ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)»

М.Е. Жуковский, И.В. Родионов, Д.А. Шабанов

ОСНОВЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

Учебное пособие

МОСКВА
МФТИ
2017

УДК 519.21
ББК 22.171я73
Ж86

Рецензенты:
В.И. Питербарг

Жуковский, М.Е., Родионов, И.В., Шабанов, Д.А.

Ж86 Основы теории случайных процессов: учебное пособие / Жуковский, М.Е., Родионов, И.В., Шабанов, Д.А. – М.: МФТИ, 2017. – 119 с.

ISBN 978-5-7417-0625-1

В учебном пособии отражено содержание курса по случайным процессам, который авторы ведут у студентов факультета инноваций и высоких технологий Московского физико-технического института (государственного университета) в пятом семестре. Каждая глава пособия содержит определенный теоретический материал, а также набор задач по соответствующей теме с решениями. В конце каждой главы приведен список упражнений. Сформулированные в пособии теоремы и теоретические утверждения необходимы для решения задач, их доказательства не приводятся. Пособие охватывает семестровый курс случайных процессов и рассчитано на студентов математических и физических специальностей высших учебных заведений.

УДК 519.21(075)
ББК 22.171я73

ISBN 978-5-7417-0625-1

Оглавление

Предисловие

Авторы считают необходимым долгом выразить признательность и благодарность своим учителям — профессорам кафедры теории вероятностей механико-математического факультета МГУ имени М. В. Ломоносова Екатерине Вадимовне Булинской и Александру Вадимовичу Булинскому. Под их руководством авторы изучали теорию случайных процессов, будучи студентами, аспирантами, а затем и сотрудниками кафедры. Многие задачи настоящего пособия были когда-то предложены для решения авторам Екатериной Вадимовной и Александром Вадимовичем. Благодаря их педагогическому таланту и стало возможным появление данной книги.

М. Е. Жуковский,
И. В. Родионов,
Д. А. Шабанов.

Введение

Теория случайных процессов сформировалась как самостоятельная математическая область в первой половине XX века. Предметом ее изучения стали вероятностные временные математические модели физических явлений, которые невозможно было описать с помощью обычных аналитических методов. Классический пример подобного явления — открытие Р. Броуном в XIX веке эффекта «броуновского движения» частиц пыльцы в воде, который не удавалось разумно описать с помощью господствовавших в то время методов ньютоновской механики. В настоящее время теория случайных процессов имеет обширные применения как в различных естественных науках (физика, биология и др.), так и в сугубо практических областях (финансы, страхование и пр.).

Пособие охватывает семестровый курс случайных процессов для математических и физических специальностей высших учебных заведений. Каждая глава пособия содержит теоретический материал и набор задач с решениями по соответствующей теме. В конце главы приведен список упражнений для самостоятельной работы. Сформулированные в пособии теоремы и теоретические утверждения необходимы для решения задач, но их доказательства не приводятся. Пособие рассчитано на читателя, хорошо владеющего основами математического анализа и теории вероятностей.

1. Основные понятия теории случайных процессов

Теория случайных процессов — это раздел теории вероятностей, связанный с математическим анализом случайных явлений, в которых присутствует *фактор времени*. Формально случайный процесс можно определить как набор случайных величин, индексированных некоторым множеством.

▲ **Определение 1.1.** Пусть T — некоторое множество. Тогда набор $X = (X_t, t \in T)$ случайных величин $X_t = X_t(\omega)$, заданных на одном и том же вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P})$ для любого $t \in T$, называется *случайным процессом* на множестве T.

Множество T из определения естественно интерпретировать как *время*, тем самым в каждый момент времени t имеется некоторая случайная величина $X_t : \Omega \to \mathbb{R}$. Отметим, что в общем случае X_t могут быть случайными элементами с различными множествами значений.

При определении случайного процесса важно добавить, что, по аналогии со случайными величинами, случайный процесс порождает σ-алгебру.

▲ **Определение 1.2.** Пусть $X = (X_t, t \in T)$ — случайный процесс. Тогда σ-алгеброй, *порожденной* этим процессом, называется минимальная σ-алгебра, содержащая все σ-алгебры $\mathcal{F}_{X_t}$, т.е.

$$\sigma(X_t, t \in T) := \sigma\left(\bigcup_{t \in T} \mathcal{F}_{X_t}\right).$$

На случайный процесс $X = (X_t(\omega), t \in T)$ можно смотреть как на функцию двух переменных $X = X(t, \omega)$, одна их которых, t, индексирует время, а вторая, ω, — случайность. При фиксированном t мы получаем функцию от ω — случайную величину X_t. Если же, наоборот, зафиксировать ω, то получится функция от t.

▲ **Определение 1.3.** При фиксированном $\omega = \omega_0$ функция

$$\tilde{X}_{\omega_0}(t) = X_t(\omega)\Big|_{\omega=\omega_0}$$

на T называется *траекторией*, или *реализацией*, случайного процесса $X = (X_t, t \in T)$.

Именно вероятностный анализ траекторий случайных процессов и составляет главный интерес в изучаемой нами теории.

Случайные процессы принято классифицировать по «природе» множества T. Пусть $X = (X_t, t \in T)$ — случайный процесс.

- Если $T = [a, b], (a, b), [a, +\infty), \mathbb{R}, \mathbb{R}_+$ и так далее (конечный или бесконечный интервал с границами или без), то процесс X называется процессом *с непрерывным временем.*

- Если $T \subset \mathbb{Z}$ (например, $T = \mathbb{Z}_+, \mathbb{Z}$), то процесс X называется процессом с *дискретным* временем.

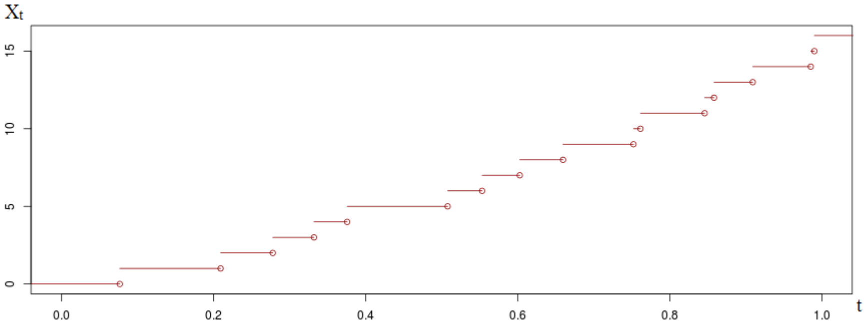
- Если $T \subset \mathbb{R}^d, d > 1$, то процесс X называется *случайным полем.*

Приведем ряд интересных примеров случайных процессов.

▲ **Пример 1.1.** Пусть $\{\xi_n, n \in \mathbb{N}\}$ — независимые случайные векторы размерности m. Тогда процесс с дискретным временем $(S_n = \xi_1 + \cdots + \xi_n, n \in \mathbb{Z}_+)$, где $S_0 = 0$, называется *случайным блужданием.*

Физическая модель: хаотическое дискретное движение в пространстве, например, прыжки кузнечика.

▲ **Пример 1.2.** Пусть $\{\xi_n : n \in \mathbb{N}\}$ — независимые одинаково распределенные случайные величины, $\xi_n \geqslant 0$, $\xi_n \neq \text{const}$ п. н., $S_n = \xi_1 + \cdots + \xi_n$, $S_0 = 0$. Тогда процесс $(X_t = \sup\{n : S_n \leqslant t\}, t \geqslant 0)$ называется *процессом восстановления*, построенным по случайным величинам $\{\xi_n, n \in \mathbb{N}\}$. Всюду далее (если не оговорено иного), говоря о процессе восстановления, построенном по случайным величинам $\{\xi_n, n \in \mathbb{N}\}$, мы будем иметь в виду, что эти случайные величины обладают описанными выше свойствами. Нарисуем траекторию процесса восстановления. Ясно, что его значения — это целочисленные неотрицательные числа, т.е. траектория не убывает. Тогда получается следующая скачкообразная картина.



▲ **Задача 1.1.** Докажите, что с вероятностью единица процесс восстановления конечен в каждый момент времени.

Решение

Пусть сначала $\mathsf{E}\xi_i > 0, \mathsf{E}\xi_i < +\infty$ для любого i.
Если $X_t = +\infty$, то $S_n \leqslant t$ для любого $n \in \mathbb{N}$.

$$\{X_t = +\infty\} = \Big\{\sup\{n : S_n \leqslant t\} = +\infty\Big\} = \bigcap_{n=1}^{\infty}\{S_n \leqslant t\}.$$

Так как $\{S_n \leqslant t\} \supset \{S_{n+1} \leqslant t\}$, то из непрерывности вероятностной меры следует

$$\mathsf{P}(X_t = +\infty) = \lim_{n\to+\infty} \mathsf{P}(S_n \leqslant t).$$

Положим $a = \mathsf{E}\xi_i$. Согласно усиленному закону больших чисел,

$$\frac{S_n}{n} \xrightarrow{\text{п.н.}} a \;\Rightarrow\; \frac{S_n}{n} \xrightarrow{\text{d}} a.$$

Тогда

$$\mathsf{P}(S_n \leqslant t) = \mathsf{P}\left(\frac{S_n}{n} \leqslant \frac{t}{n}\right) \leqslant |\text{при больших } n| \leqslant$$
$$\leqslant \mathsf{P}\left(\frac{S_n}{n} \leqslant \frac{a}{2}\right) \to \mathsf{P}\left(a \leqslant \frac{a}{2}\right) = 0.$$

Значит, для любого фиксированного $t > 0$ выполнено равенство $\mathsf{P}(X_t = +\infty) = 0$. Заметим, что X_t — неубывающий процесс (все траектории процесса являются неубывающими функциями). Следовательно,

$$\mathsf{P}(\exists t : X_t = +\infty) = \mathsf{P}(\exists n \in \mathbb{N} : X_n = +\infty) \leqslant \leqslant \sum_{n=1}^{\infty} \mathsf{P}(X_n = \infty) = 0.$$

Если же $\mathsf{E}\xi_i = +\infty$, то положим $\tilde{\xi}_i = \min(\xi_i, 1)$. Очевидно, $0 < \mathsf{E}\tilde{\xi}_i < +\infty$. Тогда $X_t \leqslant \tilde{X}_t$, где $\tilde{X}_t$ — процесс восстановления $\tilde{\xi}_i$, откуда

$$\mathsf{P}(\exists t : X_t = +\infty) \leqslant \mathsf{P}(\exists t : \tilde{X}_t = +\infty) = 0.$$

Задача решена.

Где может возникнуть процесс восстановления?

Физическая модель: «модель перегорания лампочки». Пусть ξ_n — случайная величина, равная времени работы лампочки с номером n, тогда X_t показывает, сколько раз пришлось заменить лампочку к моменту времени t.

▲ **Пример 1.3.** Модель страхования Спарре–Андерсена.

Пусть $(X_t, t \geqslant 0)$ — процесс восстановления, построенный по $\{\xi_n, n \in \mathbb{N}\}$. Пусть $\{\eta_m, m \in \mathbb{N}\}$ — независимые одинаково распределенные случайные величины, независимые с $\{\xi_n, n \in \mathbb{N}\}$, а $y_0, c > 0$ — некоторые положительные константы. Тогда процесс $Y = (Y_t, t \geqslant 0)$, где

$$Y_t = y_0 + c \cdot t - \sum_{k=1}^{X_t} \eta_k, \tag{1}$$

называется *моделью страхования Спарре–Андерсена.*

Поясним смысл параметров модели. Пусть имеется некоторая страховая компания, которая получает доход в виде постоянных страховых взносов и теряет деньги, выплачивая страховые премии случайного размера в случайные моменты времени. Тогда можно интерпретировать параметры следующим образом:

- y_0 — начальный капитал компании;
- c — скорость поступления страховых взносов;
- ξ_n — время между $(n-1)$-й и n-й страховыми выплатами;
- $S_n = \xi_1 + \cdots + \xi_n$ — время n-й выплаты;
- η_n — размер n-й выплаты;
- X_t — число выплат к моменту времени $t > 0$;
- $\sum_{k=1}^{X_t} \eta_k$ — общий размер выплат к этому моменту времени;
- Y_t — текущий капитал компании в момент времени t.

Для процесса $Z_t = \sum_{k=1}^{X_t} \eta_k$ в правой части (1) справедлив следующий усиленный закон больших чисел.

Теорема 1.1. *Пусть $(X_t, t \geqslant 0)$ — процесс восстановления, построенный по последовательности случайных величин $\{\xi_n, n \in \mathbb{N}\}$. Пусть, кроме того, $\{\eta_m, m \in \mathbb{N}\}$ — независимые одинаково распределенные случайные величины, независимые с $\{\xi_n, n \in \mathbb{N}\}$, а величины $\mu = \mathsf{E}\xi_i$, $\lambda = \mathsf{E}\eta_j$ — положительны и конечны. Положим $Z_t = \sum_{k=1}^{X_t} \eta_k$. Тогда*

$$\frac{Z_t}{t} \xrightarrow{\text{п. н.}} \frac{\lambda}{\mu}.$$

Кроме того, выполнена *элементарная теорема восстановления.*

Теорема 1.2. *В условиях теоремы* 1.1,

$$\frac{\mathsf{E}Z_t}{t} \longrightarrow \frac{\lambda}{\mu}.$$

▲ Задачи для самостоятельного решения

1. Пусть $X = X(\omega, t)$ — функция, определенная на $\Omega \times T$ и принимающая значения в $\mathbb{R}$. Пусть, кроме того, E — множество всех функций $f : T \to \mathbb{R}$. Рассмотрим на этом множестве σ-алгебру $\mathcal{E} = \sigma(\{f : f(t) \in B\}, B \in \mathcal{B}(\mathbb{R}), t \in T)$, которая называется *цилиндрической σ-алгеброй.* Докажите, что $(X_t = X(\cdot, t), t \in T)$ — случайный процесс тогда и только тогда, когда X — $(\mathcal{F}|\mathcal{E})$-измеримая функция.

2. Процессы $(X_t, t \in T)$ и $(Y_t, t \in \tilde{T})$ называются *независимыми*, если независимы порожденные ими σ-алгебры. Докажите, что $(X_t, t \in T)$ и $(Y_t, t \in \tilde{T})$ — независимы тогда и только тогда, когда для любых натуральных n, m и для любых $t_1, \dots, t_n \in T$, $\tilde{t}_1, \dots, \tilde{t}_m \in \tilde{T}$ векторы $(X_{t_1}, \dots, X_{t_n})$ и $(Y_{\tilde{t}_1}, \dots, Y_{\tilde{t}_m})$ являются независимыми.

3. Нарисуйте траекторию процесса Y_t в модели страхования Спарре–Андерсена.

4. Пусть $(X_t, t \geq 0)$ — процесс восстановления. Докажите, что $\mathsf{P}(\lim_{t \to +\infty} X_t = +\infty) = 1$.

5. Пусть $(X_t, t \geq 0)$ — процесс восстановления, построенный по независимым, но не обязательно одинаково распределенным случайным величинам $(\xi_n, n \in \mathbb{N})$.

   а) Приведите пример подобного процесса, уходящего на бесконечность за конечное время с положительной вероятностью.

   б) Приведите пример подобного процесса, не уходящего на бесконечность даже за бесконечное время.

6. Пусть $(X_t, t \geq 0)$ — процесс восстановления, построенный по случайным величинам $\{\xi_n, n \in \mathbb{N}\}$. Обозначим $\mu = \mathsf{E}\xi_i$, $\sigma^2 = \mathsf{D}\xi_i$, причем известно, что $\mu, \sigma^2 \in (0, +\infty)$. Докажите, что имеет место сходимость по распределению при $t \to +\infty$:

$$\frac{X_t - \frac{t}{\mu}}{\sigma\sqrt{t/\mu^3}} \xrightarrow{d} \mathcal{N}(0,1).$$

7. Пусть $(\xi_n, n \in \mathbb{N})$ — случайный процесс, элементы которого $\xi_1, \xi_2, \ldots$ являются независимыми случайными величинами. Обозначим

$$\mathcal{X} = \bigcap_{n=1}^{\infty} \mathcal{F}_n^{\infty}, \quad \text{где } \mathcal{F}_n^{\infty} = \sigma(\xi_{n+1}, \xi_{n+2}, \ldots),$$

остаточную σ-алгебру, т.е. пересечение всех σ-алгебр, порожденных наборами $\{\xi_k,\, k \geq n+1\}$. Докажите закон нуля или единицы: для любого события $A \in \mathcal{X}$ выполнено $\mathsf{P}(A) = 0$ или $\mathsf{P}(A) = 1$.

Указание: при доказательстве воспользуйтесь следующим утверждением. Если $\mathcal{A} \subset \mathcal{F}$ — некоторая алгебра подмножеств Ω, то для любого $\varepsilon > 0$ и любого $B \in \sigma(\mathcal{A})$ найдется такое $A \in \mathcal{A}$, что $\mathsf{P}(A \triangle B) \leqslant \varepsilon$.

Замечание: этот закон нуля или единицы применяется для доказательства различных утверждений о случайном блуждании. В частности, с помощью этого закона можно доказать, что с вероятностью единица простейшее симметричное случайное блуждание бесконечно много раз возвращается в ноль (см., например, [8]).

2. Ветвящиеся процессы Гальтона–Ватсона

В настоящей главе, а также в третьей и четвертых главах мы поговорим о важнейших примерах случайных процессов. В этой главе речь пойдет о ветвящихся процессах.

Модель ветвящихся процессов была впервые предложена Ф. Гальтоном и Г. Ватсоном в 1873 году в связи с анализом вырождения аристократических фамилий Великобритании.

Физическая модель: в дискретные моменты времени частицы распадаются на случайное количество таких же частиц. Число потомков каждой частицы имеет одно и то же распределение.

Формальная математическая модель выглядит следующим образом.

Математическая модель: пусть ξ — случайная величина со значениями в $\mathbb{Z}_+$, а $\{\xi_k^{(n)}, k, n \in \mathbb{N}\}$ — набор независимых случайных величин с тем же распределением, что и ξ. Положим

$$X_0 = 1,\ X_n = \sum_{k=1}^{X_{n-1}} \xi_k^{(n)},\ n \geqslant 1.$$

▲ **Определение 2.1.** Процесс $(X_n, n \in \mathbb{Z}_+)$ называется *ветвящимся процессом Гальтона–Ватсона* с законом размножения частиц ξ.

Интерпретация величин модели совершенно ясна: X_n — это число частиц в n-м поколении, а $\xi_k^{(n)}$ — число потомков k-й частицы из $(n-1)$-го поколения.

Нас будут интересовать два вопроса относительно ветвящихся процессов.

- Какова вероятность вырождения ветвящегося процесса $\mathsf{P}(\exists n:\ X_n = 0)$?
- Каково распределение общего числа частиц в процессе?

Ответы на эти вопросы можно найти с помощью метода производящих функций.

▲ **Определение 2.2.** Пусть ξ — случайная величина. Тогда ее *производящей функцией* называется функция

$$\varphi_\xi(z) = \mathsf{E} z^\xi,\ z \in \mathbb{R},$$

определенная для тех z, для которых конечно данное математическое ожидание.

Выделим основные свойства производящих функций.

1. $\varphi_\xi(1) = 1$;

2. $\varphi'_\xi(1) = \mathsf{E}\xi$;

3. Если ξ и η независимы, то $\varphi_{\xi+\eta}(z) = \varphi_\xi(z) \cdot \varphi_\eta(z)$.

Если ξ принимает значения в $\mathbb{Z}_+$, то имеются дополнительные свойства:

4. $\varphi_\xi(z) = \sum_{k=0}^{\infty} z^k \mathsf{P}(\xi = k)$ — степенной ряд, сходящийся абсолютно и равномерно в области $\{|z| \leqslant 1\}$;

5. $\varphi_\xi(z)$ непрерывно дифференцируема бесконечное число раз в области $\{|z| < 1\}$;

6. $\varphi_\xi(0) = \mathsf{P}(\xi = 0)$;

7. $\mathsf{P}(\xi = k) = \frac{1}{k!}\left(\frac{d^k}{dz^k}\varphi_\xi(z)\right)\Big|_{z=0}$.

▲ **Задача 2.1.** Случайная величина ξ имеет пуассоновское распределение с параметром $\lambda > 0$. Найдите ее производящую функцию.

Решение

Вычисляем по определению:

$$\varphi_\xi(z) = \mathsf{E}z^\xi = \sum_{k=0}^{\infty} z^k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(z\lambda)^k}{k!} = e^{\lambda(z-1)}.$$

Задача решена.

Пусть теперь $X = (X_n, n \in \mathbb{N})$ — это ветвящийся процесс Гальтона–Ватсона с законом размножения частиц ξ.

Лемма 2.1. *Между производящими функциями элементов процесса X имеют место следующие рекуррентные соотношения:*

1. $\varphi_{X_{n+1}}(z) = \varphi_{X_n}(\varphi_\xi(z))$;
2. $\varphi_{X_n}(z) = \underbrace{\varphi_\xi(\varphi_\xi(\ldots\varphi_\xi(}_{n \text{ раз}}z)\ldots))$;
3. $\varphi_{X_{n+1}}(z) = \varphi_\xi(\varphi_{X_n}(z))$.

Данная лемма позволяет найти уравнение для вероятности вырождения. Обозначим через q вероятность вырождения, т.е.

$$q = \mathsf{P}(\text{процесс выродился}) = \mathsf{P}(\exists n : X_n = 0).$$

Тогда выполнено следующее утверждение.

Лемма 2.2. *Вероятность вырождения q ветвящегося процесса с законом размножения частиц ξ является решением уравнения*

$$z = \varphi_\xi(z). \tag{2}$$

Уравнение (2) — это первая замечательная формула в теории ветвящихся процессов. Отметим, что

$$z = 1$$

всегда является решением (2). Однако при наличии других корней нам надо уметь отбирать среди них вероятность вырождения. Полную классификацию ситуаций дает теорема о вероятности вырождения.

Теорема 2.1 (о вероятности вырождения). *Пусть случайная величина ξ не равна единице тождественно, т.е. $\mathsf{P}(\xi = 1) \neq 1$. Обозначим $\mu = \mathsf{E}\xi$ (может быть, $\mu = +\infty$). Тогда*

1. *Если $\mu \leqslant 1$, то уравнение (2) имеет на $[0, 1]$ только одно решение: $z = 1$. В этом случае $q = 1$.*

2. *Если $\mu > 1$, то уравнение (2) имеет единственное решение $z_0 \in [0, 1)$. В этом случае $q = z_0$.*

Вывод: вероятность вырождения — это наименьший корень уравнения $z = \varphi_\xi(z)$ из отрезка $[0, 1]$.

Отметим, что теорема имеет весьма жизненную интерпретацию: если среднее число потомков в процессе не больше единицы, то процесс обречен на вымирание; в противном случае есть ненулевая вероятность бесконечного существования популяции.

▲ **Задача 2.2.** Найдите вероятность вырождения ветвящегося случайного процесса с геометрическим $\mathrm{Geom}(p)$ законом размножения частиц (т.е. $\mathsf{P}(\xi = k) = p(1-p)^k$, $k \in \mathbb{Z}_+$).

Решение

Найдем сначала производящую функцию числа частиц:

$$\varphi_\xi(z) = \mathsf{E}z^\xi = \sum_{k=0}^{\infty} z^k p(1-p)^k = \frac{p}{1 - z(1-p)}.$$

Теперь решаем уравнение $z = \varphi_\xi(z)$. Квадратное уравнение имеет два корня: $z_1 = 1$, $z_2 = p/(1-p)$. Согласно следствию из теоремы мы должны выбрать наименьший корень на отрезке $[0, 1]$. Тем самым
$q = 1$, если $p \geqslant 1/2$;
$q = p/(1-p)$, если $p < 1/2$.

Задача решена.

Общее число частиц. Обсудим теперь вопрос об общем числе частиц в ветвящемся процессе. Обозначим общее число частиц в процессе до момента времени n включительно $Y_n = 1 + X_1 + \ldots + X_n$. Тогда имеет место следующее рекуррентное соотношение между производящими функциями Y_n и Y_{n-1}, аналогичное соотношениям в лемме 2.1 для самого X_n:

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

Далее, для каждого $z \in [0, 1)$ существует предел

$$\rho(z) = \lim_{n \to +\infty} \varphi_{Y_n}(z),$$

который можно интерпретировать как производящую функцию общего числа частиц за все время существования процесса. Пусть Y — общее число частиц в ветвящемся процессе за все время (если процесс не выродился, то, конечно, $Y = +\infty$). Тогда

$$\rho(z) = \mathsf{E}(z^Y I(Y < +\infty)) = \sum_{k=0}^{+\infty} z^k \mathsf{P}(Y = k).$$

Найти $\rho(z)$ можно из второго замечательного уравнения в теории ветвящихся случайных процессов.

Лемма 2.3. *Функция $\rho(z)$ является решением уравнения*

$$\rho(z) = z\varphi_\xi(\rho(z)). \tag{3}$$

▲ **Задача 2.3.** Случайная величина ξ имеет геометрическое распределение с параметром $p \in (0, 1)$, а $(X_n, n \in \mathbb{Z}_+)$ — ветвящийся процесс с законом размножения частиц ξ. Вычислите производящую функцию общего числа частиц в процессе, а также найдите вероятность того, что всего в процессе было ровно k частиц.

Решение

Производящую функцию $\varphi_\xi(z)$ мы уже нашли в задаче 2.2:
$$\varphi_\xi(z) = \frac{p}{1 - z(1-p)}.$$

Теперь решим уравнение (3). Оно будет квадратным относительно $\rho = \rho(z)$:

$$\rho^2(1-p) - \rho + pz = 0.$$

Отсюда
$$\rho_{1,2} = \frac{1 \pm \sqrt{1 - 4p(1-p)z}}{2(1-p)}.$$

Уравнение (3) означает, что $\rho(0) = 0$, следовательно, надо взять корень со знаком минус:

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

Осталось найти вероятности $\mathsf{P}(Y = k)$, для этого надо разложить полученную функцию $\rho(z)$ в степенной ряд:

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

Стало быть, при $k \geqslant 1$

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

Задача решена.

▲ Задачи для самостоятельного решения

1. Найдите производящую функцию числа частиц в n-м поколении ветвящегося процесса, если производящая функция числа потомков одной частицы равна:

а) $pz+1-p$, б) $(1-p)/(1-pz)$, в) $1-p(1-z)^\alpha$, $\alpha\in(0,1)$.

