}$. Заметим, что $f_m(x) \geq I_G(x)$. По определению слабой сходимости имеем

$$\int f_m(x)\ d\mu = \lim_n \int f_m(x)\ d\mu_n \geq \overline{\lim}_n \mu_n(G).$$

Переходя к пределу по m и пользуясь теоремой Лебега об ограниченной сходимости, получаем $\lim_m \int f_m(x)\ d\mu = \mu(G)$, следовательно

$$\mu(G) \geq \overline{\lim}_n \mu_n(G).$$

Очевидно, свойства 1) и 2) эквивалентны.

Далее, 3) легко следует из 1) + 2), если учесть, что $\mathrm{Int}B = B \setminus \partial B \subset B \subset \overline{B}$ и

$$\mu(\mathrm{Int}B) = \mu(B) = \mu(\overline{B}),$$

если $\mu(\partial B) = 0$.

Для вывода слабой сходимости из свойства 3) заметим, что достаточно доказать свойство
$$\lim_n \int f(x)\ d\mu_n = \int f(x)\ d\mu$$
для произвольной *неотрицательной* функции $f(x)$ Пусть $M = \sup_X f(x)$. Применим известную формулу
$$\int f(x)\ d\mu_n = \int_0^M \mu_n(\{f(x) > t\})\ dt.$$
В силу непрерывности $f(x)$ имеем $\partial\{x : f(x) > t\} = \{x : f(x) = t\}$. Так как множества $\{x : f(x) = t\}$ не пересекаются при различных t, то $\mu(\{x : f(x) = t\}) > 0$ только для не более чем счетного набора значений t. Для всех остальных $\lim_n \mu_n(\{f(x) > t\}) = \mu(\{f(x) > t\})$. Переходя к пределу под знаком интеграла (пользуясь теоремой об ограниченной сходимости), получаем
$$\lim_n \int f(x)\ d\mu_n = \lim_n \int_0^M \mu_n(\{f(x) > t\})\ dt = \int_0^M \mu(\{f(x) > t\})\ dt = \int f(x)\ d\mu.$$
□

4.2. **Теорема Прохорова.** Следующий факт известен как теорема Улама.

Теорема 4.5. *Пусть μ — вероятностная борелевская мера на польском пространстве X. Для любого $\varepsilon > 0$ существует такое компактное множество $K \subset X$, что $\mu(X \setminus K) \le \varepsilon$.*

Доказательство. Для любого ε и любого $n \in \mathbb{N}$ найдем конечное объединение $A_n = \cup_{i=1}^{N(n)} B_{1/n}(x_i)$ замкнутых шаров радиса $1/n$, удовлетворяющих условию $\mu(X \setminus A_n) < \frac{\varepsilon}{2^n}$. Это можно сделать в силу того, что все пространство является X является счетным объединением таких шаров. Заметим, что A_n обладает конечной $1/n$-сетью (центры шаров из A_n). Пусть $K = \cap_{n=1}^{\infty} A_n$. Так как $K_n \subset A_n$, то K обладает конечной ε - сетью для любого $\varepsilon > 0$. Поэтому K — вполне ограниченное, замкнутое множество в полном метрическом пространстве. Следовательно, оно является компактом. Кроме того, $\mu(X \setminus K) = \mu(\cup_{n=1}^{\infty} A_n^c) \le \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon$. □

Определение 4.6. *Слабой топологией на пространстве вероятностных мер называется топология, порожденная множествами вида $O_{\mu,f,\varepsilon} = \big\{\nu : |\int_X f d\nu - \int_X f d\mu| < \varepsilon\big\}$, f — непрерывная ограниченная функция.*

Будем говорить, что множество $\mathcal{M}$ вероятностных мер слабо компактно, если из любой последовательности $\{\mu_n\} \subset \mathcal{M}$ можно выделить слабо сходящуюся подпоследовательности. Строго говоря, это свойство следует называть (относительной) слабой секвенциальной компактностью. Впрочем (см. задачи), слабая сходимость на метрических пространствах метризуема. Следовательно, в нашей ситуации эти понятия эквивалентны. Семейство мер $\mathcal{M}$ вероятностных мер на польском пространстве X называется плотным, если для любого $\varepsilon > 0$ существует такой компакт $K \subset X$, что $\mu(X \setminus K) \le \varepsilon$ для любой меры $\mu \in \mathcal{M}$.

Следующий результат (теорема Прохорова) играет исключительно важную роль в приложениях теории слабой сходимости мер. Для его доказательства понадобится еще один классический результат функционального анализа.

Теорема 4.7. *(Ф. Рисс) Пусть K — метрический компакт, $C(K)$ — пространство непрерывных функций на K, наделенное равномерной нормой. Для любого непрерывного линейного функционала $L : C(K) \to \mathbb{R}$ существует такая борелевская (в общем случае, знакопеременная) мера μ, что имеет место представление*

$$L(f) = \int_K f \ d\mu.$$

В случае, если L — неотрицательный функционал (т.е. $L(f) \geq 0$ для $f \geq 0$), то μ — неотрицательная мера.

Теорема 4.8. *(Ю.В. Прохоров) Семейство $\mathcal{M}$ вероятностных мер на польском пространстве X тогда и только тогда слабо компактно, когда оно плотно.*

Доказательство. Доказательство необходимости близко доказательству теоремы Улама. Пусть $\{x_i\}$ — счетное всюду плотное множество в X. Положим $A_{m,r} = \cup_{i \leq m} B_r(x_i)$, где $B_r(x_i) = \{x : |x - x_i| < r\}$ — открытые шары. Докажем, что

$$\lim_m \inf_{\mu \in \mathcal{M}} \mu(A_{m,r}) = 1$$

для любого $r > 0$. Предположим, что это не так. Зафиксируем $r > 0$. Тогда для некоторого $\varepsilon > 0$ и для любого m можно найти такую меру μ_m, что $\mu_m(A_{m,r}) < 1 - \varepsilon$. Пользуясь слабой компактностью $\mathcal{M}$ и переходя к подпоследовательностям, без ограничения общности можно считать, что $\{\mu_m\}$ сходится слабо к некоторой мере μ. Для меры μ существует такое натуральное число M, что $\mu(A_{M,r}) > 1 - \frac{\varepsilon}{2}$. Но в силу слабой сходимости $1 - \frac{\varepsilon}{2} < \mu(A_{M,r}) \leq \underline{\lim}_m \mu_m(A_{M,r}) \leq \underline{\lim}_m \mu_m(A_{m,r}) < 1 - \varepsilon$. Мы получили противоречие. Следовательно, для всех $r > 0$ $\lim_m \inf_{\mu \in \mathcal{M}} \mu(A_{m,r}) = 1$. Отсюда немедленно следует, что для любого $\varepsilon > 0$ и $N \in \mathbb{N}$ существует такое число $M(N)$, что

$$\mu(A_{M(N),1/N}) = \mu(\cup_{i \leq M(N)} B_{1/N}(x_i)) > 1 - \frac{\varepsilon}{2^N}$$

для любого $\mu \in \mathcal{M}$. Множество

$$\cap_{N \in \mathbb{N}} \cup_{i \leq M(N)} \overline{B}_{1/N}(x_i)$$

($\overline{B}_r(x_i) = \{x : |x - x_i| \leq r\}$) является искомым компактом.

Докажем достаточность. Предположим сначала, что $X = K$ — компакт. В пространстве $C(K)$ непрерывных функций на K найдем счетное всюду плотное множество функций $\{f_j\}$ (почему такое существует?). Так как для каждого f_j число

$$\int_K f_j \ d\mu$$

ограничено числом $\sup_K |f_j|$, то, пользуясь диагональным методом, можно найти последовательность $\{\mu_n\}$, т.ч. предел

$$\lim_n \int f_j \ d\mu_n = F(f_j)$$

существует. Заметим, что $|F(f_j)| \leq \sup_K |f_j|$ и $F(f_j) \geq 0$ для любой неотрицательной функции f_j. Следовательно, F продолжается по непрерывности до неотрицательного линейного функционала на $C(K)$. По теореме Рисса, существует вероятностная мера μ, обладающая свойством $F(f) = \int_K f \ d\mu$. При этом $\lim_n \int f_j \ d\mu_n = \int f_j \ d\mu$. Найдем произвольную непрерывную функцию f и приблизим ее по равномерной норме некоторой функцией f_j: $|f - f_j| < \varepsilon$. Тогда $\left| \int f \ d\mu - \lim_n \int f \ d\mu_n \right| \leq \left| \int f_j \ d\mu - \right.$

$\lim_n \int f_j \, d\mu_n\big| + \big|\int (f - f_j) \, d\mu\big| + \big|\lim_n \int (f - f_j) \, d\mu_n\big| \le 2\varepsilon$. В силу произвольности ε, получаем $\int f \, d\mu = \lim_n \int f \, d\mu_n$.

В случае некомпактного пространства X надо воспользоваться известным из общей топологии фактом, что X гомеоморфно вкладывается в некоторое компактное метрическое пространство Y. В качестве Y можно взять, например, $[0,1]^{\mathbb{N}}$. Таким образом, можно считать, что $X \subset Y$ и любая мера μ продолжается на все Y по формуле $\mu(A) = \mu(A \cap X)$, $A \subset Y$. Можно выделить последовательность мер $\{\mu_n\}$, слабо сходящуюся на Y к мере ν. Возьмем последовательность компактов $K_m \subset X$, т.ч.

$$\sup_{\mu \in \mathcal{M}} \mu(X \setminus K_m) \le \frac{1}{m}.$$

Тогда по свойству 1) теоремы 4.4 $\nu(K_m) \ge \overline{\lim}_n \mu_n(K_m) \ge 1 - \frac{1}{m}$ (так как $K_m \subset X \subset Y$ — компакт, и, следовательно, замкнут в Y). Поэтому $\nu(\cup_m K_m) = 1$ и $\nu(X) = 1$. Докажем, что $\{\mu_n\}$ (рассматриваемые как меры на X) слабо сходятся к $\nu|_X$. Действительно, возьмем открытое множество U в X. Оно является пересечением некоторого открытого множества $V \subset Y$ с X. В силу сходимости $\mu_n \to \nu$ имеем $\underline{\lim}_n \mu_n(U) = \underline{\lim}_n \mu_n(V) \ge \nu(V) = \nu(V \cap X) = \nu(U)$. По свойству 2) теоремы 4.4 $\mu_n \to \nu$ слабо, как меры на X. $\square$

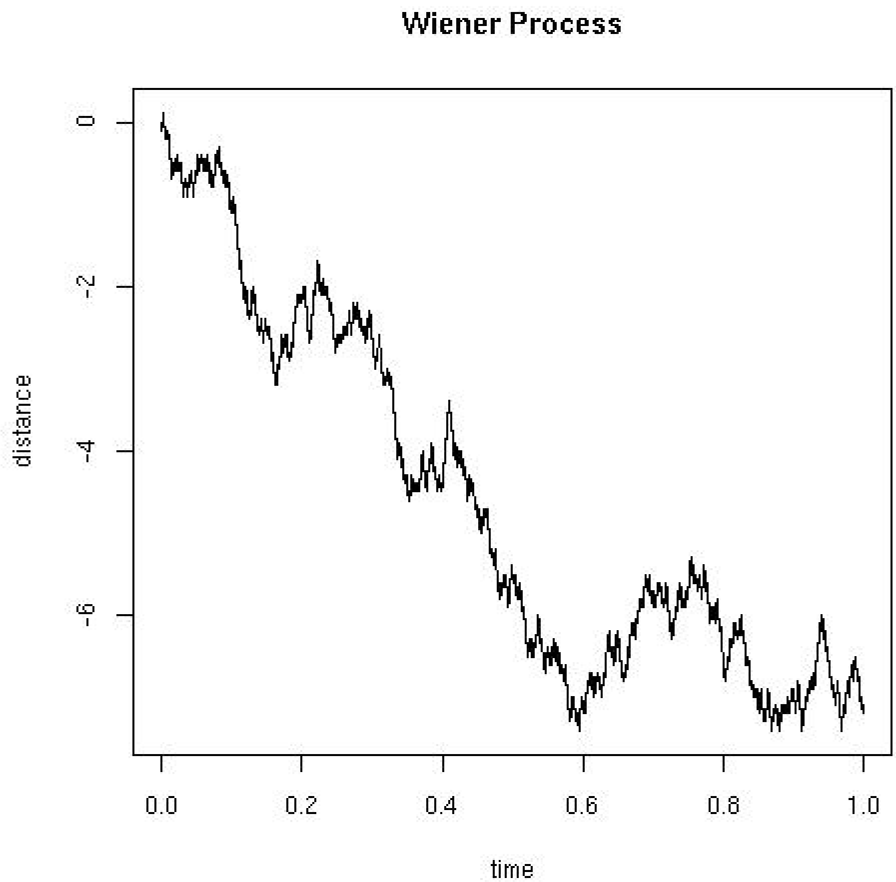
Замечание 4.9. *Достаточность также можно доказывать, используя теорему Банаха-Алаоглу о слабой компактности единичного шара.*

5. Лекция 6. Винеровский процесс.

5.1. **Определение и базовые свойства.** Перейдем к описанию основной (непрерывной) модели всей теории случайных процессов — винеровского процесса или броуновского движения.

Определение 5.1. *Случайный процесс W_t, $t \in [0,T]$ называется винеровским процессом (броуновским движением), если он обладает следующими свойствами:*

1) *Случайный вектор $(W_{t_1}, \cdots, W_{t_n})$, $0 \le t_1 < t_2 < \cdots < t_n \le T$ имеет гауссовское распределение и $W_0 = 0$ п.н.*
2)
$$\mathbb{E}W_t = 0, \ \mathbb{E}\big(W_t W_s\big) = t \wedge s$$
3) *Траектории $t \to W_t(\omega)$ непрерывны для всех $\omega \in \Omega$.*



Замечание 5.2. *Винеровский процесс является* **гауссовским** *процессом (т.е., процессом с гауссовскими конечномерными распределениями) с непрерывными траекториями.*

Вопрос о существовании винеровского процесса мы обсудим ниже. Сформулируем сначала важное эквивалентное определение.

Теорема 5.3. *Винеровский процесс является процессом с независимыми приращениями, т.е. случайные величины*

$$W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \cdots, W_{t_n} - W_{t_{n-1}}, \quad 0 \le t_0 < t_1 < t_2 < \cdots < t_n \le T$$

независимы. Кроме этого, выполнено свойство

$$\mathbb{E}W_t = 0, \quad D(W_t - W_s) = t - s, \quad s \le t. \tag{3}$$

Обратно, гауссовский процесс с непрерывными траекториями и независимыми приращениями является винеровским, если выполнено свойство (3) и $W_0 = 0$.

