М.А. МАТАЛЫЦКИЙ

ЭЛЕМЕНТЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

Учебное пособие для студентов математических специальностей учреждений, обеспечивающих получение высшего образования

Допущено Министерством образования Республики Беларусь

Учреждение образования
«Гродненский государственный университет
имени Янки Купалы» 2004

УДК 519.2
ББК 22.171
М33

Рецензенты: кафедра теории вероятностей и математической статистики Белорусского государственного университета, заведующий кафедрой, профессор, доктор физико-математических наук Н.Н. Труш,

заведующий кафедрой математического анализа Гомельского государственного университета имени Ф. Скорины, профессор, доктор физико-математических наук Ю.В. Малинковский.

Маталыцкий М.А.

М33 Элементы теории случайных процессов: Учеб. пособие / М.А.Маталыцкий. – Гродно: ГрГУ, 2004. – 326 с.

ISBN 985-417-471-9

В учебном пособии приведены теоретические сведения о марковских случайных процессах и цепях Маркова с дискретным и непрерывным временем, процессах с конечными моментами второго порядка (корреляционная теория), процессах с независимыми приращениями, стационарных и эргодических случайных процессах, стохастических интегралах и стохастических дифференциальных уравнениях. Рассмотрены вопросы применения случайных процессов при анализе математических моделей различных реальных объектов. Приведены решения более 100 различных типовых примеров.

Для студентов математических специальностей вузов, а также научных и инженерных работников, которые интересуются теорией случайных процессов и ее применениями.

УДК 519.2
ББК 22.171

ISBN 985-417-471-9

Посвящается 30-ю кафедры теории вероятностей и математической статистики Белгосуниверситета и 60-ю математического факультета Гродненского госуниверситета

ПРЕДИСЛОВИЕ

За последние десятилетия значительно увеличился объем преподавания дисциплин, использующих вероятностные и статистические методы, в высших учебных заведениях. В университетах для студентов математических специальностей, таких, как «математика», «прикладная математика», «экономическая кибернетика», «актуарная математика», «математическая экономика», «компьютерная математика», читается годовой или полутораgодовой курс теории вероятностей и математической статистики. Курс состоит из трех основных разделов: элементы теории вероятностей, элементы теории случайных процессов и элементы математической статистики. По первому и третьему разделу издано достаточно большое число учебных пособий и сборников задач, в то время как по второму разделу имеется значительный пробел в этом направлении. Предлагаемое учебное пособие позволяет сократить этот пробел. Оно подготовлено на базе курса лекций и методических разработок по университетскому курсу теории вероятностей и случайных процессов, читаемому автором для студентов математических специальностей.

Учебное пособие рассчитано на лиц, усвоивших раздел теории вероятностей и приступивших к изучению теории случайных процессов. Для облегчения такого изучения вначале в нем излагаются основные понятия и методы современной теории вероятностей, снабженные большим числом примеров. Автор надеется, что пособие будет полезно для студентов, магистрантов и аспирантов университетов, а также для различных специалистов, желающих познакомиться с основными методами и результатами теории в достаточно строгом, но не самом общем и исчерпывающем, изложении. На простых моделях изучаются наиболее характерные свойства различных типов случайных процессов. Рассматриваются теоретико-вероятностные задачи, представляющие интерес для приложений.

Автор выражает благодарность рецензентам, сделавшим ряд полезных замечаний.

ВВЕДЕНИЕ

При изучении различных явлений действительности мы сталкиваемся с процессами, предсказать развитие которых заранее не можем. Случайные процессы –удобная математическая модель функций времени, значениями которых являются случайные величины. Например: число запросов, поступающих в единицу времени на центральный сервер информационно-компьютерной сети, являясь случайной величиной, зависит от времени суток; расход электроэнергии в единицу времени также является функцией времени со случайными значениями; координата отдельной молекулы в газе, заключенном в сосуд, меняется со временем и принимает случайные значения. Таким образом, можно сказать, что случайный процесс – это семейство случайных величин, зависящих от времени.

Теория случайных процессов, возникшая в результате построения математических моделей реальных физических процессов, представляет собой наиболее содержательную и более всего используемую в приложениях часть теории вероятностей. Она находит многочисленные применения в физике, технике, экономике, биологии, медицине и других дисциплинах, а также в различных разделах математики.

