Случайные процессы

Лектор В. И. Богачев

Содержание

§ 1. Основные понятия

Случайным процессом на вероятностном пространстве $(\Omega, \mathcal{B}, P)$, где $\mathcal{B}$ — σ-алгебра подмножеств Ω, P — вероятностная мера на $\mathcal{B}$, с пространством состояний S и параметрическим множеством T, играющим роль времени, принято называть произвольное семейство случайных элементов ξ_t, где $t \in T$, со значениями в пространстве S. Мы будем в основном иметь дело с частным случаем, когда $S = \mathbb{R}^1$, т.е. $\{\xi_t\}_{t \in T}$ — просто некий набор случайных величин на Ω, индексируемый множеством T. При этом и само множество T у нас в основном будет иметь специальный вид: это будет подмножество $\mathbb{R}$, чаще всего множество неотрицательных целых чисел $\mathbb{Z}^+$, отрезок $[0, T]$ или луч $[0, +\infty)$.

Напомним, что случайная величина есть функция на Ω, для которой все множества вида $\{\omega\colon \xi(\omega) < c\}$, где $c \in \mathbb{R}$, входят в $\mathcal{B}$. Как известно, тогда в $\mathcal{B}$ входят и все множества вида $\{\omega\colon \xi(\omega) \in A\}$ для всякого борелевского множества A на прямой (борелевская σ-алгебра — наименьшая, содержащая все лучи или все открытые множества). Хотя явного описания борелевских множеств нет, таковыми обычно оказываются множества, возникающие в реальных задачах. Типичные борелевские множества: открытые, замкнутые, счетные пересечения открытых, счетные объединения замкнутых.

Польза от включения в $\mathcal{B}$ множеств вида $\{\omega\colon \xi(\omega) \in A\}$ состоит в том, что им можно приписать вероятности. В частности, возникает распределение случайной величины ξ, заданное формулой

$$P_\xi(A) := P(\{\omega\colon \xi(\omega) \in A\}).$$

Сама мера P задана на каком-то большом (или абстрактном) пространстве, но ее распределение — борелевская мера на прямой. Если даны несколько случайных величин $\xi_1, \ldots, \xi_n$ на Ω, то их совместным распределением называется борелевская мера на $\mathbb{R}^n$, заданная формулой

$$P_{\xi_1,\ldots,\xi_n}(A) := P\Big(\{\omega\colon (\xi_1(\omega), \ldots, \xi_n(\omega)) \in A\}\Big),$$

где теперь A — борелевское множество в $\mathbb{R}^n$. Конечно, предварительно надо проверить, что $\{\omega\colon (\xi_1(\omega), \ldots, \xi_n(\omega)) \in A\} \in \mathcal{B}$.

Такие распределения — частный случай общего понятия образа меры P при измеримом отображении f в измеримое пространство $(E, \mathcal{E})$, обозначаемого символом $P \circ f^{-1}$ и задаваемого на $\mathcal{E}$ формулой

$$P \circ f^{-1}(E) = P(f^{-1}(E)), \quad E \in \mathcal{E}.$$

Легко видеть, что $P\circ f^{-1}$ — вероятностная мера на $\mathcal{E}$. Для интеграла от интегрируемой по этой мере функции g имеет место формула замены переменных

$$\int_E g(y)\,P\circ f^{-1}(dy)=\int_\Omega g(f(\omega))\,P(d\omega).$$

Например,

$$\mathrm{E}\xi=\int_{-\infty}^{+\infty} tP_\xi(dt),\quad \mathrm{E}(\xi\eta)=\int_{\mathbb{R}^2} xy\,P_{\xi,\eta}(dx\,dy).$$

Независимость случайных величин $\xi_1,\ldots,\xi_n$ означает, что их совместное распределение $P_{\xi_1,\ldots,\xi_n}$ есть произведение $P_{\xi_1}\otimes\cdots\otimes P_{\xi_n}$ их распределений. Иначе говоря,

$$P(\xi_1\in A_1,\ldots,\xi_n\in A_n)=P(\xi_1\in A_1)\cdots P(\xi_n\in A_n).$$

Очень важной характеристикой случайного процесса являются его *конечномерные распределения* $P_{\xi_{t_1},\ldots,\xi_{t_n}}$ при всяческих $t_1,\ldots,t_n\in T$ (возможно, совпадающих), задаваемые той же формулой, т.е.

$$P_{\xi_{t_1},\ldots,\xi_{t_n}}(A):=P\Big(\{\omega\colon(\xi_{t_1}(\omega),\ldots,\xi_{t_n}(\omega))\in A\}\Big).$$

Многие случайные процессы, встречающиеся в приложениях, описываются через конечномерные распределения. Два важнейших примера: пуассоновский и винеровский процессы, рассматриваемые далее. Специальным образом задаются ветвящиеся и марковские процессы, в том числе цепи Маркова, о которых также пойдет речь. Наши основные примеры будут входить в широкие классы процессов: процессы с независимыми приращениями, мартингалы, стационарные процессы, для которых будет представлен ряд основных результатов. Явным образом случайные процессы с непрерывным временем задаются нечасто, но есть и такие примеры.

1.1. Пример. *Процесс восстановления* определяется так. Пусть дана последовательность неотрицательных независимых случайных величин ξ_i с общей функцией распределения $F(t)=P(\xi_i<t)$, не равных тождественно нулю. Положим $S_0=0$, $S_n=\xi_1+\cdots+\xi_n$. Термин «восстановление» возникает из трактовки ξ_i как длительности исправной работы i-го из последовательно работающих элементов. В начальный момент времени $t=0$ начинает работать первый элемент, в момент $\xi_1=S_1$ он заменяется вторым элементом, который в свою очередь заменяется следующим элементом в момент S_2 и т. д. Скажем, можно представлять себе постоянно горящую лампочку или батарейку часов. Моменты $S_1,S_2,\ldots$ называются моментами

восстановления. Число восстановлений ν_t в интервале времени $[0,t]$ равно максимальному n, для которого $S_n \leqslant t$. Это зависящее от времени случайное число ν_t и называют процессом восстановления. Ниже приведены некоторые вычисления с процессом восстановления, показывающие, что в случае экспоненциального распределения величин ξ_i процесс ν_t оказывается пуассоновским. На самом деле в этом примере рассмотрен простейший вид процессов восстановления.

Для общего процесса $\xi_t(\omega)$ возникают выборочные траектории

$$\xi_\bullet(\omega)\colon t \mapsto \xi_t(\omega)$$

В определении процесса никаких ограничений на них нет, но в реальных задачах бывает полезно знать, будут ли такие траектории ограниченными или непрерывными, будут ли у них какие-то асимптотические свойства при $t \to +\infty$ или $t \to 0$. Как мы увидим ниже, сами постановки таких вопросов требуют некоторого уточнения.

Если случайные величины ξ_t обладают средними, то появляется функция $M(t) = \mathrm{E}\xi_t$ — среднее процесса, т.е. математическое ожидание ξ_t. Если же и квадраты ξ_t интегрируемы, то можно задать важную характеристику процесса — *корреляционную функцию* (называемую также *ковариационной* или *ковариацией*):

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

В случае $t = s$ это дисперсия ξ_t, при нулевом среднем это скалярное произведение в $L^2(P)$ функций ξ_t и ξ_s (в общем случае это скалярное произведение центрированных функций $\xi_t - \mathrm{E}\xi_t$ и $\xi_s - \mathrm{E}\xi_s$). Нередко ковариационной функцией называют только слагаемое $\mathrm{E}(\xi_t\xi_s)$; понятно, что для величин с нулевым средним разницы нет. Здесь уместно напомнить неравенство Коши – Буняковского

$$|\mathrm{E}(\xi\eta)|^2 \leqslant |\mathrm{E}(\xi^2)\mathrm{E}(\eta^2)|.$$

§ 2. Распределения случайных процессов

Выше было сказано о конечномерных распределениях случайного процесса ξ_t, которые являются борелевскими вероятностными мерами на пространствах $\mathbb{R}^n$ и задаются формулами

$$P_{\xi_{t_1},\ldots,\xi_{t_n}}(A) := P\Big(\{\omega\colon (\xi_{t_1}(\omega),\ldots,\xi_{t_n}(\omega)) \in A\}\Big).$$

Здесь некоторые t_i могут совпадать, но конечномерные распределения с совпадающими моментами времени легко находятся по конечномерным распределениям с различными моментами. Например, если

$t_1 = t_2$, то мера $P_{\xi_{t_1},\xi_{t_1},\xi_{t_3},\dots,\xi_{t_n}}$ на $\mathbb{R}^{n+1}$ есть образ меры $P_{\xi_{t_1},\xi_{t_3},\dots,\xi_{t_n}}$ на $\mathbb{R}^n$ при отображении

$$(x_1, x_2, \dots, x_n) \mapsto (x_1, x_1, x_2, \dots, x_n)$$

из $\mathbb{R}^{n-1}$ в $\mathbb{R}^n$. Поэтому достаточно иметь дело с различными t_i. Более того, можно считать, что конечномерные распределения P_S заданы для всевозможных конечных подмножеств $S \subset T$. При такой интерпретации никаких кратных точек не возникает. Для $S = \{s_1, \dots, s_n\}$, где уже все s_i различны, мера P_S задается не на $\mathbb{R}^n$, а на пространстве $\mathbb{R}^S$ всех вещественных функций на S (конечно, оно изоморфно $\mathbb{R}^n$), наделенном σ-алгеброй множеств вида

$$C_A = \{x \in \mathbb{R}^S \colon (x(s_1), \dots, x(s_n)) \in A\},$$

где A — борелевское множество в $\mathbb{R}^n$, причем $P_S(C_A) = P_{\xi_{s_1},\dots,\xi_{s_n}}(A)$. Это определение не зависит от способа нумерации точек множества S, так как если взять перестановку σ точек $1, \dots, n$ и соответствующим образом переставить точки s_i, то в новом представлении множество C_A будет соответствовать множеству $C_{\sigma(A)}$, где $\sigma(A)$ — образ A при отображении $(x_1, \dots, x_n) \mapsto (x_{\sigma(1)}, \dots, x_{\sigma(n)})$. При этом

$$P_{\xi_{s_1},\dots,\xi_{s_n}}(A) = P_{\xi_{s_{\sigma(1)}},\dots,\xi_{s_{\sigma(n)}}}(\sigma(A)).$$

Это равенство достаточно проверять лишь на произведениях измеримых множеств $A = A_1 \times \cdots A_n$, для которых оно очевидно:

$$P_{\xi_{s_1},\dots,\xi_{s_n}}(A_1 \times \cdots \times A_n) = P_{\xi_{s_{\sigma(1)}},\dots,\xi_{s_{\sigma(n)}}}(A_{\sigma(1)} \times \cdots \times A_{\sigma(n)}).$$

Оказывается, все эти конечномерные распределения можно получать из бесконечномерного распределения процесса, которое определяется следующим образом.

Обозначим через $\mathbb{R}^T$ множество всех вещественных функций на множестве T (для процесса с пространством состояний S надо брать множество S^T, но здесь это не обсуждается). Если T бесконечно, то получается бесконечномерное линейное пространство. В пространстве $\mathbb{R}^T$ есть элементарные цилиндры вида

$$C_{t,A} = \{x \in \mathbb{R}^T \colon x(t) \in A\},$$

где A — борелевское множество на прямой. Например, если $T = \{1, 2\}$, $t = 1$, $A = [a, b]$, то C_t — цилиндр (полоса) с основанием $[a, b]$ на первой оси. Элементарные цилиндры порождают алгебру цилиндров вида

$$C_{t_1,\dots,t_n,A} = \{x \in \mathbb{R}^T \colon (x(t_1), \dots, x(t_n)) \in A\},$$

где A — борелевское множество теперь уже в $\mathbb{R}^n$. Чтобы увидеть, что в самом деле это алгебра, достаточно заметить, что два цилиндра с множествами точек $t_1, \ldots, t_n$ и $s_1, \ldots, s_k$ и основаниями $A \subset \mathbb{R}^n$ и $B \subset \mathbb{R}^k$ могут быть записаны с помощью объединенного множества точек $(t_1, \ldots, t_n, s_1, \ldots, s_k)$ и оснований в общем пространстве $\mathbb{R}^{n+k}$, полученных умножением исходных оснований на $\mathbb{R}^k$ и $\mathbb{R}^n$ соответственно.

Представление цилиндра $C_{t_1,\ldots,t_n,A}$ не однозначно, всегда можно добавить лишнюю точку t_{n+1} и заменить A на $A \times \mathbb{R}^1$, т.е. реальной зависимости от значений в t_{n+1} не будет, но можно выбрать наименьшее возможное число точек t_j в представлении. Для цилиндра, отличного от всего $\mathbb{R}^T$, точки в таком минимальном наборе определены однозначно, дальнейшая неединственность представления связана только с их возможной перенумерацией.

Цилиндры еще не составляют σ-алгебру. Символами $\mathcal{B}_T$ или $\sigma(\mathbb{R}^T)$ обозначают наименьшую σ-алгебру, содержащую все цилиндры (это то же самое, что наименьшая σ-алгебра, содержащая все элементарные цилиндры).

В терминах σ-алгебр, порожденных функциями, можно сказать, что $\mathcal{B}_T$ есть σ-алгебра в пространстве траекторий $\mathbb{R}^T$, порожденная всеми функциями вида $x \mapsto x(t)$, $t \in T$.

Простейший типичный случай бесконечного T есть множество $\mathbb{N}$, тогда получаем счетное произведение прямых $\mathbb{R}^\infty$, т.е. множество всех бесконечных вещественных последовательностей $x = (x_1, x_2, \ldots)$. Цилиндры в нем задаются несколькими координатами, а порожденная ими σ-алгебра $\mathcal{B}(\mathbb{R}^\infty)$ оказывается борелевской σ-алгеброй пространства $\mathbb{R}^\infty$ с его топологией произведения. Однако для несчетных T такого совпадения нет.

Множества из $\mathcal{B}_T$ не имеют какого-либо простого явного описания (даже если согласиться, что цилиндры имели явное описание), но все же некоторое описание легко указать, если использовать σ-алгебру $\mathcal{B}(\mathbb{R}^\infty)$ в пространстве последовательностей. Тогда получим описание $\mathcal{B}_T$ как всех множеств вида

$$\{x \in \mathbb{R}^T \colon (x(t_1), x(t_2), \ldots) \in E\} \tag{2.1}$$

где $\{t_j\} \subset T$, $E \in \mathcal{B}(\mathbb{R}^\infty)$. В самом деле, класс множеств такого вида замкнут относительно дополнений и счетных объединений и содержит элементарные цилиндры. Поэтому он содержит порожденную ими σ-алгебру. С другой стороны, все такие множества входят в $\mathcal{B}_T$. Действительно, это верно для $E \subset \mathbb{R}^\infty$ вида $C_{n,(a,b)} = \{x \in \mathbb{R}^\infty \colon x_n \in (a, b)\}$

причем класс множеств E, для которых это верно, есть σ-алгебра, откуда следует, что это верно для множеств из наименьшей σ-алгебры, содержащей все $C_{n,(a,b)}$, а это и есть $\mathcal{B}(\mathbb{R}^\infty)$.

Из указанного описания видно, что всякое отдельное множество из $\mathcal{B}_T$ пришло из σ-алгебры, порожденной неким счетным набором цилиндров с моментами времени t_n из некоего счетного множества $\{t_n\} \subset T$. Конечно, это ясно и без всякого описания. В самом деле, всякий элементарный цилиндр $C_{t,A}$ входит в σ-алгебру, порожденную цилиндрами $C_{t,(-\infty,r)}$, где r рационально. Далее, если для каждого счетного набора $\{t_j\}$ взять σ-алгебру $S_{\{t_j\}}$, порожденную цилиндрами $C_{t_j,A}$, то объединение всех этих σ-алгебр тоже оказывается σ-алгеброй. Обычно это не так, но в данном случае эффект в том, что для всякого счетного семейства множеств E_n, пришедших из своих σ-алгебр $S_{\{t_j^n\}}$, все эти множества входят в σ-алгебру, соответствующую счетному множеству, равному объединению всех t_j^n.

2.1. Определение. *Распределением случайного процесса $\{\xi_t\}_{t\in T}$ называется мера P_ξ на пространстве траекторий $\mathbb{R}^T$, заданная на σ-алгебре $\mathcal{B}_T$ формулой*

$$P_\xi(B) = P(\omega\colon \textit{функция } t \mapsto \xi_t(\omega) \textit{ входит в } B), \quad B \in \mathcal{B}_T. \tag{2.2}$$

Требует пояснения включение в область определения $\mathcal{B}$ меры P множества таких ω, что траектория $t \mapsto \xi_t(\omega)$ входит в B. Если B — цилиндрическое множество $C_{t_1,\ldots,t_n,A}$, то это очевидно. Тогда $P_\xi(B) = P_{t_1,\ldots,t_n}(A)$. Вся совокупность множеств B из $\mathcal{B}_T$, для которых включение верно, есть σ-алгебра, так как $\mathcal{B}$ — σ-алгебра. Из этого следует, что указанная совокупность совпадает с $\mathcal{B}_T$. Можно было бы также воспользоваться указанным выше представлением (2.1) множества B с помощью множества $E \in \mathcal{B}(\mathbb{R}^\infty)$ и счетного набора $\{t_j\}$. Тогда

$$P_\xi(B) = P\Big(\omega\colon \big(\xi_{t_1}(\omega), \xi_{t_2}(\omega), \ldots\big) \in E\Big),$$

где множество в правой части входит в $\mathcal{B}$ (это тоже следует из $\mathcal{B}$-измеримости функций ξ_{t_j} стандартным образом: класс множеств E, для которых это верно, есть σ-алгебра, причем она содержит множества вида $\{x \in \mathbb{R}^\infty\colon x_j < c\}$ при всех j и c).

Ясно, что P_ξ — вероятностная мера (если B составлено из дизъюнктных множеств B_n, то соответствующее множество точек ω есть дизъюнктное объединение множеств с траекториями из B_n).

Имеется принципиальная разница между случаями счетного и несчетного T. В первом случае для практических целей можно считать, что все реальные множества попали в $\mathcal{B}_T$, но во втором случае очень многие нужные множества траекторий в $\mathcal{B}_T$ не попадают. Например, множество из одной тождественно нулевой траектории не входит в $\mathcal{B}_T$, как и множество всех неотрицательных траекторий (эта проблема обсуждается ниже).

2.2. Определение. *Модификацией или версией случайного процесса* $\{\xi_t\}_{t\in T}$ *называется случайный процесс* $\{\eta_t\}_{t\in T}$ *с теми же* Ω *и* T*, для которого при каждом* t *верно равенство* $\xi_t(\omega) = \eta_t(\omega)$ *почти всюду* (*где множество меры нуль может зависеть от* t). *При этом эти процессы называются стохастически эквивалентными.*

Ясно, что стохастически эквивалентные процессы имеют одинаковые конечномерные распределения и равные распределения в пространстве траекторий. Обратное, конечно, неверно, хотя бы из-за того, что процессы с одинаковыми распределениями могут быть заданы на разных вероятностных пространствах.