Доказательство. Пусть W_t — винеровский процесс, $s \le t$. Очевидно,

$$\begin{aligned} D(W_t - W_s) &= \mathbb{E}(W_t - W_s)^2 - \big(\mathbb{E}(W_t - W_s)\big)^2 = \mathbb{E}(W_t^2 - 2W_tW_s + W_s^2) \\ &= t - 2s + s = t - s. \end{aligned}$$

Если $t_i < t_{i+1} \le t_j < t_{j+1}$, то

$$\begin{aligned} \mathbb{E}(W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j}) &= \mathbb{E}\big(W_{t_{i+1}}W_{t_{j+1}} - W_{t_{i+1}}W_{t_j} - W_{t_i}W_{t_{j+1}} + W_{t_i}W_{t_j}\big) \\ &= t_{i+1} - t_{i+1} - t_i + t_i = 0. \end{aligned}$$

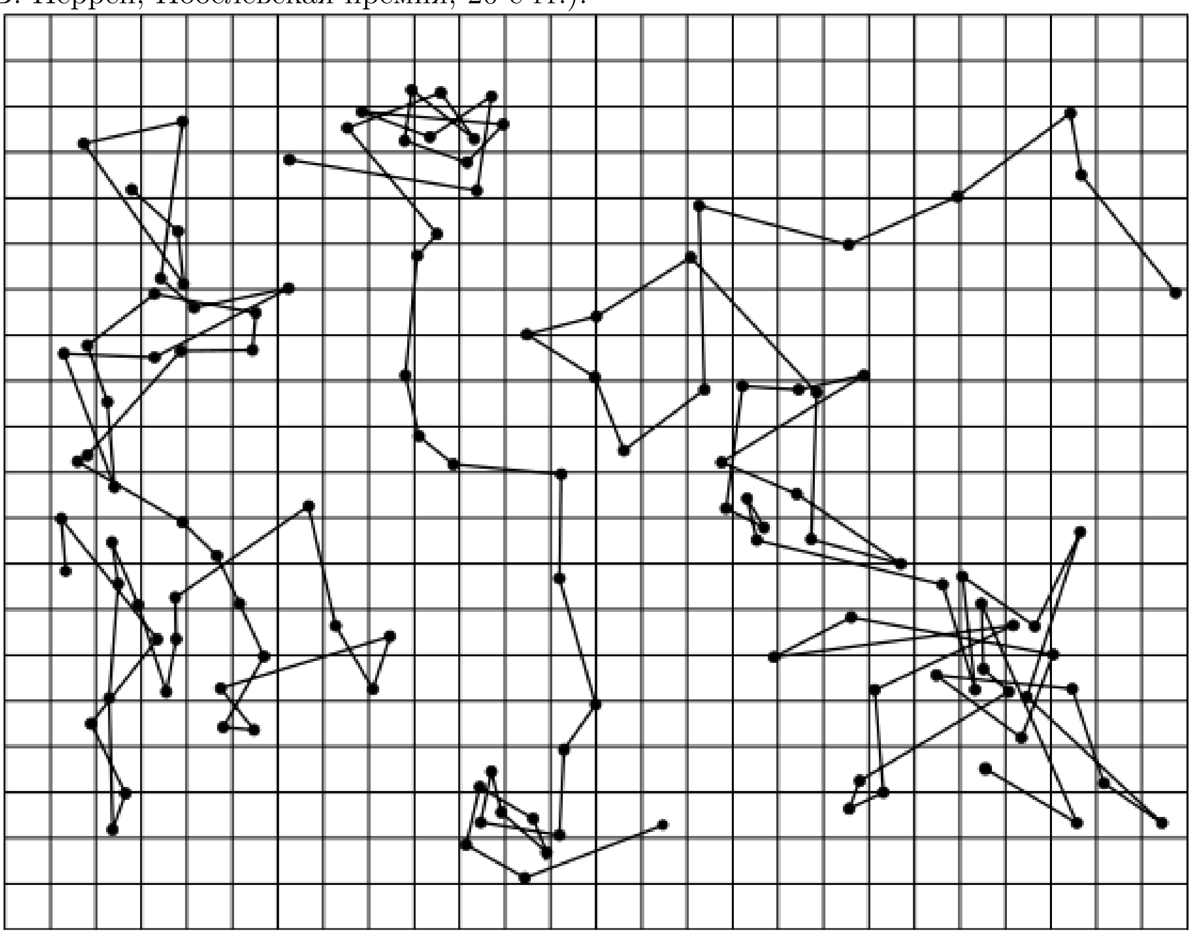
Получаем, что приращения некоррелированы. В силу гауссовости конечномерных распределений приращения независимы.

Обратно, если W_t — гауссовский процесс с независимыми приращениями, выполнено $W_0 = 0$ и выполнено (3), то

$$\begin{aligned}\mathbb{E}(W_tW_s) &= \mathbb{E}(W_t - W_s)W_s + \mathbb{E}W_s^2 = \mathbb{E}(W_t - W_s)(W_s - W_0) + \mathbb{E}(W_s - W_0)^2 \\ &= D(W_s - W_0) = s,\end{aligned}$$

если $s \le t$. □

Весьма интересна история появления винеровского процесса в физике и математике. В 1828 г. ботаник Р. Броун описал явление, впоследствие названное броуновским движением: хаотическое движение пыльцы в жидкости. Оказалось, что это движение вызвано ударами молекул. Первое описание физической модели этого явления было предложено А.Эйнштейном (1905) ("Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen") и М.Смолуховским (1906). Работы Эйнштейна привели к оценке числа Авогадро (Ж.-Б. Перрен, Нобелевская премия, 20-е гг.).



Долгое время Эйнштейн считался пионером в физико-математической теории броуновского движения, но примерно 50 лет спустя была переоткрыта работа Л. Башелье "Théorie de la spéculation", написанная в 1900 г. В этой работе Башелье фактически применил винеровский процесс к описанию ценных бумаг на французском рынке. Его работа осталась незамечена до конца 50-х годов.

Первое строгое математическое доказательство существования винеровского процесса получил Норберт Винер в 1922-23 гг. Его конструкция основывалась на довольно абстрактных методах (интеграл Даниеля) и была весьма тяжелой. Упрощение было достигнуто позже. Стандартное доказательство основано на теореме Колмогорова о построении меры по конечномерным распределениям и его же теореме о существовании непрерывной модификации процесса. Поль Леви предложил доказательство, основанное на сходимости случайных рядов (фактически, на разложении процесса в ряд Фурье со случайными коэффициентами).

5.2. **Недифференцирумость траекторий винеровского процесса.** Траектории винеровского процесса являются весьма нерегулярными функциями. Как мы скоро убедимся, они почти нигде не дифференцируемы (хотя и непрерывны).

Теорема 5.4. *Пусть* $s = t_{0,n} < t_{1,n} < \dots < t_{i_n,n} = t$ *— последовательность разбиений отрезка* $[s,t]$ *с* $\lim_n \max_i |t_{i,n} - t_{i-1,n}| = 0$*. Тогда с.в.*

$$\xi_n = \sum_{i=0}^{n} \bigl(W_{t_{i+1,n}} - W_{t_{i,n}}\bigr)^2$$

обладает свойством: $\lim_n \xi_n = t - s$ *по вероятности.*

Доказательство. Найдем $\mathbb{E}\xi_n$, $D\xi_n$.

$$\mathbb{E}\xi_n = \sum_i \mathbb{E}\bigl(W_{t_{i+1,n}} - W_{i,n}\bigr)^2 = \sum_i (t_{i+1,n} - t_{i,n}) = t - s.$$

В силу независимости приращений

$$\begin{aligned} D(\sum_i \bigl(W_{t_{i+1,n}} - W_{t_{i,n}}\bigr)^2) &= \sum_i D\bigl(W_{t_{i+1,n}} - W_{t_{i,n}}\bigr)^2 \\ &= \sum_i \mathbb{E}\bigl(W_{t_{i+1,n}} - W_{t_{i,n}}\bigr)^4 - \bigl(\mathbb{E}\bigl(W_{t_{i+1,n}} - W_{t_{i,n}}\bigr)^2\bigr)^2. \end{aligned}$$

Воспользуемся тем, что $\mathbb{E}\eta^4 = 3\sigma^4$ для $\eta \sim \mathcal{N}(0,\sigma^2)$. Получаем

$$\begin{aligned} D(\sum_i \bigl(W_{t_{i+1,n}} - W_{t_{i,n}}\bigr)^2) = 2\sum_i (t_{i+1,n} - t_{i,n})^2 &\le 2\max_i |t_{i+1} - t_i| \sum_i (t_{i+1,n} - t_{i,n}) \\ &= 2\max_i |t_{i+1} - t_i|(t-s). \end{aligned}$$

Очевидно, последняя величина стремится к нулю. Из сходимости $D(\xi_n) = \mathbb{E}(\xi_n - \mathbb{E}\xi_n)^2$ к нулю следует сходимость к нулю по вероятности с.в. $\xi_n - \mathbb{E}\xi_n$. □

Напомним, что вариацией функции f на отрезке $[a,b]$ называется величина

$$var_{[a,b]}(f) = \sup \sum_i |f(t_{i+1}) - f(t_i)|,$$

где супремум берется по всем возможным разбиениям $a = t_0 < t_1 < \dots < t_n = b$. В случае, если f имеет производную во всех точках $[a,b]$ и эта производная интегрируема, то

$$var_{[a,b]}(f) = \int_a^b |f'(t)|dt.$$

Лемма 5.5. *Докажите, что если f непрерывна и $var_{[a,b]}(f) < \infty$, то*

$$\lim_n \sum_i |f(t_{i+1}) - f(t_i)|^2 = 0$$

при стремлении к нулю максимума отрезков разбиения.

Следствие 5.6. *С вероятностью 1 траектории винеровского процесса имеют бесконечную вариацию на $[a, b]$.*

Доказательство. Воспользуемся теоремой 5.4. Найдем почти всюду сходящуюся подпоследовательность $\xi_{n_m}(\omega) \to t - s$. Но любая траектория $t \to W_t(\omega)$, для которой это выполнено, имеет неограниченную вариацию по предыдущей лемме. □

6. Лекция 7. Построение винеровского процесса.

6.1. **Регулярность траекторий.** Есть несколько классических доказательств существования винеровского процесса. Например, с помощью теоремы Колмогорова о существовании процесса и теоремы Колмогорова о существовании непрерывной модификации процесса (см. **А.Д. Вентцель**, Введение в теорию случайных процессов). Есть доказательство, использующее случайные ряды. Мы выведем существование процесса из так называемого принципа инвариантности Донскера (изложение следует **N.V. Krylov**, Introduction to the theory of random processes, AMS, Graduate Studies in Math, vol 43., 2002.).

Пусть $\{\eta_i\}$ — последовательность независимых одинаково распределенных с.в. с $\mathbb{E}\eta_i = 0$, $\mathbb{E}\eta_i^2 = 1$. Положим

$$S_k = \eta_1 + \cdots + \eta_k$$

(случайное блуждание). Определим процесс с непрерывным временем

$$\xi^n_t = S_{[nt]}/\sqrt{n} + (nt - [nt])\eta_{[nt]+1}/\sqrt{n}.$$

Мы докажем, что распределение некоторой подпоследовательности $\xi^{n_k}_t$ сходится слабо к так называемой **мере Винера** на пространстве $C([0, 1])$. В доказательстве применим теорему Прохорова.

Напомним, что пространство $C([0, 1])$ непрерывных функций на $[0, 1]$ с равномерной нормой $\|f\| = \sup_{t\in[0,1]} |f(t)|$ является сепарабельным банаховым пространством. Напомним, что борелевской σ-алгеброй $\mathcal{B}(C([0, 1]))$ называется σ-алгебра, порожденная всеми открытыми шарами.

Упражнение 6.1. *Докажите, что $\mathcal{B}(C([0, 1]))$ совпадает с σ-алгеброй, порожденной отображениями вида $f \to f(t)$, где t — произвольная точка, принадлежащая отрезку $[0, 1]$ (т.н. цилиндрическая σ-алгебра).*

Для применения теоремы Прохорова необходимо воспользоваться известным критерием компактности в пространстве непрерывных функций.

Теорема 6.2. (**Арцела-Асколи**) *Множество $K \subset C([0, 1])$ относительно компактно тогда и только тогда, когда*

1) *Существует такое число N, что $\sup_{t\in[0,1]} |x(t)| \le N \quad \forall x \in K$*
2) *для любого $\varepsilon > 0$ существует такое число $\delta > 0$, что*

$$|x(t) - x(s)| \le \varepsilon \quad \forall x \in K, \quad |t - s| \le \delta, \quad t, s \in [0, 1].$$

Лемма 6.3. *Пусть $x_t := x(t)$ — функция на $[0,1]$. Предположим, что существует константа $a > 0$ и такое целое число $n \geq 0$, что*

$$|x_{(i+1)/2^m} - x_{i/2^m}| \leq 2^{-ma}$$

для всех $m \geq n$ и $0 \leq i \leq 2^m - 1$. Тогда для всех бинарно-рациональных чисел $t, s \in [0,1]$, удовлетворяющих $|t-s| \leq 2^{-n}$, выполнено

$$|x_t - x_s| \leq N(a)|t-s|^a,$$

где $N(a) = 2^{2a+1}(2^a - 1)^{-1}$.

Доказательство. Два произвольных бинарно-рациональных числа t, s представим в виде сумм рядов (в действительности, конечных сумм)

$$t = \sum_{i=1}^{\infty} \varepsilon_1(i) 2^{-i}, \ s = \sum_{i=1}^{\infty} \varepsilon_2(i) 2^{-i}.$$

Положим

$$t_k = \sum_{i=1}^{k} \varepsilon_1(i) 2^{-i}, \ s_k = \sum_{i=1}^{k} \varepsilon_2(i) 2^{-i}.$$

Заметим, что $|t-s| \leq 2^{-k}$ тогда и только тогда, когда $|t_k - s_k| = 2^{-k}$ или $t_k = s_k$. Пусть теперь $|t-s| \leq 2^{-k}, \ k \geq n$ Используем представление $x_t = x_{t_k} + \sum_{m=k}^{\infty}(x_{t_{m+1}} - x_{t_m})$. Из этой (и аналогичной для x_s) оценки следует

$$|x_t - x_s| \leq |x_{t_k} - x_{s_k}| + \sum_{m=k}^{\infty}(|x_{t_{m+1}} - x_{t_m}| + |x_{s_{m+1}} - x_{s_m}|).$$

Положим $t_k = r2^{-k}$. Для s_k существуют три возможности: $s_k = t_k$ или $s_k = (r \pm 1)2^{-k}$. Кроме этого, заметим, что $|t_{m+1} - t_m| \leq 2^{-(m+1)}$. Поэтому пары точек t_m, t_{m+1} удовлетворяют условию леммы. По предположению леммы $|x_t - x_s| \leq 2\sum_{m=k}^{\infty} 2^{-ma} = 2^{-ka} 2^{a+1}(2^a - 1)^{-1}$. Осталось заметить, что для любых t, s со свойством $|t-s| \leq 2^{-n}$ можно принять $k = [\log_2(1/|t-s|)]$. Тогда $k \geq n$, $|t-s| \leq 2^{-k}$ и $2^{-ka} \leq 2^a |t-s|^a$. $\square$

Теорема 6.4. *Пусть ξ_t — непрерывный процесс и $\alpha > 0, \beta > 0, N \in (0, \infty)$ — такие константы, что*

$$\mathbb{E}|\xi_t - \xi_s|^{\alpha} \leq N|t-s|^{1+\beta}, \ \forall s, t \in [0,1].$$

Тогда для $0 < a < \beta\alpha^{-1}$ и любого $\varepsilon > 0$ существует такое n, что

$$P(\xi \in K_n(a)) \geq 1 - \varepsilon,$$

где

$$K_n(a) = \{x \in C([0,1]) : |x_0| \leq 2^n, |x_t - x_s| \leq N(a)|t-s|^a \quad \forall |t-s| \leq 2^{-n}\}.$$

Доказательство. Положим

$$A_n = \{\omega : |\xi_0| \geq 2^n\} \cup \{\ \omega : \sup_{m \geq n} \max_{i=0,\cdots,2^m-1} |\xi_{(i+1)/2^m} - \xi_{i/2^m}| 2^{ma} > 1\}.$$

Для $\omega \notin A_n$ имеем: $t \to \xi_t(\omega) \in K_n(a)$ в силу леммы 6.3.