2. Найдите вероятности вырождения для ветвящихся процессов с производящей функцией числа потомков одной частицы:

   а) $1-p(1-z)^\alpha$, $\alpha\in(0,1)$, б) $(1+z+z^2+z^3)/4$.

3. Найдите распределение момента вырождения N для ветвящихся процессов с производящей функцией числа потомков одной частицы:

   а) $pz+1-p$, б) $1-p(1-z)^\alpha$, $\alpha\in(0,1)$.

4. Ветвящийся процесс имеет следующий закон ξ распределения потомков одной частицы:

$$\mathsf{P}(\xi=0)=1/4,\ \ \mathsf{P}(\xi=2)=1/2,\ \ \mathsf{P}(\xi=6)=1/4.$$

   Верно ли, что вероятность вырождения будет принадлежать интервалу $(1/4,1/3)$?

5. Пусть ξ — число потомков частицы в ветвящемся процессе Гальтона–Ватсона $(X_n, n\in\mathbb{Z}_+)$. Обозначим $\mathsf{E}\xi=\mu$, $\mathsf{D}\xi=\sigma^2$. Найдите $\mathsf{E}X_n$ и $\mathsf{D}X_n$.

6. Пусть $(X_n, n\in\mathbb{Z}_+)$ — ветвящийся процесс с законом размножения частиц ξ. Обозначим через $Y_n=X_n+\ldots+X_0$ — общее число частиц в процессе за время n, а через $\varphi_{Y_n}(z)$ — его производящую функцию. Докажите, что

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

7. Пусть $(X_n, n\in\mathbb{Z}_+)$ — ветвящийся процесс с законом размножения частиц ξ. Обозначим через $Y_n=X_n+\ldots+X_0$ — общее число частиц в процессе за время n, а через $\varphi_{Y_n}(z)$ — его производящую функцию. Используя предыдущую задачу, докажите, что для всех $z\in[0,1)$ существует предел

$$\rho(z)=\lim_{n\to\infty}\varphi_{Y_n}(z),$$

   удовлетворяющий соотношению $\rho(z)=z\varphi_\xi(\rho(z))$. Чему по непрерывности функции ρ надо положить $\rho(1)$?

8. Пусть ветвящийся процесс Гальтона–Ватсона построен по случайной величине ξ, имеющей производящую функцию $\varphi(s) = 1 - \frac{1}{2}\sqrt{1-s}$. Найдите

   а) вероятность вырождения процесса;

   б) производящую функцию общего числа частиц процесса;

   в) вероятность того, что в процессе было всего 10 частиц.

3. Процессы с независимыми приращениями. Пуассоновский процесс

Следующие два примера случайных процессов, которые мы разберем (пуассоновский процесс в настоящей главе и винеровский процесс в следующей главе), являются процессами с независимыми приращениями.

▲ **Определение 3.1.** Процесс $(X_t, t \geqslant 0)$ называется *процессом с независимыми приращениями*, если для любого натурального n и любых $t_n > \ldots > t_1 \geqslant 0$ случайные величины $X_{t_n} - X_{t_{n-1}}, \ldots, X_{t_2} - X_{t_1}, X_{t_1}$ независимы в совокупности.

Отметим, что в случае дискретного времени процесс с независимыми приращениями является не чем иным, как случайным блужданием. Тем самым процессы с независимыми приращениями естественно рассматривать как обобщение случайных блужданий для непрерывного времени.

В теории случайных процессов важнейшую роль играют вопросы существования процессов. Дело в том, что совместные распределения случайных величин X_t не могут быть уж совсем произвольными, а должны быть в определенном роде согласованы. Для процессов с независимыми приращениями критерий существования звучит следующим образом.

Теорема 3.1. *Пусть для любых неотрицательных $s < t$ задано $Q_{s,t}$ — распределение вероятностей на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ с характеристической функцией $\varphi_{s,t}$, пусть Q_0 — также некоторое распределение вероятностей на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Тогда процесс $(X_t, t \geqslant 0)$ с независимыми приращениями и условиями*

$$X_t - X_s \stackrel{d}{=} Q_{s,t} \quad \forall\, 0 \leqslant s < t \text{ и } X_0 \stackrel{d}{=} Q_0$$

существует $\Leftrightarrow$ $\forall\, 0 \leqslant s < u < t$, $\forall\, \tau \in \mathbb{R}$

$$\varphi_{s,t}(\tau) = \varphi_{s,u}(\tau) \cdot \varphi_{u,t}(\tau). \qquad (4)$$

Необходимость условия (4) почти очевидна из свойств характеристических функций, но оказывается, что его достаточно для установления существования процесса.

Первым процессом с независимыми приращениями нашего курса является пуассоновский процесс.

▲ **Определение 3.2.** Случайный процесс $(N_t, t \geqslant 0)$ называется *пуассоновским процессом интенсивности* $\lambda > 0$, если выполнены три условия:

1. $N_0 = 0$ п.н.;

2. N_t имеет независимые приращения;

3. $N_t - N_s \sim \mathrm{Pois}(\lambda(t-s))$ для любых неотрицательных $t > s$.

▲ **Задача 3.1.** Докажите, что пуассоновский процесс существует.

Решение

Рассмотрим $\varphi_{s,t}$ — характеристическую функцию $\mathrm{Pois}(\lambda(t-s))$. Тогда

$$\varphi_{s,t}(\tau) = |\text{упражнение}| = e^{\lambda(t-s)(e^{i\tau}-1)}.$$

Отсюда видно, что для любых $s < u < t$ и $\tau \in \mathbb{R}$

$$\varphi_{s,u}(\tau) \cdot \varphi_{u,t}(\tau) = \varphi_{s,t}(\tau).$$

Значит, выполнено условие (4), и по теореме 3.1 о существовании процессов с независимыми приращениями пуассоновский процесс существует.

Задача решена.

Попробуем понять, как устроены траектории пуассоновского процесса. Из определения можно сделать следующие наблюдения:

- траектории целочисленные: $N_t \sim \mathrm{Pois}(\lambda t) \Rightarrow N_t \in \mathbb{Z}_+,\ t \geqslant 0$;

- траектории неубывающие: $N_t - N_s \geqslant 0,\ t > s$.

Подобные траектории, как мы помним, имели процессы восстановления. Оказывается, и пуассоновский процесс является таковым, о чем говорит теорема о его явной конструкции.

Теорема 3.2 (явная конструкция пуассоновского процесса). *Пусть* $\{\xi_n, n \in \mathbb{N}\}$ *— независимые одинаково распределенные экспоненциальные случайные величины,* $\xi_i \sim \mathrm{Exp}(\lambda), \lambda > 0$. *Положим* $S_0 = 0, S_n = \xi_1 + \ldots + \xi_n$. *Тогда процесс восстановления* $(X_t, t \geqslant 0)$,

$$X_t = \sup\{n : S_n \leqslant t\},$$

является пуассоновским процессом интенсивности λ.

Теорема 3.2 позволяет выявить ряд свойств скачков пуассоновского процесса.

- С вероятностью единица все скачки N_t имеют размер, равный единице.
- Пусть Y_n — момент n-го скачка N_t. Тогда $Y_n \sim \Gamma(\lambda, n)$ (т.е. плотность равна $p_n(x) = \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x} I\{x > 0\}$).
- $Y_n - Y_{n-1}, n \in \mathbb{N}$, — независимые экспоненциальные $\mathrm{Exp}(\lambda)$ случайные величины.

Заметим также, что все траектории процесса восстановления являются непрерывными справа. В задаче 8 самостоятельной работы к настоящей главе описаны условия, при которых это свойство всегда можно предполагать при необходимости.

Итак, при решении задач мы можем рассматривать пуассоновский процесс либо как процесс с независимыми приращениями, либо как процесс восстановления для экспоненциальных случайных величин в зависимости от того, как будет удобнее. Приведем пример применения основного определения пуассоновского процесса (определения 3.2) для нахождения ковариационной функции этого процесса.

▲ **Определение 3.3.** Пусть $(X_t, t \in T)$ — случайный процесс. Тогда

- функция $a(t) = \mathsf{E}X_t$, $t \in T$, называется *функцией среднего* процесса X_t;

- функция $R(s,t) = \mathrm{cov}(X_s, X_t)$, $s,t \in T$, называется *ковариационной функцией* процесса X_t.

▲ **Задача 3.2.** Найдите функцию среднего и ковариационную функцию пуассоновского процесса $(N_t, t \geqslant 0)$.

Решение

Заметим, что из определения 3.2 пуассоновского процесса вытекает, что $N_t \sim \mathrm{Pois}(\lambda t)$. Значит, для любого $t \geqslant 0$

$$\mathsf{E}N_t = \lambda t.$$

Далее, при $t \geqslant s$

$$\begin{aligned}\mathrm{cov}(N_t, N_s) &= \mathrm{cov}(N_t - N_s, N_s) + \mathrm{cov}(N_s, N_s) = \\ &= |N_t - N_s \text{ и } N_s \text{ независимы}| = \\ &= 0 + \mathrm{cov}(N_s, N_s) = \mathsf{D}N_s = \lambda s.\end{aligned}$$

В общем же случае мы получаем, что

$$\mathrm{cov}(N_t, N_s) = \lambda \min(t, s).$$

Задача решена.

▲ Задачи для самостоятельного решения

1. Пусть $(X_t, t \geq 0)$ — процесс восстановления, построенный по $\{\xi_n, n \in \mathbb{N}\}$. Верно ли, что процесс X_t всегда имеет независимые приращения?

2. Пусть $N^1 = (N_t^1, t \geq 0), \ldots, N^k = (N_t^k, t \geq 0)$ — независимые пуассоновские процессы (т.е. независимыми являются порожденные ими σ-алгебры), причем N^i имеет интенсивность λ_i. Докажите, что процесс $N_t = \sum_{i=1}^k N_t^i$ также является пуассоновским, и найдите его интенсивность.

3. Задан процесс $\{Y_t = \sum\limits_{j=1}^{N_t} \xi_j,\ t \geqslant 0\}$, где $(\xi_n, n \in \mathbb{N})$ — независимые одинаково распределенные случайные величины, не зависящие от пуассоновского процесса $N = (N_t, t \geqslant 0)$ интенсивности λ. Докажите, что процесс Y_t имеет независимые приращения.

4. Пусть $(\xi_n, n \in \mathbb{N})$ — независимые экспоненциальные случайные величины с параметром λ, $S_n = \xi_1 + \ldots + \xi_n$, а $N = (N_t, t \geqslant 0)$ — процесс восстановления, построенный по ним (пуассоновский процесс интенсивности λ). Для каждого $t > 0$ обозначим $V_t = S_{N_t+1} - t$ («перескок») и $U_t = t - S_{N_t}$ («недоскок»).

   а) Вычислите вероятность $\mathsf{P}(V_t > v, U_t > u)$.

   б) Докажите, что V_t и U_t — независимы и что $V_t \sim Exp(\lambda)$.

   в) Вычислите функцию распределения U_t и $\mathsf{E}U_t$.

5. Пусть $N = (N_t, t \geqslant 0)$ — пуассоновский процесс интенсивности λ. Найдите математическое ожидание числа таких его скачков на отрезке $[0, T]$, что

   а) в их правой a-окрестности нет других скачков (эта окрестность может выходить и за пределы отрезка),

   б) в их левой a-окрестности нет других скачков.

6. Пусть $(N_t, t \geqslant 0)$ — пуассоновский процесс интенсивности λ. Найдите предел п. н. N_t/t при $t \to +\infty$.

7. Докажите, что у пуассоновского процесса $(N_t, t \geq 0)$ интенсивности $\lambda > 0$ не существует непрерывной модификации, т.е. не существует такого процесса $(X_t, t \geq 0)$, что для каждого $t \geq 0$ с вероятностью единица $X_t = N_t$ и все траектории X_t непрерывны.

   Замечание: Процесс $(Y_t, t \in T)$ называется модификацией процесса $(X_t, t \in T)$, если для любого $t \in T$ выполнено $\mathsf{P}(X_t = Y_t) = 1$.

8. Пусть $X = (X_t, t \geq 0)$ — такой случайный процесс, для которого $\mathsf{P}(X_s \leq X_t) = 1$ при $0 \leq s \leq t < \infty$. Предположим, что X непрерывен по вероятности справа в каждой точке

s полупрямой $[0, \infty)$, т.е. X_t сходится по вероятности к X_s, когда $t \to s + 0$. Докажите, что у X существует модификация, имеющая п. н. неубывающие траектории, непрерывные справа.

Указание: сначала надо установить утверждение из анализа. Пусть $f(x)$ — такая функция на $[0, 1]$, что она не убывает по рациональным точкам. Докажите, что тогда функция $\tilde{f}(x)$, определенная как $\tilde{f}(x) = \lim\limits_{r \to x+0,\ r \in \mathbb{Q}} f(r)$, является непрерывной справа и неубывающей.

4. Гауссовские процессы. Винеровский процесс

В этой главе мы рассмотрим еще один важный пример случайных процессов — винеровский процесс. Этот процесс не только является процессом с независимыми приращениями, но также относится к классу гауссовских.

▲ **Определение 4.1.** Случайный процесс $(X_t, t \in T)$ называется *гауссовским*, если все его конечномерные распределения — гауссовские, т.е. для любого $n \in \mathbb{N}$ и любых $t_1, \dots, t_n \in T$ случайный вектор $(X_{t_1}, \dots, X_{t_n})$ является гауссовским.

Напомним, что распределение гауссовского вектора однозначно определяется его математическим ожиданием и матрицей ковариаций. Значит, конечномерные распределения гауссовского процесса однозначно определяются функцией среднего и ковариационной функцией. Какими же свойствами должны обладать эти функции, чтобы такой процесс существовал? Функция среднего может быть произвольной, а вот ковариационная функция обязана обладать свойством неотрицательной определенности.

▲ **Определение 4.2.** Функция $f(x, y), x, y \in T$, принимающая действительные значения, называется *неотрицательно определенной* на $T \times T$, если для любого $n \in \mathbb{N}$ и для любых $t_1, \dots, t_n \in T$, $z_1, \dots, z_n \in \mathbb{R}$ выполнено

$$\sum_{i,j=1}^{n} f(t_i, t_j) z_i z_j \geqslant 0.$$

Лемма 4.1. *Ковариационная функция $R(s, t)$ любого случайного процесса является неотрицательно определенной и симметричной.*

Оказывается, перечисленные в лемме 4.1 свойства являются не только необходимыми, но и достаточными для того, чтобы функция была ковариационной для некоторого случайного процесса. Следующая теорема говорит о том, что подобный процесс может быть выбран даже гауссовским.

Теорема 4.1 *Пусть $a : T \to \mathbb{R}$ — некоторая функция, определенная на некотором множестве T, и $R : T \times T \to \mathbb{R}$ — симметричная и неотрицательно определенная функция. Тогда существует вероятностное пространство $(\Omega, \mathcal{F}, \mathsf{P})$ и такой гауссовский процесс $(X_t, t \in T)$ на нем, что для любых $s, t \in T$*

$$a(t) = \mathsf{E}X_t \text{ и } R(s,t) = \operatorname{cov}(X_s, X_t).$$

▲ **Задача 4.1.** Докажите, что существует гауссовский процесс $(X_t, t \geqslant 0)$ с нулевой функцией среднего и ковариационной функцией $r(s,t) = \min(s,t)$.

Решение. Согласно теореме 4.1, достаточно проверить, что функция $r(s,t) = \min(s,t)$ является неотрицательно определенной. Для этого рассмотрим следующее интегральное представление:

$$\min(s,t) = \int_0^{+\infty} I_{[0,t]}(x) I_{[0,s]}(x)\,dx.$$

Тогда для любых $t_1, \ldots, t_n \in \mathbb{R}_+$, $z_1, \ldots, z_n \in \mathbb{R}$ выполнено

$$\sum_{i,j=1}^{n} \min(t_i, t_j) z_i z_j = \sum_{i,j=1}^{n} \left(\int\limits_0^{\infty} I_{[0,t_i]}(x) I_{[0,t_j]}(x)\,dx \right) z_i z_j =$$

$$= \int\limits_0^{\infty} \left(\sum_{i,j=1}^{n} (z_i I_{[0,t_i]}(x))(z_j I_{[0,t_j]}(x)) \right) dx =$$

$$= \int\limits_0^{\infty} \left(\sum_{i=1}^{n} z_i I_{[0,t_i]}(x) \right)^2 dx \geqslant 0.$$

Задача решена.

Отметим, что задачу можно решить и по-другому, явно предъявив случайный процесс с ковариационной функцией $\min(s,t)$. Подобным процессом служит пуассоновский процесс интенсивности 1. Вторым подходящим процессом (который, как мы увидим ниже, в отличие от пуассоновского процесса, гауссовский) является винеровский процесс, также называемый *процессом броуновского движения*.

▲ **Определение 4.3.** Случайный процесс $(W_t, t \geqslant 0)$ называется *винеровским* (или *процессом броуновского движения*), если выполнены следующие три условия:

1. $W_0 = 0$ п. н.;
2. W_t имеет независимые приращения;
3. $W_t - W_s \sim \mathcal{N}(0, t-s)$ для любых неотрицательных $t \geqslant s$.

▲ **Задача 4.2.** Докажите, что винеровский процесс существует.

Решение

Пусть $\varphi_{s,t}(\tau)$ — характеристическая функция распределения $\mathcal{N}(0, t-s), t \geqslant s$. Тогда

$$\varphi_{s,t}(\tau) = e^{-\frac{1}{2}\tau^2(t-s)} = e^{-\frac{1}{2}\tau^2(t-u+u-s)}.$$

Значит, для любых $s < u < t$ и $\tau \in \mathbb{R}$

$$\varphi_{s,t}(\tau) = \varphi_{s,u}(\tau) \cdot \varphi_{u,t}(\tau).$$

По критерию существования процессов с независимыми приращениями (теорема 3.1), винеровский процесс существует.

Задача решена.

Модель винеровского процесса впервые была предложена Л. Башелье в 1903 году для моделирования колебаний курсов акций на бирже. С того времени данный процесс стал одним из центральных в теории вероятностей, он активно применяется для моделирования хаотического движения.

Перейдем к обсуждению свойств винеровского процесса.

▲ **Задача 4.3.** Докажите, что винеровский процесс является гауссовским.

Решение

Пусть $t_1, \ldots, t_n \in \mathbb{R}_+$. Нам необходимо показать, что вектор $(W_{t_1}, \ldots, W_{t_n})$ является гауссовским. Будем считать, что $t_1 \leqslant \ldots \leqslant t_n$. Заметим, что, в силу определения винеровского процесса, случайные величины $W_{t_n} - W_{t_{n-1}}, \ldots, W_{t_2} - W_{t_1}, W_{t_1}$ независимы и имеют нормальное распределение. Стало быть, вектор

$$(W_{t_n} - W_{t_{n-1}}, \ldots, W_{t_2} - W_{t_1}, W_{t_1})$$

является гауссовским. Вектор же $(W_{t_1}, \ldots, W_{t_n})$ является его линейным преобразованием, значит, он тоже гауссовский, и процесс W_t является гауссовским.

Задача решена.

Так как винеровский процесс является гауссовским, то важно вычислить его ковариационную функцию и функцию среднего.

▲ **Задача 4.4.** Найдите функцию среднего и ковариационную функцию винеровского процесса $(W_t, t \geqslant 0)$.

Решение

Заметим, что из определения 4.3 винеровского процесса вытекает, что $W_t \sim \mathcal{N}(0, t)$. Значит, для любого $t \geqslant 0$

$$\mathsf{E}W_t = 0.$$

Далее, при $t \geqslant s$

$$\begin{aligned} \operatorname{cov}(W_t, W_s) &= \operatorname{cov}(W_t - W_s, W_s) + \operatorname{cov}(W_s, W_s) = \\ &= |W_t - W_s \text{ и } W_s \text{ независимы}| \\ &= 0 + \operatorname{cov}(W_s, W_s) = \mathsf{D}W_s = s. \end{aligned}$$

В общем же случае мы получаем, что

$$\operatorname{cov}(W_t, W_s) = \min(t, s).$$

Задача решена.

Гауссовость винеровского процесса позволяет сформулировать его второе эквивалентное определение.

Теорема 4.2. *Процесс $(W_t, t \geqslant 0)$ является винеровским $\Leftrightarrow$ выполнены три свойства:*

1. W_t *— гауссовский процесс;*
2. $\mathsf{E}W_t = 0$ *для любого* $t \geqslant 0$*;*
3. $\operatorname{cov}(W_s, W_t) = \min(s, t)$ *для любых* $s, t \geqslant 0$.

Отметим, что второе определение, использующее гауссовость, зачастую гораздо удобнее для проверки того, что некоторый процесс является винеровским, ведь проверить то, что процесс является гауссовским, часто бывает проще, чем проверить сильное свойство независимости приращений.

▲ **Задача 4.5.** Пусть $(W_t, t \geqslant 0)$ — винеровский процесс. Докажите, что процесс

$$B_t = t \cdot W_{\frac{1}{t}} I\{t > 0\}$$

также является винеровским.

Решение

Проверим второе определение из теоремы 4.2.
1. Для любых $t_1, \ldots, t_n \geqslant 0$ вектор $(B_{t_1}, \ldots, B_{t_n})$ является линейным преобразованием гауссовского вектора $\left(W_{\frac{1}{t_1}}, \ldots, W_{\frac{1}{t_n}}\right)$. Значит, $(B_{t_1}, \ldots, B_{t_n})$ — гауссовский вектор, а B_t — гауссовский процесс.

2. $\mathsf{E}B_t = 0$ для любого $t \geqslant 0$.
3. Для любых $t, s \geqslant 0$

$$\operatorname{cov}(B_t, B_s) = \operatorname{cov}(tW_{\frac{1}{t}}, sW_{\frac{1}{s}}) =$$

$$= ts \min\left(\frac{1}{t}, \frac{1}{s}\right) = \frac{ts}{\max(t, s)} = \min(t, s).$$

Значит, по теореме 4.2, процесс B_t является винеровским.

Задача решена.

В заключение темы обсудим свойства траекторий винеровского процесса. Сначала дадим определение модификации случайного процесса, с понятием которой мы уже сталкивались в задаче 7 самостоятельной работы к предыдущей главе.

▲ **Определение 4.4.** Процесс $(Y_t, t \in T)$ называется *модификацией* процесса $(X_t, t \in T)$, если

$$\forall t \in T \quad \mathsf{P}(X_t = Y_t) = 1.$$

Важнейшими свойствами траекторий $(W_t, t \geqslant 0)$ являются следующие.

1. Траектории винеровского процесса непрерывны с вероятностью единица, точнее, у него существует модификация, все траектории которой непрерывны (называется *непрерывной модификацией*).

2. С вероятностью единица траектория $(W_t, t \geqslant 0)$ не дифференцируема ни в одной точке $\mathbb{R}_+$.

3. Выполнен закон повторного логарифма, т.е.

$$\mathsf{P}\left(\limsup_{t\to+\infty} \frac{W_t}{\sqrt{2t\ln\ln t}} = 1\right) = 1,$$

где $\limsup\limits_{t\to+\infty} f(t) = \lim\limits_{t\to+\infty}\left(\sup\limits_{s\geqslant t} f(s)\right)$.

Перечисленные свойства показывают, что винеровский процесс действительно моделирует хаотическое движение (траектория непрерывна, но совсем не гладкая), а "размах" отклонения от нуля во времени имеет порядок $\sqrt{2t\ln\ln t}$.

▲ Задачи для самостоятельного решения

1. Пусть $(W_t, t \geqslant 0)$ — винеровский процесс. Докажите, что следующие процессы тоже винеровские:

   а) $X_t = -W_t$;

   б) $X_t = \sqrt{c}\, W_{t/c}$, $c > 0$;

   в) $X_t = W_{t+a} - W_a$, $a > 0$;

   г) $X_t = W_t\, I\{t < T\} + (2W_T - W_t) I\{t \geqslant T\}$.

2. Пусть $(Y_t, t \in [0,1])$ — гауссовский процесс с нулевой функцией среднего и ковариационной функцией $r(s,t) = \min(s,t) - st$. Докажите, что такой процесс существует и что процесс $(X_t = (t+1)Y_{t/(t+1)},\ t \geqslant 0)$ является винеровским.

3. Пусть $W_t^1, \ldots, W_t^d$ — независимые винеровские процессы. Докажите, что с вероятностью единица процесс $W_t = (W_t^1, \ldots, W_t^d)$ (многомерный винеровский процесс) выйдет из шара произвольного фиксированного радиуса r с центром в нуле пространства $\mathbb{R}^d$.

4. Докажите существование гауссовского процесса $X = (X_t, t \in \mathbb{R}_+^d)$ с нулевой функцией среднего и ковариационной функцией
$$R(s,t) = \prod_{k=1}^{d} \min(s_k, t_k),$$
   где $s = (s_1, \ldots, s_d) \in \mathbb{R}_+^d$, $t = (t_1, \ldots, t_d) \in \mathbb{R}_+^d$.

5. Пусть $(W_t, t \geqslant 0)$ — винеровский процесс. Докажите, что с вероятностью единица его траектория имеет неограничен-

ную вариацию на произвольном отрезке $[a, b] \subset \mathbb{R}_+$, т.е. что

$$\sup_T \sum_{i=1}^{n} |W_{t_{i+1}} - W_{t_i}| = +\infty \text{ п. н.},$$

где $T = \{a = t_0 < \ldots < t_n = b\}$ — разбиение отрезка $[a, b]$.

6. Пусть $(W_t, t \geq 0)$ — винеровский процесс. Для каждого $n \in \mathbb{N}$ рассмотрим разбиения отрезка $[a; b]$ точками $a = t_{n;0} < t_{n;1} < \ldots < t_{n;N(n)} = b$, где $N(n) \in \mathbb{N}$ и $\Delta T(n) := \max_{k=0;\ldots;N(n)}(t_{n;k+1} - t_{n;k}) \to 0$ при $n \to +\infty$. Положим $U_n := \sum_{k=0}^{N(n)-1}(W_{t_{n;k+1}} - W_{t_{n;k}})^2$. Найдите предел в среднем квадратическом величин U_n, когда $n \to +\infty$.

7. Последовательность положительных чисел $\{t_n, n \in \mathbb{N}\}$ такова, что $\sum_{n=1}^{\infty} t_n^{-1/2} < \infty$. Докажите, что тогда $|W_{t_n}| \to +\infty$ п.н. с ростом n. Здесь $(W_t, t \geq 0)$ — винеровский процесс.