Если никакого Ω нет, а есть только множество T и вероятностная мера P_0 на σ-алгебре $\mathcal{B}_T$, то появляется и случайный процесс: в качестве $(\Omega, \mathcal{B}, P)$ берем тройку $(\mathbb{R}^T, \mathcal{B}_T, P_0)$ и полагаем

$$\xi_t(\omega) := \omega(t), \quad \omega \in \mathbb{R}^T.$$

Конечно, так можно поступить и тогда, когда процесс все же был, а в качестве P_0 взяли его распределение в $\mathbb{R}^T$. Заданный выше процесс имеет такие же конечномерные распределения, как и исходный процесс, а также такое же распределение в $\mathbb{R}^T$.

Отметим, что стохастически эквивалентные процессы могут иметь очень разные выборочные траектории. Рассмотрим такой пример. В качестве вероятностного пространства возьмем отрезок $[0,1]$ с борелевской σ-алгеброй и стандартной мерой Лебега. Пусть $T = [0,1]$. Положим $\xi_t(\omega) = 0$, $\eta_t(\omega) = 0$ при $\omega \neq t$, $\eta_t(t) = 1$. Для каждого t имеет $\xi_t(\omega) = \eta_t(\omega)$ почти всюду (а именно при всех $\omega \neq t$), но процесс ξ имеет тождественно нулевые траектории, а каждая траектория $t \mapsto \eta_t(\omega)$ имеет скачок в точке $t = \omega$.

При задании процесса с помощью конечномерных распределений возникает, конечно, вопрос о существовании нужного процесса, а также более тонкий вопрос о существовании модификации процесса с траекториями с нужными свойствами (скажем, непрерывными или возрастающими).

§ 3. Теоремы Колмогорова

Ответ на вопрос о существовании процесса с заданными конечномерными распределениями дается теоремой Колмогорова. Для ее формулировки заметим, что конечномерные распределения процесса ξ обладают следующими свойствами согласованности:

1) проекция меры $P_{t_1,\ldots,t_n}$ на $\mathbb{R}^{n-1}$ есть мера $P_{t_1,\ldots,t_{n-1}}$, т.е.

$$P_{t_1,\ldots,t_{n-1}}(A) = P_{t_1,\ldots,t_n}(A \times \mathbb{R}^1)$$

для всех борелевских множеств $A \subset \mathbb{R}^{n-1}$;

2) для всякой перестановки σ номеров $1, \ldots, n$ и всяких борелевских множеств $A_1, \ldots, A_n$ на прямой верно равенство

$$P_{t_1,\ldots,t_n}(A_1 \times \cdots \times A_n) = P_{t_{\sigma(1)},\ldots,t_{\sigma(n)}}(A_{\sigma(1)} \times \cdots \times A_{\sigma(n)}).$$

Заметим, что если конечномерные распределения задавать не для упорядоченных наборов точек, а для конечных подмножеств $S \subset T$, как было указано выше, то остается лишь первое условие согласованности, означающее, что проекция меры P_{S_1} на $\mathbb{R}^{S_2}$ для $S_2 \subset S_1$, возникающая при проектировании $\mathbb{R}^{S_1}$ на $\mathbb{R}^{S_2}$ (сужении функций с S_1 на S_2) совпадает с мерой P_{S_2}.

Оба свойства очевидны, про второе уже говорилось ранее. Оказывается, эти очевидные условия и достаточны для существования процесса с данными распределениями.

3.1. Теорема. (Теорема Колмогорова о согласованных распределениях) *Набор борелевских вероятностных мер $P_{t_1,\ldots,t_n}$ на пространствах $\mathbb{R}^n$, где $t_i \in T$ и T — заданное множество, служит совокупностью конечномерных распределений некоторого случайного процесса с параметрическим множеством T в точности тогда, когда выполнены условия согласованности.*

Доказательство. Обоснование будет разбито на несколько этапов, часть которых отнесена к задачам. Основная идея Колмогорова — построить нужный процесс на пространстве траекторий $\mathbb{R}^T$ с σ-алгеброй $\mathfrak{B}_T$. Но нужна мера P. Пока такая мера есть на алгебре цилиндров: можно взять

$$P(C_{t_1,\ldots,t_n,A}) = P_{t_1,\ldots,t_n}(A),$$

что дает аддитивную функцию множества. При этом важно, что она корректно определена: ведь запись цилиндра не единственна, так как

$$C_{t_1,\ldots,t_n,A} = C_{t_1,\ldots,t_n,t_{n+1},A\times\mathbb{R}^1},$$

кроме того, цилиндры можно представлять иным способом путем перестановок точек t_i, что уже обсуждалось выше. Условия Колмогорова как раз ответственны за то, чтобы мера цилиндра не зависела от его представления. Теперь для построения меры на $\mathcal{B}_T$ достаточно знать счетную аддитивность P на алгебре цилиндров. Тогда продолжение на порожденную σ-алгебру дается внешней мерой P^*. На самом деле есть еще несколько более общий технический результат: достаточно иметь счетную аддитивность на полуалгебре $\mathcal{A}$ цилиндров вида $C_{t_1,A_1} \cap \cdots \cap C_{t_n,A_n}$. Алгебра получается из всевозможных конечных объединений таких пересечений. Это еще не вся алгебра цилиндров, но порожденная ей σ-алгебра тоже равна $\mathcal{B}_T$.

Откуда берется счетная аддитивность из аддитивности на полуалгебрах? Очень полезное достаточное (хотя и не необходимое) условие таково. Класс $\mathcal{K}$ множеств из алгебры или полуалгебры $\boldsymbol{\mathcal{A}}$ называется компактным, если из того, что $K_n \in \mathcal{K}$ и $\cap_{n=1}^{\infty} K_n = \varnothing$, следует, что найдется N, для которого $\cap_{n=1}^{N} K_n = \varnothing$.

Задача 1. Доказать, что любой набор компактов в $\mathbb{R}^n$ — компактный класс.

Задача 2. Доказать, что совокупность всех цилиндров в $\mathbb{R}^T$ вида $C_{t_1,\ldots,t_n,K}$, где K — компакт в $\mathbb{R}^n$, есть компактный класс.

Задача 3. Пусть P — аддитивная неотрицательная функция на алгебре или полуалгебре $\boldsymbol{\mathcal{A}}$, причем имеется такой компактный класс $\mathcal{K} \subset \boldsymbol{\mathcal{A}}$, для всякого $A \in \boldsymbol{\mathcal{A}}$ и всякого $\varepsilon > 0$ найдется такое множество $K \in \mathcal{K}$, что $K \subset A$ и $P(A \backslash K) < \varepsilon$. Доказать, что функция P счетно-аддитивна.

Задачи 2 и 3 дают счетную аддитивность нашей меры P, что обеспечивает существование счетно-аддитивного продолжения на $\mathcal{B}_T$. $\square$

Теоремы Колмогорова дает существование произведения $\bigotimes_t \mu_t$ произвольного набора вероятностных мер μ_t на прямой: в качестве $P_{t_1,\ldots,t_n}$ берем конечные произведения $\mu_{t_1} \otimes \cdots \otimes \mu_{t_n}$. Произведение вероятностных мер, причем на произвольных пространствах, можно определить без теоремы Колмогорова, но это отнюдь не банально.

Итак, теорема Колмогорова говорит, что в некотором смысле случайные процессы суть вероятностные меры на пространствах траекторий. Однако не все так просто. Реальный процесс может иметь непрерывные или возрастающие траектории или быть неотрицательным, т.е. иметь траектории из неотрицательных функций. К сожалению, все эти множества (непрерывные функции, возрастающие функции, неотрицательные функции) не входят в σ-алгебру $\mathcal{B}_T$ для несчетных множеств T. Дело в том, что всякое множество $B \in \mathcal{B}_T$ определяется по значениям функций на некотором счетном множестве: существует такое множество $\{t_j\} \subset T$, что если функция x входит в B и функция y совпадает с x во всех точках t_j, то y тоже входит в B. Это следует из сказанного выше про $\mathcal{B}_T$. Ни возрастание, ни неотрицательность, ни непрерывность нельзя определить по ограничению на счетное множество. Правда, есть еще процедура лебеговского пополнения меры P, когда к множествам из исходной σ-алгебры добавляются всевозможные подмножества множеств меры нуль. Это сильно увеличивает σ-алгебру, но не помогает с нашими множествами. Рассмотрим простейший пример. Пусть процесс с множеством $T = [0,1]$ тождественно равен нулю. У него единственная нулевая траектория, для ее описания достаточно меры Дирака, сосредоточенной в одной точке $\mathbb{R}^T$, состоящей из нулевой функции x_0. Эта мера вообще определена сразу на всех множествах формулой $\delta_{x_0}(B) = 1$, если $x_0 \in B$, $\delta_{x_0}(B) = 0$, если $x_0 \notin B$. Однако с точки зрения канонической σ-алгебры $\mathcal{B}_T$ эта точка нехороша: она не входит в пополнение $\mathcal{B}_T$. В самом деле, если бы она входила в пополнение, то представляла бы собой множество из $\mathcal{B}_T$ с добавленным подмножеством меры нуль. Сама точка x_0 не есть элемент $\mathcal{B}_T$, значит, она должны быть подмножеством множества меры нуль, что невозможно, ибо всякое содержащее ее множество из $\mathcal{B}_T$ имеет меру 1. Таким образом, существование какого-то продолжения меры на множество еще не означает, что оно измеримо относительно пополнения меры. Совершенно аналогична ситуации с множеством непрерывных траекторий. Непрерывность всех траекторий процесса ξ не означает, что множество непрерывных функций измеримо относительно пополнения P_ξ. Однако имеется трюк, позволяющий обходить эту трудность.

3.2. Лемма. *Пусть μ — вероятностная мера на σ-алгебре $\mathcal{A}$ пространства X и X_0 — такое множество, что $\mu^*(X_0) = 1$, т.е. если $X_0 \subset A \in \mathcal{A}$, то $\mu(A) = 1$, иначе говоря, в дополнении X_0 нет множеств из $\mathcal{A}$ положительной меры. Тогда можно сузить μ на X_0 в следующем смысле: на σ-алгебре $\mathcal{A}_0$ подмножеств в X_0 вида $A \cap X_0$, где $A \in \mathcal{A}$, корректно определена вероятностная мера*

$$\mu_0(A \cap X_0) := \mu(A).$$

Доказательство. Корректность определения означает, что если $A \cap X_0 = B \cap X_0$, где $A, B \in \mathcal{A}$, то $\mu(A) = \mu(B)$. Это в самом деле так, ибо $A \triangle B$ входит в $X \backslash X_0$ и потому имеет меру нуль. При этом мера μ_0 счетно-аддитивна, так если множества $A_n \cap X_0$, где $A_n \in \mathcal{A}$, попарно не пересекаются, то все множества $A_n \triangle A_k$ лежат в $X \backslash X_0$ и потому имеют меру нуль. Значит, имеет нулевую меру и их объединение D.

Тогда множества $A_n \backslash D$ попарно не пересекаются и $\mu(A_n) = \mu(A_n \backslash D)$. Значит,

$$\mu_0\Big(X_0 \cap \bigcup_{n=1}^{\infty} A_n\Big) = \mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \mu\Big(\bigcup_{n=1}^{\infty} (A_n \backslash D)\Big) = \\ = \sum_{n=1}^{\infty} \mu(A_n \backslash D) = \sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} \mu_0(A_n \cap X_0).$$

Наконец, $\mu_0 \geqslant 0$ и $\mu_0(X_0) = \mu_0(X \cap X_0) = \mu(X) = 1$. $\square$

Например, пусть μ есть распределение P_ξ процесса $\{\xi_t\}_{t \in T}$, причем для всякого t имеем $\xi_t(\omega) \geqslant 0$ при почти всех ω. Можно ли получить процесс с такими же конечномерными распределениями, но неотрицательными траекториями при всех ω? Как отмечалось выше, множество X_0 всех неотрицательных функций на T может оказаться неизмеримым относительно меры P_ξ (и даже ее пополнения), поэтому обычным способом нельзя сузить меру P_ξ на X_0 и из нее изготовить процесс. Однако $P_\xi^*(X_0) = 1$. В самом деле, пусть множество $B \in \mathfrak{B}_T$ содержит X_0. Как мы знаем, множество B определяется по значениям функций на некотором счетном множестве $\{t_j\}$. Для каждого t_j по условию имеем $P_\xi(x\colon x(t_j) < 0) = P(\omega\colon \xi_{t_j}(\omega) < 0) = 0$. Значит, множество $B_0 = \{x\colon x(t_j) \geqslant 0 \ \forall\, j\}$ имеет меру 1. При этом $B_0 \subset B$, так как иначе нашлась бы функция x, неотрицательная на $\{t_j\}$, но не входящая в B. Тогда не входит в B и равная этой функции на $\{t_j\}$ всюду неотрицательная функция x_1, которая вне $\{t_j\}$ определена нулем (здесь как раз важно то, что принадлежность к B определяется по значениям на $\{t_j\}$). Это невозможно, так как $y \in X_0 \subset B$. Доказанная лемма позволяет задать новый процесс η_t с новым вероятностным пространством $\Omega_0 = X_0$ и мерой P_0, полученной описанным сужением P_ξ, для которого уже все траектории неотрицательны.

Колмогоровым было найдено очень удобное достаточное условие существования непрерывной версии процесса, т.е. такой модификации, что почти все траектории непрерывны.

3.3. Теорема. (Теорема Колмогорова о непрерывной модификации процесса) *Пусть $T = [a, b]$ и процесс ξ_t удовлетворяет следующему условию: существуют такие числа $C, \beta, \varepsilon > 0$, что*

$$\mathrm{E}|\xi_t - \xi_s|^\beta \leqslant C|t - s|^{1+\varepsilon} \quad \forall\, t, s \in [a, b].$$

Тогда существует стохастически эквивалентный процесс, траектории которого непрерывны.

Доказательство. Можно считать, что $C = 1$ и $[a,b] = [0,1]$. Положим $q = 2^{-\varepsilon/(2\beta)}$, $\lambda = 2^{-\varepsilon/2}$. Для каждого n разделим $[0,1]$ на 2^n промежутков длины 2^{-n} и рассмотрим непрерывные процессы $\xi_n(t,\omega)$, графики которых на отрезке — ломаные с вершинами в точках $k2{-}n$, причем $\xi_n(k2^{-n},\omega) = \xi_{k2^{-n}}(\omega)$. Покажем, что для почти всех ω непрерывные функции $t \mapsto \xi_n(t,\omega)$ равномерно сходятся на отрезке. Заметим, что

$$\sup_{t\in[0,1]} |\xi_n(t,\omega) - \xi_{n-1}(t,\omega)| = \max_{k\leqslant 2^n} |\xi_{(k+1)2^{-n}}(\omega) - \xi_{k2^{-n}}(\omega)|.$$

Из неравенства Чебышёва имеем

$$P\Bigl(\omega\colon |\xi_{(k+1)2^{-n}}(\omega) - \xi_{k2^{-n}}(\omega)| \geqslant q^n\Bigr) \leqslant 2^{-n-n\varepsilon} q^{-n\beta}.$$

Поэтому

$$P\Bigl(\omega\colon \sup_{t\in[0,1]} |\xi_n(t,\omega) - \xi_{n-1}(t,\omega)| \geqslant q^n\Bigr) \leqslant 2^n 2^{-n-n\varepsilon} q^{-n\beta} = \lambda^n.$$

Так как $\lambda < 1$, ряд из левых частей сходится. По лемме Бореля–Кантелли имеет меру нуль множество точек ω, попадающих в бесконечно много множеств $\Bigl\{\omega\colon \sup_{t\in[0,1]} |\xi_n(t,\omega) - \xi_{n-1}(t,\omega)| \geqslant q^n\Bigr\}$. Иначе говоря, для почти каждого ω найдется такой номер $N(\omega)$, что $\sup_{t\in[0,1]} |\xi_n(t,\omega) - \xi_{n-1}(t,\omega)| < q^n$ при всех $n \geqslant N(\omega)$. Для таких ω последовательность $\xi_n(t,\omega)$ сходится равномерно на $[0,1]$, ибо

$$\begin{aligned}
&|\xi_{n+m}(t,\omega) - \xi_n(t,\omega)| \\
&\quad\leqslant |\xi_{n+m}(t,\omega) - \xi_{n+m-1}(t,\omega)| + \dots + |\xi_{n+1}(t,\omega) - \xi_n(t,\omega)| \\
&\qquad\qquad < q^{n+m} + \dots + q^n \leqslant \frac{q^n}{1-q}.
\end{aligned}$$

Для точек ω с таким свойством положим $\eta_t(\omega) = \lim\limits_{n\to\infty} \xi_n(t,\omega)$, для оставшихся точек из множества меры нуль положим $\eta_t(\omega) = 0$. Полученный процесс имеет непрерывные траектории. Проверим, что он является версией исходного. В точках t вида $t = k2^{-n}$ имеем $\eta_t(\omega) = \xi_t(\omega)$ при почти всех ω. Для всякой иной точки t есть сходящаяся к ней последовательность точек t_j рассмотренного вида. Для фиксированного $\delta > 0$ имеем

$$\begin{aligned}
&P\Big(\omega\colon |\xi_t(\omega) - \eta_t(\omega)| \geqslant \delta\Big)\\
&\qquad = P\Big(\omega\colon |\xi_t(\omega) - \xi_{t_j}(\omega) + \xi_{t_j}(\omega) - \eta_t(\omega)| \geqslant \delta\Big)\\
&\qquad = P\Big(\omega\colon |\xi_t(\omega) - \xi_{t_j}(\omega) + \eta_{t_j}(\omega) - \eta_t(\omega)| \geqslant \delta\Big)\\
&\leqslant P\Big(\omega\colon |\xi_t(\omega) - \xi_{t_j}(\omega)| \geqslant \delta/2\Big) + P\Big(\omega\colon |\eta_{t_j}(\omega) - \eta_t(\omega)| \geqslant \delta/2\Big)\\
&\qquad\qquad \leqslant \frac{2}{\delta}\mathbb{E}|\xi_t - \xi_{t_j}| + P\Big(\omega\colon |\eta_{t_j}(\omega) - \eta_t(\omega)| \geqslant \delta/2\Big).
\end{aligned}$$

При $t_j \to t$ оба слагаемых в правой части стремятся к нулю: первое в силу условия, второе в силу непрерывности $t \mapsto \eta_t(\omega)$. Поэтому левая часть равна нулю. Итак, получаем $\eta_t(\omega) = \xi_t(\omega)$ почти всюду. $\square$

В заключение разговора о распределениях процессов отметим, что в случае (основном для приложений) параметрического множества T, являющегося подмножеством $\mathbb{R}$, при рассмотрении конечномерных распределений $P_{\xi_{t_1},\dots,\xi_{t_n}}$ бывает удобно учитывать упорядоченность прямой и иметь дело только с наборами вида $t_1 < t_2 < \dots < t_n$. Конечно, условие симметрии 2) позволяет это сделать. Однако при этом не следует упускать из вида, что условие 1) с проекциями относится ко всем поднаборам уже упорядоченного по времени набора $t_1, \dots, t_n$, а не только укороченным $t_1, \dots, t_{n-1}$: надо проверять условие с проекциями для всех $t_{i_1}, \dots, t_{i_k}$. Скажем, в случае трех моментов $t_1 < t_2 < t_3$ надо проверять не только t_1, t_2 и t_1, но и t_2, t_3 и т.д. На самом деле достаточно проверять согласование для проекций на подпространства размерности на единицу меньше, так как их композиции дадут все проектирования.