В силу неравенства Чебышева

$$\begin{aligned}P(\xi \notin K_n(a)) = P(A_n) &\leq P(|\xi_0| \geq 2^n) + \mathbb{E}(\sup_{m\geq n} \max_{i=0,\cdots,2^m-1} |\xi_{(i+1)/2^m} - \xi_{i/2^m}|2^{ma}). \\ &\leq P(|\xi_0| \geq 2^n) + \sum_{m=n}^{\infty} \sum_{i=0}^{2^m-1} \mathbb{E}(|\xi_{(i+1)/2^m} - \xi_{i/2^m}|^{\alpha}) 2^{ma\alpha} \\ &\leq P(|\xi_0| \geq 2^n) + N \sum_{m=n}^{\infty} 2^{-m(\beta - a\alpha)}.\end{aligned}$$

□

Лемма 6.5. *Последовательность распределений ξ^n слабо относительно компактна на $C([0,1])$.*

Доказательство. Для упрощения доказательства предположим дополнительно, что $m_4 = \mathbb{E}(|\eta_k|^4) < \infty$. В силу теоремы 6.4 достаточно проверить, что

$$\mathbb{E}(|\xi_t - \xi_s|^4) \leq N|t-s|^2, \quad \forall s, t \in [0,1].$$

Положим $a_n = \mathbb{E}(S_n)^4$. В силу независимости слагаемых

$$a_{n+1} = a_n + 6n + m_4.$$

Отсюда вытекает, что

$$a_n \leq 3n^2 + nm_4.$$

Далее возможны 2 случая

1) s, t принадлежат одному интервалу $[\frac{k}{n}, \frac{k+1}{n}]$. Тогда

$$|\xi_t^n - \xi_s^n| = \sqrt{n}|\eta_{k+1}||t-s|$$

$$\mathbb{E}|\xi_t^n - \xi_s^n|^4 = n^2 m_4 |t-s|^4 = m_4|t-s|^2.$$

2) Пусть $t, s, t > s$ принадлежат разным интервалам. Положим s_1 — ближайшая точка справа к s, t_1 — ближайшая точка слева к t. Очевидно,

$$s_1 - s \leq t - s,\ t - t_1 \leq t - s,\ t_1 - s_1 \leq t - s, (t_1 - s_1) \leq \frac{1}{n}(t_1 - s_1)^2.$$

$$s_1 = \frac{1}{n}([ns]+1),\ t_1 = \frac{[nt]}{n},\ [nt] - ([ns]+1) = n(t_1 - s_1).$$

Поэтому

$$\begin{aligned}\mathbb{E}|\xi_t - \xi_s|^4 \leq &81(\mathbb{E}|\xi_t - \xi_{t_1}|^4 + \mathbb{E}|\xi_{t_1} - \xi_{s_1}|^4 + \mathbb{E}|\xi_{s_1} - \xi_s|^4) \\ &\leq 162(t-s)^2 m_4 + 81\mathbb{E}(|S_{[nt]} - S_{[ns]+1}|/\sqrt{n})^4 \\ &= 162(t-s)^2 m_4 + \frac{81}{n^2} a_{[nt]-[ns+1]}. \\ &\leq 162(t-s)^2 m_4 + 243(t-s)^2 + 81(t_1 - s_1)m_4/n \\ &\leq 243(m_4+1)|t-s|^2.\end{aligned}$$

Лемма доказана. □

6.2. **Мера Винера и принцип Донскера.** Пусть на пространстве $C([0,1])$ задана борелевская мера P. Превратим пространство $\Omega := C([0,1])$ в вероятностное, наделив его борелевской σ-алгеброй и вероятностной мерой P. Из упражнения 6.1 следует, что отображение вида

$$C([0,1]) \ni x \to (x(t_1), x(t_2), \cdots, x(t_n)) \in \mathbb{R}^n$$

является случайным вектором для любого набора $t_i \in [0,1]$

Определение 6.6. *Борелевская мера P на пространстве $C([0,1])$ называется мерой Винера, если $P(\{x(0)=0\}) = 1$ и случайный вектор*

$$(x(t_1), x(t_2), \cdots, x(t_n)), \quad t_1 < t_2 \cdots < t_n$$

имеет гауссовское распределение с нулевым средним и матрицей ковариации $C_{i,j} = \min(t_i, t_j)$.

Упражнение 6.7. *Если P — мера Винера, то отображение*

$$[0,1] \times C([0,1]) \ni (t,x) \to x(t)$$

является винеровским процессом на пространстве $[0,1] \times \Omega$, где $\Omega = C([0,1])$ наделено борелевской σ-алгеброй.

Таким образом, вопрос о построении винеровского процесса сводится к вопрсу о построении меры Винера.

Теорема 6.8. *(**Донскер**) Последовательность распределений ξ^n на $C([0,1])$ обладает подпоследовательностью, слабо сходящейся к мере Винера.*

Доказательство. Существование слабо сходящейся подпоследовательности следует из предыдущей леммы и теоремы Прохорова.

Нам остается доказать, что предельная мера P может быть только мерой Винера. Докажем, что предел распределения случайного вектора $(\xi_{t_1}, \cdots, \xi_{t_n})$ есть гауссовское распределение с нулевым средним и матрицей ковариации $C_{i,j} = \min(t_i, t_j)$.

Пусть $n = 2$ (остальные значения рассматриваются аналогично). Тогда

$$\begin{aligned}\lambda_1 \xi^n_{t_1} + \lambda_2 \xi^n_{t_2} &= (\lambda_1 + \lambda_2) S_{[nt_1]}/\sqrt{n} + \lambda_2 (S_{[nt_2]} - S_{[nt_1]})/\sqrt{n} \\ &+ \eta_{[nt_1]+1}(([nt_1 - [nt_1]])\lambda_1/\sqrt{n} + \lambda_2/\sqrt{n}) \\ &+ \eta_{[nt_2]+1}([nt_2 - [nt_2]])\lambda_2/\sqrt{n}.\end{aligned}$$

Выражение справа — сумма независимых слагаемых $A+B+C+D$. Для вычисления предела найдем (как в доказательстве ЦПТ) предел характеристических функций $\mathbb{E}\big(\exp(i\big[\lambda_1 \xi^n_{t_1} + \lambda_2 \xi^n_{t_2}\big])\big)$.

Нетрудно видеть, что $\mathbb{E}(e^{iC})$, $\mathbb{E}(e^{iD})$ стремятся к единице (почему?). Для остальных слагаемых получаем

$$\lim_n \mathbb{E}(e^{i(A+B)}) = \lim_n \varphi^{[nt_1]}(\varphi(\lambda_1 + \lambda_2)/\sqrt{n}) \cdot \varphi^{[nt_2]-[nt_1]-1}(\varphi(\lambda_2)/\sqrt{n}).$$

Аргументы, применявшиеся в доказательстве ЦПТ, дают нам в пределе соотношение

$$\lim_n \mathbb{E}(e^{i(\lambda_1 \xi^n_{t_1} + \lambda_2 \xi^n_{t_2})}) = \exp\big(-(\lambda_1^2 t_1^2 + 2\lambda_1\lambda_2 \min(t_1,t_2) + \lambda_2^2 t_2^2)/2\big).$$

А это и есть характеристическая функция распределения $\mathcal{N}(0, \min(t_1, t_2))$ $\square$

7. Лекция 8. Марковское свойство. Распределение максимума винеровского процесса.

Отметим без доказательства некоторые важные свойства винеровского процесса. Многомерный винеровским процессом называется набор n независимых винеровских процессов.

1) В размерностях 1, 2 винеровский процесс бесконечное число раз возвращается в окрестность любой точки.

2) В размерности 1 график винеровского процесса бесконечное число раз пересекает ось абсцисс на любом отрезке $[0, t]$.

3) Законы повторного логарифма

$$\overline{\lim}_{t\to\infty} \frac{|W_t|}{\sqrt{2t\log\log t}} = 1, \quad \overline{\lim}_{0\le t\le 1; t\to 0} \frac{|W_t|}{\sqrt{2t\log\frac{1}{t}}} = 1$$

почти всюду.

4) Гёльдеровость траекторий.

Свойства гёльдеровости траекторий вытекает из доказанного в предыдущем разделе.

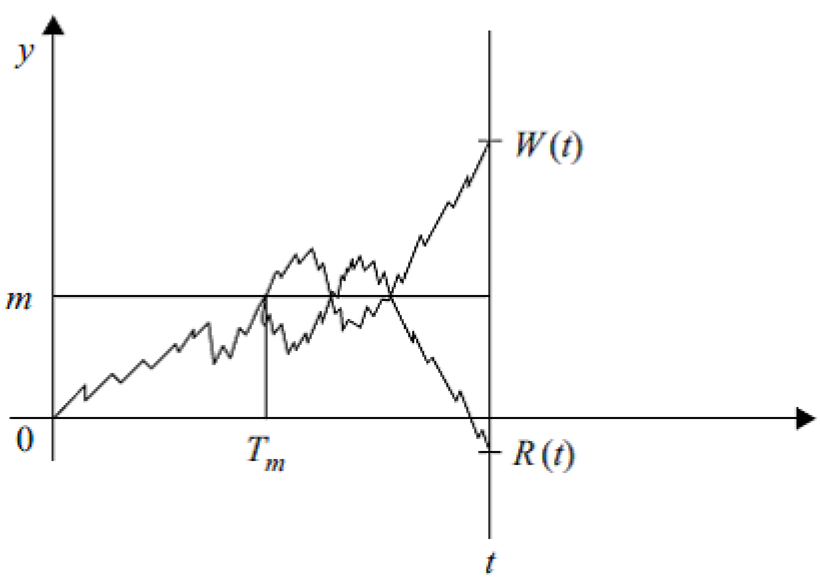
Упражнение 7.1. *Докажите, что для любого $\varepsilon > 0$ для почти любой траектории ω существует такое число $N(\omega) > 0$, что выполнено неравенство $|W_t - W_s| \le N(\omega)|t-s|^{1/2-\varepsilon}$, $s, t \in [0, 1]$.*

5) Принцип отражения

Пусть $\tau = \inf\{t : W_t > a\}$ — первый момент достижения точки $a > 0$. Тогда процесс

$$\xi_t = \begin{cases} W_t, & t \le \tau \\ 2W_\tau - W_t & t \ge \tau \end{cases}$$

винеровский.



Из принципа отражения вытекает формула для распределения τ. Действительно, в силу принципа отражения

$$P(\tau \le t, W_t \le a) = P(\tau \le t, W_t \ge a).$$

С другой стороны
$$P(\tau \le t, W_t \ge a) = P(W_t \ge a).$$
Поэтому
$$P(\tau \le t) = P(\tau \le t, W_t \le a) + P(\tau \le t, W_t \ge a) = 2P(W_t \ge a) = \sqrt{\frac{2\pi}{t}} \int_a^{+\infty} e^{-\frac{x^2}{2t}}\, dx.$$

Упражнение 7.2. *Докажите, что плотность распределения τ равна*
$$\frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}}.$$

6) Распределение максимума

Пусть $S_t = \max_{0\le s\le t} W_s$ — максимум винеровского процесса на $[0,t]$, $a > 0$. Тогда
$$P(S_t \le a) = P(|W_t| \le a)$$

Это можно вывести 1) из предыдущего свойства, 2) из того, что винеровский процесс является пределом процессов ξ^n_t и свойства симметричного случайного блуждания
$$P(\max_{1\le k\le n} S_k \ge m) = 2P(S_n > m) + P(S_n = m)$$
3) из марковского свойства (см. далее).

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задано семейство σ-алгебр $\mathcal{F}_t$, $t \in [0, T_0]$, удовлетворяющее свойству
$$\mathcal{F}_s \subset \mathcal{F}_t,\ s \le t.$$
Такое семейство будем называть потоком ("filtration" в англоязычной литературе).

Будем говорить, что $(W_t, \mathcal{F}_t)$ — винеровский процесс относительно $\mathcal{F}_t$, если W_t измерим относительно $\mathcal{F}_t$ и $W_{t+s} - W_t$ не зависит от $\mathcal{F}_t$ для любых $s, t \ge 0$, W_t — гауссовский процесс с непрерывными траекториями и свойством $\mathbb{E}(W_t) = 0, \mathbb{E}(W_t^2) = t$.

Для заданного винеровского процесса всегда существует такое семейство $\mathcal{F}_t$, для которого это свойство выполняется. В качестве такового можно взять σ-алгебру $\mathcal{F}_{\le t}$, которая строится следующим образом.

Упражнение 7.3. *Пусть $\mathcal{F}_{\le t} = \sigma\{W_s, s \in [0,t]\} = \sigma\{\omega : W_s(\omega) \in B, s \in [0,t], B \in \mathcal{B}\}$ (через $\mathcal{B}$ обозначается совокупность борелевских множеств прямой) порождена случайными величинами W_s, $s \in [0,t]$. Тогда $(W_t, \mathcal{F}_{\le t})$ — винеровский процесс относительно $\mathcal{F}_{\le t}$.*

В момент времени t σ-алгебру $\mathcal{F}_t$ естественно понимать как совокупность событий "прошлого" и "настоящего".

Непосредственно проверяется так называемое марковское свойство винеровского процесса.

Теорема 7.4. *(Марковское свойство) Пусть $s \ge 0$. Процесс $t \to W_{t+s} - W_s$, является винеровским.*

σ-алгебра, порожденная $W_{t+s} - W_s$, не зависит от $\mathcal{F}_s$.