8. Пусть $(X_t, t \geq 0)$ — гауссовский процесс со стационарными независимыми приращениями (стационарные приращения означают, что распределение случайной величины $X_t - X_s$ зависит только от $t - s$). Докажите, что найдутся такие константы $a, b \in \mathbb{R}$, что процесс $Y_t = (X_t - at)/b$ является винеровским.

9. Пусть $\{\xi_n, n \in \mathbb{Z}_+\}$ — независимые $\mathcal{N}(0, 1)$ случайные величины. Докажите, что процесс

$$X_t = \frac{\xi_0 t}{\sqrt{\pi}} + \sqrt{\frac{2}{\pi}} \sum_{k=1}^{\infty} \frac{\sin(kt)}{k} \, \xi_k$$

является винеровским на отрезке $[0, \pi]$ (ряд понимается как предел частичных сумм в L^2).

Указание: надо разложить функцию $I_{[0,t]}(x)$ в ряд Фурье по ортонормированной системе на $[0, \pi]$, составленной из $\sqrt{\frac{2}{\pi}} \cos(kx)$.

5. Марковские моменты

С точки зрения приложений полезным является анализ поведения случайного процесса в случайные моменты времени. В данном разделе мы рассмотрим важнейший подобный класс случайных моментов, называемых *марковскими*. Для этого нам понадобится ввести ряд определений.

▲ **Определение 5.1.** Пусть $(\Omega, \mathcal{F}, \mathsf{P})$ — вероятностное пространство. Множество σ-алгебр $\mathbb{F} = (\mathcal{F}_t, t \in T)$, где $T \subset \mathbb{R}$, называется *фильтрацией* (или *потоком σ-алгебр*) на $(\Omega, \mathcal{F}, \mathsf{P})$, если

$$\forall s < t,\ s, t \in T: \quad \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}.$$

Фильтрацию можно понимать как поток информации, который увеличивается с течением времени.

▲ **Определение 5.2.** Случайный процесс $(X_t, t \in T)$ называется *согласованным* с фильтрацией $\mathbb{F} = (\mathcal{F}_t, t \in T)$, если для любого $t \in T$ X_t является $\mathcal{F}_t$-измеримым, то есть

$$\sigma(X_t) = \mathcal{F}_{X_t} \subset \mathcal{F}_t.$$

Откуда может появиться фильтрация на вероятностном пространстве? Несложно понять, что с любым случайным процессом можно связать некоторую фильтрацию, которая называется его *естественной фильтрацией*.

▲ **Определение 5.3.** *Естественной фильтрацией* процесса $(X_t, t \in T), T \subset \mathbb{R}$, называется $\mathbb{F}^X = (\mathcal{F}_t^X, t \in T)$, где

$$\mathcal{F}_t^X = \sigma(X_s,\ s \leqslant t, s \in T).$$

Очевидно, что любой процесс согласован со своей естественной фильтрацией. Теперь мы готовы ввести понятие марковского момента относительно фильтрации.

▲ **Определение 5.4.** Пусть $T \subset \mathbb{R}$. Тогда отображение $\tau : \Omega \to T \cup \{+\infty\}$ называется *марковским моментом* относительно фильтрации $\mathbb{F} = (\mathcal{F}_t, t \in T)$ на $(\Omega, \mathcal{F}, \mathsf{P})$, если для любого $t \in T$ выполнено

$$\{\tau \leqslant t\} \in \mathcal{F}_t.$$

Если, к тому же, выполнено $\mathsf{P}(\tau < +\infty) = 1$, то τ называется *моментом остановки* относительно $\mathbb{F}$.

Неформальный смысл определения марковского момента состоит в следующем. Пусть τ — случайный момент наступления некоторого события в процессе $(X_t, t \in T)$. Тогда τ будет являться марковским моментом относительно естественной фильтрации процесса X_t, если для $\forall t_0 \in T$ можно однозначно сказать, наступило уже τ к моменту времени t_0 или еще нет (то есть $\tau \leqslant t_0$ или $\tau > t_0$?), зная лишь значения процесса X_t до момента времени t_0 включительно. Приведем пример, иллюстрирующий подобный принцип.

▲ **Задача 5.1.** $(X_n, n \in \mathbb{N})$ — действительный случайный процесс. Докажите, что тогда для любого $B \in \mathcal{B}(\mathbb{R})$

$$\tau_B = \min\{n : X_n \in B\}$$

является марковским моментом относительно $\mathbb{F}^X$.

Решение

Легко заметить, что

$$\{\tau_B \leqslant n\} = \bigcup_{k=1}^{n} \underbrace{\{X_k \in B\}}_{\in \mathcal{F}_k^X \subset \mathcal{F}_n^X} \in \mathcal{F}_n^X.$$

Задача решена.

Учитывая задачу 5.1, легко привести пример немарковского момента. Например, рассмотрим

$$\tau_B = \min\{n : X_{n+1} \in B\}.$$

Данный момент для произвольного процесса уже не будет марковским, ведь событие $\{X_{n+1} \in B\}$ может не принадлежать элементу фильтрации $\mathcal{F}_n^X$.

С марковским моментом можно связать сигма-алгебру событий, “произошедших до момента времени τ”. Дадим формальное определение.

▲ **Определение 5.5.** Пусть τ — марковский момент относительно $\mathbb{F} = (\mathcal{F}_t, t \in T)$. Тогда сигма-алгеброй $\mathcal{F}_\tau$ называется

$$\mathcal{F}_\tau := \{A \in \mathcal{F} : \forall t \in T \quad \{\tau \leqslant t\} \cap A \in \mathcal{F}_t\}.$$

Перечислим основные свойства $\mathcal{F}_\tau$.

1. $\mathcal{F}_\tau$ — действительно σ-алгебра, причем $\mathcal{F}_\tau \subset \mathcal{F}$.
2. Случайная величина τ (вообще говоря, *расширенная*, так как может принимать значение $+\infty$) измерима относительно $\mathcal{F}_\tau$.
3. Если $\tau = t = \mathrm{const}$, то $\mathcal{F}_\tau = \mathcal{F}_t$.

Перейдем к приложениям марковских моментов к винеровскому процессу. Сначала приведем формулировку несложного утверждения (которое частично уже вошло в самостоятельную работу к прошлой главе), называемого марковским свойством, которое показывает, что если остановить винеровский процесс в любой фиксированный момент времени и запустить время и положение заново, то новый процесс также будет винеровским.

Лемма 5.1 (марковское свойство винеровского процесса). *Пусть $(W_t, t \geqslant 0)$ — винеровский процесс. Тогда для любого $a > 0$ процесс*

$$X_t = W_{t+a} - W_a,\ t \geqslant 0,$$

является винеровским и не зависит от сигма-алгебры $\mathcal{F}_a^W$ (т.е. $\sigma(X_t, t \geqslant 0)$ не зависит от $\mathcal{F}_a^W$).

Естественный вопрос: можно ли заменить a на случайную величину? Оказывается, утверждение останется верным, если a заменить на момент остановки относительно естественной фильтрации винеровского процесса. Более того, подобный факт верен для целого класса процессов, называемых *процессами Леви.*

▲ **Определение 5.6.** Процесс $(X_t, t \geqslant 0)$ называется *процессом Леви*, если он имеет независимые стационарные приращения и $X_0 = 0$ (п. н.). *Стационарность приращений* означает, что для любых $t, s, h \in \mathbb{R}_+$ выполнено $X_t - X_s \overset{d}{=} X_{t+h} - X_{s+h}$.

Тогда имеет место следующее замечательное свойство, называемое *строго марковским.*

Теорема 5.1 (строго марковское свойство). *Пусть* $(X_t, t \geqslant 0)$ *— процесс Леви с непрерывными справа траекториями, а* τ *— момент остановки относительно* $\mathbb{F}^X$, *естественной фильтрации процесса* X. *Тогда процесс* $Y = (Y_t, t \geqslant 0)$, *где*

$$Y_t = X_{t+\tau} - X_\tau,\ t \geqslant 0,$$

имеет те же конечномерные распределения, что и X, *а также не зависит от сигма-алгебры* $\mathcal{F}_\tau$.

Как мы помним, винеровский процесс имеет непрерывную модификацию, а пуассоновский процесс — непрерывную справа. Тем самым теорема применима к ним обоим.

Вторым замечательным утверждением о моментах остановки является принцип отражения для винеровского процесса.

Теорема 5.2 (принцип отражения). *Пусть* $(W_t, t \geqslant 0)$ *— винеровский процесс,* τ *— момент остановки относительно* $\mathbb{F}^W$. *Тогда процесс* $(Z_t, t \geqslant 0)$, *где*

$$Z_t = \begin{cases} W_t, & t \leqslant \tau, \\ 2W_\tau - W_t, & t > \tau, \end{cases}$$

является винеровским.

Какие марковские моменты было бы интересно рассмотреть в связи с винеровским процессом? Рассмотрим, например, $\tau_x = \inf\{t : W_t = x\}$ — первый момент времени достижения уровня $x \in \mathbb{R}$ винеровским процессом.

▲ **Задача 5.2.** Докажите, что τ_x — момент остановки относительно естественной фильтрации винеровского процесса.

Решение

Будем считать, что траектории W_t непрерывны и $x \geqslant 0$ (в случае $x < 0$ решение аналогично). Тогда

$$\begin{aligned}
\{\tau_x > t\} &= |\text{используем непрерывность траекторий}| = \\
&= \{\forall s \leqslant t : W_s < x\} = \\
&= \bigcup_{k=1}^{\infty} \left\{\forall s \leqslant t : W_s \leqslant x - \frac{1}{k}\right\} = \\
&= \bigcup_{k=1}^{\infty} \bigcap_{\substack{s \in \mathbb{Q} \\ s \leqslant t}} \underbrace{\left\{W_s \leqslant x - \frac{1}{k}\right\}}_{\in \mathcal{F}_s^W \subset \mathcal{F}_t^W} \in \mathcal{F}_t^W.
\end{aligned}$$

Значит, раз $\{\tau_x > t\} = \Omega \setminus \{\tau_x \leqslant t\} \in \mathcal{F}_t^W$, то τ_x — марковский момент относительно $\mathbb{F}^W$. Из закона повторного логарифма вытекает, что с вероятностью единица траектория W_t растет вверх неограниченно. Значит, найдется такое t, что $W_t = x$, и, стало быть, $\mathsf{P}(\tau_x < +\infty) = 1$.

Задача решена.

С помощью случайной величины τ_x можно найти распределение максимального значения винеровского процесса на отрезке $[0, t]$. Обозначим $M_t = \max\limits_{s \leqslant t} W_s$. Заметим простую связь M_t с τ_x: для любого $x > 0$

$$\{M_t \geqslant x\} = \{\tau_x \leqslant t\}.$$

В качестве следствия мы получаем, что процесс M_t является согласованным с естественной фильтрацией винеровского процесса.

С помощью принципа отражения можно отыскать совместное распределение M_t и W_t.

Теорема 5.3. *Для любых $x, y, t \geqslant 0$ выполнено:*

$$\mathsf{P}(W_t < y - x, M_t \geqslant y) = \mathsf{P}(W_t > y + x).$$

Из данной теоремы можно найти распределение случайной величины $M_t = \max\limits_{s \in [0,t]} W_s$.

Теорема 5.4 (Башелье). *Выполнено равенство по распределению*

$$M_t \overset{d}{=} |W_t|.$$

▲ **Задача 5.3.** Докажите теорему Башелье.

Решение

$$\begin{aligned}\mathsf{P}(M_t \geqslant y) &= \mathsf{P}(M_t \geqslant y,\ W_t < y) + \mathsf{P}(M_t \geqslant y,\ W_t \geqslant y) = \\ &= |\text{теорема 5.3 с } x = 0 \text{ для первой вероятности,} \\ &\{W_t \geqslant y\} \subset \{M_t \geqslant y\} \text{ для второй}| = \\ &= \mathsf{P}(W_t > y) + \mathsf{P}(W_t \geqslant y) = 2\mathsf{P}(W_t \geqslant y) = \\ &= \mathsf{P}(|W_t| \geqslant y).\end{aligned}$$

Задача решена.

▲ Задачи для самостоятельного решения

1. Пусть задана фильтрация $\mathbb{F} = (\mathcal{F}_n, n \in \mathbb{N})$, а $\tau_1, \tau_2, \ldots$ — марковские моменты относительно $\mathbb{F}$. Докажите, что случайные величины
$$\sum_{k=1}^{m} \tau_k, \quad \prod_{k=1}^{m} \tau_k, \quad \sup_k \tau_k, \quad \inf_k \tau_k$$
тоже являются марковскими моментами относительно $\mathbb{F}$.

2. Пусть задана фильтрация $\mathbb{F} = (\mathcal{F}_t, t \geq 0)$, а $\tau_1, \tau_2, \ldots$ — марковские моменты относительно $\mathbb{F}$. Докажите, что
$$\sum_{k=1}^{m} \tau_k, \quad \max_{k=1,\ldots,m} \tau_k, \quad \min_{k=1,\ldots,m} \tau_k$$

тоже являются марковскими моментами относительно $\mathbb{F}$.

3. а) Пусть τ — марковский момент относительно фильтрации $\mathbb{F} = (\mathcal{F}_t, t \in T)$, где $T \subset \mathbb{R}_+$. Докажите, что τ является случайной величиной, измеримой относительно $\mathcal{F}_\tau$.

   б) Пусть τ — марковский момент относительно фильтрации $\mathbb{F} = (\mathcal{F}_n, n \in \mathbb{N})$, а случайный процесс $(X_n, n \in \mathbb{N})$ согласован с $\mathbb{F}$. Докажите, что X_τ является $\mathcal{F}_\tau$-измеримым (считаем, что $X_\tau = +\infty$, если $\tau = +\infty$).

4. Пусть задана фильтрация $\mathbb{F} = (\mathcal{F}_t, t \in T)$, $T \subset \mathbb{R}$, на $(\Omega, \mathcal{F}, \mathsf{P})$. Докажите, что для марковских моментов σ, τ относительно $\mathbb{F}$ и для любого $A \in \mathcal{F}_\tau$ выполнено
$$A \cap \{\tau \le \sigma\} \in \mathcal{F}_\tau \cap \mathcal{F}_\sigma.$$
   Замечание: в качестве следствия получаем, что если выполнено неравенство $\tau \le \sigma$, то $\mathcal{F}_\tau \subset \mathcal{F}_\sigma$.

5. Пусть $(W_t, t \ge 0)$ — винеровский процесс, а τ_x — первый момент достижения уровня x, т.е. $\tau_x = \inf\{t : W_t = x\}$. С помощью теоремы Башелье вычислите плотность τ_x и $\mathsf{E}\tau_x$.

6. Пусть $(W_t, t \ge 0)$ — винеровский процесс. Для некоторого $y > 0$ положим $\tau = \min\{t : W_t = y\}$. Найдите плотность случайной величины $Y_a = \sup_{t \in [\tau, \tau + a]} W_t$.

7. Пусть $(W_t, t \ge 0)$ — винеровский процесс. Докажите, что для любых $0 \le a < b \le c < d$ с вероятностью единица выполнено
$$\max_{s \in [a,b]} W_s \ne \max_{s \in [c,d]} W_s.$$
   Замечание. Тем самым разумно определена величина
$$T_{[a,b]} = \arg \max_{s \in [a,b]} W_s,$$
   момент времени достижения максимума на отрезке $[a, b]$.

8. Найдите распределение случайной величины $T = T_{[0,1]}$ из предыдущей задачи. А именно, докажите, что для $t \in (0, 1)$
$$\mathsf{P}(T \le t) = \frac{1}{\pi} \arcsin \sqrt{t}.$$

9. Пусть $(W_t, t \geqslant 0)$ — винеровский процесс, а $u > s > 0$. Найдите

$$\mathsf{P}\left(W_t \text{ не имеет нулей на отрезке } [s, u]\right).$$

6. Мартингалы

Большая часть оставшихся глав настоящего учебного пособия посвящена различным классам случайных процессов, наиболее важным с точки зрения приложений. В предыдущих главах мы уже рассмотрели процессы с независимыми приращениями и гауссовские процессы. В этой главе речь пойдет о мартингалах.

▲ **Определение 6.1.** Пусть $T \subset \mathbb{R}$. Процесс $(X_t, t \in T)$ называется *мартингалом относительно фильтрации* $\mathbb{F} = (\mathcal{F}_t, t \in T)$, если

1. X_t согласован с $\mathbb{F}$;
2. X_t является L^1-*процессом*, т.е. для любого $t \in T$ выполнено неравенство $\mathsf{E}|X_t| < +\infty$;
3. $\mathsf{E}(X_t|\mathcal{F}_s) = X_s$ п. н. для любых $s \leqslant t,\ s, t \in T$.

▲ **Определение 6.2.** Если вместо условия 3 выполнено

$$\mathsf{E}(X_t|\mathcal{F}_s) \geqslant X_s \text{ п. н.} \quad \forall s \leqslant t,\ s, t \in T,$$

то процесс X_t называется *субмартингалом* относительно $\mathbb{F}$.

Если же вместо условия 3 выполнено

$$\mathsf{E}(X_t|\mathcal{F}_s) \leqslant X_s \text{ п. н.} \quad \forall s \leqslant t,\ s, t \in T,$$

то X_t называется *супермартингалом* относительно $\mathbb{F}$.

Субмартингалы и супермартингалы — это мартингалы “наполовину”. Для мартингалов условие 3 означает, что функция среднего является постоянной, $\mathsf{E}X_t = \mathsf{E}X_s\ \forall s, t \in T$.

▲ **Определение 6.3.** *Мартингалами* (без указания фильтрации) будем называть мартингалы относительно их естественной фильтрации.

Какие же процессы являются мартингалами? Следующая теорема дает критерий мартингальности для процессов с независимыми приращениями.

Теорема 6.1. *Пусть* $(X_t, t \in T)$ — L^1*-процесс, имеющий независимые приращения. Тогда* X_t — *мартингал* $\Leftrightarrow$ $\mathsf{E}X_t = \mathrm{const}$, *т.е. функция среднего постоянна.*

Несложно проверить, что если функция среднего не убывает, то процесс с независимыми приращениями является субмартингалом. Если же она не возрастает, то тогда — супермартингалом.

Теорема 6.1 сразу дает нам несколько хороших примеров мартингалов.

1. Винеровский процесс $(W_t, t \geqslant 0)$ — мартингал.
2. $(S_n, n \in \mathbb{Z}_+)$ — случайное блуждание, $S_n = \xi_1 + \ldots + \xi_n$, ξ_i — независимые случайные величины, $\mathsf{E}|\xi_i| < +\infty$. Тогда S_n — мартингал $\Leftrightarrow$ для любого i $\mathsf{E}\xi_i = 0$.
3. Пуассоновский процесс $(N_t, t \geqslant 0)$ — субмартингал.
4. Пусть ξ_n, $n \in \mathbb{N}$ — независимые неотрицательные случайные величины, $S_n = \xi_1 \cdot \ldots \cdot \xi_n$. Тогда S_n — мартингал $\Leftrightarrow$ $\mathsf{E}\xi_i = 1$ для всех i.

▲ **Задача 6.1.** Пусть ξ — случайная величина, $\mathsf{E}|\xi| < +\infty$, $\mathbb{F} = (\mathcal{F}_t, t \in T)$ — фильтрация. Докажите, что процесс

$$X_t = \mathsf{E}(\xi|\mathcal{F}_t)$$

является мартингалом относительно $\mathbb{F}$.

Решение

Первые два условия очевидны.
Пусть $t > s$.

$$\mathsf{E}(X_t|\mathcal{F}_s) = \mathsf{E}\left(\mathsf{E}(\xi|\mathcal{F}_t)\Big|\mathcal{F}_s\right) = |\text{т.к. } \mathcal{F}_s \subset \mathcal{F}_t| = \\ = \mathsf{E}(\xi|\mathcal{F}_s) = X_s \ (\text{п. н.}).$$

Задача решена.

Мартингалы из задачи 6.1 принято называть *мартингалами Леви.*

▲ **Задача 6.2.** Пусть $(X_t, t \in T)$ — мартингал относительно $\mathbb{F} = (\mathcal{F}_t, t \in T)$, а $h(x)$ — такая выпуклая книзу функция, что $\mathsf{E}|h(X_t)| < \infty$. Докажите, что процесс

$$Y_t = h(X_t)$$

является субмартингалом относительно $\mathbb{F}$.

Решение

Заметим, что функция h является борелевской, раз она выпуклая. В условии задачи сказано, что существует конечное математическое ожидание Y_t, значит, выполнено условие 2 из определения мартингала. Условие согласованности 1 очевидно из согласованности процесса X_t. Покажем, что выполнено условие 3. Для этого воспользуемся неравенством Йенсена для условного математического ожидания:

$$\mathsf{E}(Y_t|\mathcal{F}_s) = \mathsf{E}(h(X_t)|\mathcal{F}_s) \geqslant h(\mathsf{E}(X_t|\mathcal{F}_s)) = h(X_s) = Y_s \text{ (п. н.)}.$$

Задача решена.

В случае дискретного времени свойство (3) из определения 6.1 достаточно проверять только для соседних моментов времени.

▲ **Задача 6.3.** Докажите, что L^1-процесс $(X_n, n \in \mathbb{Z}_+)$, согласованный с фильтрацией $\mathbb{F} = (\mathcal{F}_n, n \in \mathbb{Z}_+)$, является мартингалом относительно $\mathbb{F}$ тогда и только тогда, когда для любого n выполнено равенство $\mathsf{E}(X_n|\mathcal{F}_{n-1}) = X_{n-1}$ (п. н.).

Решение

В прямую сторону все очевидно. Пусть $n > k$, тогда, используя свойства условного математического ожидания, получаем, что

$$\begin{aligned}\mathsf{E}(X_n|\mathcal{F}_k) &= \mathsf{E}(\mathsf{E}(X_n|\mathcal{F}_{n-1})|\mathcal{F}_k) = \\ &= \mathsf{E}(X_{n-1}|\mathcal{F}_k) = \ldots \\ &= \mathsf{E}(X_{k+1}|\mathcal{F}_k) = X_k \text{ (п.н.)}.\end{aligned}$$

Задача решена.

Рассмотрим два важных свойства мартингалов. Первым является разложение Дуба–Мейера для произвольных процессов с дискретным временем. Для его формулировки нам понадобится понятие предсказуемого процесса.

▲ **Определение 6.4.** Процесс $(X_n, n \in \mathbb{N})$ является *предсказуемым* относительно фильтрации $(\mathcal{F}_n, n \in \mathbb{Z}_+)$, если X_n измерим относительно $\mathcal{F}_{n-1}$ для всех $n \in \mathbb{N}$.

Теорема 6.2 (разложение Дуба–Мейера). *Пусть $(X_n, n \in \mathbb{Z}_+)$ — L^1-процесс, согласованный с фильтрацией $\mathbb{F} = (\mathcal{F}_n, n \in \mathbb{Z}_+)$. Тогда существует единственное разложение вида*

$$X_n = M_n + A_n,$$

где M_n — мартингал относительно $\mathbb{F}$, а $(A_n, n \in \mathbb{N})$ — предсказуемый процесс относительно $\mathbb{F}$, $A_0 = 0$ п. н.

Заметим в связи с теоремой, что процесс $(X_n, n \in \mathbb{Z}_+)$ является субмартингалом относительно $\mathbb{F}$ тогда и только тогда, когда в его разложении Дуба–Мейера последовательность A_n является п. н. неубывающей.

Второе свойство связывает мартингалы с марковскими моментами и называется *теоремой об остановке.*

Теорема 6.3 (теорема об остановке). *Пусть $(X_n, n \in \mathbb{Z}_+)$ – L^1-процесс, согласованный с фильтрацией $\mathbb{F} = (\mathcal{F}_n, n \in \mathbb{Z}_+)$. Тогда X_n является мартингалом (субмартингалом) относительно $\mathbb{F}$ $\Leftrightarrow$ для любых τ, σ — ограниченных почти наверное марковских моментов относительно $\mathbb{F}$ таких, что $\tau \leqslant \sigma$ п. н., выполнено*

$$\mathsf{E}X_\tau = (\leqslant)\ \mathsf{E}X_\sigma.$$

Теорема об остановке означает, что среднее значение в случайные моменты времени у мартингала остается тем же, что и в неслучайные. Более того, это свойство эквивалентно свойству мартингальности.

Можно ли отказаться от ограниченности марковских моментов? Оказывается, нет.

▲ **Задача 6.4.** Приведите пример мартингала $(X_n, n \in \mathbb{Z}_+)$ и таких марковских моментов τ, σ относительно его естественной фильтрации, что $\mathsf{E}X_\tau \neq \mathsf{E}X_\sigma$.

Решение

Рассмотрим простейшее симметричное случайное блуждание на прямой $(S_n, n \in \mathbb{Z}_+)$ и положим

$$\tau = \min\{n : S_n = 1\}, \quad \sigma = \min\{n : S_n = 2\}.$$

Тогда S_n — мартингал, но

$$\mathsf{E}S_\tau = 1 < 2 = \mathsf{E}S_\sigma.$$

Задача решена.

Тем не менее теорема об остановке применима в случае, когда марковские моменты не ограничены, но ограничен сам процесс до некоторого момента времени.

Лемма 6.1 *Пусть $(X_n, n \in \mathbb{Z}_+)$ — мартингал относительно $\mathbb{F} = (\mathcal{F}_n, n \in \mathbb{Z}_+)$, а τ — момент остановки относительно $\mathbb{F}$, причем существует $c > 0$ такое, что*

$$\forall n \in \mathbb{Z}_+ \quad |X_{\min(\tau,n)}| \leqslant c.$$

Тогда $\mathsf{E}X_\tau = \mathsf{E}X_0$.

Перейдем к обсуждению аналогов теоремы об остановке для случая непрерывного времени. Здесь марковские моменты нужно заменить на опциональные.

▲ **Определение 6.5.** Отображение $\tau : \Omega \to [0, +\infty]$ называется *опциональным моментом* относительно фильтрации $\mathbb{F} = (\mathcal{F}_t, t \geqslant 0)$, если

$$\forall t \geqslant 0 \quad \{\omega : \tau(\omega) < t\} \in \mathcal{F}_t.$$

Примерами опциональных моментов могут служить первые моменты попадания процесса в открытые множества.

▲ **Задача 6.5.** Пусть $(X_t, t \geqslant 0)$ — процесс с непрерывными справа траекториями, а $G \subset \mathbb{R}$ — открытое множество. Тогда

$$\tau = \inf\{t : X_t \in G\}$$

является опциональным моментом относительно $\mathbb{F}^X$.