Задача 4. Пусть $T = [a, b]$ и процесс ξ_t таков, что $P(\xi_t \geqslant \xi_s) = 1$ для всех $t \geqslant s$ из T. Доказать, что имеется стохастически эквивалентный процесс, почти все траектории которого возрастающие.

§ 4. Процессы с независимыми приращениями

4.1. Определение. *Процесс $\{\xi_t\}_{t\in T}$, где $T \subset \mathbb{R}$, называется процессом с независимыми приращениями, если для всех точек $t_0 \leqslant t_1 \leqslant \dots \leqslant t_n$ из T случайные величины $\xi_{t_1} - \xi_{t_0}, \dots, \xi_{t_n} - \xi_{t_{n-1}}$ независимы.*

Если $\xi_{t_0} = 0$, то конечномерное распределение $P_{\xi_{t_1},\dots,\xi_{t_n}}$ такого процесса получается линейным преобразованием из распределения случайного вектора $(\xi_{t_1},\dots,\xi_{t_n}-\xi_{t_{n-1}})$, представляющего собой произведение распределений приращений в силу их независимости. Это линейное преобразование имеет вид $y_1 = x_1$, $y_2 = x_1 + x_2, \dots$, $y_n = x_1 + \cdots + x_n$.

Из теоремы Колмогорова можно получить критерий существования процесса с независимыми приращениями с заданными начальным распределением и распределениями приращений. Этот критерий удобнее сформулировать через характеристические функции. Напомним, что характеристической функцией случайной величины ξ называется комплексная функция $\varphi_\xi(y) = \mathrm{E}\exp(iy\xi)$ на прямой. Характеристической функцией случайного вектора $\xi = (\xi_1,\dots,\xi_n)$ называется комплексная функция $\varphi_\xi(y) = \mathrm{E}\exp[i(y_1\xi_1 + \cdots + y_n\xi_n)]$ аргумента $y = (y_1,\dots,y_n)$ из $\mathbb{R}^n$. Характеристическая функция вероятностной меры μ на $\mathbb{R}^n$ задается формулой

$$\varphi_\mu(y) = \int_{\mathbb{R}^n} \exp[i(y,x)]\,\mu(dx).$$

Равенство двух мер на $\mathbb{R}^n$ равносильно совпадению их характеристических функций. Если мера μ на $\mathbb{R}^n$ проектируется на подпространство $x_m = 0$, то характеристическая функция проекции есть просто сужение φ_μ на это подпространство. Поэтому критерий Колмогорова существования процесса на $T \subset \mathbb{R}$ с заданными конечномерными распределениями $P_{t_1,\dots,t_n}$ с упорядоченными наборами $t_1 < \cdots < t_n$ можно сформулировать так.

4.2. Следствие. *Пусть для каждого набора $t_1 < \cdots < t_n$ в T дана борелевская вероятностная мера $P_{t_1,\dots,t_n}$ на $\mathbb{R}^n$ с характеристической функцией $\varphi_{t_1,\dots,t_n}$. Существование процесса с конечномерными распределениями $P_{t_1,\dots,t_n}$ при $t_1 < \cdots < t_n$ из T равносильно тому, что при всех $m \leqslant n$ верно равенство*

$$\varphi_{t_1,\dots,t_n}|_{y_m=0} = \varphi_{t_1,\dots,t_{m-1},t_{m+1},\dots,t_n}(y_1,\dots,y_{m-1},y_{m+1},\dots,y_n).$$

Для процессов с независимыми приращениями получаем такой критерий.

4.3. Теорема. *Пусть на прямой заданы борелевская вероятностная мера Q_0 и для каждой пары s,t с $0 \leqslant s < t$ борелевская вероятностная мера $Q_{s,t}$. Существование процесса с независимыми приращениями ξ_t на $[0,+\infty)$, для которого ξ_0 имеет распределение Q_0 и* *разности $\xi_t - \xi_s$ при $t > s$ имеют распределения $Q_{s,t}$ с характеристическими функциями $\varphi_{s,t}$, равносильно тождеству*

$$\varphi_{s,t} = \varphi_{s,u}\varphi_{u,t} \quad \textit{при} \ \ 0 \leqslant s < u < t.$$

Доказательство. Если такой процесс существует, то выполнено равенство $\xi_t - \xi_s = (\xi_t - \xi_u) + (\xi_u - \xi_t)$, где слагаемые независимы, поэтому характеристическая функция суммы есть произведение характеристических функций слагаемых.

Пусть тождество верно. Проверим сначала, что процесс есть для начального распределения, сосредоточенного в нуле, т.е. Q_0 — мера Дирака в нуле, ищем процесс с $\xi_0 = 0$. Распределение $P_{t_1,\ldots,t_n}$ при $0 < t_1 < \cdots < t_n$ зададим как образ произведения

$$Q_{0,t_1} \otimes Q_{t_1,t_2} \otimes \cdots \otimes Q_{t_{n-1},t_n}$$

при указанном выше линейном отображении

$$y_1 = x_1, y_2 = x_1 + x_2, \ldots, y_n = x_1 + \cdots + x_n.$$

Если же $t_1 = 0$, то в качестве $P_{0,t_2,\ldots,t_n}$ возьмем меру $P_{t_2,\ldots,t_n}$, считая ее мерой на $\mathbb{R}^n$, сосредоточенной на плоскости с нулевой первой координатой.

Заметим, что при линейном отображении L мера μ с характеристической функцией φ_μ переходит в меру с характеристической функцией $\varphi_\mu(L^*y)$, где L^* — сопряженный оператор. В самом деле, характеристическая функция меры-образа равна

$$\int_{\mathbb{R}^n} \exp(i(y, Lx))\,\mu(dx) = \int_{\mathbb{R}^n} \exp(i(L^*y, x))\,\mu(dx).$$

Произведение мер Q_{t_{i-1},t_i} имеет в качестве характеристической функции произведение $\prod_{i=1}^n \varphi_{Q_{t_{i-1},t_{i-1}}}(y_i)$. Поэтому характеристическая функция меры $P_{t_1,\ldots,t_n}$ есть

$$\varphi_{t_{n-1},t_n}(y_n)\varphi_{t_{n-2},t_{n-1}}(y_n + y_{n-1}) \cdots \varphi_{0,t_1}(y_n + \cdots + y_1).$$

Нам надо проверить условие из следствия. Оно сводится к случаю $t_1 > 0$. При $y_m = 0$ это произведение содержит прежние сомножители плоть до сомножителя $\varphi_{t_m,t_{m+1}}(y_n + \cdots + y_{m+1})$, затем идет $\varphi_{t_{m-1},t_m}(y_n + \cdots + y_{m+1})$ с таким же аргументом из-за обнуления y_m, а далее идут сомножители с аргументами вида

$$y_n + \cdots + y_{m+1} + y_{m-1} + \cdots\,,$$

где в индексах точка t_m появляется дважды: в соседних сомножителях $\varphi_{t_m,t_{m+1}}$ и φ_{t_{m-1},t_m}. Однако по нашему условию выполнено равенство

$$\varphi_{t_m,t_{m+1}}\varphi_{t_{m-1},t_m} = \varphi_{t_{m-1},t_{m+1}}$$

в общей точке $y_n + \dots + y_{m+1}$. Это дает условие из следствия. Итак, существует нужный процесс с нулевым начальным значением. Теперь к этому процессу можно прибавить независимую с ним случайную величину с распределением Q_0. На уровне характеристических функций это сводится к домножению полученных характеристических функций с наборами, включающими $t_1 = 0$, на характеристическую функцию меры Q_0. $\square$

Отметим, что меры $Q_{s,t}$ определяются по двумерным распределениям процесса: такая мера есть образ двумерного распределения P_{ξ_s,ξ_t} при линейной функции $l(x,y) = y - x$. Если известна характеристическая функция φ_{ξ_s,ξ_t} двумерного распределения, то $\varphi_{s,t}(y) = \varphi_{\xi_s,\xi_t}(-y,y)$.

Важнейшие примеры процессов с независимыми приращениями — рассматриваемые ниже пуассоновский и винеровский процессы, а также процессы Леви.

§ 5. Пуассоновский процесс

5.1. Определение. *Пуассоновским процессом с параметром интенсивности $\lambda > 0$ называется процесс N_t на $T = [0, +\infty)$ со следующими свойствами: процесс N_t имеет независимые приращения, $N_0 = 0$, при $s < t$ приращение $N_t - N_s$ имеет распределение Пуассона с параметром $\lambda(t-s)$, т.е. $P(N_t - N_s = k) = (\lambda(t-s))^k e^{-\lambda(t-s)}/k!$, $k = 0, 1, \ldots$.*

Из определения следует, что N_t имеет распределение Пуассона с параметром λt. Ясно также, что $N_t \geqslant 0$ и $N_t \geqslant N_s$ при $t > s$ п.н.

Существование пуассоновского процесса следует из доказанной выше общей теоремы про процессы с независимыми приращениями. В самом деле, характеристическая функция приращения равна

$$\varphi_{s,t}(\tau) = \sum_{k\geqslant 0} e^{ik\tau}(\lambda(t-s))^k e^{-\lambda(t-s)}/k! = \exp\Bigl((e^{i\tau} - 1)\lambda(t-s)\Bigr).$$

Поэтому тождество $\varphi_{s,t} = \varphi_{s,u}\varphi_{u,t}$ при $0 \leqslant s < u < t$ выполнено.

Однако этот очень важный для теории и приложений процесс можно задать явно как процесс восстановления. Напомним, что экспоненциальное распределение с параметром $\lambda > 0$ имеет плотность распределения, равную нулю на $(-\infty, 0]$ и $\lambda e^{-\lambda x}$ на $(0, +\infty)$.

5.2. Теорема. *Пусть η_t — процесс восстановления, построенный из независимых случайных величин ξ_n с экспоненциальным распределением с параметром λ по формуле $\eta_t(\omega) = \sup\{n\colon S_n(\omega) \leqslant t\}$, где*

$S_n = \xi_1 + \dots + \xi_n$, $S_0 = 0$, $t > 0$, $\eta_0 = 0$. Тогда η_t — пуассоновский процесс интенсивности λ.

Доказательство. Сначала мы найдем распределение случайного вектора $(S_1, \dots, S_n)$. Оно получается из распределения $(\xi_1, \dots, \xi_n)$, равного степени распределения ξ_1, линейным преобразованием с матрицей такого вида: в строке с номером k стоит 1 до позиции k, затем 0 на остальных. Ее определитель равен 1, поэтому плотность распределения равна $p_{\xi_1,\dots,\xi_n}(x_1, x_2 - x_1, \dots, x_n - x_{n-1})$, что равно

$$p_{\xi_1}(x_1)p_{\xi_1}(x_2-x_1)\cdots p_{\xi_1}(x_n-x_{n-1}) = \prod_{j=1}^{n} \lambda e^{-\lambda(x_j - x_{j-1})} I_{\{x_j > x_{j-1}\}} = \\ = \lambda^n e^{-\lambda x_n} I_{\{x_n > x_{n-1} > \dots > x_1\}}.$$

Зафиксируем моменты времени $t_0 = 0 < t_1 < \dots < t_n$ и неотрицательные целые числа $k_0 = 0 \leqslant k_1 \leqslant \dots \leqslant k_n$. Для нашего процесса вычислим вероятность

$$P(\eta_{t_n} - \eta_{t_{n-1}} = k_n - k_{n-1}, \dots, \eta_{t_2} - \eta_{t_1} = k_2 - k_1, \eta_{t_1} = k_1).$$

Надо убедиться, что она распадается в произведение вероятностей указанных событий, причем в случае $n = 2$ имеет нужный для пуассоновости вид. Решающее наблюдение состоит в том, что это есть вероятность события

$$S_1, \dots, S_{k_1} \in (0, t_1),\ S_{k_1+1}, \dots, S_{k_2} \in (t_1, t_2), \dots, \\ S_{k_{n-1}+1}, \dots, S_{k_n} \in (t_{n-1}, t_n),\ S_{k_n+1} > t_n.$$

В силу сделанного вычисления это есть интеграл

$$\int_U \lambda^{k_n+1} e^{-\lambda x_{k_n+1}} I_{\{0 < x_1 < \dots < x_{k_n+1}\}}\, dx_1 \cdots dx_{k_n+1}$$

по области U в $\mathbb{R}^{k_n+1}$, заданной соответствующими включениям для S_j неравенствами $x_1, \dots, x_{k_1} \in (0, t_1), \dots, x_{k_n+1} > t_n$. Если $n = 2$, то получаем интеграл по области $x_1, \dots, x_{k_1} \in (0, t_1)$, $x_{k_1+1}, \dots, x_{k_2} \in (t_1, t_2)$, $x_{k_n+1} > t_n$. Так как $x_{k_n+1} > t_n$ и $x_{k_n} < t_n$, то интеграл по x_{k_n+1} отделяется и дает множитель $\lambda^{k_n} e^{-\lambda t_n}$. Оставшийся интеграл есть объем области в $\mathbb{R}^{k_n}$, заданной указанными выше неравенствами без участия x_{k_n+1}. Можно проверить, что этот объем равен произведению $\prod_{j=1}^n (t_j - t_{j-1})^{k_j - k_{j-1}}/(k_j - k_{j-1})!$. Из этого вытекает как пуассоновость распределения каждого приращения, так и их независимость. $\square$

Разумеется, при таком подходе нуждается в обосновании существование независимой последовательности случайных величин с распределением Пуассона, но это проще доказать, чем общую теорему Колмогорова.

Пуассоновский процесс имеет среднее λt и такую же дисперсию.

Задача 5. Доказать, что ковариация пуассоновского процесса с параметром λ равна $\lambda \min(t, s)$.

§ 6. Гауссовские процессы

Напомним, что случайная величина называется гауссовской, если ее распределение либо сосредоточено в некоторой точке a (есть мера Дирака δ_a), либо имеет плотность $(2\pi\sigma)^{-1/2}\exp(-(x-a)^2/2\sigma)$ с некоторыми $a \in \mathbb{R}$ и $\sigma > 0$. Параметр $a = \mathrm{E}\xi$ — среднее меры, параметр $\sigma = \mathrm{E}(\xi - a)^2$ — дисперсия ξ. Мера со средним 0 и дисперсией 1 называется стандартной. Характеристическая функция гауссовской меры (или гауссовской случайной величины) с параметрами a, σ равна $\varphi(y) = \exp(iay - \sigma y^2/2)$.

Борелевская вероятностная мера γ на $\mathbb{R}^n$ называется гауссовской, если таковы ее образы при линейных функциях, т.е. всякая линейная функция $c_1x_1 + \ldots + c_nx_n$ должна быть гауссовской случайной величиной. Важно, что это не только координатные функции, но и все их линейные комбинации. Стандартной гауссовской мерой на $\mathbb{R}^n$ называется степень стандартной одномерной гауссовской меры, ее плотность равна $(2\pi)^{-n/2}\exp(-|x|^2/2)$. Легко проверить, что произведение гауссовских мер тоже является гауссовской мерой.

Из определения вытекает, что характеристическая функция меры γ имеет вид

$$\varphi(y) = \exp(i(y, a) - (Ky, y)/2),$$

где a — вектор, K — симметричный неотрицательно определенный оператор. При этом

$$(y, a) = \int (y, x)\,\gamma(dx), \quad (Ku, v) = \int (u, x-a)(v, x-a)\,\gamma(dx).$$

Обратно, функция указанного вида является характеристической для гауссовской меры, полученной как образ стандартной гауссовской меры при аффинном отображении $x \mapsto \sqrt{K}x + a$. Поскольку оператор K имеет собственный ортонормированный базис, получаем, что всякая гауссовская мера распадается в произведение одномерных гауссовских мер в подходящем ортонормированном базисе.

Случайный вектор $\xi = (\xi_1, \ldots, \xi_n)$ называется гауссовским, если его распределение в $\mathbb{R}^n$ есть гауссовская мера. Это равносильно тому, что все линейные комбинации $c_1\xi_1 + \cdots + c_n\xi_n$ — гауссовские случайные величины (гауссовости лишь самих компонент ξ_i для этого недостаточно).

Важным свойством гауссовских случайных векторов является то, что независимость их компонент равносильна некоррелированности, т.е. равенству $\mathrm{E}(\xi_i - \mathrm{E}\xi_i)(\xi_j - \mathrm{E}\xi_j) = 0$ при $i \neq j$. Это проверяется с помощью преобразования Фурье.

6.1. Определение. *Случайный процесс называется гауссовским, если все его конечномерные распределения гауссовские.*

Из сказанного выше следует, что гауссовость процесса ξ_t равносильна тому, что все случайные величины $c_1\xi_{t_1} + \cdots + c_n\xi_{t_n}$ с постоянными c_i являются гауссовскими. Это требование сильнее, нежели гауссовость каждой отдельной величины ξ_t.

Поскольку гауссовская мера определяется средним и ковариацией, то гауссовский процесс определяется средним $M(t) = \mathrm{E}\xi_t$ и ковариацией $K(s,t) = \mathrm{E}(\xi_t\xi_s) - \mathrm{E}\xi_t\mathrm{E}\xi_s$.

Ковариация симметрична и неотрицательно определена в следующем смысле:

$$\sum_{i,j\leqslant n} K(t_i,t_j)x_ix_j \geqslant 0$$

для всех $x_i \in \mathbb{R}$. Иначе говоря, матрица с элементами $K(t_i,t_j)$ симметрична и неотрицательно определена. Симметричность очевидна, а неотрицательность указанной квадратичной формы следует из равенства

$$\sum_{i,j\leqslant n} K(t_i,t_j)x_ix_j = \mathrm{E}\Big|\sum_{i=1}^{n} x_i(\xi_{t_i} - \mathrm{E}\xi_{t_i})\Big|^2.$$

Верно и обратное.

6.2. Теорема. *Пусть даны функция M на T и симметричная функция K на $T\times T$, неотрицательно определенная в указанном выше смысле. Тогда существует гауссовский процесс, для которого M есть среднее, а K есть ковариация.*

Доказательство. Здесь также удобнее пользоваться следствием 4.2, чем самой теоремой Колмогорова. Для фиксированных элементов $t_1, \ldots, t_n \in T$ функция

$$\varphi_{t_1,\ldots,t_n}(y) = \exp\Big(i \sum_{j=1}^{n} M(t_j) y_j - \frac{1}{2} \sum_{j,m \leqslant n} K(t_j, t_m) y_j y_m\Big)$$

служит характеристической для гауссовской меры на пространстве $\mathbb{R}^n$ со средним $(M(t_1), \ldots, M(t_n))$ и ковариационной матрицей $(K(t_j, t_m))$. При этом выполнено нужное условие согласованности: ограничение $\varphi_{t_1,\ldots,t_n}$ на плоскость $y_m = 0$ совпадает с функцией φ для набора без точки t_m. □

Если $M(t) = 0$ и $K(t,t) = 1$, $K(t,s) = 0$ при $t \neq s$, то получаем гауссовский процесс с независимыми значениями, но этот процесс редко бывает полезен. Гораздо чаще используется винеровский процесс, который вводится не так просто.