Фундаментальным свойством винеровского процесса является то, что теорема выше оказывается верной не только для фиксированной константы t, но также для некоторого класса случайных величин, называемых марковскими моментами (сильное марковское свойство).

Определение 7.5. *Случайная величина $\tau \in [0, +\infty]$ называется марковским моментом (stopping time), если для любого t событие $\{\tau \le t\}$ принадлежит $\mathcal{F}_t$.*

Неформально можно сказать, что τ — случайная величина, "не зависящая от будущего".

Упражнение 7.6. *1) Пусть $a \in \mathbb{R}$. Проверьте, что $\tau_a = \inf\{t \ge 0 : W_t \ge a\}$ (считаем, что $\tau_a = \infty$, если $W_t < a$ для всех t) — марковский момент.*

2) Пусть $a < 0 < b$. Проверьте, что $\tau = \inf\{t \ge 0 : W_t \notin (a, b)\}$ — марковский момент.

Ниже мы будем иметь дело с винеровским процессом $(W_t, \mathcal{F}_t)$ относительно потока σ-алгебр $\mathcal{F}$. Как мы уже видели, "сдвинутый" на постоянную величину s процесс

$$B_t = W_{t+s} - W_s$$

является винеровским и не зависит от $\mathcal{F}_s$.

Далее мы рассматриваем марковский момент τ, обладающий свойством $P(\tau < \infty) = 1$. С каждым таким τ можно связать две σ-алгебры (через $\mathcal{B}$ обозначается совокупность борелевских множеств прямой)

$$\mathcal{F}_{\le\tau} = \sigma\{\omega : W_{\min(s,\tau)}(\omega) \in B, s \ge 0, B \in \mathcal{B}\}$$
$$\mathcal{F}_{\ge\tau} = \sigma\{\omega : W_{s+\tau}(\omega) - W_s(\omega) \in B, s \ge 0, B \in \mathcal{B}\}.$$

Теорема 7.7. *(Сильное марковское свойство) Пусть τ — марковский момент относительно потока σ-алгебр $\mathcal{F}_t$, $P(\tau < \infty) = 1$. Процесс $B_t = W_{t+\tau} - W_\tau$ является винеровским.*

σ-алгебры $\mathcal{F}_{\le\tau}$ и $\mathcal{F}_{\ge\tau}$ независимы.

Доказательство. Для доказательства первого утверждения достаточно показать, что для $0 \le t_1 \le \cdots \le t_n$ и ограниченной непрерывной функции g

$$\mathbb{E}g(B_{t_1}, \cdots, B_{t_k}) = \mathbb{E}g(W_{t_1}, \cdots, W_{t_k}).$$

Предположим, что τ принимает не более чем счетное множество значений:

$$\tau = \sum_n r_n I_{\{\tau = r_n\}}$$

Тогда

$$\begin{aligned}\mathbb{E}g(B_{t_1}, \cdots, B_{t_k}) &= \mathbb{E}\big(\sum_n g(B_{t_1}, \cdots, B_{t_k}) I_{\{\tau=r_n\}}\big)\\ &= \mathbb{E}\big(\sum_n g(W_{t_1+r_n} - W_{r_n}, \cdots, W_{t_k+r_n} - W_{r_n}) I_{\{\tau=r_n\}}\big).\end{aligned}$$

Заметим, что в силу марковости τ, случайная величина $I_{\{\tau=r_n\}}$ измерима относительно $\mathcal{F}_{r_n}$, а с.в. $W_{t_i+r_n} - W_{r_n}$ — независима от $\mathcal{F}_{r_n}$. Поэтому последнее выражение равно (в силу независимости и марковского свойства)

$$\mathbb{E}\big(\sum_n g(W_{t_1+r_n} - W_{r_n}, \cdots, W_{t_k+r_n} - W_{r_n})\big) P(\tau = r_n) = \sum_n \mathbb{E}\Big[g(W_{t_1}, \cdots, W_{t_k}) P(\tau = r_n)\Big]$$
$$= \mathbb{E}\Big[g(W_{t_1}, \cdots, W_{t_k})\Big] \cdot \sum_n P(\tau = r_n) = \mathbb{E}\big(g(W_{t_1}, \cdots, W_{t_k})\big).$$

В общем случае приблизим τ последовательностью с дискретным множеством значений

$$\tau_n(\omega) = (k+1)2^{-n}, \quad k2^{-n} < \tau(\omega) \le (k+1)2^{-n}.$$

Очевидно, $\tau \le \tau_n \le \tau + 2^{-n}$. Заметим, что $\{\omega : \tau_n(\omega) > t\} = \{\omega : \tau(\omega) > 2^{-n}[2^n t]\} \in \mathcal{F}_{2^{-n}[2^n t]} \subset \mathcal{F}_t$. В силу непрерывности траекторий W_t и функции g, а также теоремы Лебега

$$\begin{aligned}\mathbb{E}g(B_{t_1}, \cdots, B_{t_k}) &= \lim_n \mathbb{E}g(W_{t_1+\tau_n} - W_{\tau_n}, \cdots, W_{t_k+\tau_n} - W_{\tau_n}) \\ &= \lim_n \mathbb{E}g(W_{t_1}, \cdots, W_{t_k}) = \mathbb{E}g(W_{t_1}, \cdots, W_{t_k}).\end{aligned}$$

Первая часть теоремы доказана.

Для доказательства независимости σ-алгебр рассмотрим пару непрерывных ограниченных функций f, g. Достаточно доказать, что

$$\mathbb{E}\big(f(W_{t_1\wedge\tau}, \cdots, W_{t_k\wedge\tau}) \cdot g(B_{t_1}, \cdots, B_{t_m})\big) = \mathbb{E}\big(f(W_{t_1\wedge\tau}, \cdots, W_{t_k\wedge\tau}) \cdot \mathbb{E}g(B_{t_1}, \cdots, B_{t_m})\big)$$

Как и выше, доказательство сводится к случаю, когда τ принимает дискретное множество значений.

$$\begin{aligned}&\mathbb{E}\Big(f(W_{t_1\wedge\tau}, \cdots, W_{t_k\wedge\tau}) \cdot g(B_{t_1}, \cdots, B_{t_m})\Big) = \sum_n \mathbb{E}\Big(f(W_{t_1\wedge\tau}, \cdots, W_{t_k\wedge\tau}) \cdot g(B_{t_1}, \cdots, B_{t_m}) I_{\{\tau=r_n\}}\Big) \\ &\quad = \sum_n \mathbb{E}\Big(f(W_{t_1\wedge r_n}, \cdots, W_{t_k\wedge r_n}) \cdot g(W_{t_1+r_n} - W_{r_n}, \cdots, W_{t_k+r_n} - W_{r_n}) I_{\{\tau=r_n\}}\Big).\end{aligned}$$

В силу марковского свойства последнее выражение равно

$$\begin{aligned}&\sum_n \mathbb{E}\Big(f(W_{t_1\wedge r_n}, \cdots, W_{t_k\wedge r_n}) \cdot I_{\{\tau=r_n\}}\Big)\mathbb{E}g(W_{t_1+r_n} - W_{r_n}, \cdots, W_{t_k+r_n} - W_{r_n}) \\ &\quad = \sum_n \mathbb{E}\Big(f(W_{t_1\wedge r_n}, \cdots, W_{t_k\wedge r_n}) \cdot I_{\{\tau=r_n\}}\Big)\mathbb{E}g(W_{t_1}, \cdots, W_{t_k}) \\ &\quad = \mathbb{E}f(W_{t_1\wedge\tau}, \cdots, W_{t_k\wedge\tau}) \cdot \mathbb{E}g(W_{t_1}, \cdots, W_{t_k}) = \mathbb{E}f(W_{t_1\wedge\tau}, \cdots, W_{t_k\wedge\tau}) \cdot \mathbb{E}g(B_{t_1}, \cdots, B_{t_k}).\end{aligned}$$

Что и требовалось доказать. □

Следствие 7.8. *(Распределение максимума) Положим* $S_t = \sup_{s\in[0,t]} W_s$. *Тогда*

$$P(S_t \le a) = P(|W_t| \le a) = \frac{2}{\sqrt{2\pi t}} \int_0^a e^{-\frac{x^2}{2t}}\, dx.$$

Доказательство.

$$P(S_t \ge a) = P(S_t \ge a, W_t < a) + P(W_t \ge a).$$

Заметим, что $W_{\tau_a} = a$. Далее, в силу марковского свойства

$$P(S_t \ge a, W_t < a) = P(\tau_a \le t, W_{\tau_a+(t-\tau_a)} - W_{\tau_a} < 0) = \int_0^t P(W_{s+(t-s)} - W_s < 0 | \tau_a = s)\, dF_{\tau_a}(s)$$

Далее, заметим, что $W_t - W_s$ не зависит от $\mathcal{F}_s$ при $s \in [0, t]$, а с.в. τ_a — марковский момент. Следовательно, $\{\tau_a \le s\} \in \mathcal{F}_s$ и $P(W_{s+(t-s)} - W_s < 0 | \tau_a = s) = P(W_{s+(t-s)} - W_s < 0)$. Так как $P(W_{s+(t-s)} - W_s < 0) = 1/2$, последняя величина равна

$$\frac{1}{2}\int_0^t dF_{\tau_a}(s) = \frac{1}{2}P(S_t \ge a).$$

Отсюда следует, что $P(S_t \ge a) = 2P(W_t \ge a) = P(|W_t| \ge a)$. □

8. Лекция 9. Мартингалы.

8.1. **Определение. Примеры. Неравенство Иенссена.**

Определение 8.1. *Пусть $\mathcal{F}_t$ — поток σ-алгебр. Случайный процесс $\{\xi_t\}$ с $\mathbb{E}|\xi_t| < \infty$ называется мартингалом относительно $\mathcal{F}_t$, если $\mathbb{E}(\xi_t|\mathcal{F}_s) = \xi_s$, $s \le t$.*

Аналогично определяется мартингал с дискретным временем $\{\xi_n\}$: $\mathbb{E}(\xi_n|\mathcal{F}_m) = \xi_m$, $m \le n$.

Из определения следует, в частности, что с.в. ξ_t измерима относительно $\mathcal{F}_t$ при всех t, т.е. ξ_t — согласованный процесс.

Пример 8.2. *Согласованный процесс ξ_t с постоянным средним $\mathbb{E}\xi_t = C$, приращения $\xi_t - \xi_s$ которого не зависят от $\mathcal{F}_s$, $t > s$, является мартингалом относительно потока σ-алгебр, порожденных ξ_t.*

Доказательство. $\mathbb{E}(\xi_t|\mathcal{F}_s) = \mathbb{E}(\xi_t - \xi_s|\mathcal{F}_s) + \mathbb{E}(\xi_s|\mathcal{F}_s) = \xi_s + \mathbb{E}(\xi_t - \xi_s) = \xi_s$. □

Пример 8.3. *Пусть $(W_t, \mathcal{F}_t)$ — винеровский процесс. Тогда $W_t^2 - t$ — мартингал относительно $\mathcal{F}_t$.*

Доказательство.

$$\begin{aligned}\mathbb{E}\Big(W_t^2 - t|\mathcal{F}_s\Big) &= \mathbb{E}\big((W_t - W_s)^2 + 2W_s(W_t - W_s) + W_s^2|\mathcal{F}_s\Big) - t \\ &= \mathbb{E}(W_t - W_s)^2 + 2W_s\mathbb{E}\big((W_t - W_s)|\mathcal{F}_s\big) + W_s^2 - t \\ &= t - s + 2W_s\mathbb{E}(W_t - W_s) + W_s^2 - t = W_s^2 - s.\end{aligned}$$

□

Пример 8.4. *Пусть X — случайная величина с конечным первым моментом. Тогда $X_t = \mathbb{E}(\xi|\mathcal{F}_t)$ — мартингал относительно $\mathcal{F}_t$.*

Определение 8.5. *Пусть $\mathcal{F}_t$ — поток σ-алгебр. Случайный согласованный процесс $\{\xi_t\}$ называется субмартингалом (супермартингалом) относительно $\mathcal{F}_t$, если $\mathbb{E}(\xi_t|\mathcal{F}_s) \ge \xi_s$ ($\mathbb{E}(\xi_t|\mathcal{F}_s) \le \xi_s$), $s \le t$.*

Упражнение 8.6. *Докажите следующую эквивалентную формулировку: случайный согласованный процесс $\{\xi_t\}$ называется мартингалом (субмартингалом) относительно $\mathcal{F}_t$, если*

$$\mathbb{E}(\xi_t \cdot \eta) = \mathbb{E}(\xi_s \cdot \eta) \quad \Big(\mathbb{E}(\xi_t \cdot \eta) \ge \mathbb{E}(\xi_s \cdot \eta)\Big)$$

для любой неотрицательной ограниченной функции η, измеримой относительно $\mathcal{F}_s$.

Теорема 8.7. *Пусть f — выпуклая функция, ξ_t — мартингал. Тогда $f(\xi_t)$ — субмартингал, если $\mathbb{E}|f(\xi_t)| < \infty$.*

Доказательство. Идея доказательства состоит в применении известного неравенства для выпуклых функций

$$f(x) \ge f(y) + C(x - y),$$

выполненного для некоторой константы C, зависящей только от x. Если f — гладкая функция по x, то $C(x) = f'(x)$. В противном случае множество таких C содержит более одного элемента и образует так называемый субдифференциал функции f. В качестве C можно взять, например, правую производную f. Имеем:

$$f(\xi_t) \ge f(\xi_s) + C(\xi_s)(\xi_t - \xi_s).$$

Умножим неравенство на неотрицательную функцию η, измеримую относительно $\mathcal{F}_s$, и возьмем математическое ожидание от обеих частей:

$$\mathbb{E}\big(f(\xi_t)\eta\big) \geq \mathbb{E}\big(f(\xi_s)\eta\big) + \mathbb{E}\big(C(\xi_s)(\xi_t - \xi_s)\eta\big) = \mathbb{E}\big(f(\xi_s)\eta\big) + \mathbb{E}\big(C(\xi_s)\xi_t\eta\big) - \mathbb{E}\big(C(\xi_s)\xi_s\eta\big).$$

Далее заметим, что $\mathbb{E}\big(C(\xi_s)\eta\xi_t\big) = \mathbb{E}\big(C(\xi_s)\eta\mathbb{E}(\xi_t|F_s)\big) = \mathbb{E}\big(C(\xi_s)\eta\xi_s\big)$ в силу того, что ξ_t — мартингал. Отсюда получаем искомое неравенство $\mathbb{E}\big(f(\xi_t)\eta\big) \geq \mathbb{E}\big(f(\xi_s)\eta\big)$.

