ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА

teach-in

ЛЕКЦИИ УЧЕНЫХ МГУ

ТЕОРИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

СЕРДОБОЛЬСКАЯ МАРИЯ ЛЬВОВНА

ФИЗФАК МГУ

БЛАГОДАРИМ ЗА ПОДГОТОВКУ КОНСПЕКТА
СТУДЕНТКУ ФИЗИЧЕСКОГО ФАКУЛЬТЕТА МГУ
ГЕОРГИЕВСКУЮ ЕКАТЕРИНУ ПАВЛОВНУ

Оглавление

Лекция 1. Основные понятия теории вероятности. Случайный процесс.

Литература.

- Б.М. Миллер, А.Р. Панков, Теория случайных процессов, Физматлит, М., 2002
- В. Феллер, Введение в теорию вероятностей и ее приложения (том 1), Мир, М., 1984
- А.Н. Ширяев, Вероятность, МЦНМО, М., 2004
- Ю.А. Розанов, Введение в теорию случайных процессов, Наука, М., 1985

Понятие случайного процесса.

Обсудим объект рассмотрения теории случайных процессов. В теории вероятностей есть понятие случайной величины. Эта математическая конструкция в ряде случаев оказывается очень удобной и адекватной математической моделью достаточно большого количества реальных физических числовых характеристик. Например, мы производим некоторое измерение. В результате мы получаем некоторое число. До измерения нам неизвестно, чему равно это число. Во-вторых, при повторении эксперимента мы (скорее всего) получим какое-то другое число. Таким образом, числовая характеристика, которую мы получаем при измерениях, носит неопределенный характер. Так как мы хотим работать с такой величиной, получать какие-то математически обоснованные выводы. Для этого делаем предположение: считаем, что детерминированность (неопределенность) данной характеристики описывается законами теории вероятностей. Тогда мы неизбежно приходим к понятию случайной величины. Для этого нужно определить **вероятностное пространство** (Ω, F, P), где
Ω – множество элементарных исходов
F – сигма-алгебра подмножеств в Ω (некоторая система подмножеств множества элементарных исходов)
$P(\cdot)$ - вероятность, то есть некоторый функционал, который каждому подмножеству A, содержащемуся в F, сопоставляет некоторое число, причем вероятность существует у тех и только у тех подмножеств, которые принадлежат сигма-алгебре событий.

Далее мы определяем **случайную величину (с.в.)** как некоторый функционал, который сопоставляет элементарному исходу действительное число:

$$\xi(\cdot) : \Omega \longrightarrow \mathbb{R} \tag{1.1}$$

Тем самым мы моделируем неопределенность результата наблюдения. Например, у нас есть некоторый случайный эксперимент. Когда мы его проводим, у нас реализуются разные элементарные исходы. Подставляя их в аргумент функционала ξ, получаем разные числа. Далее мы хотели бы считать разные вероятности, ассоциированные со случайной величиной. Поэтому на случайную величину налагается следующее требование

$$\forall x \in \mathbb{R} \quad \{\omega : \xi(\omega) < x\} \in \mathcal{F} \tag{1.2}$$

Это равносильно тому, что существует вероятность множества тех элементарных исходов, для которых $\xi(\omega) < x$

$$P\{\omega \in \Omega : \xi(\omega) < x\} \equiv P(\xi < x), \quad x \in \mathbb{R} \tag{1.3}$$

Это условие и определяет математическое понятие случайной величины. С физической точки зрения, оно позволяет моделировать определенные характеристики.

Замечание. В правой части (1.3) пропала зависимость от элементарного исхода. Дело в том, что когда мы работаем со случайными величинами, мы практически никогда не знаем, как они зависят от элементарного исхода. Более того, нам не известно, как устроено само вероятностное пространство. Поэтому данное вероятностное пространство после задания случайной величины переходит в некое новое вероятностное пространство, в котором роль элементарного исхода играет не ω, а конкретное значение случайной величины (действительное число). Математически это выражается в том, что рассматриваемая случайная величина порождает новое вероятностное пространство, в котором роль множества элементарных исходов играет действительная прямая. В качестве сигма-алгебры событий выступает сигма-алгебра борелевских подмножеств:

$$(\Omega, \mathcal{F}, P) \xrightarrow{\xi} \left(\mathbb{R}, \mathcal{B}, P_\xi\right) \tag{1.4}$$

При этом вероятность

$$P_\xi\ \mathbb{R} \to [0, 1] \tag{1.5}$$

задается так, что

$$P_\xi(-\infty, x) = P(\xi < x) \quad \forall x \in \mathbb{R} \tag{1.6}$$

В реальных физических задачах подавляющее большинство характеристик зависит от некоторых параметров (время, точка в пространстве и прочее). Соответственно, для каждого значения такого параметра мы получаем новую случайную величину. Таким образом, возникает потребность задания математической модели, в которой фигурирует не одна, а семейство случайных величин. При этом каждый член этого семейства зависит от параметра, который принимает, вообще говоря, бесконечное множество значений (например, время). Такой математический объект и есть *случайный процесс*.

Определение случайного процесса.

Определение. Случайный процесс – это семейство случайных величин

$$\{\xi(t)\}_{t\in\mathbb{T}} \tag{1.7}$$

индексированных параметром t, который пробегает множество $\mathbb{T}$, которое содержит бесконечное множество значений.

Аналогично случаю случайных величин, для расчета вероятностей, ассоциированных со случайным процессом, необходимо выполнение следующего условия . Для каждого натурального n, для любого набора из n x-в действительных чисел, для любых n значений параметра t должна существовать вероятность

$$P\left(\xi(t_1) < x_1, \ldots, \xi(t_n) < x_n\right) \tag{1.8}$$

Далее сделаем несколько упрощающих предположений, которые не сильно ограничат общность рассуждений, но при этом сократят формулы. Во-первых, будем считать, что все случайные величины в заданном семействе, заданы на одном и том же вероятностном пространстве. В противном случае пришлось бы каждый раз проверять выполнение условия (1.8). В (1.8) в скобках, по сути, указано пересечение нескольких событий. Вероятность пересечения существует, если существует вероятность каждого из этих событий. Предположение, что случайные величины заданы на одном вероятностном пространстве, автоматически обеспечивает выполнение данного условия.

Во-вторых, будем считать, что параметр t – действительное число, а множество $\mathbb{T}$ – какая-то часть действительной прямой. В этом случае будем называть параметр t *временем*.

n-мерная функция распределения случайного процесса.

Определение. Функция, аргумент которой лежит в указанном множестве

$$F_\xi^{(n)}: \ (\mathbb{R}\otimes\mathbb{T}) \to \mathbb{R} \tag{1.9}$$

(то есть, область определения этой функции – множество пар чисел: x из $\mathbb{R}$ и t из $\mathbb{T}$, причем пар n штук. Область значения этой функции – действительная прямая.) Пусть эта функция определена через

$$F_\xi^{(n)}\left(x_1, t_1; x_2, t_2; \ldots; x_n, t_n\right) = P\left(\xi(t_1) < x_1, \ldots, \xi(t_n) < x_n\right) \tag{1.10}$$

для всех $n = 1, 2\ldots$ при всех $x_1, \ldots, x_n \in \mathbb{R}$, $t_1, \ldots, t_n \in \mathbb{T}$. Тогда данная функция называется **n-мерной функцией распределения случайного процесса** $\xi(t), t \in \mathbb{T}$.

В сущности в определении идет речь о совместном распределении n случайных величин.

Существуют краткие формы записи

$$F_{\xi}^{(n)}\left(x_1, t_1; \ldots; x_n, t_n\right) \equiv F^{(n)} \equiv F\left(x_1, t_1; \ldots; x_n, t_n\right) \tag{1.11}$$

Свойства n-мерной функции распределения:

1. Пусть $\{k_1, \ldots, k_n\}$. Произвольная перестановка индексов $\{1, \ldots, n\}$ (вместе с соответствующей перестановкой аргументов функции распределения) не меняет значение функции распределения

$$F^{(n)}\left(x_{k_1}, t_{k_1}; \ldots; x_{k_n}, t_{k_n}\right) \equiv F^{(n)}\left(x_1, t_1; \ldots; x_n, t_n\right) \tag{1.12}$$

   Это свойство в сущности отражает коммутативность пересечения.

2. Зафиксируем все x_k, кроме одного (например, для $k = 1, \ldots, n-1$) и зафиксируем t_k для всех $k = 1, \ldots, n$. Тогда получим функцию одной переменной

$$F^{(n)}\left(x_1, t_1; \ldots, x_{n-1}, t_{n-1}; x_n, t_n\right) = \Phi(x_n) \tag{1.13}$$

   Эта функция обладает почти всеми свойствами обычной функции распределения: она непрерывна слева, не убывает и при $n > 1$ имеет предельные значения

$$\Phi\left(x_n\right)_{x_n \to -\infty} = 0 \tag{1.14}$$

$$\Phi\left(x_n\right)_{x_n \to +\infty} = F^{(n-1)}\left(x_1, t_1; \ldots, x_{n-1}, t_{n-1}\right) \quad \forall t_n \tag{1.15}$$

Любое семейство функции распределения обладает такими свойствами. Верно и обратное утверждение: если семейство некоторых функций удовлетворяет данным свойствам, то они будут функциями распределения некоторого случайного процесса (*теорема Колмогорова*).

Пример. В случае $n = 2$

$$F\left(x_1, t_1; x_2, t_2\right) = F\left(x_2, t_2; x_1, t_1\right) \quad \forall x_1, x_2 \in \mathbb{R}, \quad t_1, t_2 \in \mathbb{T} \tag{1.16}$$

Далее (в основном) будем считать, что времена упорядочены, то есть для $\mathbb{T} = \{t \geqslant 0\}$ выполняется $0 = t_0 < t_1 < t_2 < \ldots < t_n$.

Плотность вероятности случайного процесса.

Функция распределения сама по себе не очень информативна, так как в значение функции распределения вносят вклад все значения случайной величины, которые меньше x. Поэтому данная конструкция является не слишком удобной для того, чтобы определять, как именно устроено распределение случайной величины. Поэтому вводят дополнительную характеристику, которая более информативна с точки зрения того, какие значения более или менее вероятны. Эта конструкция называется *плотностью вероятности*.

Определение. Если для n-мерной функции распределения $F_\xi^{(n)}(\cdot)$ при любых $x_1, \dots, x_n \in \mathbb{R}$ и $t_1, \dots, t_n \in \mathbb{T}$ выполняется

$$P\left(x \leqslant \xi(t) < x + dx\right) = F_\xi^{(n)}\left(x - dx, t\right) - F_\xi^{(n)}\left(x, t\right) = p_\xi^{(n)}(x, t) dx \tag{1.17}$$

то случайный процесс имеет **n-мерную плотность распределения вероятности** $p_\xi^{(n)}(\cdot)$.

Здесь $x \leqslant \xi(t) < x + dx$ означает, что $x_k \leqslant \xi(t_k) < x_k + dx_k$ для всех $k = \overline{1, n}$. Если равенство (1.17) не выполняется, что плотность вероятности не существует. Возможна ситуация, в которой существуют плотности вероятности до какого-то порядка (например, 1-го и 2-го), а дальше равенство (1.17) не выполнено (то есть равенство не выполнено для $k = 3$. отсюда будет следовать, что и все остальные плотности не существуют.)

Таким образом, случайный процесс – это некий функционал, у которого два аргумента ω, t. Соответственно, мы можем рассмотреть ситуацию, в которой один из аргументов фиксирован, а другой – нет.

Траектория и сечение случайного процесса.

Пусть $\omega \in \Omega$ фиксирован. Тогда числовая функция

$$\xi_\omega(t), t \in \mathbb{T} : \mathbb{T} \to \mathbb{R} \tag{1.18}$$

Тогда случайность пропадает, получаем обычную числовую функцию. Эту функцию (1.18) называют **траекторией (реализацией) случайного процесса.**

Аналогично, фиксируем $t \in \mathbb{T}$. Тогда получим случайною величину

$$\xi(t) : \Omega \to \mathbb{R} \tag{1.19}$$

которая называется **сечением случайного процесса** в точке t.

Замечание. Отметим, что под записью $\xi(t)$ подразумевается одна случайная величина, то есть сечение случайного процесса. Если же мы указываем множество аргумента t, то есть $\xi(t),\ t \in \mathbb{T}$, то речь идет о случайном процессе.

Для объяснения понятий траектории и сечения обратимся к частотной интерпретации. Любая вероятностная модель моделируется частотой. Пусть мы многократно повторяем свой случайный эксперимент. Тогда доля случаев, когда реализовалось рассматриваемое событие среди всех повторений эксперимента должна быть вероятности этого события. Перенесем это понятие на случай случайного процесса. Рассмотрим движение большого количества идентичных частиц, движущихся в некотором объёме. Их очень много, поэтому это статистический ансамбль, в рамках которого можно отождествлять вероятность и частоту. Тогда $\xi(\omega, t)$ – координата частицы с номером ω в момент времени t (аналог вариативности случайно величины в зависимости от элементарного исхода – это зависимость от номера частицы). Тогда, если мы фиксируем элементарный исход, то мы рассматриваем конкретную частицу, её движение. В результате получаем некоторую кривую – траекторию движения отдельной частицы. Она же и будет траекторией случайного процесса рис. 1.1.

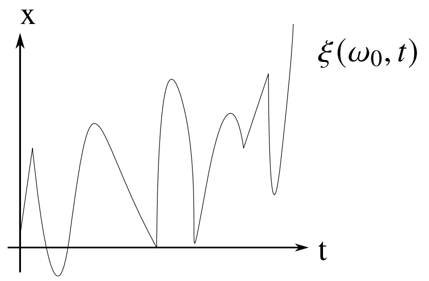

Рис. 1.1: Траектория случайного процесса.

Если же перейдем ко всему ансамблю, то нужно будет нарисовать траектории всех частиц, который движутся (рис. 1.2).

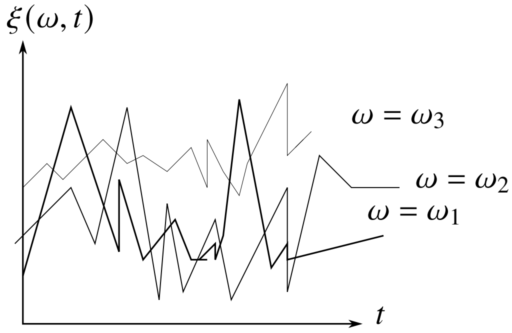

Рис. 1.2: Траектории частиц в выделенном объёме

Чтобы получить сечение случайного процесса, нужно рассечь график 1.2 прямой постоянного времени. Тогда получим огромное количество точек пересечения конкретных траекторий с вертикальной прямой (рис. 1.3). Формально сечение – это множество точек, которое образовано пересечением каждой конкретной траектории с указанной вертикальной прямой. Таких точек очень много. Каждая такая точка не несёт никакой информации о поведении сечения в конкретный момент времени. Аналогично конкретная реализация случайной величины ничего не говорит о случайной величине. Можно построить гистограмму (рис. 1.4).

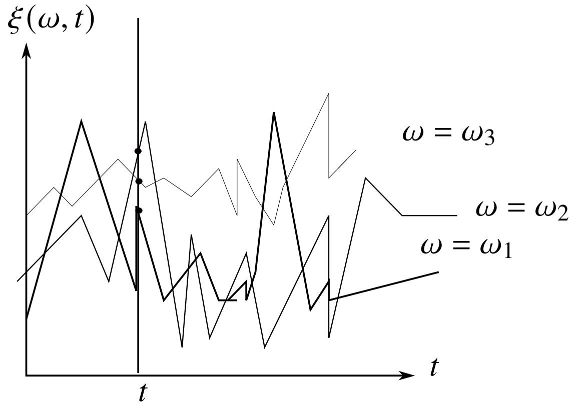

Рис. 1.3: Сечение случайного процесса

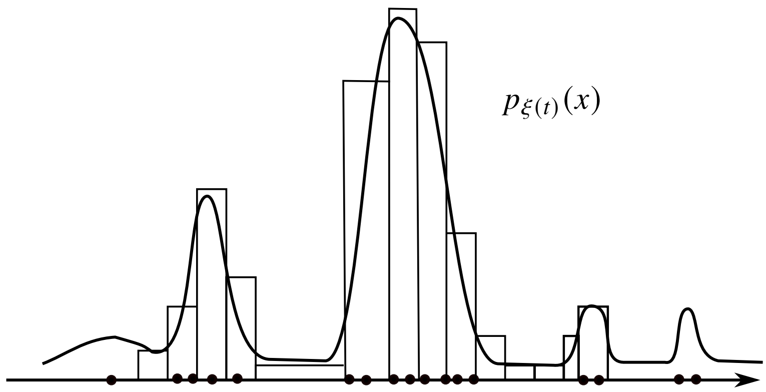

Рис. 1.4: Пример гистограммы

Высота каждого столбика – доля частиц среди всех частиц в ансамбле, у которых координаты лежат с диапазоне, обозначенным шириной столбика. Далее, если измельчать эту гистограмму, то она перейдет в плавную кривую. Эта кривая, с вероятностей точки зрения – аналог плотности вероятности случайной величины $\xi(t_0)$, то есть одномерная плотность вероятности как функция от x при фиксированном t. Эта функция характеризует распределение случайной величины, которая является сечением случайного процесса.

Таким образом, мы получили траекторию – объект математического анализа и сечение, которое характеризуется распределением случайной величины (то есть используем теорию вероятности).

Замечание. Достаточно часто будет возникать потребность анализировать распределение случайных величин как функции от t. Когда мы исследуем случайный процесс, мы интересуемся в первую очередь его динамикой (как меняется поведение системы в зависимости от времени). Скорее всего эта динамика будет описываться какими-то дифференциальными уравнениями. Любое дифференциально уравнение, как и любой интеграл – это предельный переход. Соответственно, должна быть какая-то наука, которая объясняет, как переходить к пределам, когда речь идет о случайных объектах. В случае обычной функции существуют различные способы задания производной. В случае случайной величины для вычисления $\frac{d\xi}{dt}$ нужно переходить к пределу в последовательности случайных величин. Из теории вероятности известно, что предел последовательности случайных величин определяется по-разному. При этом нам надо определить предел так, чтобы выполнялись физические предположения и чтобы определение было математически корректным. Также может возникнуть другая ситуация. Например, пусть у нас есть некоторые физические предположения о рассматриваемой системе. Возникает вопрос, нельзя ли вывести какие-нибудь дифференциальные уравнения на вероятностные характеристики. В этом случае возникает так называемое *уравнение Фокера-Планка*. Это уравнение в частных производных для плотностей вероятности случайного процесса.

Моменты случайного процесса.

Моменты случайного процесса (моменты функции случайного процесса) в сущности есть моменты сечений случайного процесса.

Если при каждом $t \in \mathbb{T}$ существует математическое ожидание величины $\xi(t)$, существует числовая функция

$$m(\cdot) : \mathbb{T} \to \mathbb{R} \tag{1.20}$$

которая называется **математическим ожиданием (момент I порядка) случайного процесса**. Нужно рассчитать математическое ожидание одного сечения:

$$m(t) = M\xi(t),\ t \in \mathbb{T} \tag{1.21}$$

Напомним, что для расчёта математического ожидания используется одномерная функция распределения (так как речь идёт только об одном сечении):

$$M\xi(t) = \int\limits_{-\infty}^{\infty} x dF^{(1)}(x,t) \tag{1.22}$$

Математическое ожидание считается существующим, если сходится интеграл от модуля

$$\int\limits_{x\in\mathbb{R}} |x| dF^{(1)}(x,t) < \infty \tag{1.23}$$

Здесь речь идет об интеграле Лебега-Стильтьеса, при этом

$$dF^{(1)}(x,t) = F^{(1)}(x+dx,t) - F^{(1)}(x,t) \tag{1.24}$$

Рассмотрим **момент II порядка**.

Определение. Рассмотрим функцию

$$d^2 : \mathbb{T} \to \mathbb{R}_+ = \{x \geqslant 0\} \tag{1.25}$$

Зададим эту функцию как дисперсию сечения случайного процесса

$$d^2(t) = D\xi(t) = M\left(\xi(t) - M\xi(t)\right)^2,\ t \in \mathbb{T} \tag{1.26}$$

эта функция называется **дисперсией случайного процесса.**

Второй момент случайного процесса характеризуется не только дисперсией, но и коэффициентом ковариации.

Определение. Пусть функция

$$R : \mathbb{T} \otimes \mathbb{T} \to \mathbb{R} \tag{1.27}$$

такая, что

$$\begin{aligned} R(t,s) \overset{def}{=} cov\left(\xi(t), \xi(s)\right) \overset{def}{=} M\left(\xi(t) - M\xi(t)\right)\left(\xi(s) - M\xi(s)\right) = \\ = M\xi(t)\xi(s) - M\xi(t) \cdot M\xi(s) \end{aligned} \tag{1.28}$$

– ковариационная функция случайного процесса.

Моменты более высоких порядков в теории случайных процессов практически никогда не используются.

Свойства моментов случайного процесса.

Верны все свойства математического ожидания (линейность, неравенства).

Существуют также другие формулы для вычисления дисперсии и ковариационной функции

$$D\xi(t) = M\xi(t)^2 - (M\xi(t))^2 \tag{1.29}$$

$$R(t,s) = M\xi(t)\xi(s) - M\xi(t)M\xi(s) \tag{1.30}$$

Из определения следует, что ковариационная функция симметрична по своим аргументом

$$R(t,s) = R(s,t),\ \forall t, s \in \mathbb{T} \tag{1.31}$$

На диагонали она совпадает с дисперсией:

$$D\xi(t) = R(t,t) \tag{1.32}$$

При этом ковариациаонная функция обладает свойством неотрицательной определенности, то есть для любых неслучайных $x_1, .., x_n \in \mathbb{R}$ и любых $t_1, \ldots, t_n \in \mathbb{T}$

$$\sum_{i,j=1}^{n} R(t_i, t_j) x_i x_j \geqslant 0 \tag{1.33}$$

В сущности это свойство неотрицательности матрицы ковариации. Докажем это. Положим (центрируем случайную величину)

$$\xi^0(t) = \xi(t) - M\xi(t) \tag{1.34}$$

тогда по определению можно расписать сумму

$$\sum_{i,j=1}^{n} R(t_i, t_j) x_i x_j = \sum_{i,j=1}^{n} M\xi^0(t_i) M\xi^0(t_j) x_i x_j = M\left(\sum_{i,j=1}^{n} \xi^0(t_i)\xi^0(t_j) x_i x_j\right) =$$
$$= M\left(\sum_{i=1}^{n} x_i \xi^0(t_i) \cdot \sum_{n=j=1}^{n} x_j \xi^0(t_j)\right) = M\left(\sum_{i=1}^{n} x_i \xi^0(t_i)\right)^2 \geqslant 0 \tag{1.35}$$

Совместные распределения случайного процесса.

Чтобы рассчитать коэффициент ковариации, нужно знать двумерное распределение случайного процесса (в определении ковариациаонной функции участвует два сечения, чтобы рассчитать математические ожидания, нужно будет знать, как совместно распределены две случайные величины). Задача нахождения совместного распределения довольна сложна. Поэтому в теории вероятностей вводят упрощающее понятие – независимость случайных величин. В теории случайных процессов тоже вводится некоторое упрощающее предположение, основанное на независимости. Однако, независимость сечений не соответствует реальности, так как в этом случае мы утверждаем, что сечения независимы, сколь бы ни были близки времена (то есть мы рассматриваем два бесконечно близких сечения). Очевидно, реальности такое предположение не выполняется. Поэтому рассмотрим независимость другого типа.

Определение. Рассмотрим два момента времени $0 \leqslant s < t$. **Приращение случайного процесса** – случайная величина $\xi(t) - \xi(s)$.

Определение. **Процесс с независимыми приращениями** – это случайный процесс $\xi(t),\ t \geqslant 0$, если для любого $n = 2, 3, \ldots$ и для любых $t_0, t_1, \ldots, t_n$, упорядоченных: $0 \leqslant t_0 < t_1 < \ldots < t_n$ приращения $\xi(t_k) - \xi(t_{k-1}),\ k = \overline{1, n}$ (как случайные величины) независимы в совокупности. Таким образом, мы требуем независимость не сечений как таковых, а изменений случайного процесса в непересекающиеся промежутки времени.

Определение. Процесс $\xi(t), t \geqslant \mathbb{T}$ называется **однородным** (во времени), если *распределение* случайной величины $\xi(t) - \xi(s)$ (приращение) при $t > s$ зависит только от разности времен $t - s$, но не зависит от t, sв отдельности.

Определение. Случайный процесс $\xi(t),\ t \geqslant 0$ **начинается в нуле**, если $\xi(0) = 0$=0 с вероятностью единица.

Предложение 1. Если случайный процесс $\xi(t),\ t \geqslant 0$ имеет независимые приращения и начинается в нуле, то его ковариационная функция $R(t, s),\ t, s \geqslant 0$ на самом деле зависит не от двух аргументов, а от одного – $min(t, s)$.

Доказательство. Центрируем сечения:

$$\xi^0(t) = \xi(t) - M\xi(t) \tag{1.36}$$

Пусть $0 < s < t$. Тогда у нас появляется два приращения – от 0 до s и от s до t. Для этих приращений введем обозначения

$$\Delta\xi_1 = \xi(s) - \xi(0) = \xi(s) \tag{1.37}$$

$$\Delta\xi_2 = \xi(t) - \xi(s) \tag{1.38}$$

Тогда, по условию $\Delta\xi_1$ и $\Delta\xi_2$ независимы как приращения случайного процесса. Центрируем приращения

$$\Delta\xi_1^0 = \xi^0(s) - \xi^0(0) \tag{1.39}$$

$$\Delta\xi_2^0 = \xi^0(t) - \xi^0(s) \tag{1.40}$$

Тогда

$$\Delta\xi_1^0 = \Delta\xi_1 - (M\xi(s) - M\xi(0)) \tag{1.41}$$

$$\Delta\xi_2^0 = \Delta\xi_2 - (M\xi(t) - M\xi(s)) \tag{1.42}$$

Эти случайные величины тоже являются независимыми как функции от функции от независимых случайных величин. Выразим ковариационную функцию через указанные центрированные приращения. По условию процесс начинается в нуле, поэтому

$$\xi^0(s) = \Delta\xi_1^0 \tag{1.43}$$

Второе сечение есть сумма приращений

$$\xi^0(t) = \xi^0(s) + \xi^0(t) - \xi^0(s) = \Delta\xi_1^0 + \Delta\xi_2^0 \tag{1.44}$$

Тогда, так как математические ожидания центрированных приращений равны нулю:

$$M\Delta\xi_1^0 = M\xi^0(s) = 0 \tag{1.45}$$

$$M\Delta\xi_2^0 = M\xi^0(t) - M\xi^0(s) = 0 \tag{1.46}$$

Получаем (по определению) выражение для ковариационной функции

$$\begin{aligned} R(t,s) &\overset{def}{=} M\xi^0(t)\xi^0(s) = M\Delta\xi_1^0\left(\Delta\xi_1^0 + \Delta\xi_2^0\right) = M\left(\Delta\xi_1\right)^2 + M\Delta\xi_1^0\Delta\xi_2^0 = \\ &= \text{воспользуемся независ. приращений} = M\left(\Delta\xi_1^0\right)^2 - M\Delta\xi_1^0 \cdot M\Delta\xi_2^0 = \\ &= M\left(\xi^0(s)\right)^2 + 0 = D\xi(s) \end{aligned} \tag{1.47}$$

Таким образом, получили, что в рассматриваемом случае ковариационная функция равна дисперсии сечения в меньший из двух момент времени. Далее, если поменяем местами t и s, аналогично получим

$$R\left(t,s\right) = D\xi(t), \;\; 0 < t < s \tag{1.48}$$

Это можно записать в виде

$$R\left(t,s\right) = D\xi(t_0), \;\; t_0 = min(t,s) \tag{1.49}$$

Лекция 2. Процесс Пуассона.

Повторение.

Определение. Приращение случайного процесса $\xi(t),\ t \geqslant 0$ – это случайная величина $\xi(t)-\xi(s)$ при $0 \leqslant s < t$. То есть, приращение – это то, как изменилось значение случайного процесса от момента случайного s к моменту времени t.

Определение. Случайный процесс $\xi(t),\ t \geqslant 0$ называется **процессом с независимыми приращениями**, если для любого $n = 2, 3, \ldots$ и для любых $t_0, t_1, \ldots, t_n$, таких, что $0 \leqslant t_0 < t_1 < \ldots < t_n$ случайные величины $\xi(t_k) - \xi(t_{k-1}),\ k = 1, \ldots, n$, независимы (в совокупности).

Определение. Случайный процесс $\xi(t),\ t \geqslant 0$ называется **однородным (во времени)**, если для любых $0 \leqslant s < t$ *распределение* приращения $\xi(t) - \xi(s)$ зависит только от $t - s$.

Определение. Случайный процесс $\xi(t),\ t \geqslant 0$ **начинается в нуле**, если $\xi(0) = 0$ с вероятностью 1.

Далее мы будем рассматривать процесс с описанными свойствами – процесс Пуассона.

Процесс Пуассона

Определение. Процесс Пуассона – случайный процесс $\xi(t),\ t \geqslant 0$ со следующими свойствами:

1. $\xi(t), t \geqslant 0$ – процесс с независимыми приращениями
2. $\xi(t), t \geqslant 0$ – процесс, однородный во времени
3. $\xi(t)$ начинается в нуле, то есть $P\left(\xi(0) = 0\right) = 1$
4. распределение приращения $\xi(t)-\xi(s), t > s$ есть распределение Пуассона с параметром $\lambda(t-s)$, т.е.
$$P\left(\xi(t) - \xi(s) = m\right) = \frac{(\lambda(t-s))^m}{m!}e^{-\lambda(t-s)}, m = 0, 1, ... \tag{2.1}$$
$\lambda = const, 0 < \lambda < \infty$

Отметим, что условие 4 согласовано с условием 2. Формула распределения (2.1) действительно зависит только от разности времени, то есть процесс является однородным во времени. Более того, условие 4 согласовано также и с условием 3. Если мы будем стягивать друг к другу t и s, то получим, что пери всех $m \neq 0$ формула (2.1) будет давать нам нулевую вероятность, а в нуле (при $m = 0$) сечение будет равно 0 с вероятностью 1.

Свойства процесса Пуассона.