6.3. Пример. Посмотрим, что требуется для задания гауссовского процесса с независимыми приращениями и нулевым средним. Как мы видели, нужны характеристические функции $\varphi_{s,t}$ разностей $\xi_t - \xi_s$, причем они должны удовлетворять некоему тождеству. В гауссовском случае $\varphi_{s,t}(y) = \exp(-y^2\sigma(s,t)/2$, где $\sigma(s,t)$ — дисперсия разности. Нужное тождество приобретает вид

$$\sigma(s,t) = \sigma(s,u) + \sigma(u,t), \quad s \leqslant u \leqslant t.$$

Отметим еще, что гауссовость приращений процесса с независимыми приращениями и $\xi_0 = 0$ влечет его гауссовость, так как гауссовскими оказываются все линейные комбинации $c_1\xi_{t_1} + c_2(\xi_{t_2} - \xi_{t_1}) + \cdots + c_n(\xi_{t_n} - \xi_{t_{n-1}})$, а в таком виде можно представить любую линейную комбинацию $\alpha_1\xi_{t_1} + \alpha_2\xi_{t_2} + \cdots + \alpha_n\xi_{t_n}$.

Если еще предположить, что $\varphi_{s,t}$ зависит только от разности, то $\sigma(s,t) = \psi(t-s)$, поэтому $\psi(t-s) = \psi(u-s) + \psi(t-u)$. Иначе говоря, $\psi(x) + \psi(y) = \psi(x+y)$ при $x, y \geqslant 0$. Среди непрерывных (или хотя бы измеримых) функций этому условию удовлетворяют функции $\psi(x) = kx$. Выбор $k = 1$ дает винеровский процесс.

§ 7. Винеровский процесс

7.1. Определение. *Винеровским процессом называется процесс* $\{w_t\}_{t \geqslant 0}$ *с независимыми приращениями, для которого* $w_0 = 0$ *п.н. и* $w_t - w_s$ *при* $t > s \geqslant 0$ *имеет гауссовское распределение со средним* 0 *и дисперсией* $t - s$.

Аналогично определяется винеровский процесс на отрезке $[0,T]$.

Винеровский процесс часто называют броуновским движением, так как он хорошо описывает реальное тепловое движение частиц в жидкости, открытое английским ботаником Р. Броуном (1773–1858), который на самом деле, конечно, Браун (Robert Brown), как и отец Браун у Честертона.

Таким образом, $\mathrm{E}w_t = 0$, $\mathrm{E}\exp(i\lambda(w_t - w_s)) = \exp(-(t-s)\lambda^2/2)$ при $t > s$, приращения $w_{t_1}, w_{t_2} - w_{t_1}, \ldots, w_{t_n} - w_{t_{n-1}}$ при $0 \leqslant t_1 \leqslant t_2 \leqslant t_n$ независимы.

Сами случайные величины w_t являются гауссовскими с нулевыми средними и дисперсиями t, но не независимыми. Из рассмотренного выше примера вытекает, что винеровский процесс существует.

Нередко к требуемым свойствам винеровского процесса добавляют еще такое:

почти все траектории $w_\bullet(\omega)$ непрерывны.

Однако теорема Колмогорова о непрерывных модификациях обеспечивает это. В самом деле,

$$\mathrm{E}(w_t - w_s)^4 = 3(t-s)^2.$$

Для построения винеровского процесса можно применять и саму теорему Колмогорова, но для этого надо вычислить его ковариационную функцию и проверить ее неотрицательную определенность. Вычисление простое: при $t > s$ имеем $\mathrm{E}(w_t - w_s)^2 = t - s$, левая часть есть $t + s - 2\mathrm{E}(w_t w_s)$. Итог:

$$\mathrm{E}(w_t w_s) = \min(t,s).$$

Неотрицательную неопределенность можно не проверять, так как такая же ковариация у пуассоновского процесса с параметром 1 (задача 5).

Из определения нетрудно найти конечномерные распределения, так как при $0 < t_1 < \cdots < t_n$ распределение вектора из приращений $(w_{t_1}, w_{t_2} - w_{t_1}, \ldots, w_{t_n} - w_{t_{n-1}})$ равно произведению одномерных распределений этих приращений, то искомое распределение самого вектора $(w_{t_1}, w_{t_2}, \ldots, w_{t_n})$ получается линейным преобразованием, обратным оператору $Tx = (x_1, x_2 - x_1, x_n - x_{n-1})$. Вычисление дает такой ответ для плотности:

$$p_{t_1,\ldots,t_n}(x_1,\ldots,x_n) = \prod_{i=1}^{n} \bigl(2\pi(t_i - t_{i-1})\bigr)^{-1/2} \exp\Bigl(-\sum_{i=1}^{n} \frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}\Bigr),$$

где $x_0 = 0$, $t_0 = 0$.

В следующем параграфе приведены некоторые сведения о траекториях винеровского процесса, но сейчас приведем полезное утверждение (к сожалению, весьма трудно доказываемое) о представлении винеровского процесса на отрезке в виде функционального ряда.

7.2. Теорема. *Пусть $\{e_n\}$ — ортонормированный базис в вещественном пространстве $L^2[0,T]$ и $\{\xi_n\}$ — последовательность независимых стандартных гауссовских случайных величин. Тогда ряд*

$$\sum_{n=1}^{\infty} \xi_n(\omega)\varphi_n(t), \quad \varphi_n(t) = \int_0^t e_n(s)\,ds$$

равномерно сходится на $[0,T]$ при п.в. ω, причем его сумма $w_t(\omega)$ есть винеровский процесс.

Трудность в этой теореме — обоснование сходимости ряда. Однако если удовольствоваться сходимостью в L^2 вместо равномерной, то легко видеть, что она имеет место для стандартного тригонометрического базиса в $L^2[0,2\pi]$. В самом деле, для $\sin nt$ получаем $\varphi_n(t) = cn^{-1}(\cos nt - 1)$, для $\cos nt$ при $n > 0$ получаем $\varphi_n(t) = cn^{-1}\sin nt$. Таким образом, нам нужна сходимость в L^2 ряда вида

$$\sum_{n=1}^{\infty} n^{-1}\xi_n(\omega)u_n(t),$$

где $\{u_n\}$ — ортонормированный базис. Для сходимости такого ряда нужна сходимость ряда из квадратов коэффициентов, т.е. условие

$$\sum_{n=1}^{\infty} n^{-2}\xi_n(\omega)^2 < \infty.$$

Сходимость имеет место почти наверное, так как сходится ряд из интегралов в силу того, что интеграл от ξ_n^2 равен 1. Далее можно считать, что сходимость имеет место при всех ω. Такой способ дает некий процесс $w_t(\omega)$, хотя не очевидно, почему у него непрерывные траектории. Но можно это проигнорировать и просто проверить, что это процесс с независимыми приращениями и нужными распределениями $w_t - w_s$. Правда, сходимость в L^2 еще не задает процесс при всех t. Эту проблему можно обойти так. Из сказанного видно, что ряд сходится также в L^2 по произведению меры Лебега и P. Как известно, тогда есть подпоследовательность частичных сумм $\sum_{n=1}^{N_k} n^{-1}\xi_n(\omega)u_n(t)$, которая сходится почти всюду на $[0,2\pi] \times \Omega$. По теореме Фубини имеется такое множество $S \subset [0,2\pi]$, что его дополнение имеет меру нуль и при каждом $t \in S$ указанные частичные суммы сходятся при почти всех ω. Конечный предел этих сумм обозначим через $w_t(\omega)$ для всех пар (t, ω), для которых есть сходимость. Ниже вычислено, что $\mathrm{E}|w_t - w_s|^2 = t - s$, поэтому для $t \notin S$ можно задать w_t как предел в $L^2(P)$ функций W_{s_n}, где $s_n \in S$ и $s_n \to t$ (предел не зависит от выбора такой последовательности). Теперь процесс определен при всех $t \in [0, 2\pi]$. Ясно, что среднее нулевое. Как найти математическое ожидание $(w_t - w_s)^2$? Непосредственное вычисление с учетом независимости ξ_n дает ряд

$$\sum_{n=1}^{\infty} \left(\int_s^t e_n(\tau)\, d\tau \right)^2.$$

Что это за ряд? Замечаем, что суммируются квадраты скалярных произведений $(I_{[s,t]}, e_n)_{L^2}$. В силу равенства Парсеваля получаем квадрат нормы $I_{[s,t]}$, т.е. нужную разность $t - s$. Для проверки независимости приращений замечаем, что процесс-то гауссовский из-за гауссовости и независимости ξ_n. Еще вспоминаем, что у гауссовского вектора независимость компонент следует из некоррелированности. Поэтому остается проверить, что $\mathrm{E}[(w_{t_4} - w_{t_3})(w_{t_2} - w_{t_2})] = 0$ при $t_1 \leqslant t_2 \leqslant t_3 \leqslant t_4$. Как и выше, слева получаем

$$\sum_{n=1}^{\infty} \int_{t_3}^{t_4} e_n(\tau)\, d\tau \int_{t_1}^{t_2} e_n(\tau)\, d\tau = \sum_{n=1}^{\infty} (I_{[t_3,t_4]}, e_n)_{L^2} (I_{[t_1,t_2]}, e_n)_{L^2}.$$

В правой части узнаем $(I_{[t_3,t_4]}, I_{[t_1,t_2]})_{L^2}$, что равно нулю, поскольку имеем $[t_3, t_4] \cap [t_1, t_2] = \varnothing$ или $\{t_2\}$. Есть и другие способы построения винеровского процесса, но никакой из них не сравнится по простоте со способом Колмогорова.

Задача 6. Пусть w_t — винеровский процесс. Процесс $\xi_t = e^{-t} w_{e^{2t}}$ называется процессом Орнштейна–Уленбека. Доказать, что это гауссовский процесс, найти его ковариационную функцию.

Задача 7. Пусть $H \in (0, 1]$. Доказать, что функция

$$K_H(s, t) = \frac{1}{2}\Big(|s|^{2H} + |t|^{2H} - |s - t|^{2H} \Big),$$

где $s, t \in \mathbb{R}$ или $s, t \geqslant 0$, является ковариационной для гауссовского процесса B_t^H с нулевым средним. Этот процесс называется дробным броуновским движением с параметром или индексом Х■рста H. При $H = 1/2$ и $t \geqslant 0$ это винеровский процесс. Для доказательства сначала проверить, что равносильное описание такого гауссовского процесса дается соотношением $B_0^H = 0$,г $\mathrm{E}(B_t^H - B_s^H)^2 = |t - s|^{2H}$. Поэтому достаточно подобрать кривую X_t в гильбертовом пространстве, для которой $\|X_t - X_s\|^2 = |t-s|^{2H}$.

§ 8. Некоторые свойства винеровского процесса

Теорема Колмогорова дает версию винеровского процесса с почти наверное непрерывными траекториями. На самом деле доказательство дает даже больше: почти всякая траектория $w_\bullet(\omega)$ г■льдерова всякого порядка $\alpha < 1/2$, т.е.

$$|w_{t+h}(\omega) - w_t(\omega)| \leqslant C_\alpha(\omega)|h|^\alpha,$$

где случайная величина C_α интегрируема в любой степени. Почему $\alpha < 1/2$ и не годится $1/2$? Это видно из того, что, например, нормированное приращение $h^{-1/2}(w_{t_i+h} - w_{t_i})$ есть стандартная гауссовская случайная величина. На дизъюнктных отрезках такие величины независимы, поэтому вероятность того, что они одновременно не больше фиксированного R, стремится к нулю с ростом числа отрезков. На самом деле нет г■льдеровости даже в фиксированной точке, что видно из следующей теоремы («закон повторного логарифма»).

8.1. Теорема. *Для винеровского процесса w_t с вероятностью* 1 *верны равенства*

$$\limsup_{t\to+\infty} \frac{|w_t(\omega)|}{\sqrt{2t\ln\ln t}} = 1, \quad \limsup_{t\to 0} \frac{|w_t(\omega)|}{\sqrt{2t\ln|\ln t|}} = 1.$$

Из второго равенства следует, что для фиксированной точки t_0 с вероятностью 1 имеем

$$\limsup_{h\to 0} \frac{|w_{t_0+h}(\omega) - w_{t_0}(\omega)|}{\sqrt{2h\ln|\ln h|}} = 1.$$

Из этого ясно, что почти всякая броуновская траектория не дифференцируема в точке t_0, но верно даже больше: почти всякая броуновская траектория не дифференцируема ни в одной точке.

Близкие свойства у так называемого броуновского моста, заданного на $[0,1]$ формулой

$$b_t(\omega) = w_t(\omega) - tw_1(\omega).$$

Отсутствие точек дифференцируемости влечет неограниченность вариации траектории винеровского процесса на всяком нетривиальном подотрезке. Однако это обстоятельство удивительным образом не мешает определить стохастический интеграл

$$\int_0^T f(t)\,dw_t$$

по винеровскому процессу для неслучайной функции $f \in L^2[0,T]$. Такой интеграл нельзя задать как интеграл Стилтьеса, но для ступенчатой функции f так сделать можно: если $f(t) = c_k$ при $t \in [t_{k-1}, t_k)$, где $0 = t_0 < t_1 < \cdots < t_n = T$, то положим

$$I(f) := \int_0^T f(t)\,dw_t := c_1 w_{t_1} + c_2(w_{t_2} - w_{t_1}) + \cdots + c_n(w_{t_n} - w_{t_{n-1}}).$$

Тогда $I(f)$ — гауссовская случайная величина со средним 0 и дисперсией $c_1^2 t_1 + c_2^2(t_2 - t_1) + \cdots + c_n^2(t_n - t_{n-1})$, равной $\|f\|^2_{L^2[0,T]}$. Кроме того,

$$(I(f), I(g))_{L^2(P)} = (f,g)_{L^2[0,T]}.$$

На линейном подпространстве ступенчатых функций в $L^2[0,T]$ получен линейный оператор со значениями в $L^2(P)$, сохраняющий скалярное произведение. Это подпространство всюду плотно в $L^2[0,T]$. Поэтому оператор I продолжается по непрерывности до линейного оператора $I\colon L^2[0,T] \to L^2(P)$, тоже сохраняющего скалярное произведение. Продолжение и берется в качестве стохастического интеграла.

Задача 8. Доказать, что для винеровского процесса с вероятностью 1 суммы $\sum_{j=0}^{2^n}(w_{t_{j+1}} - w_{t_j})^2$ при $n \to \infty$ стремятся к 1, где $t_j = j2^{-n}$.

Задача 9. Доказать, что броуновский мост является гауссовским процессом, найти его ковариацию, выяснить, имеет ли он независимые приращения.

Задача 10. Доказать, что для всякой функции $f \in L^2[0,T]$ ее стохастический интеграл $I(f)$ по винеровскому процессу есть гауссовская случайная величина со средним 0 и дисперсией $\|f\|^2_{L^2[0,T]}$.

§ 9. Процессы Леви

9.1. Определение. *Случайный процесс $\{\xi_t\}_{t\in T}$, где $T \subset \mathbb{R}^1$, называется стохастически непрерывным, если $P(|\xi_t - \xi_s| \geqslant \varepsilon) \to 0$ при $t \to s, t, s \in T$, для всякого $\varepsilon > 0$.*

Равносильное свойство: $\mathrm{E}[|\xi_t - \xi_s|/(1 + |\xi_t - \xi_s|)] \to 0$ при $t \to s$. Не требуется наличия матожидания и дисперсии, но если дисперсия есть, то достаточным условием стохастической непрерывности является непрерывность ковариации.

9.2. Определение. *Процессом Леви называется стохастически непрерывный случайный процесс $\{\xi_t\}_{t\geqslant 0}$ с независимыми приращениями и $\xi_0 = 0$, для которого распределение $\xi_t - \xi_s$ зависит только от разности $t - s$.*

Можно показать, что такой процесс имеет модификацию с траекториями, непрерывными справа и имеющими пределы слева.

Пуассоновский и винеровский процессы являются процессами Леви. Если взять два таких независимых процесса, то их сумма тоже будет процессом Леви.

§ 10. Условные вероятности и условные математические ожидания

Здесь кратко напоминаются основные определения и факты, связанные с условными математическими ожиданиями. В элементарной теории вероятностей (с конечными или счетными пространствами) условная вероятность события B при условия события A ненулевой вероятности задается равенством $P(B|A) = P(A \cap B)/P(A)$. В таком духе понимается условная вероятность $P(B|\eta(\omega) = x)$ при условии равенства случайной величины числу x. Появляется здесь и условное математическое ожидание. Однако в более общих ситуациях оказываются полезными условные математические ожидания и вероятности, не задаваемые множествами положительной меры.

Пусть $(\Omega, \mathfrak{B}, P)$ — вероятностное пространство, $\mathcal{A}$ — под-σ-алгебра основной σ-алгебры $\mathfrak{B}$.

10.1. Определение. *Пусть ξ — интегрируемая случайная величина. Ее условным математическим ожиданием относительно $\mathcal{A}$ называется интегрируемая функция $\mathrm{E}^{\mathcal{A}}\xi$, обозначаемая также через $\mathrm{E}(\xi|\mathcal{A})$, измеримая относительно $\mathcal{A}$ и удовлетворяющая равенству*

$$\mathrm{E}(\eta\xi) = \mathrm{E}(\eta\mathrm{E}^{\mathcal{A}}\xi)$$

для всякой ограниченной функции η, измеримой относительно $\mathcal{A}$.

Второе обозначение $\mathrm{E}(\xi|\mathcal{A})$ оказывается часто удобным, хотя оно чуть длиннее и менее удобно для подстановки аргументов.

В терминах интегралов имеем

$$\int_\Omega \eta\xi\, dP = \int_\Omega \eta\mathrm{E}^{\mathcal{A}}\xi\, dP.$$

Так как ограниченные $\mathcal{A}$-измеримые функции равномерно приближаются простыми $\mathcal{A}$-измеримыми функциями (т.е. с конечным числом значений), это равносильно тождеству

$$\int_A \xi \, dP = \int_A \mathrm{E}^{\boldsymbol{\mathcal{A}}} \xi \, dP \quad \forall \, A \in \boldsymbol{\mathcal{A}}.$$

Если $\xi \in L^2(P)$, то определяющее равенство можно записать как

$$(\xi - \mathrm{E}^{\boldsymbol{\mathcal{A}}} \xi, \eta)_{L^2} = 0$$

для всех ограниченных $\boldsymbol{\mathcal{A}}$-измеримых функций η. Таким образом, разность $\xi - \mathrm{E}^{\boldsymbol{\mathcal{A}}}\xi$ должна быть ортогональна в $L^2(P)$ линейному подпространству, соответствующему таким функциям. Значит, она ортогональна и его замыканию $L^2_{\boldsymbol{\mathcal{A}}}$ в гильбертовом пространстве $L^2(P)$. Это позволяет задать $\mathrm{E}^{\boldsymbol{\mathcal{A}}}\xi$ как ортогональную проекцию ξ на $L^2_{\boldsymbol{\mathcal{A}}}$. Правда, $L^2(P)$ и $L^2_{\boldsymbol{\mathcal{A}}}$ состоят не из индивидуальных функций, а из классов эквивалентности. Однако это не помеха: нетрудно проверить, что всякий элемент из $L^2_{\boldsymbol{\mathcal{A}}}$ имеет представителя, измеримого относительно $\boldsymbol{\mathcal{A}}$. Так как функция $\eta = 1$ измерима относительно $\boldsymbol{\mathcal{A}}$, из определения очевидно свойство

$$\int_\Omega \mathrm{E}^{\boldsymbol{\mathcal{A}}} \xi \, dP = \int_\Omega \xi \, dP.$$

Для функций из $L^2(P)$ получается линейность $\mathrm{E}^{\boldsymbol{\mathcal{A}}}\xi$ по ξ, а также условие $\mathrm{E}^{\boldsymbol{\mathcal{A}}}\xi \geqslant 0$ п.в. при $\xi \geqslant 0$ п.в. Значит, $\mathrm{E}^{\boldsymbol{\mathcal{A}}}\xi_1 \leqslant \mathrm{E}^{\boldsymbol{\mathcal{A}}}\xi_2$ п.в., если $\xi_1 \leqslant \xi_2$ п.в. Это позволяет продолжить $\mathrm{E}^{\boldsymbol{\mathcal{A}}}$ на $L^1(P)$. Достаточно сделать это при $\xi \geqslant 0$. Для срезок $\min(\xi, n)$ уже есть $\psi_n = \mathrm{E}^{\boldsymbol{\mathcal{A}}} \min(\xi, n)$, причем $\psi_n \leqslant \psi_{n+1}$ п.в. Кроме того, $\mathrm{E}\psi_n = \mathrm{E}\min(\xi, n) \leqslant \mathrm{E}\xi$. По теореме о монотонной сходимости функция $\psi = \lim\limits_{n\to\infty} \psi_n$ конечна почти всюду и интегрируема. Легко видеть, что она имеет $\boldsymbol{\mathcal{A}}$-измеримую версию и может быть взята в качестве $\mathrm{E}^{\boldsymbol{\mathcal{A}}}\xi$. Продолжение $\mathrm{E}^{\boldsymbol{\mathcal{A}}}\xi$ на $L^1(P)$ остается линейным по ξ.