В этом доказательстве есть пробел — мы неявно использовали интегрируемость функции $C(\xi_s)\xi_t\eta$. Для того, чтобы его устранить, можно дополнительно потребовать, чтобы $\eta C(\xi_s)$ было ограничено (например, взять $\eta = \varphi I_{A_N}$, где $A_N = \{|C(\xi_s)| \leq N\}$). Тогда вычисления выше верны и мы получаем $\mathbb{E}\big(f(\xi_t)\varphi I_{A_N}\big) \geq \mathbb{E}\big(f(\xi_s)\varphi I_{A_N}\big)$. Переходя к пределу $N \to \infty$ и применяя теорему Лебега, получаем искомое соотношение $\mathbb{E}\big(f(\xi_t)\varphi\big) \geq \mathbb{E}\big(f(\xi_s)\varphi\big)$. □

8.2. **Мартингалы и марковские моменты. Мартингальные неравенства.**

Определение 8.8. *Случайная величина $\tau \in [0, +\infty]$ называется марковским моментом (stopping time), если для любого t событие $\{\tau \leq t\}$ принадлежит $\mathcal{F}_t$.*

Неформально можно сказать, что τ — случайная величина, "не зависящая от будущего".

Упражнение 8.9. *Пусть $U \subset \mathbb{R}$ — открытое множество, W_t — винеровский процесс. Докажите, что $\tau = \inf\{t > 0 : W_t \notin U\}$ — марковский момент. Указание: представьте U в виде возрастающих замкнутых множеств K_n. Тогда $\{\omega : \tau(\omega) \leq t\} = \cap_m \cup_{r \in \mathbb{Q}, r<t} \{\omega : W_r \notin K_m\}$.*

Лемма 8.10. *Пусть τ — марковский момент с конечным числом значений $t_1 < t_2 < \cdots < t_n$, ξ_t — субмартингал. Выполнено неравенство*

$$\mathbb{E}\xi_{t_1} \leq \mathbb{E}\xi_\tau \leq \mathbb{E}\xi_{t_n}.$$

Доказательство. Имеем: $\mathbb{E}\xi_\tau = \sum_{t_k} \mathbb{E}(\xi_{t_k} I_{\{\tau = t_k\}}) = \mathbb{E}(\xi_{t_1}) - \mathbb{E}(\xi_{t_1} I_{\tau > t_1}) + \mathbb{E}(\xi_{t_2} I_{\tau > t_1}) - \mathbb{E}(\xi_{t_2} I_{\tau > t_2}) + \cdots + \mathbb{E}(\xi_{t_n} I_{\tau > t_{n-1}})$. Осталось заметить, что по определению субмартингала $\mathbb{E}(\xi_{t_{k+1}} I_{\tau > t_k}) \geq \mathbb{E}(\xi_{t_k} I_{\tau > t_k})$, следовательно $\mathbb{E}\xi_\tau \geq \mathbb{E}(\xi_{t_1})$. Второе неравенство доказывается аналогично □

Следствие 8.11. *1) Если τ — марковский момент с конечным числом значений $t_1 < t_2 < \cdots < t_n$, ξ_t — мартингал, то выполнено точное равенство*

$$\mathbb{E}\xi_{t_1} = \mathbb{E}\xi_\tau = \mathbb{E}\xi_{t_n}.$$

2) Если ξ_t — неотрицательный субмартингал, то

$$P(\max_{1 \leq k \leq n} \xi_{t_k} \geq \varepsilon) \leq \frac{\mathbb{E}\xi_{t_n}}{\varepsilon}.$$

Доказательство. Первое утверждение очевидно, для доказательства второго заметим, что если $\tau = \min\{t_k : \xi_{t_k} \geq \varepsilon\}$ ($\tau = t_n$, если таких t_k нет), то согласно лемме 8.10 и неравенству Чебышева

$$P(\max_{1 \leq k \leq n} \xi_{t_k} \geq \varepsilon) = P(\xi_\tau \geq \varepsilon) \leq \frac{\mathbb{E}(\xi_\tau)}{\varepsilon} \leq \frac{\mathbb{E}\xi_{t_n}}{\varepsilon}$$

□

Естественно предположить, что доказанные выше соотношения можно распространить на более общие типы марковских моментов. Это действительно можно сделать в предположении односторонней непрерывности траекторий процессов и ограничений на интегрируемость моментов. Следующая теорема является частным случаем так называемой optional stopping theorem или OS-теоремы. В доказательстве нам потребуется следующий вспомогательный результат.

Задача 8.12. *Пусть η — с.в., $\mathbb{E}(|\eta|) < \infty$. Существует такая четная выпуклая функция $f : \mathbb{R} \to [0, +\infty)$, что $\mathbb{E}f(\eta) < \infty$, f возрастает на $[0, +\infty)$ и $\lim_{t\to+\infty} \frac{f(t)}{t} = +\infty$. Указание: воспользуйтесь тем, что $\mathbb{E}|\eta| < \infty$ равносильно условию $\sum_{n=0}^{\infty} nP(n \le |\eta| < n+1) < \infty$.*

Теорема 8.13. *Пусть ξ_t — мартингал с непрерывными справа траекториями, τ — марковский момент. Тогда соотношение*

$$\mathbb{E}\xi_\tau = \mathbb{E}\xi_0$$

имеет место, если $\tau \le K$ п.н. для некоторого числа $K > 0$.

Доказательство. Приблизим τ последовательностью марковских моментов с конечным числом значений $\tau_n = \min\Big[2^{-n}\big(1 + [2^n\tau]\big), K\Big]$. Имеем $\mathbb{E}\xi_0 = \mathbb{E}\xi_{\tau_n}$. В силу непрерывности справа траекторий $\xi_{\tau_n} \to \xi_\tau$. Теорема будет доказана, если мы сможем перейти к пределу под знаком интеграла. Для этого нам достаточно доказать, что семейство с.в. $\{\xi_{\tau_n}\}$ *равномерно интегрируемо*, т.е.

$$\lim_{N\to\infty} \sup_n \mathbb{E}\big(|\xi_{\tau_n}| I_{|\xi_{\tau_n}|>N}\big) = 0.$$

Заметим сначала, что $|\xi_t|$ — субмартингал, поэтому в силу следствия 8.11

$$\sup_n P(|\xi_{\tau_n}| > N) \le P(\sup_n |\xi_{\tau_n}| > N) \le \frac{\mathbb{E}(|\xi_K|)}{N}.$$

Поэтому

$$\lim_{N\to\infty} \sup_n P(|\xi_{\tau_n}| > N) = 0. \tag{4}$$

Согласно задаче 8.12, $\mathbb{E}f(\xi_K) < \infty$ для некоторой неотрицательной четной выпуклой функции f с $\lim_{t\to+\infty} \frac{f(t)}{t} = +\infty$. Процесс $f(\xi_t) = f(|\xi_t|)$ — субмартингал, поэтому в силу леммы 8.10

$$\mathbb{E}f(\xi_{\tau_n}) \le \mathbb{E}f(\xi_K).$$

Далее, в силу неравенства Иенсена (примененного к вероятностной мере $\frac{1}{P(|\xi_{\tau_n}|>N)} I_{|\xi_{\tau_n}|>N} \cdot P$)

$$f\Big(\frac{\mathbb{E}\big(|\xi_{\tau_n}| I_{|\xi_{\tau_n}|>N}\big)}{P(|\xi_{\tau_n}| > N)}\Big) \le \frac{\mathbb{E}\Big(f(|\xi_{\tau_n}|) I_{|\xi_{\tau_n}|>N}\Big)}{P(|\xi_{\tau_n}| > N)} \le \frac{\mathbb{E}\Big(f(|\xi_{\tau_n}|)\Big)}{P(|\xi_{\tau_n}| > N)} \le \frac{\mathbb{E}\big(f(|\xi_K|)\big)}{P(|\xi_{\tau_n}| > N)}.$$

В силу того, что f возрастает на $[0, +\infty)$, получаем

$$\mathbb{E}\big(|\xi_{\tau_n}| I_{|\xi_{\tau_n}|>N}\big) \le P(|\xi_{\tau_n}| > N) f^{-1}\Big(\frac{\mathbb{E}\big(f(|\xi_K|)\big)}{P(|\xi_{\tau_n}| > N)}\Big).$$

Заметим, что из условия $\lim_{t\to+\infty} \frac{f(t)}{t} = +\infty$ следует, что $\lim_{x\to 0+} x f^{-1}(C/x) = C \lim_{y\to+\infty} \frac{f^{-1}(Cy)}{Cy} = C \lim_{t\to+\infty} \frac{t}{f(t)} = 0$. Поэтому, в силу (4).

$$\lim_{N\to\infty} \sup_n \mathbb{E}\big(|\xi_{\tau_n}| I_{|\xi_{\tau_n}|>N}\big) = 0$$

Равномерная интегрируемость доказана.

Наконец, докажем, что из равномерной интегрируемости следует искомый предельный переход. Имеем:

$$\mathbb{E}\xi_{\tau_n} = \mathbb{E}(\xi_{\tau_n} I_{\{|\xi_{\tau_n}|\le N\}}) + \mathbb{E}(\xi_{\tau_n} I_{\{|\xi_{\tau_n}|>N\}}).$$

Пусть N таково, что $P(|\xi_\tau| = N) = 0$. Тогда $\lim_n \mathbb{E}(\xi_{\tau_n} I_{|\xi_{\tau_n}|\le N}) \to \mathbb{E}(\xi_\tau I_{|\xi_\tau|\le N})$ по теореме о мажорированной сходимости. В силу равномерной интегрируемости $\lim_N \sup_n \mathbb{E}(|\xi_{\tau_n}| I_{|\xi_{\tau_n}|>N}) = 0$. Отсюда легко следует искомое соотношение $\lim_n \mathbb{E}\xi_{\tau_n} = \mathbb{E}\xi_\tau$. □

Следующее обобщение следствия 8.11 (2) несложно и мы его опустим (см. микротеорему 7.3.5, **А.Д. Вентцель**, "Случайные процессы").

Теорема 8.14. *(***Неравенство Колмогорова***). Пусть ξ_t — неотрицательный субмартингал с непрерывными справа траекториями. Тогда $\forall \varepsilon > 0$*

$$P(\sup_{t\in[0,T]} \xi_t \ge \varepsilon) \le \frac{\mathbb{E}\xi_T}{\varepsilon}.$$

Задача 8.15. *(***Задача о разорении с точки зрения мартингалов***). Вы приходите в казино с k\$. За одну партию вы выигрываете 1\$ с вероятностью p и проигрываете 1\$ с вероятностью $q = 1 - p$, $p > 1/2$. Если у вас остается 0\$, вы уходите. Если вы приобретаете K\$, где $K > k$ — некоторая фиксированная сумма, вы тоже уходите. Найти вероятность того, что вы уйдете ни с чем.*

Решение Пусть $X_i = 1$ в случае выигрыша в i-й партии и $X_i = -1$ в случае проигрыша. Положим: $S_n = k + X_1 + \cdots + X_n$. Несложно доказать, что с.в. $\xi_n = \left(\frac{q}{p}\right)^{S_n}$ — мартингал (как произведение независимых неотрицательных с.в. со средним 1). Пусть τ — первый момент достижения процессом S_n состояния 0 или K. По OS-теореме

$$\left(\frac{q}{p}\right)^k = \xi_0 = \mathbb{E}\xi_\tau = \left(\frac{q}{p}\right)^K P(S_\tau = K) + P(S_\tau = 0).$$

Учитывая соотношение $P(S_\tau = K) + P(S_\tau = 0) = 1$, легко находим вероятность разорения $P(S_\tau = 0)$.

Для обоснования выкладок покажем, что $\tau < \infty$ почти всюду. Заметим, что последовательность $T_n = S_n - n(p-q) = k + \sum_{i=1}^n (X_i - \mathbb{E}X_i)$ является мартингалом. По неравенству Чебышева

$$P(\tau > N) \le P(\inf_{n\le N} S_n > 0) \le P(T_N > (p-q)N) \le \frac{\mathbb{E}(T_N^2)}{(q-p)^2 N^2} = \frac{k^2 + ND(X_1)}{(q-p)^2 N^2}.$$

Следовательно, $P(\tau < \infty) = 1$.

Докажем теперь, что $\mathbb{E}\xi_\tau = \left(\frac{q}{p}\right)^k$. Положим: $\tau_n = \min(\tau, n)$. Достаточно заметить, что $\lim_n \tau_n = \tau$ и случайные величины ξ_{τ_n} равномерно ограничено единицей. Поэтому по теорем Лебега о мажорированной сходимости

$$\mathbb{E}\xi_\tau = \lim_n \mathbb{E}\xi_{\tau_n} = \left(\frac{q}{p}\right)^k.$$

9. Лекция 10. Стохастическое интегрирование.

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задано семейство σ-алгебр $\mathcal{F}_t$, $t \in [0, T_0]$, удовлетворяющее свойству

$$\mathcal{F}_s \subset \mathcal{F}_t,\ s \leq t.$$

Такое семейство будем называть потоком ("filtration" в англоязычной литературе).

Будем говорить, что $(W_t, \mathcal{F}_t)$ — винеровский процесс относительно $\mathcal{F}_t$, если W_t — винеровский процесс, W_t измерим относительно $\mathcal{F}_t$ и $W_{t+s} - W_t$ не зависит от $\mathcal{F}_t$ для любых $s, t \geq 0$.

Для заданного винеровского процесса всегда существует такое семейство $\mathcal{F}_t$, для которого это свойство выполняется. В качестве такового можно взять σ-алгебру $\mathcal{F}_{\leq t}$, которая строится следующим образом.

Упражнение 9.1. *Пусть $\mathcal{F}_{\leq t} = \sigma\{W_s, s \in [0,t]\} = \sigma\{\omega : W_s(\omega) \in B, s \in [0,t], B \in \mathcal{B}\}$ (через $\mathcal{B}$ обозначается совокупность борелевских множеств прямой) порождена случайными величинами W_s, $s \in [0,t]$. Тогда $(W_t, \mathcal{F}_{\leq t})$ — винеровский процесс относительно $\mathcal{F}_{\leq t}$.*

В момент времени t σ-алгебру $\mathcal{F}_t$ естественно понимать как совокупность событий "прошлого" и "настоящего".

Мы обсудим построение **стохастического интеграла** или **интеграла Ито**

$$\int_0^T f(t,\omega)\ dW_t$$

от случайных функций. Потребность в этом объекте возникает в различных задачах физики и техники.

Замечание 9.2. *Всюду далее будем предполагать, что все функции вида $f(t,\omega)$, $t \in [0,T], \omega \in \Omega$ измеримы относительно пополнения σ-алгебры $\mathcal{B}(\mathbb{R}) \times \mathcal{F}$ по мере $\lambda \times P$, где λ — мера Лебега на отрезке $[0,T]$. В частности, для таких функций определены двойные интегралы $\int_{[0,T]\times\Omega} f(s,\omega)\ ds\ dP$ по произведению мер $\lambda\ \times P$. По теореме Фубини такие интегралы сводятся к повторным интегралам вида $\mathbb{E}\int_0^T f(t,\omega)\ dt$.*

Впервые интегралы такого вида появились в работе Винера, Пэли и Зигмунда (1933). Там они были определены для неслучайных дифференцируемых функций путем интегрирования по частям

$$\int_0^T f(s)\ dW_s = f(T)W_T - \int_0^T f'(s)W_s\ ds.$$

В более сложной ситуации попытки определить стохастический интеграл для каждой фиксированной траектории наталкиваются на трудность, связанную с тем, что почти все траектории $W_t(\omega)$ имеют бесконечные вариации на отрезке. Поэтому конструкция стохастического интеграла (К. Ито, 1942) осуществляется совершенно другим путем.