Приведем некоторые исторические замечания. Первое математическое описание случайного процесса, называемого в настоящее время винеровским, или процессом броуновского движения, дал Л.Башелье в докладе, представленном им Парижской академии в 1900 г. [9]. Он предложил использовать этот процесс в качестве модели колебаний цены активов, стремился получить аналитические выражения для стоимости различных типов опционов и сравнить их с наблюдаемыми рыночными ценами опционов. Опцион является примером финансовой производной и дает его владельцу право купить указанное число долей акций по определенной цене в указанную дату или до нее. Вообще понятие случайного процесса принадлежит XX веку и связано с именами А.Н.Колмогорова, А.Я.Хинчина, Е.Е.Слуцкого, Н.Винера. То, что теория случайных процессов принадлежит к категории наиболее быстро развивающихся математических дисциплин, в значительной мере определяется ее глубокими связями с практикой [3].

В 1905 г. двумя известными физиками, М.Смолуховским и А.Эйнштейном, была разработана теория броуновского движения, исходящая из теоретико-вероятностных предпосылок. Она привела математику к порогу создания теории случайных процессов. В исследованиях датского ученого А.К.Эрланга была начата новая важная область поисков, связанных с изучением загрузки телефонных сетей. Эти работы оказали значительное влияние не только на решение чисто телефонных задач, но и на формирование элементов теории случайных процессов, в частности, процессов гибели и размножения. Такие процессы позднее применялись при исследовании динамики биологических популяций, именно от задач биологии и пошло наименование этого частного типа случайных процессов.

Изучение явления диффузии средствами теории вероятностей было предпринято в 1914 г. известными физиками – М.Планком и А.Фоккером. Н.Винер, основатель кибернетики, в середине 20-х годов при изучении броуновского движения ввел в рассмотрение процессы, названные винеровскими. Следует также отметить работы русского математика, профессора МГУ А.А.Маркова по изучению цепных зависимостей (цепи Маркова) и работы Е.Е.Слуцкого по теории случайных функций.

В 1931 г. была опубликована известная большая статья А.Н.-Колмогорова «Об аналитических методах в теории вероятностей», в которой были заложены основы теории марковских процессов: в ней получены прямые и обратные дифференциальные уравнения, которые управляют вероятностями перехода случайных процессов без последействия. В этой же работе был дан набросок теории скачкообразных процессов без последействия, подробное развитие которой позднее (1936 г.) было дано В.Фелером, получившим интегро-дифференциальное уравнение для скачкообразных марковских процессов. В 1934 г. в работе А.Я.Хинчина осуществлено построение основ стационарных случайных процессов на базе физических задач. Он ввел понятие стационарного процесса в узком и широком смыслах. Вышеупомянутые работы следует считать началом построения общей теории случайных процессов, они послужили основой для последующих исследований Г.Крамера, Г.Вальда, А.Н.Колмогорова и многих других известных ученых. Более подробно история развития теории случайных процессов изложена в [3].

Охарактеризуем ряд основных задач теории случайных процессов, большинство из которых рассматривается в данном учебном пособии.

1. Одной из основных задач является построение математической модели, допускающее строгое или формальное определение случайного процесса, и исследование общих свойств этой модели.

2. Важной задачей является классификация случайных процессов. Существующая классификация в теории случайных процессов заключается в выделении из всей совокупности таких процессов некоторых классов, допускающих более или менее конструктивное описание [2]. Каждый класс характеризуется тем, что достаточно дополнительно задать лишь конечное число функциональных характеристик, чтобы выделить из всего класса отдельный случайный процесс. Иногда рассматривают классы процессов, допускающих единообразное решение определенного набора задач. Можно отметить следующие широкие классы процессов: а) марковские процессы, включая, естественно, цепи Маркова, б) процессы с конечными моментами второго порядка (гильбертовы процессы), в) процессы с независимыми приращениями, г) стационарные в узком и широком смыслах случайные процессы, в частности, гауссовский и винеровский процессы; д) эргодические процессы.