Отметим, что $\mathrm{E}^{\boldsymbol{\mathcal{A}}}\xi$ единственно с точностью до переопределения на множестве из $\boldsymbol{\mathcal{A}}$ меры нуль. Для квадратично интегрируемых функций это ясно из единственности проекций, а для лишь интегрируемых следует из того, что если f и g измеримы относительно $\boldsymbol{\mathcal{A}}$ и имеют равные интегралы по всем множествам $A \in \boldsymbol{\mathcal{A}}$, то $f = g$ п.в. В самом деле, взяв $A = \{f > g\}$, заключаем, что $P(A) = 0$, иначе интеграл по A от f будет больше интеграла от g. Затем берем $A = \{f < g\}$.

Если $\boldsymbol{\mathcal{A}}$ есть σ-алгебра, порожденная случайной величиной ψ, то всякая $\boldsymbol{\mathcal{A}}$-измеримая функция f имеет вид $f(\omega) = g(\psi(\omega))$, где g —

борелевская функция. В этом случае под выражением $\mathrm{E}(\xi|\psi = x)$ понимается функция $g(x)$, указанным образом представляющая функцию $f(\omega) = \mathrm{E}^{\mathcal{A}}\xi(\omega)$ как $g(\psi(\omega))$.

Для $\mathcal{A}$-измеримых функций ξ имеем $\mathrm{E}^{\mathcal{A}}\xi = \xi$ по определению, в частности $\mathrm{E}^{\mathcal{A}}1 = 1$, но есть еще одно простое свойство близкого характера: если ограниченная функция ζ измерима относительно $\mathcal{A}$, то ее можно выносить из-под условного ожидания:

$$\mathrm{E}^{\mathcal{A}}(\zeta\xi) = \zeta\mathrm{E}^{\mathcal{A}}\xi. \tag{10.1}$$

В самом деле, для всякой $\mathcal{A}$-измеримой ограниченной функции η функция $\zeta\mathrm{E}^{\mathcal{A}}\xi$, измеримая относительно $\mathcal{A}$, удовлетворяет равенству

$$\mathrm{E}(\eta\zeta\mathrm{E}^{\mathcal{A}}\xi) = \mathrm{E}(\eta\zeta\xi),$$

так как $\eta\zeta$ измерима относительно $\mathcal{A}$. Значит, $\zeta\mathrm{E}^{\mathcal{A}}\xi$ годится в качестве $\mathrm{E}^{\mathcal{A}}(\zeta\xi)$.

Равенство (10.1) можно записать в виде

$$\mathrm{E}^{\mathcal{A}}(g\mathrm{E}^{\mathcal{A}}f) = \mathrm{E}^{\mathcal{A}}g\mathrm{E}^{\mathcal{A}}f$$

для всех ограниченных $\mathfrak{B}$-измеримых g. В таком виде оно характеризует условное ожидание: можно доказать, что если ортогональный проектор T в $L^2(P)$ обладает свойствами $T1 = 1$ и $T(gTf) = TgTf$ для всех ограниченных f и g, то найдется такая под-σ-алгебра $\mathcal{A}$, что $Tf = \mathrm{E}^{\mathcal{A}}f$.

Если $\mathcal{A}_0 \subset \mathcal{A}$ — еще меньшая под-σ-алгебра, то

$$\mathrm{E}^{\mathcal{A}_0}(\mathrm{E}^{\mathcal{A}}\xi) = \mathrm{E}^{\mathcal{A}_0}\xi.$$

Для множеств принято писать $P(B|\mathcal{A}) = \mathrm{E}^{\mathcal{A}}I_B$. Если $\mathcal{A}$ порождается функцией ψ, то появляется выражение $P(B|\psi = x)$.

Как это связано со школьным $P(B|A)$? Если конечное вероятностное пространство разбито на множества $A_1, \ldots, A_k$ положительной меры, то можно взять σ-алгебру $\mathcal{A}$, ими порожденную. Она состоит из их конечных объединений. Измеримые относительно нее функции постоянны на этих множествах. Тогда для множества B из данного пространства $\mathrm{E}(B|\mathcal{A})$ есть функция, которая на A_j принимает значение $P(B|A_j) = P(B \cap A_j)/P(A_j)$. Если ее проинтегрировать по A_j, то получится $P(B \cap A_j)$.

Задача 11. Пусть интегрируемая случайная величина ξ независима с σ-алгеброй $\mathcal{A}$, т.е. со всеми $\mathcal{A}$-измеримыми функциями. Доказать, что $\mathrm{E}^{\mathcal{A}}\xi$ совпадает с числом $\mathrm{E}\xi$.

Для условных математических ожиданий аналогично обычным интегралам верно неравенство Йенсена: если V — выпуклая функция и интегрируемая случайная величина ξ такова, что $V(\xi)$ тоже интегрируема, то

$$V(\mathrm{E}^{\mathcal{A}}\xi) \leqslant \mathrm{E}^{\mathcal{A}}V(\xi) \quad \text{п.в.}$$

В частности, для $V(t) = |t|^p$ при $p \geqslant 1$ получаем

$$|\mathrm{E}^{\mathcal{A}}\xi|^p \leqslant \mathrm{E}^{\mathcal{A}}|\xi|^p \quad \text{п.в.}$$

Проверка неравенства Йенсена сводится к случаю ограниченных функций, а затем с помощью равномерных приближений к случаю простых функций, когда оно очевидно: если $\xi = c_1 I_{B_1} + \cdots + c_n I_{B_n}$, где множества B_i дизъюнктны, то

$$\mathrm{E}^{\mathcal{A}}\xi = c_1 \mathrm{E}^{\mathcal{A}} I_{B_1} + \cdots + c_n \mathrm{E}^{\mathcal{A}} I_{B_n},$$

где $\mathrm{E}^{\mathcal{A}} I_{B_i} \geqslant 0$ и $c_1 \mathrm{E}^{\mathcal{A}} I_{B_1} + \cdots + \mathrm{E}^{\mathcal{A}} I_{B_n} = 1$ п.в., ибо $I_{B_1} + \cdots + I_{B_n} = 1$. В силу выпуклости получаем п.в.

$$\begin{aligned} V(\mathrm{E}^{\mathcal{A}}\xi) &= V(c_1 \mathrm{E}^{\mathcal{A}} I_{B_1} + \cdots + c_n \mathrm{E}^{\mathcal{A}} I_{B_n}) \\ &\leqslant (\mathrm{E}^{\mathcal{A}} I_{B_1}) V(c_1) + \cdots + (\mathrm{E}^{\mathcal{A}} I_{B_n}) V(c_n), \end{aligned}$$

что совпадает с $\mathrm{E}^{\mathcal{A}}V(\xi)$, ибо $V(\xi) = V(c_1) I_{B_1} + \cdots + V(c_n) I_{B_n}$.

10.2. Замечание. Независимость случайной величины ξ с под-σ-алгеброй $\mathcal{A}$, т.е. независимость ξ и η для всякой $\mathcal{A}$-измеримой случайной величины η, равносильна равенству

$$\mathrm{E}^{\mathcal{A}}[f(\xi)] = \mathrm{E}[f(\xi)]$$

для всех ограниченных борелевских (или непрерывных) функций f на прямой, так как она есть тождество

$$\mathrm{E}(I_{\xi \in B} I_A) = P(\xi \in B) P(A), \quad B \in \mathcal{B}(\mathbb{R}),\ A \in \mathcal{A},$$

а это тождество равносильно равенству

$$\mathrm{E}[f(\xi) g] = \mathrm{E} f \mathrm{E} g$$

для всех ограниченных борелевских функций f на прямой и всех ограниченных $\mathcal{A}$-измеримых функций g, что и есть указанное равенство.

§ 11. Фильтрации и мартингалы

В теории случайных процессов оказывается полезным рассматривать не одну заданную σ-алгебру, как это бывает в базовых постановках теории вероятностей (тех, где вообще нужны σ-алгебры), и даже не две, как в условных ожиданиях, а целые бесконечные семейства. Говорят, что на вероятностном пространстве $(\Omega, \mathcal{B}, P)$ задана *фильтрация*, если задано семейство под-σ-алгебр $\mathcal{F}_t$ в $\mathcal{B}$, индексируемых параметром t из некоего множества T на прямой (или иного упорядоченного множества), причем $\mathcal{F}_s \subset \mathcal{F}_t$ при $s \leqslant t$.

Практический смысл множеств из $\mathcal{F}_t$ в приложениях часто бывает таков: это события, о наступлении или ненаступлении которых известно к моменту времени t.

11.1. Определение. *Случайный процесс $\{\xi_t\}_{t\in T}$ называется согласованным с фильтрацией $\{\mathcal{F}_t\}_{t\in T}$, если при каждом t случайная величина ξ_t измерима относительно $\mathcal{F}_t$.*

Всякий случайный процесс $\{\xi_t\}_{t\in T}$ согласован с фильтрацией, им порожденной, когда в качестве $\mathcal{F}_t$ берется σ-алгебра, порожденная всеми случайными величинами ξ_s с $s \leqslant t$. Для обсуждения марковских процессов такую σ-алгебру удобно обозначать символом $\mathcal{F}_{\leqslant t}$.

11.2. Определение. *Случайный процесс $\{\xi_t\}_{t\in T}$ называется мартингалом относительно фильтрации $\{\mathcal{F}_t\}_{t\in T}$, если он с ней согласован, случайные величины ξ_t интегрируемы и при $s \leqslant t$ п.в. верно равенство*

$$\mathrm{E}(\xi_t|\mathcal{F}_s) = \xi_s.$$

Если процесс квадратично интегрируем, то геометрическая интерпретация такова: есть возрастающее семейство σ-алгебр и кривая в гильбертовом пространстве $L^2(P)$, причем проекция ξ_t на подпространство для s есть значение кривой в s.

Заметим, что

$$\mathrm{E}\xi_t = \mathrm{E}\xi_s,$$

ибо $\mathrm{E}\xi_t = \mathrm{E}(\mathrm{E}(\xi_t|\mathcal{F}_s)) = \mathrm{E}\xi_s$.

Если $T \subset [0, +\infty)$, $0 \in T$ и $\xi_0 = 0$ или хотя бы $\mathrm{E}\xi_0 = 0$, то $\mathrm{E}\xi_t = 0$ для всех t.

Важный пример мартингала: последовательность $S_n = \xi_1 + \cdots + \xi_n$, где ξ_i независимы и имеют нулевые математические ожидания, $T = \mathbb{N}$, $\mathcal{F}_n$ порождена $\xi_1, \ldots, \xi_n$. В самом деле, при $m > n$ имеет место равенство $\mathrm{E}(\xi_m|\mathcal{F}_n) = \mathrm{E}\xi_m = 0$ в силу задачи 11 выше.

Если $\{\xi_t\}_{t\in T}$ — мартингал, причем $\xi_t \in L^2(P)$, то при $s < t$ ортогональная проекция ξ_t на подпространство $\mathcal{F}_s$-измеримых функций есть ξ_s, т.е. разность $\xi_t - \xi_s$ ортогональна этому подпространству, откуда

$$\mathrm{E}(\xi_t \xi_s) = \mathrm{E}(\xi_s \xi_s).$$

Поэтому приращения ортогональны в $L^2(P)$: при $t_1 < t_2 \leqslant t_3 < t_4$ имеем

$$\mathrm{E}[(\xi_{t_4} - \xi_{t_3})(\xi_{t_2} - \xi_{t_1})] = 0.$$

С другой стороны, верен такой факт.

11.3. Теорема. *Если $\{\xi_t\}_{t \geqslant 0}$ — процесс с независимыми приращениями, причем $\mathrm{E}\xi_t$ — постоянная, то это мартингал относительно порожденной им фильтрации.*

Доказательство. При $s < t$ имеем

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

При этом $\mathrm{E}(\xi_s | \mathcal{F}_s) = \xi_s$, $\mathrm{E}(\xi_t - \xi_s | \mathcal{F}_s) = \mathrm{E}(\xi_t - \xi_s) = 0$ в силу задачи 11, ибо $\xi_t - \xi_s$ и ξ_τ независимы при $\tau \leqslant s$ (в силу независимости $\xi_t - \xi_s$ и $\xi_\tau - \xi_0$ и постоянства ξ_0), значит, случайная величина $\xi_t - \xi_s$ является независимой с $\mathcal{F}_s$. $\square$

Необходимость постоянства $\mathrm{E}\xi_t$ для мартингальности отмечалась выше.

Из доказанного следует, что винеровский процесс — мартингал, а пуассоновский нет.

Следующий пример охватывает очень широкий круг мартингалов, встречающихся в приложениях.

11.4. Пример. Пусть даны фильтрация $\{\mathcal{F}_t\}_{t \geqslant 0}$ и интегрируемая случайная величина ξ. Тогда процесс $\xi_t = \mathrm{E}(\xi | \mathcal{F}_t)$ — мартингал. Это следует из равенства $\mathrm{E}^{\mathcal{F}_s} \mathrm{E}^{\mathcal{F}_t} \xi = \mathrm{E}^{\mathcal{F}_s} \xi$ при $s < t$.

Задача 12. Пусть $\{\xi_n\}_{n \in \mathbb{N}}$ — мартингал относительно $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ и $\mathrm{E}|\xi_n|^2 \leqslant C$. Доказать, что найдется функция $\xi \in L^2(P)$, для которой $\xi_n = \mathrm{E}^{\mathcal{F}_n} \xi$.

Указанный пример мартингала примечателен следующим.

11.5. Теорема. (Теорема Дуба о сходимости мартингалов) *Пусть $\{\xi_n\}_{n \in \mathbb{N}}$ — мартингал относительно $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$. Этот мартингал сходится в $L^1(P)$ в точности тогда, когда есть такая функция $\xi \in L^1(P)$, что $\xi_n = \mathrm{E}^{\mathcal{F}_n} \xi$. При этом есть и сходимость почти всюду.*

Равносильным условием является равномерная интегрируемость ξ_n, т.е. условие

$$\lim_{R\to+\infty}\sup_n\int_{\{|\xi_n|\geqslant R\}}|\xi_n|\,dP=0.$$

Задача 13. Пусть $\{\xi_t\}_{t\geqslant 0}$ — процесс с независимыми приращениями, $\xi_t\in L^2(P)$, причем $\xi_0=0$, $\mathrm{E}\xi_t=0$. Доказать, что $\xi_t^2-\mathrm{E}(\xi_t^2)$ — мартингал относительно фильтрации, порожденной $\{\xi_t\}_{t\geqslant 0}$. В частности, если w_t — винеровский процесс, то w_t^2-t — мартингал.

Если в определении мартингала равенство заменить неравенством $\mathrm{E}(\xi_t|\mathcal{F}_s)\leqslant\xi_s$, то получится определение *супермартингала*, а противоположное неравенство $\mathrm{E}(\xi_t|\mathcal{F}_s)\geqslant\xi_s$ даст определение *субмартингала*.

Из неравенства Йенсена для условных ожиданий получаем, что если ξ_t — мартингал и V — выпуклая функция, для которой $V(\xi_t)$ интегрируемы, то $V(\xi_t)$ — субмартингал. Например, взяв винеровский процесс w_t, получаем субмартингалы $|w_t|$, w_t^2, $\exp w_t$.

Задача 14. Доказать, что пуассоновский процесс — субмартингал.

§ 12. Предсказуемые процессы и разложение Дуба – Мейера

12.1. Определение. *Случайный процесс $\{\xi_n\}_{n\in\mathbb{Z}_+}$ называется предсказуемым относительно фильтрации $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$, если ξ_n оказывается $\mathcal{F}_{n-1}$-измеримым для всех n.*

Это свойство сильнее согласованности.

12.2. Пример. Мартингал $\{\xi_n\}_{n\in\mathbb{Z}_+}$ предсказуем только в случае $\xi_n=\xi_0$ п.в. В частности, если $\xi_0=0$ п.в., то $\xi_n=0$ п.в.

В самом деле, в силу $\mathcal{F}_{n-1}$-измеримости ξ_n имеем

$$\xi_n=\mathrm{E}^{\mathcal{F}_{n-1}}\xi_n=\xi_{n-1}.$$

Значит, $\xi_n=\xi_{n-1}=\cdots=\xi_0$.

Следующий результат показывает, что согласованные процессы с натуральным временем — суммы мартингалов и предсказуемых процессов.

12.3. Теорема. (Разложение Дуба – Мейера) *Пусть процесс $\{\xi_n\}_{n\in\mathbb{Z}_+}$ согласован с фильтрацией $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$, причем все ξ_n интегрируемы. Тогда существует единственное с точностью до эквивалентности разложение*

$$\xi_n=\mu_n+\pi_n,$$

где $\{\mu_n\}$ — мартингал относительно данной фильтрации, а процесс $\{\pi_n\}$ предсказуем, $\pi_0=0$.

Доказательство. Положим $\pi_0 = 0$,
$$\pi_n = \sum_{k=1}^{n} \mathrm{E}(\xi_k - \xi_{k-1} | \mathcal{F}_{k-1}).$$
Предсказуемость очевидна. Покажем, что $\mu_n = \xi_n - \pi_n$ — мартингал. Согласованность и интегрируемость есть. Достаточно проверить, что $\mathrm{E}(\mu_n | \mathcal{F}_{n-1}) = \mu_{n-1}$. Мы имеем:
$$\mathrm{E}(\mu_n|\mathcal{F}_{n-1}) = \mathrm{E}(\xi_n|\mathcal{F}_{n-1}) - \pi_n =$$
$$= \mathrm{E}(\xi_{n-1}|\mathcal{F}_{n-1}) - \sum_{k=1}^{n-1} \mathrm{E}(\xi_k - \xi_{k-1}|\mathcal{F}_{k-1}) = \xi_{n-1} - \pi_{n-1} = \mu_{n-1},$$
что и требовалось.