Шаг 1. Стохастический интеграл для простых функций.

Простой функцией будем считать функции вида

$$f(t,\omega) = \sum_{k=0}^{n} f_k(\omega) I_{(t_k, t_{k+1}]}, \tag{5}$$

где $0 = t_0 < t_1 < \cdots < t_{n+1} = T$ — разбиение отрезка $[0,T]$, а каждая с.в. f_k измерима относительно $\mathcal{F}_{t_k}$ и обладает конечным вторым моментом $\mathbb{E}f_k^2 < \infty$.

Замечание 9.3. *Дополнительное условие измеримости f_k относительно $\mathcal{F}_{t_k}$ существенно отличает понятие простой функции в смысле, указанном выше, от стандартного понятия ступенчатой функции, известного из анализа.*

Пример 9.4. *Функция $\sum_{i=0}^n W_{t_k} I_{(t_k,t_{k+1}]}$ — простая, а $\sum_{i=0}^n W_{t_{k+1}} I_{(t_k,t_{k+1}]}$ — нет.*

Для простых функций положим:

$$\int_0^t f(s,\omega)\ ds = \sum_{i=0}^n f_k(\omega)(W_{t_{k+1}} - W_{t_k}).$$

Лемма 9.5. *(Изометрия стохастического интеграла). Для любой простой функции f*

$$\mathbb{E}\Bigl(\int_0^T f(t,\omega)\ dW_t\Bigr)^2 = \mathbb{E}\Bigl[\int_0^T f^2(t,\omega) dt\Bigr].$$

Доказательство. Без ограничения общности можно стчитать, что функции f, g строятся с помощью одного и того же разбиения.

$$\mathbb{E}\Bigl(\int_0^T f(t,\omega)\ dW_t\Bigr)^2 = \sum_{i,j} \mathbb{E}\Bigl(f_i f_j (W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j})\Bigr).$$

Несложно видеть, что для $i < j$

$$\mathbb{E}\Bigl(f_i f_j (W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j}) = \mathbb{E}\Bigl[\Bigl(f_i f_j (W_{t_{i+1}} - W_{t_t})\Bigr]\mathbb{E}\Bigl(W_{t_{j+1}} - W_{t_j}\Bigr) = 0.$$

Поэтому

$$\mathbb{E}\Bigl(\int_0^T f(t,\omega)\ dW_t\Bigr)^2 = \sum_{i=1}^n \mathbb{E}(f_i^2 (W_{t_{i+1}} - W_{t_i})^2) = \sum_{i=1}^n \mathbb{E}(f_i^2)\mathbb{E}(W_{t_{i+1}} - W_{t_i})^2$$
$$= \mathbb{E}\Bigl(\sum_{i=1}^n f_i^2 (t_{i+1} - t_i)\Bigr) = \mathbb{E}\Bigl[\int_0^T f^2(t,\omega) dt\Bigr].$$

□

В частности, для пары простых функций f, g

$$\mathbb{E}\bigl(\int_0^T f(t,\omega)\ dW_t \cdot \int_0^T g(t,\omega)\ dW_t\bigr) = \mathbb{E}\int_0^T fg\ dt,$$

$$\mathbb{E}\bigl(\int_0^T f(t,\omega)\ dW_t - \int_0^T g(t,\omega)\ dW_t\bigr)^2 = \mathbb{E}\int_0^T (f-g)^2\ dt.$$

Из последнего равенства следует, что определение интеграла от простой функции не зависит от способа ее представления в виде (5).

Замечание 9.6. *Нетрудно убедиться, что стохастический интеграл линеен: $\int_0^T (c_1 f + c_2 g)\ dW_t = c_1 \int_0^T f\ dW_t + c_2 \int_0^T g\ dW_t$.*

Как мы видим, стохастический интеграл является линейным изометрическим оператором из пространства простых функций, наделенных скалярным произведением $(f,g) \to \mathbb{E}(\int_0^T fg\ dt)$, в гильбертово пространство $L^2(P)$ случайных величин с конечным вторым моментом и скалярным произведением $(f,g) \to \mathbb{E}(fg)$. Поэтому стохастический интеграл естественно продолжить на пополнение простых функций по норме $f \to \Big(\mathbb{E}(\int_0^T f^2\ dt)\Big)^{1/2}$, т.е.

$$\int_0^T f\ dW_t = \lim_{n\to\infty} \int_0^T f_n\ dW_t,$$

где $\{f_n\}$ — такая последовательность простых функций, т.ч. $\lim_n \mathbb{E} \int_0^T \big(f_n - f\big)^2 dt = 0$. Свойство изометричности стохастического интеграла обеспечивает корректность этого определения.

Таким образом, мы определили стохастический интеграл на пополнении $N(0,T)$ простых функций по норме $f \to \Big(\mathbb{E}(\int_0^T f^2\ dt)\Big)^{1/2}$. Очевидно, функции из $N(0,T)$

1) измеримы в смысле замечания 9.2.

2) обладают свойством

$$\mathbb{E}\Big(\int_0^T f^2\ dt\Big) < \infty.$$

Функции из $N(0,T)$ обладают дополнительным важным свойством:

3) Функции из $N(0,T)$ измеримы относительно пополнения σ-алгебры, порожденной множествами вида $([0,T]\cap B)\times F$, где B — борелевское подмножество множества (t,∞), $F \subset \mathcal{F}_t$.

Последнее свойство следует из того, что им обладают простые функций.

Существует простой достаточный признак принадлежности $N(0,T)$.

Теорема 9.7. *Пусть функция $f(t,\omega)$ удовлетворяет свойствам 1)-2) и измерима относительно $\mathcal{F}_t$ при любом $t \in [0,T]$ (т.е., f* **согласована** *с $\mathcal{F}_t$). Тогда $f \in N(0,T)$.*

Доказательство. Достаточно построить последовательность простых функций $f_n \to f$ в $L^2(P)$. Предположим сначала, что f ограничена и имеет непрерывные траектории $t \mapsto f(t,\omega)$ для любого $\omega \in \Omega$. Положим:

$$f_n(t,\omega) = \sum_j f(t_j,\omega)(t_j, t_{j+1}].$$

Очевидно, функция f_n простая. В силу непрерывности путей $t \to f(t,\omega)$ имеем:

$$\sum_j \int_{t_j}^{t_{j+1}} \big(f(t_j,\omega) - f(t,\omega)\big)^2 dt \to 0.$$

В силу ограниченности функции f имеем: $\mathbb{E}\Big(\sum_j \int_{t_j}^{t_{j+1}} \big(f(t_j,\omega) - f(t,\omega)\big)^2 dt\Big) \to 0$ по теореме Лебега.

Предположим теперь, что f ограничена. В силу предыдущего шага достаточно приблизить f случайными функциями с непрерывными траекториями. Для этого рассмотрим такую гладкую неотрицательную функцию φ, что $\varphi = 0$ на $(-\infty,-1] \cup [0,+\infty)$ и $\int_{-\infty}^{\infty} \varphi_n(x)\ dx = 1$. Положим $\varphi_n = n\varphi(nx)$ и

$$f_n(t,\omega) = \int_0^T \varphi_n(s-t) f(s,\omega)\ ds = \int_0^t \varphi_n(s-t) f(s,\omega)\ ds.$$

Случайная функция f_n обладает непрерывными траекториями, ограничена и измерима относительно $\mathcal{F}_t$ при фиксированном t (для доказательства этого воспользуйтесь стандартным определением интеграла Лебега, как предела интегралов от простых функций). Поэтому f_n интегрируема согласно предыдущему шагу. Далее, как хорошо известно из анализа (см, например, В.И. Богачев, **Введение в теорию меры**, 1-й т., теорема 4.2.4), $\int_0^T (f_n(s) - f(s))^p ds \to 0$ для всех $p \geq 1$ (в частности, для $p = 2$). Опять в силу теоремы Лебега (используем ограниченность f и оценку $|\int_0^T g dx| \leq \sup_{[0,T]} g \cdot T$), имеем: $\mathbb{E} \int_0^T (f_n(s) - f(s))^2 ds \to 0$. Таким образом, теорема доказана для ограниченных функций.

На последнем этапе приблизим произвольную функцию $f \in N([0,T])$ ограниченными функциями $f_n = I_{\{|f|<n\}} \cdot f + n \cdot I_{\{f \geq n\}} - n \cdot I_{\{f \leq -n\}}$. Сходимость $f_n \to f$ в $L^2(P)$ следует из теоремы Лебега, так как $|f_n| \leq f$. $\square$

Пример 9.8.

$$\int_0^t W_s \, dW_s = \frac{1}{2} W_t^2 - \frac{1}{2} t.$$

Доказательство. Для любого разбиения $0 = t_0 < t_1 < \cdots < t_n = t$ имеем:

$$\begin{aligned}
W_{t_n}^2 &= (W_{t_n} - W_{t_{n-1}})^2 + 2W_{t_{n-1}}(W_{t_n} - W_{t_{n-1}}) + W_{t_{n-1}}^2. \\
&= (W_{t_n} - W_{t_{n-1}})^2 + 2W_{t_{n-1}}(W_{t_n} - W_{t_{n-1}}) + (W_{t_{n-1}} - W_{t_{n-2}})^2 \\
&+ 2W_{t_{n-2}}(W_{t_{n-1}} - W_{t_{n-2}}) + W_{t_{n-2}}^2 \\
&= \cdots \\
&= \sum_{i=1}^n (W_{t_i} - W_{t_{i-1}})^2 + 2 \sum_{i=1}^n W_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}).
\end{aligned}$$

При измельчении разбиения первая сумма, как мы видели, стремится в среднеквадратичном к t, в то время как вторая сумма сходится (по определению) к стохастическому интегралу $2 \int_0^t W_s \, dW_s$. Последний существует в силу предыдущей теоремы. В пределе получаем нужное равенство. $\square$

10. Лекция 11. Стохастическое интергирование (продолжение).

10.1. **Стохастический интеграл как мартингал.** Будем предполагать, что процесс $f(t, \omega)$ согласован и $\mathbb{E}(\int_0^T f^2(s, \omega) \, ds) < \infty$.

Теорема 10.1. *Существует такая версия стохастического интеграла с переменным верхним пределом $\eta_t = \int_0^t f(s, \omega) \, dW_s$, что*

1) *η_t — мартингал*
2) *для п.в. ω траектории η_t непрерывны.*

Доказательство.

Упражнение 10.2. *Докажите, что теорема верна для простой функции.*

Приблизим теперь функцию f простыми $f = \lim_n f_n$ по $L^2([0,T] \times P)$-норме. Положим

$$I_t^{(n)} = \int_0^t f_n \, dW_t.$$

В силу упражнения $I_t^{(n)}$ — мартингал с непрерывными траекториями. По неравенству Колмогорова и свойству изометрии

$$P(\sup_{0 \leq t \leq T} |I_t^{(n)} - I_t^{(m)}| > \varepsilon) \leq \frac{1}{\varepsilon^2} \mathbb{E}(|I_T^{(n)} - I_T^{(m)}|^2) = \frac{1}{\varepsilon^2} \mathbb{E} \int_0^T (f_n - f_m)^2 \, dt \to 0$$

при $n, m \to \infty$. Выберем теперь подпоследовательность (которую для простоты обозначим снова через $I_t^{(n)}$), удовлетворяющую неравенствам

$$P(\sup_{0\le t\le T} |I_t^{(n)} - I_t^{(n+1)}| > 2^{-n}) < 2^{-n}.$$

По лемме Бореля-Кантелли

$$P(\sup_{0\le t\le T} |I_t^{(n)} - I_t^{(n+1)}| > 2^{-n} \text{ для бесконечного числа } n) = 0.$$

Отсюда следует, что для п.в. ω существует $N(\omega)$, что

$$\sup_{0\le t\le T} |I_t^{(n)} - I_t^{(n+1)}| < 2^{-n}$$

для всех $n > N(\omega)$. Следовательно, $I_t^{(n)}$ сходится равномерно на $[0,T]$ к случайной функции I_t для почти всех ω. Очевидно, I_t имеет непрерывные траектории. Кроме того, так как $I_t^{(n)} \to \int_0^t f_n\, dW_t$ в $L^2(P)$ для всех t, то I_t является версией $\int_0^t f_n\, dW_t$. То, что I_t является мартингалом, следует из того, что $I_t^{(n)}$ — мартингал и $L^2(P)$ сходимости $I_t^{(n)} \to I_t$ для всех $t \in [0,T]$. □

10.2. **Формула Ито.** Изложение следует книге **N. V. Krylov**, Introduction to the theory of stochastic processes. GMS 43. 2002.

Будем говорить, что согласованный процесс ξ_t имеет стохастический дифференциал

$$d\xi_t = \sigma_t dW_t + bdt,$$

если существует представление

$$\xi_t = \xi_0 + \int_0^t \sigma_s dW_s + \int_0^t bds.$$

Лемма 10.3.

$$d\big(\xi_t\eta_t\big) = \eta_t d\xi_t + \xi_t d\eta_t + d\eta_t \cdot d\xi_t,$$

где

$$dtdt = dtdW_t = 0, \quad dW_t \cdot dW_t = dt.$$

Фактически требуется доказать, что

$$\xi_t\eta_t = \xi_0\eta_0 + \int_0^t (\xi_s\tilde{\sigma}_s + \eta_s\sigma_s)dW_s + \int_0^t \big(\xi_s\tilde{b}_s + \eta_s b_s + \tilde{\sigma}_s\cdot\sigma_s\big)ds. \tag{6}$$

при

$$\xi_t = \xi_0 + \int_0^t \sigma_s dW_s + \int_0^t b_s ds, \quad \eta_t = \eta_0 + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \tilde{b}_s ds.$$

Заметим, что интегралы в правой части не всегда определены (и лемму тогда применять нельзя). Достаточные условия существования, например, таковы: $\mathbb{E}\int_0^t(|b_s|^2 + |\tilde{b}_s|^2)\, ds < \infty$ и $|\sigma(t,\omega)| + |\tilde{\sigma}(t,\omega)| < C$. В этом случае из изометрии для стохастического интеграла и неравенства Коши-Буняковского следует, что $\mathbb{E}(\xi_t^2) \le C(\mathbb{E}\xi_0^2 + t\mathbb{E}(\int_0^t b_s^2 ds) + \mathbb{E}\int_0^t \sigma_s^2\, ds) < \infty$. Похожим образом оценивается правая часть (6). Конечно, достаточные условия существования интегралов могут быть и другими. Например, для существования правой части (6) достаточно потребовать конечность

$$\mathbb{E}(\int_0^t |\xi_s|^2|\tilde{\sigma}_s|^2 + |\eta_s|^2|\sigma_s|^2)ds,$$

$$\mathbb{E}\int_0^t (|\xi_s||\tilde{b}_s| + |\eta_s||b_s| + |\tilde{\sigma}_s|\cdot|\sigma_s|)\ ds.$$

Доказательство. Мы докажем лемму для случая $\mathbb{E}\int_0^t(|b_s|^2 + |\tilde{b}_s|^2)\ ds < \infty$ и $|\sigma_t| + |\tilde{\sigma}_t| < C$.