Решение

Достаточно заметить, что в силу непрерывности справа траекторий процесса

$$\{\tau \geqslant t\} = \bigcap_{s \in \mathbb{Q},\ s < t} \{X_s \notin G\} \in \mathcal{F}_t^X.$$

Задача решена.

Теорема 6.4. *Пусть $(X_t, t \geqslant 0)$ — мартингал относительно $\mathbb{F} = (\mathcal{F}_t, t \geqslant 0)$ с непрерывными справа траекториями. Тогда $\forall\ \tau, \sigma : \tau \leqslant \sigma$ (п. н.) — ограниченных опциональных моментов относительно $\mathbb{F}$ выполнено*

$$\mathsf{E}X_\tau = \mathsf{E}X_\sigma.$$

▲ Задачи для самостоятельного решения

1. Пусть $(W_t, t \geqslant 0)$ — винеровский процесс. Докажите, что процесс $Y_t = W_t^2 - t$ является мартингалом относительно естественной фильтрации процесса W_t.

2. Пусть $(W_t, t \geq 0)$ — винеровский процесс. Найдите все такие пары $(\alpha, \beta) \in \mathbb{R}^2$, что процесс

$$\left(X_t = \exp\left\{\alpha W_t + \beta t\right\},\ t \geq 0\right)$$

является мартингалом (субмартингалом, супермартингалом) относительно естественной фильтрации процесса W_t.

3. Пусть $\xi_1, \ldots, \xi_n, \ldots$ — такая последовательность случайных величин, что для любого n существует плотность $f_n(x_1, \ldots, x_n)$ случайного вектора $(\xi_1, \ldots, \xi_n)$. Пусть $\eta_1, \ldots, \eta_n, \ldots$ — другая последовательность случайных величин, причем также для любого n существует плотность $g_n(x_1, \ldots, x_n)$ случайного вектора $(\eta_1, \ldots, \eta_n)$. Докажите, что процесс
$$X_n = \frac{g_n(\xi_1, \ldots, \xi_n)}{f_n(\xi_1, \ldots, \xi_n)}$$
является мартингалом относительно фильтрации $(\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n),\ n \in \mathbb{N})$.

4. Пусть $(W_t, t \geqslant 0)$ — винеровский процесс, а τ — момент остановки относительно его естественной фильтрации. Докажите, что процесс
$$(X_t = W_{\min(t,\tau)},\ t \geqslant 0)$$
является мартингалом относительно естественной фильтрации процесса W_t.

   Указание: надо аппроксимировать τ марковскими моментами с конечным числом значений.

5. Рассмотрим ветвящийся процесс Гальтона–Ватсона $(X_n, n \in \mathbb{Z}_+)$ с законом размножения частиц $\mathrm{Pois}(3)$. Найдите разложение Дуба–Мейера для данного процесса.

6. Докажите, что если τ — марковский момент относительно $\mathbb{F} = (F_t, t \geqslant 0)$, то τ является и опциональным моментом относительно $\mathbb{F}$.

7. Пусть $(W_t, t \geqslant 0)$ — винеровский процесс, а момент τ равен $\min\{t :\ |W_t| = 1\}$. Вычислите $\mathsf{E}\tau$.

8. Пусть $(S_n, n \in \mathbb{N})$ — простейшее случайное блуждание с вероятностью шага вправо p. Пусть $a < x < b$ — целые числа, а процесс $X_n = x + S_n$, где $n \geqslant 1$. Обозначим $\tau = \min\{n : S_n \in \{a, b\}\}$ — момент выхода процесса X_n из полосы. Вычислите $\mathsf{E}\tau$.

7. Марковские процессы

В настоящей главе мы рассмотрим один из наиболее известных классов случайных процессов — марковские процессы.

▲ **Определение 7.1.** Процесс $(X_t, t \in T)$, $T \subset \mathbb{R}$, называется *марковским* относительно фильтрации $\mathbb{F} = (\mathcal{F}_t, t \geqslant 0)$, если X_t согласован с $\mathbb{F}$ и для любых $s \leqslant t$ $(s, t \in T)$ и любой ограниченной борелевской функции f выполнено

$$\mathsf{E}\big(f(X_t)\big|\mathcal{F}_s\big) = \mathsf{E}\big(f(X_t)\big|X_s\big).$$

Марковскими процессами без указания фильтрации называются марковские процессы относительно своей естественной фильтрации.

Смысл определения марковского процесса состоит в том, что распределение "будущего" (X_t) зависит от "прошлого" $(\mathcal{F}_s)$ только через "настоящее" (X_s).

Какие процессы являются марковскими? Для ответа на подобный вопрос полезно применять эквивалентные определения марковского процесса.

Теорема 7.1. *Следующие свойства эквивалентны:*

1. $(X_t, t \in T)$ — *марковский процесс;*
2. *для любых m, $s_1 < \cdots < s_m < s < t$, и любой непрерывной ограниченной борелевской функции f выполнено*
$$\mathsf{E}\big(f(X_t)\big|X_s, X_{s_m}, \ldots, X_{s_1}\big) = \mathsf{E}\big(f(X_t)\big|X_s\big);$$
3. *для любых m, $s_1 < \cdots < s_m < s < t$, $B \in \mathcal{B}(\mathbb{R})$ выполнено*
$$\mathsf{P}\big(X_t \in B\big|X_s, X_{s_m}, \ldots, X_{s_1}\big) = \mathsf{P}\big(X_t \in B\big|X_s\big).$$

Первый класс процессов, который полностью входит в класс марковских процессов, — это процессы с независимыми приращениями.

Теорема 7.2. *Процессы с независимыми приращениями являются марковскими.*

Теорема 7.2 сразу дает нам несколько хороших примеров марковских процессов:

- винеровский процесс $(W_t, t \geqslant 0)$;
- пуассоновский процесс $(N_t, t \geqslant 0)$;
- случайное блуждание на прямой.

▲ **Задача 7.1.** Докажите, что ветвящийся процесс Гальтона–Ватсона $(X_n, n \in \mathbb{Z}_+)$ является марковским.

Решение

Будем пользоваться определением 3 из теоремы 7.1.

$$\begin{aligned}
\mathsf{P}(X_n = j|X_{n-1} = m, \ldots, X_1 = a) = \\
&= \mathsf{P}\left(\sum_{i=1}^{X_{n-1}} \xi_i^{(n)} = j \middle| X_{n-1} = m, \ldots, X_1 = a\right) = \\
&= \mathsf{P}\left(\sum_{i=1}^{m} \xi_i^{(n)} = j \middle| X_{n-1} = m, \ldots, X_1 = a\right) = \\
&= |\xi_i^{(n)} \text{ не зависят от } X_k,\ k \leqslant n-1| = \\
&= \mathsf{P}\left(\sum_{i=1}^{m} \xi_i^{(n)} = j\right) = |\text{аналогично}| = \\
&= \mathsf{P}(X_n = j|X_{n-1} = m).
\end{aligned}$$

Значит, $\mathsf{P}(X_n = j|X_{n-1}, \ldots, X_1) = \mathsf{P}(X_n = j|X_{n-1})$ и процесс является марковским.

Задача решена.

В отличие от процессов с независимыми приращениями гауссовские процессы далеко не всегда являются марковскими. Как мы помним, распределение гауссовских процессов зависит от ковариационной функции. Естественно ожидать, что свойство марковости должно формулироваться как некоторое условие на ковариационную функцию. Это условие сформулировано в следующей теореме.

Теорема 7.3 *Пусть $(X_t, t \in T)$, $T \subseteq \mathbb{R}$, — гауссовский процесс с ковариационной функцией $R(s,t)$. Тогда X_t является марковским тогда и только тогда, когда для любых $s \leqslant u \leqslant t$, $u, s, t \in T$, выполнено равенство*

$$R(s,u) \cdot R(u,t) = R(s,t) \cdot R(u,u).$$

Обширные исследования марковских процессов связаны с изучением переходных функций.

▲ **Определение 7.2.** Пусть $T \subset \mathbb{R}$. Функция $P(s,x,t,B)$, где $s, t \in T, s \leqslant t, x \in \mathbb{R}, B \in \mathcal{B}(\mathbb{R})$, называется *переходной функцией*, если

1. При фиксированных s, x, t функция $P(s,x,t,\cdot)$ является вероятностной мерой на $\mathbb{R}$.
2. При фиксированных s, t, B функция $P(s,\cdot,t,B)$ является борелевской.
3. При $t = s$
$$P(s,x,s,B) = \delta_x(B) = I\{x \in B\}.$$
4. При $s \leqslant u \leqslant t$, $s, u, t \in T$, $x \in \mathbb{R}$, $B \in \mathcal{B}(\mathbb{R})$ выполняется *уравнение Колмогорова–Чепмена*:
$$P(s,x,t,B) = \int_{\mathbb{R}} P(s,x,u,dy) \cdot P(u,y,t,B).$$

▲ **Определение 7.3.** Марковский процесс $(X_t, t \in T)$ *обладает переходной функцией* $P(s,x,t,B)$, если

$$\mathsf{P}(X_t \in B | X_s = x) = P(s,x,t,B)$$

P_{X_s}-п. н. относительно x (т.е. *почти для всех* x из распределения P_{X_s} на $\mathbb{R}$).

Смысл данного определения состоит в том, что у марковского процесса с переходной функцией существует «хороший» вариант условного распределения, стоящего в левой части равенства. Вопрос о том, существует ли у марковского процесса переходная функция, является очень сложным и до конца не изученным. Впрочем, следующая задача показывает, что сама функция $\mathsf{P}(X_t \in B | X_s = x)$ близка к тому, чтобы обладать всеми искомыми свойствами.

▲ **Задача 7.2.** Докажите, что функция $\mathsf{P}(X_t \in B | X_s = x)$ марковского процесса X_t обладает свойствами 1 – 4 почти для всех x по мере P_{X_s}.

Решение

Свойства 1 – 3 следуют из свойств условных распределений. Докажем лишь четвертое свойство. Обозначим $P(s, x, t, B) = \mathsf{P}(X_t \in B | X_s = x)$. Тогда для любых $s \leqslant u \leqslant t$:

$$\begin{aligned}
P(s, X_s, t, B) &= \mathsf{P}(X_t \in B | X_s) = \mathsf{E}(I\{X_t \in B\} | X_s) = \\
&= \mathsf{E}(\mathsf{E}(I\{X_t \in B\} | X_u, X_s) | X_s) = \\
&= \mathsf{E}(\mathsf{E}(I\{X_t \in B\} | X_u) | X_s) = \\
&= \mathsf{E}(\mathsf{P}(X_t \in B | X_u) | X_s) = \\
&= \mathsf{E}(P(u, X_u, t, B) | X_s) \quad (\text{п. н.}).
\end{aligned}$$

Значит,

$$P(s, x, t, B) = \mathsf{E}(P(u, X_u, t, B) | X_s = x) =$$

$$= |\text{обозначим через } Q$$

$$\text{условное распределение } X_u \text{ при } X_s = x| =$$

$$\int_{\mathbb{R}} P(u, y, t, B) Q(dy) = |\text{т.к. } Q(A) = P(s, x, u, A)| =$$

$$= \int_{\mathbb{R}} P(u, y, t, B) \cdot P(s, x, u, dy) \quad (\mathsf{P}_{X_s}\text{-п. н.}).$$

Задача решена.

Задача 7.2 означает, что для произвольных марковских процессов имеет место лишь ослабленный вариант свойства 4, справедливый не для всех x, а лишь для почти всех по мере P_{X_s}.

В заключение дадим определение переходной плотности.

▲ **Определение 7.4.** Если переходная функция $P(s, x, t, B)$ марковского процесса X_t имеет вид

$$P(s, x, t, B) = \int_B p(s, x, t, y) dy,$$

то $p(s, x, t, y)$ называется *переходной плотностью* процесса X_t.

В следующих главах мы рассмотрим специальные подклассы марковских процессов — марковские цепи.

▲ Задачи для самостоятельного решения

1. Пусть $(X_t, t \in T)$ — марковский процесс, $T \subset \mathbb{R}_+$. Пусть для любого $t \in T$ задана борелевская функция h_t. Рассматривается случайный процесс $Y_t = (h_t(X_t), t \in T)$. Докажите, что если h_t — биекция для любого $t \in T$ (считаем, что в этом случае h_t^{-1} — тоже борелевская), то Y_t — тоже марковский процесс. Приведите пример марковского процесса X_t и борелевских функций h_t, при которых процесс Y_t не является марковским.

2. Пусть $(X_t, t \in T)$, $(Y_t, t \in T)$ — независимые марковские процессы. Верно ли, что процесс $(X_t + Y_t, t \in T)$ тоже марковский?

3. Верно ли, что $Y = (Y_n = X_0 + \ldots + X_n, n \in \mathbb{Z}_+)$ (общее число частиц, существовавших в ветвящемся процессу Гальтона–Ватсона $X = (X_n, n \in \mathbb{Z}_+)$) является марковским процессом а) относительно естественной фильтрации процесса X, б) относительно естественной фильтрации процесса Y?

4. Пусть $(W_t, t \geqslant 0)$ — винеровский процесс. Является ли марковским процесс $X_t = W_t^2$?

5. Пусть $(W_t, t \geqslant 0)$ — винеровский процесс, и для всех $x \geq 0$ задана случайная величина $\tau_x = \min\{t : \ W_t = x\}$. Докажите, что процесс $\tau = (\tau_x, x \geqslant 0)$ является марковским.

6. Пусть $(W_t, t \geqslant 0)$ — винеровский процесс. Вычислите его переходную плотность.

7. Пусть $(X_n, n \geqslant 0)$ — независимые случайные величины с равномерным распределением на множестве $\{-1, 0, 1\}$. Рассмотрим процесс

$$Y_n = X_0X_1 + X_1X_2 + \ldots + X_{n-1}X_n.$$

   Докажите, что X_n является мартингалом относительно фильтрации

$$\mathbb{F} = (\sigma(X_1, \ldots, X_n), n \in \mathbb{Z}_+),$$

   но не является марковским процессом относительно нее.

8. Марковские цепи с дискретным временем

Марковские цепи — это марковские процессы, принимающие не более чем счетное число значений. Пусть $\mathcal{X}$ — не более чем счетное множество. Для удобства всегда можно считать, что либо $\mathcal{X} = \{1, \ldots, N\}$, либо $\mathcal{X} = \mathbb{N}$.

▲ **Определение 8.1.** Случайный процесс $(X_n, n \in \mathbb{Z}_+)$ со значениями в $\mathcal{X}$ называется *марковской цепью*, если для любого $m \in \mathbb{N}$, для любых $k_1 < \ldots < k_m < k < n$ и для любых $a_1, \ldots, a_m, i, j \in \mathcal{X}$ выполнено (*марковское свойство*)

$$\mathsf{P}(X_n = j | X_k = i, X_{k_m} = a_m, \ldots, X_{k_1} = a_1) = \mathsf{P}(X_n = j | X_k = i)$$

всегда, когда $\mathsf{P}(X_k = i, X_{k_m} = a_m, \ldots, X_{k_1} = a_1) > 0$.

Определение означает, что «будущее» процесса (X_n) и его «прошлое» $(X_{k_m}, \ldots, X_{k_1})$ независимы при фиксированном «настоящем» (X_k). Также из теоремы 7.1 следует, что марковская цепь — это марковский процесс с не более чем счетным числом значений. Формальное утверждение, объясняющее обозначенную выше независимость, сформулировано ниже.

Лемма 8.1. *Пусть $(X_n, n \in \mathbb{Z}_+)$ — марковская цепь. Для $a_0, \ldots, a_n \in \mathcal{X}$ обозначим*

$$\begin{aligned} A &= \{X_n = a_n, \ldots, X_{k+1} = a_{k+1}\}, && \text{(будущее)} \\ B &= \{X_k = a_k\}, && \text{(настоящее)} \\ C &= \{X_{k-1} = a_{k-1}, \ldots, X_0 = a_0\}. && \text{(прошлое)} \end{aligned}$$

Тогда $\mathsf{P}(AC|B) = \mathsf{P}(A|B) \cdot \mathsf{P}(C|B)$.

В теории марковских цепей принята следующая терминология.

- Множество значений цепи $\mathcal{X}$ называется *фазовым пространством.* Элементы цепи называются *состояниями* цепи.

- Условные вероятности $p_{ij}(k,n) := \mathsf{P}(X_n = j | X_k = i)$ при $k < n$, $i,j \in \mathcal{X}$ называются *переходными вероятностями* марковской цепи.

- Переходные вероятности при фиксированных $k < n$ образуют *матрицу переходных вероятностей* $P(k,n) = \big\| p_{ij}(k,n) \big\|_{i,j \in \mathcal{X}}$, где i — номер строки матрицы, а j — номер столбца.

- Вектор $\Pi(0) = \big(p_j(0),\ j \in \mathcal{X}\big)$ распределения случайной величины X_0, где $p_j(0) = \mathsf{P}(X_0 = j)$, называется *начальным распределением цепи.*

- Вектор $\Pi(n) = \big(p_j(n),\ j \in \mathcal{X}\big)$, где $p_j(n) = \mathsf{P}(X_n = j)$, является *распределением цепи в момент времени n.*

Как мы увидим далее, векторы $\Pi(n)$ удобно понимать как вектор–строки.

Легко понять, что конечномерные распределения марковской цепи однозначно определяются переходными вероятностями и начальным распределением.

▲ **Задача 8.1.** Пусть $(X_n, n \in \mathbb{Z}_+)$ — марковская цепь с начальным распределением $\Pi(0)$ и переходными вероятностями $p_{ij}(k,n)$. Вычислите $\mathsf{P}(X_{k_n} = a_n, \ldots, X_{k_1} = a_1)$ при $k_n > \ldots > k_1$.

Решение

Согласно марковскому свойству, определению условной вероятности и формуле полной вероятности,

$$\mathsf{P}(X_{k_n} = a_n, \ldots, X_{k_1} = a_1) =$$

$$= \prod_{j=1}^{n} \mathsf{P}(X_{k_j} = a_j | X_{k_{j-1}} = a_{j-1}, \ldots, X_{k_1} = a_1) \mathsf{P}(X_{k_1} = a_1) =$$

$$= \prod_{j=1}^{n} \mathsf{P}(X_{k_j} = a_j | X_{k_{j-1}} = a_{j-1}) \mathsf{P}(X_{k_1} = a_1) =$$

$$= \prod_{j=1}^{n} p_{a_{j-1}a_j}(k_{j-1}, k_j) \mathsf{P}(X_{k_1} = a_1) =$$

$$= \left(\prod_{j=1}^{n} p_{a_{j-1}a_j}(k_{j-1}, k_j) \right) \times$$

$$\times \sum_{a_0 \in \mathcal{X}} \mathsf{P}(X_{k_1} = a_1 | X_0 = a_0) \mathsf{P}(X_0 = a_0) =$$

$$= \left(\prod_{k=1}^{n} p_{a_{j-1}a_j}(k_{j-1}, k_j) \right) \sum_{a_0 \in \mathcal{X}} p_{a_0 a_1}(0, k_1) p_{a_0}(0).$$

Задача решена.

Перечислим свойства переходных вероятностей.

Лемма 8.2. *Переходные вероятности марковской цепи удовлетворяют следующим свойствам:*

1. $0 \leqslant p_{ij}(k, l) \leqslant 1$;
2. *для любых* i, k, l
$$\sum_{j \in \mathcal{X}} p_{ij}(k, l) = 1;$$
3. $p_{ij}(k, k) = \delta_{ij}$ *(т.е. 1, если* $i = j$, *и 0 иначе)*;
4. *выполнены уравнения Колмогорова–Чепмена: для любых* $k < l < n$, $i, j \in \mathcal{X}$
$$p_{ij}(k, n) = \sum_{\alpha \in \mathcal{X}} p_{i\alpha}(k, l) \cdot p_{\alpha j}(l, n).$$

▲ **Задача 8.2.** Выведите уравнения Колмогорова–Чепмена.

Решение

По формуле полной вероятности,

$$
\begin{aligned}
p_{ij}(k,n) = \mathsf{P}(X_n = j | X_k = i) = \\
&= \sum_{\alpha \in \mathcal{X}} \mathsf{P}(X_n = j, X_l = \alpha | X_k = i) = \\
&= \sum_{\alpha \in \mathcal{X}} \mathsf{P}(X_n = j | X_l = \alpha, X_k = i) \cdot \mathsf{P}(X_l = \alpha | X_k = i) = \\
&= |\text{марковское свойство}| = \\
&= \sum_{\alpha \in \mathcal{X}} \mathsf{P}(X_n = j | X_l = \alpha) \cdot \mathsf{P}(X_l = \alpha | X_k = i) = \\
&= \sum_{\alpha \in \mathcal{X}} p_{i\alpha}(k,l) \cdot p_{\alpha j}(l,n).
\end{aligned}
$$

Задача решена.

Данные свойства являются не только необходимыми, но и достаточными для того, чтобы при заданных функциях с такими свойствами существовала марковская цепь с равными этим функциям переходными вероятностями. Об этом говорит теорема о существовании марковских цепей.

Теорема 8.1. *Пусть $\mathcal{X}$ — не более чем счетное множество, Π — распределение вероятностей на $\mathcal{X}$, а функции $p_{ij}(k,n)$, $i,j \in \mathcal{X}$, $k < n \in \mathbb{Z}_+$, удовлетворяют свойствам 1 – 4 из леммы 8.2. Тогда существует марковская цепь с фазовым пространством $\mathcal{X}$, начальным распределением Π и переходными вероятностями $p_{ij}(k,n)$.*

Важнейшим классом марковских цепей являются однородные марковские цепи.

▲ **Определение 8.2.** Марковская цепь называется *однородной*, если $p_{ij}(k,n)$ зависит только от $i, j, n-k$, т.е.

$$p_{ij}(k,n) = p_{ij}(0, n-k) =: p_{ij}(n-k).$$

В этом случае функции $p_{ij}(k)$ называются *переходными вероятностями за k шагов*, а $p_{ij} = p_{ij}(1)$ — *переходными вероятностями (за один шаг)*.

Приведем примеры однородных марковских цепей.

1. Простейшее случайное блуждание.

$$\begin{gathered}\mathsf{P}(S_n = j|S_{n-1} = i, S_{n-2} = a_{n-2}, \ldots, S_1 = a_1) = \\ = \mathsf{P}(S_{n-1} + \xi_n = j|S_{n-1} = i, S_{n-2} = a_{n-2}, \ldots, S_1 = a_1) = \\ = \mathsf{P}(\underbrace{\xi_n = j - i|S_{n-1} = a_{n-1}, S_{n-2} = a_{n-2}, \ldots, S_1 = a_1}_{\text{независимы}}) = \\ = \mathsf{P}(\xi_n = j - i),\end{gathered}$$

   которая не зависит от n, так как все ξ_k — одинаково распределенные. Значит, это однородная марковская цепь.

2. Ветвящиеся процессы Гальтона–Ватсона.

$$\begin{gathered}\mathsf{P}(X_n = j|X_{n-1} = i, \ldots, X_1 = a_1) = \\ = \mathsf{P}\left(\sum_{k=1}^{X_{n-1}} \xi_k^{(n)} = j|X_{n-1} = i, \ldots, X_1 = a_1\right) = \\ = \mathsf{P}\left(\sum_{k=1}^{i} \xi_k^{(n)} = j\right)\end{gathered}$$

   — не зависит от n. Значит, это однородная марковская цепь.

Введем ряд обозначений для матриц переходных вероятностей:

- $P(k) = \|p_{ij}(k)\|_{i,j \in \mathcal{X}}$,
- $P = P(1) = \|p_{ij}\|_{i,j \in \mathcal{X}}$.

Матрица $P(k)$ называется *матрицей переходных вероятностей за k шагов*, а P — *матрицей переходных вероятносей* (*за* 1 *шаг*).

Отметим ряд простых свойств матриц переходных вероятностей, понимая под векторами распределений векторы-строки.

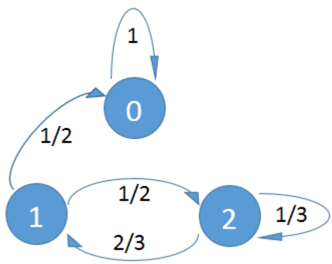
Лемма 8.3. *Для любых $k, l \in \mathbb{N}$ выполнены следующие равенства*:

1. $P(k+l) = P(k) \cdot P(l)$.
2. $P(k) = P^k$.
3. $\Pi(k+l) = \Pi(k) \cdot P(l)$.
4. $\Pi(k) = \Pi(0) \cdot P^k$.

Лемма показывает, что распределение однородной марковской цепи определяется лишь начальным распределением и переходными вероятностями за один шаг. Саму же цепь зачастую удобно изображать с помощью ориентированного графа. Например, пусть $\mathbb{X} = \{0, 1, 2\}$ и

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 2/3 & 1/3 \end{pmatrix}.$$

Данной матрице соответствует следующий граф:



Перейдем к обсуждению предельных свойств однородных марковских цепей. Для этого введем понятия стационарного и предельного распределений для матриц переходных вероятностей.

▲ **Определение 8.3.** Распределение $\Pi = (\pi_j, j \in \mathcal{X})$ вероятностей на $\mathcal{X}$ называется *стационарным* для матрицы переходных вероятностей P (или для марковской цепи), если $\Pi = \Pi \cdot P$, т.е.

для любого $j \in \mathcal{X}$ выполнено равенство $\pi_j = \sum_{\alpha \in \mathcal{X}} \pi_\alpha p_{\alpha j}$.

Термин "стационарность" хорошо иллюстрирует тот факт, что если начальное распределение Π является стационарным, то $\forall n \ \Pi(n) = \Pi$. Действительно, $\Pi(n) = \Pi(0) \cdot P^n = \Pi \cdot P^n = \Pi$, т.е. распределение цепи не меняется со временем.

▲ **Определение 8.4.** Распределение вероятностей Π называется *предельным* для матрицы переходных вероятностей P, если для любого начального распределения $\Pi(0)$ выполнено

$$\Pi(n) = \Pi(0) \cdot P^n \to \Pi, \ n \to \infty.$$

Предельное распределение — это то распределение, к которому распределение цепи сходится, независимо от начального. Система в некотором смысле "забывает", из какого начального состояния она "стартовала", и сходится в пределе к одному и тому же распределению. Конечно, предельное распределение гораздо важнее и интереснее, чем стационарное, однако второе искать намного проще и удобнее — достаточно решить одно линейное уравнение. Однако оказывается, что во многих случаях эти два распределения совпадают. Связывает их утверждение, называемое эргодической теоремой.