Пусть $\xi_n = \mu'_n + \pi'_n$ — еще одно разложение с указанными свойствами. Надо проверить, что $\mu_n = \mu'_n$ п.в. Остается воспользоваться примером выше, который говорит, что предсказуемый мартингал, стартующий из нуля, равен нулю, ибо разность мартингалов — мартингал, разность предсказуемых процессов предсказуема. □

§ 13. Марковские моменты и теорема об остановке

13.1. Определение. *Пусть $T \subset \mathbb{R}$. Случайная величина τ со значениями в $T \cup \{+\infty\}$ называется марковским моментом относительно фильтрации $\{\mathcal{F}_t\}_{t\in T}$, если имеет место включение $\{\omega\colon \tau(\omega) \leqslant t\} \in \mathcal{F}_t$ для всех t. Если $\tau < +\infty$ п.в., то τ называется моментом остановки.*

13.2. Пример. Для процесса $\{\xi_n\}_{n\in\mathbb{N}}$ и порожденной им фильтрации для фиксированного борелевского множества $B \subset \mathbb{R}$ положим
$$\tau_B(\omega) = \min\{n\colon \xi_n(\omega) \in B\}, \text{если такие } n \text{ есть, иначе } \tau_B(\omega) = +\infty.$$
Тогда τ_B — марковский момент, ибо для всякого $t \in \mathbb{N}$ множество $\{\tau_B \leqslant t\}$ равно объединению множеств $\{\omega\colon \xi_n(\omega) \in B\}$, $n = 1, \ldots, t$, входящих в $\mathcal{F}_{\leqslant t}$.

Если $\mathcal{F}_t$ интерпретировать как класс событий, о наступлении или ненаступлении которых известно к моменту t, то марковость можно понимать так: если при каком-то ω марковский момент принял значение t, то к моменту t это уже известно. В случае фильтрации, порожденной случайным процессом $\{\xi_t\}$, наглядный смысл таков: по наблюдениям значений процесса до момента t включительно можно сказать, попало ли τ в множество B до момента времени t.

Следующая теорема об остановке говорит о неизменности матожидания мартингала в ограниченный момент остановки.

13.3. Теорема. *Пусть $\{\xi_n\}_{n\in\mathbb{Z}_+}$ — мартингал относительно фильтрации $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$ и τ — ограниченный момент остановки. Тогда имеем $\mathrm{E}\xi_\tau = \mathrm{E}\xi_0$. В частности, если $\mathrm{E}\xi_0 = 0$, то $\mathrm{E}\xi_\tau = 0$.*

Доказательство. Пусть $\tau \leqslant m$. Тогда $|\xi_\tau| \leqslant |\xi_0| + \cdots + |\xi_m|$, поэтому $\mathrm{E}\xi_\tau$ существует. Для фиксированного $j \leqslant m$ имеем

$$\mathrm{E}(\xi_\tau I_{\{\tau=j\}}) = \mathrm{E}(\xi_j I_{\{\tau=j\}}) = \mathrm{E}(\xi_{j+1} I_{\{\tau=j\}}) = \cdots = \mathrm{E}(\xi_m I_{\{\tau=j\}}),$$

ибо $\{\tau = j\} \in \mathcal{F}_j$. Сумма по j от 0 до m дает слева $\mathrm{E}\xi_\tau$, а справа $\mathrm{E}\xi_m = \mathrm{E}\xi_0$. □

13.4. Пример. Пусть ξ_n — независимые случайные величины, причем $P(\xi_n = 1) = P(\xi_n = -1) = 1/2$, $S_0 = 0$, $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$. Последовательность S_n — мартингал относительно фильтрации $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$, который можно трактовать как суммарный выигрыш (возможно, отрицательный) при бросании монеты. Математическое ожидание выигрыша нулевое, причем так будет и при фиксации выигрыша в ограниченный момент остановки. Результат меняется, если выигрыш фиксировать в неограниченный момент остановки. Например, пусть τ — первый момент достижения суммой значения 2. Тогда $S_\tau = 2$ п.в., значит, $\mathrm{E}S_\tau = 2$. В качестве ограниченного момента остановки σ здесь можно взять первый момент достижения 2 до момента времени 100, положив $\sigma = \min(\tau, 100)$.

Предположим теперь, что выигрыш можно определять с помощью некоторой стратегии по формуле $X_n = X_{n-1} + \xi_n f_n(X_1, \ldots, X_{n-1})$, $X_0 = 0$, выбрав какие-то функции f_n. Тогда $\{X_n\}$ по-прежнему остается мартингалом и $\mathrm{E}X_\tau = 0$ в ограниченные моменты остановки.

Теорема об остановке переносится на непрерывное время в такой формулировке.

Пусть $\{\mathcal{F}_t\}_{t\geqslant 0}$ — фильтрация. Отображение $\tau\colon \Omega \to [0, +\infty]$ называется опциональным моментом, если $\{\omega\colon \tau(\omega) < t\} \subset \mathcal{F}_t$ для всех $t \geqslant 0$.

Марковский момент τ является опциональным, поскольку имеем $\{\tau < t\} = \bigcup_{n=1}\{\tau \leqslant t - 1/n\} \in \mathcal{F}_t$, ибо $\{\tau \leqslant t - 1/n\} \in \mathcal{F}_{t-1/n} \subset \mathcal{F}_t$.

13.5. Пример. Пусть процесс $\{\xi_t\}_{t\geqslant 0}$ согласован с фильтрацией $\{\mathcal{F}_t\}_{t\geqslant 0}$ и имеет непрерывные справа траектории. Тогда для открытого множества U момент $\tau_U = \inf\{t\colon \xi_t \in U\}$ является опциональным. Действительно, $\tau_U < t$ в точности тогда, когда есть $s < t$ с $\xi_s \in U$, причем из-за непрерывности справа такое s можно взять рациональным. Поэтому $\tau_U < t$ есть счетное объединение множеств $\{\xi_s \in U\} \in \mathcal{F}_s$ по рациональным $s < t$.

13.6. Теорема. *Пусть* $\{\xi_t\}_{t\geqslant 0}$ — *мартингал относительно фильтрации* $\{\mathcal{F}_t\}_{t\geqslant 0}$, *имеющий непрерывные справа траектории. Тогда для всякого ограниченного опционального момента* τ *имеем* $\mathrm{E}\xi_\tau = \mathrm{E}\xi_0$.

§ 14. Задача о разорении игрока: мартингальный подход к решению

Пусть ξ_n, где $n \in \mathbb{N}$, — независимые случайные величины, причем $P(\xi_n = 1) = p$, $P(\xi_n = -1) = q = 1 - p$.

Пусть $a < x < b$ — целые числа, $S_n = x + \xi_1 + \cdots + \xi_n$, $S_0 = x$.

Суммы S_n интерпретируются как суммарные выигрыши в моменты времени n с начальным капиталом x.

Момент окончания игры $\tau = \min\{n\colon S_n \in \{a, b\}\}$, вероятность разорения есть $P(S_\tau = a)$. Это вероятность падения капитала до критического уровня a ранее достижения уровня b. Предельный уровень b можно интерпретировать как предельный капитал второго игрока или банка, если игра идет с банком (при достижении b банк закрывается).

Ясно, что τ — марковский момент, но почему он конечен п.в. (т.е. является моментом остановки)? Если $\tau(\omega) = +\infty$, то $S_n(\omega)$ никогда не попадает в a и b, но вероятность этого нулевая. Это видно из того, что если $S_n(\omega) \neq a, b$ при всех n, то $a < S_n(\omega) < b$, ибо длина шага равна 1. Можно непосредственно проверить, что вероятность постоянного нахождения в (a, b) нулевая, но без всякой проверки это ясно из центральной предельной теоремы, ведь иначе с положительной вероятностью мы бы имели $S_n/\sqrt{n} \to 0$. Итак, τ — момент остановки. Поэтому в a или b сумма обязательно попадет, но надо знать вероятность попадания сначала в a.

В случае $p = q = 1/2$ последовательность S_n — мартингал относительно фильтрации порожденной $\{\xi_n\}_{n\geqslant 1}$. Мы хотели бы сказать, что $\mathrm{E}S_\tau = \mathrm{E}S_0 = x$, но момент τ не является ограниченным, так что теорема об остановке не работает. Эта теорема применима к ограниченным моментам $\tau_k = \min(\tau, k)$, которые остаются марковскими. Для них $\mathrm{E}S_{\tau_k} = x$. Так как τ_k возрастает к τ, то $S_{\tau_k} \to S_\tau$ п.в. Заметим, что $|S_{\tau_k}| \leqslant |a| + |b|$, так как $S_\tau \in \{a, b\}$, $S_n \in (a, b)$, если $n < \tau$. По теореме Лебега $\mathrm{E}S_{\tau_k} \to \mathrm{E}S_\tau$. Итак, $\mathrm{E}S_\tau = x$. При этом $\mathrm{E}S_\tau = aP(S_\tau = a) + bP(S_\tau = b)$, откуда

$$P(S_\tau = a) = \frac{b - x}{b - a}.$$

Случай $p \neq q$. В этом случае мартингалом оказывается последовательность $\Pi_n = (q/p)^{S_n}$. В самом деле: в силу независимости ξ_n

имеем

$$\mathrm{E}^{\mathcal{F}_n}\Pi_{n+1} = \Pi_n \mathrm{E}^{\mathcal{F}_n}(q/p)^{\xi_{n+1}} = \Pi_n \mathrm{E}(q/p)^{\xi_{n+1}} = \Pi_n,$$

так как $\mathrm{E}(q/p)^{\xi_{n+1}} = p(q/p) + q(q/p)^{-1} = 1$. Как и выше, с учетом оценки $|X_{\min(n,\tau)}| \leqslant (q/p)^a + (q/p)^b$ получаем равенство $\mathrm{E}\Pi_\tau = \mathrm{E}\Pi_0 = (q/p)^x$, откуда находим

$$P(S_\tau = a) = \frac{(q/p)^b - (q/p)^x}{(q/p)^b - (q/p)^a}.$$

§ 15. Марковские процессы

До сих пор речь шла о случайных процессах со значениями в $\mathbb{R}$, но при обсуждении марковских процессов и марковских цепей удобнее говорить о процессах с более общим пространством состояний X, наделенном σ-алгеброй $\mathcal{X}$. Тогда процесс состоит из измеримых отображений $\xi_t\colon (\Omega, \mathcal{B}) \to (X, \mathcal{X})$. Хотя пространства состояний из реальных задач борелевски изоморфны прямой, они обычно обладают какой-то естественной структурой, которую невыгодно терять. Например, при рассмотрении случайного блуждания по целочисленной решетке на плоскости было бы неудобно отождествлять ее с $\mathbb{Z}$.

Пусть $T \subset \mathbb{R}$. Как и ранее, для всякого процесса $\{\xi_t\}_{t\in T}$ со значениями в X появляются порожденные им σ-алгебры $\mathcal{F}_{=t} = \sigma(\xi_t)$, а также более широкие σ-алгебры $\mathcal{F}_{\leqslant t} = \sigma(\xi_s\colon s \leqslant t)$ и $\mathcal{F}_{\geqslant t} = \sigma(\xi_s\colon s \geqslant t)$, интерпретируемые как «прошлое» и «будущее».

Далее положим

$$P(A|\mathcal{F}) := \mathrm{E}(I_A|\mathcal{F}).$$

15.1. Определение. *Процесс $\{\xi_t\}_{t\in T}$ называется марковским, если для всякого $t \in T$ всех $A \in \mathcal{F}_{\leqslant t}$ и $B \in \mathcal{F}_{\geqslant t}$ почти наверное*

$$P(A \cap B|\mathcal{F}_{=t}) = P(A|\mathcal{F}_{=t})P(B|\mathcal{F}_{=t}). \tag{15.1}$$

Как обычно, тождество для множеств равносильно тождеству

$$\mathrm{E}(fg|\mathcal{F}_{=t}) = \mathrm{E}(f|\mathcal{F}_{=t})\mathrm{E}(g|\mathcal{F}_{=t})$$

для ограниченных $\mathcal{F}_{t\leqslant}$-измеримых функций f и ограниченных $\mathcal{F}_{t\geqslant}$-измеримых функций g.

Смысл равенства — условная независимость будущего и прошлого при фиксированном настоящем. Для таких процессов будущее зависит от прошлого только через настоящее. Приведем еще одно равносильное условие. Как обычно, равенство условных средних есть равенство почти всюду.

15.2. Лемма. *Тождество* (15.1) *равносильно тождеству*
$$P(B|\mathcal{F}_{\leqslant t}) = P(B|\mathcal{F}_{=t}) \quad \forall\, B \in \mathcal{F}_{\geqslant t}. \tag{15.2}$$
Иначе говоря,
$$\mathrm{E}(f|\mathcal{F}_{\leqslant t}) = \mathrm{E}(f|\mathcal{F}_{=t}) \tag{15.3}$$
для всякой ограниченной $\mathcal{F}_{\geqslant t}$-измеримой функции f.

Доказательство. Если верно (15.1), то для $A \in \mathcal{F}_{\leqslant t}$, $\in \mathcal{F}_{\geqslant t}$ имеем
$$\mathrm{E}(I_A I_B) = \mathrm{E}(I_{A\cap B}) = \mathrm{E}(P(A\cap B|\mathcal{F}_{=t})) = \mathrm{E}(P(A|\mathcal{F}_{=t})P(B|\mathcal{F}_{=t})),$$
что равно $\mathrm{E}(I_A P(B|\mathcal{F}_{=t}))$ из-за $\mathcal{F}_{=t}$-измеримости $P(B|\mathcal{F}_{=t})$. Это дает равенство (15.2).

Пусть выполнено (15.2). Надо проверить, что для $A \in \mathcal{F}_{\leqslant t}$ левая и правая части (15.1) дают равные интегралы по множествам $C \in \mathcal{F}_{=t}$. Интеграл левой части равен $P(A\cap B\cap C)$, ибо $I_C P(A\cap B|\mathcal{F}_{=t}) = P(A\cap B\cap C|\mathcal{F}_{=t})$ в силу включения $C \in \mathcal{F}_{=t}$. Интеграл от правой части равен (также с учетом включения $C \in \mathcal{F}_{=t}$)
$$\int I_C P(A|\mathcal{F}_{=t})P(B|\mathcal{F}_{=t})\,dP = \int P(A|\mathcal{F}_{=t})P(C\cap B|\mathcal{F}_{=t})\,dP$$
$$= \int I_A P(C\cap B|\mathcal{F}_{=t})\,dP = \int I_A P(C\cap B|\mathcal{F}_{\leqslant t})\,dP.$$
Так как $A \in \mathcal{F}_{\leqslant t}$, то это тоже равно $P(A\cap B\cap C)$. □

В основном определении и равносильном описании из предыдущей леммы использовались множества из исходного вероятностного пространства или функции на нем. В следующей характеризации марковости рассматриваются функции на фазовом пространстве X (которое никак не фигурировало в определении) или множества из X.

15.3. Теорема. *Случайный процесс ξ_t со значениями в X является марковским в точности тогда, когда для всякой ограниченной $\mathcal{X}$-измеримой функции g на X при всех $s,t \in T$ с $s \leqslant t$ верно равенство*
$$\mathrm{E}(g(\xi_t)|\mathcal{F}_{\leqslant s}) = \mathrm{E}(g(\xi_t)|\mathcal{F}_{=s}). \tag{15.4}$$
Равносильное условие $P(\xi_t \in E|\mathcal{F}_{\leqslant s}) = P(\xi_t \in E|\mathcal{F}_{=s})\ \ \forall\, E \in \mathcal{X}$.

Доказательство. Если верно (15.3), то верно и (15.4), ибо функция $f = g(\xi_t)$ измерима относительно $\mathcal{F}_{\geqslant s}$ при $t \geqslant s$.

Пусть верно (15.4). Покажем, что верно и (15.3). Сначала рассмотрим функции f вида $f = f_1(\xi_{t_1})\cdots f_n(\xi_{t_n})$, где $t \leqslant t_1 < t_2 < \cdots < t_n$

и функции f_i на X измеримы относительно $\mathcal{X}$. Проверим (15.3) индукцией по n. При $n = 1$ это есть (15.4). Пусть равенство верно при некотором n, $t_n < t_{n+1}$, функция f_{n+1} ограничена и $\mathcal{X}$-измерима. Тогда

$$\begin{aligned} &\mathrm{E}\big((f_1(\xi_{t_1})\cdots f_{n+1}(\xi_{t_{n+1}})|\mathcal{F}_{\leqslant t}\big) \\ &\qquad = \mathrm{E}\Big(\mathrm{E}\big(f_1(\xi_{t_1})\cdots f_{n+1}(\xi_{t_{n+1}})|\mathcal{F}_{\leqslant t_n}\big)\Big|\mathcal{F}_{\leqslant t}\Big) \\ &\qquad\quad = \mathrm{E}\Big(f_1(\xi_{t_1})\cdots f_n(\xi_{t_n})\mathrm{E}\big(f_{n+1}(\xi_{t_{n+1}})|\mathcal{F}_{\leqslant t_n}\big)\Big|\mathcal{F}_{\leqslant t}\Big), \end{aligned}$$

ибо функции $f_j(\xi_{t_j})$ измеримы относительно $\mathcal{F}_{\leqslant t_n}$ при $j \leqslant n$. В силу (15.4) мы имеем

$$\mathrm{E}\big(f_{n+1}(\xi_{t_{n+1}})|\mathcal{F}_{\leqslant t_n}\big) = \mathrm{E}\big(f_{n+1}(\xi_{t_{n+1}})|\mathcal{F}_{=t_n}\big),$$

причем функция справа имеет вид $\varphi(\xi_{t_n})$, где функция $\varphi\colon X \to \mathbb{R}$ измерима относительно $\mathcal{X}$, как показано в лемме ниже. По предположению индукции

$$\mathrm{E}\big(f_1(\xi_{t_1})\cdots f_n(\xi_{t_n})\varphi(\xi_{t_n})|\mathcal{F}_{\leqslant t}\big) = \mathrm{E}\big(f_1(\xi_{t_1})\cdots f_n(\xi_{t_n})\varphi(\xi_{t_n})|\mathcal{F}_{=t}\big).$$

С этой же функцией совпадает $\mathrm{E}\big(f_1(\xi_{t_1})\cdots f_{n+1}(\xi_{t_{n+1}})|\mathcal{F}_{=t}\big)$, так как $\mathcal{F}_{=t} \subset \mathcal{F}_{\leqslant t_n}$ и потому

$$\begin{aligned} &\mathrm{E}\big(f_1(\xi_{t_1})\cdots f_{n+1}(\xi_{t_{n+1}})|\mathcal{F}_{=t}\big) \\ &\qquad = \mathrm{E}\Big(\mathrm{E}\big(f_1(\xi_{t_1})\cdots f_{n+1}(\xi_{t_{n+1}})|\mathcal{F}_{\leqslant t_n}\big)\Big|\mathcal{F}_{=t}\Big) \\ &\qquad\quad = \mathrm{E}\Big(f_1(\xi_{t_1})\cdots f_n(\xi_{t_n})\mathrm{E}\big(f_{n+1}(\xi_{t_{n+1}})|\mathcal{F}_{\leqslant t_n}\big)\Big|\mathcal{F}_{=t}\Big). \end{aligned}$$

Итак, равенство (15.4) верно для функций указанного вида, а тогда и для их линейных комбинаций.