3. Задача отыскания для различных классов случайных процессов аналитического аппарата, дающего возможность находить вероятность характеристики процессов, тесно связана с предыдущей. Для простейших вероятностных характеристик такой аппарат создан и использует, как правило, теорию обыкновенных дифференциальных уравнений и уравнений с частными производными, а также интегро-дифференциальные и интегральные уравнения, разностные уравнения, преобразования Фурье.

4. Изучение различных преобразований случайных процессов также является важной задачей теории случайных процессов. Эти преобразования используются для того, чтобы с их помощью изучать сложные процессы путем сведения их к более простым. К такой задаче можно отнести и анализ стохастических дифференциальных и интегральных уравнений, в которые входят случайные процессы.

5. Задача определения значений некоторого функционала от процесса по значениям других функционалов от этого же процесса играет также важную роль в формировании ряда разделов теории случайных процессов. Примером такой задачи является задача предсказания, позволяющая определить значение процесса в некоторые будущие моменты времени, наблюдая процесс в течение определенного промежутка времени.

Опишем кратко некоторые основные области применения различных классов случайных процессов в настоящее время. Марковские процессы широко используются при разработке математических моделей информационно-компьютерных систем и сетей, в математической, в частности, финансовой экономике, в математической биологии, в теории каскадов космических частиц. В этой же теории применяются процессы с независимыми приращениями. Стационарные в узком и широком смыслах случайные процессы имеют широкое применение в радиоэлектронике и теории информации, а гауссовские процессы – также в радиоэлектронике и молекулярной теории газов. Вообще в последнее время методы теории случайных процессов находят все новые области применения, и сейчас ни одна из естественных наук и многие гуманитарные науки не избежали влияния этой теории.

ГЛАВА 1. ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ

§1. СЛУЧАЙНЫЕ СОБЫТИЯ. КЛАССИЧЕСКОЕ ОПРЕДЕЛЕНИЕ ВЕРОЯТНОСТИ

Определение. Возможные события, порождаемые комплексом условий, называются элементарными, если:

а) они различны (т.е. осуществление одного означает неосуществление любого другого);

б) после выполнения комплекса условия обязательно происходит одно из них.

Обозначим через $\Omega = \{\omega_1, \omega_2, ..., \omega_n, ...\}$ пространство элементарных событий.

Определение. Любое объединение элементарных событий называется случайным событием, $B \subseteq \Omega$.

Событие B осуществляется тогда, когда происходит одно из элементарных событий $\omega \in B$. В этом смысле пространство Ω может рассматриваться тоже как событие. Т.к. одно из элементарных событий происходит всегда, то и событие Ω происходит всегда, поэтому оно является достоверным событием. Событие, не содержащее ни одного элементарного события, является невозможным и обозначается $\varnothing$.

Таким образом, мы пришли к описанию случайных событий как множеств, получающихся объединением элементарных событий. В связи с этим для определения соотношений между случайными событиями в теории вероятностей принят язык теории множеств, который приобретает своеобразную вероятностную трактовку. Поясним некоторые из них при помощи таблицы 1.

Таблица 1

Язык теории множеств	Соотношения между событиями на языке теории множеств
$A = B \bigcup C$	Событие A (объединение событий B и C) происходит тогда и только тогда, когда происходит хотя бы одно из событий B и C.
$A = B \bigcap C$	Событие A (пересечение событий B и C) происходит тогда и только тогда, когда происходят и событие B, и событие C.
$B \bigcap C = \varnothing$	События B и C являются несовместными. Если событие C происходит, то событие B не происходит.
$C \subseteq B$	Событие C влечет за собой событие B.
$A = \Omega \setminus B$, $\left(A = \overline{B}\right)$	Событие A является дополнительным (противоположным) по отношению к событию B. Событие A происходит тогда и только тогда, когда не происходит событие B.

1. $A \bigcup B = B \bigcup A,\ A \bigcap B = B \bigcap A$. Эти равенства следуют из определений.