Теорема 8.2 (эргодическая теорема). *Пусть $(X_n, n \in \mathbb{Z}_+)$ — однородная марковская цепь с матрицей переходных вероятностей P и фазовым пространством $\mathcal{X} = \{1, \ldots, N\}$. Если существует такое n_0, что для любых $i, j \in \mathcal{X}$*

$$p_{ij}(n_0) > 0,$$

то существуют числа $(\pi_1, \ldots, \pi_N)$ со следующими свойствами: для любого $j \in \mathcal{X}$

$$\pi_j > 0, \quad \sum_{j \in \mathcal{X}} \pi_j = 1, \quad \lim_{n \to \infty} p_{ij}(n) = \pi_j \ \forall i \in \mathcal{X}.$$

▲ **Определение 8.5.** Распределение $\Pi = (\pi_1, \ldots, \pi_N)$ из эргодической теоремы называется *эргодическим* для P.

Следующая лемма показывает, что эргодическое распределение является одновременно и стационарным, и предельным.

Лемма 8.4. *В условиях эргодической теоремы эргодическое распределение является и предельным, и единственным стационарным.*

Приведем пример применения эргодической теории.

▲ **Задача 8.3.** Имеются две урны, содержащие в начальный момент времени соответственно k_1 и k_2 шаров, причем $k_1+k_2 = k$. В каждый момент времени n с вероятностью $1/k$ выбирается один из этих шаров и перекладывается из той урны, где он лежал, в другую. Пусть X_n — это число шаров в первой урне в момент времени n. Докажите, что $(X_n, n \in \mathbb{Z}_+)$ — однородная марковская цепь, и найдите ее стационарное распределение. Является ли оно предельным?

Решение

Заметим, что если $X_{n-1} = i$, то X_n может равняться только $i+1$ или $i-1$. Более того, согласно определению модели,

$$\mathsf{P}(X_n = i + 1 | X_{n-1} = i, \dots, X_1 = a_1) = \frac{k-i}{k},$$

ведь мы должны были выбрать шар из второй урны. Эта же вероятность равна $\mathsf{P}(X_n = i + 1 | X_{n-1} = i)$, что и доказывает марковость процесса.

Для нахождения стационарного распределения выпишем переходные вероятности за один шаг:

$$p_{ij} = \begin{cases} \frac{i}{k} & \text{если } j = i - 1; \\ \frac{k-i}{k} & \text{если } j = i + 1; \\ 0 & \text{иначе.} \end{cases}$$

Тем самым нам необходимо решить следующую систему относительно $(\pi_0, \dots, \pi_k)$:

$$\pi_0 = \frac{1}{k}\pi_1,\ \pi_i = \frac{k-i+1}{k}\pi_{i-1} + \frac{i+1}{k}\pi_{i+1},\ i = 1,\ldots,k-1,$$

$$\pi_0 + \ldots + \pi_k = 1.$$

Решение системы задается соотношением $\pi_i = \binom{k}{i}\pi_0$ для всех $i \in \{0, 1, \ldots, k\}$. Действительно, рассуждая по индукции, получаем, что при выполнении формулы для π_{i-1} и π_i

$$\pi_{i+1} = \frac{1}{i+1}\left(k\pi_i - (k-i+1)\pi_{i-1}\right) =$$

$$= \frac{1}{i+1}\left(k\binom{k}{i} - (k-i+1)\binom{k}{i-1}\right)\pi_0 =$$

$$= \frac{k(k-i)}{i(i+1)}\binom{k-1}{i-1}\pi_0 = \binom{k}{i+1}\pi_0.$$

Из последнего уравнения системы находим $\pi_i = \binom{k}{i}2^{-k}$.
Осталось понять, что происходит с предельным распределением. Заметим, что эргодическая теорема не применима, так как ни в какой степени матрица переходных вероятностей не имеет все элементы положительными. Это подсказывает, что предельного распределения не существует.
Действительно, пусть Π — это предельное распределение. Рассмотрим марковскую цепь $Y_n = X_{2n}$, тогда у нее предельное распределение также существует и равно Π. Но в силу построения модели, все элементы Y_n всегда имеют ту же четность, что и $Y_0 = X_0$. Стало быть, если $\mathsf{P}(X_0 = 0) = 1$, то предельное распределение должно быть сосредоточено только в состояниях с четными номерами, а если $\mathsf{P}(X_0 = 1) = 1$, то — в нечетных. Тем самым имеется явная зависимость от начального распределения, что противоречит самому определению предельного распределения.

Задача решена.

▲ Задачи для самостоятельного решения

1. Докажите, что процесс $(X_n, n \in \mathbb{Z}_+)$ со значениями в не более чем счетном множестве $\mathcal{X}$ является марковской цепью тогда и только тогда, когда для любого $n \in \mathbb{Z}_+$ и любых $a_{n+1}, \ldots, a_0 \in \mathcal{X}$ выполнено

$$\mathsf{P}\left(X_{n+1} = a_{n+1} | X_n = a_n, \ldots, X_0 = a_0\right) =$$
$$= \mathsf{P}\left(X_{n+1} = a_{n+1} | X_n = a_n\right)$$

всегда, когда вероятности условий положительны.

2. Пусть $\{\xi_n, n \in \mathbb{N}\}$ — независимые одинаково распределенные случайные величины с плотностью $p(x)$, положительной при всех $x \in \mathbb{R}$. Укажите, являются ли следующие случайные последовательности цепями Маркова:

   а) $\xi = (\xi_n, n \in \mathbb{N})$;

   б) $S = (S_n = \xi_1 + \ldots + \xi_n, n \in \mathbb{Z}_+)$, где $S_0 = 0$;

   в) $S^+ = (S_n I(S_n \geq 0), n \in \mathbb{Z}_+)$;

   г) $T = (T_n = \max\{S_0, \ldots, S_n\}, n \in \mathbb{Z}_+)$.

   Для тех, которые окажутся цепями Маркова, укажите переходные вероятности за один шаг.

3. Пусть $\{\xi_n, n \in \mathbb{Z}_+\}$ — марковская цепь с фазовым пространством $S = \{1, 2, 3\}$, начальным состоянием $\xi_0 = 1$ п. н. и матрицей переходных вероятностей

$$\begin{pmatrix} 3/7 & 3/7 & 1/7 \\ 1/11 & 2/11 & 8/11 \\ 1/11 & 4/11 & 6/11 \end{pmatrix}.$$

   Положим $\eta_n = I\{\xi_n = 1\} + 2I\{\xi_n \neq 1\}$. Докажите, что η_n — тоже марковская цепь, и найдите ее матрицу переходов.

4. Марковская цепь $(\xi_n, n \in \mathbb{Z}_+)$ имеет начальное состояние $\xi_0 = 0$ и переходные вероятности

$$\mathsf{P}(\xi_{n+1} = k + 1 | \, \xi_n = k) = p, \quad \mathsf{P}(\xi_{n+1} = k | \, \xi_n = k) = 1 - p,$$

   где $k, n \in \mathbb{N}$, $p \in [0, 1]$. Найдите распределение ξ_n. Докажите, что последовательность $\tau_0 = 0$, $\tau_k = \min\{n : \ \xi_n = k\}$

также является цепью Маркова и найдите ее переходные вероятности.

5. Цепь Маркова ξ_n имеет начальное состояние $\xi_0 = 0$ и переходные вероятности

$$\mathsf{P}(\xi_{n+1} = k+1|\xi_n = k) = a^{-k}, \quad \mathsf{P}(\xi_{n+1} = k|\xi_n = k) = 1-a^{-k},$$

где $k, n \in \mathbb{Z}_+$, $a > 1$. Найдите $\mathsf{E}a^{\xi_n}$, $\mathsf{D}a^{\xi_n}$.

6. Приведите пример такой однородной марковской цепи с дискретным временем, что

   а) у нее есть несколько стационарных распределений, но нет предельного;

   б) у нее нет стационарного распределения, но есть пределы переходных вероятностей при $n \to \infty$.

   Докажите, что если однородная марковская цепь с дискретным временем имеет несколько стационарных распределений, то их, на самом деле, бесконечно много.

7. Пусть $(\xi_n, n \in \mathbb{Z}_+)$ — независимые одинаково распределенные случайные величины со значениями $\{0, 1, 2, 3\}$ и следующим распределением:

$$\mathsf{P}(\xi_n = 0) = 1/7, \ \ \mathsf{P}(\xi_n = 1) = 2/7, \ \ \mathsf{P}(\xi_n = 2) = 3/7,$$

$$\mathsf{P}(\xi_n = 3) = 1/7.$$

   Рассматриваются процессы $X_n = \xi_1 + \ldots + \xi_n \pmod 4$, $n \in \mathbb{N}$ (остаток от деления суммы на 4) и $Y_n = \xi_1 \cdot \ldots \cdot \xi_n \pmod 4$. Докажите, что $(X_n, n \in \mathbb{N})$ и $(Y_n, n \in \mathbb{N})$ являются однородными марковскими цепями, и найдите их предельные распределения.

9. Марковские цепи с непрерывным временем

Настоящий раздел посвящен марковским цепям с непрерывным временем. Определение практически ничем не отличается от случая дискретного времени. Пусть $\mathcal{X}$ — не более чем счетное множество.

▲ **Определение 9.1.** Случайный процесс $(X_t, t \geqslant 0)$ со значениями в $\mathcal{X}$ называется *марковской цепью*, если для любого натурального n, для любых $a_1, \ldots, a_n, i, j \in \mathcal{X}$ и для любых моментов $t \geqslant s > s_n > \cdots > s_1 \geq 0$

$$\mathsf{P}(X_t = j | X_s = i, X_{s_n} = a_n, \ldots, X_{s_1} = a_1) = \mathsf{P}(X_t = j | X_s = i)$$

всегда, когда $\mathsf{P}(X_s = i, X_{s_n} = a_n, \ldots, X_{s_1} = a_1) > 0$.

Терминология повторяет дискретный случай.

- Множество $\mathcal{X}$ называется *фазовым пространством*, или *пространством состояний* марковской цепи.
- Условные вероятности

  $$p_{ij}(s,t) := \mathsf{P}(X_t = j | X_s = i)$$

  называются *переходными вероятностями* марковской цепи.
- Из переходных вероятностей можно составить *матрицы переходных вероятностей* $P(s,t) = \big\|p_{ij}(s,t)\big\|_{i,j\in\mathcal{X}}$, где i — номер строки матрицы, j — номер столбца.
- $\Pi(0) = \big(p_j(0),\ j \in \mathcal{X}\big)$, где $p_j(0) = \mathsf{P}(X_0 = j)$ — *начальное распределение цепи.*
- $\Pi(t) = \big(p_j(t),\ j \in \mathcal{X}\big)$, где $p_j(t) = \mathsf{P}(X_t = j)$ — *распределение цепи в момент времени* $t \geqslant 0$.

Как и в дискретном случае, конечномерные распределения однозначно определяются переходными вероятностями и начальным распределением. Свойства переходных вероятностей остаются теми же.

Лемма 9.1. *Переходные вероятности марковской цепи удовлетворяют следующим свойствам:*

1. $0 \leqslant p_{ij}(s,t) \leqslant 1$;
2. $\sum\limits_{j\in\mathcal{X}} p_{ij}(s,t) = 1 \; \forall i \in \mathcal{X}$;
3. $p_{ij}(s,s) = \delta_{ij}$;
4. *выполнены уравнения Колмогорова–Чепмена*: *для любых* $s \leqslant u \leqslant t$

$$p_{ij}(s,t) = \sum_{\alpha\in\mathcal{X}} p_{i\alpha}(s,u)p_{\alpha j}(u,t).$$

Как и в дискретном случае, перечисленные свойства не только необходимы, но и достаточны для существования марковской цепи.

Теорема 9.1. *Пусть $\mathcal{X} \subset \mathbb{R}$ — не более чем счетное множество, $\Pi(0) = \big(p_j(0), j \in \mathcal{X}\big)$ — распределение вероятностей на $\mathcal{X}$, а для любых $i,j \in \mathcal{X}$ $p_{ij}(s,t)$, $0 \leqslant s \leqslant t$, — функции, обладающие свойствами 1)–4) из леммы* 9.1. *Тогда существует такая марковская цепь $(X_t, t \geqslant 0)$ с фазовым пространством $\mathcal{X}$, что $\Pi(0)$ является её начальным распределением, а $p_{ij}(s,t)$ являются её переходными вероятностями.*

Нас снова будут интересовать предельные распределения однородных марковских цепей.

▲ **Определение 9.2.** Марковская цепь $(X_t, t \geqslant 0)$ называется *однородной*, если её переходные вероятности $p_{ij}(s,t)$ зависят только от i, j и $t-s$, т.е.

$$p_{ij}(s,t) = p_{ij}(0, t-s) =: p_{ij}(t-s).$$

Функция $p_{ij}(t)$ называется *переходной вероятностью за время* t, а

$$P(t) = \left\|p_{ij}(t)\right\|_{i,j\in\mathcal{X}}$$

— *матрицей переходных вероятностей за время* t.

Заметим, что, в силу леммы 9.1, матрицы $P(t)$ являются *стохастическими*, т.е. их элементы неотрицательны и сумма элементов в каждой строке равна единице. Более того, уравнения Колмогорова–Чепмена обеспечивают для них полугрупповое свойство.

▲ **Определение 9.3.** *Стохастической полугруппой* называется набор стохастических матриц $(P(t), t \geqslant 0)$, если

1. $P(0) = \mathrm{I}$ (единичная матрица);

2. $P(t+s) = P(t)\cdot P(s) \quad \forall s,t \geqslant 0.$

Тем самым, в отличие от случая дискретного времени, нам приходится иметь дело с целой стохастической полугруппой матриц вместо одной матрицы переходных вероятностей за один шаг, что усложняет изучение стационарных и предельных распределений.

▲ **Определение 9.4.** Распределение $\Pi = (\pi_j, j \in \mathcal{X})$ вероятностей на $\mathcal{X}$ называется *стационарным* распределением для стохастической полугруппы $(P(t), t \geqslant 0)$, если для любого $t \geqslant 0$ выполнено равенство

$$\Pi = \Pi \cdot P(t),$$

т.е. $\pi_j = \sum_\alpha \pi_\alpha p_{\alpha j}(t)$.

Тем самым нахождение стационарного распределения требует решения бесконечной системы линейных уравнений.

▲ **Задача 9.1.** Докажите, что если стационарное распределение Π взять в качестве начального, то для любого $t \geqslant 0$ $\Pi(t) = \Pi$, т.е. распределение цепи не будет меняться со временем.

Решение

Заметим, что для любого $j \in \mathcal{X}$

$$p_j(t) = \mathsf{P}(X_t = j) = \sum_{\alpha \in \mathcal{X}} \mathsf{P}(X_t = j | X_0 = \alpha) \mathsf{P}(X_0 = \alpha) = \\ = \sum_{\alpha \in \mathcal{X}} p_{\alpha j}(t) \pi_\alpha = \pi_j.$$

Задача решена.

▲ **Определение 9.5.** Распределение $\Pi = (\pi_j, j \in \mathcal{X})$ вероятностей на $\mathcal{X}$ называется *предельным* для стохастической полугруппы $(P(t), t \geqslant 0)$, если для любого начального распределения $\Pi(0)$ выполнено:

$$\Pi(t) = \Pi(0) \cdot P(t) \to \Pi \text{ при } t \to +\infty,$$

т.е. сходимость $p_j(t) \to \pi_j$ имеет место для любого $j \in \mathcal{X}$.

Стационарные и предельные распределения связывает эргодическая теорема.

Теорема 9.2 (эргодическая теорема) *Пусть $(P(t), t \geqslant 0)$ — стохастическая полугруппа. Если существуют $j_0 \in \mathcal{X}$ и $h, \delta > 0$ такие, что для любого $i \in \mathcal{X}$*

$$p_{ij_0}(h) \geqslant \delta,$$

то существует набор неотрицательных чисел $(\pi_j, j \in \mathcal{X})$ таких, что для любых $i, j \in \mathcal{X}$

$$\lim_{t \to +\infty} p_{ij}(t) = \pi_j, \text{ причем } |p_{ij}(t) - \pi_j| \leqslant (1-\delta)^{[t/2]}.$$

В отличие от случая дискретного времени, набор чисел π_j из эргодической теоремы не всегда является стационарным и предельным распределением. В качестве следствия из эргодической теоремы сформулируем следующую лемму о свойствах $\Pi = (\pi_j, j \in \mathcal{X})$.

Лемма 9.2. *В условиях эргодической теоремы набор* $\Pi = (\pi_j, j \in \mathfrak{X})$ *удовлетворяет следующим свойствам.*

1) $\Pi(t) \to \Pi$ *при* $t \to +\infty$, *т.е. для каждого* j *выполнено* $p_j(t) \to \pi_j, t \to +\infty$.

2) *выполнены равенства*

$$\Pi = \Pi \cdot P(t),\ \forall t > 0,$$

или для любого $j \in \mathfrak{X}$

$$\pi_j = \sum_{\alpha} \pi_\alpha \cdot p_{\alpha j}(t).$$

3) *либо* $\sum_j \pi_j = 1$ (*и тогда* $\Pi = (\pi_j, j \in \mathfrak{X})$ — *это стационарное и предельное распределение*), *либо* $\sum_j \pi_j = 0$ (*возможно только для бесконечного* $\mathfrak{X}$).

Лемма показывает, что мы снова можем искать предельное распределение (если Π действительно распределение) как стационарное, но остался вопрос о том, как найти само стационарное распределение. Нам нужен аналог матрицы переходных вероятностей за один шаг. Подобную роль в случае непрерывного времени играет инфинитезимальная матрица марковской цепи.

▲ **Определение 9.6.** Пусть $(P(t), t \geqslant 0)$ — стохастическая полугруппа. Тогда матрица

$$Q = \frac{d^+}{dt} P(t)\Big|_{t=0}$$

называется *инфинитезимальной* матрицей марковской цепи. Её элементы выражаются формулой:

$$q_{ij} = \lim_{h \to 0+} \frac{p_{ij}(h) - p_{ij}(0)}{h}.$$

Инфинитезимальная матрица — это правая производная стохастической полугруппы в нуле. Но возникает естественный вопрос о существовании такой производной. Оказывается, что при наличии непрерывности в нуле стохастической полугруппы производная всегда существует.

▲ **Определение 9.7.** Стохастическая полугруппа $(P(t), t \geqslant 0)$ называется *стандартной*, если

$$P(t) \to P(0) = \mathrm{I} \text{ (единичная матрица)},$$

при $t \to 0+$, т.е. для всех $i, j \in \mathcal{X}$ выполнено $\lim_{t\to 0+} p_{ij}(t) = \delta_{ij}$.

Теорема 9.3. *Пусть $(P(t), t \geqslant 0)$ — стандартная стохастическая полугруппа. Тогда у нее существует инфинитезимальная матрица Q, причем для любых $i \neq j$ $0 \leqslant q_{ij} < +\infty$, а $q_{ii} \in [-\infty, 0]$.*

Теорема говорит о том, что внедиагональные элементы матрицы Q всегда неотрицательны и конечны, а диагональные элементы всегда неположительные, но могут равняться $-\infty$. Матрица Q в некотором смысле определяет всю цепь целиком. По ней можно восстановить переходные вероятности с помощью дифференциальных уравнений Колмогорова. Для формулировки соответствующих теорем нам понадобится понятие консервативности цепи.

▲ **Определение 9.8.** Марковская цепь называется *консервативной*, если все элементы инфинитезимальной матрицы Q конечны и для любого $i \in \mathcal{X}$

$$\sum_{j\in\mathcal{X}} q_{ij} = 0.$$

Свойство консервативности очень естественно, ведь мы помним, что сумма $\sum_{j\in\mathcal{X}} p_{ij}(t)$ всегда равна единице, поэтому производная суммы равна нулю. Для консервативной цепи выполнены обратные уравнения Колмогорова.

Теорема 9.4 (обратные уравнения). *Пусть дана консервативная марковская цепь со стохастической полугруппой $(P(t), t \geqslant 0)$ и инфинитезимальной матрицей Q. Тогда для любого $t > 0$*

$$P'(t) = Q \cdot P(t),$$

или, что то же самое,

$$p'_{ij}(t) = \sum_{\alpha} q_{i\alpha} \cdot p_{\alpha j}(t).$$

Прямые уравнения требуют более жестких условий, но они и более полезны.

Теорема 9.5 (прямые уравнения). *Пусть $(P(t), t \geqslant 0)$ – стохастическая полугруппа с инфинитезимальной матрицей Q, причем все элементы Q конечны и, кроме того, для любых $i \neq j$ имеет место представление*

$$p_{ij}(h) = q_{ij} \cdot h + \alpha_{ij}(h),$$

где $\frac{\alpha_{ij}(h)}{h} \to 0$ при $h \to 0$ равномерно по всем $i \in \mathcal{X}$. Тогда для любого $t > 0$

$$P'(t) = P(t) \cdot Q,$$

или, что то же самое,

$$p'_{ij}(t) = \sum_{\alpha} p_{i\alpha}(t) \cdot q_{\alpha j}.$$

Отметим, что доказательство обеих систем дифференциальных уравнений нетривиально только для случая бесконечного фазового пространства.

▲ **Задача 9.2.** Докажите, что если фазовое пространство конечно, то для стандартной марковской цепи выполнены обе системы дифференциальных уравнений Колмогорова.

Решение

Согласно уравнениям Колмогорова–Чепмена, для любых $i, j \in \mathcal{X}$, $t, s \geq 0$

$$p_{ij}(t+s) = \sum_{\alpha \in \mathcal{X}} p_{i\alpha}(s) p_{\alpha j}(t).$$

Далее, все переходные вероятности $p_{ij}(t)$ по теореме 9.3 дифференцируемы в нуле. Раз фазовое пространство конечно, то все диагональные элементы тоже конечны. Остается продифференцировать вышеприведенное равенство либо по s в нуле (получатся обратные уравнения), либо по t в нуле (получатся прямые уравнения).

Задача решена.

Прямые уравнения Колмогорова позволяют вывести уравнения и для распределения цепи $\Pi(t)$ в произвольный момент времени (напомним, что мы понимаем $\Pi(t)$ как вектор-строку).

Лемма 9.3. *В условиях теоремы* 9.5 *выполнены уравнения* $\Pi'(t) = \Pi(t) \cdot Q$, *или, что то же самое,*

$$p_j'(t) = \sum_{\alpha \in \mathcal{X}} p_\alpha(t) q_{\alpha j}.$$

Наконец, последняя лемма позволяет найти стационарное распределение как решение одного линейного уравнения.

Лемма 9.4. *В условиях теоремы* 9.5 *и эргодической теоремы* 9.2 *выполнено равенство* $\Pi \cdot Q = 0$, *или, что то же самое, для любого* j

$$\sum_{\alpha \in \mathcal{X}} \pi_\alpha \cdot q_{\alpha j} = 0.$$

В заключение приведем пример применения последних лемм для конкретной инфинитезимальной матрицы.

▲ **Задача 9.3.** Инфинитезимальная матрица однородной марковской цепи имеет вид

$$Q = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix},$$

где $\lambda, \mu > 0$ — положительные числа. Найдите стационарное распределение и распределение цепи в момент времени t при начальном распределении $p_1(0) = 1,\ p_2(0) = 0$.

Решение

Выпишем уравнения из леммы 9.3:

$$p_1'(t) = -\lambda p_1(t) + \mu p_2(t),$$

$$p_2'(t) = -\mu p_2(t) + \lambda p_1(t).$$

Заметим, что второе уравнение несущественно, ведь $p_1(t) + p_2(t) = 1$. Отсюда получаем, что

$$p_1'(t) = -(\lambda + \mu) p_1(t) + \mu.$$

Общее решение подобного дифференциального уравнения имеет вид

$$p_1(t) = C\, e^{-(\lambda+\mu)t} + \frac{\mu}{\lambda + \mu}.$$

Из начальных условий находим, что $C = \frac{\lambda}{\lambda+\mu}$. Для нахождения стационарного распределения надо решить систему $\Pi Q = 0$ при условии, что $\pi_1 + \pi_2 = 1$. Решением, очевидно, является $\pi_1 = \frac{\mu}{\lambda+\mu}$, $\pi_2 = \frac{\lambda}{\lambda+\mu}$. Легко видеть, что именно эти два числа являются пределами для $p_1(t)$ и $p_2(t)$.

Задача решена.

▲ Задачи для самостоятельного решения

1. Докажите, что пуассоновский процесс интенсивности λ является однородной марковской цепью. Найдите его переходные вероятности, инфинитезимальную матрицу и стационарное распределение.

2. Всякий ли процесс восстановления является марковской цепью?

3. Пусть $n \times n$ матрица $Q = (q_{ij})_{i,j=1}^n$ такова, что $q_{ij} \geqslant 0$ при $i \neq j$ и $\sum_{j=1}^n q_{ij} = 0$ для любого $i = 1, \ldots, n$. Докажите, что тогда матрицы $P(t) = \exp\{tQ\}$ образуют стохастическую полугруппу.

4. Докажите, что переходные вероятности $p_{ij}(t)$ стандартной стохастической полугруппы равномерно непрерывны на $\mathbb{R}_+$.

5. Переходные вероятности марковской цепи $(X_t, t \geqslant 0)$ с фазовым пространством $\mathfrak{X} = \{1, 2, 3\}$ имеют вид:

   $p_{11}(h) = 1 - \lambda h + o(h)$, $p_{12}(h) = \lambda h + o(h)$, $p_{13}(h) = o(h)$ при $h \to 0+$;

   $p_{21}(h) = o(h)$, $p_{22}(h) = 1 - \mu h + o(h)$, $p_{23}(h) = \mu h + o(h)$ при $h \to 0+$;

   $p_{31}(h) = \nu h + o(h)$, $p_{32}(h) = o(h)$, $p_{33}(h) = 1 - \nu h + o(h)$ при $h \to 0+$.

   Докажите, что такая цепь удовлетворяет условию эргодической теоремы. Найдите ее инфинитезимальную матрицу и стационарное распределение.

6. Система “массового обслуживания” состоит из прибора и ремонтного устройства. Прибор работает случайное время, имеющее экспоненциальное распределение с параметром λ. Ремонт прибора занимает случайное время, имеющее экспоненциальное распределение с параметром μ. Обозначим

   $$p_1(t) = P(\text{прибор работает в момент времени } t),$$

   $$p_2(t) = P(\text{прибор ремонтируется в момент времени } t).$$

   Найдите $p_i(t)$, $i \in \{1, 2\}$, при условии $p_1(0) = p_2(0) = 1/2$, предполагая, что процесс образует марковскую цепь.