Рассмотрим одно сечение случайного процесса. Его можно записать в виде

$$\xi(t) = \xi(t) - \xi(0) \in \mathbb{P}(\lambda t) \tag{2.2}$$

При этом эта случайная величина имеет распределение Пуассона с параметром λt. Отсюда сразу можно найти моменты процесса. Запишем явный вид распределения

$$P(\xi(t) = m) = \frac{(\lambda t)^m}{m!} e^{-\lambda t},\ m = 0, 1, ... \tag{2.3}$$

Известно, что математическое ожидание распределения Пуассона равно параметру Пуассона. Соответственно, математическое ожидание случайного процесса линейно зависит от времени

$$M\xi(t) = \lambda t \tag{2.4}$$

Дисперсия Пуассонова процесса равна его математическому ожиданию, поэтому

$$D\xi(t) = \lambda t \tag{2.5}$$

Тогда, по доказанному выше свойству, получаем выражения для ковариационной функции случайного процесса

$$R(t, s) = \lambda \cdot min(t, s) \tag{2.6}$$

так как процесс однородный с независимыми приращениями и начинается в нуле.

Замечание. Пусть $\xi(t)$ описывает некоторую физическую характеристику. Тогда эта характеристика должна быть безразмерной. Это видно из формул (2.4) и (2.5). Отметим также, что параметр λ имеет размерность обратного времени.

n-мерное распределение.

Далее посчитаем $n-$мерные распределения. У нас дискретные сечения, соответственно, нужно посчитать $n-$мерное дискретное распределение. То есть, будем считать совместную вероятность сечений для разных времен $0 = t_0 < t_1 < \ldots < t_n$ и для неотрицательных целых чисел $m_1, \ldots, m_n \in \{0, 1, 2, \ldots\}$, $m_0 = 0$. Для указанных параметров посчитаем вероятность того, что сечение в момент времени t_1 примет значение m_1, в момент времени t_2 – значение m_2 и т.д. Для краткости обозначим эту вероятность $P^{(n)}$, где n – количество событий под знаком вероятности, m – обобщенный индекс, составленный из числовых параметров m_k, t – обобщенный аргумент, составленный из t_k:

$$P(\xi(t_1) = m_1, \xi(t_2) = m_2, \ldots, \xi(t_n) = m_n) = P^{(n)}(m, t) \tag{2.7}$$

Известно, что для любого $k = 1, 2, \dots$ приращение имеет распределение Пуассона

$$\xi(t_k) - \xi(t_{k-1}) \in \mathbb{P}\left(\lambda(t_k - t_{k-1})\right) \tag{2.8}$$

При этом мы сразу можем сделать вывод, что это приращение неотрицательно с вероятностью единица

$$\xi(t_k) - \xi(t_{k-1}) \geqslant 0 \ \text{ с вероятностью } 1 \tag{2.9}$$

так как в распределении Пуассона присутствуют только неотрицательные значения. Соответственно, вероятность (2.7) имеет смысл считать только тогда, когда m_k постепенно возрастают. В противном случае, если хотя бы для одного $k = 1, 2, \dots$ нарушено неравенство $m_{k-1} \leqslant m_k$ вероятность будет равна нулю $P^{(n)}(m, t) = 0$. Поэтому далее будем искать эту вероятность только для упорядоченных m_k. Пусть $0 = m_0 \leqslant m_1 \leqslant m_2 \leqslant \dots \leqslant m_n$. Для краткости обозначим

$$\Delta\xi_k = \xi(t_k) - \xi(t_{k-1}), \quad k = 1, 2, \dots \tag{2.10}$$

Тогда событие

$$\xi(t_1) = m_1, \ \ \xi(t_2) = m_2, \dots, \xi(t_n) = m_n \tag{2.11}$$

равносильно событию

$$\Delta\xi_1 = m_1 - m_0, \ \ \Delta\xi_2 = m_2 - m_1, \dots, \Delta\xi_n = m_n - m_{n-1} \tag{2.12}$$

так как если мы знаем сечения, то мы знаем и приращения. Обратно, если сечения устроены так, как в (2.12), то для сечений получим (2.11). Таким образом, эти события равносильны, то есть их вероятности равны. При этом нам известно, что $\Delta\xi_1, \dots, \Delta\xi_n$ – независимые случайные величины. При этом нам известно их распределение

$$\Delta\xi_k \in \mathbb{P}\left(\lambda(t_k - t_{k-1})\right) \tag{2.13}$$

следовательно

$$\begin{aligned} P\left(\xi(t_1) = m_1, \xi(t_2) = m_2, \dots, \xi(t_n) = m_n\right) = \\ = P\left(\Delta\xi_1 = m_1 - m_0, \Delta\xi_2 = m_2 - m_1, \dots, \Delta\xi_n = m_n - m_{n-1}\right) = \\ = \prod_{k=1}^{n} \frac{\left(\lambda(t_k - t_{k-1})\right)^{m_k - m_{k-1}}}{(m_k - m_{k-1})!} e^{-\lambda(t_k - t_{k-1})} \end{aligned} \tag{2.14}$$

Таким образом, мы получили n-мерное распределение процесса Пуассона. Оно устроено как произведение пуассоновых вероятностей относительно разностей времен и разности значений.

Замечание. Перепишем полученную формулу как

$$P^{(n)}(m,t)=\prod_{k=1}^{n}\frac{(\lambda\Delta t_k)^{\Delta m_k}}{(\Delta m_k)!}e^{-\lambda\Delta t_k} \tag{2.15}$$

где $\Delta m_k = m_k - m_{k-1}$, $\Delta t_k = t_k - t_{k-1}$ для $k = 1,\ldots,n$ ($m_0 = 0, t_0 = 0$).

Преобразуем экспоненциальный множитель. Очевидно

$$\prod_{k=1}^{n} e^{-\lambda\Delta t_k} = e^{\Delta t_1+\ldots+\Delta t_n} = e^{-\lambda t_n} \tag{2.16}$$

Также можно преобразовать первый множитель (учитывая, что $\Delta m_1 + \Delta m_n = m_n$)

$$P^{(n)}(m,t)=\frac{m_n!}{\Delta m_1!\ldots\Delta m_n!}p_1^{\Delta m_1}\ldots p_n^{\Delta m_n}\cdot\frac{(\lambda t_n)^{m_n}}{m_n!}e^{-\lambda t_n} \tag{2.17}$$

где

$$p_k=\frac{\lambda\Delta t_k}{\lambda t_n} \tag{2.18}$$

причем

$$p_1+\ldots+p_n=1 \tag{2.19}$$

Таким образом, получили формулу (2.17), в которой второй множитель – пуассонова вероятность от последнего времени и последнего значения. Первый множитель называется *полиномиальной вероятностью*. В сущности это обобщение биномиальной вероятности, то есть это вероятность того, что при m_n испытаниях произошло Δm_k исходов типа k. Если бы мы рассматривали двумерное распределение, то этот множитель представлял бы собой обычный биномиальный член.

Пуассонов поток требований.

Рассмотрим некоторую физическую модель, приводящую к математической конструкции процесса Пуассона. Рассмотрим следующую ситуацию. Пусть в некоторый случайный момент времени t поступает требование (события, которые происходят в случайные моменты времени, будем называть требованиями). Такая ситуация (произойдет событие или нет) управляется теорией вероятности. Длительность самих событий мы считаем нулевой. Будем считать, что $\xi(t)$ – число требований в промежуток $[0,t)$.
Огромное количество процессов имеет такую структуру – мы считаем количество некоторых событий, а сами события происходят в непредсказуемые заранее моменты времени. Например, данную модель можно применять к системам обслуживания, во многих физических задачах (радиоактивный распад, вспышки на солнце и многое другое). Количество требований к моменту времени t есть случайная величина. Если мы параметром t пробегаем все время, то получаем, что $\xi(t)$ – случайный процесс. Наложим на него

дополнительное условие.

Определение. Поток требований $\xi(t), t \geqslant 0$ называется **пуассоновым**, если он удовлетворяет следующим требованиям

1. *имеет независимые приращения*. То есть, если мы рассматриваем количество требований на непересекающихся промежутках времени, то количество требований на этих промежутках будут независимыми случайными величинами.

2. *однородный во времени*. Это означает, что распределение количества требований зависит только от длины промежутка.

3. *начинается в нуле*, то есть отсчет требований мы начинаем с нуля.

4. при $h \to 0$ распределение приращения $\Delta\xi(h) = \xi(t+h) - \xi(t)$ (количество требований, которые поступили с момента времени t вплоть до момента времени $t+h$) удовлетворяет следующим требованиям

   (a) $P(\Delta\xi(h) = 0) = 1 - \lambda h + o(h)$

   (b) $P(\Delta\xi(h) = 1) = \lambda h + o(h)$

   (c) $P(\Delta\xi(h) > 1) = o(h)$

   Здесь $\lambda = const,\ 0 < \lambda < \infty$.

Замечание. Заметим, что многие реальные процессы не всегда удовлетворяют первым трем условиям теоремы. Например, в случае подсчёта количества вызовов на телефонную станцию мы получаем неоднородный во времени случайный процесс. Вспышки на солнце (в случае рассмотрения больших промежутков времени) тоже являются неоднородным во времени случайным процессом. Однако, если рассматривать эти небольшие промежутки времени, данные процессы можно считать однородными во времени.

Обсудим заданные условия. В сущности условие (a) удовлетворяет ситуации, в которой промежуток времени очень мал. Понятно, что в промежуток нулевой длины не может поступить требование. То есть, в случае нулевого временного промежутка вероятность того, что количество требований равно нулю, равна единице. Однако, если промежуток времени (то есть h) очень мал, но не равен нулю, то существует малая вероятность того, что ни одно требование не поступит, уменьшается на некоторую малую поправку, то есть вероятность. записанная в условии (a), будет меньше 1 на λh.

Второе условие (b) соответствует тому, что за малый промежуток времени поступит ровно одно требование. Вероятность этого события мала, имеет первый порядок малости по h. Все остальные же варианты (условие (c)) маловероятны, их вероятность исчезающе мала.

Таким образом, данные условия описывают редкий поток событий. В малые промежутки времени скорее всего не происходит ничего или поступает одно требование. Отметим, что данные выражения по сути есть разложения вероятности $P\left(\xi(t)=m\right)=P_m(t)$ как функции от t в ряд Тейлора в точке $t=0$.

Понятно, что описанный пуассонов поток требований очень похож на поток пуассона, за исключением последнего свойства. На самом деле можно показать, что в сущности эти два понятия идентичны.

Теорема. Пуассонов поток требований $\xi(t), t \geqslant 0$ есть процесс Пуассона.

Доказательство теоремы. Покажем, что количество требований (распределение сечения пуассонового потока требований) имеет распределение Пуассона: $\xi(t) \in \mathbb{P}(\lambda t)$ при $t>0$ $(\xi(0)=0 \; c \; P=1)$.
Зафиксируем $t = fix, t > 0$. Разобьём промежуток $[0,t]$ разобьём на n одинаковых промежутков длины h:

$$[0) = \bigcup_{k=1}^{n} \Big[t_{k-1} - t_k\Big), \;\; t_k - t_{k-1} = \frac{t}{n} \overset{def}{=} h, \;\; n \in \mathbb{N}, \; t_0 = 0 \tag{2.20}$$

причём n пока что фиксировано.

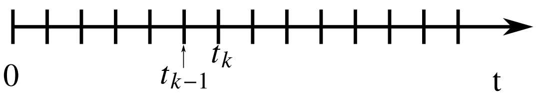

Рис. 2.1: Деление временного отрезка

Рассмотрим приращение пуассонового потока требований на каждом из маленьких промежутков

$$\Delta\xi_k = \xi(t_k) - \xi(t_{k-1}) \tag{2.21}$$

При этом данная величина есть число требований в $[t_{k-1}, t_k)$, $k = 1, \ldots, n$.

Тогда, согласно условию независимости приращений, $\Delta\xi_1, \ldots, \Delta\xi_k$ независимы. При этом они одинаково распределены (так как процесс однородный во времени, а длина промежутка одинакова для всех промежутков). Распределение $\Delta\xi_k$ удовлетворяет асимптотическим формулам $(a)-(c)$.

Введём события $A_k^{(n)}$ вида:

$$A_k^{(n)}: \quad \text{в промежутке } [t_{k-1}, t_k) \text{ поступило } \leqslant 1 \text{ требования} \Longleftrightarrow \{\Delta\xi_k = 0, 1\} \tag{2.22}$$

здесь n обозначает количество промежутков времени.

Введем событие пересечения

$$A^{(n)} \bigcap_{k=1}^{n} A_k^{(n)} \Leftrightarrow \text{ в каждом промежутке поступило } \leqslant 1 \text{ требования} \tag{2.23}$$

То есть в каждом из промежутков разбиения произошло не более одного требования. Рассмотрим также событие

$$B^{(n)} = \overline{A^{(n)}} = \bigcup_{k=1}^{n} B_k^{(n)}, \quad B_k^{(n)} = \overline{A_k^{(n)}} = \{\Delta\xi_k > 1\} \tag{2.24}$$

Мы вводим данные события, чтобы привязаться к событиям, заданным в условиях определения.

Разложим вероятность по полной группе событий:

$$P\left(\xi(t) = m\right) = P\left(\{\xi(t) = m\} \bigcap A^{(n)}\right) + P\left(\{\xi(t) = m\} \bigcap B^{(n)}\right) = P_1^{(n)} + P_2^{(n)} \tag{2.25}$$

Тогда вероятность $P_1^{(n)}$ связана с тем, что на каждом интервале поступило не более одного требования, а вероятность, а вероятность $P_2^{(n)}$ соответствует тому, что хотя бы на одном из промежутков времени поступило два или более промежутков времени.

Рассмотрим второе слагаемое:

$$P_2^{(n)} = P\left(\{\xi(t) = m\} \bigcap B^{(n)}\right) \overset{1}{\leqslant} P\left(B^{(n)}\right) \overset{2}{=} P\left(\bigcup_{k=1}^{n} B_k^{(n)}\right) \overset{3}{\leqslant} \sum_{k=1}^{n} P\left(B_k^{(n)}\right) \overset{4}{=}$$
$$= \sum_{k=1}^{n} P\left(\Delta\xi_k > 1\right) \overset{5}{=} n \cdot o(h) \overset{6}{=} n \cdot o\left(\frac{t}{n}\right) \underset{\substack{n\to\infty \\ t=fix}}{\longrightarrow} 0 \tag{2.26}$$

Пояснения:
$\overset{1}{\leqslant}$ – вероятность пересечение не больше, чем вероятность одного события
$\overset{2}{=}$ – вероятность одного события есть вероятность объединения, так как (2.9)
$\overset{3}{\leqslant}$ – вероятность объединения не больше, чем сумма вероятностей
$\overset{4}{=}$ – согласно (2.24)
$\overset{5}{=}$ и $\overset{6}{=}$ – запишем асимптотику и учтём, что $\Delta t = \frac{t}{n}$

Таким образом, вероятность второго слагаемого в (2.12) можно сделать сколь угодно малой, если n достаточно велико.

Рассмотрим первое слагаемое в (2.12). Очевидно, что

$$\xi(t) = \Delta\xi_1 + \ldots + \Delta\xi_n \tag{2.27}$$

Рассмотрим пересечение событий

$$\{\xi(t)=m\}\bigcap A^{(n)} \iff \begin{cases} \text{ровно } m \text{ из } \Delta\xi_1 \ldots \Delta\xi_n \text{ равны } 1 \\ \text{ровно } n-m \text{ из } \Delta\xi_1 \ldots \Delta\xi_n \text{ равны } 0 \end{cases} \tag{2.28}$$

То есть, сумма приращений равна m, и каждое приращение равно или 0, или 1. Таким образом, событие из (2.28) отвечает тому, что m случайных величин принимает значение 1, а $n-m$ случайных величин принимает значение 0. Если $\Delta\xi_1 = 1, \ldots, \Delta\xi_m = 1$, а остальные $\Delta\xi_k = 0$, то при фиксированных $k_1, \ldots, k_n$

$$P\left(\Delta\xi_{k_1} = 1, \ldots, \Delta\xi_{k_m} = 1, \Delta\xi_k = 0 \text{ при прочих } k\right) = p_n^m \cdot q_n^{n-m} \tag{2.29}$$

где

$$p_n = P(\Delta\xi_k = 1) = \lambda\frac{t}{n} + o\left(\frac{t}{n}\right) \tag{2.30}$$

$$q_n = P(\Delta\xi_k = 0) = 1 - \lambda\frac{t}{n} + o\left(\frac{t}{n}\right) \tag{2.31}$$

Видно, что $p_n \to 0$ при $n \to \infty$ и

$$\lim_{n\to\infty} np_n = \lim_{n\to\infty} n\left(\lambda\frac{t}{n}\right) = \lambda \cdot t \tag{2.32}$$

По теореме Пуассона

$$P_1^{(n)} = P\left(\{\xi(t)\} = m \cap A^{(n)}\right) = C_n^m p_n^m q_n^{n-m} \underset{n\to\infty}{\longrightarrow} \frac{(\lambda t)^m}{m!}e^{-\lambda t} \cdot 1 \tag{2.33}$$

Этот же результат можно объяснить и другим способом: если мы гарантируем, что на каждом из маленьких промежутков происходит не более одного события, то это эквивалентно схеме независимых испытаний (успех – на данном маленьком промежутке поступило одно требование, неудача – ничего не поступило). Процесс с независимыми приращениями, значит, испытания независимы. Таким образом, имеем n независимых испытаний, причём вероятность успеха в каждом единичном испытании одинаковая (так как процесс однородный). То есть, речь идёт о схеме Бернулли.

Итак, при фиксированных $t \geqslant 0$ и $m \geqslant 0$ для любого $n = 1, 2, \ldots$

$$P(\xi(t) = m) = P_1^{(n)} + P_2^{(n)} \tag{2.34}$$

С другой стороны, при $n \to \infty$

$$P_1^{(n)} + P_2^{(n)} \to \frac{(\lambda t)^m}{m!}e^{-\lambda t} + 0 = \frac{(\lambda t)^m}{m!}e^{-\lambda t} \tag{2.35}$$

То есть, вероятность (2.34) сколь угодно мало отличается от вероятности (2.35), если n достаточно большое. Однако, левая и правая части никак от n не зависят, поэтому

$$P\left(\xi(t)=m\right)=\frac{(\lambda t)^m}{m!}e^{-\lambda t} \tag{2.36}$$

То есть вероятность задается пуассоновым выражением. Соответственно, получаем, что сечение имеет распределение Пуассона. В силу однородности любое приращение имеет такое распределение. Все остальные свойства процесса Пуассона выполнены. Теорема доказана.

Понятно, что верно и обратное утверждение. Процесс Пуассона есть пуассонов поток требований. То есть, если бы будем рассматривать процесс Пуассона как процесс подсчета чего-то, что мы назовем требованиями, то мы получим пуассонов поток требований.

Траектория процесса Пуассона.

Рассмотрим процесс Пуассона как поток требований (рис. 2.2). Отсчет требований начинается в момент времени 0. До момента поступления первого требования значение сечения равно нулю. Мы договорились, что мы подсчитываем требования, не включая момент t, поэтому, если τ_1 – момент поступления первого требования, то в этот момент сечение все еще равно нулю. Далее происходит скачок на единицу. Далее, при поступлении каждого нового требования будем получать все новые скачки. Таким образом, траектория пуансонова процесса представляет собой лесенку, непрерывную слева. Отметим, что высота ступенек остается одной и той же, а ширина ступенек меняется.

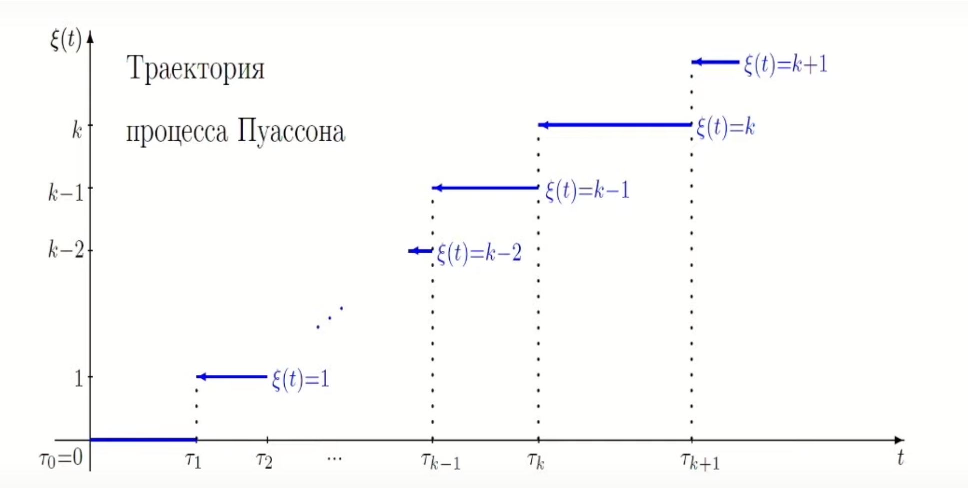

Рис. 2.2: Иллюстрация к объяснению.

Отсюда можно получить

$$\xi(t)=k \Leftrightarrow \tau_k < t \leqslant \tau_{k+1} \tag{2.37}$$

Первое событие соответствует тому, что к моменту времени t мы насчитали k требований. Во втором событии (которое выражается через неравенства) t – фиксированный числовой параметр, который не является случайным. Случайные здесь только границы интервала. Рассмотрим эту ситуацию в терминах частоты. Пусть дана некоторая числовая ось. На эту числовую ось мы куда-то ставим числовой параметр t и жестко его фиксируем. У нас есть некоторое разнообразие элементарных исходов-интервалов. Все эти интервалы наносим на числовую ось. Некоторые из них накрывают значение t, а некоторые – нет. Соответственно, вероятность рассматриваемого события есть доля тех интервалов, которые накрыли значение t. Данная идея широко применяется в математической статистике (*интервальное оценивание*). С другой стороны, данное событие можно рассматривать как стандартные неравенства для случайных величин τ_k и τ_{k+1}. Далее при расчете мы будем рассматривать вероятности с точки зрения сечений процесса Пуассона и относительно случайных времён поступления требований.

Распределения времён поступления требований.

Траектория процесса Пуассона есть кусочно-постоянная функция, скачки которой равны 1 и происходят в случайные моменты времени $\tau_1, \tau_2, \ldots$. Эта функция непрерывная слева, так как $\xi(t)$ – количество требований на $[0, t)$, поэтому $\xi(\tau_k) = k - 1$.

Случайные величины $\tau_1, \tau_2, \ldots$ принимают значения из $\mathbb{R}_+ = \{t \geqslant 0\}$ и они упорядочены с вероятностью 1, то есть $\tau_1 < \tau_2 < \ldots$.

Отметим эквивалентность событий. Если $\tau_k \geqslant t$, то $k-$е требование поступило самое раннее в момент времени t. Следовательно, на промежутке $[0, t)$ число требований заведомо $< k$, то есть $\xi(t) < k$. Соответственно, к моменту времени t и позже k-е требование может поступить только в момент времени t или позже. Соответственно, эквивалентны события

$$\tau_k \geqslant t \Leftrightarrow \xi(t) < k \tag{2.38}$$

Перейдем к дополнениям

$$\tau_k < t \Leftrightarrow \xi(t) \geqslant k \tag{2.39}$$

Отметим, что τ распределены абсолютно непрерывно, поэтому для них (в описании событий) непринципиально, строгие неравенства мы пишем, или нет. А ξ распределено дискретно, поэтому для нее события со строгим неравенством с точки зрения вероятности существенно отличается от события с нестрогим неравенством.

Таким образом, получаем распределение первого времени

$$F_{\tau_1} = P(\tau_1 < t) = P(\xi(t) \geqslant 1) = 1 - P(\xi(t) = 0) = 1 - \frac{(\lambda t)^0}{0!} e^{-\lambda t} = 1 - e^{-\lambda t} \tag{2.40}$$

Мы рассматриваем вероятность того, что первое требование поступило раньше, чем момент времени t. Это эквивалентно тому, что в момент времени t поступило требований заведомо больше или равно единицы. Мы знаем, как распределено $\xi(t)$. Таким образом, получили функцию распределения времени поступления первого требования. Эта функция непрерывна в каждой точке. Это гарантирует существование плотности вероятности

$$p_{\tau_1}(t) = F'_{\tau_1}(t) = \lambda e^{-\lambda t}, \quad t > 0 \tag{2.41}$$

При этом

$$F_{\tau_1}(t) = 0 \text{ при } t \leqslant 0, \quad p_{\tau_1}(t) = 0 \text{ при } t < 0 \tag{2.42}$$

Таким образом, $\tau_1 \in \mathbb{E}(\lambda)$ – экспоненциальное распределение с параметром λ.

Отметим, что математическое ожидание процесса Пуассона равно λt (2.4). Соответственно, при $t = 1$ математическое ожидание равно λ. Соответственно, λ по смыслу есть средняя интенсивность потока требований (количество требований, поступивших в единицу времени).

Лекция 3. Процесс Винера.

Дискретные случайный блуждания.

Рассмотрим дискретные случайные блуждания. Пусть есть некоторая точечная частица, которая совершает одномерное движение по действительной прямой $R = \{-\infty < x < \infty\}$. Пусть $\xi(t)$ – координата частицы в момент времени t. При этом те моменты времени, в которые мы наблюдаем частицу, мы будем считать дискретными. В начальный момент времени $t = 0$ частица находится в некоторой фиксированной точке. Потом она начинает движение влево или вправо с постоянной скоростью и в течение промежутка времени длины 1 она с постоянной скоростью движется по прямой. Спустя этот промежуток времени, то есть в момент времени $t = 1$ частица или продолжает движение в том же направлении или разворачивается и продолжает движение в обратном направлении. Далее аналогично. Получаем некоторое блуждание по прямой. При этом мы считаем, что изменения направления носят случайный характер.

Далее задаем процесс – дискретные случайные блуждания. Это *случайный процесс с дискретным временем*, то есть $T = \{0, 1, \dots\}$. На этот процесс налагаем требования

1. приращения координаты частицы

$$\Delta\xi_k = \xi(k) - \xi(k-1), k = 1, \dots, n \tag{3.1}$$

– независимые случайные величины для любого $n \geqslant 2$

2. частица движется или в том же направлении, или поворачивает, то есть изменение координаты частицы за время 1 либо +1 с вероятностью p, либо -1 с вероятностью q и $p + q = 1$, то есть частица не совершает никаких других движений

$$P(\Delta\xi_k = 1) = p, \quad P(\Delta\xi_k = -1) = q, \quad p + q = 1 \tag{3.2}$$

3. в начальный момент времени частица находилась в некоторой точке на действительной прямой

$$P(\xi(0) = m) = 1, \quad m - \text{ некоторое целое число} \tag{3.3}$$

Таким образом, мы получили процесс с независимыми приращениями, однородный, начинающийся в m. Однородность в данном случае заключается в том, что вероятности в условии 2 не зависят от k – неважно, на каком шаге происходит разворот (или не разворот) частицы. Выбор между направлениями следующего движения не зависит от момента времени.

Траектория данного процесса есть набор дискретных точек (k, ξ_k), $k = 0, 1, \dots$. Вторая координата дискретна, так как частица в результате своих движений может оказаться только в точке с целочисленной координатой. В результате получим рис. 3.1.

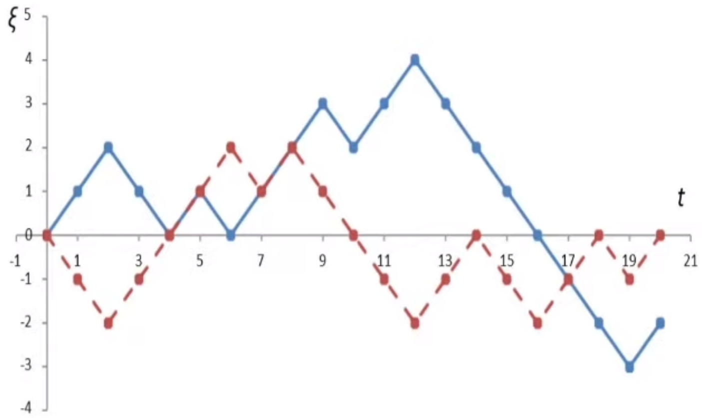

Рис. 3.1: Траектория случайных блужданий.

Здесь изображены две возможные траектории случайного процесса. Обе траектории начинаются в нуле. Частица, движущаяся по синей траектории, начала движение вправо, оказалась в точке 1. Далее она продолжает движение в том же направлении и оказывается в точке с координатой 2. Далее она поворачивает и делает два шага влево. Далее аналогично. Таким образом, такого сорта траектория представляет собой ломаную. Каждое звено ломаной по оси x изменяется на $\Delta x = \pm 1$. Шаг дискретизации по времени тоже единичный, $\Delta t = 1$. Тангенс угла наклона каждого звена ломаной (скорость частицы) $\frac{\Delta x}{\Delta t} = \pm 1$.

Попробуем посчитать вероятность того, что частица, выйдя из некоторой точки m через n шагов окажется в точке, смещенной на d единиц. Пусть

$$\xi_n = \xi(n), \quad n = 0, 1, 2, \ldots, \quad P(\xi_0 = m) \tag{3.4}$$

Найдем $P(\xi_n = m + d)$ для каждого $d \in \mathbb{Z}$.

Пусть частица сделала ровно k прыжков вправо и $n - k$ прыжков влево. Тогда можно найти, на сколько она сместилась относительно начала

$$d = 1 \cdot k + (-1) \cdot (n - k) = 2k - n \tag{3.5}$$

Отсюда получаем

$$k = \frac{n + d}{2} \tag{3.6}$$

Понятно, что количество прыжков, которое частица могла сделать вправо из общего количества прыжков, должно быть целым числом и не должно быть меньше 0 и больше n. Это налагает условия на n и d. То есть, если задано число шагов, то смещение не может быть произвольным. Значит,

$$P\left(\xi_n = m + d\right) = 0 \tag{3.7}$$

выполнено если и только если $k \in \{0, 1, \dots, n\}$. Соответственно, максимально возможное смещение равно $k = n$. Случай $k = 0$ отвечает всем прыжкам влево и тогда смещение равно $-n$. Кроме того, условие, что k – целое, означает, что n и d должны быть одной четности. Учитывая данные ограничения, получим формулу для расчета вероятности того, что $\times i_n = m + d$:

$$P\left(\xi_n = m + d\right) = P\left(\xi_n = m + d \middle| \xi_0 = m\right) = P\left(\xi_n = m + d, \xi_0 = m\right) = C_n^k p^k q^{n-k} \tag{3.8}$$

Здесь условная и безусловная вероятность одинаковы, так как условие происходит всегда (см. выше). При этом условная вероятность равна вероятности пересечения, так как вероятность наступления события $\xi_0 = m$ равна 1. Поэтому здесь записано три возможных записи одной и той же вероятности. Чтобы посчитать эту вероятность, учтем, что всего есть n прыжков. Чтобы сместиться на d единиц, мы должны сделать ровно $k = \frac{n+d}{2}$ прыжков вправо. Если n и d заданы, то это однозначно определяет количество прыжков вправо. Получаем типичную схему независимых испытаний: каждый прыжок независим. Если мы считаем успехом прыжок вправо, то для того, чтобы сместиться на расстояние d, нам нужно сделать $\frac{n+d}{2}$ прыжков вправо, то есть в схеме из n независимых испытаний Бернулли должно быть ровно k успехов. В написанной формуле p – вероятность прыжка вправо, q – вероятность прыжка влево. Таким образом, мы получаем, что процесс является однородным во времени, так как распределение случайной величины ξ_n определяется только d и n, то есть не зависит от начального состояния. По сути мы получили распределение координаты частицы на $n-$мм шаге.