2. $\overline{A \bigcup B} = \overline{A} \bigcap \overline{B}$. Доказательство следует из следующей цепочки импликаций:

$$\omega \in \overline{A \bigcup B} \Rightarrow \omega \notin A \bigcup B \Rightarrow \omega \notin A, \omega \notin B \Rightarrow \omega \in \overline{A}, \omega \in \overline{B} \Rightarrow \omega \in \overline{A} \bigcap \overline{B}$$

и наоборот

$$\omega \in \overline{A} \bigcap \overline{B} \Rightarrow \omega \notin A, \omega \notin B \Rightarrow \omega \notin A \bigcup B \Rightarrow \omega \in \overline{A \bigcup B}.$$

3. $\overline{A \bigcap B} = \overline{A} \bigcup \overline{B}$.

4. $A \subset B \Rightarrow \overline{B} \subset \overline{A}$. Это следует из того, что если $\omega \in \overline{B}$, то $\omega \notin B$, поэтому $\omega \notin A$ и, значит, $\omega \in \overline{A}$.

Из соотношений 2 – 4 следует, что если задана некоторая конструкция из событий, ее дополнение можно выразить, заменив в ней все события на противоположные, символы объединения, пересечения и включения – на символы пересечения, объединения и обратный к включению соответственно. Это свойство известно под названием закона де Моргана, например, $\overline{(A \cup B) \cap C} = \left(\overline{A} \cap \overline{B}\right) \cup \overline{C}$.

Определение *(классическое определение вероятности)*. Пусть пространство элементарных событий состоит из конечного числа равновозможных элементарных событий $\Omega = \{\omega_1, \omega_2, ..., \omega_N\}$ и пусть случайное событие A состоит из n элементарных событий: $A = \{\omega_{j_1}, \omega_{j_2}, ..., \omega_{j_n}\}$, $\omega_{j_i} \in \Omega$, $i = \overline{1, n}$. Тогда вероятностью события A называется число

$$P(A) = \frac{n}{N}.$$

Данная формула называется формулой классической вероятности.

Пример 1.1. Монету бросают дважды. Найти вероятность того, что хотя бы один раз монета упадёт гербом вверх.

Решение. Пространством элементарных событий является множество $\Omega = \{ГГ, ГЦ, ЦГ, ЦЦ\}$. Здесь, например, ГЦ означает, что при первом бросании появился герб, а при втором – цифра. Таким образом $N = 4$. Пусть A = {хотя бы один раз появится герб}, тогда

$$A = \{ГГ, ГЦ, ЦГ\} \Rightarrow n = 3 \Rightarrow P(A) = \frac{3}{4}.$$

Основной проблемой при решении задач с использованием формулы классической вероятности является подсчет числа способов, которыми могло произойти то или иное событие. В связи с этим такие задачи решаются, как правило, методами комбинаторики.

Часто используется следующее очевидное правило (основной принцип комбинаторики): если некий выбор A можно осуществить m различными способами, а некоторый другой выбор B можно осуществить n способами, то выбор A и B (A или B) можно осуществить mn $(m+n)$ способами.

При этом классическое определение вероятности можно дать другими словами.

Определение. Рассмотрим эксперимент, имеющий N одинаково возможных исходов (любой мыслимый результат эксперимента называется элементарным событием). Предположим, что событию A благоприятствует n из этих исходов (оно состоит из n элементарных событий). Тогда справедлива формула классической вероятности.

При решении задач часто используют размещения, перестановки и сочетания. Если дано множество $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$, то размещением (сочетанием) из n элементов по k называется любое упорядоченное (неупорядоченное) подмножество из k элементов множества Ω. При $k = n$ размещение называется перестановкой из n элементов.

Пусть, например, дано множество $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Размещениями из трех элементов этого множества по два являются $(\omega_1, \omega_2), (\omega_1, \omega_3), (\omega_2, \omega_3), (\omega_2, \omega_1), (\omega_3, \omega_1), (\omega_3, \omega_2)$; сочетаниями: $(\omega_1, \omega_2), (\omega_1, \omega_3), (\omega_2, \omega_3)$. Два сочетания отличаются друг от друга хотя бы одним элементом, а размещения отличаются либо самими элементами, либо порядком их следования.

Число сочетаний из n элементов по k вычисляется по формуле

$$C_n^k = \frac{A_n^k}{P_k} = \frac{n!}{k!(n-k)!},$$

где $A_n^k = \frac{n!}{(n-k)!} = n(n-1)...(n-k+1)$ –

число размещений из n элементов по k, а $P_k = k!$ – число перестановок из k элементов. Справедливость соотношения $A_n^k = k! C_n^k$ следует из того, что число всех k-элементных подмножеств множества Ω равно C_n^k и каждое такое подмножество можно упорядочить $k!$ способами.