7. Пусть $X_1, \ldots, X_n$ — независимые неотрицательные одинаково распределенные случайные величины с функцией распределения $F(x)$. Является ли процесс $Y = (Y(x), x \geqslant 0)$,

где
$$Y(x) = \frac{1}{n}\sum_{j=1}^{n} I\{X_j \leqslant x\}$$
— это эмпирическая функция распределения, марковской цепью? Если да, то найдите ее переходные вероятности и проверьте на однородность.

10. Линейные преобразования в пространстве L^2

Еще два класса процессов, о которых мы поговорим в учебном пособии, это L^2-процессы и стационарные процессы. Для определения обоих и обсуждения их свойств нам понадобится ряд фактов о пространстве L^2 случайных величин.

▲ **Определение 10.1.** Пусть $(\Omega, \mathcal{F}, \mathsf{P})$ — вероятностное пространство. Пространством $L^2(\Omega, \mathcal{F}, \mathsf{P})$ называется пространство случайных величин на $(\Omega, \mathcal{F}, \mathsf{P})$ с конечным вторым моментом:

$$L^2(\Omega, \mathcal{F}, \mathsf{P}) = \{\xi \text{ – случайная величина на } (\Omega, \mathcal{F}, \mathsf{P}) : \mathsf{E}|\xi|^2 < +\infty\}.$$

Перечислим важнейшие свойства пространства L^2.

1. L^2 — линейное пространство: для любых $a, b \in \mathbb{R}$, $\xi, \eta \in L^2$ имеем $a\xi + b\eta \in L^2$.

2. Функция $\|\xi\| = \sqrt{\mathsf{E}\xi^2}$ обладает свойствами нормы:

   $\|\xi + \eta\| \leqslant \|\xi\| + \|\eta\|$ (неравенство треугольника),

   $\|a\xi\| = |a|\|\xi\|$.

   Если вместо случайных величин рассматривать их классы эквивалентности п. н.: $\tilde{\xi} = \{\eta \in L^2 : \xi = \eta \text{ п. н.}\}$, то $\|\cdot\|$ станет настоящей нормой, т.е. $\|\tilde{\xi}\| = 0 \;\Leftrightarrow\; \tilde{\xi} = 0$ (где 0 — это класс $\{\eta \in L^2 : 0 = \eta \text{ п.н.}\}$).

3. Сходимость в пространстве L^2: $\xi_n \xrightarrow{L^2} \xi$, если $\mathsf{E}|\xi_n - \xi|^2 \to 0$ (здесь имеется ввиду, что $\|\xi_n - \xi\|^2 = \mathsf{E}|\xi_n - \xi|^2$).

   Пространство L^2 является полным относительно такой сходимости: любая фундаментальная последовательность из L^2 имеет предел, т.е. если $\|\xi_n - \xi_m\| \to 0, \; n, m \to \infty$, то $\exists \xi \in L^2 : \xi_n \xrightarrow{L^2} \xi$.

4. Функция $\langle \xi, \eta \rangle = \mathsf{E}\xi\eta$ играет роль скалярного произведения.

   Если $\xi, \eta \in L^2$, то $\xi\eta \in L^1$ (т.е. $\mathsf{E}\xi\eta$ конечно) и выполнено неравенство Коши–Буняковского:

   $$|\langle \xi, \eta \rangle| = |\mathsf{E}\xi\eta| \leqslant \sqrt{\mathsf{E}\xi^2 \mathsf{E}\eta^2} = \|\xi\| \cdot \|\eta\|.$$

Нам особенно будет важно свойство непрерывности скалярного произведения в L^2.

Лемма 10.1 (непрерывность скалярного произведения). *Пусть* $\xi_n \xrightarrow{L^2} \xi, \eta_n \xrightarrow{L^2} \eta$ *при* $n \to \infty$, *где* $\xi_n, \eta_n, \xi, \eta \in L^2$. *Тогда*

$$\langle \xi_n, \eta_n \rangle \to \langle \xi, \eta \rangle, \quad n \to \infty.$$

В этой главе мы в основном будем изучать L^2-процессы.

▲ **Определение 10.2.** Случайный процесс $(X_t, t \in T)$ называется L^2-*процессом*, если $X_t \in L^2$ для любого $t \in T$.

Перейдем к обсуждению свойств L^2-процессов, заданных на $T = [a, b]$. Первым из них является непрерывность, среди видов которой мы выделим стохастическую непрерывность и непрерывность в среднем квадратичном. Заметим, что здесь и ниже, когда речь идет о (различных видах) непрерывности на $[a, b]$ процесса, то концы отрезка (a и b) могут быть бесконечными. Кроме того, вместо отрезка (интервала) можно рассматривать и множества (a, b), $(a, b]$, $[a, b)$. Все определения и утверждения теорем сохранятся.

▲ **Определение 10.3.** Случайный процесс $(X_t, t \in [a, b])$ называется *стохастически непрерывным* в точке $t_0 \in [a, b]$, если $X_t \xrightarrow{\mathsf{P}} X_{t_0}$ (сходится по вероятности) при $t \to t_0$.

▲ **Определение 10.4.** Случайный процесс $(X_t, t \in [a, b])$ называется *непрерывным в среднем квадратичном* (в с/к) в точке $t_0 \in [a, b]$, если $X_t \xrightarrow{L^2} X_{t_0}$ при $t \to t_0$.

Раз сходимость по вероятности слабее сходимости в L^2, то из сходимости в среднем квадратичном вытекает сходимость по вероятности. Когда случайный процесс непрерывен в среднем квадратичном? Ответ дает следующая теорема.

Теорема 10.1 (критерий непрерывности в L^2). *Пусть $(X_t, t \in [a,b])$ — L^2-процесс. Тогда X_t непрерывен в с/к в точке t_0 $\Leftrightarrow$ функция $K(s,t) = \mathsf{E}X_sX_t$ непрерывна в точке (t_0, t_0).*

Обобщением теоремы 10.1 является следующий критерий непрерывности на целом отрезке.

▲ **Задача 10.1.** Пусть $(X_t, t \in [a,b])$ — L^2-процесс. Тогда X_t непрерывен в с/к на $[a,b]$ (т.е. в любой точке отрезка) $\Leftrightarrow$ $K(s,t) = \mathsf{E}X_sX_t$ непрерывна на $[a,b] \times [a,b]$.

Решение

В обратную сторону достаточно применить теорему 10.1. Для доказательства в прямую сторону заметим, что если $s \to s_0$, $t \to t_0$, то $X_s \xrightarrow{L^2} X_{s_0}$, $X_t \xrightarrow{L^2} X_{t_0}$. В силу непрерывности скалярного произведения получаем, что $\mathsf{E}X_sX_t \to \mathsf{E}X_{s_0}X_{t_0}$, что и означает непрерывность $K(s,t)$ в произвольной точке квадрата $[a,b] \times [a,b]$.

Задача решена.

В качестве следствия можно заметить, например, что винеровский и пуассоновский процессы непрерывны в среднем квадратичном на $\mathbb{R}_+$.

Для полноты изложения сформулируем критерий стохастической непрерывности в терминах двумерной сходимости по распределению.

Теорема 10.2. *Процесс $(X_t, t \in [a,b])$ стохастически непрерывен в точке t_0 $\Leftrightarrow$ $(X_t, X_{t_0}) \xrightarrow{d} (X_{t_0}, X_{t_0})$.*

Перейдем к дифференцированию случайных процессов. Заметим, что здесь и ниже, когда речь идет о (различных видах) дифференцируемости на (a,b) процесса, то концы отрезка (a и b) могут быть бесконечными.

▲ **Определение 10.5.** Процесс $(X_t, t \in (a,b))$ называется *дифференцируемым по вероятности* в точке $t \in (a,b)$, если у выражения $\dfrac{X_{t+h} - X_t}{h}$ существует предел по вероятности при $h \to 0$. Он называется *производной по вероятности* и обозначается $(\mathsf{P})\frac{d}{dt}X_t$:

$$\frac{X_{t+h} - X_t}{h} \xrightarrow{\mathsf{P}} (\mathsf{P})\frac{d}{dt}X_t.$$

▲ **Определение 10.6.** Процесс $(X_t, t \in (a,b))$ называется *дифференцируемым в среднем квадратичном* в точке $t \in (a,b)$, если у выражения $\dfrac{X_{t+h} - X_t}{h}$ существует предел в L^2 при $h \to 0$. Он называется *производной в среднем квадратичном* и обозначается $(L^2)\frac{d}{dt}X_t$:

$$\frac{X_{t+h} - X_t}{h} \xrightarrow{L^2} (L^2)\frac{d}{dt}X_t.$$

Покажем сразу, что дифференцируемость является достаточно редким свойством.

▲ **Задача 10.2.** Докажите, что винеровский процесс $(W_t, t \geqslant 0)$ не дифференцируем ни по вероятности, ни в среднем квадратичном.

Решение

Покажем, что выражение $(W_{t+h} - W_t)/h$ не имеет даже предела по распределению. Действительно,

$$(W_{t+h} - W_t)/h \sim \mathcal{N}(0, |h|^{-1}).$$

При растущей дисперсии нормальные случайные величины не могут иметь предела по распределению (это следует, например, из метода характеристических функций). Стало быть, нет предела ни по вероятности, ни в среднем квадратичном.

Задача решена.

Когда же существует производная? Критерий дифференцируемости в одной точке формулируется достаточно громоздко, поэтому мы приведем лишь критерий непрерывной дифференцируемости на целом интервале (т.е. того, что процесс $(L^2)\frac{d}{dt}X_t$ существует и является непрерывным на (a,b)).

Теорема 10.3. *L^2-процесс $(X_t, t \in (a,b))$ является непрерывно дифференцируемым в среднем квадратичном на (a,b) $\Leftrightarrow$ функция $K(s,t) = \mathsf{E}X_sX_t$ имеет непрерывную вторую смешанную производную $\dfrac{\partial^2}{\partial s \partial t}K(s,t)$ на $(a,b)\times(a,b)$.*

Отметим, что снова критерий формулируется в терминах функции $K(s,t) = \mathsf{E}X_sX_t$, которая для процессов с нулевым средним совпадает с ковариационной.

Из свойств пространства L^2 вытекает, что L^2-производная L^2-процесса, в свою очередь, является L^2-процессом. Сформулируем лемму о свойствах L^2-производной.

Лемма 10.2. *Пусть $(X_t, t \in (a,b))$ — L^2-процесс, дифференцируемый в среднем квадратичном. Тогда (всюду ниже подразумевается производная в (L^2))*

1. *Для любой случайный величины $\xi \in L^2$ справедливо равенство $\mathsf{E}\left(\frac{d}{dt}X_t \cdot \xi\right) = \frac{d}{dt}\left(\mathsf{E}X_t\xi\right)$.*

2. $\mathsf{E}\dfrac{d}{dt}X_t = \dfrac{d}{dt}\mathsf{E}X_t.$

3. $\mathsf{E}\left(\dfrac{d}{dt}X_t \cdot \dfrac{d}{ds}X_s\right) = \dfrac{\partial^2}{\partial s\partial t}\mathsf{E}X_tX_s.$

4. $\operatorname{cov}\left(\dfrac{d}{dt}X_t, \dfrac{d}{ds}X_s\right) = \dfrac{\partial^2}{\partial s\partial t}\operatorname{cov}(X_t, X_s).$

Заметим, что для вычисления дисперсии у производной надо вычислить вторую смешанную производную ковариационной функции, а затем положить $s = t$.

Перейдем к интегрированию в среднем квадратичном. Для предела в L^2 введем специальное обозначение: если $\xi_n \xrightarrow{L^2} \xi$ и все эти случайные величины — элементы L^2, то $\xi = \text{l.i.m.}_{n\to+\infty}\xi_n$.

▲ **Определение 10.7.** Процесс $(X_t, t \in [a,b])$ называется *интегрируемым в среднем квадратичном* на $[a,b]$, если существует предел в L^2 интегральных сумм $\sum_{i=1}^n X_{s_i}(t_i - t_{i-1})$, где $T = \{a = t_0 < t_1 < \dots < t_n = b\}, s_i \in [t_{i-1}, t_i]$ – размеченное разбиение $[a,b]$, при $\Delta T = \max\limits_i(t_i - t_{i-1}) \to 0$. Подобный предел называется *интегралом от* X_t *по* $[a,b]$ и обозначается

$$(L^2)\int_a^b X_t dt := \text{l.i.m.}_{\Delta T\to 0}\sum_{i=1}^n X_{s_i}(t_i - t_{i-1}).$$

На вопрос об интегрируемости отвечает критерий, сформулированный в следующей теореме.

Теорема 10.4 *Пусть $(X_t, t \in [a,b])$ — L^2-процесс. Тогда X_t интегрируем в среднем квадратичном на $[a,b]$ $\Leftrightarrow$ функция $K(s,t) = \mathsf{E}X_sX_t$ интегрируема по Риману на $[a,b]\times[a,b]$.*

Заметим, что критерий, как и в случаях с непрерывностью и дифференцируемостью, дается в терминах условий на функцию $K(s,t) = \mathsf{E}X_sX_t$. В качестве простого следствия из теоремы получаем, что непрерывность в среднем квадратичном на отрезке влечет и интегрируемость в среднем квадратичном, ведь непрерывные функции интегрируемы по Риману. Например, винеровский и пуассоновский процессы интегрируемы на любом конечном отрезке из $\mathbb{R}_+$.

Свойства интегралов аналогичны свойствам производных. Как и в случае производных, из свойств пространства L^2 следует, что интеграл L^2-процесса принадлежит L^2.

Лемма 10.3 *Пусть $(X_t, t \in [a,b])$ — L^2-процесс, интегрируемый в среднем квадратичном. Тогда (всюду ниже подразумевается интеграл в (L^2))*

1. $\forall \xi \in L^2$ *выполнено*

$$\mathsf{E}\left(\int\limits_a^b X_t dt \cdot \xi\right) = \int\limits_a^b \mathsf{E}(X_t \cdot \xi) dt.$$

2. $\mathsf{E}\int\limits_a^b X_t dt = \int\limits_a^b \mathsf{E} X_t dt.$

3. *Пусть даны* $[c', d'] \subset [a,b]$, $[c,d] \subset [a,b]$, *тогда*

$$\mathsf{E}\left(\int\limits_c^d X_t dt \int\limits_{c'}^{d'} X_s ds\right) = \int\limits_c^d \int\limits_{c'}^{d'} (\mathsf{E} X_t X_s) ds dt.$$

4. Для $[c', d'] \subset [a,b]$, $[c,d] \subset [a,b]$ имеем

$$\operatorname{cov}\left(\int\limits_c^d X_t dt, \int\limits_{c'}^{d'} X_s ds\right) =$$

$$= \int\limits_c^d \int\limits_{c'}^{d'} \operatorname{cov}(X_t, X_s) ds dt.$$

В заключение рассмотрим пример с винеровским процессом.

▲ **Задача 10.3.** Пусть $(W_t, t \geqslant 0)$ — винеровский процесс. Найдите функцию среднего и ковариационную функцию процесса $X_t = \int_0^t W_s ds$, $t \geqslant 0$.

Решение

Применяем лемму 10.3.

$$\mathsf{E}X_t = \mathsf{E}\int_0^t W_s ds = \int_0^t \mathsf{E}W_s ds = 0.$$

$$\operatorname{cov}(X_t, X_s) = \operatorname{cov}\left(\int_0^t W_{t'}dt', \int_0^s W_{s'}ds'\right) =$$

$$= \int_0^t \int_0^s \operatorname{cov}(W_{t'}, W_{s'})dt'ds' = \int_0^t \int_0^s \min(t', s')dt'ds' =$$

(предположим $t > s$)

$$= \int_s^t \int_0^s s'dt'ds' + 2\int_0^s \int_0^{t'} s'dt'ds' =$$

$$= (t-s)\frac{s^2}{2} + \frac{s^3}{3} = \frac{ts^2}{2} - \frac{s^3}{6}.$$

Задача решена.

▲ Задачи для самостоятельного решения

1. Пусть $\xi = (\xi_t, t \in [0,1])$ — случайный процесс такой, что все ξ_t, $t \in T$, независимы в совокупности, одинаково распределены и нетривиальны (отличны от констант). Является ли процесс ξ стохастически непрерывным хоть в одной точке $[0,1]$?

2. Приведите пример процесса $(X_t, t \in [0,1])$, который является стохастически непрерывным на $[0,1]$, но не интегрируемым по вероятности на $[0,1]$ (интеграл по вероятности определяется как сходимость римановских интегральных сумм по вероятности).

   Замечание. Данный пример показывает, что сходимость по вероятности достаточно плоха, ведь из непрерывности не следует интегрируемость, в отличие от сходимости в L^2.

3. Является ли пуассоновский процесс $(N_t, t \geqslant 0)$ дифферен-

цируемым а) по вероятности, б) в среднем квадратичном?

4. Пусть $(W_t, t \geqslant 0)$ — винеровский процесс. Найдите распределение случайной величины $X_t = \int_0^t W_s ds$. Докажите, что процесс $(X_t, t \geqslant 0)$ является гауссовским.

5. Найдите для каждого $t \geq 0$ совместное распределение величин W_t и X_t из предыдущей задачи.

6. Задан случайный процесс $X_t = \int\limits_0^t e^{-W_s} ds$, где W_s — винеровский процесс. Найдите функцию среднего и ковариационную функцию процесса X_t.

7. Пусть $(W_t, t \geqslant 0)$ — винеровский процесс. Вычислите для $t > 0$ предел в L^2 при $n \to \infty$ у выражения
$$\sum_{i=1}^{n} W_{t(i-1)/n} \left(W_{ti/n} - W_{t(i-1)/n} \right).$$

11. Стационарные случайные процессы

В настоящем разделе мы начинаем изучение стационарных случайных процессов. Различают два вида стационарности: в узком и широком смыслах.

Пусть $T = \mathbb{R}, \mathbb{Z}, \mathbb{R}_+, \mathbb{Z}_+$ (в общем случае — полугруппа).

▲ **Определение 11.1.** Случайный процесс $(X_t, t \in T)$ называется *стационарным в узком смысле*, если для любых $n \in \mathbb{N}$, $t_1, \ldots, t_n, h \in T$ выполнено равенство распределений:

$$(X_{t_1+h}, \ldots, X_{t_n+h}) \stackrel{d}{=} (X_{t_1}, \ldots, X_{t_n}).$$

Смысл определения состоит в том, что процесс всегда один и тот же для любого $h \in T$, если мы начинаем изучать процесс с момента времени h, то его свойства остаются теми же самыми, как если бы мы начали смотреть на него в начальный момент времени. Процесс $Y = (Y_t, t \in T)$, где $Y_t = X_{t+h}$, имеет то же распределение, что и исходный процесс.

▲ **Определение 11.2.** Процесс $(X_t, t \in T)$ называется *стационарным в широком смысле*, если:

1. X_t — L^2-процесс;

2. $\mathsf{E}X_t = a = \mathrm{const}$ для любого $t \in T$;

3. $cov(X_t, X_s) = \mathrm{cov}(X_{t+h}, X_{s+h})$ для любых $t, s, h \in T$.

Стационарность в широком смысле означает всего лишь, что функция среднего постоянна, а ковариационная функция зависит лишь от разности аргументов $t - s$.

Приведем ряд примеров.

1. $T = \mathbb{Z}_+$ и $X_n = \xi$ для любого целого неотрицательного n — стационарный в узком смысле процесс.

2. $T = \mathbb{Z}_+, X_n = \xi_n$, где ξ_n — независимые одинаково распределенные случайные величины. Тогда X_n стационарен в узком смысле.

3. Пусть ξ, η — одинаково распределенные различные (т.е. $\mathsf{P}(\xi = \eta) < 1$) невырожденные случайные величины. Положим
$$X_n = \begin{cases} \xi, & n = 0, 1 \pmod 3, \\ \eta, & n = 2 \pmod 3. \end{cases}$$
Тогда $X_n \overset{d}{=} X_m$ для любых n, m, все элементы процесса одинаково распределены, но он не стационарен в узком смысле, так как $(X_0, X_1) = (\xi, \xi)$, $(\xi, \eta) = (X_1, X_2)$ и $(\xi, \xi) \overset{d}{\neq} (\xi, \eta)$.

Последний пример показывает, что для стационарности в узком смысле недостаточно одинаковой распределенности всех элементов процесса.

Зададимся вопросом о том, как связаны между собой два понятия стационарности. Следующая лемма показывает, что если выполнено условие принадлежности пространству L^2, то стационарность в узком смысле влечет стационарность в широком.

Лемма 11.1. *Стационарный в узком смысле L^2-процесс стационарен и в широком смысле.*

Из стационарности в широком смысле в общем случае также не вытекает стационарность в узком смысле.

▲ **Задача 11.1.** Пусть ξ_1, ξ_2 — независимые случайные величины, $\mathsf{P}(\xi_i = \pm 1) = \frac{1}{2}$. В каком смысле стационарен процесс
$$X_t = \xi_1 \sin t + \xi_2 \cos t, \ t \in \mathbb{R}?$$

Решение

Заметим, что $\mathsf{E}X_t = 0$ для любого $t \in \mathbb{R}$. Найдем ковариационную функцию:
$$\operatorname{cov}(X_t, X_s) = \operatorname{cov}(\xi_1 \sin t + \xi_2 \cos t, \xi_1 \sin s + \xi_2 \cos s) =$$
$$= \sin t \sin s \, \mathsf{D}\xi_1 + \cos t \cos s \, \mathsf{D}\xi_2 =$$
$$= \sin t \sin s + \cos t \cos s = \cos(t - s).$$

Значит, процесс X_t стационарен в широком смысле. В узком же смысле он не стационарен, так как не все его элементы одинаково распределены. Например, $X_0 = \xi_2$ принимает значения ± 1, а $X_{\frac{\pi}{4}} = \frac{1}{\sqrt{2}}(\xi_1 + \xi_2)$ — значения $0, \pm\sqrt{2}$.

Задача решена.

Как мы помним, для гауссовских процессов распределения зависят от ковариационной функции и функции среднего. Естественно ожидать, что для них понятия стационарности совпадают.

Теорема 11.1. *Пусть $(X_t, t \in T)$ — гауссовский процесс. Он стационарен в узком смысле тогда и только тогда, когда он стационарен в широком смысле.*

В теории марковских цепей слово "стационарность" употреблялось для стационарного распределения. Согласно задаче 9.1, если начальное распределение было стационарным, то распределение цепи не меняется со временем. Следующая теорема показывает, что в этой ситуации имеет место более сильное утверждение.

Теорема 11.2. *Пусть $(X_t, t \geqslant 0)$ — марковская однородная цепь с фазовым пространством $\mathfrak{X}$, стохастической полугруппой переходных вероятностей $(P(t), t \geqslant 0), P(t) = \|p_{ij}(t)\|$ и начальным распределением $\Pi = (\pi_j, j \in \mathfrak{X})$, которое является стационарным для стохастической полугруппы $P(t)$. Тогда X_t — стационарный в узком смысле случайный процесс.*

Приведем еще один пример стационарного в узком смысле процесса.

▲ **Задача 11.2.** Пусть $f(x)$ — периодическая функция с периодом $T > 0$, а случайная величина $\xi \sim U(0, T)$ имеет равномерное распределение на отрезке $[0, T]$. Докажите, что случайный процесс $X_t = f(\xi + t)$, $t \in \mathbb{R}$, является стационарным в узком смысле.

Решение

Пусть $t_1, \dots, t_m, h \in \mathbb{R}$, рассмотрим случайный вектор $(X_{t_1+h}, \dots, X_{t_m+h})$. Заметим, что случайность у нас только одна — это случайная величина ξ. Найдем характеристическую функцию этого вектора и покажем, что она не зависит от h.

$$\varphi_{t_1+h,\dots,t_m+h}(\lambda_1, \dots, \lambda_m) = \mathsf{E} e^{i \sum_{j=1}^m \lambda_j X_{t_j+h}} =$$

$$= \frac{1}{T} \int_0^T \exp\left(i \sum_{j=1}^m \lambda_j f(t_j + h + x) \right) dx =$$

$$= \frac{1}{T} \int_h^{T+h} \exp\left(i \sum_{j=1}^m \lambda_j f(t_j + y) \right) dy$$

(функция под интегралом периодическая с периодом T)

$$= \frac{1}{T} \int_0^T \exp\left(i \sum_{j=1}^m \lambda_j f(t_j + y) \right) dy =$$

$$= \varphi_{t_1,\dots,t_m}(\lambda_1, \dots, \lambda_m).$$

Значит, векторы $(X_{t_1+h}, \dots, X_{t_m+h})$ и $(X_{t_1}, \dots, X_{t_m})$ одинаково распределены, поэтому процесс стационарен в узком смысле.

Задача решена.

Перейдем к более детальному обсуждению стационарных в широком смысле процессов. Пусть $T = \mathbb{Z}$ или $\mathbb{R}$, а $(X_t, t \in T)$ — стационарный в широком смысле процесс. Тогда его ковариационную функцию $R(s, t)$ можно рассматривать как функцию одной переменной,

$$R(s, t) = R(s - t), \text{ или } R(t) = R(t, 0) = \operatorname{cov}(X_t, X_0).$$

Как мы помним, ковариационная функция любого случайного процесса (будучи функцией двух аргументов) является неотрицательно определенной. Для функции одной переменной также можно ввести понятие неотрицательной определенности, причем сразу в комплексном смысле.

▲ **Определение 11.3.** Функция $r(s,t)$, $s,t \in T$, называется *неотрицательно определенной (в комплексном смысле)* на $T \times T$, если для любых n, $t_1, \ldots, t_n \in T$, $z_1, \ldots, z_n \in \mathbb{C}$ выполнено неравенство

$$\sum_{i,j=1}^{n} r(t_i, t_j) z_i \overline{z_j} \geqslant 0.$$

▲ **Определение 11.4.** Пусть $T = \mathbb{Z}$ или $\mathbb{R}$. Тогда функция $(r(t), t \in T)$ называется *неотрицательно определенной (в комплексном смысле)*, если неотрицательно определена функция двух переменных $\tilde{r}(s,t) = r(s-t)$.

Отметим, что, в силу леммы 4.1 и задачи 3 самостоятельной работы к настоящей главе, ковариационная функция любого случайного процесса является неотрицательно определенной в комплексном смысле, а, стало быть, такова и функция $R(t)$ у нашего стационарного процесса X_t. Оказывается, неотрицательно определенные функции одной переменной имеют особое представление. Начнем мы со случая дискретного времени.