Полученную формулу можно интерпретировать в терминах траекторий. Частица выбирает одну из возможных траекторий из $(0, m)$ в $(n, m + d)$ (первая координата – время, вторая – координата). Возможны те и только те траектории, в которых ровно $k = \frac{n+d}{2}$ смещений вправо и $n - k = \frac{n-d}{2}$ смещений влево. Вероятность прохода по одной из таких траекторий равна $p^k q^{n-k}$, так как каждый шаг вправо она делает с вероятностью p, влево – с вероятностью q, шаги независимы, соответственно будет произведение стольких p и q, сколько было шагов вправо и влево соответственно. При этом существует C_n^k таких траекторий.

Вероятность непопадания в ноль.

Пусть точка стартует из m и $m > 0$, то есть начинаем движение где-то справа от нуля. Дальше происходит положительное смещение на d единиц (через n шагов частица окажется в точке, которая тоже находится строго справа от нуля).

$$\xi_0 = m > 0, \;\; \xi_n = m + d > 0 \tag{3.9}$$

Найти вероятность того, что всюду на протяжении блужданий между точками m и $m+d$ частица ни разу не попала в ноль:

$$P_n^+(m, m+d) \overset{def}{=} P\left(\xi_n = m+d, \xi_{n-1} \neq 0, \ldots, \xi_1 \neq 0 \middle| \xi_0 = m\right) \tag{3.10}$$

Здесь нижний индекс означает количество шагов. Аргументы P^+ – начальная и конечная точки. Значок + означает, что частица не заходит в ноль. В сущности это вероятность того, что частица, стартуя из точки m, не заходила в ноль ни в какой момент времени и дошла до $m+d$. Отметим, что нам задано количество шагов – n и смещение – d. Это значит, что мы можем рассчитать количество шагов вправо. Оно равно $k = \frac{n+d}{2}$ вне зависимости от того, заходила ли частица в ноль. Соответственно, вероятность прохода по траектории из $x_0 = m$ в $x_n = m+d$ за n шагов равна $p^k q^{n-k}$. Однако в данном случае меняется количество траекторий. Нам нужны только те траектории, которые не заходят в ноль.

$$P_n^+(m, m+d) = N_n^+(m, m+d) \cdot p^k q^{n-k}, \quad k = \frac{n+d}{2} \tag{3.11}$$

где $N_n^+(m, m+d)$ – число траекторий из $(0, m)$ в $(n, m+d)$, не пересекающих горизонтальную ось и не касающихся этой оси. Таким траекторий меньше, чем всех возможных, поэтому можно записать

$$N_n^+(m, m+d) = C_n^k - N_n^-(m, m+d) \tag{3.12}$$

где C_n^k – полное число траекторий из $(0, m)$ в $(n, m+d)$, а $N_n^-(m, m+d)$ – число этих траекторий, для которых $\xi_k = 0$ хотя бы для одного $k = 1, \ldots, n$.

Найдем $N_n^-(m, m+d)$ с помощью так называемого *принципа отражения*. Сопоставим случайным блужданиям $\xi_n,\ n = 0, 1, 2, \ldots$ некоторый новый случайный процесс, который задан следующим образом. До некоторого момента n_0 новый случайный процесс равен $-\xi_k$, то есть мы отражаем значение ξ_k относительно горизонтальной оси вниз до некоторого момента n_0. А после момента n_0 новый случайный процесс в точности совпадает с прежним.

$$\widetilde{\xi_k} = -\xi_k,\ k \leqslant n_0,\ \ \widetilde{\xi_k} = \xi_k,\ \ k > n_0 \tag{3.13}$$

Здесь n_0 – первый момент времени, когда частица пришла в ноль $\xi_{n_0} = 0$.

Итак, мы считаем количество траекторий, проходящих через ноль. Значит, в какой-то момент времени, хотя бы в один, частица заходит в ноль. Обозначим через n_0 первый момент времени из всех, когда частица побывала в нуле. С точки зрения случайного процесса это минимальный номер шага, для которого $\xi_{n_0} = 0$. Данную ситуацию можно проиллюстрировать (рис. 3.2). Здесь синяя траектория – это блуждания частица из точки с координатой $m = 2$. Эта траектория проходит через ноль в момент времени $n_0 = 10$. Траектории ξ мы сопоставляем траекторию $\widetilde{\xi}$. Здесь эта траектория фиолетовая. До момента 10 эта траекториями представляет собой верхнюю траекторию, симметрично отраженную вниз.

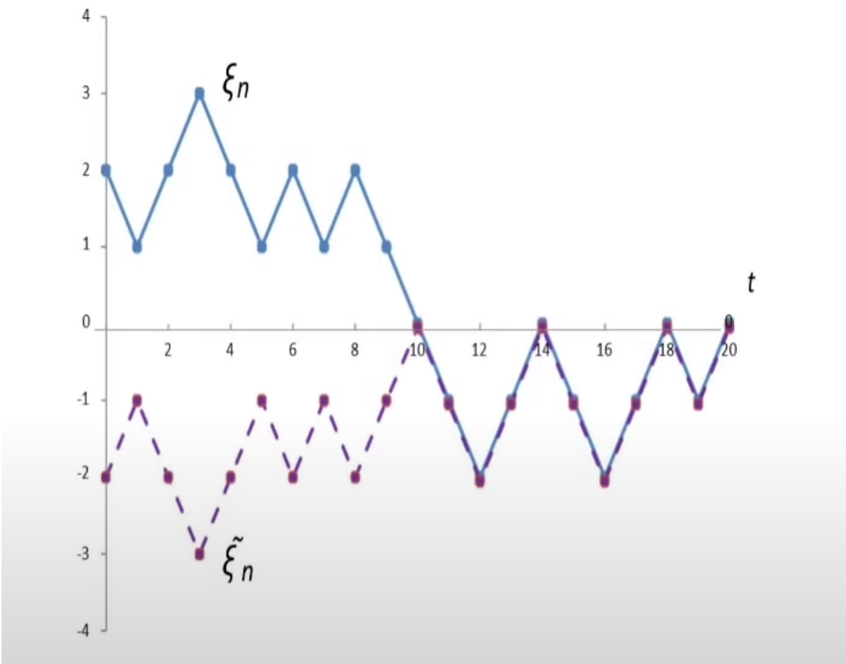

Рис. 3.2: Иллюстрация в примеру.

Соответственно, мы получили, что любой траектории, которая заходит в ноль, можно сопоставить траекторию, идущую от точки $-m$. Обратно, рассмотрим любое блуждание из $-m$ в $m + d$. Известно, что $m > 0$ и $m + d > 0$. Поскольку траектории изображаются непрерывными линиями, то обязательно в какой-то момент времени это любое блуждание пройдет через ноль. Соответственно, если такую траекторию отразить обратно наверх, то получим траекторию, заведомо проходящую через ноль. То есть, между указанными траекториями устанавливается взаимно-однозначное соответствие. То есть, блуждание за n шагов из m в $m + d$, проходящее через ноль и любое блуждание из $-m$ в $m + d$ находятся во взаимно-однозначном соответствии. Значит, количество траекторий из точки m в точку $m + d$ за n шагов, хотя бы один раз заходящих в ноль, равно общему количеству траекторий из точки $-m$ в $m + d$:

$$N_n^-(m, m+d) = C_n^{\overline{k}} \tag{3.14}$$

где

$$\overline{k} = \frac{n + 2m + d}{2} = k + m \in \{0, \ldots, n\} \tag{3.15}$$

– количество прыжков вправо, которое нужно совершить, чтобы перейти из точки $-m$ в точку $m + d$ за n шагов.

Таким образом, вероятность перехода из точки m в точку $m + d$ за n шагов без захода в ноль равна

$$P_m^+(m, m+d) = \left(C_n^k - C_n^{\overline{k}}\right) p^k q^{n-k} \tag{3.16}$$

Примем *соглашение*: коэффициент C_n^k определен только тогда, когда k принимает значения только $k \in \{0, 1, \ldots, n\}$. В противном случае будем считать, что $C_n^k = 0$. Аналогично будем считать для $C_n^{\overline{k}}$.

Модель случайных блужданий очень распространена. Например, ее можно использовать для описания азартных игр. Пусть существует некоторая игра, за кон которой мы получаем или тратим один рубль. Тогда наш капитал будет представлять случайные блуждания, если выигрыш и проигрыш утроены в соответствии с описанными выше законами. При этом условие не захода в ноль соответствует тому, что капитал остается строго положительным на протяжении всей игры.

Процесс Винера.

Попробуем перевести дискретную модель блуждания в непрерывную. Перейдем к пределу в нашей модели, но этот предел мы будем рассматривать только для симметричных случайных блужданий, то есть для тех, у которых скачок вправо и скачок влево равновероятны. Рассмотрим предел, при котором шаг по времени и шаг по пространству стремятся к нулю: $\Delta t \to 0$, $\Delta x \to 0$. При этом стремление к нулю должно быть согласовано. Потребуем, чтобы дисперсия получившегося в пределе случайного процесса была ненулевой и не была бесконечной. То есть, для фиксированного $t = n \cdot \Delta t$ при $n \to \infty$

$$0 < D\xi(t) < \infty \tag{3.17}$$

Понятно, что координата частицы есть сумма последовательных смещений частицы:

$$\xi(t) = \sum_{k=1}^{n} \Delta\xi_k \tag{3.18}$$

При этом

$$P\left(\Delta\xi_k = \Delta x\right) = P\left(\Delta\xi_k = -\Delta x\right) = \frac{1}{2} \tag{3.19}$$

Согласно нашему предположению, $\Delta\xi_1, \ldots, \Delta\xi_n$ – независимые случайный величины.

Тогда

$$M\xi_k = 0 \tag{3.20}$$

$$D\Delta\xi_k = \left(\Delta\xi_k\right)^2, \quad t = n \cdot \Delta t \tag{3.21}$$

Тогда

$$D\xi(t) = n\left(\Delta x\right)^2 = t \cdot \frac{(\Delta x)^2}{\Delta t} \tag{3.22}$$

Полученная величина не должна уходить в ноль и в бесконечность при любом фиксированном t. Чтобы добиться этого, наложим дополнительное условие

$$\frac{(\Delta x)^2}{\Delta t} \to \sigma^2, \ \ 0 < \sigma^2 < \infty \tag{3.23}$$

Тогда получим, что

$$D\xi(t) \to \sigma^2 t \tag{3.24}$$

Теорема. Пусть

$$\Delta t \to 0, \ \ \Delta x \to 0, \ \ n \to \infty, \ \ t = n\Delta t \text{ фиксировано} \tag{3.25}$$

Пусть $\xi(t)$ – координата частицы на n-м прыжке при дискретных симметричных случайных блужданиях. Шаг по времени в этих случайных блужданиях равен Δt, шаг по пространству – Δx. Тогда, если

$$\frac{(\Delta x)^2}{\Delta t} = \sigma^2, \ \ 0 < \sigma^2 < \infty \tag{3.26}$$

то по распределению

$$\xi(t) \to w(t) \tag{3.27}$$

где случайная величина

$$w(t) \in \mathbb{N}\left(0, \sigma^2 t\right) \tag{3.28}$$

имеет нормальное распределение с параметрами 0 и $\sigma^2 t$.

Замечание. Данная теорема утверждает, что если мы переходим от дискретных случайных блужданий с заданными условиями с помощью предельного перехода в каждый фиксированный момент времени t, в сущности мы делим промежуток $[0, t]$ на n одинаковых промежутков длины Δt. При этом $\Delta t = \frac{t}{n}$ и $n \to \infty$. Однако одновременно с этим мы стремим $\Delta x \to 0$ так, чтобы выполнялось (3.26). При таком предельном переходе случайная величина $\xi(t)$ будет сходиться по распределению к нормальной случайной величине с нулевым средним и дисперсией $\sigma^2 t$.

Доказательство. Имеет место центральная передельная теорема. У нас есть случайная величина $\xi(t)$ – это сумма независимых, одинаково распределённых случайных величин. Центрируем эту сумму и нормируем. Далее мы делим на $\sqrt{n}$, чтобы указанная величина имела дисперсию 1.

$$\frac{\xi(t)}{\sqrt{\sigma^2 t}} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{\Delta\xi_k - M\left(\sum\limits_{k=1}^{n} \Delta\xi_k\right)}{\sqrt{D\left(\sum\limits_{k=1}^{n} \Delta\xi_k\right)}} \tag{3.29}$$

По центральной предельной теореме полученная величина должна сходится по распределению к стандартной нормальной случайной величине

$$\frac{\xi(t)}{\sqrt{\sigma^2 t}} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{\Delta\xi_k - M\left(\sum\limits_{k=1}^{n} \Delta\xi_k\right)}{\sqrt{D\left(\sum\limits_{k=1}^{n} \Delta\xi_k\right)}} \xrightarrow[n\to\infty]{d} w \in \mathbb{N}(0,1) \tag{3.30}$$

Поэтому

$$\xi(t) \xrightarrow{d} \sqrt{\sigma^2 t} \cdot w_0(t) \in \mathbb{N}\left(0, \sigma^2 t\right) \tag{3.31}$$

Соответственно, рассмотренный предел симметричных случайных блужданий приводит нас к случайному процессу, у которого сечение распределено согласно (3.31).

Теорема доказана.

Отметим, что $\frac{\Delta x}{\Delta t}$ – мгновенная скорость движения частицы. Соответственно, $\frac{(\Delta x)^2}{(\Delta t)^2}$ – квадрат мгновенной скорости. При этом квадрат скорости должен быть равен $\frac{\sigma^2}{\Delta t}$. Тогда, если $\Delta t \to 0$, то квадрат мгновенной скорости стремится к бесконечности. То есть частица в каждый момент времени мгновенно куда-то перелетает.

Процесс Винера.

Определение. **Процесс Винера** – это случайный процесс $w(t)$, $t \geqslant 0$, который удовлетворяет следующим условиям

1. это процесс с независимыми приращениями:

   $\forall 0 < t_1 < \dots . < t_n$ $w(t_k) - w(t_{k-1})$, $k = 1, \dots, n-$независимые случайные величины $(t_0 = 0)$

2. это однородный во времени процесс, то есть распределение случайной величины $w(t+s) - w(s)$ зависит только от t.

3. начинается в нуле

4. приращение распределено нормально:

$$w(t+s) - w(s) \in \mathbb{N}(0, \sigma^2 t), \quad s, t \geqslant 0 \tag{3.32}$$

   то есть плотность вероятности

$$p_{w(t+a)-w(s)}(x) = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{1}{2\sigma^2 t}x^2}, \quad -\infty < x < \infty \tag{3.33}$$

Отметим, что при $t \to 0$ плотность стремится к дельта-функции, то есть к тому, что $w(0) = 0$ с вероятностью 1. Это означает, что предельный переход при $t \to 0$ согласован с условием 3.

Из условия 4 следует, что сечение процесса распределено нормально:

$$w(t) = w(t) - w(0) \in \mathbb{N}(0, \sigma^2 t) \tag{3.34}$$

Отсюда получаем:

$$Mw(t) = 0 \tag{3.35}$$

$$Dw(t) = \sigma^2 t \tag{3.36}$$

А ковариационная функция (для любого однородного процесса с независимыми приращениями)

$$R(t, s) = \sigma^2 \cdot min(t, s) \tag{3.37}$$

Совместное распределение случайного процесса.

Найдем n-мерное распределение. Пусть $0 = t_0 < t_1 < \ldots < t_n$. Найдем совместную функцию распределения

$$F\left(x_1, t_1; x_2, t_2; \ldots; x_n t_n\right) \stackrel{def}{=} P\left(w(t_1) < x_1, \ldots, w(t_n) < x_n\right) \tag{3.38}$$

Так как у нас есть независимость приращений, разумно перейти к приращениям. Дальше рассчитываем вероятность как интеграл от совместной плотности приращений

$$F\left(x_1, t_1; x_2, t_2; \ldots; x_n t_n\right) \stackrel{def}{=} P\left(w(t_1) < x_1, \ldots, w(t_n) < x_n\right) =$$
$$= P\left(\Delta w < x_1, \ldots, \Delta w_1 + \Delta w_n < x_n\right) = \int\limits_{z_1 < x_1} \ldots \int\limits_{z_1 + \ldots + z_n < x_n} p_1(z_1) \ldots p_n(z_n) dz_1 \ldots dz_n \tag{3.39}$$

где $p_k(\cdot)$ – плотности вероятности приращения Δw_k. Они известны, так как речь идет о процессе Винера

$$p_k(z) = \frac{1}{\sqrt{2\pi\sigma^2\left(t_k - t_{k-1}\right)}} e^{-\frac{z^2}{2\sigma^2 |t_k - t_{k-1}|}}, \quad -\infty < z < \infty \tag{3.40}$$

Преобразуем полученный ответ. Совершим линейную замену переменных

$$y_1 = z_1, \; y_2 = z_1 + z_2, \ldots, y_n = z_1 + \ldots + z_n \tag{3.41}$$

Якобиан перехода равен единице. Обратное преобразование имеет вид

$$z_1 = y_1 = \Delta y_1, \; z_2 = y_2 - y_1 = \Delta y_2, \; \ldots, \; z_n = y_n - y_{n-1} = \Delta y_n \tag{3.42}$$

Нам нужно посчитать вероятность, ассоциированную с сечениями. От сечений перешли к приращениям. Записали интеграл через плотности вероятности приращений. Дальше преобразуем область и тем самым возвращаемся от приращений к сечениям. Тогда получаем

$$\int\limits_{x_1<x_1} \dots \int\limits_{z_1+\ldots+z_n<x_n} p_1(z_1)\dots p_n(z_n)dz_1\dots dz_n = \int\limits_{y_1<x_1} \dots \int\limits_{y_n<x_n} p_1(\Delta y_1)\dots p_n(\Delta y_n)dy_1\dots dy_n \tag{3.43}$$

Тогда функция распределения

$$F\left(x_1, t_1; \ldots; x_n, t_n\right) = \int\limits_{-\infty}^{x_1} \dots \int\limits_{-\infty}^{x_n} p_1(\Delta y_1)\dots p_n(\Delta y_n)dy_1\dots dy_n \tag{3.44}$$

Таким образом, функция распределения есть n-кратный интеграл, под знаком которого стоит совместная плотность вероятности сечения. Если положить $\Delta x_k = x_k - x_{k-1}$ для $k = 1, \ldots, n$ и $x_0 = 0$

$$p\left(x_1, t_1; \ldots; x_{n-1}, t_{n-1}; x_n, t_n\right) = p_1(\Delta x_1)\dots p_n(\Delta x_n) \tag{3.45}$$

В явном виде

$$p\left(x_1, t_1; x_2, t_2; \ldots; x_n, t_n\right) = \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2(t_k - t_{k-1})}} e^{-\frac{(x_k - x_{k-1})^2}{2\sigma^2(t_k - t_{k-1})}} \tag{3.46}$$

для любых $x_1, \ldots, x_n \in \mathbb{R}$ и $0 = t_0, t_1 < \ldots < t_n$, $t_0 = 0$ и $x_0 = 0$.

Все полученные формулы получаются на основе предположения, что все приращения независимы и имеют известное нам распределение.

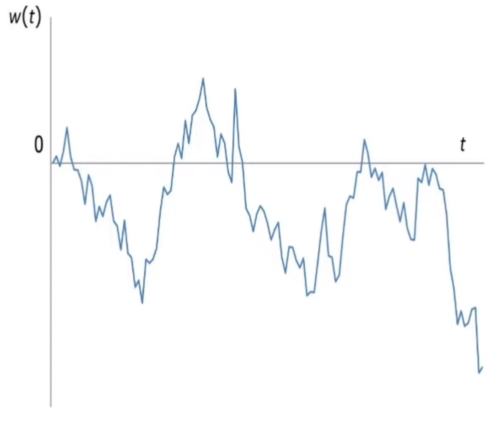

Рис. 3.3: Траектория Винеровского процесса.

Траектории Винеровского процесса

Пример траектории винеровского процесса изображён на рис. 3.3. Отметим, что у частицы бесконечно большая скорость. С точки зрения траектории это означает, что траектория в каждой точке имеет вертикальную касательную. Таким образом, получаем траекторию, непрерывную в каждой точке, но при этом недифференцируемую ни в какой точке.

Лекция 4. Процесс Винера.

Процесс Винера.

Определение. Процесс Винера – этьо случайный процесс, удовлетворяющий следующим условия

1. он имеет независимые приращения
2. является однородным во времени
3. начинается в нуле
4. приращение
$$\Delta w(t-s) = w(t) - w(s) \in \mathbb{N}\left(0, \sigma^2(t-s)\right) \tag{4.1}$$
то есть
$$p_{\Delta w(t-s)}(x) = \frac{1}{\sqrt{2\pi\sigma^2(t-s)}} e^{-\frac{x^2}{2\sigma^2(t-s)}}, \quad -\infty < x < \infty, t > 0 \tag{4.2}$$
следовательно
$$w(t) = w(t) - w(0) \in \mathbb{N}\left(\sigma^2 t\right) \tag{4.3}$$

Броуновское движение хорошо описывается процессом Винера. Частица совершает *одномерное броуновское движение*, если её координата описывается процессом Винера $w(t), \ \ t > 0$.

Время первого попадания в точку.

Рассмотрим некоторые другие характеристики, описывающие поведение траектории этого процесса. Для начала рассмотрим *время первого попадания в точку* $x > 0$. Рассмотрим следующую ситуацию. Пусть есть некоторая броуновская частица. Она выходит из нуля. Дальше она начинает двигаться по числовой прямой. Выделим на этой прямой некоторую точку $x > 0$ и найдем момент времени, когда частица в первый раз дошла до этой точки. Для определенности будем считать x положительным. Нормальное распределение симметрично, поэтому решение для отрицательных x будет аналогично (вероятности будут те же).

Пусть $w(t), t \geqslant 0$ – процесс Винера, $x > 0$ фиксирован. Пусть случайная величина $\tau(x)$ равна времени, когда броуновская частица в первый раз пришла в точку x:

$$\tau(x) = t, \text{ если } w(s) < x \text{ для всех } 0 \leqslant s < t \text{ и } \ w(t) = x \tag{4.4}$$

Будем искать распределение заданной случайной величины $\tau(x)$ при фиксированном $x > 0$. Понятно, что искать мы будем при $t > 0$ и скорее всего распределение этой случайной величины будет абсолютно непрерывным. Для того, чтобы найти это распределение, воспользуемся следующим тождеством (без доказательства)

$$P\left(w(t) < x \middle| t > \tau(x)\right) = P\left(w(t) > x \middle| t > \tau(x)\right) = \frac{1}{2} \tag{4.5}$$

Здесь мы рассматриваем ситуацию, в которой частица хотя бы один раз побывала в точке x (так как $t > \tau(x)$). Это мы знаем точно. Мы рассматриваем вероятности того, будет ли частица больше или меньше x в эти поздние моменты времени. Известно, что вероятность того, что $w(t) = x$ равна нулю. Поэтому можно считать, что указанные две вероятности противоположны. Мы принимаем, что эти две вероятности равны. Это означает, что после точки $(\tau(x), x)$ траектории распределены симметрично относительно x. Это условие можно сравнить со следующим условием

$$P\left(w(t) < 0\right) = P\left(w(t) > 0\right) = \frac{1}{2} \tag{4.6}$$

Тогда распределение траектории симметрична относительно нуля, так как $w(t) \in \mathbb{N}\left(0, \sigma^2 t\right)$.

Поясним сказанное с помощью иллюстрации (рис. 4.1).

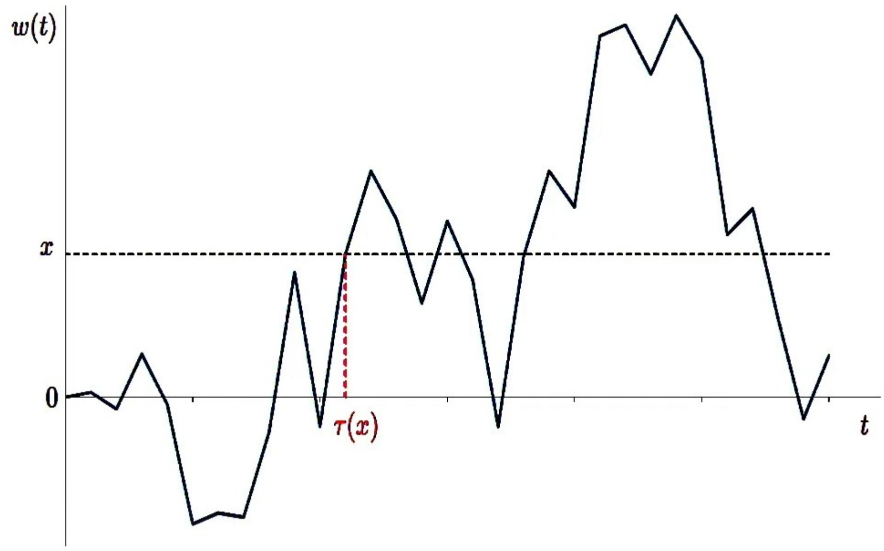

Рис. 4.1: Иллюстрация к объяснению.

Траектория, указанная на рис. 4.1 – траектория броуновского движения. Далее ставим некоторый уровень x – фиксированная горизонтальная кривая. Нас интересует первый раз, когда траектория пересекла этот уровень. Далее мы рассматриваем часть траекторий после этого момента. Тогда будем считать, что распределение траекторий симметрично относительно данной горизонтальной оси. Безусловная вероятность симметрична относительно нуля (в каждой точке сечение распределено симметрично относительно нуля), однако мы рассматриваем условную вероятность, которая симметрична относительно уровня x.

Таким образом, для вероятности имеем

$$P\left(w(t) < x \middle| t > \tau(x)\right) = \frac{1}{2} \tag{4.7}$$

Заметим, что имеет место взаимосвязь

$$\tau(x) \geqslant t \implies w(t) \leqslant x \tag{4.8}$$

то есть правое событие влечёт второе. Событие $\tau(x) \geqslant t$ означает, что в момент, когда частица достигла уровня x, находится или в t или позже. Значит, момент времени, когда частица достигла уровня x, находится или в точке t на числовой оси или позже. Если $\tau(x) = t$, то $w(t) = x$. Если $\tau(x) > t$, значит, к моменту времени t мы еще не дошли до уровня x, значит, $w(t) < x$. Поэтому

$$w(t) > x \implies \tau(x) < t \tag{4.9}$$

Траектория непрерывна и начинается в нуле. В момент времени t частица оказывается правее x, значит, когда-то раньше она уже была в точке x.

Подставим полученное следствие в выражение для условной вероятности. Тогда получим

$$\frac{1}{2} = P\left(w(t) > x \middle| \tau(x) < t\right) = \frac{P\left(w(t) > x, \tau(x) < t\right)}{P\left(\tau(x) < t\right)} = \frac{P\left(w(t) > x\right)}{P\left(\tau(x) < t\right)} \tag{4.10}$$

Здесь первое равенство есть само предположение (4.5). Далее пишем формулу для условной вероятности. Далее пользуемся тем, что первое из событий под знаком условной вероятности влечёт за собой второе, значит второе выполнено автоматически и его можно не писать. Отсюда получаем выражение для функции распределения

$$F_\tau(t; x) = 2P\left(w(t) > x\right) \tag{4.11}$$

Нам известно распределение $w(t)$, поэтому

$$F_\tau(t;x) = 2P\left(w(t) > x\right) = 2\int\limits_x^\infty \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{x^2}{2\sigma^2 t}} dz = 2\int\limits_{\frac{z}{\sqrt{\sigma^2 t}}}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du =$$

$$= 2\left(1 - \int\limits_{-\infty}^{\frac{x}{\sqrt{\sigma^2 t}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du\right) = 2\left(1 - \Phi\left(\frac{x}{\sqrt{\sigma^2 t}}\right)\right) \tag{4.12}$$

где $\Phi(\cdot)$ – интеграл вероятности. Для $-\infty < a < +\infty$ он задается как интеграл от плотности стандартного нормального распределения

$$\Phi(a) = \int\limits_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \tag{4.13}$$

Соответственно, его производная есть сама плотность

$$\Phi'(a) = \frac{1}{\sqrt{2\pi}} e^{-\frac{a^2}{2}} \tag{4.14}$$

Свойства этой функции аналогичны свойствам функции распределения. На $-\infty$ она равна нулю, на $+\infty$ она равна единице, функция неубывает, является гладкой функцией, бесконечное число раз недифференцируемой, ее график центрально-симметричен.

Таким образом, получаем, что функция распределения случайной величины τ как функция от t (x – параметр) имеет вид

$$F_\tau(\tau; x) = 2\left(1 - \Phi\left(\frac{x}{\sqrt{\sigma^2 t}}\right)\right) \tag{4.15}$$

Единственный возможный разрыв этой функции в нуле. Устремим t к нулю справа. Тогда аргумент Φ стремится к $+\infty$. Тогда $\Phi \to 1$. Соответственно, функция распределения стремится к нулю. Соответственно, точка $t = 0$ не вызывает никаких проблем. Функция распределения всюду непрерывна. Значит, можно найти ее плотность вероятности

$$p_\tau(t; x) = \frac{\partial F_\tau(t; x)}{\partial t} = -2\Phi'(a)\Big|_{a=\frac{x}{\sqrt{\sigma^2 t}}} \cdot \frac{\partial}{\partial t} \frac{x}{\sqrt{\sigma^2 t}} = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{x^2}{2\sigma^2 t}} \cdot \frac{x}{t} \tag{4.16}$$

Рассмотрим полученную функцию не как функцию от x, как функцию от t. По t плотность имеет порядок $t^{-\frac{3}{2}}$ и имеет особенность в нуле.