В силу линейности достаточно сконцентрироваться на случаях $d\xi_t = bdt,\ d\xi_t = \sigma dW_t$ и (аналогично для η_t).

1) $d\xi_t = bdt, d\eta_t = \tilde{b}dt$. Следует из стандартного правила анализа или из аргументов пункта 2).

2) $d\xi_t = \sigma dW_t,\ d\eta_t = \tilde{b}dt$. Пусть $0 = t_{m_0} \le t_{m_1} \cdots \le t_{mk_m} = t$ разбиение $[0,t]$. Положим

$$k_m(s) = t_{mi},\ \tilde{k}_m(s) = t_{m,i+1}, \quad \text{если} \quad s \in [t_{mi}, t_{m,i+1}).$$

Очевидно, $k_m(s), \tilde{k}_m(s) \to s$ равномерно на $[0,t]$. Из формулы $ab-cd = (a-c)d+(b-d)a$ получаем

$$\begin{aligned}
\xi_t\eta_t - \xi_0\eta_0 &= \sum_{i=0}^{k_m-1}(\xi_{t_{m,i+1}}\eta_{t_{m,i+1}} - \xi_{t_{m,i}}\eta_{t_{m,i}}) \\
&= \sum_{i=0}^{k_m-1}\eta_{t_{mi}}\int_{t_{mi}}^{t_{m,i+1}}\sigma_s dW_s + \sum_{i=0}^{k_m-1}\xi_{t_{m,i+1}}\int_{t_{mi}}^{t_{m,i+1}}\tilde{b}_s ds \\
&= \int_0^t \eta_{k_m(s)}\sigma_s dW_s + \int_0^t \xi_{\tilde{k}_m(s)}\tilde{b}_s ds.
\end{aligned}$$

В силу непрерывности траекторий процесса ξ, мы получаем, что почти всюду

$$\Big|\int_0^t \xi_{\tilde{k}_m(s)}\tilde{b}_s ds - \int_0^t \xi_s\tilde{b}_s ds\Big| \le \sup_{s\le t}|\xi_{\tilde{k}_m(s)} - \xi_s|\int_0^t |\tilde{b}_s|\ ds \to 0.$$

Из изометрии стохастического интеграла следует, что

$$\int_0^t \eta_{k_m(s)}\sigma_s dW_s \to \int_0^t \eta_s\sigma_s dW_s$$

в $L^2(P)$ (и по мере) для всех $t\in[0,T]$. Действительно,

$$\mathbb{E}\Big(\int_0^t(\eta_{k_m(s)} - \eta_s)\sigma_s dW_s\Big)^2 = \mathbb{E}\int_0^t(\eta_{k_m(s)} - \eta_s)^2\sigma_s^2 ds \le C^2\mathbb{E}\int_0^t(\eta_{k_m(s)} - \eta_s)^2 ds.$$

Заметим теперь, что

$$\eta_s - \eta_{k_m(s)} = \int_{k_m(s)}^s \tilde{\sigma}_r dW_r + \int_{k_m(s)}^s b_r dr.$$

Поэтому

$$\mathbb{E}(\eta_s - \eta_{k_m(s)})^2 \le 2\mathbb{E}\Big(\int_{k_m(s)}^s \tilde{\sigma}_r^2 dr + (s - k_m(s))\int_{k_m(s)}^s b_r^2 dr\Big) \le C(s - k_m(s)).$$

Очевидно, $\mathbb{E}\int_0^t(\eta_{k_m(s)} - \eta_s)^2 ds \to 0$.

3) $d\xi_t = \sigma dW_t,\ d\xi_t = \tilde{\sigma}dW_t$. Зафиксируем $t\in[0,T]$ и докажем, что

$$\xi_t\eta_t = \xi_0\eta_0 + \int_0^t(\xi_s\tilde{\sigma}_s + \eta_s\sigma_s)dW_s + \int_0^t \tilde{\sigma}_s\cdot\sigma_s\ ds.$$

Докажем утверждение в ситуации, что $\sigma, \tilde{\sigma}$ — простые согласованные функции вида $\sum_i \zeta(t_i, \omega) I_{[t_i, t_{i+1})}$, $\mathcal{T} = \{t_1, \cdots, t_r\}$ (общий случай получается с помощью аппроксимаций и оценок). Как и выше, рассмотрим последовательность разбиений, причем будем считать, что все точки t_i входят в эти разбиения. Применим формулу

$$ab - cd = (a-c)d + (b-d)c + (a-c)(b-d).$$

Имеем для всех точек $t_q < t_{mr} < t_{q+1}$, $t_q, t_{q+1} \in \mathcal{T}$

$$\begin{aligned}
\xi_{t_{q+1}}\eta_{t_{q+1}} - \xi_{t_{q+1}}\eta_{t_{q+1}} &= \sum_r (\xi_{t_{m,r+1}}\eta_{t_{m,r+1}} - \xi_{t_{mr}}\eta_{t_{mr}}) \\
&= \sum \eta_{t_{mr}} \int_{t_{mr}}^{t_{m,r+1}} \sigma_s dW_s + \sum \xi_{t_{mr}} \int_{t_{mr}}^{t_{m,r+1}} \tilde{\sigma}_s dW_s \\
&+ \sum \int_{t_{mr}}^{t_{m,r+1}} \sigma_s dW_s \int_{t_{mr}}^{t_{m,r+1}} \tilde{\sigma}_s dW_s \\
&= \int_{t_q}^{t_{q+1}} \eta_{k_m(s)} \sigma_s dW_s + \int_{t_q}^{t_{q+1}} \xi_{k_m(s)} \tilde{\sigma}_s dW_s \\
&+ \sigma_{t_q} \tilde{\sigma}_{t_{q+1}} \sum (W_{t_{m,r+1}} - W_{t_{mr}})^2.
\end{aligned}$$

В силу оценок выше и сходимости квадратичной вариации винеровского процесса, это выражение сходится по вероятности к

$$\int_{t_q}^{t_{q+1}} \eta_s \sigma_s dW_s + \int_{t_q}^{t_{q+1}} \xi_s \tilde{\sigma}_s dW_s + \sigma_{t_q} \tilde{\sigma}_{t_q} (t_{q+1} - t_q).$$

Осталось заметить, что $\sigma_{t_q} \tilde{\sigma}_{t_q} (t_{q+1} - t_q) = \int_{t_q}^{t_{q+1}} \sigma_s \tilde{\sigma}_s ds$ (в силу ступенчатости $\sigma, \tilde{\sigma}$) и взять сумму по q. □

Из доказанной леммы следует **формула Ито**

$$df(t, \xi_t) = f_t(t, \xi_t) dt + f_x(t, \xi_t) d\xi_t + \frac{1}{2} f_{xx}(t, \xi_t) (d\xi_t)^2.$$

Равенство выполнено, если правая часть определена.

Приведем эвристические аргументы для ее обоснования. Достаточно показать, что

1) Формула выполнена для функций вида $f(t)(aW_t + b)$ (проверяется непосредственно)

2) Если формула выполнена для $u(t, x)$, $v(t, x)$, то она выполнена и для произведений $u(t, x)v(t, x)$. Действительно,

$$\begin{aligned}
d(u(t, \xi_t) v(t, \xi_t)) &= du(t, \xi_t) \cdot v(t, \xi_t) + u(t, \xi_t) \cdot dv(t, \xi_t) + du(t, \xi_t) dv(t, \xi_t) \\
&= (u_t + u_x d\xi_t + \frac{1}{2} u_{xx} (d\xi_t)^2) v + (v_t + v_x d\xi_t + \frac{1}{2} v_{xx} (d\xi_t)^2) u \\
&+ u_x v_x (d\xi_t)^2 \\
&= (uv)_t(t, \xi_t) dt + (uv)_x(t, \xi_t) d\xi_t + \frac{1}{2} (uv)_{xx}(t, \xi_t) (d\xi_t)^2.
\end{aligned}$$

Из 2) немедленно следует, что формула выполнена для многочленов от переменных t, x. Отсюда можно вывести (например, приближая непрерывно дифференцируемую функцию u вместе с производными равномерно на любом шаре многочленами и контролируя рост возникающих процессов), что она выполнена для непрерывно дифференцируемых функций.

Пример 10.4. *Докажите, что процесс*

$$\xi_t = e^{W_t - \frac{1}{2}t}$$

является мартингалом.

Доказательство. Найдем $d\xi_t$. По формуле Ито $d\xi_t = e^{W_t-\frac{1}{2}t}dW_t$. Поэтому $\xi_t = 1 + \int_0^t e^{W_s-\frac{1}{2}s}dW_s = 1 + \int_0^t \xi_s dW_s$. Процесс является мартингалом, потому что является стохастическим интегралом. □

11. Лекция 12. Стохастические дифференциальные уравнения.

11.1. **Процесс Орнштейна-Уленбека.** Винеровский процесс как модель движения взвешеных частиц в жидкости не был вполне удовлетворительным с физической точки зрения. В 30-х годах была предложена другая модель, учитывавшая ньютоновское взаимодействие частиц. Поведение частицы описывалось уравнением Ланжевена

$$dv(t) = -\beta v(t)dt + dW_t,$$

где $\beta > 0$ — некоторый коэффициент, v — скорость частицы. Формально это уравнение можно переписать таким образом:

$$ma = m\frac{dv(t)}{dt} = -m\beta v(t) + m\frac{dW_t}{dt}.$$

В математике и физике решение этого уравнения получило название процесса Орнштейна-Уленбека.

Процессом Орнштейна-Уленбека назывется решение стохастического дифференциального уравнения

$$d\xi_t = -\xi_t dt + \sqrt{2}dW_t, \quad \xi_0 = x.$$

Для решения этого уравнения найдем дифференциал процесса $e^t\xi_t$:

$$d(e^t\xi_t) = e^t\xi_t dt + e^t(-\xi_t dt + \sqrt{2}dW_t) = \sqrt{2}e^t dW_t.$$

Следовательно,

$$e^t\xi_t = x + \sqrt{2}\int_0^t e^s dW_s$$

и

$$\xi_t = xe^{-t} + \sqrt{2}e^{-t}\int_0^t e^s dW_s.$$

Из этого соотношения видно, что ξ_t — гауссовский процесс. Следовательно, его конечномерные распределения полностью определяются средним $\mathbb{E}\xi_t$ и функцией ковариации

$$K(s,t) = \mathbb{E}(\xi_t - \mathbb{E}\xi_t)(\xi_s - \mathbb{E}\xi_s).$$

Имеем:

$$\mathbb{E}\xi_t = xe^{-t},$$

$$\mathbb{E}(\xi_t - \mathbb{E}\xi_t)(\xi_s - \mathbb{E}\xi_s) = 2e^{-t-s}\mathbb{E}\Big(\int_0^t e^u dW_u \int_0^s e^u dW_u\Big) = 2e^{-t-s}\int_0^{\min(t,s)} e^{2u}du$$
$$= e^{-t-s}\big(e^{2\min(t,s)} - 1\big)$$

(мы пользуемся изометрией стохастического интеграла). В частности, $D\xi_t = (1 - e^{-2t})$.

Полугруппа Орнштейна-Уленбека T_tf определяется следующим образом:

$$T_tf(x) = \mathbb{E}f(\xi_t).$$

Зная распределение ξ_t, нетрудно написать явное представление T_t

$$T_t f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(xe^{-t} + \sqrt{1-e^{-2t}}y) e^{-\frac{y^2}{2}} dy.$$

Задача 11.1. *Используя явное представление полугруппы Орнштейна-Уленбека доказать следующие ее свойства*

1) (T_t — полугруппа)

$$T_{t+s} f = T_t(T_s f)$$

2) (Генератор T_t) $u(t,x) = T_t f$ является решением параболического уравнения в частных производных

$$u_t = u_{xx} - xu_x, \quad u(0,x) = f(x)$$

3) (Инвариантная мера) стандартное гауссовское распределение $\gamma = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\ dy$ является инвариантной мерой относительно T_t:

$$\int T_t f \cdot g\ d\gamma = \int f \cdot T_t g\ d\gamma.$$

4) (Эргодическое свойство)

$$\lim_{t\to\infty} T_t f = \int f\ d\gamma.$$

11.2. **Теорема существования и единственности.** Итак, на примере процесса Орнштейна-Уленбека мы убедились в полезности изучения стохастических дифференциальных уравнений. Стохастическим дифференциальным уравнением называется уравнение вида

$$d\xi_t = \sigma(\xi_t) dW_t + \beta(\xi_t) dt, \quad \xi_{t_0} = \eta,$$

где η — некоторая случайная величина, измеримая относительно $\mathcal{F}_{t_0}$. Это уравнение следует понимать как интегральное

$$\xi_t = \eta + \int_{t_0}^{t} \sigma(\xi_s) dW_s + \int_{t_0}^{t} \beta(\xi_s) ds. \tag{7}$$

Функции σ и β носят названия коэффициент диффузии и коэффициент сноса.

Теорема 11.2. *Предположим, что σ и β удовлетворяют условию Липшица, т.е. для некоторой константы K*

$$|\sigma(x) - \sigma(y)| \le K|x-y|, \quad |\beta(x) - \beta(y)| \le K|x-y|.$$

Тогда для любого отрезка $[t_0, T]$ и любого начального значения η с $\mathbb{E}\eta^2 < \infty$ существует единственное (с точностью до стохастической эквивалентности) решение (7) на $[t_0, T]$.

Замечание 11.3. *Доказательство ниже приведено для случая $\eta = x_0$, $t_0 = 0$. Доказательство общего случая совершенно аналогично.*

Доказательство. (набросок, подробно см., например, в **Вентцель А.Д.**, Курс теории случайных процессов. М. Наука, 1975.).

Существование. Решения строятся методом последовательных итераций:

$$\xi_t^{(n)} = x_0 + \int_0^t \sigma(\xi_s^{(n-1)})\ dW_s + \int_0^t \beta(\xi_s^{(n-1)})\ ds. \tag{8}$$

С помощью изометрии Ито и неравенства Коши-Буняковского доказываем оценки

$$\mathbb{E}\big|\xi_t^{(1)} - x_0\big|^2 \le CK^2(t+t^2) \le CK^2T(1+t)$$

$$\mathbb{E}\big|\xi_t^{(n)} - \xi_t^{(n-1)}\big|^2 \le CK^2(1+t)\int_0^t \mathbb{E}\big|\xi_s^{(n-1)} - \xi_s^{(n-2)}\big|^2\, ds$$

(C зависит только от x_0^2).