Теорема 11.3 (Герглотц). *Функция $(R(n), n \in \mathbb{Z})$ является неотрицательно определенной (в комплексном смысле) тогда и только тогда, когда существует конечная мера G на $\big([-\pi,\pi], \mathcal{B}([-\pi,\pi])\big)$ такая, что для любого $n \in \mathbb{Z}$*

$$R(n) = \int_{-\pi}^{\pi} e^{in\lambda} G(d\lambda).$$

▲ **Определение 11.5.** Мера G из теоремы Герглотца называется *спектральной мерой* для процесса X_n: для любого $n \in \mathbb{Z}$

$$\operatorname{cov}(X_n, X_0) = R(n) = \int_{-\pi}^{\pi} e^{in\lambda} G(d\lambda).$$

Если мера G имеет плотность $g(\lambda)$, то $g(\lambda)$ называется *спектральной плотностью* X_n.

Как можно найти спектральную плотность? В силу того, что

$$R(n) = \int_{-\pi}^{\pi} e^{in\lambda} g(\lambda) d\lambda,$$

то $R(n)$ — это коэффициенты Фурье для функции $g(\lambda)$. Стало быть, надо просто сложить обратно ряд Фурье:

$$g(\lambda) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} R(n) e^{-in\lambda}.$$

Приведем пример стационарного процесса, называемого *белым шумом*.

▲ **Определение 11.6.** Случайный процесс $(\varepsilon_n, n \in \mathbb{Z})$ называется *белым шумом*, если $\mathsf{E}\varepsilon_n = 0$ для любого n и $\mathsf{E}\varepsilon_n\varepsilon_m = \delta_{nm}$ для всех $n, m \in \mathbb{Z}$.

▲ **Задача 11.3.** Вычислите спектральную плотность белого шума.

Решение

Найдем коэффициенты $R(n)$:

$$R(n) = \mathsf{E}\varepsilon_n\varepsilon_0 = \delta_{n0} = I\{n = 0\}.$$

Отсюда,

$$g(\lambda) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} R(n) e^{-in\lambda} = \frac{1}{2\pi}.$$

Следовательно, $g(\lambda)$ — это плотность равномерного распределения на отрезке $[-\pi, \pi]$.

Задача решена.

В случае непрерывного времени аналогом теоремы Герглотца является теорема Бохнера–Хинчина.

Теорема 11.4 (Бохнер–Хинчин). *Пусть $(R(t), t \in \mathbb{R})$ непрерывна в нуле. Тогда $R(t)$ — неотрицательно определена (в комплексном смысле) тогда и только тогда, когда найдется конечная мера G на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ такая, что для любого $t \in \mathbb{R}$ выполнено*

$$R(t) = \int_{\mathbb{R}} e^{it\lambda} G(d\lambda).$$

▲ **Определение 11.7.** Мера G из теоремы Бохнера–Хинчина называется *спектральной* для процесса X_t:

$$cov(X_t, X_0) = R(t) = \int_{\mathbb{R}} e^{it\lambda} G(d\lambda).$$

Если G имеет плотность $g(\lambda)$, то $g(\lambda)$ называется *спектральной плотностью* случайного процесса X_t.

Как можно найти спектральную плотность? Здесь помогает теория преобразования Фурье и формула обращения из него.

Теорема 11.5 (формула обращения). *Если $\int\limits_{\mathbb{R}} |R(t)| dt < +\infty$, то*

$$g(\lambda) = \frac{1}{2\pi} \int\limits_{\mathbb{R}} e^{-it\lambda} R(t) dt.$$

▲ Задачи для самостоятельного решения

1. Пусть $N = \{N(t), t \geqslant 0\}$ — пуассоновский процесс интенсивности λ, а случайная величина η не зависит от N, причем $\mathsf{P}(\eta = 1) = \mathsf{P}(\eta = -1) = 1/2$. Является ли процесс $X_t = \eta(-1)^{N_t}$ стационарным и в каком смысле?

2. Пусть f — периодическая функция на $\mathbb{R}$ с периодом $T > 0$. Случайная величина ξ равномерно распределена на $[0, T]$. Случайный вектор (ζ, η) не зависит от ξ. Докажите, что процесс $X_t = \zeta \cdot f(\eta t + \xi)$ стационарен в узком смысле.

3. Докажите, что действительная функция $r(s,t)$ неотрицательно определена (в комплексном смысле) $\Leftrightarrow$ она неотрицательно определена (в действительном смысле) и симметрична.

4. Пусть $W_t^{(1)}$ и $W_t^{(2)}$ — независимые винеровские процессы. Для любого $t \in \mathbb{R}$ положим
$$X_t = W_t^{(1)} I\{t \geqslant 0\} + W_{-t}^{(2)} I\{t < 0\}.$$
Докажите, что процесс $Y_t = \frac{1}{h}\left(X_t - X_{t-h}\right)$ является стационарным в широком смысле. Найдите его ковариационную функцию и спектральную плотность.

5. Пусть $\{\varepsilon_n, n \in \mathbb{Z}\}$ — белый шум. Положим
$$X_n = \frac{1}{2}\,\varepsilon_{n-1} + \frac{1}{4}\,\varepsilon_{n-2},\ n \in \mathbb{Z}.$$
Вычислите спектральную плотность процесса X_n.

6. Являются ли следующие функции ковариационными функциями некоторых гауссовских процессов $(X_t, t \in \mathbb{R})$:

   а) $r(s,t) = e^{-\alpha|s-t|}$, $\alpha > 0$;

   б) $r(s,t) = 2\cos(t-s)$;

   в) $r(s,t) = (1-(t-s)^2) I\{|s-t| \leq 1\}$;

   г) $r(s,t) = \exp\{ia(s-t) - \frac{(t-s)^2\sigma^2}{2}\}$?

   Для тех процессов, которые существуют, выясните, существуют ли у них спектральные плотности.

7. Пусть $(X_t, t \geq 0)$ — непрерывный в среднем квадратичном стационарный в широком смысле процесс. Докажите, что если у него существует производная в L^2, то она является стационарным в широком смысле процессом. Более того, если $\mathsf{E}X_t \neq 0$, то ни для какой случайной величины ξ процесс $(\xi + \int_0^t X_s ds, t \geq 0)$ не является стационарным в широком смысле.

12. Спектральное представление

Настоящая глава посвящена спектральному представлению стационарных в широком смысле процессов. Это представление помогает работать со стационарными в широком смысле процессами, так как оно позволяет, в некотором смысле, разделить параметр времени и случайность. Начнем мы с понятия *ортогональной случайной меры.*

▲ **Определение 12.1.** Пусть $(X_t, t \in T)$ и $(Y_t, t \in T)$ — два (действительных) случайных процесса, тогда $Z_t = X_t + iY_t$ называется *комплексным случайным процессом.* В этом случае скалярное произведение в L^2 и ковариационная функция определяются немного по-другому:

$$\langle Z_s, Z_t \rangle = \mathsf{E} Z_s \overline{Z_t},$$

$$R(s,t) = \operatorname{cov}(X_s, X_t) = \mathsf{E}(Z_s - \mathsf{E}Z_s)\overline{(Z_t - \mathsf{E}Z_t)}$$

Пусть $(\Omega, \mathcal{F}, \mathsf{P})$ — вероятностное пространство, и задано $L^2(\Omega) = L^2(\Omega, \mathcal{F}, \mathsf{P})$ — пространство комплексных случайных величин $\xi : \mathsf{E}|\xi|^2 < +\infty$, а $\mathcal{K}$ — некоторое полукольцо подмножеств (определение полукольца см., например, в [2]) на множестве Λ.

▲ **Определение 12.2.** Отображение $Z : \mathcal{K} \to L^2(\Omega, \mathcal{F}, \mathsf{P})$ называется *ортогональной случайной мерой на* $\mathcal{K}$, если выполнены следующие условия.

1. Если $A, B \in \mathcal{K}$ не пересекаются, то

$$\mathsf{E} Z(A)\overline{Z(B)} = 0$$

(свойство ортогональности).

2. Если $B = \bigsqcup\limits_{i=1}^{\infty} B_i$, $B, B_i \in \mathcal{K}$, то

$$Z(B) = \sum_{i=1}^{\infty} Z(B_i) \text{ п. н.},$$

где сходимость ряда понимается как сходимость в L^2. Ортогональная случайная мера Z называется *центрированной*, если $\mathsf{E}Z(B) = 0$ для любого $B \in \mathcal{K}$.

Ортогональная случайная мера сопоставляет каждому элементу $\mathcal{K}$ случайную величину. С ней можно также связать обычную меру, называемую *структурной*.

▲ **Определение 12.3.** Функция μ на $\mathcal{K}$, определенная по правилу $\mu(B) = \mathsf{E}|Z(B)|^2$, называется *структурной мерой* для ортогональной случайной меры Z.

▲ **Задача 12.1.** Докажите, что структурная мера действительно является мерой на $\mathcal{K}$.

Решение

Покажем, что μ счетно-аддитивна на $\mathcal{K}$. Пусть $B = \bigsqcup\limits_{i=1}^{\infty} B_i$, $B, B_i \in \mathcal{K}$. Тогда для любого натурального числа N, пользуясь аддитивностью и ортогональностью Z, получаем следующее равенство:

$$\mathsf{E}\left|Z(B) - \sum_{i=1}^{N} Z(B_i)\right|^2 = \mathsf{E}|Z(B)|^2 + \sum_{i=1}^{N} \mathsf{E}\left|Z(B_i)\right|^2 -$$

$$- \sum_{i=1}^{N} \mathsf{E}\left(Z(B)\overline{Z(B_i)} + Z(B_i)\overline{Z(B)}\right) =$$

$$= \mathsf{E}|Z(B)|^2 + \sum_{i=1}^{N} \mathsf{E}\left|Z(B_i)\right|^2 - 2\sum_{i=1}^{N} \mathsf{E}\left|Z(B_i)\right|^2 =$$

$$= \mu(B) - \sum_{i=1}^{N} \mu(B_i).$$

Левая часть равенства стремится к нулю, поэтому и правая также стремится к нулю, что и означает счетную аддитивность μ.

Задача решена.

Отметим еще одно важное свойство структурной меры.

▲ **Задача 12.2.** Докажите, что для любых $A, B \in \mathcal{K}$ выполняется равенство $\mathsf{E}Z(A)\overline{Z(B)} = \mu(A \cap B)$.

Решение

По свойству полукольца, A и B имеют следующее представление:

$$A = (A \cap B) \sqcup \bigsqcup_{j=1}^{m} C_j, \ B = (A \cap B) \sqcup \bigsqcup_{i=1}^{n} D_i,$$

где C_j и D_i — непересекающиеся элементы $\mathcal{K}$. Тогда в силу ортогональности и аддитивности Z,

$$\mathsf{E} Z(A)\overline{Z(B)} = \mathsf{E}|Z(A \cap B)|^2 = \mu(A \cap B).$$

Задача решена.

Как показывает задача 1 самостоятельной работы к настоящей главе, указанное свойство является и достаточным для того, чтобы отображение Z было ортогональной случайной мерой. Приведем пример применения данного принципа для построения ортогональной случайной меры на полукольце $\mathcal{K}_{\mathbb{R}_+}$ полуинтервалов на $\mathbb{R}_+$.

▲ **Задача 12.3.** Пусть $(W_t, t \geqslant 0)$ — винеровский процесс. Докажите, что отображение $Z((a,b]) = W_b - W_a$ является ортогональной случайной мерой на $\mathcal{K}_{\mathbb{R}_+} = \{(a,b] : 0 \leqslant a < b < +\infty\}$.

Решение

Найдем скалярное произведение $W_t - W_s$ и $W_b - W_a$ при $t > s$, $b > a$. В силу независимости приращений и центрированности винеровского процесса, оно равно нулю, если $(s,t]$ и $(a,b]$ не пересекаются. Если же, например, $s < a < t < b$, то

$$\mathsf{E}(W_t - W_s)(W_b - W_a) = \mathsf{E}(W_t - W_a)^2 = t - a.$$

Аналогично проверяется, что в общем случае

$$\mathsf{E}(W_t - W_s)(W_b - W_a) = \mathrm{mes}((s,t] \cap (a,b]),$$

где mes — это мера Лебега. Остается применить задачу 1 самостоятельной работы к настоящей главе.

Задача решена.

Данная задача является иллюстрацией теоремы (которую мы сформулируем ниже) о том, что все ортогональные случайные меры на $\mathcal{K}_{\mathbb{R}_+}$ получаются подобным образом с помощью процессов с ортогональными приращениями.

▲ **Определение 12.4.** Процесс $(X_t, t \in T)$, $T \subset \mathbb{R}$, называется *процессом с ортогональными приращениями*, если $\forall s < u < t \in T : \mathsf{E}(X_t - X_u)\overline{X_s} = 0$. Процесс $(X_t, t \in T)$ называется *центрированным*, если $\mathsf{E}X_t = 0$ для любого $t \in T$.

Теорема 12.1. *1. Пусть Z — ортогональная случайная мера на $\mathcal{K}_{\mathbb{R}_+}$. Тогда процесс $(X_t, t \geqslant 0)$, заданный по формуле $X_t = Z(0, t]$, является процессом с ортогональными приращениями, непрерывным справа в среднем квадратичном.*
2. Пусть $(X_t, t \geqslant 0)$ — L^2-процесс с ортогональными приращениями, непрерывный справа в среднем квадратичном. Тогда отображение $Z : \mathcal{K}_{\mathbb{R}_+} \to L^2(\Omega)$, где $Z(a, b] = X_b - X_a$, является ортогональной случайной мерой на $\mathcal{K}_{\mathbb{R}_+}$.

Как и по обычной мере, по ортогональной случайной мере можно построить интеграл, называемый *стохастическим*. Это понятие нам понадобится для определения спектрального представления стационарных в широком смысле процессов.

Пусть Z — ортогональная случайная мера на полукольце $\mathcal{K}$ подмножеств Λ, причем $\Lambda \in \mathcal{K}$. Пусть μ — структурная мера, $\mu(\Lambda) > 0$.

По теореме о продолжении меры, μ можно продолжить до меры на $\sigma(\mathcal{K})$ — минимальной сигма-алгебре, содержащей $\mathcal{K}$. Мера же Z, как показывает следующая лемма, продолжается на $\alpha(\mathcal{K})$ — минимальную алгебру, содержащую полукольцо $\mathcal{K}$.

Лемма 12.1. *Пусть $A \in \alpha(\mathcal{K})$ и $A = \bigsqcup_{k=1}^{n} C_k$, $C_k \in \mathcal{K}$. Положим $Z(A)$ как $Z(A) = \sum_{k=1}^{n} Z(C_k)$. Тогда подобное продолжение корректно и дает ортогональную случайную меру на $\alpha(\mathcal{K})$.*

Опишем построение стохастического интеграла. Как и в случае интеграла Лебега, сначала определяется интеграл от простых функций.

▲ **Определение 12.5.** Функция $f : \Lambda \to \mathbb{C}$ называется *простой*, если f представима в виде

$$f(\lambda) = \sum_{k=1}^{n} c_k \mathrm{I}_{B_k}(\lambda),$$

где $c_k \in \mathbb{C},\ B_k \in \alpha(\mathcal{K})$ и $B_1, \ldots, B_n$ — разбиение Λ.

- *Стохастическим интегралом* от простой функции f, заданной равенством $f(\lambda) = \sum_{k=1}^{n} c_k \mathrm{I}_{B_k}(\lambda)$, по ортогональной случайной мере Z называется

$$J(f) = \sum_{k=1}^{n} c_k Z(B_k).$$

- Пусть f — произвольная функция, $f \in L^2(\Lambda)$, где $L^2(\Lambda) = L^2(\Lambda, \sigma(\mathcal{K}), \mu)$. Пусть $f_n \xrightarrow{L^2(\Lambda)} f$ — последовательность простых функций. Тогда *стохастическим интегралом* от функции f по ортогональной случайной мере Z называется

$$J(f) = \mathrm{l.i.m.}_{n\to\infty} J(f_n),$$

т.е. $\mathsf{E}|J(f) - J(f_n)|^2 \to 0$.

Следующая теорема подтверждает корректность определения.

Теорема 12.2. *Стохастический интеграл существует для всех $f \in L^2(\Lambda)$, и его определение корректно.*

Каковы же полезные свойства данного интеграла? Выделим важнейшие из них.

1. **Линейность**. Для любых $\alpha, \beta \in \mathbb{C},\ f, g \in L^2(\Lambda)$ выполнено

$$J(\alpha f + \beta g) = \alpha J(f) + \beta J(g).$$

2. **Изометричность**. Для любых $f, g \in L^2(\Lambda)$ выполнено

$$\langle J(f), J(g) \rangle_{L^2(\Omega)} = \mathsf{E} J(f)\overline{J(g)} = \int_\Lambda f\overline{g} d\mu = \langle f, g \rangle_{L^2(\Lambda)}.$$

3. **Согласованность с** Z. Для любого $B \in \alpha(\mathcal{K})$

$$J(\mathrm{I}_B) = Z(B).$$

Сформулируем итоговую теорему.

Теорема 12.3. *Стохастический интеграл по ортогональной случайной мере Z является линейным изометрическим биективным отображением между $L^2(\Lambda)$ и некоторым линейным подпространством $L^2_Z \subset L^2(\Omega)$. При этом Z продолжается до ортогональной случайной меры на $\sigma(\mathcal{K})$ по правилу $Z(A) = J(\mathrm{I}_A),\ A \in \sigma(\mathcal{K})$.*

Для стохастического интеграла из теоремы 12.3 принято следующее обозначение:

$$J(f) = \int_\Lambda f(\lambda) Z(d\lambda).$$

Отметим, что при $\Lambda \notin \mathcal{K}$, но $\Lambda = \bigsqcup_{n=1}^\infty \Lambda_n$, где $\Lambda_n \in \mathcal{K}$, построение стохастического интеграл проходит полностью аналогично. В этом случае структурную меру μ можно продолжить до конечной или σ-конечной меры на $\sigma(\mathcal{K})$.

Построение интеграла для $f \in L^2(\Lambda, \sigma(\mathcal{K}), \mu)$ происходит в два этапа:

1. определяем стохастический интеграл $J_n(f_n)$, где $f_n = f\Big|_{\Lambda_n}$, для всех n;

2. для самой f полагаем $J(f) = \sum_{n=1}^\infty J_n(f_n)$, где ряд понимается как сходящийся в $L^2(\Omega)$.

Свойства остаются теми же самыми, а Z продолжается до ортогональной случайной меры на $\mathcal{G} = \{A \in \sigma(\mathcal{K}) : \mu(A) < \infty\}$ по правилу $Z(A) = J(\mathrm{I}_A)$.

Перейдем к обсуждению связи стохастического интегрирования и стационарных в широком смысле случайных процессов. Ключевую роль здесь играет теорема Карунена.

Теорема 12.4 (Карунен). *Пусть $(X_t, t \in T)$ — центрированный L^2-процесс, причем его ковариационная функция допускает* факторизацию, *т.е. существует измеримое пространство $(\Lambda, \mathcal{A})$ и конечная мера μ на нем такая, что для любых $t, s \in T$ выполнено равенство*

$$\mathrm{cov}(X_t, X_s) = \int_\Lambda f_t(\lambda)\overline{f_s(\lambda)}\mu(d\lambda),$$

где $\{f_t(\lambda), t \in T\}$ — некоторый набор функций из $L^2(\Lambda, \mathcal{A}, \mu)$.

Если система $\{f_t(\lambda),\ t \in T\}$ — полная в $L^2(\Lambda, \mathcal{A}, \mu)$, то существует центрированная ортогональная случайная мера Z на $\mathcal{A}$ такая, что для любого $t \in T$ выполняется равенство

$$X_t = \int_\Lambda f(t, \lambda) Z(d\lambda). \tag{5}$$

При этом μ будет являться структурной мерой для Z.

Представление вида (5) принято называть *спектральным.* Его польза состоит в том, что зависимости от времени (t) и случайности (Z) разделены, случайность едина для всех, со временем меняется лишь некоторая неслучайная функция. Теорема Карунена говорит о том, что оно имеет место при факторизации ковариационной функции. Для стационарных случайных процессов подобную факторизацию с $f_t = e^{it\lambda}$ обеспечивают теорема Герглотца и теорема Бохнера–Хинчина из предыдущей главы. В случае дискретного времени теорема о спектральном представлении звучит следующим образом.

Теорема 12.5 (о спектральном представлении). *Пусть $(X_n, n \in \mathbb{Z})$ — центрированный стационарный в широком смысле случайный процесс. Тогда существует центрированная ортогональная случайная мера Z на $\mathcal{B}([-\pi,\pi])$ такая, что для любого $n \in \mathbb{Z}$ выполнено*

$$X_n = \int_{-\pi}^{\pi} e^{in\lambda} Z(d\lambda).$$

Приведем пример применения спектрального представления.

▲ **Задача 12.4.** Пусть $(X_n, n \in \mathbb{Z})$ — центрированный стационарный в широком смысле случайный процесс, а

$$X_n = \int_{-\pi}^{\pi} e^{in\lambda} Z(d\lambda)$$

— его спектральное представление. Докажите, что

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \xrightarrow{L^2} Z(\{0\}) \text{ при } n \to \infty.$$

Решение

В силу спектрального представления, имеем

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \overset{L^2}{=} \int_{-\pi}^{\pi} \left(\frac{1}{n}\sum_{k=0}^{n-1} e^{ik\lambda}\right) Z(d\lambda).$$

Далее, сумма под интегралом равна 1 при $\lambda = 0$, иначе

$$\frac{1}{n}\sum_{k=0}^{n-1} e^{ik\lambda} = \frac{1 - e^{in\lambda}}{n(1 - e^{i\lambda})} \to 0, \quad n \to \infty.$$

Последнее выражение ограничено по модулю единицей, поэтому по теореме Лебега

$$\left(\frac{1}{n}\sum_{k=0}^{n-1} e^{ik\lambda}\right) I\{\lambda \neq 0\} \overset{L^2(\Lambda)}{\longrightarrow} 0, \quad n \to \infty,$$

где $\Lambda = [-\pi, \pi]$, $\sigma(\mathcal{K}) = \mathcal{B}([-\pi, \pi])$, μ — структурная мера для Z. В силу изометрии стохастического интеграла, мы получаем, что

$$\int_{-\pi}^{\pi} \left(\frac{1}{n}\sum_{k=0}^{n-1} e^{ik\lambda}\right) I\{\lambda \neq 0\} Z(d\lambda) \overset{L^2(\Omega)}{\longrightarrow} 0, \quad n \to \infty.$$

Значит,

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \overset{L^2}{=} \int_{-\pi}^{\pi} \left(\frac{1}{n}\sum_{k=0}^{n-1} e^{ik\lambda}\right) I\{\lambda \neq 0\} Z(d\lambda) + \\ + \int_{-\pi}^{\pi} I\{\lambda = 0\} Z(d\lambda) \overset{L^2(\Omega)}{\longrightarrow} Z(\{0\})$$

в силу того, что второй интеграл просто равен $Z(\{0\})$.

Задача решена.

В случае непрерывного времени для получения спектрального представления необходимо дополнительное условие непрерывности в среднем квадратичном.

Теорема 12.6 (о спектральном представлении). *Пусть $(X_t, t \in \mathbb{R})$ — центрированный стационарный в широком смысле непрерывный в среднем квадратичном случайный процесс. Тогда существует центрированная ортогональная случайная мера Z на $\mathcal{B}(\mathbb{R})$ такая, что для любого $t \in \mathbb{R}$ выполнено*

$$X_t = \int_{\mathbb{R}} e^{it\lambda} Z(d\lambda).$$

Спектральное представление наиболее ярко проявляет себя в случае линейных преобразований случайных процессов.

▲ **Задача 12.5.** Пусть $(X_t, t \in \mathbb{R})$ — центрированный стационарный в широком смысле случайный процесс, а $X_t = \int_{\mathbb{R}} e^{it\lambda} Z(d\lambda)$ — его спектральное представление. Найдите спектральное представление процесса $Y_t = (L^2)\frac{d}{dt}X_t = \frac{d}{dt}X_t$ (Y_t является стационарным в широком смысле — см. задачу 7 самостоятельной работы предыдущего раздела).

Решение

Естественно ожидать, что нам надо просто продифференцировать по t под знаком интеграла. Обоснуем подобный переход. По определению производной,

$$Y_t = \frac{d}{dt}X_t = \text{l.i.m.}_{n\to 0}\frac{X_{t+h} - X_t}{h} =$$

$$= \text{l.i.m.}_{h\to 0}\int_{\mathbb{R}} \frac{e^{i(t+h)\lambda} - e^{it\lambda}}{h} Z(d\lambda).$$

В силу изометрии стохастического интеграла, можно переставить предел в $L^2(\Omega)$ и интеграл, при этом мы получим интеграл от предела подинтегральной функции в $L^2(\mathbb{R}) = L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. Он имеет вид

$$\text{l.i.m.}_{h\to 0}\frac{e^{i(t+h)\lambda} - e^{it\lambda}}{h}.$$

Но при фиксированном t этот предел равен $(i\lambda)e^{it\lambda}$, что соответствует пределу почти всюду в $L^2(\mathbb{R})$. Пределы почти всюду и в среднем квадратичном обязаны совпадать (почти всюду), если они оба существуют (так как из обеих сходимостей следует сходимость по мере). Следовательно,

$$Y_t = \int_{\mathbb{R}} (i\lambda)e^{it\lambda} Z(d\lambda).$$

Задача решена.

Покажем теперь, что происходит со спектральной плотностью при взятии производных.

▲ **Задача 12.6.** Пусть в условии предыдущей задачи процесс

$(X_t, t \in \mathbb{R})$ имеет спектральную плотность $g(\lambda)$. Найдите спектральную плотность процесса $Y_t = \frac{d}{dt} X_t$.