Наконец, для всякой ограниченной $\mathcal{F}_{\geqslant t}$-измеримой функции f найдется последовательность таких линейных комбинаций f_j, сходящаяся к f в $L^2(P)$. Соответствующие условные средние тоже сходятся в $L^2(P)$, поэтому равенство (15.4) остается в силе и для f. □

Выше был использован такой факт.

15.4. Лемма. *Пусть $\mathcal{A}$ — σ-алгебра, порожденная отображением $\xi\colon \Omega \to X$, т.е. класс множеств вида $\xi^{-1}(E)$, $E \in \mathcal{X}$. Тогда всякая $\mathcal{A}$-измеримая функция $f\colon \Omega \to \mathbb{R}$ имеет вид $\varphi\circ\xi$, где функция $\varphi\colon X \to \mathbb{R}$ измерима относительно $\mathcal{X}$.*

ДОКАЗАТЕЛЬСТВО. Достаточно проверить это для ограниченных функций. Для индикаторов множеств из $\mathcal{A}$ утверждение верно по определению. Значит, оно верно для простых функций. Для общей ограниченной $\mathcal{A}$-измеримой функции f найдется равномерно сходящаяся к ней последовательность простых $\mathcal{A}$-измеримых функций f_n. Тогда $f_n = \varphi_n \circ \xi$, где функции φ_n на X ограничены и $\mathcal{X}$-измеримы. Эти функции можно взять равномерно ограниченными. Множество X_0 всех точек x, где существует $\varphi(x) = \lim\limits_{n\to\infty} \varphi_n(x)$, входит в $\mathcal{X}$. Вне него доопределим функцию φ нулем. Это дает $\mathcal{X}$-измеримую функцию. При этом $\varphi(\xi(\omega) = \lim\limits_{n\to\infty} \varphi_n(\xi(\omega)) = \lim\limits_{n\to\infty} f_n(\xi(\omega)) = f(\xi(\omega)$ для всех ω. $\square$

Марковский процесс с целочисленным временем называют *цепью Маркова*. Если при этом еще и S счетно или конечно, то такой процесс называют *дискретной цепью Маркова*, а в случае конечного S — *конечной цепью Маркова*.

При широких условиях марковский процесс обладает переходной функцией, удовлетворяющей так называемому уравнению Чэпмена – Колмогорова. А именно справедливо следующее утверждение.

15.5. ТЕОРЕМА. *Пусть S — борелевское множество в $\mathbb{R}^n$ с борелевской σ-алгеброй $\mathcal{S}$ (или измеримое пространство, изоморфное такому множеству) и $\{\xi_t\}_{t\in T}$ — марковский процесс с пространством состояний S, причем T — какое-либо из множеств $[0,+\infty)$, $(-\infty,+\infty)$, $\mathbb{N}$, $\mathbb{Z}$. Тогда этот процесс обладает переходной функцией, т.е. имеется функция $P(s,x,t,E)$ четырех аргументов, где $s,t \in T$, $x \in S$, $E \in \mathcal{S}$, для которой*

(i) *при фиксированных s,t,x функция $E \mapsto P(s,x,t,E)$ является вероятностной мерой на $\mathcal{S}$,*

(ii) *при фиксированных s,t,E функция $x \mapsto P(s,x,t,E)$ измерима относительно $\mathcal{S}$,*

(iii) *$E \mapsto P(s,x,s,E) = \delta_x(E)$ есть мера Дирака в точке x,*

(iv) *выполнено уравнение Чэпмена – Колмогорова*

$$P(s,x,t,E) = \int_S P(u,y,t,E)P(s,x,u,dy), \quad s \leqslant u \leqslant t, \qquad (15.5)$$

(v) *при фиксированных s,t,E с $s \leqslant t$ почти наверное*

$$P(s,\xi_s,t,E) = P(\xi_t \in E|\mathcal{F}_{=s}).$$

Переходная функция дает техническую интерпретацию «вероятности процесса попасть в множество E в момент t при нахождении в точке x в момент времени s».

Из леммы вытекает равенство

$$P(s, \xi_s, t, E) = P(\xi_t \in E | \mathcal{F}_{\leqslant s}).$$

Условие (v) можно записать в виде

$$P(s, x, t, E) = P(\xi_t \in E | \xi_s = x).$$

В случае счетного множества состояний и $P(\xi_s = x) > 0$ правая часть понимается обычным образом, а в общем случае она понимается как $\theta(x)$, где θ — борелевская функция, которая представляет условную вероятность в виде $P(\xi_t \in E | \sigma(\xi_s)) = \varphi(\xi_s)$.

Смысл равенства в (15.5) том, что попадание в E в момент t при нахождении в x в момент s слагается из переходов в различные y в промежуточный момент u и последующих переходов из y в E.

Если переходные функции $P(s, x, t, E)$ зависят только от разности $t - s$, то процесс называют *однородным*. Тогда можно считать, что задана функция $P(x, t, E)$ трех аргументов, интерпретируемая как «вероятность попасть в множество E в момент времени t при нахождении процесса в x при $t = 0$.»

Конечномерные распределения марковского процесса выражаются через переходную функцию и распределения μ_t случайных элементов ξ_t по формуле

$$P((\xi_{t_1}, \dots, \xi_{t_n}) \in C) = \int_X \mu_{t_1}(dx_1) \int_X P(t_1, x_1, t_2, dx_2) \cdots \\ \cdots \int_X I_C(x_1, \dots, x_n)\, P(t_{n-1}, x_{n-1}, t_n, dx_n), \quad (15.6)$$

где порядок интегрирования — с конца (сначала по x_n, затем по x_{n-1} и т.д.).

С помощью теоремы Колмогорова о согласованных распределениях можно проверить, что для заданных переходной функции и начального распределения существует марковский процесс. Однако более эффективный способ задания марковских процессов связан с задаваемыми ими полугруппами. Обсуждение этого вопроса выходит за рамки курса. Ниже с помощью теоремы Колмогорова будет проверено лишь существование однородной марковской цепи с конечным числом состояний.

Как связаны марковские процессы с двумя другими нашими основными классами процессов: процессами с независимыми приращениями и мартингалами?

15.6. Теорема. *Процесс с независимыми приращениями является марковским.*

В частности, винеровский и пуассоновский процессы — марковские.

Задача 15. Проверить непосредственно марковость винеровского и пуассоновского процессов, вычислив $\mathrm{E}(f(w_t)|\mathcal{F}_{=s})$, $s < t$, а также аналогичное условное среднее для N_t.

Задача 16. Пусть w_t — винеровский процесс. Верно ли, что $|w_t|$ — марковский процесс?

С мартингалами соотношение более сложное. Вещественный марковский процесс может не быть мартингалом уже из-за неинтегрируемости, но и интегрируемость не поможет.

Задача 17. Если вещественный случайный процесс ξ_t является марковским, то таков и процесс $f(\xi_t)$, где f — гомеоморфизм прямой, в частности, процесс ξ_t^3. При этом для винеровского процесса w_t процесс w_t^3 не является мартингалом. В случае марковского процесса ξ_t в фазовом пространстве $(X, \mathcal{X})$ и измеримого изоморфизма $f\colon X \to X$ процесс $f(\xi_t)$ — марковский.

Задача 18. Привести пример мартингала, не являющегося марковским процессом.

Задача 19. Пусть $\{\xi_t\}$ — марковский процесс со значениями в $\mathbb{R}$, причем случайные величины ξ_t интегрируемы. Доказать, что этот процесс — мартингал относительно порожденной им фильтрации $\{\mathcal{F}_{\leqslant t}\}$ в точности тогда, когда $\mathrm{E}(\xi_t|\mathcal{F}_{=s}) = \xi_s$ при $s \leqslant t$.

§ 16. Цепи Маркова

Обратимся теперь к цепям Маркова, т. е. марковским процессам с дискретным временем $T = \mathbb{N}$ и счетным числом состояний $X = \{x_j\}$, причем будем рассматривать однородные процессы. Для них переходные функции определяются вероятностями перехода $p_{i,j}$ из x_i в x_j за один шаг, т.е.

$$p_{i,j} \geqslant 0, \quad \sum_{j=1}^{\infty} p_{i,j} = 1, \quad i \in \mathbb{N}.$$

Если дано начальное распределение процесса π_0, а именно мера на X с $\pi_0(x_j) \geqslant 0$ и $\sum_j \pi_0(x_j) = 1$, то вероятность нахождения процесса в x_j в момент $t = 1$ равна

$$\sum_i p_{i,j}\pi_0(x_i),$$

так как в x_j можно попасть из каждого x_i с вероятностью $p_{i,j}$, а вероятность нахождения в x_i была $\pi_0(x_i)$. Если распределение π_0 отождествить с вектор-строкой $(\pi_0(1), \pi_0(2), \ldots)$, а распределение цепи π_1 в момент 1 с вектор-строкой $(\pi_1(1), \pi_1(2), \ldots)$, то получим соотношение

$$\pi_1 = \pi_0 P, \quad P = (p_{i,j})_{i,j\geqslant 1}$$

где на матрицу умножаем справа. Поэтому для распределения π_n процесса в момент времени n получаем

$$\pi_n = \pi_{n-1}P = \pi_0 P^n.$$

Если распределения отождествлять с вектор-столбцами, то надо использовать транспонированную матрицу P^* и умножать на нее слева. Изобретатель марковских цепей А. А. Марков применил их к исследованию чередования гласных и согласных в словах русского языка.

Хотя интуитивно не вызывает сомнений существование марковских цепей, проверим это, так как в самом описании цепи фигурируют лишь переходные вероятности и начальное распределение.

16.1. Теорема. *Пусть даны матрица переходов $P = (p_{i,j})_{i,j\leqslant m}$ и начальное распределение π_0 на $X = \{x_1, \ldots, x_m\}$. Тогда существует цепь Маркова $\{\xi_n\}$ с фазовым пространством X, для которой π_0 является распределением ξ_0 и $p_{i,j} = \mathrm{E}(\xi_{n+1} = x_j | \xi_n = x_i)$.*

Доказательство. Проверка легко проводится с помощью теоремы Колмогорова, которая была у нас для вещественных процессов, но в данном случае можно считать, что $X \subset \mathbb{R}$, скажем, считать, что $x_i = i$, и пользоваться этой формулировкой. Нам нужно предъявить согласованные конечномерные распределения на пространствах $\mathbb{R}^n$, фактически вместо всего пространства $\mathbb{R}^n$ достаточно рассматривать конечное множество $\{1, \ldots, m\}^n$. Мера $P_{1,\ldots,n}$ на его подмножествах задается такой формулой на отдельных точках:

$$P_{1,\ldots,n}(i_1, \ldots, i_n) = \pi_0(i_1)p_{i_1,i_2} \cdots p_{i_{n-1},i_n}, \tag{16.1}$$

причем в качестве P_0 берется распределение π_0. Затем эта мера распространяется на все конечные множества. Меры $P_{k_1,\ldots,k_l}$ при $k_i \leqslant n$ задаются как проекции меры $P_{1,\ldots,n}$ на соответствующие $\mathbb{R}^l$. Например,

$$P_{2,3}(i,j) = \sum_k P_{1,2,3}(k,i,j) = \sum_k \pi_0(k)p_{k,i}p_{i,j},$$

$$P_2(i) = \sum_k P_{1,2}(k,i) = \sum_k \pi_0(k)p_{k,i}.$$

Ясно, что проекция $P_{1,\ldots,n+1}$ на $\mathbb{R}^n$ равна $P_{1,\ldots,n}$, ибо $\sum_j p_{i,j} = 1$. Из этого вытекает согласованность заданной системы мер. Все эти равенства — частные случаи формулы (15.6) для марковских процессов, с помощью которой можно строить процесс в общем случае. Теорема Колмогорова дает существование процесса $\{\xi_n\}$ с такими конечномерными распределениями, но нам нужна еще и марковость, а также нужно равенство $P(\xi_n = j|\xi_{n-1} = i) = p_{i,j}$. Для этого достаточно проверить, что

$$P(\xi_m = i, \xi_n = j|\xi_t = l) = P(\xi_m = i|\xi_t = l)P(\xi_n = j|\xi_t = l)$$

при всех $m < t < n$, а также

$$P(\xi_n = j, \xi_{n-1} = i) = p_{i,j}P(\xi_{n-1} = i).$$

Последнее верно, так как по определению $P(\xi_n = j, \xi_{n-1} = i)$ есть сумма произведений $\pi_0(i_1)p_{i_1,i_2}\cdots p_{i_{n-2},i}p_{i,j}$ по всем $i_1, \ldots, i_{n-2}$, причем $P(\xi_{n-1} = i)$ есть сумма таких же произведений без множителя $p_{i,j}$. Из доказанного следует, что

$$P(\xi_{k+n} = j|\xi_k = i) = p_{i,j}^{(n)},$$

где $p_{i,j}^{(n)}$ — элементы матрицы P^n.

Таким образом, нужно проверить равенство

$$P(\xi_m = i, \xi_t = l, \xi_n = j)P(\xi_t = l) = P(\xi_m = i, \xi_t = l)P(\xi_n = j, \xi_t = l).$$

Из предыдущей формулы получаем

$$P(\xi_n = j, \xi_t = l) = P(\xi_n = j|\xi_t = l)P(\xi_t = l) = p_{l,j}^{(n-t)}P(\xi_t = l).$$

Значит, надо показать, что

$$P(\xi_m = i, \xi_t = l, \xi_n = j) = P(\xi_m = i, \xi_t = l)p_{l,j}^{(n-t)}.$$

При $n = t + 1$ это равенство имеет вид

$$P(\xi_m = i, \xi_t = l, \xi_{t+1} = j) = P(\xi_m = i, \xi_t = l)p_{l,j}$$

и вытекает из (16.1). Если оно верно при некотором $n > t$, то верно и при $n + 1$, так как

$$\begin{aligned} P(\xi_m = i, \xi_t = l, \xi_{n+1} = j) &= \sum_k P(\xi_m = i, \xi_t = l, \xi_n = k)p_{k,j} \\ &= \sum_k P(\xi_m = i, \xi_t = l)p_{l,k}^{(n-t)}p_{k,j} = P(\xi_m = i, \xi_t = l)p_{l,j}^{(n+1-t)}, \end{aligned}$$

поскольку $\sum_k p_{l,k}^{(n-t)}p_{k,j} = p_{l,j}^{(n+1-t)}$. $\square$

Вместо теоремы Колмогорова можно построить случайные величины ξ_n с заданным начальным распределением и заданными вероятностями перехода $p_{i,j}$ по индукции, используя последовательность независимых случайных величин η_n, равномерно распределенных в $[0,1]$. В качестве ξ_0 берем с.в. с распределением π_0 на $1,\ldots,m$. Затем $\xi_1(\omega)$ задаем с помощью значения $\eta_1(\omega)$ так: делим $[0,1]$ на m промежутков с длинами $\delta_1,\ldots,\delta_m$, $\delta_k=\pi_0(1)p_{1,k}+\cdots+\pi_0(m)p_{m,k}$, а номер $\xi_1(\omega)$ есть номер того промежутка, в который попало число $\eta_1(\omega)$. Если величина ξ_n со значениями в $1,\ldots,m$ уже определена, то ξ_{n+1} задаем так: делим $[0,1]$ на m промежутков с длинами $p_{\xi_n,j}$ и берем номер промежутка, в который попадает $\eta_{n+1}(\omega)$. Длины промежутков подбираются так, чтобы обеспечить вероятности значений $\xi_{n+1}=j$, $j=1,\ldots,m$, заданные равенством $\pi_{n+1}=\pi_0P^{n+1}$. Обычно, однако, сами случайные величины ξ_n используются редко, а имеют дело с начальным распределением и переходными вероятностями.

Случай счетного фазового пространства аналогичен. Из сказанного явствует, что однородную цепь Маркова с начальным распределением π_0 и переходными вероятностями $p_{i,j}$ можно определять без ссылки на общие марковские процессы как последовательность случайных величин ξ_n, $n=0,1,\ldots$, со значениями в счетном пространстве $X=\{x_i\}$, удовлетворяющую таким условиям: π_0 есть распределение ξ_0 и

$$\begin{aligned} P(\xi_{n+1}=j|\xi_0=x_{i_0},\xi_1=x_{i_1},\ldots,\xi_n=x_{i_n}) \\ =P(\xi_{n+1}=j|\xi_n=x_{i_n})=\pi_{i_n,j}. \end{aligned}$$

Последнее условие как раз и есть марковость (будущее зависит от прошлого только через настоящее).

§ 17. Стационарные распределения конечной цепи Маркова и эргодичность

Вероятностная мера π на X называется стационарной или инвариантной меры цепи, если

$$\pi=\pi P.$$

В случае конечного X такая мера всегда существует. В самом деле, множество вероятностных мер на $X=\{x_1,\ldots,x_m\}$ можно отождествить с выпуклым симплексом S в $\mathbb{R}^m$, состоящим из векторов с неотрицательными компонентами с единичной суммой. Матрица P задает линейный оператор в $\mathbb{R}^m$, переводящий S в S. По теореме Боля–Брауэра он обладает неподвижной точкой. В общем случае инвариантная мера не единственна, но в следующей теореме, называемой *эргодической*, выделен важный случай, когда она единственна, причем при любом начальном распределении распределения π_n сходятся к инвариантному. В данном конечном случае все разумные виды сходимости мер совпадают, так что можно говорить о сходимости по вариации (по метрике $\|\mu - \nu\| = \sum |\mu(x_i) - \nu(x_i)|$) или на каждом элементе из X.