Посчитаем вероятность того, что время достижения уровня конечно

$$P\left(\tau(x) < \infty\right) = \int\limits_0^\infty p_\tau(t; x) dt \tag{4.17}$$

После подставления плотности получим

$$P\left(\tau(x) < \infty\right) = \int\limits_0^\infty p_\tau(t; x) dt = \int\limits_0^\infty \frac{x}{t} \cdot \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{x^2}{2\sigma^2 t}} dt = \int\limits_0^\infty \frac{1}{\sqrt{2\pi\beta^2}} t^{-\frac{3}{2}} e^{-\frac{1}{t}\frac{1}{2\beta^2}} dt, \quad \beta^2 = \frac{\sigma^2}{x^2} \tag{4.18}$$

Сделаем замену переменной

$$\frac{1}{t} = u^2 \in (\infty, 0) \tag{4.19}$$

Тогда пределы интегрирования меняются местами и

$$t = u^{-2}, \;\; dt = -2u^{-3}du, \;\; t^{-\frac{3}{2}} = u^3 \tag{4.20}$$

Тогда для любого фиксированного $x > 0$ получаем

$$P\left(\tau(x) < \infty\right) = \int\limits_0^\infty \frac{1}{\sqrt{2\pi\beta^2}} u^3 e^{-\frac{u^2}{2\beta^2}} 2u^{-3} du = 2\int\limits_0^\infty \frac{1}{\sqrt{2\pi\beta^2}} e^{-\frac{u^2}{2\beta^2}} du = 1 \tag{4.21}$$

Таким образом, любой уровень $x > 0$ достижим за конечное время. Эту идею можно рассмотреть с точки зрения частот. Пусть есть огромное количество траекторий движения броуновских частиц. Далее мы фиксируем некоторый уровень x и смотрим, какие траектории достигли уровня x за конечное время. Оказывается, что этому условию будут удовлетворять почти все траектории ансамбля.

Вычислим математическое ожидание времени достижения уровня $x > 0$

$$M\tau(x) = \int\limits_0^\infty t \cdot \frac{x}{t} \cdot \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{x^2}{2\sigma^2 t}} dt \tag{4.22}$$

После вычисления интеграла получим (в $C(x)$ находится все, что не зависит от t)

$$M\tau(x) = \int\limits_0^\infty t \cdot \frac{x}{t} \cdot \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{x^2}{2\sigma^2 t}} dt = C(x) \cdot \int\limits_0^\infty \frac{1}{\sqrt{t}} e^{-\frac{x^2}{2\sigma^2 t}} dt \tag{4.23}$$

Отметим, что экспонента под знаком интеграла является конечной величиной и при $t \to 0$ и при $t \to \infty$. Однако, есть множитель $\frac{1}{\sqrt{t}}$. В нуле эта особенность интегрируема. Однако, на бесконечности этот интеграл ведет себя как

$$\frac{1}{\sqrt{t}} e^{-\frac{x^2}{2\sigma^2 t}} = O\left(t^{-\frac{1}{2}}\right) \text{ при } t \to \infty \tag{4.24}$$

Таким образом, мы имеем дело с расходящимся интегралом. По порядку величины подинтегральная функция на бесконечности ведет себя как $t^{-\frac{1}{2}}$ (медленно убывающая функция). Эта функция убывает не настолько быстро, чтобы дать конечное значение интеграла. Таким образом, математическое ожидание времени достижения уровня бесконечно

$$M\tau(x) = \infty \tag{4.25}$$

Полаем следующий парадоксальный факт. Каждый уровень достигается с вероятностью единица за конечное время. Однако, среднее время достижения любого сколь угодно малого уровня бесконечно. Отметим, что у плотности вероятности на хвосте (то есть при $t \to \infty$) довольно большие значения. Это значит, что среди всех траекторий существует очень много траекторий, у которых время достижения уровня большое. Из-за этого "тяжелого хвоста"плотности математическое ожидание времени достижения уровня оказывается равным бесконечности.

Задача.

Пусть $t > 0$, $t = fix$. Введем случайный процесс

$$w^*(t) = \max_{0\leqslant x\leqslant t} w(s) \tag{4.26}$$

По определению этот процесс равен максимальному значению траектории на промежутке $[0, t]$. Здесь мы фиксируем t, а не s и смотрим распределение максимума траекторий на сегменте.

Найти функцию распределения

$$F^*(x;t) = P(w^*(t) < x) \tag{4.27}$$

То есть, мы ищем, как распределена самая далёкая точка от нуля, в которой побывала частица.

Известно, что траекторию непрерывна, если рассматривать функцию $w(\cdot)$ как функцию от s. Поэтому максимум на сегменте достигается. Рассмотрим события

$$w^*(t) < x \implies w(s) < x \text{ для всех } s \in [0,t] \implies \tau(x) > t \tag{4.28}$$

Если максимум меньше x, значит, любое значение меньше x в каждой точке сегмента. Отсюда следует, что время, когда частица побывала в точке x, будет позже, чем момент времени t. Соответственно, первое из указанных событий влечет за собой последнее.

Верно обратное

$$\tau(x) > t \implies w(s) < x \text{ для всех } s \in [0,t] \implies w^*(t) < x \tag{4.29}$$

Если $\tau(x) > t$, значит, момент, когда частица первый раз оказалась в точке x, наступил позже, чем t. Значит, до момента времени t, включая сам этот момент, частица еще не была в точке x. То есть ее координата меньше x. Каждая $w(s) < x$, следовательно, максимум тоже меньше x.

Таким образом, получили, что событие $w^*(t) < x$ эквивалентно с точки зрения случайности событию $\tau(x) > t$. Поэтому

$$F^*(x;t) = P(w^*(t) < x) = P(\tau(x) > t) = 1 - P(\tau(x) \leqslant t) = 1 - P(\tau(x) < t) \tag{4.30}$$

Учтем, что $\tau(\cdot)$ распределено абсолютно непрерывно, в последнем равенстве переходим от нестрогого неравенства к строгому. В последнем выражении по сути стоит функцию распределения $\tau(\cdot)$. Поэтому

$$F^*(x;t) = 1 - 2\left(1 - \Phi\left(\frac{x}{\sqrt{\sigma^2 t}}\right)\right) = 2\Phi\left(\frac{x}{\sqrt{\sigma^2 t}}\right) - 1 \tag{4.31}$$

Отметим, что функция распределения $F^*(x;t)$ – функция от x, t здесь параметр. Но аналитическое выражение для этой функции выражается через функцию распределения τ в точке t, то есть t – естественный аргумент функции распределения.

Найдем плотность распределения

$$p^*(x;t) = \frac{\partial F^*(x;t)}{\partial x} = 2\frac{1}{\sqrt{2\pi}}e^{-\frac{a^2}{2}}\Big|_{a=\frac{x}{\sqrt{\sigma^2 t}}} \cdot \frac{1}{\sqrt{\sigma^2 t}} = 2\cdot\sqrt{\frac{1}{2\pi\sigma^2 t}}e^{-\frac{x^2}{2\sigma^2 t}} \tag{4.32}$$

Таким образом, для $x > 0$, $t > 0$ плотность как функция от x и t как параметра имеет вид удвоенной нормальной плотности сечения $w(t)$. То есть, плотность вероятности максимума лежит выше плотности $w(t)$.

Отметим, что параметр t здесь задает отрезок, на котором вычисляется максимум. Попробуем посчитать вероятность того, что максимум меньше или равен нуля $(t > 0)$:

$$P\left(w^*(t) \leqslant 0\right) = \lim_{x\to+0} F^*(x;t) \tag{4.33}$$

Подставим полученное выражение для функции распределения

$$P\left(w^*(t) \leqslant 0\right) = \lim_{x\to+0} 2\Phi\left(\frac{x}{\sqrt{\sigma^2 t}}\right) - 1 = 2\Phi(0) - 1 = 0 \tag{4.34}$$

Если мы стремим $x \to +0$, то $\Phi(0) \to \frac{1}{2}$.

Таким образом, максимум не принимает отрицательные значения и ноль. Значит,

$$P\left(w^*(t) > 0\right) = 1 \tag{4.35}$$

то есть на любом промежутке $[0, t]$ максимальное значение координаты положительно с вероятностью 1. Например, берем некоторый сколь угодно малый промежуток времени и чуть-чуть сдвигаемся от нуля. За это очень малое время частица уже успела побывать в точке с положительной координатой. Этот факт согласуется с тем, что траектория вине ровского процесса – очень дрожащая кривая, у нее нет производной ни в какой точке. На любом сколько угодно малом промежутке частица посещает точки с положительными координатами. Аналогичный результат можно получить и ля случая $x < 0$, так как траектории симметричны относительно нуля. Таким образом, частица на сколько угодно малом промежутке времени посещает точки и с положительными, и с отрицательными координатами. В результате получаем очень зазубренную траекторию.

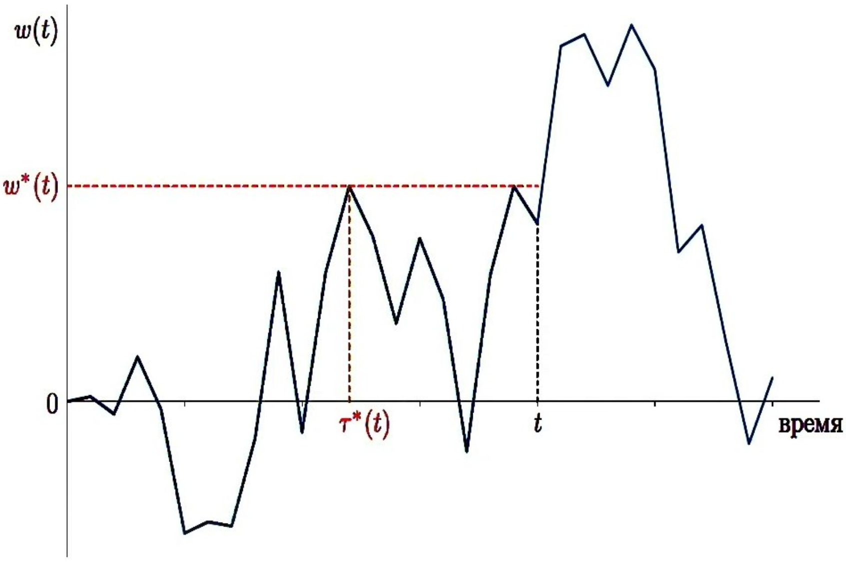

Рис. 4.2: Иллюстрация к объяснению.

Проиллюстрируем сказанное с помощью рис. 4.2. На рис. 4.2 нарисована траекторию движения броуновской частицы. Рассмотрим только ту часть траектории, которая лежит до момента времени t. Рассмотрим максимум данной траектории на этом промежутке. Нас интересует распределение максимума. Тогда получаем, что почти все траектории (за исключением тех, которые реализуются с вероятностью ноль) имеют положительный максимум на отрезке $[0;t]$. Здесь $\tau^*(t)$ – момент первого достижения максимального значения на сегменте $[0;t]$. Рассмотрим распределение данной случайной величины. Она описывается законом арксинуса.

Закон арксинуса.

Пусть случайная величина $\tau^*(t)$ равна первому моменту времени на промежутке $[0;t]$, когда траектория достигла своего первого максимального значения. Тогда плотность вероятности и функция распределения этой случайной величины

$$p_{\tau^*}(u) = \frac{1}{\pi} \cdot \frac{1}{\sqrt{u(t-u)}}, \quad 0 < u < t \tag{4.36}$$

а ее функция распределения

$$F_{\tau^*}(u) = \frac{2}{\pi} \arcsin\sqrt{\frac{u}{t}}, \quad 0 < u < t \tag{4.37}$$

Отметим, что плотность неограничена на краях диапазона изменений u. С точностью до du плотность это вероятность того, что τ^* примет значение в точке u. Оказывается, что велика вероятность того, что u будет близко к нулю или к t. Соответственно, максимум траектории достигается скорее на концах отрезка чем в середине. При этом плотность вероятности симметрична относительно середины (точка $\frac{t}{2}$).

Замечание. Ранее мы говорили о том, что существует способ решения задач, заключающийся в отражении траекторий случайных блужданий и замене исходного случайного процесса блужданий на эквивалентный ему по статистическим свойствам "отраженный"процесс (*принцип отражения*). Аналогичный принцип можно использовать и в случае винеровского процесса.

Принцип отражения. Если $w(t)$, $t \geqslant 0$ – винеровский процесс, $\tau(x)$ – время первого достижения траекторией уровня $x > 0$, то существует процесс

$$\widetilde{w}(t) = \begin{cases} w(t), \quad 0 \leqslant t \leqslant \tau(x) \\ 2x - w(t), t \geqslant \tau(x) \end{cases} \tag{4.38}$$

вторая строка здесь по сути означает симметричное отражение относительно уровня x. Суть заключается в следующем. До момента времени $\tau(x)$ процессы $\widetilde{w}(t)$ и $w(t)$ ведут себя одинаково. А после момента времени $\tau(x)$ траектория $\widetilde{w}(t)$ симметрично отражена относительно уровня x.

Такой процесс $\widetilde{w}(t)$ стохастически эквивалентен процессу $w(t)$. Это означает, что все вероятностные характеристики процесса $\widetilde{w}(t)$ такие же, как и у процесса $w(t)$.

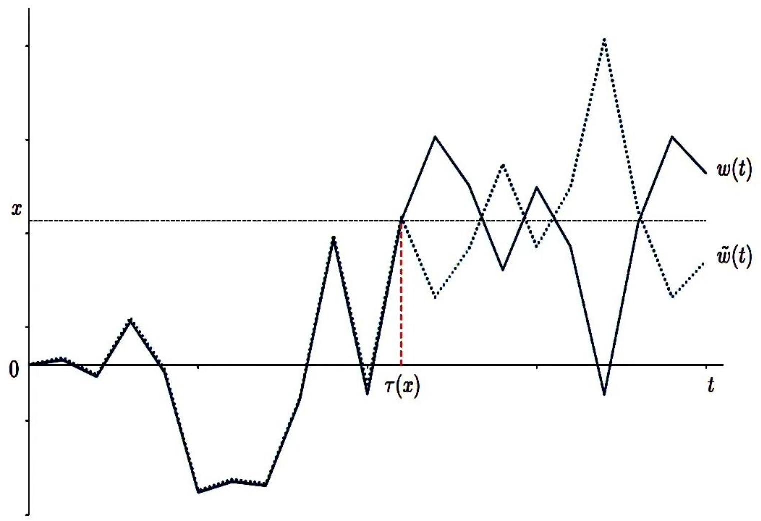

Рис. 4.3: Принцип отражения.

Проиллюстрируем сказанное рис. 4.3. Сплошная линия – $w(t)$. Далее мы ставим некоторый уровень x и смотрим момент первого достижения этого уровня. Это приходит в момент времени $\tau(x)$. До этого момента сплошная и пунктирная траектории совпадают. А далее симметрично отражаем траекторию относительно продолженного уровня x. Видно, что для сплошной и пунктирной траекторий момент первого достижения уровня x одинаковый.

Итоги.

Процесс Винера – это процесс с непрерывным временем и непрерывным множеством значений, который получается как предел последовательности дискретных случайных блужданий, которые случаются через промежутки времени Δt как прыжки на расстояние Δx влево или вправо. Тогда, при определённом соотношении на Δx и Δt получаем непрерывный предел по времени и по пространству – винеровский процесс. Все сечения этого процесса распределены нормально. Дисперсия пропорциональная времени, в котором мы рассматриваем сечение.

Лекция 5. Марковский случайный процесс.

Марковский процесс.

Приведём определение из теории вероятностей. Рассмотрим последовательность дискретно распределённых случайных величин $\xi_0, \xi_1, ..$, каждая из которых принимает значение из множества $X = \{x_1, ..., x_r\}$. Распределение любой из этих случайных величин задается формулами

$$\sum_{j=1}^{r} P\left(\xi_k = x_j\right) = 1 \tag{5.1}$$

– случайные величины ξ_k никаких значений, кроме x_j, не принимают. При этом мы допускаем, что некоторые из вероятностей могут быть нулевыми (хотя бы одна ненулевая)

$$P\left(\xi_k = x_j\right) \geqslant 0 \tag{5.2}$$

Здесь $j = 1, \ldots, r$.

Определение. Описанная выше последовательность называется **цепью Маркова с состояниями** $x_1, \ldots, x_r$ $(2 \leqslant r < \infty)$, если для любого $n \geqslant 1$ и любых $x_{i_0}, \ldots, x_{i_{n-1}}, x_{i_n}$ выполнено равенство

$$P\left(\xi_n = x_{i_n} | \xi_{n-1} = x_{i_{n-1}}, .., \xi_0 = x_{i_0}\right) = P\left(\xi_{i_n} = x_{i_n} | \xi_{n-1} = x_{i_{n-1}}\right) \stackrel{def}{=} \pi_{i_{n-1}, i_n} \tag{5.3}$$

В левой части этого равенства стоит выражение для условного распределения случайной величины ξ_n при условии, что все предыдущие значения фиксированы. То есть, мы налагаем требование, чтобы данное условное распределение совпадало с условным распределением этой же случайной величины при условии фиксирования только одно (предыдущее) значение ξ_{n-1}. Такое условие называется *условием Маркова*.

Последовательность случайных величин – это процесс с дискретным временем. Далее рассмотрим процесс с непрерывным временем.

Определение. Случайный процесс $\xi(t), \quad t \geqslant 0$ называется **марковским процессом с состояниями** $x_1, \ldots, x_r$, если для любых $n \geqslant 1$ и любых упорядоченных моментов времени $0 = t_0 < t_1 < .... < t_n \quad (n = 2, 3, ...)$ случайные величины, равные сечениям процесса в моменты времени $\xi_0 = \xi(t_0), \xi_1 = \xi(t_1), \xi_2 = \xi(t_2), ..., \xi_n = \xi(t_n)$ удовлетворяет условию марковости

$$P\left(\xi_n = x_{i_n} | \xi_{n-1} = x_{i_{n-1}}, .., \xi_1 = x_{i_1}\right) = P\left(\xi_{i_n} = x_{i_n} | \xi_{n-1} = x_{i_{n-1}}\right) \tag{5.4}$$

То есть это процесс, у которого любые n последовательных сечений образуют цепь Маркова. Далее будем считать, что число состояний r конечно.

Определение. Условная вероятность

$$P\left(\xi(t) = x_j \middle| \xi(s) = x_i\right) \tag{5.5}$$

для $0 \leqslant s < t$ называется **вероятностью перехода** из $i-$го состояния в $j-$е за время $t - s > 0$.

Таким образом, если рассматриваемый процесс описывает некоторую динамическую систему, то мы точно знаем, что в момент времени s эта система находилась в $i-$м состоянии и нас интересует вероятность того, что за время $t - s$ эта система перейдет из состояния с номером i в состояние с номером j.

Сделаем дополнительное предположение. Будем считать, что эта вероятность перехода зависит не от двух времен, а только от их разности

$$P\left(\xi(t) = x_j | \xi(s) = x_i\right) = \pi_{ij}(t - s) \tag{5.6}$$

что соответствует однородности вероятности перехода. Положим по определению

$$\pi_{ij}(0) = \delta_{ij} \tag{5.7}$$

в силу того, что $\pi_{ij}(0)$ соответствует ситуации, когда моменты времени в аргументе и в условии вероятности совпадают. Понятно, что если $i \neq j$, то события до и после черты условной вероятности несовместны

$$P\left(\xi(t) = x_j \middle| \xi(t) = x_i\right) = 0 \text{ при } i \neq j \tag{5.8}$$

Если же $i = j$, то вероятность равна единице

$$P\left(\xi(t) = x_j \middle| \xi(t) = x_i\right) = 1 \text{ при } i = j \tag{5.9}$$

Определение. Матрица $\pi(t)$ размера $r \times r$ с элементами

$$\pi_{ij}(t) = P\left(\xi(s+t) = x_j \middle| \xi(s) = x_i\right), \quad i, j = 1, \ldots, r \tag{5.10}$$

называется **матрицей перехода за время t**. При этом

$$\pi(0) = 1 \tag{5.11}$$

Таким образом, получаем некоторый набор матриц, которые различны в разных t. эти матрицы перехода являются одной из двух составляющих, которые полностью определяют распределение марковского процесса.

Случайный процесс полностью задается семейством своих конечномерных распределений. Поэтому при задании какого-то процесса мы в первую очередь ищем совместное распределение. Для марковского процесса любое конечномерное распределение полностью определяется двумя наборами данных: матрицами перехода $\pi(t)$, $t > 0$ и начальными вероятностями (то есть распределением процесса в начальный момент времени)

$$P\left(\xi(0) = x_i\right) = a_i, \quad i = 1, \ldots, r \tag{5.12}$$

где

$$\sum_{i=1}^{r} a_i = 1 \tag{5.13}$$

n-мерное распределение марковского процесса: для $0 = t_0 < t_1 < \ldots < t_n$ положим (для краткости)

$$\xi(t_k) = \xi_k, \quad k = 0, 1, \ldots, n \tag{5.14}$$

то есть сечение в момент времени t_k это просто $k-$я случайная величина.

Тогда $n-$мерное распределение задается совместной вероятностью

$$P\left(\xi_0 = x_{i_0}, \xi_1 = x_{i_1}, \ldots, \xi_{n-1} = x_{i_{n-1}}, \xi_n = x_{i_n}\right) \tag{5.15}$$

Эту совместную вероятность мы будем искать по стандартной формуле теории вероятностей (вероятность пересечения расписывается через условные вероятности). Вероятность пересечения есть вероятность последнего события при условии, что фиксированы все остальные. Однако, чтобы равенство было верным, нужно добавить множители в виде вероятностей всех остальных событий. Вероятность для $n-1$ случайной величины расписываем по аналогичной структуре и так далее. В результате получаем

$$\begin{gathered} P\left(\xi_0 = x_{i_0}, \xi_1 = x_{i_1}, \ldots, \xi_{n-1} = x_{i_{n-1}}, \xi_n = x_{i_n}\right) = \\ = P\left(\xi_n = x_{i_n} \middle| \xi_{n-1} = x_{i_{n-1}}, \ldots, \xi_0 = x_{i_0}\right) \times P\left(\xi_{n-1} = x_{i_{n-1}} \middle| \xi_{n-1} = x_{i_{n-2}}, \ldots, \xi_0 = x_{i_0}\right) \times \ldots \\ \ldots \times P\left(\xi_1 = x_{i_1} \middle| \xi_0 = x_{i_0}\right) \times P\left(\xi_0 = x_{i_0}\right) \end{gathered} \tag{5.16}$$

Далее воспользуемся марковским условием. Формулы, написанные красным в (5.16) можно выбросить из общего выражения, так как для набора случайных величин ξ_k выполнено марковское условие. Таким образом, если мы ищем распределение случайной величины ξ_n при фиксированных предыдущих случайных величинах, то все предыдущие, кроме ξ_{n-1}, можно выбросить. Соответственно, в каждой условной вероятности останется только одно событие после черты условной вероятности. Тогда вместо условных вероятностей мы можем писать вероятности перехода, а самый последний множитель – начальная вероятность. Тогда, в терминах вероятностей перехода и начальных вероятностей получим

$$P\left(\xi_0 = x_{i_0}, \xi_1 = x_{i_1}, \ldots, \xi_{n-1} = x_{i_{n-1}}, \xi_n = x_{i_n}\right) = \\ = a_{i_0} \cdot \pi_{i_0 i_1}(\Delta t_1) \ldots \pi_{i_{n-1} i_{n-1}}(\Delta t_{n-1})\, \pi_{i_{n-1} i_n}(\Delta t_n) \tag{5.17}$$

где

$$\Delta t_k = t_k - t_{k-1} \quad k = 1, \ldots, n \tag{5.18}$$

(5.17) есть совместная вероятность значений сечений. Отметим, что сама формула представляет собой произведение $n-1$ сомножителя. При этом в каждых двух соседних множителях правый индекс левого множителя такой же, как левый индекс следующего множителя. То есть структура индексов имеет цепочечный вид. Отсюда и термин – цепь Маркова.

Отметим также, что сечения случайного процесса *не являются* независимыми случайными величинами. Да, вероятность произведения есть произведение каких-то вероятностей, но эти множители не зависят отдельно от индексов каждого сечения. Они зависят от двух индексов двух соседних сечений. Поэтому вероятность произведения не является произведением вероятностей отдельных событий, зависящих только от одного индекса. Таким образом, сечения марковского процесса не являются независимыми. С другой, стороны, вероятность зависит от разности времен.

Матрица перехода.

Перейдем к изучению марковского процесса. Для этого нам нужно знать начальные вероятности и матрицы перехода. Будем рассматривать матрицы перехода как семейство матичнозначных функций

$$\pi(t) : \mathbb{R}_+ \to Mat_r, t \geqslant 0 \tag{5.19}$$

где Mat_r – множество матриц размера $r \times r$.

Отметим, что в нуле

$$\pi(0) = I \tag{5.20}$$

Условимся, что в $\pi_{ij}(t)$ i – номер строки, j – номер столбца.

Обсудим **свойства матриц перехода**:

1. свойство стохастичности матрицы для любого $t > 0$. Это значит, что для любой матрицы перехода все ее элементы неотрицательны $\pi_{ij}(t) \geqslant 0$ и сумма элементов вдоль любой строки матрицы $\pi(t)$ равна 1

$$\sum_{j=1}^{r} \pi_{ij}(t) = 1, \quad i = 1, \ldots, r \tag{5.21}$$

В сущности это нормировка условного распределения

$$\pi_{ij}(t) = P\left(\xi(t) = x_j \middle| \xi(0) = x_i\right), \quad j = 1, \ldots, r \tag{5.22}$$

Здесь $x_i = fix$. Вероятности (5.22) задают распределение сечения в момент времени t.

Таким образом, стохастичность матрицы – это фундаментальное свойство вероятности, которое говорит о том, что сумма всех вероятностей полной группы событий равна единице.

2. Для матриц перехода выполняется **уравнение Чепмена-Колмогорова**:

$$\pi\,(t+s) = \pi(s)\cdot\pi(t) = \pi(t)\cdot\pi(s), \quad t, s \geqslant 0 \tag{5.23}$$

Отсюда в том числе следует, что матрицы перехода при разных временах коммутируют.

Само уравнение доказывается из простейших формул теории вероятности и марковского свойства. Рассмотрим случай $0 < s < t + s$. Таким образом, имеем три последовательных момента времени. Тогда

$$\pi_{ij}(t+s) = P\left(\xi(t+s) = x_i \middle| \xi(0) = x_i\right) \tag{5.24}$$

Далее подставляем формулу полной вероятности

$$\pi_{ij}(t+s) = \frac{P\left(\xi(0) = x_i, \xi(t+s) = x_j\right)}{P\left(\xi(0) = x_i\right)} \tag{5.25}$$

Подставляем в совместную вероятность в числителе полную группу событий по промежуточному состоянию (момент времени s). Тогда получаем сумму по всем возможным промежуточным состояниям

$$\pi_{ij}(t+s) = \frac{\sum_{k=1}^{r} P\left(\xi(0) = x_i, \xi(s) = x_k, \xi(t+s) = x_j\right)}{P\left(\xi(0) = x_i\right)} \tag{5.26}$$

Воспользуемся формулой для совместной вероятности трех сечений

$$\pi_{ij}(t+s) = \frac{\sum_{k=1}^{r} a_i \pi_{ik}(s)\pi_{kj}(t)}{a_i} = \sum_{k=1}^{r} \pi_{ik}(s)\pi_{kj}(t) = [\pi(s)\pi(t)]_{ij} \tag{5.27}$$

Выбирая $0 < t < t + s$ аналогично получим

$$\pi(t+s) = \pi(t)\pi(s), \quad t, s > 0 \tag{5.28}$$

В случае $t = 0$ или $s = 0$ уравнение Чепмена-Колмогорова выполнено тривиально. Таким образом, уравнение Чепмена-Колмогорова выполнено для всех t и s.

Отметим, что полученное уравнение есть свойство чисто из теории вероятностей, так как в нем участвуют матрицы в фиксированные моменты времени. Нас же будет интересовать зависимость этих матриц от t. Рассмотрим уравнение (5.23) так, как будто $\pi(\cdot)$ – обычная функция действительной переменной (не матрица). Тогда уравенение Чепмена-Колмогорова – функциональное (нужно найти все функции, которые удовлетворяют этому уравнению). Среди его решений есть экспонента $\pi = e^{\lambda t}$, где λ – некоторая константа. Тогда производная $\pi(t)$ будет пропорциональна самой функции, то есть $\pi'(t) = \lambda\pi(t)$. Оказывается, уравнение Чепмена-Колмогорова приводит ровно к такому же следствию в общем матричном случае.

Далее условимся, что пределы в матричных равенствах мы понимаем как поэлементные:

$$\lim_{h\to+0} \pi(h) = A \iff \lim_{h\to+0} \pi_{ij}(h) = A_{ij} \tag{5.29}$$

Производная тоже как предел понимается как поэлементный, то есть производная от матрицы это матрица, составленная из производных её элементов

$$\dot{\pi}(t) \equiv \frac{d\pi}{dt} = D \iff \dot{\pi}_{ij}(t) \equiv \frac{d\pi_{ij}}{dt} = D_{ij} \tag{5.30}$$

Интеграл от матрицы – это матрица, каждый элемент которой есть интеграл от соответствующего элемента матрицы

$$\int_{t_1}^{t_2} \pi(t)dt = V \iff \int_{t_1}^{t_2} \pi_{ij}(t)dt = V_{ij} \tag{5.31}$$

для всех $i, j = 1, \ldots, r$.

3. требование (правой) непрерывности матрицы перехода $\pi(\cdot)$ в нуле $t = 0$:

$$\pi_{ij}(h) \to \delta_{ij}, \;\; \pi(h) \to \pi(0) = I \text{ при } h \to +0 \tag{5.32}$$

Семейство матричнозначных функций определено только при неотрицательных t, поэтому непрерывность в нуле может быть только правой.

Отсюда получаем следующие свойства.

Свойство 1.
Матричнозначная функция $\pi(\cdot) : \mathbb{R}_+ \to \mathbb{M}_r$ непрерывна при $t > 0$.