Решение

Пусть спектральное представление Y_t имеет вид

$$Y_t = \int_{\mathbb{R}} e^{it\lambda} Z_1(d\lambda).$$

По предыдущей задаче $Z_1(d\lambda) = (i\lambda) Z(d\lambda)$. Тогда

$$\begin{aligned}
\operatorname{cov}(Y_t, Y_0) &= \mathsf{E} \int_{\mathbb{R}} e^{it\lambda} Z_1(d\lambda) \cdot \overline{\int_{\mathbb{R}} Z_1(d\lambda)} = \\
&= \mathsf{E} \int_{\mathbb{R}} e^{it\lambda} (i\lambda) Z(d\lambda) \cdot \overline{\int_{\mathbb{R}} (i\lambda) Z(d\lambda)} = \\
&= \int_{\mathbb{R}} e^{it\lambda} (i\lambda) \overline{(i\lambda)} \mu(d\lambda) = \\
&= \int_{\mathbb{R}} e^{it\lambda} (i\lambda) \overline{(i\lambda)} G(d\lambda) = \int_{\mathbb{R}} e^{it\lambda} \lambda^2 g(\lambda) d\lambda.
\end{aligned}$$

Значит, спектральная плотность Y_t равна $\lambda^2 g(\lambda)$. Здесь через μ мы обозначили структурную меру для Z, а через G — спектральную меру для X_t, а также воспользовались тем, что они совпадают (см. лемму 12.2).

Задача решена.

Последнее утверждение, которое использовалось в решении задачи, столь важно, что стоит выделить его в отдельную лемму.

Лемма 12.2. *В рамках теорем о спектральном представлении (теоремы 12.5 и 12.6), спектральная мера процесса X_t совпадает со структурной мерой ортогональной случайной меры Z.*

В заключение отметим, что, как и дифференцирование, интегрирование в L^2 позволяет себя переставлять со стохастическим интегралом.

▲ Задачи для самостоятельного решения

1. Пусть Λ — множество, $\mathcal{A}$ — алгебра его подмножеств, а μ — мера на $\mathcal{A}$. Пусть отображение $Z : \mathcal{A} \to L^2(\Omega, \mathcal{F}, \mathsf{P})$ удовлетворяет равенству

$$\mathsf{E}Z(B)\overline{Z(C)} = \mu(B \cap C) \text{ для любых } B, C \in \mathcal{A}.$$

Докажите, что Z есть ортогональная мера на $\mathcal{A}$.

2. Пусть $\lambda_1, \ldots, \lambda_k$ — числа из отрезка $[-\pi, \pi]$, а $\Phi_1, \ldots, \Phi_k$ — центрированные попарно некоррелированные случайные величины. Докажите, что процесс $(X_n, n \in \mathbb{Z})$, где

$$X_n = \sum_{j=1}^{k} e^{i\lambda_j n}\Phi_j,$$

является стационарным в широком смысле, и найдите его спектральное представление.

3. Пусть $\{X_n, n \in \mathbb{Z}\}$ — стационарная в широком смысле последовательность со средним a и ковариационной функцией $R(n)$. Докажите, что

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{L^2} a$$

тогда и только тогда, когда

$$\frac{1}{n}\sum_{k=1}^{n} R(k) \longrightarrow 0.$$

4. Пусть центрированный стационарный в широком смысле процесс $(X_n, n \in \mathbb{Z})$ удовлетворяет равенству $X_0 = X_N$ п. н. для некоторого $N \in \mathbb{N}$. Докажите, что тогда для всех $n \in \mathbb{Z}$ X_n имеет вид

$$X_n = \sum_{k=-\lfloor N/2 \rfloor}^{\lfloor (N-1)/2 \rfloor} e^{i\frac{2\pi k}{N}n}\Phi_k,$$

где Φ_k — центрированные попарно некоррелированные случайные величины из L^2.

5. Пусть $P(x) = a_0+a_1x+\ldots+a_nx^n$ — многочлен, а $P\left(\frac{d}{dt}\right)$ — соответствующий оператор дифференцирования в L^2. Пусть $(\xi_t, t \in \mathbb{R})$ — стационарный в широком смысле процесс с известным спектральным представлением. Стационарный в широком смысле процесс $(X_t, t \in \mathbb{R})$ имеет спектральное представление и, кроме того, удовлетворяет уравнению

$$P\left(\frac{d}{dt}\right) X_t = \xi_t.$$

Найдите спектральное представление для X_t. При каких условиях на многочлен P решение уравнения единственно?

6. Стационарный процесс $(Y_t, t \in \mathbb{R})$ удовлетворяет равенству $dY_t/dt = X_t$, где стационарный процесс $(X_t, t \in \mathbb{R})$ имеет спектральную плотность $f(\lambda) = \lambda^2 I\{|\lambda| < 1\}$. Найдите $\operatorname{cov}(Y_1, Y_0)$.

7. Случайный процесс $(X_t, t \in \mathbb{R})$ является центрированным и стационарным в широком смысле. Его спектральная плотность равна $f(\lambda) = |\lambda|\, I_{[-2,2]}(\lambda)$. Используя спектральное представление, найдите спектральную плотность процесса Y_t, удовлетворяющего уравнению $\frac{d^2}{dt^2} Y_t + 5Y_t = X_t$. Вычислите $\mathsf{D}Y_1$.

8. Пусть $(X_t, t \in \mathbb{R})$ — стационарный в широком смысле процесс, а

$$X_t = m + \int\limits_{\mathbb{R}} e^{it\lambda} Z(d\lambda),$$

— его спектральное представление. Докажите, что

$$(L^2) \lim_{b-a\to+\infty} \frac{1}{b-a} \left(\int_a^b X_t dt \right) = m + Z(\{0\}).$$

13. Элементы стохастического исчисления Ито

Последняя глава посвящена введению в одну из самых важных областей теории случайных процессов — стохастическое исчисление Ито.

В основе построения интеграла Ито лежит понятие винеровского процесса относительно фильтрации.

▲ **Определение 13.1.** Пусть $\mathbb{F} = (\mathcal{F}_t, t \geqslant 0)$ — фильтрация на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P})$. Случайный процесс $(W_t, t \geqslant 0)$ называется *винеровским относительно* $\mathbb{F}$, если

1. $W_0 = 0$ п.н.;
2. W_t согласован с $\mathbb{F}$;
3. $W_t - W_s$ не зависит от $\mathcal{F}_s$ для любых $t > s \geqslant 0$;
4. $W_t - W_s \sim \mathcal{N}(0, t-s)$;
5. W_t имеет непрерывные траектории.

Пусть $(W_t, t \geqslant 0)$ — винеровский процесс относительно фильтрации $\mathbb{F} = (\mathcal{F}_t, t \geqslant 0)$. Рассмотрим следующий набор подмножеств

$$\mathcal{K} = \{(s,t] \times B : \ 0 \leqslant s < t, \ B \in \mathcal{F}_s\}.$$

Легко проверить, что $\mathcal{K}$ является полукольцом на $(0, +\infty) \times \Omega$. Рассмотрим следующее отображение $\xi : \mathcal{K} \to L^2(\Omega, \mathcal{F}, \mathsf{P})$:

$$\xi((s,t] \times B) = I_B(W_t - W_s) \text{ при } 0 \leqslant s < t, \ B \in \mathcal{F}_s.$$

Следующая лемма связывает нас со стохастическим интегралом из предыдущего раздела.

Лемма 13.1. *Отображение ξ является центрированной ортогональной случайной мерой на $\mathcal{K}$ со структурной мерой $\mu = \mathrm{mes} \times \mathsf{P}$, где mes — это мера Лебега на $(0, +\infty)$.*

В предыдущем разделе нами был построен стохастический интеграл по ортогональной случайной мере. Тогда и по ξ такой интеграл существует, и именно он и называется *интегралом Ито.*

▲ **Определение 13.2.** Для любой функции $f \in L^2((0,+\infty)\times \times\Omega, \sigma(\mathcal{K}), \mathrm{mes}\times \mathsf{P})$ определен стохастический интеграл по ортогональной случайной мере ξ, называемый *стохастическим интегралом Ито.* Для него используется специальное обозначение

$$\mathcal{I}(f) = \int_0^{+\infty} f dW_t := \int\limits_{(0,+\infty)\times\Omega} f(t,\omega)\xi(dt,d\omega).$$

▲ **Определение 13.3.** Сигма-алгебра $\sigma(\mathcal{K})$ называется *предсказуемой сигма-алгеброй* на $(0,+\infty)\times\Omega$ и обозначается $\mathcal{PRED}$. Случайные процессы $f = (f(t), t > 0)$, измеримые (как отображение $f : (0,+\infty)\times\Omega \to \mathbb{R}$) относительно $\mathcal{PRED}$, называются *предсказуемыми.*

Заметим, что в интеграле Ито, как в стохастическом интеграле по ортогональной случайной мере, пространство элементарных исходов является частью $\Lambda = (0,+\infty)\times\Omega$, тем самым, мы интегрируем случайный процесс по случайной мере.

Базовые свойства интеграла Ито являются теми же, что и у интеграла по ортогональной случайной мере.

1. **Линейность**. Для любых $\alpha,\beta \in \mathbb{R}$ и функций $f, g \in L := L^2((0,+\infty)\times\Omega, \mathcal{PRED}, \mathrm{mes}\times\mathsf{P})$ выполнено
$$\mathcal{I}(\alpha f + \beta g) = \alpha\mathcal{I}(f) + \beta\mathcal{I}(g).$$

2. **Изометричность**. Для любых $f, g \in L$ выполнено
$$\langle \mathcal{I}(f), \mathcal{I}(g)\rangle_{L^2(\Omega,\mathcal{F},\mathsf{P})} = \langle f, g\rangle_L, \text{ т.е.}$$
$$\mathsf{E}\mathcal{I}(f)\mathcal{I}(g) = \int_0^{\infty} (\mathsf{E} f(t,\omega)g(t,\omega))\, dt.$$

3. **Центрированность**. Для любого $f \in L$ выполнено
$$\mathsf{E}\mathcal{I}(f) = 0.$$

4. Для любых $0 \leqslant s < t$, $B \in \mathcal{F}_s$ имеет место равенство

$$\mathcal{I}(I\{(s,t] \times B\}) = I_B(W_t - W_s).$$

Рассмотрим пример вычисления интеграла Ито чуть более сложных функций.

▲ **Задача 13.1.** Пусть $0 = t_0 < t_1 < \ldots < t_n$ и даны случайные величины $f_1, \ldots, f_n$, причем f_i является $\mathcal{F}_{t_{i-1}}$-измеримой для всех $i = 1, \ldots, n$. Докажите, что тогда для ступенчатого процесса $f(t,\omega) = \sum_{i=1}^n f_i(\omega) I\{t_{i-1} < t \leqslant t_i\}$ выполнено

$$\mathcal{I}(f) = \sum_{i=1}^n f_i(W_{t_i} - W_{t_{i-1}}).$$

Решение

В силу линейности, достаточно доказать, что для любого $i = 1, \ldots, n$

$$\mathcal{I}\left(f_i(\omega) I\{t_{i-1} < t \leqslant t_i\}\right) = f_i(W_{t_i} - W_{t_{i-1}}).$$

Заметим, что, в силу свойства 4, формула верна при $f_i = I_B$ для $B \in \mathcal{F}_{t_{i-1}}$. Рассмотрим последовательность простых $\mathcal{F}_{t_{i-1}}$-измеримых функций f_i^n, сходящихся в $L^2(\Omega)$ к f_i. Тогда

$$\mathcal{I}\left(f_i^n(\omega) I\{t_{i-1} < t \leqslant t_i\}\right) = f_i^n(W_{t_i} - W_{t_{i-1}}).$$

Далее, последовательность $f_i^n(\omega) I\{t_{i-1} < t \leqslant t_i\}$ сходится в L к $f_i(\omega) I\{t_{i-1} < t \leqslant t_i\}$. Осталось показать, что правая часть последнего равенства сходится к $f_i(W_{t_i} - W_{t_{i-1}})$ в $L^2(\Omega)$. Здесь достаточно воспользоваться независимостью $W_{t_i} - W_{t_{i-1}}$ и $\mathcal{F}_{t_{i-1}}$:

$$\mathsf{E}\left((f_i^n - f_i)(W_{t_i} - W_{t_{i-1}})\right)^2 =$$

$$= \mathsf{E}(f_i^n - f_i)^2 \mathsf{E}\left(W_{t_i} - W_{t_{i-1}}\right)^2 \to 0.$$

Задача решена.

Для предсказуемых процессов можно рассмотреть интеграл Ито по отрезку $[0,t]$ и получить случайный процесс как функцию верхнего предела. Пусть $f = (f(t,\omega), t > 0)$ — предсказуемый процесс. Тогда положим для $t > 0$

$$\mathcal{I}_t(f(s,\omega)) = \int_0^t f(s,\omega)dW_s := \mathcal{I}(f(s,\omega)I\{s \leqslant t\}).$$

Как показывает следующая теорема, данный процесс обладает весьма интересными свойствами.

Теорема 13.1. *Пусть $f = (f(t,\omega), t > 0)$ — предсказуемый процесс. Тогда процесс $(\mathcal{I}_t(f), t > 0)$ является мартингалом относительно $\mathbb{F}$, имеющим непрерывные траектории с вероятностью 1.*

Подобные замечательные свойства процесса $(\mathcal{I}_t(f), t > 0)$ позволили построить аналог классического дифференциального и интегрального исчислений. Однако прежде чем переходить к стохастическому исчислению, необходимо прояснить ответы на два вопроса: какие процессы являются предсказуемыми и каковы могут быть способы вычисления интеграла Ито на отрезке. Здесь важны две леммы.

Лемма 13.2. *Если случайный процесс $(X_t, t > 0)$ согласован с фильтрацией $\mathbb{F}$ и имеет непрерывные слева траектории, то он является предсказуемым.*

Лемма 13.3. *Пусть $(X_t, t > 0)$ — предсказуемый процесс, непрерывный в среднем квадратичном на отрезке $[0,t]$. Тогда*

$$\int_0^t X_s dW_s = \mathrm{l.i.m.}_{\Delta(T)\to 0} \sum_{k=1}^n X_{t_{k-1}}(W_{t_k} - W_{t_{k-1}}),$$

где $T = \{0 = t_0 < \ldots < t_n = t\}$ — разбиение $[0,t]$, $\Delta(T) = \max\limits_{k=1,\ldots,n}(t_k - t_{k-1})$.

Отметим, что условие взятия крайней левой точки t_{k-1} в лемме 13.3 весьма принципиально, при выборе других точек и предел будет другим.

▲ **Определение 13.4.** Пусть $X = (X_t, t \geqslant 0)$ — предсказуемый процесс. Говорят, что он имеет *стохастический дифференциал*

$$dX_t = g(t, \omega)dt + f(t, \omega)dW_t, \tag{6}$$

если

- процессы $g(t, \omega)$ и $f(t, \omega)$ также предсказуемы;
- процесс $f(s, \omega)I\{s \leqslant t\}$ принадлежит L для любого конечного t;
- функция $g(s, \omega)$ интегрируема на любом отрезке $[0, t]$ с вероятностью 1;
- для любого $t > 0$ с вероятностью 1 выполняется равенство

$$X_t = X_0 + \int_0^t g(s, \omega)ds + \int_0^t f(s, \omega)dW_s.$$

Фундаментальным результатом стохастического исчисления является формула Ито, позволяющая находить стохастический дифференциал функции от случайного процесса.

Теорема 13.2 (формула Ито). *Пусть $(X_t, t \geqslant 0)$ — предсказуемый процесс, имеющий стохастический дифференциал (6). Пусть функция $H(t,x) : [0,+\infty) \times \mathbb{R} \to \mathbb{R}$ такова, что существуют непрерывные частные производные $\partial H/\partial t$, $\partial^2 H/\partial x^2$. Пусть также для любого $t > 0$ процесс*

$$h(s,\omega) = f(s,\omega)\frac{\partial H}{\partial x}(s, X_s) I\{s \leqslant t\}$$

принадлежит L. Тогда процесс $Y_t = H(t, X_t)$ также имеет стохастический дифференциал, равный

$$dY_t = \frac{\partial H(t,X_t)}{\partial t}dt + \frac{\partial H(t,X_t)}{\partial x}dX_t + \\ +\frac{1}{2}\frac{\partial^2 H(t,X_t)}{\partial x^2} f^2 dt.$$

Поясним формулу. Первые два слагаемых соответствуют обычному правилу дифференцирования функций многих переменных. Но в случае стохастического дифференциала случайность дает еще одно слагаемое вида $\frac{1}{2}\frac{\partial^2 H(t,X_t)}{\partial x^2}(dX_t)^2$. Здесь $(dX_t)^2$ можно понимать так: возводим формально (6) в квадрат и полагаем

$$(dt)^2 = 0, \;\; dt \cdot dW_t = 0, \;\; (dW_t)^2 = dt.$$

Отсюда $(dX_t)^2 = f^2 dt$.

Приведем пример применения формулы Ито.

▲ **Задача 13.2.** С помощью формулы Ито вычислите

$$\int_0^t W_s dW_s.$$

Решение

Положим $Y_t = H(t, W_t)$ и найдем такую H, чтобы было выполнено равенство $dY_t = W_t dW_t$. По формуле Ито, имеем

$$dY_t = \frac{\partial H(t, W_t)}{\partial t} dt + \frac{\partial H(t, W_t)}{\partial x} dW_t + \frac{1}{2} \frac{\partial^2 H(t, W_t)}{\partial x^2} dt,$$

ведь в нашем случае $X_t = W_t$ и, стало быть, $f(s, \omega) = 1$. Приравнивая коэффициенты при dW_t и dt, получаем

$$\frac{\partial H}{\partial x} = x, \ \frac{\partial H}{\partial t} + \frac{1}{2} \frac{\partial^2 H}{\partial x^2} = 0.$$

Из первого равенства следует, что $H(t, x) = x^2/2 - C(t)$. Из второго равенства находим, что $C(t) = t/2 + C$. Начальное условие — нулевое, поэтому $C = 0$. В итоге,

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

Задача решена.

Отметим многомерный вариант формулы Ито. Рассмотрим несколько предсказуемых процессов X_t^i, $i = 1, \ldots, k$, имеющих стохастический дифференциал вида (6). Если функция $H(t, x_1, \ldots, x_k)$ удовлетворяет обобщению условий теоремы 13.2, то

$$dH(t, X_t^1, \ldots, X_t^k) = \frac{\partial H}{\partial t} dt + \sum_{i=1}^{k} \frac{\partial H}{\partial x_i} dX_t^i + \frac{1}{2} \sum_{i,j=1}^{k} \frac{\partial^2 H}{\partial x_i \partial x_j} f_i f_j dt.$$

Нам осталось разобрать вопрос о существовании решений стохастических дифференциальных уравнений. Будем рассматривать стохастическое дифференциальное уравнение

$$dX_t = b(t, X_t) dt + \sigma(t, X_t) dW_t, \ t \leqslant T, \tag{7}$$

с начальным условием $X_0 = Z$ для некоторой $\mathcal{F}_0$-измеримой случайной величины Z.

▲ **Определение 13.5.** Согласованный с фильтрацией $\mathbb{F}$ процесс $X = (X_t, t \in [0,T])$ называется *сильным решением* стохастического дифференциального уравнения (7), если он с вероятностью 1 имеет непрерывные траектории и для любого $t \leqslant T$ п. н. выполнено равенство

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s.$$

При каких условиях на функции b и σ решение существует? Потребуем выполнение двух условий: пусть существуют такие $L = L(T)$ и $c = c(T)$, что для любых $x, y \in \mathbb{R}, t \leqslant T$

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leqslant L|x-y|; \tag{8}$$

$$|b(t,x)|^2 + |\sigma(t,x)|^2 \leqslant c(1+x^2). \tag{9}$$

В данных условия выполнена следующая теорема о существовании и единственности решения уравнения (7).

Теорема 13.3. *Пусть функции b и σ измеримы и удовлетворяют условиям (8) и (9). Пусть случайная величина Z является $\mathcal{F}_0$-измеримой и $\mathsf{E}Z^2 < \infty$. Тогда существует единственное сильное решение стохастического дифференциального уравнения (7) с начальным условием $X_0 = Z$, причем функция $\mathsf{E}|X_t|^2$ ограничена на отрезке $[0,T]$.*

Рассмотрим в качестве примера уравнение Ланжевена

$$dV_t = -aV_t dt + \sigma dW_t, \ a, \sigma > 0.$$

▲ **Задача 13.3.** Докажите, что решение уравнения Ланжевена с начальным условием $V_0 = Z$ задается формулой

$$V_t = Ze^{-at} + \sigma \int_0^t e^{-a(t-s)} dW_s.$$

Решение

Положим $X_t = \int_0^t e^{as} dW_s$ и $Y_t = \sigma e^{-at} X_t$. Тогда по формуле Ито

$$\begin{aligned} dY_t &= -a\sigma e^{-at} X_t dt + \sigma e^{-at} dX_t = \\ &= -aY_t dt + \sigma e^{-at} e^{at} dW_t = \\ &= -aY_t dt + \sigma dW_t. \end{aligned}$$

Снова по формуле Ито

$$d(Ze^{-at}) = -aZe^{-at} dt,$$

откуда мы и получаем, что процесс $V_t = Ze^{-at} + Y_t$ удовлетворяет уравнению Ланжевена

$$dV_t = -aV_t dt + \sigma dW_t,$$

а также начальному условию $V_0 = Z$, ведь $Y_0 = 0$.

Задача решена.

▲ Задачи для самостоятельного решения

1. Пусть $0 = t_0 < t_1 < \ldots < t_n$ и даны случайные величины $f_1, \ldots, f_n$, причем f_i является $\mathcal{F}_{t_i}$-измеримой для всех $i = 1, \ldots, n$. Докажите, что тогда для ступенчатого процесса $f(t, \omega) = \sum_{i=1}^n f_i(\omega) I\{t_{i-1} < t \leqslant t_i\}$ процесс $\mathcal{I}_t(f)$ является мартингалом с п. н. непрерывными траекториями.

2. Пусть τ — момент остановки относительно естественной фильтрации винеровского процесса $(W_t, t \geqslant 0)$, причем $\mathsf{E}\tau < +\infty$. Докажите, что

   а) $\int_0^{+\infty} I\{t \leqslant \tau\} dW_t = W_\tau$, если τ принимает лишь конечное число значений;

   б) $\int_0^{+\infty} I\{t \leqslant \tau\} dW_t = W_\tau$, если τ — произвольный;

   в) $\mathsf{E}W_\tau^2 = \mathsf{E}\tau$.

3. Приведите пример стохастического дифференциального

уравнения с нулевым начальным условием, решение которого не единственно.

4. Решите стохастические дифференциальные уравнения:

   а) $dX_t = X_t dW_t,\ X_0 = 1$;

   б) $dX_t = \frac{1}{2} X_t dt + 2X_t dW_t,\ X_0 = 1$.

5. Докажите, что если $V(0)$ не зависит от винеровского процесса $(W_t, t \geqslant 0)$ и $V(0) \sim \mathcal{N}(0, \sigma^2/(2a))$, то решение $V = (V_t, t \geqslant 0)$ уравнения Ланжевена $dV_t = -aV_t dt + \sigma dW_t$ является гауссовским процессом с ковариационной функцией $\frac{\sigma^2}{2a} e^{-a|s-t|}$.

6. Пусть $X_t = (X_t^1, X_t^2)$ — двумерный процесс, задающийся стохастическим дифференциальным уравнением

$$dX_t = -\frac{1}{2} X_t dt + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} X_t dW_t,$$

   где $(W_t, t \geq 0)$ – одномерный винеровский процесс.

   а) Докажите, что $(X_t^1)^2 + (X_t^2)^2 = \mathrm{const}$ п. н.

   б) Решите уравнение в предположении, что $X_0^1 = 1$, $X_0^2 = 0$.

7. Пусть $(X_t, t \geqslant 0)$ — процесс из L. Докажите, что тогда процесс

$$Y_t = \exp\left(\int_0^t X_s dW_s - \int_0^t X_s^2 ds \right)$$

   удовлетворяет соотношению $\mathsf{E} Y_t \leqslant 1$ для всех $t > 0$ (предполагаем, что все необходимые процессы интегрируемы).

8. Пусть $(X_t, t \geqslant 0)$ — ограниченный процесс из L, т.е. существует такая $C > 0$, что $|X_t| \leqslant C$ для всех $t \geqslant 0$. Докажите, что тогда процесс

$$Y_t = \exp\left(\int_0^t X_s dW_s - \frac{1}{2} \int_0^t X_s^2 ds \right),\ t \geqslant 0,$$

   является мартингалом.

Заключение

В пособии изложены основы теории случайных процессов, разобраны наиболее изученные и богатые приложениями классы случайных процессов. Объем представленного материала, в целом, превышает программу стандартного семестрового курса и, скорее, соответствует годовому курсу. Тем самым, пособие может быть использовано и в рамках курсов повышенной сложности.

Заинтересованному в дальнейшем изучении теории случайных процессов читателю авторы рекомендуют, в первую очередь, ознакомиться с содержанием учебников и монографий, перечисленных в списке литературы. Так, например, в изданиях [2], [8] можно найти расширенный материал по теории мартингалов, в [2] и [7] — по марковским процессам, а книги [3]-[5] содержат значительный материал по стохастическому исчислению.

Литература

1. *Боровков А.А.* Теория вероятностей. 4-е изд. М.: Едиториал УРСС, 2003.

2. *Булинский А.В., Ширяев А.Н.* Теория случайных процессов. М.: Физматлит, 2005.

3. *Вентцель А. Д.* Курс теории случайных процессов. 2-е изд. М.: Наука. Физматлит, 1996.

4. *Крылов Н.В.* Введение в теорию случайных процессов. Часть 1. М.: изд-во МГУ, 1986.

5. *Крылов Н.В.* Введение в теорию случайных процессов. Часть 2. М.: изд-во МГУ, 1987.

6. *Севастьянов Б.А.* Курс теории вероятностей и математической статистики. 2-е изд. М.-Ижевск: Институт компьютерных исследований, 2004.

7. *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 1, Т. 2. М.: Мир, 1984.

8. *Ширяев А.Н.* Вероятность: В 2-х кн. 6-е издание, испр. М.: МЦНМО, 2017.

Учебное издание

ОСНОВЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

Учебное пособие

Составители
Жуковский Максим Евгеньевич
Родионов Игорь Владимирович
Шабанов Дмитрий Александрович

Редактор *О.П. Котова*. Корректор *Л.В. Себова*
Компьютерная верстка *Е.А. Казеннова*

Подписано в печать 24.03.2017. Формат 60×82 $^1/_{16}$.
Усл. печ. л. 5. Уч.-изд.л. 3,0. Тираж 200 экз. Заказ №.

Федеральное государственное автономное образовательное
учреждение высшего образования «Московский
физико-технический институт(государственный университет)»
141700, Московская обл., г. Долгопрудный, Институтский пер., 9
E-mail: rio@mipt.ru

Отдел оперативной полиграфии «Физтех-полиграф»
141700, Московская обл., г. Долгопрудный, Институтский пер., 9
E-mail: polygraph@mipt.ru