17.1. Теорема. *Пусть X конечно и при некотором n_1 все элементы матрицы P^{n_1} положительны. Тогда инвариантная вероятностная мера μ цепи единственна, $\mu(x_i) > 0$ для всех i, причем к ней сходятся меры $\pi_n = \pi_0 P^n$ для всякого начального распределения цепи.*

Доказательство. Зафиксируем какую-либо инвариантную вероятностную меру μ. Для упрощения обозначений будем считать, что $n_1 = 1$, общий случай аналогичен. Пусть $M = (m_{i,j})$ — матрица, в которой все строки совпадают с μ, т.е. с $(\mu(s_1), \ldots, \mu(s_m))$. Для этой матрицы $\pi M = \mu$ для всех вероятностных мер π. Найдется такое $\delta \in (0,1)$, что $p_{ij} \geqslant \delta m_{ij}$ при всех i, j. Положим $\lambda = 1 - \delta$. Тогда $P - \delta M \geqslant 0$ и P можно представить в виде выпуклой комбинации

$$P = (1 - \lambda) M + \lambda Q$$

матрицы M и некоторой матрицы $Q = (q_{ij})$ с $q_{ij} \geqslant 0$ и $\sum_j q_{ij} = 1$, т.е. $Q = \lambda^{-1}(P - \delta M)$. При таком выборе оказывается, что

$$P^n = (1 - \lambda^n) M + \lambda^n Q^n.$$

Действительно, непосредственно проверяется, что

$$M^2 = M \quad \text{и} \quad MP = PM = M.$$

Поэтому $QM = M$. Значит, рассуждая по индукции, имеем

$$\begin{aligned} P^{n+1} = ((1-\lambda^n)M + \lambda^n Q^n)P = (1-\lambda^n)M + \lambda^n Q^n((1-\lambda)M + \lambda Q) \\ = (1 - \lambda^{n+1})M + \lambda^{n+1} Q^{n+1}, \end{aligned}$$

что завершает индуктивный шаг. Таким образом, с учетом равенства $\pi_0 M - \mu M = 0$ получаем

$$\begin{aligned} \pi_0 P^n - \mu = \pi_0 P^n - \mu P^n = (\pi_0 - \mu) P^n = (\pi_0 - \mu)((1-\lambda^n)M + \lambda^n Q^n) \\ = \lambda^n (\pi_0 - \mu) Q^n, \end{aligned}$$

что стремится к нулю при $n \to \infty$, так как $\|(\pi_0 - \mu) Q^n\| \leqslant 2$. Если вероятностная мера ν также инвариантна, то $\nu = P^n \nu \to \mu$, откуда $\nu = \mu$. Наконец, $P^{n_1} \mu(x_i) > 0$ для всех i. $\square$

Отметим, что условие положительности некоторой степени P и необходимо для справедливости заключения теоремы. В самом деле, если всякая матрица P^n имеет нулевой элемент, то для некоторых i_0, j_0

для бесконечной последовательности n_k матрица P^{n_k} имеет нулевой элемент с индексами i_0, j_0. Тогда для меры π, равной нулю на всех x_i с $i \neq i_0$ и $\pi(x_{i_0}) = 1$, получаем $\pi P^{n_k}(x_{j_0}) = 0$, поэтому предел не может быть положительным на всех x_j.

§ 18. Ветвящиеся процессы Гальтона – Ватсона

Ветвящиеся процессы Гальтона – Ватсона — классика теории вероятностей, возникшая в Англии еще в XIX веке в исследовании по вырождению знатных британских семейств, проведенном англиканским священником Уотсоном (Watson, не путать с доктором Уотсоном) и естествоиспытателем широкого профиля Гальтоном (Galton).

С точки зрения общей теории процесс Гальтона – Ватсона представляет собой однородную цепь Маркова Z_n с фазовым пространством $X = \mathbb{Z}_+$, начальным распределением π_0 и переходными вероятностями вида

$$p_{k,j} = P(Z_{n+1} = j | Z_n = k) = p_j^{*k},\ k > 0, \quad p_{0,j} = 0,\ j > 0, \quad p_{0,0} = 1,$$

где

$$p_j^{*k} = \sum_{j_1 + \cdots + j_k = j} p_{j_1} p_{j_2} \cdots p_{j_k},$$

причем $(p_0, p_1, p_2, \ldots)$ — заданное вероятностное распределение на $\mathbb{Z}_+$.

Если взять независимые случайные величины $\xi_1, \xi_2, \ldots$ с общим распределением $P(\xi = m) = p_m$, то числа p_j^{*k} можно записать в виде

$$p_j^{*k} = P(\xi_1 + \cdots + \xi_k = j).$$

Далее предполагается, что $\pi_0(1) = 1$, т.е. $\pi_0(k) = 0$ при $k \neq 1$.

Интерпретация процесса Гальтона – Ватсона такова. В момент 0 есть одна частица. Ее продолжительность жизни 1, в момент 1 она порождает случайное число потомков с заданным распределением

$$P(\xi = m) = p_m.$$

Каждый из потомков также имеет продолжительность жизни 1 и независимо от других потомков производит в момент времени 2 случайное число потомков с тем же распределением и т.д.

Число частиц в n-м поколении есть Z_n. Взяв серию $\{\xi_k^{(n)}\}_{k,n \in \mathbb{N}}$ независимых последовательностей независимых случайных величин с распределением ξ (т.е. все они независимы в совокупности), можно

представить Z_n как
$$Z_n = \sum_{k=1}^{Z_{n-1}} \xi_k^{(n)},\ n \geqslant 1, Z_0 = 1.$$
где число потомков k-й частицы из поколения $n-1$ равно $\xi_k^{(n)}$.

Нас интересует вероятность вырождения
$$q = P(\exists\, n\colon Z_n = 0).$$
Ясно, что число q есть предел возрастающей последовательности чисел $q_n = P(Z_n = 0)$.

Для исследования этого вопроса полезен аналог преобразования Фурье.

18.1. Определение. *Производящей функцией случайной величины ξ со значениями в $\mathbb{Z}_+$ называется функция*
$$\varphi_\xi(z) = \mathrm{E}z^\xi = \sum_{k\geqslant 0} z^k P(\xi = k), \quad z \in \mathbb{C},\ |z| \leqslant 1.$$

Ясно, что указанный ряд сходится абсолютно в замкнутом единичном круге и задает аналитическую функцию внутри, непрерывную на замкнутом круге, причем $\varphi_\xi(1) = 1$, $\varphi'_\xi(1) = \mathrm{E}\xi$.

Для $z = e^{it}$ получаем характеристический функционал (преобразование Фурье). Если ξ_1 и ξ_2 независимы, то $\varphi_{\xi_1+\xi_2} = \varphi_{\xi_1}\varphi_{\xi_2}$.

18.2. Лемма. *Справедливо равенство*
$$\varphi_{Z_n}(z) = \varphi_{Z_{n-1}}(\varphi_\xi(z)).$$

Доказательство. Заметим, что $|\varphi_\xi(z)| \leqslant 1$ при $|z| \leqslant 1$. Кроме того,
$$\mathrm{E}\Big(z^{Z_n}|Z_{n-1} = m\Big) = \mathrm{E}\Big(z^{\sum_{k\leqslant Z_{n-1}} \xi_k^{(n)}}|Z_{n-1} = m\Big)$$
$$= \mathrm{E}\Big(z^{\sum_{k\leqslant m} \xi_k^{(n)}}|Z_{n-1} = m\Big) = \mathrm{E}\Big(z^{\sum_{k\leqslant m} \xi_k^{(n)}}\Big) = \varphi_\xi(z)^m,$$
так как Z_{n-1} зависит от ξ_k^l с $l \leqslant n-1$ и потому оказывается независимой со всеми $\xi_k^{(n)}$. Следовательно,
$$\mathrm{E}(z^{Z_n}|Z_{n-1}) = (\varphi_\xi(z))^{Z_{n-1}}.$$
Итак,
$$\varphi_{Z_n}(z) = \mathrm{E}z^{Z_n} = \mathrm{E}(\mathrm{E}(z^{Z_n}|Z_{n-1})) = \mathrm{E}(\varphi_\xi(z))^{Z_{n-1}},$$

что дает анонсированное равенство. □

Из леммы следует, что
$$\varphi_{Z_n}(z) = \varphi_\xi(\varphi_\xi(\cdots(\varphi_\xi(z))),$$
где стоит n итераций. В свою очередь это дает равенство
$$\varphi_{Z_n}(z) = \varphi_\xi(\varphi_{Z_{n-1}}(z)).$$

§ 19. Вероятности вырождения процесса Гальтона – Ватсона

19.1. Лемма. *Вероятность вырождения q удовлетворяет уравнению*
$$q = \varphi_\xi(q).$$

Доказательство. В силу сказанного выше
$$q_n = P(Z_n = 0) = \varphi_{Z_n}(0) = \varphi_\xi(\varphi_{Z_{n-1}}(0)) = \varphi_\xi(q_{n-1}).$$
Левая часть стремится к q, правая — к $\varphi_\xi(q)$. □

Однако уравнение $z = \varphi_\xi(z)$ может иметь не один корень, скажем, $\varphi_\xi(1) = 1$.

19.2. Теорема. *Пусть $P(\xi=1) < 1$. Если ξ имеет среднее $m_\xi \leqslant 1$, то уравнение $z = \varphi_\xi(z)$ имеет на отрезке $[0,1]$ единственный корень 1 и тогда $q = 1$.*

Если среднего нет или оно больше 1, то на $[0,1)$ имеется единственный корень z_0 уравнения $z = \varphi_\xi(z)$ и тогда $q = z_0$.

Таким образом, q всегда есть наименьший корень этого уравнения на $[0,1]$.

Доказательство. Так как
$$\varphi'_\xi(x) = \sum_{k\geqslant 1} kz^{k-1}P(\xi = k), \quad \varphi''_\xi(x) = \sum_{k\geqslant 2} k(k-1)z^{k-2}P(\xi = k),$$
то на полуинтервале $[0,1)$ имеем $\varphi''_\xi(z) \geqslant 0$, поэтому функция φ_ξ выпукла. При этом $\varphi_\xi(0) = P(\xi = 0) \geqslant 0$, $\varphi_\xi(1) = 1$.

Если $P(\xi = 0) = 1$, то $\varphi_\xi = 1$ и все ясно. Если $\varphi''_\xi = 0$ на $[0,1)$ и $P(\xi = 0) < 1$, то $\varphi_\xi(z) = P(\xi = 0) + zP(\xi = 1)$, поэтому $q = 1$ — единственный корень. В остальных случаях $\varphi''_\xi > 0$ на $(0,1)$, поэтому кроме корня 1 может быть только еще один корень $z_0 \in [0,1)$.

Если $\varphi'_\xi(1) = \mathrm{E}\xi \leqslant 1$, то второго корня нет, ибо в этом случае $\varphi_\xi(z) \geqslant z$ при $z \in [0,1]$ из-за выпуклости.

Пусть $\mathrm{E}\xi > 1$ или $\mathrm{E}\xi = +\infty$. Тогда есть ровно один корень z_0 в $[0,1)$. Покажем, что $q = z_0$. Для этого проверим, что $q_n \leqslant z_0$ при всех n. Это верно при $n = 0$. При последующих итерациях также получаются точки из множества $\{z \in [0,1]\colon \varphi_\xi(z) \geqslant z\} = [0, z_0]$. Действительно, $\varphi_\xi(q_n) = q_{n+1} \geqslant q_n$. Значит, предел q также попадет в это множество. □

19.3. Пример. (i) Для бинарного расщепления частиц, когда в конце жизни частица либо не имеет потомков с вероятностью $1-p$, либо производит двух потомков с вероятностью p, вероятность вырождения — корень уравнения $1 - p + px^2 = x$. В результате получаем $q = (1-p)/p$, если $p > 1/2$, и $q = 1$ в противном случае.

(ii) Случай геометрического распределения $P(\xi = k) = (1-p)p^k$, $p \in (0,1)$, получаем уравнение $x = (1-p)/(1-xp)$, дающее такой же результат. Здесь $m_\xi = p/(1-p)$, поэтому при $m_\xi = 6/5$ вероятность вырождения довольно высока ($= 5/6 > 0.8$).

§ 20. Модель страхования Крамера – Лундберга

Пусть N_t — пуассоновский процесс интенсивности $\lambda > 0$, $\{\eta_n\}$ — последовательность независимых одинаково распределенных неотрицательных случайных величин, независимая с N_t. Для фиксированных чисел $\eta_0 > 0$, $c > 0$ введем процесс

$$X_t = \eta_0 + ct - \sum_{k=1}^{N_t} \eta_k, \quad t \geqslant 0.$$

Величина X_t интерпретируется как капитал страховой компании в момент t, η_0 — начальный капитал, c — скорость поступления страховых взносов, N_t — количество страховых исков от начала до момента времени t, η_k — сумма выплат по k-у страховому случаю.

Момент разорения есть

$$\tau = \inf\{t\colon\ X_t < 0\}.$$

Задача состоит в оценивании $P(\tau < +\infty)$.

Предполагается, что для $a = \mathrm{E}\eta_1 > 0$ верно неравенство $c - \lambda a > 0$ и $\psi(v) = \mathrm{E}e^{v\eta_1} < \infty$ при всех $v > 0$.

Можно проверить, что процесс X_t имеет независимые приращения и непрерывные справа траектории. Кроме того, при $t > s$ имеем

$$\mathrm{E}\exp(-v(X_t - X_s)) = \exp((t-s)g(v)), \quad g(v) = \lambda(\psi(v) - 1) - vc.$$

Из этого вытекает, что процесс

$$M_t = \exp(-vX_t - tg(v))$$

является мартингалом относительно фильтрации $\mathcal{F}_t$, порожденной X_t, с непрерывными справа траекториями. Мартингальность проверяется так:

$$\begin{aligned}\mathrm{E}(M_t|\mathcal{F}_s) &= \mathrm{E}\Big(\exp(-v(X_t - X_s) - vX_s - tg(v))|\mathcal{F}_s\Big)\\ &= \exp(-vX_s - tg(v))\mathrm{E}\Big(\exp(-v(X_t - X_s))|\mathcal{F}_s\Big)\\ &= \exp(-vX_s - tg(v))\mathrm{E}\Big(\exp(-v(X_t - X_s))\Big),\end{aligned}$$

где в последнем равенстве использована независимость приращений. Так как математическое ожидание справа равно $\exp((t-s)g(v))$, получаем M_s. Для всякого $t > 0$ величина $\min(\tau, t)$ является ограниченным моментом остановки. Поэтому

$$\begin{aligned}e^{-v\eta_0} &= \mathrm{E}M_0 = \mathrm{E}M_{\min(\tau,t)} \geqslant \mathrm{E}(M_{\min(\tau,t)}I_{\{\tau\leqslant t\}}) = \mathrm{E}(M_\tau I_{\{\tau\leqslant t\}})\\ &= \mathrm{E}(\exp(-vX_\tau - \tau g(v))I_{\{\tau\leqslant t\}}) \geqslant \mathrm{E}(\exp(-\tau g(v))I_{\{\tau\leqslant t\}}),\end{aligned}$$

так как $X_\tau \leqslant 0$ в силу непрерывности справа. Поскольку при $\tau \leqslant t$ имеем $\exp(\tau g(v)) \leqslant \max_{s\in[0,t]} \exp(sg(v))$, то приходим к оценке

$$P(\tau \leqslant t) \leqslant e^{-v\eta_0} \max_{s\in[0,t]} e^{sg(v)}.$$

Если $g(v) > 0$, то $\max$ при $s = t$. Если $g(v) \leqslant 0$, то $\max = 1$ при $s = 0$. Возьмем наибольшее v, для которого $g(v) \leqslant 0$.

Для этого проведем анализ функции

$$g(v) = \lambda(\psi(v) - 1) - vc, \quad \psi(v) = \mathrm{E}e^{v\eta_1}, \quad v \geqslant 0.$$

Имеем:

$$g(0) = 0, \quad g'(v) = \lambda\psi'(v) - c, \quad g''(v) = \lambda\psi''(v),$$

При этом $g'(0) = \lambda\mathrm{E}\eta_1 - c < 0$, $g''(v) > 0$ при $v > 0$, т.е. функция g выпукла, на некотором интервале справа от нуля функция g отрицательна, но $g(v) \to +\infty$ при $v \to +\infty$. Значит, имеется единственный положительный корень v_0 уравнения $g(v) = 0$.

Приходим к такому заключению.

20.1. Теорема. *Имеет место оценка*

$$P(\tau < +\infty) \leqslant e^{-v_0\eta_0},$$

где v_0 — единственный положительный корень уравнения $g(v) = 0$.

Иногда проще найти точку минимума g, где $\psi'(v) = c/\lambda$ (хотя тогда оценка будет хуже). Например, если $P(\xi_1 = 0) = P(\xi_1 = 1) = 1/2$, то $\psi(v) = (1 + e^v)/2$ и $\psi'(v) = e^v/2$, т.е. точка минимума есть $\ln(2c/\lambda)$. Это дает $P(\tau < +\infty) \leqslant \exp(-\eta_0 \ln(2c/\lambda)) = \left(\frac{\lambda}{2c}\right)^{\eta_0}$.

Программа курса «Случайные процессы»

1. Общее понятие случайного процесса (случайной функции), траектории случайного процесса. Пример: процессы восстановления.

2. Пространство траекторий, цилиндрическая сигма-алгебра, меры на пространстве траекторий.

3. Конечномерные распределения случайного процесса, условия согласованности. Распределения случайных процессов в пространстве траекторий. Теорема Колмогорова о существовании случайного процесса с заданными конечномерными распределениями (формулировка).

4. Процессы с независимыми приращениями. Критерий существования процесса с независимыми приращениями в терминах характеристических функций.

5. Пуассоновский процесс постоянной интенсивности как процесс с независимыми приращениями, доказательство существования. Явная конструкция пуассоновского процесса как процесса восстановления для экспоненциальных случайных величин.

6. Ковариационная функция случайного процесса, ее симметричность и неотрицательная определенность. Гауссовские случайные процессы. Доказательство существования гауссовского процесса с заданными средним и ковариацией.

7. Винеровский процесс (процесс броуновского движения). Доказательство существования.

8. Модификация случайного процесса. Теорема Колмогорова о существовании непрерывной модификации (формулировка). Доказательство существования непрерывной модификации винеровского процесса. Представление винеровского процесса в виде функционального ряда с независимыми случайными коэффициентами (формулировка).

9. Броуновский мост. Недифференцируемость траекторий винеровского процесса (формулировка). Закон повторного логарифма для винеровского процесса (формулировка).

10. Условные вероятности и условные математические ожидания (определение и основные свойства).

11. Понятие фильтрации на вероятностном пространстве, согласованность случайного процесса с фильтрацией. Мартингалы, субмартингалы и супермартингалы. Критерий мартингальности для процессов с независимыми приращениями. Примеры мартингалов и субмартингалов. Теорема Дуба о сходимости мартингалов (формулировка).

12. Разложение Дуба–Мейера для согласованных процессов с дискретным временем.

13. Марковские моменты и моменты остановки. Теорема об остановке для дискретного времени. Теорема об остановке для непрерывного времени (формулировка).

14. Задача о разорении игрока: мартингальный подход к решению.

15. Марковские процессы, равносильность различных описаний марковского свойства. Переходные функции, уравнения Чэпмена – Колмогорова (формулировка). Однородные марковские процессы. Примеры.

16. Цепи Маркова с дискретным временем и счетным числом состояний. Теорема о существовании конечных марковских цепей (формулировка). Однородные цепи Маркова. Существование стационарных распределений однородной конечной цепи Маркова.

17. Условие эргодичности конечной цепи Маркова.

18. Ветвящиеся процессы Гальтона – Ватсона. Производящие функции случайных величин, их основные свойства. Соотношение между производящими функциями числа частиц в n-м и $(n+1)$-м поколениях.

19. Вывод уравнения для вероятности вырождения процесса Гальтона – Ватсона. Теорема о вероятности вырождения.

20. Модель страхования Крамера – Лундберга и оценка вероятности разорения (формулировка основных свойств возникающих при анализе процессов без доказательства).

ЛИТЕРАТУРА

1. Булинский А. В., Ширяев А. Н. Теория случайных процессов. М., Физматлит, 2005.

2. Ватутин В.А. Ветвящиеся процессы и их применения. Лекционные курсы НОЦ. Т. 8. М., МИАН, 2008.

3. Вентцель А.Д. Курс теории случайных процессов. М., 1975.

4. Гасников А.В. и др. Лекции по случайным процессам. М., МФТИ, 2019.

5. Натан А.А., Горбачев О.Г., Гуз С.А. Основы теории случайных процессов: уч. пособие по курсу «Случайные процессы». М.: МЗ Пресс, МФТИ, 2003.