Математически свойство записывается следующим образом

$$\lim_{h\to 0} \pi(h+t) = \pi(t), \quad t > 0 \tag{5.33}$$

Доказательство.

Докажем непрерывность справа: так как

$$\pi(h+t) = \pi(t)\cdot\pi(h) \tag{5.34}$$

и

$$\pi(h) \to I \text{ при } h \to +0 \tag{5.35}$$

получаем

$$\pi_{ij}(t+h) = \sum_{k=1}^{r} \pi_{ik}(t)\pi_{kj}(h) \underset{h\to+0}{\longrightarrow} \sum_{k=1}^{r} \pi_{ik}(t)\delta_{kj} = \pi_{ij}(t) \tag{5.36}$$

Отметим, что в данном предельном переходе очень важна конечность числа состояний. Так как сумма конечная, мы можем занести предельный переход под знак суммы. Таким образом, правая непрерывность непосредственно вытекает из уравнения Чепмена-Колмогорова.

Докажем непрерывность слева. Для $t > 0$ и $0 < h < t$ имеем

$$\pi(t) - \pi(t-h) = \pi(t-h)\left(\pi(h) - I\right) \tag{5.37}$$

запишем поэлементно, учтя, что

$$0 \leqslant \pi_{ik}(t-h) \leqslant 1 \tag{5.38}$$

Тогда получим

$$\begin{gathered}\left|\pi_{ij}(t) - \pi_{ij}(t-h)\right| = \left|\sum_{k=1}^{r} \pi_{ik}(t-h)\left(\pi_{kj}(h) - \delta_{kj}\right)\right| \leqslant \\ \leqslant \sum_{k=1}^{r} \pi_{ik}(t-h)\left|\pi_{kj}(h) - \delta_{kj}\right| \leqslant \sum_{k=1}^{r} \left|\pi_{kj}(h) - \delta_{kj}\right| \underset{h\to+0}{\longrightarrow} 0\end{gathered} \tag{5.39}$$

Здесь мы сделали оценку модуля суммы (он не больше суммы модулей). Далее мы воспользовались тем, что $\pi_{ik} \leqslant 1$. Далее, так как сумма конечна и каждое ее слагаемое стремится к нулю, получаем, что все выражение стремится к нулю.

Таким образом, мы доказали непрерывность матрицы перехода в любой точке t.

Свойство 2.

При любом $t \geqslant 0$ матрица $\pi(t)$ невырожденная.

Доказательство.

Для того, чтобы доказать это свойство, воспользуемся тем, что конечномерная матрица невырождена тогда и только тогда, когда у нее ненулевой определитель. При этом детерминант матрицы $det\ \pi(h)$ является непрерывной функцией ее r^2 элементов. Таким образом, чтобы посчитать определитель, нам нужно знать все элементы матрицы. Определитель есть сумма произведений этих элементов (собственно поэтому он и является непрерывной функцией элементов). Тогда можно утверждать, что

$$det\ \pi(h) \to det\ I = 1 \text{ при } h \to +0 \tag{5.40}$$

Тогда, при достаточно малых h определитель ненулевой

$$det\ \pi(h) \neq 0 \text{ при } 0 < h < h_0 \tag{5.41}$$

Таким образом, мы имеем невырожденность матрицы при малых значениях аргумента. Воспользуемся уравнением Чепмена-Колмогорова в следующем равенстве

$$det\ \pi(t) = det\ \pi\left(n \cdot \frac{t}{n}\right) = det\ \pi^n\left(\frac{t}{n}\right) = \left(det\ \pi\left(\frac{t}{n}\right)\right)^n \neq 0 \text{ при } t > 0 \tag{5.42}$$

если $n \in \mathbb{N}$ выбрано так, что $\frac{t}{n} < h_0$.

Здесь мы воспользуетесь тем, что матрица от суммы есть произведение матриц (уравнение Чепмена-Колмогорова во втором знаке равенства). Таким образом, получаем n-ю степень определителя, который заведомо не равен нулю, если мы выберем n достаточно большим, чтобы аргумент матрицы был меньше h_0. Таким образом, определитель матрицы не равен нулю при малых t в силу того, что матрица стремится к единичной. При произвольных t определитель тоже не равен нулю, согласно полученному выше уравнению.

Дифференцируемость матрицы перехода.

Покажем, что из уравнения Чепмена-Колмогорова вытекает, что матрица дифференцируема. Уравнение Чепмена-Колмогорова довольно жестко фиксирует вид матрицы перехода. Соответственно, из уравнения Чепмена-Колмогорова вытекает не только вид матрицы, но и ее свойства.

Будем понимать производную матрицы как поэлементную

$$\dot{\pi}_{ij}(t) = \frac{d\pi_{ij}}{dt} \stackrel{def}{=} \lim_{h\to 0} \frac{\pi_{ij}(t+h) - \pi_{ij}(t)}{h}, \quad i, j = 1, \dots, r \tag{5.43}$$

Лемма 1. Зададим матричнозначную функцию

$$J\,(t_1;t_2) = \int\limits_{t_1}^{t_2} \pi(u)du, \quad 0 < t_1 < t_2 \tag{5.44}$$

Найдётся $\delta > 0$, определяющееся только матрицей $\pi(t_1)$, такое, что если длина интервала интегрирования $t_2 - t_1 < \delta$, то существует обратная матрица

$$R(t_1, t_2) = J^{-1}(t_1, t_2) \tag{5.45}$$

Отметим, что критерий малости определятся только элементами матрицы $\pi(t_1)$.

Доказательство.
Известно, что матричнозначная функция

$$\pi(\cdot) : \mathbb{R}_+ \to Mat_r \tag{5.46}$$

является непрерывной функцией. Значит, для каждого матричного элемента можно записать теорему о среднем

$$\frac{1}{t_2 - t_1} \int\limits_{t_1}^{t_2} \pi_{ij}(u)du = \pi_{ij}(t_1 + h_{ij}), \quad 0 \leqslant h_{ij} \leqslant= t_2 - t_1 \tag{5.47}$$

Это означает, что интеграл равен значению функции в некоторой промежуточной точке. Понятно, что для разных i и j добавка должна быть разной. В правой части уравнения получаем матрицу перехода, для которой моменты времени для каждого элемента разные. Обозначим

$$\pi_{ij}(t_1 + h_{ij}) = \widetilde{\pi} \tag{5.48}$$

Тогда получаем, что матричный элемент матрицы $\widetilde{\pi}$ стремится к $\pi_{ij}(t_1)$, если $h_{ij} \to +0$

$$\widetilde{\pi}_{ij} = \pi_{ij}(t_1 + h) \to \pi_{ij}(t_1) \text{ если } \max_{1\leqslant i,j\leqslant r} h_{ij} \to +0 \tag{5.49}$$

При этом

$$\max_{1\leqslant i,j\leqslant r} h_{ij} \leqslant t_2 - t_1 \tag{5.50}$$

Поэтому

$$det\,\widetilde{\pi} \to det\,\pi(t_1) \neq 0 \tag{5.51}$$

Таким образом,

$$\forall\varepsilon > 0 \;\; \exists\delta = \delta(\varepsilon): \quad \left|det\,\widetilde{\pi} - det\,\pi(t_1)\right| < \varepsilon, \text{ если } \max_{1\leqslant i,j\leqslant r} h_{ij} < \delta \tag{5.52}$$

Выберем в указанном неравенстве ε так, чтобы определитель $\widetilde{\pi}$ был отличен от нуля. Для этого достаточно выбрать

$$0 < \varepsilon < \left|det\,\pi(t_1)\right| \tag{5.53}$$

Тогда $det\,\widetilde{\pi} \neq 0$. Тогда

$$det\,J\,(t_1, t_2) = (t_2 - t_1)^r\,det\,\widetilde{\pi} \neq 0, \quad \text{если } 0 < t_2 - t_1 < \delta\,(\pi\,(t_1)) \tag{5.54}$$

Теорема о правой производной матрицы перехода.

Теорема. Функция $\pi(\cdot)$ имеет правую производную в точке $t = 0$

$$\lim_{h\to+0} \frac{\pi(h) - \pi(0)}{h} \overset{def}{=} \Lambda, \quad \Lambda_{ij} \in \mathbb{R} \tag{5.55}$$

Доказательство. Запишем для $\forall\, 0 < t_1 < t_2,\ \forall h > 0$ интеграл

$$\Big(\pi(h) - I\Big)\int\limits_{t_1}^{t_2} \pi(u)du = \int\limits_{t_1}^{t_2} \pi(h)\pi(u)du - \int\limits_{t_1}^{t_2} \pi(u)du \overset{1}{=} \int\limits_{t_1}^{t_2} \pi(u+h)du - \int\limits_{t_1}^{t_2} \pi(u)du \overset{2}{=}$$

$$\overset{2}{=} \int\limits_{t_1+h}^{t_2+h} \pi(u)du - \int\limits_{t_1}^{t_2} \pi(u)du \tag{5.56}$$

Выкладки:
$\overset{1}{=}$ – по уравнению Чепмена-Колмогорова
$\overset{2}{=}$ – произведём сдвиг переменных: $u + h \mapsto u$ в первом слагаемом

Пусть $0 < h < t_2 - t_1$. Тогда $t_1 < t_1 + h < t_2 < t_2 + h$, то есть нижний предел интегрирования $t_1 + h$ залезает в предел интегрирования $[t_1, t_2]$.

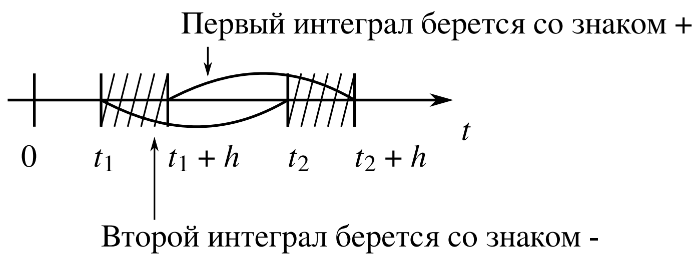

Рис. 5.1: Иллюстрация к вычислению интеграла

Тогда разность интегралов вычисляется по заштрихованной области на рис. 5.1.

Соответственно

$$\int_{t_1+h}^{t_2+h} \pi(u)du - \int_{t_1}^{t_2} \pi(u)du = \int_{t_2}^{t_2+h} \pi(u)du - \int_{t_1}^{t_1+h} \pi(u)du \tag{5.57}$$

В обозначениях леммы при $0 < t_1 < t_2$ и $0 < h < t_2 - t_1$ имеем

$$\Big(\pi(h) - I\Big)J(t_1, t_2) = \int_{t_2}^{t_2+h} \pi(u)du - \int_{t_1}^{t_1+h} \pi(u)du \tag{5.58}$$

Пусть $t_2 - t_1 < \delta$ (из леммы). Тогда

$$\frac{\pi(h) - I}{h} = \left(\frac{1}{h}\int_{t_2}^{t_2+h} \pi(u)du - \frac{1}{h}\int_{t_1}^{t_1+h} \pi(u)du\right)R(t_1, t_2) \tag{5.59}$$

при любом $t_1 > 0$, $t_2 - t_1 < \delta$, $h < t_2 - t_1$, где $\delta > 0$ определяется элементами матрицы $\pi(t_1)$ и не зависит от h. Таким образом, мы можем взять любое t_1, найти по нему δ, и дальше рассматривать малые h. Тогда, при $h \to +0$

$$\lim_{h\to+0}\left(\frac{1}{h}\int_{t_2}^{t_2+h} \pi(u)du - \frac{1}{h}\int_{t_1}^{t_1+h} \pi(u)du\right) = \Big(\pi(t_2) - \pi(t_1)\Big) \tag{5.60}$$

Здесь мы воспользовались теоремой о среднем.

Далее, если мы выберем и фиксируем $t_1, t_2 > 0$, $t_2 - t_1 < \delta$, то

$$\lim_{h\to+0} \frac{\pi(h) - I}{h} = \Big(\pi(t_2) - \pi(t_1)\Big)R(t_1, t_2) \overset{def}{=} \Lambda \tag{5.61}$$

Отметим, что Λ на самом деле не зависит от t_1 и t_2. Как бы мы не выбирали t_1 и t_2, Λ будет равно пределу, записанному в (5.61). Этот предел не зависит от t_1 и t_2.

Таким образом, мы доказали, что дифференциальное отношение при $h \to +0$ стремится к некоторой конкретной матрице, существует правая производная $\pi(\cdot)$ в нуле.

Таким образом,

$$\Lambda = \lim_{h\to+0} \frac{\pi(h) - \pi(0)}{h} \tag{5.62}$$

Теорема о дифференцируемости $\pi(\cdot)$.

Теорема.
Матричнозначная функция $\pi(\cdot)$ непрерывно дифференцируема при $t > 0$ и удовлетворяет

уравнениям Колмогорова

$$\dot{\pi}(t) = \Lambda \cdot \pi(t) = \pi(t) \cdot \Lambda \tag{5.63}$$

где $\dot{\pi}(\cdot)$ непрерывна при $t > 0$, а Λ – матрица из предыдущей теоремы.

Доказательство.
Для $t > 0$

$$\frac{\pi(t+h) - \pi(t)}{h} = \frac{\pi(h) - \pi(t)}{h} \cdot \pi(t) = \pi(t) \cdot \frac{\pi(h) - \pi(0)}{h} \tag{5.64}$$

Отсюда

$$\lim_{h \to +0} \frac{\pi(t+h) - \pi(t)}{h} = \Lambda \cdot \pi(t) = \pi(t) \cdot \Lambda \tag{5.65}$$

Таким образом, уравнение Колмогорова выполнено для правой производной. Найдем левую производную. Зафиксируем $t > 0$ и выберем $0 < h < t$. Тогда

$$\frac{\pi(t) - \pi(t-h)}{h} = \frac{\pi(h) - \pi(0)}{h} \pi(t-h) = \pi(t-h) \frac{\pi(h) - \pi(0)}{h} \tag{5.66}$$

следовательно

$$\lim_{h \to +0} \frac{\pi(t) - \pi(t-h)}{h} = \Lambda \pi(t) = \pi(t) \Lambda \tag{5.67}$$

Соответственно, $\pi(\cdot)$ удовлетворяет уравнению Колмогорова.

Непрерывность матричнозначной функции $\dot{\pi}(t), t > 0$ следует из непрерывности $\pi(\cdot)$ и любого из уравнений Колмогорова.

Теорема доказана.

Замечание. Отметим, что к уравнениям Колмогорова можно добавить начальное условие:

$$\pi(0) = I \tag{5.68}$$

Тогда получаем задачу Коши для матрицы $\pi(t)$. Ее можно записать как систему из r^2 уравнений относительно матричных элементов. Однако эти уравнения не будут независимы. Сумма по строке матрицы равна 1. Поэтому матричные элементы в каждой строке не независимы: мы можем взять один матричный элемент и выразить его через все остальные в каждой строке. Поэтому на самом деле система будет иметь меньшее количество уравнений. В итоге мы получим систему обыкновенных дифференциальных уравнений, линейных с постоянными коэффициентами и заданным начальным условием. Таким образом, зная матрицу Λ, мы всегда можем определить матрицу $\pi(t)$ в любой момент времени. То есть, зная Λ и начальные состояния, мы можем определить динамику марковского процесса в любой момент времени.

Система уравнений Колмогорова для вероятностей состояний.

Из уравнений Колмогорова для вероятностей перехода вытекают уравнения Колмогорова для вероятностей состояний. Вероятности перехода – это условные вероятности, а мы можем задать абсолютные вероятности. Пусть

$$p_i(t) = P\Big(\xi(t) = x_i\Big), \; i = 1, ..., r \;\; t > 0 \tag{5.69}$$

– вероятность того, что в момент времени t марковский процесс был в состоянии x_i. Запишем формулу полной вероятности

$$p_j(t) = P\Big(\xi(t) = x_j\Big) = \sum_{i=1}^{r} P\Big(\xi(t) = x_j | \xi(0) = x_i\Big) \cdot P\Big(\xi(0) = x_i\Big) = \sum_{i=1}^{r} \pi_{ij}(t) a_i \tag{5.70}$$

где $P\Big(\xi(0) = x_j\Big) = a_i$

Находим производную правой части и применяем уравнения Колмогорова

$$\dot{p}_j(t) = \sum_{i=1}^{r} \dot{\pi}_{ij}(t) a_i \stackrel{1}{=} \sum_{i=1}^{s} \left(\sum_{k=1}^{r} \pi_{ik}(t) \Lambda_{kj} \right) a_i = \sum_{k=1}^{r} \Lambda_{kj} \left(\sum_{i=1}^{r} \pi_{ik}(t) a_i \right) = \sum_{k=1}^{r} \Lambda_{kj} p_k(t) \tag{5.71}$$

$\stackrel{1}{=}$ – воспользуемся системой уравнений Колмогорова

Таким образом, получили

$$\dot{p}_j(t) = \sum_{k=1}^{r} p_k(t) \Lambda_{kj}, \quad j = 1, \ldots, r \tag{5.72}$$

или

$$\dot{p} = p\Lambda \tag{5.73}$$

где

$$p(t) = \Big(p_1(t) \; p_2(t) \; ..... \; p_s(t)\Big) \tag{5.74}$$

– вектор-строка.

Итак, для вероятностей состояний получаем систему линейных обыкновенных дифференциальных уравнений с постоянными коэффициентами. Существует начальное условие, не все уравнения независимы. Решив данную систему, мы можем найти распределение сечения в каждый момент времени t.

В ряде случаев уравнения Колмогорова удобно записать в другом виде.

Заметим, что

$$\sum_{k=1}^{r}\left(\pi_{jk}(h)-\delta_{jk}\right)=\sum_{k=1}^{r}\pi_{jk}(h)-1=0 \tag{5.75}$$

так как выполнено условие стохастичности

$$\sum_{k=1}^{r}\pi_{jk}(h)=1 \tag{5.76}$$

Отсюда

$$\sum_{k=1}^{r}\Lambda_{jk}=\lim_{h\to+0}\frac{1}{h}\sum_{k=1}^{r}\left(\pi_{jk}(h)-\delta_{jk}\right)=0 \tag{5.77}$$

Соответственно, для любой матрицы Λ сумма по строке равна нулю, так как эта матрица есть предел матриц, у которых сумма по строке равна нулю. Отсюда получаем, что диагональный элемент матрицы можно записать как обратную сумму недиагональных элементов

$$\Lambda_{jj}=-\sum_{k:k\neq j}\Lambda_{jk} \tag{5.78}$$

Тогда, если возьмем систему ранений Колмогорова для вероятностей состояний

$$\dot{p}_j(t)=\sum_{k=1}^{r}p_k(t)\Lambda_{kj} \tag{5.79}$$

то получим

$$\dot{p}_j(t)=p_j(t)\Lambda_{jj}+\sum_{k:k\neq j}p_k(t)\Lambda_{kj}=-\sum_{k:k\neq j}p_j(t)\Lambda_{jk}+\sum_{k:k\neq j}p_k(t\Lambda_{kj}) \tag{5.80}$$

Учтем, что $\Lambda=\dot{\pi}(0)$ и при малых $h>0$ запишем

$$\frac{p_j(t+h)-p_j}{h}\approx-\sum_{k:k\neq j}p_j(t)\frac{\pi_{jk}(h)-\pi_{jk}(0)}{h}+\sum_{k:k\neq j}p_k(t)\frac{\pi_{kj}(h)-\pi_{jk}(0)}{h} \tag{5.81}$$

Поскольку

$$\pi_{jk}(0)=\pi_{kj}(0)=\delta_{kj}=0 \text{ при } k\neq j \tag{5.82}$$

то получаем, что (5.81) эквивалентно

$$\frac{p_j(t+h)-p_j}{h}\approx-\sum_{k:k\neq j}p_j(t)\pi_{jk}(h)+\sum_{k:k\neq j}p_k(t)\pi_{kj}(h) \tag{5.83}$$

Заменим вероятности $p_j(t)$ на частоты $\frac{N_j(t)}{N}$, где $N_j(t)$ – число частиц, в момент времени

t находящихся в x_j. Здесь мы считаем, что у нас есть некий ансамбль частиц, каждая из которых может находится в момент времени t в одном из r состояний.Тогда имеем

$$\pi_{jk}(h) = P\left(\xi(t+h) = x_k \middle| \xi(t) = x_j\right) \approx \frac{N_{j\to k}}{N_j(t)} \tag{5.84}$$

– доля частиц, которые перешли из j-го в k-е состояние по отношению к количеству тех частиц, которые в момент времени t находились в j-м состоянии.

Аналогично

$$\pi_{kj}(h) \approx \frac{N_{k\to j}}{N_k(t)} \tag{5.85}$$

где $N_{j\to k}(h)$ и $N_{k\to j}(h)$ – количество частиц, совершивших за время h переход из x_j в x_k и из x_k в x_j соответственно.

Тогда получим

$$\begin{aligned} N_j(t+h) - N_j(t) \approx -\sum_{n=k:k\neq j} N_j(t)\frac{N_{j\to k}(h)}{N_j(t)} + \sum_{k:k\neq j} N_k(t)\frac{N_{k\to j}(h)}{N_k(t)} = \\ = -\sum_{k:k\neq j} N_{j\to k}(h) + \sum_{k:k\neq j} N_{k\to j}(h) = -N_j^{out}(h) + N_j^{in}(h) \end{aligned} \tag{5.86}$$

Эта формула показывает, как изменилось количество частиц, которые находились в состоянии j за время h.

Лекция 6. Теория второго порядка. Среднеквадратичные свойства случайных процессов.

Некоторые понятия линейной алгебры.

Определение. Комплексное линейное пространство $\mathcal{L}$ – это множество некоторых элементов с линейными операциями суммы элементов $x + y$ и умножения элемента на число $\alpha x \equiv \alpha \cdot x$, где $x, y \in \mathcal{L}$ и $\alpha \in \mathbb{C}$ произвольны. При этом соответствия

$$x, y \to x + y \in \mathcal{L} \tag{6.1}$$

и

$$\alpha, x \to \alpha\xi \in \mathcal{L} \tag{6.2}$$

однозначны. Это означает, что для данных двух элементов сумму можно найти единственным образом, а для данного элемента и числа их произведение можно найти единственным образом.

Справедливы следующие **свойства линейных операций** (аксиомы и теоремы). Операции коммутативны:

$$x + y = y + x \tag{6.3}$$

ассоциативны по сложению

$$(x + y) + z = x + (y + z) \tag{6.4}$$

и по умножению

$$\alpha \cdot (\beta x) = \alpha\beta \cdot x \tag{6.5}$$

и дистрибутивны

$$\alpha (x + y) = \alpha x + \alpha y \tag{6.6}$$

$$(\alpha + \beta) x = \alpha x + \beta x \tag{6.7}$$

Существует и единственен элемент

$$0 \in \mathcal{L} \tag{6.8}$$

такой, что

$$0 + x = x \tag{6.9}$$

для всех $x \in \mathcal{L}$.

Для любого $x \in \mathcal{L}$ существует и единственен элемент $-x \in \mathcal{L}$, такой что

$$x + (-x) = 0 \tag{6.10}$$

Этот элемент $-x$ называется **обратным** к x. При этом обратный элемент позволяет определить разность элементов пространства

$$x + (-y) \stackrel{def}{=} x - y \tag{6.11}$$

Справедливы равенства, связывающие умножение на число и нулевой и обратный элементы:

$$0 \cdot x = 0 \tag{6.12}$$

$$1 \cdot x = x \tag{6.13}$$

$$(-1) \cdot x = -x \tag{6.14}$$

Все описанные манипуляции с линейными пространствами по сути переносит наше естественное представление об естественных операциях с числами на некоторые абстрактные объекты.

Определение. Элементы $x_1, \ldots, x_n \in \mathcal{L}$ называются **линейно независимыми**, если линейная комбинация есть нулевой элемент пространства

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = 0 \tag{6.15}$$

тогда и только тогда, когда все коэффициенты равны нулю

$$\alpha_1 = \ldots = \alpha_n = 0 \tag{6.16}$$

Понятие линейной независимости элементов позволяет ввести следующую характеристику пространства.

Определение. **Размерность** пространства это

$$dim\ \mathcal{L} = d \tag{6.17}$$

(где $1 \leqslant d < \infty$), если найдутся d линейно независимых элементов, а любые $d+1$ элементов линейно зависимы. Если для любого $d = 1, 2, \ldots$ в $\mathcal{L}$ найдутся d линейно независимых элементов, то

$$dim\ \mathcal{L} = \infty \tag{6.18}$$

Отсюда получаем понятие базиса.

Определение. **Базис** d-мерного пространства – это линейно независимые элементы $e_1, \ldots, e_d$, такие что любой элемент пространства можно разложить как линейную комбинацию этих базисных элементов

$$\forall x \in \mathcal{L} \quad \exists \gamma_1, \ldots, \gamma_d \in \mathbb{C}: \quad x = \sum_{k=1}^{d} \gamma_k e_k \tag{6.19}$$

Существование базиса в конечномерном пространстве напрямую вытекает из определения размерности. Если найдутся d линейно независимых элементов, то они и будут базисными. Однако, далее мы будем иметь дело с бесконечномерными пространствами. Тогда нужно понять, как устроены понятие базиса в случае бесконечной размерности. В этом случае мы бы хотели получить что-то вроде (6.19), где $d = \infty$. Однако в этом случае появляются некоторые математические проблемы. Например, почему мы уверены, что в пространстве существует именно счетный базис? Мы не можем быть уверены, что разложение по базису имеет вид суммы. С другой стороны, если есть возможность представления в виде суммы, то мы имеем дело с рядом. Ряд – предел последовательности частичных сумм. Однако, последовательность частичных сумм – последовательность элементов пространства. Значит, нужно знать, чему равен предел, а для этого необходимо определить сходимость в данном линейном пространстве. Понятие сходимости можно определять разными способами, причем способы определения сходимости образуют некоторую иерархию по тому, насколько абстрактным образом вводится эта сходимость. Например, можно ввести сходимость, которая никак не связана с числовыми характеристиками, а основывается только на свойствах множеств. Можно приближать понятие сходимости к нашему представлению о сходимости действительных чисел. Самый высокий по конкретизации уровень сходимости – определение сходимости через норму элемента. Мы же будем определять норму через скалярное произведение.

Определение. Скалярное произведение в комплексном линейном пространстве $\mathcal{L}$ это правило, которое для любых $x, y \in \mathcal{L}$ единственным образом сопоставляет комплексное число $(x, y) \in \mathbb{C}$, так что

- скалярное произведение элемента на себя есть вещественное неотрицательное число

$$(x, x) \in \mathbb{R}, \quad (x, x) \geqslant 0 \tag{6.20}$$

  если и только если $x = 0$, то $(x, x) = 0$

- скалярное произведение эрмитово

$$(x, y) = \overline{(y, x)} \tag{6.21}$$

  где черта – комплексное сопряжение

- скалярное произведение линейной комбинации относительно первого сомножителя представляется в виде

$$(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1 (x_1, y) + \alpha_2 (x_2, y) \tag{6.22}$$

для любых $\alpha_1, \alpha_2 \in \mathbb{C}$.

Следствия из аксиом: в случае линейной комбинации во втором сомножителе

$$(x, \alpha_1 y_1 + \alpha_2 y_2) = \overline{\alpha_1}(x, y_1) + \overline{\alpha_2}(x, y_2) \tag{6.23}$$

В случае действительных чисел получаем

$$(x_1 - x_2, y) = (x_1, y) - (x_2, y) \tag{6.24}$$

и для второго сомножителя

$$(x, y_1 - y_2) = (x, y_1) - (x, y_2) \tag{6.25}$$

Скалярное произведение в пространстве удовлетворяет **неравенству Коши-Буняковского**

$$|(x, y)| \leqslant \sqrt{(x, x)} \cdot \sqrt{(y, y)} \tag{6.26}$$

Отметим, что неравенство Коши-Буняковского возникает во многих математических дисциплинах. Очень многие конструкции в математических науках на самом деле представляют собой скалярное произведение. Значит, имеет место и неравенство Коши-Буняковского.

Отметим, что равенство в неравенстве (6.26) достигается только в том случае, когда x и y пропорциональны

$$|(x, y)| = \sqrt{(x, x)} \cdot \sqrt{(y, y)} \leftrightarrow x = \alpha y \tag{6.27}$$

при некотором $\alpha \in \mathbb{C}$.

Определение. Элементы $x, y \in \mathcal{L}$, такие, что

$$(x, y) = 0 \tag{6.28}$$

называются **ортогональными**.

Определение. Скалярное произведение порождает **норму** – функционал, который каждому элементу сопоставляет вещественное число

$$\|x\| = \sqrt{(x, x)} \in \mathbb{R} \tag{6.29}$$

Из аксиом скалярного произведения получаем, что норма удовлетворяет следующим свойствам

- норма неотрицательна и равна нулю тогда и только тогда, когда рассматривается нулевой элемент пространства

$$\|x\| \geqslant 0; \quad \|x\| = 0 \ \leftrightarrow \ x = 0 \tag{6.30}$$

- в случае умножения элемента на число норма их произведения удовлетворяет

$$\|\alpha x\| = |\alpha| \cdot \|x\| \tag{6.31}$$

так как

$$\|\alpha x\|^2 = (\alpha x, \alpha x) = \alpha\overline{\alpha}\,(x, x) = |\alpha|^2 \cdot \|x\|^2 \tag{6.32}$$

- справедливо

$$\|x + y\| \leqslant \|x\| + \|y\| \tag{6.33}$$

– неравенство треугольника

Отметим, что не всегда норма порождена скалярным произведением. Тогда указанные свойства вводятся на уровне аксиом.

Очевидно, нормы самого элемента и обратного к нему совпадают

$$\| -x\| = \|x\| \tag{6.34}$$

Тогда неравенство Коши-Буняковского можно записать в виде

$$|(x, y)| \leqslant \|x\| \cdot \|y\| \tag{6.35}$$

Если в пространстве определена норма, можно задать сходимость последовательности элементов пространства. Последовательность $\{x_n\} \subset \mathcal{L}$ **сходится**, если существует $x \in \mathcal{L}$, такой, что

$$\|x_n - x\| \underset{n\to\infty}{\longrightarrow} 0 \tag{6.36}$$

Отметим, что последовательность норм в данном случае – обычная последовательность действительных чисел. Мы пишем

$$x \Leftrightarrow x_n \text{ или } x = \lim x_n \tag{6.37}$$

Определение. Непрерывность нормы: если $x_n \to x$, то $\|x_n\| \to \|x\|$. Фактически это вытекает из неравенства треугольника

$$\|x_n\| \leqslant \|x_n - x\| + \|x\| \tag{6.38}$$

или
$$\|x\| \leqslant \|x - x_n\| + \|x_n\| \tag{6.39}$$
Отсюда получаем
$$\|x_n\| - \|x\| \leqslant \|x_n - x\| \tag{6.40}$$
для (6.38) и
$$\|x_n\| - \|x\| \geqslant -\|x - x_n\| = -\|x_n - x\| \tag{6.41}$$
для (6.39). Тогда объединяя два полученных неравенства в одно, получим неравенство для модуля разности норм
$$\Big|\|x_n\| - \|x\|\Big| \leqslant \|x_n - x\| \to 0 \tag{6.42}$$
Итак, модуль разности норм не больше нормы разности.