Индукция:

$$\mathbb{E}\big|\xi_t^{(n)} - \xi_t^{(n-1)}\big|^2 \le (CK^2)^n T\frac{(1+t)^{n+1}}{n!}. \tag{9}$$

Следовательно, $\sum_{n=1}^{\infty}\mathbb{E}^{1/2}\big|\xi_t^{(n)} - \xi_t^{(n-1)}\big|^2 < \infty$ и ряд $\sum_{n=1}^{\infty}\xi_t^{(n)} - \xi_t^{(n-1)}$ сходится в $L^2(P)$. Докажем, что сходимость равномерна по t. Действительно,

$$P\Big(\max_t |\xi_t^{(n+1)} - \xi_t^{(n)}| \ge 1/2^n\Big) \le P\Big(\max_{u\le t}\big|\int_0^u \sigma(\xi_s^{(n)}) - \sigma(\xi_s^{(n-1)})\, dW_s\big| \ge 1/2^{n+1}\Big) + P\Big(\max_u\big|\int_0^u \beta(\xi_s^{(n)}) - \beta(\xi_s^{(n-1)})\, ds\big| \ge 1/2^{n+1}\Big)$$

По неравенству Колмогорова для субмартингалов первое слагаемое не превосходит

$$2^{2(n+1)}\mathbb{E}\Big(\big|\int_0^t \sigma(\xi_s^{(n)}) - \sigma(\xi_s^{(n-1)})\, dW_s\big|\Big)^2 \le 2^{2(n+1)}K^2\int_0^t \mathbb{E}\big|\xi_t^{(n)} - \xi_t^{(n-1)}\big|^2\, ds.$$

Второе не превосходит (неравенство Чебышева + неравенство Коши-Буняковского)

$$P\Big(\int_0^t |\beta(\xi_s^{(n)}) - \beta(\xi_s^{(n-1)})|\, ds \ge 1/2^{n+1}\Big) \le 2^{2(n+1)}\mathbb{E}\Big(\int_0^t |\beta(\xi_s^{(n)}) - \beta(\xi_s^{(n-1)})|\, ds\Big)^2 \le 2^{2(n+1)}K^2 t\int_0^t \mathbb{E}\big|\xi_s^{(n)} - \xi_s^{(n-1)}\big|^2\, ds.$$

Равномерная сходимость почти всюду следует из (9) и леммы Бореля-Кантелли.

Итак, $\xi_t^{(n)} \to \xi_t$ в $L^2(P)$ и равномерно на отрезке для почти всех траекторий. Переходя к пределу в (8) (обоснуйте), получаем

$$\xi_t = x_0 + \int_0^t \sigma(\xi_s)\, dW_s + \int_0^t \beta(\xi_s)\, ds.$$

Итак, ξ_t — решение СДУ, траектории непрерывны п.в., ξ_t — $\mathcal{F}_t$-измеримо.

Единственность: пусть есть два различных решения ξ, η.

$$\mathbb{E}|\xi_t - \eta_t|^2 \le C(K^2, T, x_0)\int_0^t \mathbb{E}(\xi_s - \eta_s)^2\, ds.$$

Единственность следует из леммы: если $0 \le f(t) \le C\int_0^t f(s)\, ds$, то $f(s) = 0$. □

12. Лекция 13.

12.1. **Уравнение Колмогорова.** В этой лекции мы обсудим связь винеровского процесса с уравнениями в частных производных.

Теорема 12.1. *Пусть $f(x)$ — дважды непрерывно дифференцируемая функция, ограниченная вместе со своими производными до порядка 2 включительно. Тогда функция*

$$u(t,x) = \mathbb{E}(f(x+W_t))$$

является решением уравнения теплопроводности (heat equation)

$$u_t = \frac{1}{2} u_{xx} \tag{10}$$

с начальным условием $u(0,x) = f(x)$.

Доказательство. Соотношение $u_t = \frac{1}{2}u_{xx}$ следует из явного представления

$$u(x,t) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} f(x+y) e^{-\frac{y^2}{2t}}\, dy = \int_{-\infty}^{\infty} f(x+y) p_t(y)\, dy,$$

где

$$p_t(y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}}.$$

Явные выкладки показывают, что

$$\frac{d}{dt} p_t = \frac{1}{2} \frac{\partial^2}{\partial y^2} p_t, \quad t \neq 0.$$

Дифференцируя интеграл по параметру и интегрируя по частям, получаем

$$\begin{aligned}
\frac{d}{dt} u(x,t) &= \frac{d}{dt} \int_{-\infty}^{\infty} f(x+y) p_t(y)\, dy = \frac{1}{2} \int_{-\infty}^{\infty} f(x+y) \frac{\partial^2}{\partial y^2} p_t(y)\, dy \\
&= \frac{1}{2} \int_{-\infty}^{\infty} \frac{\partial^2}{\partial y^2} f(x+y) p_t(y)\, dy = \frac{1}{2} \int_{-\infty}^{\infty} \frac{\partial^2}{\partial x^2} f(x+y) p_t(y)\, dy \\
&= \frac{1}{2} \frac{\partial^2}{\partial x^2} \int_{-\infty}^{\infty} f(x+y) p_t(y)\, dy = \frac{1}{2} \frac{\partial^2}{\partial x^2} u(t,x).
\end{aligned}$$

Далее, $u(0,x) = \lim_{t\to 0} u(t,x) = \lim_{t\to 0} \mathbb{E}(f(x+W_t)) = \mathbb{E}(f(x)) = f(x)$ (последнее соотношение также следует из слабой сходимости $p_t(y)dy \to \delta_0$, $t \to 0$).

Для $t = 0$ надо воспользоваться соотношением

$$\begin{aligned}
\frac{d}{dt} u(t,x)|_{t=0} &= \lim_{t\to 0} \frac{u(t,x) - u(0,x)}{t} = \lim_{t\to 0} \frac{1}{t} \int \big(f(x+y) - f(x)\big) p_t(y)\, dy \\
&= \lim_{t\to 0} \frac{1}{t} \int \big(f(x+\sqrt{t}y) - f(x)\big) p_1(y)\, dy.
\end{aligned}$$

Далее, заметим, что $\int y p_1(y)\, dy = 0$, поэтому

$$\begin{aligned}
\frac{d}{dt} u(t,x)|_{t=0} &= \lim_{t\to 0} \frac{1}{t} \int \big(f(x+\sqrt{t}y) - f(x) - \sqrt{t} f'(x) y\big) p_1(y)\, dy \\
&= \int \frac{1}{2} f_{xx}(x) y^2 p_1(y)\, dy = \frac{1}{2} f_{xx}(x).
\end{aligned}$$

□

Уравнение теплопроводности является моделью распространения тепла в телах и процесса диффузии (другое название (10) — уравнение диффузии). Рассмотрим узкий длинный цилиндр, направленный вдоль оси x, с жидкостью и взвесью мелких частиц в ней, движущимся согласно закону винеровского процесса. Пусть $f(x,t)\Delta x$ — число частиц, заключенных в участке цилиндра между x и $x + \Delta x$ (f можно интерпретировать как плотность) в момент t. Число частиц в этом участке через малый период времени Δt равно

$$f(x, t+\Delta t)\Delta x = \Delta x \int_{-\infty}^{\infty} f(x-s,t) p_{\Delta t}(s) ds.$$

Действительно, это следует из того, что частица, находящаяся на расстоянии s от $[x, x+\Delta x]$, попадает туда через промежуток времени Δt с вероятностью $\approx \Delta x \cdot p_{\Delta t}(s)$. Усредняя по вероятности, получаем среднее число частиц. Разлагая в ряд по t выражение в левой части и в ряд по x выражение в правой части, получаем

$$f(x,t) + \Delta t f_t(t,x) + \cdots = \int_{-\infty}^{\infty} \Big(f(x,t) - s f_x(x,t) + \frac{1}{2} s^2 f_{xx}(x,t) + \cdots \Big) p_{\Delta t}(s) ds.$$

Принимая во внимание, что $\int_{-\infty}^{\infty} s p_{\Delta t}(s) ds = 0$, $\int_{-\infty}^{\infty} s^2 p_{\Delta t}(s) ds = \Delta t$, получаем искомое уравнение $f_t(t,x) = \frac{1}{2} f_{xx}(t,x)$.

12.2. **Марковское свойство решений СДУ. Уравнение Колмогорова.** Наша ближайшая задача — вывести обобщение доказанного факта, известного как **уравнение Колмогорова**. Мы не сможет обсудить полное доказательство этого уравнения, так как оно требует применения довольно сложной техники. Поэтому мы ограничимся обсуждением основной идеи доказательства.

Рассмотрим решение дифференциального стохастического уравнения

$$\xi_t = x + \int_0^t \sigma(s, \xi_s) dW_s + \int_0^t b(s, \xi_s) ds. \tag{11}$$

Определим функцию

$$u(t,x) = \mathbb{E} f(\xi_t).$$

Чтобы подчеркнуть, что процесс стартует из точки x, пишут также

$$u(t,x) = \mathbb{E}^x (f(\xi_t)).$$

По формуле Ито

$$df(\xi_t) = \sigma(t, \xi_t) f'(\xi_t) dW_t + Lf(\xi_t) dt,$$

где

$$Lf(x) = \frac{1}{2} \sigma^2(t,x) f_{xx}(x) + b(t,x) f_x(x).$$

Следовательно

$$f(\xi_t) = f(\xi_s) + \int_s^t \sigma(r, \xi_r) f'(\xi_r) dW_r + \int_s^t Lf(\xi_r) dr.$$

Более общим образом, для марковского момента τ имеем

$$f(\xi_\tau) = f(\xi_s) + \int_s^\tau \sigma(r, \xi_r) f'(\xi_r) dW_r + \int_s^\tau Lf(\xi_r) dr.$$

Взяв математическое ожидание от обеих частей получим **формулу Дынкина**

$$\mathbb{E} f(\xi_\tau) = u(s,x) + \mathbb{E} \int_s^\tau Lf(\xi_r) dr$$

(мы учли, что стохастический интеграл является мартингалом и имеет нулевое математическое ожидание).

Пример 12.2. *Пусть τ — первый момент выхода винеровского процесса из отрезка (a,b), $a < 0, b > 0$. Тогда $\mathbb{E}(b - W_\tau)(W_\tau - a) = 0$ по определению τ. По формуле Дынкина*

$$0 = \mathbb{E}(b - W_\tau)(W_\tau - a) = -ab - \mathbb{E} \int_0^\tau \ dt.$$

Следовательно $\mathbb{E}\tau = -ab$.

Замечание 12.3. *Марковские моменты требуют аккуратного обращения. Формально формула Дынкина следует из мартингальности процесса. Аналогично из мартингальности $\int_0^t f dW_s$ и изометрии Ито следуют* **соотношения Вальда***:*

$$\mathbb{E}\Big(\int_0^\tau f dW_s\Big) = 0, \quad \mathbb{E}\Big(\int_0^\tau f dW_s\Big)^2 = \mathbb{E}\Big(\int_0^\tau f^2 ds\Big)^2.$$

Но они выполнены не всегда (например, $\mathbb{E}W_\tau = a \neq 0$, τ — момент достижения процессом W_t точки a). Достаточным условием выполнимости этих равенств является $\mathbb{E}\Big(\int_0^\tau f^2 ds\Big)^2 < \infty$.

Чтобы обосновать рассуждение предыдущего примера, заметим, что для любого $T > 0$ имеем

$$\mathbb{E}(\min(\tau, T)) = \mathbb{E}W^2_{\min(\tau,T)} \le (-a+b)^2$$

(первое равенство вытекает из ограниченности $\min(\tau, T)$). Устремляя T к бесконечности, получаем по теореме о монотонной сходимости, что $\mathbb{E}\tau < \infty$. Отсюда следуют искомые соотношения.

Продолжим обсуждение диффузионных процессов. Выполено следующее **эволюционное** свойство

$$\xi_{t+s}(x) = \xi_s(\xi_t(x)).$$

Из него следует следующая эвристическая формула

$$u(t+s, x) = \mathbb{E}^x f(\xi_{t+s}) = \mathbb{E}^x(\mathbb{E}^{\xi_s}(f(\xi_t))) = \mathbb{E}^x(u(t, \xi_s)).$$

Это **полугрупповое** свойство интуитивно ясно (процесс в момент времени t "забывает" о своем прошлом и развивается далее точно также, как и процесс, стартующий из (случайной!) точки ξ_t). Это одна из форм **марковских** свойств диффузионных процессов. Мы не будем останавливаться на подробных доказательствах. Несмотря на интуитивную ясность, технические обоснования утверждений такого рода довольно сложны.

Замечание 12.4. *Введем линейный оператор $T_t f(x) = \mathbb{E}^x f(\xi_t)$ на пространстве функций. Соотношение выше удобно записывать в следующей форме*

$$T_{t+s} f(x) = T_t(T_s f(x)).$$

Применяя полугрупповое свойство и формулу Дынкина, получаем **уравнение Колмогорова**

$$\begin{aligned}\partial_t u(t,x) &= \lim_{s\to 0} \frac{u(t+s,x) - u(t,x)}{s} = \lim_{s\to 0} \frac{1}{s}\big(\mathbb{E}^x u(t, \xi_s) - u(t,x)\big)\\ &= \lim_{s\to 0} \frac{1}{s}\mathbb{E}^x\Big(\int_0^s Lu(t, \xi_r) dr\Big) = Lu(t,x).\end{aligned}$$

Мы использовали в последнем равенстве соотношение $\xi_t = x$. Итак, функция $u(t,x)$ является решением следующего **уравнения в частных производных**

$$u_t(t,x) = \frac{1}{2}\sigma^2(t,x) u_{xx}(t,x) + b(t,x) u_x(t,x).$$

Список литературы

[1] Богачев В.И., Основы теории меры, тт. 1-2, Москва-Ижевск, НИЦ Регулярная и хаотическая динамика, 2006.
[2] Вентцель А.Д., Курс теории случайных процессов. М. Наука, 1975.
[3] Кельберт М.Я., Сухов Ю.М.. Вероятность и статистика в примерах и задачах. 2-й том, МЦНМО 2009.
[4] Ширяев А.Н., Вероятность. тт. 1-2. М. МЦНМО, 2007.
[5] Ширяев А.Н., Задачи по теории вероятностей. М. МЦНМО, 2006.
[6] Grimmet G., Stirzaker D., One thousand exercises in probability, Oxford University Press, 2001.
[7] Grimmet G., Stirzacker D., Probability and random processes.Oxford University Press, 2001.
[8] Stirzaker D., Elementary probability, Cambrige University Press, 2003.