Определение. Последовательность $\{x_n\} \subset \mathcal{L}$ называется **фундаментальной**, если выполнено условие сходимости вида
$$\|x_n - x_m\| \to 0 \text{ при } n, m \to \infty \tag{6.43}$$
то есть
$$\forall \varepsilon > 0 \;\; \exists N = N(\varepsilon): \;\; \|x_n - x_m\| < \varepsilon \;\; \forall n, m > N(\varepsilon) \tag{6.44}$$
В силу неравенства треугольника
$$\|x_n - x_m\| \leqslant \|x_n - x\| + \|x - x_m\| \tag{6.45}$$
следует, что *сходимость всегда влечет фундаментальность*. В обратном случае утверждение в общем случае неверно. Однако, если оно верно, то пространство называется полным.

Определение. Пространство $\mathcal{L}$ называется **полным**, если для любой фундаментальной последовательности $\{x_n\} \subset \mathcal{L}$ существует некий элемент $x \in \mathcal{L}$, к которому эта последовательность сходится
$$\|x_n - x\| \to 0 \tag{6.46}$$
В полном пространстве сходимость эквивалентна фундаментальности. Отметим, что это свойство полного пространства очень важно. Например, пусть нам нужно исследовать последовательность на сходимость. Тогда нам нужно определить не сам предел последовательности, а его существование в принципе. Тогда, не имея свойства полноты, нужно будет проверять сходимость к нулю сходимость нормы $\|x_n - x\|$. Однако, в случае полноты пространства нужно исследовать норму $\|x_n - x_m\|$. Тогда, если мы докажем, что такая норма стремится к нулю и при этом пространство является полным, то получим, что

существует некий элемент, к которому эта последовательность сходится. Таким образом, в полном пространстве вместо сходимости можно доказывать фундаментальность.

Определение. Базис бесконечномерного пространства – это линейно независимые $e_1, e_2, \ldots$ такие, что любой элемент пространства можно представить в виде ряда

$$\forall x \in \mathcal{L} \;\; \exists \gamma_1, \gamma_2, \ldots \in \mathbb{C} : \quad x = \sum_{k=1}^{\infty} \gamma_k e_k \tag{6.47}$$

при этом ряд понимается как предельный переход

$$\left\| x - \sum_{k=1}^{n} \gamma_k e_k \right\| \underset{n\to\infty}{\longrightarrow} 0 \tag{6.48}$$

Определение. Базис $e_1, e_2, \ldots,$ называется **ортонормированным**, если

$$\left(e_j, e_k\right) = \delta_{jk} \tag{6.49}$$

то есть

$$\|e_k\| = 1 \text{ и } (e_k, e_j) = 0 \text{ при всех } k \neq j \tag{6.50}$$

"Почти все" бесконечномерные пространства со скалярным произведением имеют счетный ортонормированный базис.

Гильбертовы случайные величины.

Рассмотрим случайные величины, имеющие комплексное значение

$$\xi : \; \Omega \to \mathbb{C} \tag{6.51}$$

то есть комплекснозначная случайная величина – двумерная случайная величина

$$\xi = \xi_R + i\xi_I \tag{6.52}$$

или в полярных координатах

$$\xi = \rho e^{i\varphi} \tag{6.53}$$

где $\xi_R, \xi_I, \rho, \varphi : \; \Omega \to \mathbb{R}$ суть обычные случайные величины.

По свойству линейности можно записать математическое ожидание комплекснозначной случайной величины

$$M\xi \overset{def}{=} M\xi_R + iM\xi_I \tag{6.54}$$

где вещественная часть математического ожидания есть математическое ожидание вещественной части случайной величины

$$Re\, M\xi = M\xi_n = MRe\, \xi \tag{6.55}$$

аналогично для мнимой части

$$Im\, M\xi = M\xi_I = MIm\, \xi \tag{6.56}$$

Математическое ожидание комплексно-сопряженной случайной величины

$$M\overline{\xi} = M\xi_R - iM\xi_I = \overline{M\xi} \tag{6.57}$$

Математическое ожидание модуля

$$M|\xi|^2 = MRe^2\xi + MIm^2\xi \tag{6.58}$$

Полезное неравенство:

$$|M\xi| \leqslant M|\xi| \leqslant \sqrt{M|\xi|^2} \tag{6.59}$$

В сущности эти неравенства представляют собой частные случаи неравенств треугольника и Коши-Буняковского. Обоснуем это. Математическое ожидание есть модуль интеграла

$$|M\xi| = \left| \int\limits_{-\infty}^{\infty} xdF_\xi(x) \right| \leqslant \int\limits_{-\infty}^{\infty} |x|dF_\xi(x) = M|\xi| \tag{6.60}$$

Интеграл – это предел последовательности частичных сумм.Модуль суммы не больше суммы модулей, что есть неравенство треугольника. В правой же части (6.60) записано математическое ожидание модуля.

Второе же неравенство в (6.59) представляет собой неравенство для дисперсии

$$(M|\xi|)^2 - M|\xi|^2 = -D|\xi| \leqslant 0 \tag{6.61}$$

В сущности это неравенство Коши-Буняковского. Оно эквивалентно неравенству

$$M|\xi| \leqslant \sqrt{M|\xi|^2} \tag{6.62}$$

Отсюда получаем, что неравенства (6.59) гарантируют, что если выражение в правой части (6.59) существует (математическое ожидание под знаком корня – конечный интеграл), то математическое ожидание модуля и модуль математического ожидания также будут конечны. Это утверждение позволяет нам ввести пространство гильбертовых случайных величин.

Определение. Пространство гильбертовых случайных величин – это линейное пространство

$$\mathcal{H} = \left\{\text{с.в. } \xi : \ \Omega \to \mathbb{C}, \ \ \text{такие, что } M\xi = 0, \ \ M|\xi|^2 < \infty\right\} \tag{6.63}$$

со скалярным произведением и нормой

$$(\xi, \eta) = cov\,(\xi, \eta) \overset{def}{=} M\xi\overline{\eta} - M\xi \cdot \overline{M\eta} = M\xi\overline{\eta} \tag{6.64}$$

Отметим, что здесь во вторых сомножителях появляется комплексное сопряжение.

Скалярное произведение задает норму

$$\|\xi\| = \sqrt{(\xi, \xi)} = \sqrt{M\xi\overline{\xi}} = \sqrt{M|\xi|^2} = \sqrt{D\xi} \tag{6.65}$$

Можно показать, что при таком определении размерность пространства гильбертовых случайных величин бесконечна

$$dim\,\mathcal{H} = \infty \tag{6.66}$$

Ортогональность в $\mathcal{H}$ – это некоррелированность случайных величин, так как коэффициент корреляции, равный нулю, соответствует некоррелированности.

Теорема (без доказательства). Пространство $\mathcal{H}$ полно, то есть условие

$$\|\xi_n - \xi_m\| \to 0 \tag{6.67}$$

эквивалентно существованию случайной величины $\xi \in \mathcal{H}$, такой что

$$\|\xi_n - \xi\| \to 0 \tag{6.68}$$

Замечание. Указанная теорема о полноте жестко связана с тем, что для определения математического ожидания используется интеграл Лебега. Если определять математическое ожидание через интеграл Лебега, то соответствующее пространство будет полным, а если через интеграл Римана – то неполным. В частности поэтому в теории вероятностей используется лебегова мера.

Далее все случайные величины принадлежат пространству $\mathcal{H}$.

Лемма 1 (непрерывность скалярного произведения). Если $\xi_n \to \xi$ и $\overline{\xi}_m \to \xi$, то их скалярное произведение

$$\left(\xi_n, \overline{\xi}_m\right) \to (\xi, \xi) \tag{6.69}$$

при $n, m \to \infty$.

Доказательство. Имеем

$$\|\xi_n - \xi\| \to 0 \tag{6.70}$$

$$\|\overline{\xi}_m - \xi\| \to 0 \tag{6.71}$$

тогда можно написать оценку для разности между скалярными произведениями

$$\left|\left(\xi_n, \overline{\xi}_m\right) - (\xi, \xi)\right| = \left|\left(\xi_n, \widetilde{\xi}_m\right) - (\xi_n, \xi) - (\xi_n, \xi) - (\xi, \xi)\right| = \left|\left(\xi_n, \widetilde{\xi}_m\right) + (\xi_n - \xi, \xi)\right| \leqslant$$
$$\leqslant \left|\left(\xi_n, \widetilde{\xi}_m - \xi\right)\right| + |(\xi_n - \xi, \xi)| \leqslant \|\xi_n\| \cdot \|\widetilde{\xi}_m - \xi\| + \|\xi_n - \xi\| \cdot \|\xi\| \tag{6.72}$$

Здесь мы сперва воспользовались неравенством треугольника, а потом – неравенством Коши-Буняковского.

Видно, что второе слагаемое

$$\|\xi_n - \xi\| \cdot \|\xi\| \underset{n\to\infty}{\longrightarrow} 0 \tag{6.73}$$

Для первого слагаемого

$$\xi_n \to \xi \tag{6.74}$$

следовательно,

$$\|\xi_n\| \to \|\xi\| \tag{6.75}$$

в силу непрерывности нормы. Поэтому имеет место сходимости числовой последовательности $\{\|\xi_n\|\}$, а значит, она ограничена. Отсюда

$$\|\xi_n\| \cdot \|\widetilde{\xi}_m - \xi\| \leqslant const \cdot \|\widetilde{\xi}_m - \xi\| \underset{n,m\to\infty}{\longrightarrow} 0 \tag{6.76}$$

Выбирая n и m достаточно большими, мы можем сделать

$$\left|\left(\xi_n, \widetilde{\xi}_m\right) - (\xi, \xi)\right| \tag{6.77}$$

сколь угодно малым.

Лемма 2. Если существует сходимость

$$(\xi_n, \xi_m) \underset{n,m\to\infty}{\longrightarrow} A \tag{6.78}$$

где A – некоторое комплексное число, то последовательность $\{\xi_n\}$ имеет предел.

Доказательство. Условия леммы не дают никакой информации о самом пределе. Значит, мы будем доказывать не сходимость последовательности, а её фундаментальности.

Для любого $\varepsilon > 0$ найдется номер N, такой, что

$$|(\xi_n, \xi_m) - A| \leqslant \varepsilon \tag{6.79}$$

для всех $n, m > N$ (по определению сходимости). Тогда, если считать, что $n = m$

$$|(\xi_n, \xi_n) - A| \leqslant \varepsilon \tag{6.80}$$

для всех $n > N$, то есть для нормы имеет место сходимость

$$\|\xi_n\|^2 \underset{n\to\infty}{\longrightarrow} A \tag{6.81}$$

Отсюда понятно, что число А должно быть вещественным как предел последовательности неотрицательных вещественных чисел: $A \in \mathbb{R}$, значит, $A = \overline{A}$. Следовательно,

$$\|\xi_n - \xi_m\|^2 = \|\xi\|^2 - (\xi_n, \xi_m) - \overline{(\xi_n, \xi_m)} + \|\xi_m\|^2 \to A - A - \overline{A} + A = 0 \tag{6.82}$$

Отсюда получаем, что последовательность $\{\xi_n\}$ фундаментальна, поэтому имеет предел. Лемма доказана.

Гильбертовы случайные процессы.

Перенесем конструкции для гильбертовых случайных величин на случайные процессы. Бцдем рассматривать комплекснозначные случайные процессы.

Определение. Комплекснозначный случайный процесс $\xi(t), t \in \mathbb{R}$, называется **гильбертовым**, если любое сечение есть гильбертова случайная величина $\xi(t) \in \mathcal{H}$.

Определение. Случайный процесс $\xi(t), t \in \mathbb{R}$ называется **непрерывным в среднем квадратичном (с.к.-непрерывным)** в точке t, если

$$M\,|\xi(t+h) - \xi(t)|^2 \underset{h\to 0}{\longrightarrow} 0 \tag{6.83}$$

то есть для любой последовательности $\{h_n\} \subset \mathbb{R}$, $h_n \to 0$

$$M\,|\xi(t+h_n) - \xi(t)|^2 = \|\xi(t+h_n) - \xi(t)\|^2 \underset{n\to\infty}{\longrightarrow} 0 \tag{6.84}$$

Таким образом, получаем связь между сходимостью по норме в пространстве $\mathcal{H}$ и свойством непрерывности случайного процесса в среднем квадратичном. Значит, мы исследуем непрерывность процесса в терминах среднего квадрата разности.

Пусть

$$R(t,s) = cov\left(\xi(t), \xi(s)\right) = M\xi(t)\overline{\xi(s)}, \quad s,t \in \mathbb{R} \tag{6.85}$$

– ковариационная функция. Эта функция носит эрмитов характер:

$$R(t,s) = \overline{R(s,t)} = (\xi(t), \xi(s)) \tag{6.86}$$

На диагонали функция равна дисперсии

$$R(t,t) = D\xi(t) = \|\xi(t)\|^2 \tag{6.87}$$

$R(\cdot)$ как функция двух переменных непрерывна в точке (t,s), если выполнено условие

$$\left|R\left(t + h_n, s + \widetilde{h}_m\right) - R(t,s)\right| \to 0 \tag{6.88}$$

при $h_n \to 0$, $\widetilde{h}_m \to 0$. При этом сходящиеся к нулю последовательности $\{h_n\}$ и $\left\{\widetilde{h}_m\right\}$ произвольны.

Теорема. Гильбертов случайный процесс $\xi(t), t \in \mathbb{R}$, с.к.- непрерывен в точке t тогда и только тогда, когда функция $R(\cdot)$ непрерывна в точке (t,t) (рассматриваем непрерывность на диагонали).

Доказательство. Пусть $\{h_n\}$ и $\left\{\widetilde{h}_m\right\}$ произвольны, $h_n \to 0$, $\widetilde{h}_m \to 0$. Положим (случайная величина, равная сечению)

$$\xi_n = \xi\left(t + h_n\right) \tag{6.89}$$

и

$$\widetilde{\xi}_m = \xi(t + \widetilde{h}_m) \tag{6.90}$$

и для краткости будем обозначать $\xi = \xi(t)$.

Тогда, если дана с.к.-непрерывность в точке t, то имеют место два предельных перехода

$$\|\xi_n - \xi\| \to 0 \tag{6.91}$$

и

$$\|\widetilde{\xi}_m - \xi\| \to 0 \tag{6.92}$$

по норме пространства $\mathcal{H}$. Тогда, если записать ковариационную функцию как скалярное произведение, получим (согласно Лемме 1)

$$R\left(t + h_n, t + \widetilde{h}_m\right) = \left(\xi_n, \widetilde{\xi}_m\right) \to (\xi, \xi) = R(t,t), \quad n,m \to \infty \tag{6.93}$$

Итак, в одну сторону теорема доказана. Наоборот, если для любых сходящихся к нулю $\{h_n\}$ и $\left\{\widetilde{h}_m\right\}$

$$\left|R\left(t+h_n, t+\widetilde{h}_m\right) - R(t,t)\right| \underset{n,m\to\infty}{\longrightarrow} 0 \tag{6.94}$$

то (при совпадающих последовательностях $\widetilde{h}_m = h_m$ для всех m)

$$\|\xi_n\|^2 = (\xi_n, \xi_n) = R(t+h_n, t+h_n) \to R(t,t) = (\xi, \xi) = \|\xi\|^2, \quad n \to \infty \tag{6.95}$$

а также (при $\widetilde{h}_m = 0$ при всех m)

$$(\xi_n, \xi) = R(t+h_n, t) \to R(t,t) = \|\xi\|^2, \quad n \to \infty \tag{6.96}$$

Отсюда получаем

$$\|\xi_n - \xi\|^2 = \|\xi_n\|^2 - (\xi_n, \xi) - \overline{(\xi_n, \xi)} + \|\xi\|^2 \to \|\xi\|^2 - \|\xi\|^2 - \|\xi\|^2 + \|\xi\|^2 = 0 \tag{6.97}$$

Другими словами,

$$M\left|\xi(t+h_n) - \xi(t)\right|^2 \to 0 \tag{6.98}$$

для любой сходящейся к нулю последовательности $\{h_n\}$.

Теорема доказана.

Можно показать, что непрерывность на диагонали влечет непрерывность на всех плоскости. Если $R(\cdot)$ непрерывна в точках (t,t) и (s,s), то она непрерывна в точке (t,s). Для доказательна этого утверждения можно воспользоваться неравенством Коши-Буняковского. Пусть $h_n \to 0$, $\widetilde{h}_m \to 0$. Введем случайные величины

$$\xi_n = \xi(t+h_n) \tag{6.99}$$

и

$$\widetilde{\eta}_m = \xi(s+\widetilde{h}_m) \tag{6.100}$$

При этом

$$\xi = \xi(t), \quad \eta = \xi(s) \tag{6.101}$$

Тогда имеют место сходимости

$$\|\xi_n - \xi\|^2 \to 0 \tag{6.102}$$

и

$$\|\widetilde{\eta}_m - \eta\|^2 \to 0 \tag{6.103}$$

следовательно, при $n, m \to \infty$

$$\begin{aligned} \left|R(t+h_n, s+\widetilde{h}_m) - R(t,s)\right| = \left|(\xi_n, \widetilde{\eta}_m) - (\xi, \eta)\right| = \left|(\xi_n, \widetilde{\eta}) - (\xi_n, \eta) + (\xi_n, \eta) - (\xi, \eta)\right| \leqslant \\ \leqslant \left|(\xi_n, \widetilde{\eta}_m - \eta)\right| + \left|(\xi_n - \xi, \eta)\right| \leqslant \|\xi_n\| \cdot \|\widetilde{\eta}_m - \eta\| + \|\xi_n - \xi\| \cdot \|\eta\| \end{aligned} \tag{6.104}$$

Здесь последовательность нормы ограничена

$$\|\xi_n\| \leqslant C \tag{6.105}$$

Одно слагаемое – бесконечно малая

$$\|\widetilde{\eta}_m - \eta\| \underset{m\to\infty}{\longrightarrow} 0 \tag{6.106}$$

Последнее слагаемое – произведение бесконечно малой на число

$$\|\xi_n - \xi\| \cdot \|\eta\| \underset{n\to\infty}{\longrightarrow} 0 \tag{6.107}$$

Поэтому получаем, что модуль разности стремится к нулю.

Лекция 7. Теория второго порядка.

Гильбертовы случайные процессы.

Определение. Случайный процесс $\xi(t), t \in \mathbb{R}$ называется **дифференцируемым в среднем квадратичном (с.к.-дифференцируемым)** в точке t, если существует гильбертова случайная величина

$$\dot{\xi} = \dot{\xi}(t) \in \mathcal{H} \tag{7.1}$$

такая, что

$$M\left|\frac{\xi(t+h)-\xi(t)}{h} - \dot{\xi}(t)\right|^2 = \left\|\frac{\xi(t+h)-\xi(t)}{h} - \dot{\xi}(t)\right\|^2 \xrightarrow[h\to 0]{} 0 \tag{7.2}$$

Далее свяжем средне квадратичную дифференцируемость процесса со свойствами ковариационной функции этого процесса. Для этого нам нужно исследовать вторую производную ковариационнной функции. Введем первое и второе разностные отношения для $\frac{\partial^2 R(t,s)}{\partial s \partial t}$

$$\frac{\partial R(t,s)}{\partial t} \approx \frac{R(t+h,s)-R(t,s)}{h} \tag{7.3}$$

Далее вводим приближение вида

$$\frac{\partial}{\partial s}\frac{\partial R(t,s)}{\partial t} \approx \frac{\frac{\partial R(t,s+\widetilde{h})}{\partial t} - \frac{\partial R(t,s)}{\partial t}}{\widetilde{h}} \approx \frac{R(t+h,s+\widetilde{h}) - R(t,s+\widetilde{h}) - R(t+h,s) + R(t,s)}{h\widetilde{h}} \tag{7.4}$$

Мы говорим, что $R(\cdot)$ имеет вторую смешанную производную в точке на плоскости (t,s), если для любых сходящихся к нулю последовательностей $\{h_n\}$ и $\left\{\widetilde{h}_m\right\}$ существует предел, который и называется второй смешанной производной в указанной точке

$$\lim_{n,m\to\infty} \frac{R(t+h_n,s+\widetilde{h}_m) - R(t,s+\widetilde{h}_m) - R(t+h_n,s) + R(t,s)}{h_n\widetilde{h}_m} \overset{def}{=} \frac{\partial^2 R(t,s)}{\partial s \partial t} \tag{7.5}$$

Отметим, что в математическом анализе вместо указанного двойного предела вводят последовательный предел. При этом в случае повторных пределов получаем чувствительность к замене порядка при переходе к пределу. В функциональном анализе вообще говоря нельзя менять местами производные. Однако в нашем случае определение симметрично по t и s и потому таких проблем не возникает.

Среднеквадратичная дифференцируемость случайного процесса.

Теорема. Случайный процесс $\xi(t), t \in \mathbb{R}$, средне квадратично-дифференцируем в точке t тогда и только тогда, когда его ковариационноая функция $R(\cdot)$ имеет вторую смешанную производную в точке (t,t).

Отметим, что в данной теореме мы связываем дифференцируемость процесса со свойствами ковариационной функции. С каким? Логично предположить, что дифференцируемость процесса должна быть связана с дифференцируемостью ковариационной функции. Ковариационноая функция – функция двуз переменных, а случайный процесс – функция одной переменной. Значит, речь идет не о первой производной. Логично рассмотреть вторую смешанную производную. С другой стороны, процесс зависит от t, значит, стоит рассмотреть производную в точке, в которой она нечувствительна к замене t на s.

Доказательство. Пусть $h_n \to 0$ и $\widetilde{h}_m \to 0$. Введём сокращенные обозначения

$$\xi_n = \xi\,(t + h_n) \tag{7.6}$$

и

$$\widetilde{\xi}_m = \xi(t + \widetilde{h}_m) \tag{7.7}$$

А предельное отношение обозначим

$$\xi = \xi(t) \tag{7.8}$$

Дифференциальную разность обозначим через

$$D^{(2)}_{h_n,\widetilde{h}_m} = \frac{R(t+h_n, t+\widetilde{h}_m) - R(t, t+\widetilde{h}_m) - R(t+h_n, t) + R(t,t)}{h_n\widetilde{h}_m} \tag{7.9}$$

Введем также случайные величины

$$\Delta_n = \frac{\xi(t+h_n) - \xi(t)}{h_n} \tag{7.10}$$

и

$$\widetilde{\Delta}_m = \frac{\xi(t+\widetilde{h}_m) - \xi(t)}{\widetilde{h}} \tag{7.11}$$

В этих обозначениях

$$D^{(2)}_{h_n,\widetilde{h}_m} = \frac{(\xi_n, \widetilde{\xi}_m) - \left(\xi, \widetilde{\xi}_m\right) - (\xi_n, \xi) + (\xi, \xi)}{h_n\widetilde{h}_m} = \left(\frac{\xi_n - \xi}{h_n}, \frac{\widetilde{\xi}_m - \xi}{\widetilde{h}_m}\right) \tag{7.12}$$

то есть

$$D^{(2)}_{h_n,\widetilde{h}_m} = \left(\Delta_n, \widetilde{\Delta}_m\right) \tag{7.13}$$

Условие того, что дифференциальное отношение сходится к величине $\dot{\xi}$ перепишем в виде

$$M\left|\frac{\xi(t+h_n) - \xi(t)}{h_n} - \dot{\xi}(t)\right|^2 = \|\Delta_n - \dot{\xi}\|^2 \tag{7.14}$$

где

$$\dot{\xi} = \dot{\xi}(t) \tag{7.15}$$

Если $\xi(t), t \in \mathbb{R}$, с.к.-дифференцируем в точке t, то при $n, m \to \infty$

$$\|\Delta_n - \dot{\xi}\|^2 \to 0 \tag{7.16}$$

и

$$\|\widetilde{\Delta}_m - \dot{\xi}\|^2 \to 0 \tag{7.17}$$

Отсюда по свойству непрерывности скалярного произведения получаем

$$D^{(2)}_{h_n,\widetilde{h}_m} = \left(\Delta_n, \widetilde{\Delta}_m\right) \to \left(\dot{\xi}, \dot{\xi}\right) \tag{7.18}$$

Тем самым мы доказали, что ковариационная функция имеет вторую смешанную производную на диагонали. Заметим также, что эта производная равна

$$\frac{\partial^2 R(t,t)}{\partial s \partial t} = \left(\dot{\xi}, \dot{\xi}\right) = M\dot{\xi}(t)\overline{\dot{\xi}(t)} = M|\dot{\xi}(t)|^2 \tag{7.19}$$

Итак, мы доказали теорему в одну сторону: если процесс дифференцируем, то существует вторая смешанная производная. Рассмотрим вторую часть задачи – дана дифференцируемость ковариационной функции. Если для любых сходящихся к нулю $\{h_n\}$ и $\left\{\widetilde{h}_m\right\}$ имеет место сходимость к некоторому числу D, которое и является смешанной производной

$$D^{(2)}_{h_n,\widetilde{h}_m} = \left(\Delta_n, \widetilde{\Delta}_m\right) \to \frac{\partial^2 R(t,t)}{\partial s \partial t} \overset{def}{=} D, \quad n, m \to \infty \tag{7.20}$$

то (при $\widetilde{h}_m = h_m$ для всех m) мы получаем

$$\|\Delta_n\|^2 = (\Delta_n, \Delta_n) \to \frac{\partial^2 R(t,t)}{\partial s \partial t} = D, \qquad n \to \infty \tag{7.21}$$

Отсюда $D \in \mathbb{R}$, так как к нему сходится последовательность вещественных чисел. Мы хотим доказать фундаментальность последовательности. Значит, нужно рассмотреть квадрат нормы разности

$$\|\Delta_n - \Delta_m\|^2 = \|\Delta_n\|^2 - (\Delta_n, \Delta_m) - \overline{(\Delta_n, \Delta_m)} + \|\Delta_m\|^2 \to D - D - D + D = 0 \tag{7.22}$$

Поэтому получаем $\{\Delta_n\}_{n=\overline{1,\infty}}$ – фундаментальна. Следовательно, имеет предел в $\mathcal{H}$. Обозначим его как $\dot{\xi}$.

Покажем, что предел не зависит от выбора последовательности $\{h_n\}$ в $\{\Delta_n\}$. Выберем какие-то $h_n \to 0$ и $\widetilde{h}_m \to 0$ и построим соответствующие

$$\Delta_n = \frac{\xi(t+h_n) - \xi(t)}{h_n} \tag{7.23}$$

$$\widetilde{\Delta}_m = \frac{\xi(t+\widetilde{h}_m) - \xi(t)}{\widetilde{h}_m} \tag{7.24}$$

При этом существует пределы $\dot{\xi}, \widetilde{\dot{\xi}} \in \mathcal{H}$, такие, что $\Delta_n \to \dot{\xi}$ и $\widetilde{\Delta}_m \to \widetilde{\xi}$.

Отметим, что для любых последовательностей $\{h_n\}$ и $\left\{\widetilde{h}_m\right\}$, сходящихся к нулю, имеет место сходимость (по определению наличия второй производной)

$$D^{(2)}_{h_n,\widetilde{h}_m} = \left(\Delta_n, \widetilde{\Delta}_m\right) \to D \in \mathbb{R} \tag{7.25}$$

при $n, m \to \infty$. Тогда получаем

$$\|\Delta_n - \widetilde{\Delta}_n\|^2 = (\Delta_n, \Delta_n) - \left(\Delta_n, \widetilde{\Delta}_m\right) - \overline{\left(\Delta_n, \widetilde{\Delta}_m\right)} + \left(\widetilde{\Delta}_m, \widetilde{\Delta}_m\right) \to D - D - D + D = 0 \tag{7.26}$$

Дальше, с помощью неравенства треугольника оцениваем

$$\|\dot{\xi} - \widetilde{\dot{\xi}}\| \leqslant \|\dot{\xi} - \Delta_n\| + \|\Delta_n - \widetilde{\Delta}_m\| + \|\widetilde{\Delta}_m - \widetilde{\dot{\xi}}\| \to 0, \qquad n, m \to \infty \tag{7.27}$$

Отметим, что стремление к нулю в (7.27) обусловлено стремлением к нулю каждого слагаемого. Это значит, что при выборе достаточно больших номеров n и m, можно выбрать выражение в правой части сколь угодно малым. Таким образом, выражение в левой части можно сделать меньше любого положительного числа. Это значит

$$\|\dot{\xi} - \widetilde{\dot{\xi}}\|^2 = 0 \tag{7.28}$$

то есть

$$\dot{\xi} = \widetilde{\dot{\xi}} \tag{7.29}$$

с вероятностью единица. Итак, теорема доказана.

Среднеквадратичная интегрируемость случайного процесса.

Исследуем среднеквадратичную интегрируемость случайного процесса на $[a, b]$. Введем интегральную сумму (это случайная величина)

$$\Upsilon_n = \sum_{k=1}^{n} \xi(t_k^*)\,(t_k - t_{k-1}) \tag{7.30}$$

где точки разбиения $t_0, t_1, \ldots, t_n$ и выборочные точки $t_1^*, \ldots, t_n^*$, таковы что

$$a = t_0 < t_1 < \ldots < t_n = b \tag{7.31}$$

при этом максимальная длина разбиения

$$\underset{1 \leqslant k \leqslant n}{max} \, (t_k - t_{k-1}) \underset{n \to \infty}{\longrightarrow} 0 \tag{7.32}$$

а сегмент разбиения

$$t_{k-1} \leqslant t_k^* \leqslant t_k, \quad k = 1, \ldots, n \tag{7.33}$$

Все последующие интегральные суммы строятся при этих условиях.

Определение. Случайный процесс $\xi(t), t \in \mathbb{R}$, называется **интегрируемым в среднем квадратичном (с.к.-интегрируемым)** на отрезке $[a, b]$, если существует случайная величина $\gamma \in \mathcal{H}$, такая что для любой последовательности интегральных сумм имеет место сходимость интегральных сумм к Υ

$$M \left| \sum_{k=1}^{n} \xi(t_k^*) \, (t_k - t_{k-1}) - \Upsilon \right|^2 = \left\| \sum_{k=1}^{n} \xi(t_k^*) \, (t_k - t_{k-1}) - \gamma \right\|^2 \underset{n \to \infty}{\longrightarrow} 0 \tag{7.34}$$

Сходимость не зависит от способа разбиения и выбора промежуточных точек.

По определению положим

$$\Upsilon = \int\limits_a^b \xi(t) dt \tag{7.35}$$

Далее будем связывать интегрируемость процесса с интегрируемостью ковариационной функции. Ковариационная функция – функция двух переменных. Нам нужно, чтобы интегрируемость была инвариантна относительно замены t на s в аргументе функции. Значит, возможно два варианта – интегрировать в квадрате или на диагонали. Удобнее интегрировать в квадрате.

Итак, функция $R(\cdot)$ интегрируема в области, которая является квадратом $[a, b] \times [a, b]$, если для любых двух разбиений $\{t_k\}_{k=\overline{0,n}}$ и $\left\{s_j\right\}_{j=\overline{0,m}}$ и любых соответствующих двух наборов выборочных точек $\left\{t_k^*\right\}_{k=\overline{1,n}}$ и $\left\{s_j^*\right\}_{j=\overline{1,m}}$ составляем двойную интегральную сумму вида

$$I_{n,m} \overset{def}{=} \sum_{k=1}^{n} \sum_{j=1}^{m} R\left(t_k^*, s_j^*\right) (t_k - t_{k-1}) \left(s_j - s_{j-1}\right) \underset{n,m \to \infty}{\longrightarrow} I \overset{def}{=} \int\limits_a^b \int\limits_a^b R(t, s) dt ds \tag{7.36}$$

Здесь мы потребовали, чтобы эта числовая последовательность сходилась к некоторому комплексному числу, которое и называется интегралом от ковариационной функции. Получим связь между среднеквадратичной интегрируемостью случайного процесса и интегрируемостью ковариационной функции в виде следующей теоремы.

Теорема. Случайный процесс $\xi(t), t \in \mathbb{R}$, среднеквадратично интегрируем на отрезке $[a, b]$ тогда и только тогда, когда ковариационная функция $R(\cdot)$ интегрируема в квадрате $[a, b] \times [a, b]$.

Доказательство (аналогично с.к.-дифференцируемости).
Пусть процесс с.к.-дифференцируем на отрезке $[a, b]$. Составим две произвольные интегральные суммы

$$\Upsilon_n = \sum_{k=1}^{n} \xi(t_k^*)\,(t_k - t_{k-1}) \tag{7.37}$$

$$\widetilde{\Upsilon}_m = \sum_{j=1}^{m} \xi(s_j^*)\,\left(s_j - s_{j-1}\right) \tag{7.38}$$

Тогда, по условию, если процесс с.к.-интегрируем, обе интегральные суммы сходятся к одной и той же случайной величине

$$\|\Upsilon_n - \Upsilon\| \to 0, \quad \|\widetilde{\Upsilon}_m - \Upsilon\| \to 0 \text{ при } n, m \to \infty \tag{7.39}$$

При этом

$$I_{n,m} = \sum_{k=1}^{n}\sum_{j=1}^{m} R(t_k^*, s_j^*)(t_k - t_{k-1})(s_j - s_{j-1}) = \sum_{k=1}^{n}\sum_{j=1}^{m} M\xi(t_k^*)\overline{\xi(s_j^*)}\,(t_k - t_{k-1})\,\left(s_j - s_{j-1}\right) =$$

$$= \sum_{k=1}^{n}\sum_{j=1}^{m} \left(\xi_k^*, \widetilde{\xi}_j^*\right)(t_k - t_{k-1})\left(s_j - s_{j-1}\right) = \left(\sum_{k=1}^{n} \xi_k^*\,(t_k - t_{k-1}), \sum_{j=1}^{m} \widetilde{\xi}_j^*\left(s_j - s_{j-1}\right)\right) = \left(\Upsilon_n, \widetilde{\Upsilon}_m\right) \tag{7.40}$$

Отсюда получаем

$$\int_a^b\int_a^b R(t,s)dtds \overset{def}{=} \lim_{m,n\to\infty} I_{n,m} = (\Upsilon, \Upsilon) \tag{7.41}$$

Наоборот, пусть для любых интегральных сумм

$$I_{n,m} = \left(\Upsilon_n, \widetilde{\Upsilon}_m\right) \to I \tag{7.42}$$

при $n, m \to \infty$. Тогда распишем квадрат нормы разности

$$\|\Upsilon_n - \Upsilon_m\|^2 = \|\Upsilon_n\|^2 - (\Upsilon_n, \Upsilon_m) - \overline{(\Upsilon_n, \Upsilon_m)} + \|\Upsilon_m\|^2 \underset{n,m'\to\infty}{\longrightarrow} (I - I - I + I) = 0 \tag{7.43}$$

Таким образом, $\{\Upsilon_n\}_{n=\overline{1,\infty}}$ фундаментальна и имеет предел.

Покажем, что предел не зависит от способа разбиения. Составим две разные интегральные суммы Υ_n и $\widetilde{\Upsilon}_n$, тогда $\Upsilon_n \to \Upsilon$ и $\widetilde{\Upsilon}_m \to \widetilde{\Upsilon}$ для некоторых $\Upsilon, \widetilde{\Upsilon} \in \mathcal{H}$. При этом

$$\|\Upsilon_n - \widetilde{\Upsilon}_m\|^2 = \left(\Upsilon_n, \widetilde{\Upsilon}_m\right) - \overline{\left(\Upsilon_n, \widetilde{\Upsilon}_m\right)} + \left(\widetilde{\Upsilon}_m, \widetilde{\Upsilon}_m\right) = I_{nn} - I_{nm} - \overline{I_{nm}} + I_{nm} \tag{7.44}$$

а из интегрируемости ковариационной функции следует, что все слагаемые в правой части стремятся к I, поэтому

$$\|\Upsilon_n - \widetilde{\Upsilon}_m\|^2 \to \quad \text{при} \;\; n, m \to \infty \tag{7.45}$$

Отсюда получаем

$$\|\Upsilon - \widetilde{\Upsilon}\| \leqslant \|\Upsilon - \Upsilon_n\| + \|\Upsilon_n - \widetilde{\Upsilon}_m\| + \|\widetilde{\Upsilon}_m - \widetilde{\Upsilon}\| \to 0 \;\; \text{при} \;\; n, m \to \infty \tag{7.46}$$

это возможно, только если

$$\|\Upsilon - \widetilde{\Upsilon}\|^2 = 0 \tag{7.47}$$

то есть

$$\Upsilon = \widetilde{\Upsilon} \tag{7.48}$$

с вероятностью единица. Итак, теорема доказана.

В математическом анализе есть утверждение: если функция непрерывна, то она интегрируема. Можно ли утверждать, что если случайный процесс среднеквадратично непрерывен, то он среднеквадратично интегрируем? Выясним это. Ранее мы доказали, что случайный процесс среднеквадратично непрерывен тогда и только тогда, когда его ковариационная функция непрерывная на диагонали, а следовательно, и на всей плоскости. Итак, если процесс среднеквадратично непрерывен, то его ковариационная функция как функция двух переменных, непрерывна на всей плоскости. Для того, чтобы процесс был среднеквадратично интегрируем, необходимо и достаточно, чтобы его ковариационная функция была интегрируема на плоскости. Отсюда получаем, что указанное следствие имеет место, так как свойства процесса идентичны свойствам его ковариационной функции.

Случай ненулевого математического ожидания случайного процесса.

Довольно часто мы имеем дело с процессами, у которых ненулевое среднее. Рассмотрим случай ненулевого математического ожидания случайного процесса. Пусть

$$M\xi(t) = a(t) \not\equiv 0 \tag{7.49}$$

и

$$M|\xi(t) - a(t)|^2 < \infty \tag{7.50}$$

для всех $t \in \mathbb{R}$.

Положим

$$\xi^o(t) = \xi(t) - a(t) \tag{7.51}$$

тогда этот центрированный случайный процесс будет гильбертовым при указанных условиях $\xi^o \in \mathcal{H}$. Ковариационная функция процесса с ненулевым средним определяется как коэффициент ковариации

$$R(t,s) = cov\left(\xi(t), \xi(s)\right) \overset{def}{=} M\xi^o(t)\overline{\xi^o(s)} \tag{7.52}$$

где $R^o(\cdot)$ – ковариационная функция процесса $\xi^o(t), t \in \mathbb{R}$.

Замечание. Пусть $\left\{\xi_n^o\right\} \subset \mathcal{H}$, а $\{a_n\}$ – произвольная неслучайная последовательность комплексных чисел. Тогда с учетом

$$M\xi_n^o = 0 \tag{7.53}$$

получаем

$$M|\xi_n^o - a_n|^2 = M|\xi_n^o|^2 + a_n \cdot M\xi_n^o + \overline{a_n \cdot M\xi_n^o} + |a_n|^2 = M|\xi_n^o|^2 + |a_n|^2 \tag{7.54}$$

отсюда получаем, что сходимость

$$M|\xi_m^o + a_n|^2 \to 0 \tag{7.55}$$

возможна если и только если

$$\xi_n^o \xrightarrow{\text{с.к.}} 0 \tag{7.56}$$

и последовательность комплексных чисел как числовая последовательность тоже сходится к нулю:

$$a_n \to 0 \text{ при } n \to \infty \tag{7.57}$$

Замечание. Если

$$\xi_n \xrightarrow{\text{с.к.}} \xi \tag{7.58}$$

то можно записать оценку

$$|M\xi_n - M\xi| = |M\left(\xi_n - \xi\right)| \leqslant \sqrt{M|\xi_n - \xi|^2} \underset{n\to\infty}{\longrightarrow} 0 \tag{7.59}$$

Другими словами, имеет место сходимость математических ожиданий

$$M\xi_n \to M\xi \tag{7.60}$$

Также если $M|\xi|^2 < \infty$, то

$$|M\xi|^2 \leqslant M|\xi|^2 < \infty \tag{7.61}$$

В результате для приращения случайного процесса имеем

$$M|\xi(t+h_n)-\xi(t)|^2 = M|\xi^o(t+h_n)-\xi^o(t)+(a(t+h_n)-a(t))\,|^2 = \\ = M|\xi^o(t+h_n)-\xi^o(t)|^2+|a(t+h_n)-a(t)|^2 \tag{7.62}$$

Отсюда получаем, что сходимость

$$\xi(t+h_n)\xrightarrow{\text{с.к.}}\xi^o(t) \tag{7.63}$$

тогда и только тогда, когда

- есть среднеквадратичная сходимость

$$\xi^o(t+h_n)\xrightarrow{\text{с.к.}}\xi^o(t) \tag{7.64}$$

- есть сходимость числовой последовательности

$$a(t+h_k)\to a(t)\ \text{ при }\ n\to\infty \tag{7.65}$$

Теорема. Случайный процесс $\xi(t), t\in\mathbb{R}$, с математическим ожиданием

$$M\xi(t)=a(t) \tag{7.66}$$

среднеквадратично непрерывен в точке t тогда и только тогда, когдаковариационная функция $R(\cdot)\equiv R^o(\cdot)$ непрерывна на диагонали, то есть в точке (t,t) и функция $a(\cdot)$ непрерывная в точке t.

Для дифференциальных отношений для нецентрированного

$$\Delta_n=\frac{\xi(t+h_n)-\xi(t)}{h_n} \tag{7.67}$$

и центрированного процессов

$$\Delta_n^o=\frac{\xi^o(t+h_n)-\xi^o(t)}{h_n} \tag{7.68}$$

Запишем дифференциальное отношение

$$(Da)_n=\frac{a(t+h_n)-a(t)}{h_n} \tag{7.69}$$

Тогда

$$\Delta_n-\Delta_m=\Delta_n^o-\Delta_m^o+(Da)_n-(Da)_m \tag{7.70}$$

Получаем

$$M|\Delta_n - \Delta_m|^2 = M|\Delta_n^o - \Delta_m^o|^2 + |\,(Da)_n - (Da)_m\,|^2 \tag{7.71}$$

В результате получаем, что последовательность $\{\Delta_n\}$ среднеквадратично фундаментальна (следовательно, сходится) тогда и только тогда, когда $\left\{\Delta_n^o\right\}$ среднеквадратично фундаментальна (а значит, сходится) и $\{(Da)_n\}$ фундаментальна (а значит, сходится).

Теорема. Случайный процесс $\xi(t), t \in \mathbb{R}$, с математическим ожиданием

$$M\xi(t) = a(t) \tag{7.72}$$

среднеквадратично дифференцируем в точке t тогда и только тогда, когда его ковариационная функции $R(\cdot)$ имеет вторую смешанную производную в точке (t, t) и $a(\cdot)$ дифференцируема в точке t.

При этом операции взятия среднеквадратичной производной и вычисления математического ожидания можно менять местами:

$$M\dot{\xi}(t) = \frac{d}{dt}M\xi(t) \tag{7.73}$$

Действительно, если

$$\Delta \xrightarrow{\text{с.к.}} \dot{\xi} \tag{7.74}$$

при $h_{\to}0$, то математическое ожидание

$$M\Delta_n \to M\dot{\xi} \tag{7.75}$$

Другими словами,

$$M\dot{\xi}(t) = \lim_{n\to\infty} M\frac{\xi(t+h_n) - \xi(t)}{h_n} = \lim_{n\to\infty}\frac{a(t+h_n) - a(t)}{h_n} = \dot{a}(t) \tag{7.76}$$

Итак, математическое ожидание производной есть производная математического ожидания. Отметим, что и математическое ожидание (интеграл), и производная – это предельные переходы. Значит, мы доказали, что в указанном случае два предельных перехода можно менять местами. Аналогично получаются следующие утверждения.

Теорема. Случайный процесс $\xi(t), t \in \mathbb{R}$ с математическим ожиданием

$$M\xi(t) = a(t) \tag{7.77}$$

среднеквадратично интегрируем на отрезке $[a, b]$ тогда и только тогда, когда ковариационная функция $R(\cdot)$ интегрируема в квадрате $[a, b] \times [a, b]$ и $a(\cdot)$ интегрируема на отрезке $[a, b]$. При этом математическое ожидание интеграла есть интеграл от математического ожидания

$$M\int_a^b \xi(t)dt = \int_a^b M\xi(t)dt = \int_a^b a(t)dt \tag{7.78}$$

Лекция 8. Стохастическая ортогональная мера.

Стохастичекая ортогональная мера на отрезке $[a, b)$.

На прошлой лекции мы обсудили теорию второго порядка, которая позволяет строить среднеквадратичные производные и интегралы от случайного процесса. Если мы моделируем физическое явление с помощью случайного процесса, который зависит от времени, логично, что нам нужно знать динамику этого процесса по времени. Чтобы охарактеризовать эту динамику, мы фактически пользуемся таким понятием, как дифференциал процесса. Дифференциал – это бесконечно малое изменение. Берем некоторый малый конечный промежуток времени, рассчитываем, как меняется характеристика на этом промежутке, потом делаем некоторые математические построения. После этого устремляем длину промежутка времени к нулю. То, что получаем в результате, называется дифференциалом. Однако, в реальных адекватных моделях случайный процесс зачастую зависит не столько от времени, сколько от каких-то других параметров, которые сами зависят от времени. Тогда динамика случайного процесса будет определяться динамикой этих параметров. При этом нас может интересовать изменение процесса как функции от этих параметров. Тогда возникает такое понятие, как производная одного случайного процесса по другому случайному процессу. Значит, нам нужно построить такую математическую теорию, которая сможет интерпретировать, задавать и позволять работу с таким объектном. Для того, чтобы задать такой объект, нужно определить дифференциал, который стоит в знаменателе производной. Отметим, что любой дифференциал – это некоторая мера, которая определяет величину промежутка, по которому мы задаем интегрирования или дифференцирования. Более того, нам нужно учесть, что мера имеет случайный характер.

Пусть $-\infty < a < b < \infty$ и

$$\mathbb{D} = \{\Delta x = [x_1, x_2) \subset [a, b)\}, \quad -\infty < a < b < \infty \tag{8.1}$$

Рассмотрим случайный процесс, заданный на указанном множестве

$$Z(\Delta x): \quad \Omega \to \mathbb{C}, \ \Delta x \in \mathbb{D} \tag{8.2}$$

Отметим, что в данном случае случайный процесс в качестве аргумента получает множество.

Определение. Назовем случайный процесс $Z(\Delta x), \Delta x \in \mathbb{D}$ **стохастической ортогональной мерой**, если

- он является гильбертовым, то есть любое его сечение является гильбертовой случайной величиной. Это значит, что математическое ожидание равно нулю

$$MZ(\Delta x) = 0 \tag{8.3}$$

а математическое ожидание квадрата модуля конечно

$$D(\Delta x) = M|Z(\Delta x)|^2 < \infty \tag{8.4}$$

Функция, аргументом которой является интервал

$$m(\cdot) : \ \mathbb{D} \to \mathbb{R}_+ \tag{8.5}$$

задаваемая как математическое ожидание квадрата модуля:

$$m(\Delta x) = M|Z(\Delta x)|^2, \ \ \Delta x \in \mathbb{D} \tag{8.6}$$

называется **структурной функцией меры**.

- если мы берем два непересекающихся интервала$\Delta x_1 \cap \Delta x_2 = \emptyset$, то соответствующие сечения являются некоррелированными

$$cov\left(Z(\Delta x_1), Z\left(\Delta x_2\right)\right) = MZ(\Delta x_1)\overline{Z(\Delta x_2)} = (Z(\Delta x_1), Z(\Delta x_2)) = 0 \tag{8.7}$$

- *конечная аддитивность с вероятностью единица*. Пусть дан некоторый интервал. Разобьем его на непересекающиеся части

$$\Delta x = \Delta x_1 + \ldots + \Delta x_n \tag{8.8}$$

где $\Delta x_k \cap \Delta x_j = \emptyset$ при $k \neq j$, тогда мера всего большого интервала есть сумма мер всех частей

$$Z(\Delta x) = Z\left(\sum_{k=1}^{n} \Delta x_k\right) = \sum_{k=1}^{n} Z\left(\Delta x_k\right) \ \ \text{(с вероятностью 1)} \tag{8.9}$$

- *счетная аддитивность в среднем квадратичном*. Разобьем интервал на счетное число частей

$$\Delta x = \Delta x_1 + \Delta x_2 + \ldots \tag{8.10}$$

Тогда мера множества есть бесконечная сумма мер частей множества

$$Z(\Delta x) = X\left(\sum_{k=1}^{\infty} \Delta x_k\right) = \lim_{n\to\infty} \sum_{k=1}^{n} Z(\Delta x_k) \tag{8.11}$$

где предел понимается как среднеквадратичный (по норме $\mathcal{H}$)

$$M\left|Z\left(\sum_{k=1}^{\infty} \Delta x_k\right) - \sum_{k=1}^{n} Z(\Delta x_k)\right|^2 = \left\|Z\left(\sum_{k=1}^{\infty} \Delta x_k\right) - \sum_{k=1}^{n} Z(\Delta x_k)\right\|^2 \underset{n\to\infty}{\longrightarrow} 0 \tag{8.12}$$

Отметим, что мера является стохастической, так как она носит случайный характер (первое условие), ортогональной, так как значения мер на непересекающихся интервалах ортогональны как случайные величины в пространстве гильбертовых случайных величин (второе условие). При этом заданная величина называется именно мерой, так как она обладает свойствами конечной и счётной аддитивности. Однако, данная мера все же не является мерой в обычном понимании. Изначально мера (объем, площадь, вероятность) – неотрицательная величина. Однако в нашем случае мера не только может быть отрицательной, но и комплексной. "Настоящая"счётно-аддитивная мера – это структурная функция $m(\cdot)$. Воспользуемся свойством конечной аддитивности меры Z и докажем аддитивность структурной функции:

$$\begin{aligned} m\left(\Delta x_1+\Delta x_2\right) &= \|Z\left(\Delta x_1\right)+Z\left(\Delta x_2\right)\|^2 = \|Z\left(\Delta x_1\right)+Z\left(\Delta x_2\right)\|^2 = \\ &= \|Z\left(\Delta x_1\right)\|^2 + (Z(\Delta x_1), Z(\Delta x_2)) + \overline{(Z(\Delta x_1), Z(\Delta x_2))} + \|Z(\Delta x_2)\|^2 = \\ &= \|Z(\Delta x_1)\|^2 + \|Z(\Delta x_2)\|^2 = m(\Delta x_1) + m(\Delta x_2) \end{aligned} \tag{8.13}$$

Аналогично можно показать аддитивность структурной функции для любого конечного числа интервалов. Счётная аддитивность структурной функции есть следствие счётной аддитивности стохастической меры. Пусть интервал представлен в виде счетного числа непересекающихся интервалов

$$\Delta x = \Delta x_1 + \Delta x_2 + \dots \tag{8.14}$$

Положим

$$\zeta \overset{def}{=} Z(\Delta x) = Z\left(\sum_{k=1}^{\infty} \Delta x_k\right) \tag{8.15}$$

а для частичных сумм

$$\zeta_n \overset{def}{=} Z\left(\sum_{k=1}^{n} \Delta x_k = \sum_{k=1}^{n} Z(\Delta x_k)\right) \tag{8.16}$$

Условие счётной аддитивности стохастической меры принимает вид

$$M\left|\zeta-\zeta_n\right|^2 = \|\zeta-\zeta_n\|^2 \underset{n\to\infty}{\longrightarrow} 0 \tag{8.17}$$

Это влечет

$$\|\zeta_n\| \to \|\zeta\| \tag{8.18}$$

или, что эквивалентно

$$M|\zeta_n|^2 = \|\zeta_n\|^2 \to \|\zeta\|^2 = M|\zeta|^2 \tag{8.19}$$

при $n \to \infty$.

По определению структурной функции

$$M|\zeta|^2 = \lim M|\zeta_n|^2 \tag{8.20}$$

означает, что

$$m(\Delta x) = M|Z(\Delta x)|^2 = \lim_{n\to\infty} M\left|\sum_{k=1}^{n}|Z(\Delta x_k)|\right|^2 = \lim_{n\to\infty} M \sum_{k,\widetilde{k}=1}^{n} Z(\Delta x_k)\overline{Z(\Delta x_{\widetilde{k}})} =$$
$$= \lim_{n\to\infty}\sum_{k=1}^{n} M|Z(\Delta x_k)|^2 = \lim_{n\to\infty}\sum_{k=1}^{n} m(\Delta x_k) = \sum_{k=1}^{\infty} m(\Delta x_k) \tag{8.21}$$

Итак, получили счетную аддитивность структурной функции. Мы построили ортогональную стохастическую меру с определенными свойствами. Показали, почему этот случайный процесс можно понимать как меру. При этом его структурная функция обладает свойствами обычной меры.

Интеграл по случайной мере от неслучайной функции.

Наша конечная цель – построение производной и интеграла по случайному процессу. Поэтому далее нас будет интересовать не столько мера, сколько интеграл по этой мере. Чтобы построить этот интеграл, рассмотрим несколько этапов. Сначала рассмотрим интеграл от неслучайной функции по случайной мере, а потом построим интеграл от случайной функции по случайной мере.

Предположим, что найдется

$$F(\cdot) : [a, b) \to \mathbb{R}_+ \tag{8.22}$$

такая, что для любого интервала

$$\Delta x = [x_1, x_2) \subset [a, b) \tag{8.23}$$

значение структурной функции можно представить в виде разности некоторой функции на концах рассматриваемого интервала

$$m(\Delta x) = \int\limits_{\Delta x} dF(x) = F(x_2) - F(x_1) \tag{8.24}$$

Отметим, что здесь подразумевается интеграл Лебега-Стилтьеса. Мера в этом интеграле порождена функцией (8.22). Эта функция F нужна для аккуратного определения множества функций, которые мы будем интегрировать.

Введем линейное пространство $\mathcal{L}^2$ функций из $[a, b)$ в $\mathbb{C}$ со скалярным произведением

$$(g_1, g_2) = \int_a^b g_1(x)\overline{g_2(x)}dF(x) \tag{8.25}$$

и нормой

$$\|g\|^2 = \int_a^b |g(x)|^2 dF(x) < \infty \tag{8.26}$$

Все записанные здесь интегралы – интегралы Лебега-Стильтьеса. Норма здесь квадратичная, поротому в обозначении пространства $\mathcal{L}^2$ появляется 2.

Теорема (без доказательства). Пространство $\mathcal{L}^2$ полно.

Это означает, что если есть некоторая фундаментальная последовательность функций этого пространства, то она обязательно сходится но норме пространства $\mathcal{L}^2$.

Всюду далее будем считать, что функции g являются функциями из пространства $\mathcal{L}^2$.

Сперва рассмотрим **интеграл от кусочно-постоянных функций.** Рассмотрим разбиение интервала $[a, b)$ точками

$$a = x_0 < x_1 < \ldots < x_n = b \tag{8.27}$$

и функцию, кусочно-постоянную на отрезках

$$\Delta x_k = [x_{k-1}, x_k), \quad k = 1, \ldots, n : \tag{8.28}$$

Тогда зададим функцию в виде

$$g(x) = \sum_{k=1}^{n} c_k \chi_k(x), \quad c_1, \ldots, c_n \in \mathbb{C} \tag{8.29}$$

где

$$\chi_k(x) \equiv \chi_{\Delta x_k}(x) = \begin{cases} 1, & x \in \Delta x_k \\ 0, & x \notin \Delta x_k \end{cases} \tag{8.30}$$

– индикаторная функция интервала.

Для такой функции g положим по определению

$$\mathcal{I} = \int_a^b g(x)Z(dx) = \int_a^b \sum_{k=1}^{n} c_k \chi_k(x) Z(dx) = \sum_{k=1}^{n} c_k Z(\Delta x_k) \tag{8.31}$$

При этом по построению

$$M\mathcal{I} = 0 \tag{8.32}$$

Таким образом, получили интеграл от кусочно-постоянной функции. Существование интеграла как случайной величины очевидно, так как это просто линейная комбинация случайных величин.

Покажем, что интеграл от суммы или разности кусочно-постоянных функций есть соответственно сумма или разность интегралов. Для начала заметим, что сумма или разность кусочно-постоянных функций тоже есть кусочно-постоянная функция.

Пусть даны две кусочно-постоянные функции

$$g(x) = \sum_{k=1}^{n} c_k \chi_{\Delta x_k}(x) \tag{8.33}$$

$$h(x) = \sum_{j=1}^{n} a_j \chi_{\Delta z_j}(x) \tag{8.34}$$

где точки разбиения соответственно

$$a = x_0 < x_1 < \ldots < x_n = b \tag{8.35}$$

$$a = z_0 < z_1 < \ldots < z_m = b \tag{8.36}$$

Возьмем объединение множеств точек разбиения

$$\{y_0, y_1, \ldots, y_r\} = \{x_0, x_1, \ldots, x_n\} \cup \{z_0, z_1, \ldots, z_m\} \tag{8.37}$$

при этом

$$a = y_0 < y_1 < \ldots < y_t = b \tag{8.38}$$

где y_l совпадает с одним из x_k или z_k и

$$max\ (m, n) \leqslant r \leqslant m + n \tag{8.39}$$

Тогда функцию g можно считать кусочно-постоянной на интервалах Δy_l, а не на интервалах Δx_k

$$g(x) = \sum_{k=1}^{n} c_k \chi_{\Delta x_k}(x) = \sum_{l=1}^{r} \hat{c}_l \chi_{\Delta y_l}(x) \tag{8.40}$$

Если $y + l = z_j$ лежит строго внутри одного из $[x_{k-1}, x_k)$, то $\hat{c}_l = \hat{c}_{l+1}$. Аналогично можно считать

$$h(x) = \sum_{l=1}^{r} \hat{a}_l \chi_{\Delta y_l}(x) \tag{8.41}$$

Тогда интеграл от суммы или разности

$$\int_a^b (g(x) \pm h(x))\, Z(dx) = \int_a^b \sum_{l=1}^r (\hat{c}_l \pm \hat{a}_l)\, \chi_{\Delta y_l}(x) Z(dx) = \sum_{l=1}^r (\hat{c}_l \pm \hat{a}_l)\, Z(\Delta y_l) =$$

$$= \sum_{l=1}^r \hat{c}_l Z(\Delta y_l) \pm \sum_{l=1}^r \hat{a}_l Z(\Delta y_l) = \int_a^b g(x) Z(dx) \pm \int_a^b h(x) Z(dx) \quad (8.42)$$

Итак, мы показали, что множество кусочно-постоянных функций обладает замкнутости относительно линейных операций (линейные операции не выводят из множества). То есть, линейная комбинация кусочно-постоянных функций есть кусочно-постоянная функция. Тогда и интеграл от линейной комбинации кусочно-постоянных функций равен соответствующей линейной комбинации интегралов.

Лемма 1. Пусть есть некоторая кусочно-постоянная функция

$$g(x) = \sum_{k=1}^n c_k \chi_k(x) \quad (8.43)$$

и интеграл от нее

$$\mathcal{I} = \sum_{k=1}^n c_k Z(\Delta x_k) \quad (8.44)$$

Тогда эти два объекта как элементы соответствующих пространств, имеют одинаковую норму

$$\|\mathcal{I}\|^2_{\mathcal{H}} = \|g\|^2_{\mathcal{L}^2} = \sum_{k=1}^n \int_{\Delta x_k} |c_k|^2 dF(x) \quad (8.45)$$

Доказательство. Имеем в силу ортогональности меры

$$\|\mathcal{I}\|^2 = M\left|\sum_{k=1}^n c_k Z(\Delta x_k)\right|^2 = \left\|\sum_{k=1}^n c_k Z(\Delta x_k)\right\|^2 = \sum_{k=1}^n |c_k|^2 \cdot \|Z(\Delta x_k)\|^2 =$$

$$= \sum_{k=1}^n |c_k|^2 \cdot M|Z(\Delta x_k)|^2 = \sum_{k=1}^n |c_k|^2 m(\Delta x_k) = \sum_{k=1}^n |c_k|^2 \int_{\Delta x_k} dF(x) = \sum_{k=1}^n \int_{\Delta x_k} |c_k|^2 dF(x) \quad (8.46)$$

Таким образом, мы доказали лемму в одну сторону.

С другой стороны,

$$\|g\|^2 = \int\limits_a^b |g(x)|^2 dF(x) = \int\limits_a^b \left|\sum_{k=1}^n c_k \chi_k(x)\right|^2 dF(x) = \sum_{k,\tilde{k}=1}^n \int\limits_a^b c_k \tilde{c}_k \chi_k(x) \overline{\chi_{\tilde{k}}(x)} dF(x) =$$

$$= \sum_{k=1}^n \xi_j \int\limits_{\Delta x_k} |c_k|^2 dF(x) \qquad (8.47)$$

где мы учли, что

$$\tilde{\chi}_k(x)\overline{\tilde{\chi}_{\tilde{k}}(x)} \equiv 0 \ \text{ при } \ k \neq \tilde{k} \qquad (8.48)$$

а при $k = \tilde{k}$

$$\int\limits_a^b |c_k|^2 |\chi_k(x)|^2 dF(x) = \int\limits_a^b |c_k|^2 \chi_k(x) dF(x) = \int\limits_{\Delta x_k} |c_k|^2 dF(x) \qquad (8.49)$$

Отметим, что на самом деле равенство норм является следствием ортогональности слагаемых в мере.

Лемма доказана.

Далее воспользуемся построенной конструкцией для построения интеграла от более сложных функций. Пусть функцию

$$g(\cdot) : [a, b) \to \mathbb{C} \qquad (8.50)$$

можно представить как следующий предел кусочно-постоянных функций по норме пространства $\mathcal{L}^2$:

$$g(x) = \lim_{n\to\infty} \sum_{k=1}^n g(x_k^*) \chi_{\Delta x_k}(x) \qquad (8.51)$$

где точки разбиения

$$a = x_0 < x_1 < \ldots < x_n = b \qquad (8.52)$$

точки на сегменте

$$x_k^* \in [x_{k-1}, x_k) \qquad (8.53)$$

для $k = 1, \ldots, n$ и максимальная длина сегмента разбиения стремится к нулю

$$\max_{1\leqslant k\leqslant k\leqslant n} (x_k - x_{k-1}) \xrightarrow[n\to\infty]{} 0 \qquad (8.54)$$

Введем следующее обозначение для частичной суммы

$$g_n(x) = \sum_{k=1}^{n} g(x_k^*)\chi_{\Delta x_k}(x) \tag{8.55}$$

Мы знаем, что для такой функции существует интеграл

$$\mathcal{I}_n = \int_a^b g_n(x)Z(dx), \quad n = 1, 2, \ldots \tag{8.56}$$

В силу сходимости

$$\|g_n - g\| \underset{n\to\infty}{\longrightarrow} 0 \tag{8.57}$$

имеем фундаментальности последовательности

$$\|g_n - g_m\| \underset{n,m\to\infty}{\longrightarrow} 0 \tag{8.58}$$

Тогда, в силу доказанной леммы и того, что $g_n - g_m$ – кусочно-постоянная функция, получаем

$$\|\mathcal{I}_n - \mathcal{I}_m\|^2 = \|g_n - g_m\|^2 \underset{n,m\to\infty}{\longrightarrow} 0 \tag{8.59}$$

Из полноты пространства $\mathcal{H}$ вытекает, что существует случайная величина (предел), которую мы и будем называть интегралом от функции g

$$\mathcal{I} \overset{def}{=} \int_a^b g(x)Z(dx) = \lim_{n\to\infty} \int_a^b g_n(x)Z(dx) \tag{8.60}$$

Итак, мы представили функцию g как предел кусочно-постоянных функций, интеграл от которых нам известен. Далее мы интеграл от предела задали как предел интегралов.

Проверим, что определение корректно. Покажем, что предел не зависит от представления функции $g(\cdot)$ в виде (8.51). Пусть при надлежащих условиях на разбиения функция представима в двух видах

$$g = \lim_{n\to\infty} g_n = \lim_{m\to\infty} \widetilde{g}_m \tag{8.61}$$

где

$$g_n(x) = \sum_{k=1}^{n} g(x_k^*)\chi_{\Delta x_k}(x) \tag{8.62}$$

$$\widetilde{g}_m(x) = \sum_{j=1}^{m} g(z_j^*)\chi_{\Delta z_j}(x) \tag{8.63}$$

Другими словами,

$$\|g - g_n\| \to 0 \ \text{при} \ n \to \infty \tag{8.64}$$

$$\|g - \widetilde{g}_m\| \to 0 \ \text{при} \ m \to \infty \tag{8.65}$$

Следовательно, существуют пределы интегралов $\mathcal{I}, \widetilde{\mathcal{I}} \in \mathcal{H}$, такие, что

$$\mathcal{I} = \lim_{n\to\infty} \int_a^b g_n(x) Z(dx) = \lim_{n\to\infty} \mathcal{I}_n \tag{8.66}$$

$$\widetilde{\mathcal{I}} = \lim_{n\to\infty} \int_a^b \widetilde{g}_n(x) Z(dx) = \lim_{m\to\infty} \widetilde{\mathcal{I}}_m \tag{8.67}$$

Функция $g_n - \widetilde{g}_m$ кусочно-постоянная, поэтому для нормы разности имеет место выражение

$$\|\mathcal{I}_n - \widetilde{\mathcal{I}}_m\| = \|g_n - \widetilde{g}_m\| \tag{8.68}$$

Отсюда получаем оценку для нормы разности

$$\begin{aligned} \|\mathcal{I} - \widetilde{\mathcal{I}}\| \leqslant \|\mathcal{I} - \mathcal{I}_n\| + \|\mathcal{I}_n - \widetilde{\mathcal{I}_m}\| + \|\widetilde{\mathcal{I}_m} - \widetilde{\mathcal{I}}_n\| = \|\mathcal{I} - \mathcal{I}_n\| + \|g_n - \widetilde{g}_m\|| + \|\widetilde{\mathcal{I}_m} - \widetilde{\mathcal{I}}\| \leqslant \\ \leqslant \|\mathcal{I} - \mathcal{I}_n\| + \|g_n - g\| + \|g - \widetilde{g}_m\| + \|\widetilde{\mathcal{I}_m} - \widetilde{\mathcal{I}}\| \to 0, \qquad m, n \to \infty \end{aligned} \tag{8.69}$$

Это означает, что мы можем сделать выражение в правой части сколько угодно малым, если выберем достаточно большие номера n и m. Значит, можно сделать сколь угодно малой норму разности в левой части. Однако она не зависит от n и m, значит, сколь угодно малой она будет только в случае

$$\|\mathcal{I} - \widetilde{\mathcal{I}}\|^2 = 0 \tag{8.70}$$

тем самым

$$\mathcal{I} = \widetilde{\mathcal{I}} \tag{8.71}$$

с вероятностью единица.

По сути мы задали интеграл как предел последовательности интегральных сумм но норме пространства $\mathcal{H}$

$$\int_a^b g(x) Z(dx) = \lim_{n\to\infty} \sum_{k=1}^{n} g_n(x_k^*) Z(\Delta x_k) \tag{8.72}$$

где сегменты и точки разбиения

$$\Delta x_k = [x_{k-1}, x_k), \quad x_k^* \in \Delta x_k, \quad k = 1, \ldots, n \tag{8.73}$$

$$a = x_0 < x_1 < \ldots < x_n = b \tag{8.74}$$

при этом максимальная длина сегмента разбиения стремится к нулю

$$\max_{1 \leqslant k \leqslant n} (x_k - x_{k-1}) \underset{n \to \infty}{\longrightarrow} 0 \tag{8.75}$$

Интеграл по случайной мере от случайной функции.

Перейдем к построению интеграла от случайной функции по случайной мере. В сущности для этого нам нужно заменить в (8.72) неслучайные числа $g_n\left(x_k^*\right)$ на сечения случайного процесса в точках x_k^*.

Пусть $\xi(x), x \in [a, b)$, есть гильбертов случайный процесс, тогда будем рассматривать интеграл как предел по норме пространства $\mathcal{H}$ вида

$$\int\limits_a^b \xi(x) Z(dx) = \lim_{n \to \infty} \sum_{k=1}^{n} \xi(x_k^*) Z(\Delta x_k) \tag{8.76}$$

Однако полученная конструкция существенно сложнее предыдущей. Здесь появляется два случайных процесса, поэтому есть проблема совместного распределения $\xi(x), \quad x \in [a, b)$ и $Z(\Delta x), \quad \Delta x \in \mathbb{D}$. Наличие совместного распределения приводит к существенным математическим проблемам. Дело в том, что до сих пор все формулы были линейными по случайным величинам. Здесь же линейность пропадает и все предыдущие выкладки могут быть неверными. Можно упростить задачу и рассмотреть только те случаи, в которых случайные процессы независимы. В этом случае теория существенно упрощается, однако такой подход бессмысленнен в отношении построения реальных моделей. Вообще говоря, все построенные конструкции нужны нам для построения производной или интеграла одного случайного процесса по другому. Однако, эти понятия имеют смысл только в случае функциональной зависимости рассматриваемых случайных процессов. Если случайные процессы зависимы функционально, то они зависимы и статистически.

С другой стороны, можно формально предположить, что нам известно совместное распределение случайных процессов. В этом случае можно построить некую теорию, однако она будет очень сложна и неудобна для использования. Значит, нам нужно уйти от полной независимости к произвольной, совместной зависимости ξ и Z.

Интеграл Ито.

Далее мы будем рассматривать стохастическую меру, заданную некоторым специфическим образом. Интеграл по этой мере называется интегралом Ито. В этом интеграле мера порождена процессом Винера.

Пусть $w(t), t \geqslant 0$, – стандартный процесс Винера ($\sigma^2 = 1$).

Введем стохастическую меру на интервале $[0, T)$: для $\Delta t = [t_1, t_2)$

$$Z_w(\Delta t) = w(t_2) - w(t_1) = \Delta w(t_2 - t_1), \quad 0 \leqslant t_1 < t_2 < T \tag{8.77}$$

Независимость приращений ортогональность

$$m(\Delta t) = M\left|\Delta w\left(t_2 - t_1\right)\right|^2 = t_2 - t_1 \tag{8.78}$$

аддитивность стохастической меры. Интеграл от $\xi(t), t \in [0, T]$ задаем как (среднеквадратичный) предел:

$$\int\limits_0^T \xi(t) dw(t) = \lim_{n\to\infty} \sum_{k=1}^{n} \xi(t_k^*) \Delta w_k \tag{8.79}$$

$$\Delta w_k = w(t_k) - w(t_{k-1}) \tag{8.80}$$

Лекция 9. Стохастические дифференциалы.

Интеграл по случайной мере от случайной функции.

Ранее мы пытались построить интеграл от случайного процесса по случайной мере. Пусть $\xi(t),\ x \in [a, b)$ есть гильбертов случайный процесс, тогда

$$\int_a^b \xi(x)Z(dx) = \lim_{n\to\infty} \sum_{k=1}^{n} \xi(x_k^*)Z(\Delta x_k) \tag{9.1}$$

по норме пространства $\mathcal{H}$

Проблема совместного распределения $\xi(x), x \in [a, b)$ и $Z(\Delta x), \Delta x \in \mathbb{D}$

Интеграл Ито.

Далее мы будем задавать случайную меру через процесс Винера. В этом случае интеграл от случайной функции называется интегралом Ито.

Пусть $w(t), t \geqslant 0$, – стандартный процесс Винера ($\sigma^2 = 1$). Введем стохастическую меру на интервале $[0, T)$: для $\Delta t = [t_1, t_2)$ как приращение процесса Винера на этом интервале

$$Z_w(\Delta t) = w(t_2) - w(t_1) = \Delta w(t_2 - t_1), \quad 0 \leqslant t_1 < t_2 < T \tag{9.2}$$

Из независимости приращений следует ортогональность

$$m(\Delta t) = M\left|\Delta w\ (t_2 - t_1)\right|^2 = t_2 - t_1 \tag{9.3}$$

Отсюда получаем аддитивность стохастической меры. Интеграл от $\xi(t), t \in [0, T]$ задаем как (среднеквадратичный) предел последовательности интегральных сумм

$$\int_0^T \xi(t)dw(t) = \lim_{n\to\infty} \sum_{k=1}^{n} \xi(t_k^*)\Delta w_k \tag{9.4}$$

где

$$\Delta w_k = w(t_k) - w(t_{k-1}) \tag{9.5}$$

Мы работаем со случайными величинами, поэтому пределы понимаются в смысле пространства гильбертовых случайных величин.

Ранее мы заметили, что существует проблема в определении интеграла, связанная с совместным распределением двух случайных величин. В данном случае эти величины – $\xi(t)$ и $w(t)$. Рассматривать ситуацию их полной независимости бессмысленно, однако, нам все еще нужно как-то охарактеризовать статистическую зависимость между этими случайными процессами.

Сделаем следующее предположение. Пусть *вероятностные свойства сечения* $\xi(t)$ (то есть статистическое поведение этого сечения) *не зависят от вероятностных свойств процесса Винера* $w(s)$ *при* $s > t$, *но, возможно, зависят от процесса Винера* $w(t)$ *при* $0 \leqslant s \leqslant t$. Такое предположение довольно физично. Мы пытаемся исследовать динамику процесса $\xi(t)$ (который представляет собой результат наблюдения) учитывая, что этот процесс функционально зависит от процесса $w(t)$. Вряд ли можно ожидать, что результат наблюдения в момент времени t связан с какими-то будущими моментами. С другой стороны, мы допускаем, что ξ функционально определяется процессом w во все моменты времени вплоть до момента времени t. Это означает, что процесс ξ не опережает процесс Винера w.

Пример. Найти интеграл от процесса Винера по мере от процесса Винера

$$\int_0^T w(t)dw(t)$$

Составим интегральную сумму как

$$\mathcal{I}_n = \sum_{j=1}^{n} w(t_{j-1}) \left[w(t_j) - w(t_{j-1})\right] \qquad (t_j^* = t_{j-1}) \tag{9.6}$$

В квадратных скобках стоит мера интервала разбиения, а подынтегральная функция взята в выборочной точке, равной левому концу интервала. Значит, построенная интегральная сумма довольно конкретного вида (она зависит от выбранной точки). Заметим, что интеграл от случайного процесса есть случайная величина. Значит, нельзя ничего сказать о конкретном значении этого интеграла. Скорее всего нам нужно получить ответ в форме: этот интеграл есть случайная величина с известным распределением. Значит, нам нужно показать, что последовательность интегральных сумм (9.6) сходится и исследовать распределение предельной случайной величины. Для этого нам нужно преобразовать слагаемые в сумме (9.6), чтобы можно было что-то сказать о ее распределении. Имеет место очевидная цепочка равенств

$$2b(a-b) = 2ab - 2b^2 = a^2 - b^2 - (a-b)^2 \tag{9.7}$$

Тогда можно преобразовать

$$w(t_{j-1}) \left[w(t_j) - w(t_{j-1})\right] = \frac{1}{2}\left[w^2(t_j) - w^2(t_{j-1})\right] - \frac{1}{2}\left[w(t_j) - w(t_{j-1})\right]^2 \tag{9.8}$$

Тогда интегральная сумма разбивается на две

$$\mathcal{I}_n = \frac{1}{2}\sum_{j=1}^{n}\left[w^2(t_j) - w^2(t_{j-1})\right] + \frac{1}{2}\sum_{j=1}^{n}\left(\Delta w_j\right)^2 = \frac{1}{2}\Big(w^2(t_n) - w^2(t_0)\Big) + \frac{1}{2}\sum_{j=1}^{n}\left(\Delta w_j\right)^2 \quad (9.9)$$

Первую из этих сумм просто посчитать. Под знаком суммы стоит разность выражений в двух соседних точках. При раскрытии суммы все слагаемые, кроме первого и последнего, сократятся. Вторая сумма – случайная величина, представляющая собой сумму квадратов приращений процесса Винера. Математическое ожидание этой суммы есть сумма математического ожидания, так как сумма конечна. При этом

$$M\sum_{j=1}^{n}\left(\Delta w_j\right)^2 = \sum_{j=1}^{n} M\left(\Delta w_j\right)^2 = \sum_{j=1}^{n}\left(t_j - t_{j-1}\right) = T \quad (9.10)$$

В силу независимости приращений процесса Винера дисперсия суммы есть сумма дисперсий. Учтем это и распишем дисперсию квадрата приращения

$$D\sum_{j=1}^{n}\left(\Delta w_j\right)^2 = \sum_{j=1}^{n} D\left(\Delta w_j\right)^2 = \sum_{j=1}^{n}\left[M\left(\Delta w_j\right)^4 - \left(M\left(\Delta w_j\right)^2\right)^2\right] \quad (9.11)$$

Находим $M(\Delta w_j)^4$, где

$$\Delta w_j \in N\left(0, t_j - t_{j-1}\right) \quad (9.12)$$

поэтому

$$M(\Delta w_j)^4 = 3\left(t_j - t_{j-1}\right)^2 \quad (9.13)$$

Тогда получаем оценку для дисперсии

$$\begin{aligned} D\left(\Delta w_j\right)^2 &= 3\left(t_j - t_{j-1}\right)^2 - \left(t_j - t_{j-1}\right)^2 = 2\left(t_j - t_{j-1}\right)^2 \leqslant \\ &\leqslant 2\left[\max_{j=1,\dots,n}\left(t_j - t_{j-1}\right)\right]\cdot\left(t_j - t_{j-1}\right) = 2h_n\left(t_j - t_{j-1}\right) \end{aligned} \quad (9.14)$$

где

$$h_n = \max_{j=1,\dots,n}\left(t_j - t_{j-1}\right) \quad (9.15)$$

– фактически шаг разбиения. Отсюда получаем, что математическое ожидание

$$M\sum_{j=1}^{n}\left(\Delta w_j\right)^2 = T \quad (9.16)$$

а для дисперсии имеем оценку

$$D\sum_{j=1}^{n}\left(\Delta w_j\right)^2 \leqslant 2h_n\sum_{j=1}^{n}\left(t_j - t_{j-1}\right) = 2h_nT \underset{n\to\infty}{\longrightarrow} 0 \tag{9.17}$$

Таким образом

$$M\left[\sum_{j=1}^{n}\left(\Delta w_j\right)^2 - T\right]^2 \to 0 \tag{9.18}$$

или

$$\sum_{j=1}^{n}\left(\Delta w_j\right)^2 \xrightarrow{\text{с.к.}} T \tag{9.19}$$

Тогда получаем, что для интегральной суммы имеет место

$$\mathcal{I}_n = \frac{1}{2}\left(w^2(t_n) - w^2(t_0)\right) - \frac{1}{2}\sum_{j=1}^{n}\left(\Delta w_j\right)^2 = \frac{w^2(T)}{2} - \frac{1}{2}\sum_{j=1}^{n}\left(\Delta w_j\right)^2 \xrightarrow{\text{с.к.}} \frac{w^2(T)}{2} - \frac{T}{2} \tag{9.20}$$

Тогда получаем, что интеграл от процесса Винера равен

$$\int_0^T w(t)dw(t) = \frac{w^2(T)}{2} - \frac{T}{2} \tag{9.21}$$

Можно показать, что добавочное слагаемое $-\frac{T}{2}$ связано с выбором $t_j^* = t_{j-1}$. При другом выборе получается другое добавочное слагаемое.

Стохастические дифференциальные уравнения.

Итак, мы построили теорию, позволяющую исследовать динамику случайных процессов. При этом наша теория учитывает зависимость рассматриваемого случайного процесса от других случайных параметров. Таким образом, для изучения динамики, нужно решать дифференциальные уравнения, в которых участвуют разные случайные процессы.

Рассмотрим стохастическое дифференциальное уравнение

$$\begin{cases} d\xi = \alpha(t)dt + \beta(t)dw(t) \\ \xi(0) = \xi_0 \end{cases} \quad 0 \leqslant t < T \tag{9.22}$$

где $\alpha(t), \beta(t), t \geqslant 0$ – некоторые случайные процессы. Его решение есть

$$\xi(t) = \xi_0 + \int_0^t \alpha(t)dt + \int_0^t \beta(\xi)dw(t) \tag{9.23}$$

где первый интеграл понимается как среднеквадратичный, а второй – как стохастический интеграл по мере, порожденной процессом Винера.

Отметим, что указанные уравнения представляют собой широко использующуюся модель наблюдений. Например, в каждый момент времени $\xi(t)$ – результат нашего наблюдения. Мы интересуемся динамикой ξ. В уравнении (9.22) есть два слагаемых. Рассмотрим первое слагаемое. Функция $\alpha(t)$ описывает наши физические представления о динамике нашей физической характеристики. В сущности это означает, что величина $\alpha(t)$ описывает изменение того, что мы измеряем. Обратно, рассмотрим второе слагаемое отдельно от первого. Вспомним, что каждое сечение процесса Винера – нормальная случайная величина. Значит, в отсутствие первого слагаемого случайная величина ξ определяется нормальной случайной величиной в момент времени t. Нормальная случайная величина – наиболее распространенная и естественная модель погрешности наблюдения. Таким образом, в данном дифференциальном уравнении есть некая полезная часть результата, которая дает информацию о физической природе процесса, а есть погрешность. В случае такой модели результат наблюдения будет динамически изменяться по закону (9.22). Тогда можно сказать, что $\alpha(t)$ – фактическая производная зависимости $\xi(t)$ как полезного сигнала от t.

Примеры.

Пример 1. Рассмотрим простейшее уравнение

$$d\xi = a(t)dt + b(t)dw(t) \tag{9.24}$$

где $a(\cdot)$, $b(\cdot)$ – неслучайные функции. Тогда

$$\xi(t) = \xi_0 + \int_0^t a(t)dt + \int_0^t b(t)dw(t) \tag{9.25}$$

Найдем второй интеграл. Рассмотрим интегральную сумму Ито

$$\mathcal{I}_n = \sum_{j=1}^{n} b\left(t_{j-1}\right)\Delta w_j \tag{9.26}$$

Таким образом, получаем, что интегральная сумма есть сумма независимых нормальных случайных величин с указанными коэффициентами. Учтем, что

$$\Delta w_j \in \mathbb{N}\left(0, t_j - t_{j-1}\right), \quad j = 1, \dots, n \tag{9.27}$$

независимы. Тогда имеем, что частичная сумма для второго интеграла имеет известное распределение

$$\mathcal{I}_n \in \mathbb{N}\left(0, B_n^2\right) \tag{9.28}$$

где дисперсия

$$B_n^2 = \sum_{j=1}^{n} b\left(t_{j-1}\right)\left(t_j - t_{j-1}\right) \underset{n}{\longrightarrow} \int_0^t b^2(s)ds = B^2 \tag{9.29}$$

Отсюда получаем, что

$$\mathcal{I}_n \to \mathcal{I} \in \mathbb{N}\left(0, B^2\right) \tag{9.30}$$

по распределению. Можно показать, что имеет место и среднеквадратичная сходимость. Итак, для указанного дифференциального уравнения процесс $\xi(t)$ представляет собой сумму (9.25), где ξ_0 – некоторая случайная величина, начальное значение случайного процесса. Второе слагаемое – неслучайная добавка. Это та часть наблюдения, которая связана с неслучайным дифференциалом $a(t)dt$ (то, что раньше мы назвали динамикой полезного сигнала). Третье слагаемое представляет собой шум, который искажает сигнал. Итак,

$$\int_0^t b(t)dw(t) \in \mathbb{N}\left(0, B^2\right) \tag{9.31}$$

Пример 2. Вернемся к общему стохастическому дифференциальному уравнению

$$\begin{cases} d\xi = \alpha(t)dt + \beta(t)dw(t) \\ \xi(0) = \xi_0 \end{cases}, \quad 0 \leqslant t < T \tag{9.32}$$

Справедлива следующая **формула Ито**.

Пусть случайный процесс есть некоторая функция действительных переменных

$$\eta(t) = u\left(\xi(t), t\right) \tag{9.33}$$

где $u(x,t),\ x \in \mathbb{R}, t \geqslant 0$ – дважды непрерывно дифференцируемая функция. Будем считать, что $\xi(t)$ удовлетворяет задаче (9.32). Дифференциал процесса (9.33) можно записать в виде

$$d\eta = \frac{\partial u}{\partial t}(\xi,t)dt + \frac{\partial u}{\partial x}(\xi,t)d\xi + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(\xi,t)\cdot\beta^2(t)dt \tag{9.34}$$

Этот дифференциал не совпадает с обычным дифференциалом для неслучайной функции. Откуда взялось дополнительное слагаемое? Отметим, что это дополнительное слагаемое связано с добавкой, которая определяет амплитуду шума.

Пусть для простоты в функции u нет явной зависимости от t, а в (9.32) нет первого слагаемого

$$u = u(x), \alpha = 0 \tag{9.35}$$

Пусть также

$$d\xi = b(t)dw(t) \tag{9.36}$$

где $b(t)$ – неслучайная функция. Рассмотрим приращение

$$\eta(t + dt) - \eta(t)$$

для

$$\eta(t) = u\left(\xi(t)\right) \tag{9.37}$$

Разложим функцию $u(\cdot)$ в ряд Тейлора:

$$du = u(x + dx) - u(x) \approx u'(x)dx + \frac{1}{2}u''(x)\left(dx\right)^2 \tag{9.38}$$

Применим полученное разложение к случайным процессам. Имеем

$$\eta(t) = u\left(\xi(t)\right) \tag{9.39}$$

где в роли x выступает

$$x = \xi(t) \tag{9.40}$$

При этом дифференциал процесса описывается формулой

$$dx = d\xi = b(t)dw(t) \tag{9.41}$$

Тогда получаем

$$d\eta(t) \approx u'(\xi)d\xi + \frac{1}{2}u''(\xi)\left(d\xi\right)^2 \tag{9.42}$$

Отметим, что в аналогичном случае в математическом анализе второе слагаемое, как и все остальные, отбрасывались, так как они стремятся к нулю быстрее, чем $d\xi$. Однако здесь другой случай. Отметим, что для приращения

$$(d\xi)^2 = b^2(t)\left[dw(t)\right]^2 \tag{9.43}$$

имеем

$$D\left[dw(t)\right]^2 = \left(M\left[dw(t)\right]^4 - [dt]^2\right) = \left(3\,[dt]^2 - [dt]^2\right) = 2\,[dt]^2 \tag{9.44}$$

поэтому

$$D\left[b(t)dw(t)\right]^2 \sim [dt]^2 \sim 0 \tag{9.45}$$

в первом порядке по dt. Посчитаем математическое ожидание

$$M\,(d\xi)^2 = b^2(t)dt \sim dt \tag{9.46}$$

то есть $(d\xi)^2$ имеет первый порядок малости по dt. Поэтому мы не можем выбросить второе слагаемое в (9.42). Отсюда в формуле Ито и появляется дополнительное слагаемое.

Рассмотрим способы использования формулы Ито в решении задач. Если случайный процесс есть функция

$$\eta(t) = g\,(w(t), t) \tag{9.47}$$

то есть

$$\xi = dw, \ \ d\xi = dw, \ \ \alpha = 0, \ \ \beta = 1 \tag{9.48}$$

Тогда из формулы Ито

$$d\eta = \left(\frac{\partial g}{\partial t}(w,t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(w,t)\right)dt + \frac{\partial g}{\partial x}(w,t)dw(t) \tag{9.49}$$

Пример. Решить уравнение

$$d\xi(t) = a\xi(t)dt + b\xi(t)dw(t) \tag{9.50}$$

где $a, b = const$. Отметим, что здесь шум пропорционален сигналу. Попробуем применить формулу Ито в данном случае.

Будем искать решение в виде

$$\xi(t) = g\,(w(t), t) \tag{9.51}$$

Тогда, по формуле Ито

$$\begin{gathered} d\xi = \left(\frac{\partial g}{\partial t}(w,t) + \frac{b^2}{2}\frac{\partial^2 g}{\partial x^2}(w,t)\right)dt + \frac{\partial g}{\partial x}(w,t)dw(t) = \\ = a\xi(t)dt + b\xi(t)dw(t) = ag\,(w(t),t)\,dt + bg\,(w(t),t)\,dw(t) \end{gathered} \tag{9.52}$$

Приравниваем множители при dt и dw, получаем для $x \in \mathbb{R}, \ t \geqslant 0$

$$\frac{\partial g}{\partial t}(x,t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(x,t) = ag(x,t) \tag{9.53}$$

$$\frac{\partial g}{\partial x}(x,t) = bg(x,t) \tag{9.54}$$

Из уравнения (9.54)

$$g(x,t) = C(t)e^{bx} \tag{9.55}$$

тогда, поставив это в (9.53), получим

$$\dot{C}(t) + \frac{b^2}{2}C = aC \tag{9.56}$$

Отсюда

$$C(t) = Ce^{at - \frac{b^2}{2}t} \tag{9.57}$$

Тогда получаем ответ

$$\xi(t) = g\left(w(t), t\right) = Ce^{at - \frac{b^2}{2}t} = \xi(0)e^{at - \frac{b^2}{2}t}e^{bw(t)} \tag{9.58}$$

Пример. Решить уравнение

$$(1+t)d\xi(t) = -\xi(t)dt + dw(t) \tag{9.59}$$

Попробуем решить упрощенную неслучайную задачу

$$(1+t)dx(t) = -xdt \tag{9.60}$$

получим

$$d\ln x = -d\ln(1+t) \tag{9.61}$$

откуда

$$x = \frac{C}{(1+t)} \tag{9.62}$$

Исходя из этого, будем искать решение в виде

$$\xi(t) = \frac{\eta(t)}{1+t} \tag{9.63}$$

Тогда начальное условие

$$\xi(0) = \eta(0) \tag{9.64}$$

Запишем дифференциал

$$d\xi(t) = \frac{d\eta(t)}{1+t} - \frac{\eta(t)}{(1+t)^2}dt \tag{9.65}$$

Отсюда получаем

$$d\eta(t) - \frac{\eta(t)}{1+t}dt = (1+t)d\xi(t) = -\frac{\eta(t)}{1+t}dt + dw(t) \tag{9.66}$$

Сравним левую и правую части равенства (9.67). Тогда получим

$$d\eta(t) = dw(t) \tag{9.67}$$

В результате получаем

$$\eta(t) = \eta(0) + w(t) \tag{9.68}$$

Значит, решение имеет вид

$$\xi(t) = \frac{\xi(0) - w(t)}{1 + t} \tag{9.69}$$

ФИЗИЧЕСКИЙ
ФАКУЛЬТЕТ
МГУ ИМЕНИ
М.В. ЛОМОНОСОВА

teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