Рассмотрим перестановки с повторениями. Пусть из элементов $\omega_1, \omega_2, ..., \omega_i$ образуются конечные последовательности, содержащие n членов, в которых ω_1 повторяется k_1 раз, ω_2 – k_2 раза, …, $\omega_i - k_i$ раз, $k_1 + k_2 + ... + k_i = n$. Такие последовательности называются перестановками с повторениями. Две перестановки считаются одинаковыми, если они совпадают порядком расположения элементов и считаются различными, если у них различный порядок расположения элементов. Число различных перестановок с повторениями равно

$$\mathrm{P}_n(k_1, k_2, ..., k_i) = \frac{n!}{k_1! k_2! ... k_i!}.$$

Пример 1.2. Какова вероятность того, что из шести отмеченных чисел в карточке «Спортлото» (игра 6 из 49) k чисел будут выигрышными, $k = \overline{0, 6}$.

Решение. В данном примере эксперимент состоит в том, что случайным образом отмечаются 6 чисел из 49 в карточке «Спортлото». Поэтому равновозможными элементарными событиями будут наборы из шести отмеченных чисел. Т.к. для определения того, произойдет или не произойдет событие A, – среди отмеченных чисел k чисел выигрышные, – порядок чисел не существен, то в качестве равновозможных элементарных событий достаточно рассматривать неупорядоченные наборы 6 чисел из 49. Следовательно, число равновозможных элементарных событий равно C_{49}^6. Событие A состоит из наборов 6 чисел, k из которых выигрышные, а $6 - k$ проигрышные. Набор из k выигрышных чисел можно выбрать C_6^k способами, а набор $6 - k$ проигрышных чисел можно выбрать C_{43}^{6-k} способами. Тогда по основному принципу комбинаторики набор из k выигрышных и $6-k$ проигрышных чисел можно выбрать $C_6^k C_{43}^{6-k}$ способами, следовательно,

$$P(A)=\frac{C_6^k C_{43}^{6-k}}{C_{49}^6}.$$

Например, для $k=6$ имеем $P(A)\approx\left(14\cdot 10^6\right)^{-1}$.

Геометрическая вероятность является расширением понятия классической вероятности на случай несчетного множества элементарных событий. В случае, когда Ω – несчетное множество, вероятность определяется не на элементарных событиях, а на их множествах.

Пусть равновозможные элементарные события ω являются точками Ω – ограниченного множества n-мерного евклидова пространства, имеющего меру Лебега $\mu(\Omega)$. Рассмотрим систему А измеримых по Лебегу подмножеств Ω. Для любого случайного события $A\in$ А его вероятностью назовем число

$$P(A)=\frac{\mu(A)}{\mu(\Omega)},$$

где $\mu(A)$ – мера Лебега – действительная, неотрицательная, счетно-аддитивная функция множеств, т.е. такая, что а) $\mu(A)\geq 0$, б) $\mu(\varnothing)=0$, в) $A\subseteq B\Rightarrow \mu(A)\leq\mu(B)$,

г) $A\bigcap B=\varnothing\Rightarrow\mu(A\bigcup B)=\mu(A)+\mu(B)$ (в частных случаях – это длина, площадь, объем). Это определение называют геометрическим определением вероятности.

Пример 1.3. На одной стороне ленты магнитофонной кассеты длиной 10 м записан гимн студентов-математиков Беларуси длиной 2 м, а на другой – 14-я соната Бетховена длиной 3 м, причем их местоположение неизвестно. Случайным образом с обеих сторон ленты был поврежден (стерт) участок длиной 1 м, начинающийся на расстоянии 5 м от начала ленты. Найти вероятности следующих событий: A={гимн и соната не повреждены}, B={гимн поврежден, а соната – нет}, C={соната повреждена, а гимн – нет}, D={и гимн и соната повреждены}.

Решение. Из того, что положение гимна и сонаты совершенно неизвестно, делаем предположение, что любое положение начала каждого из них, при котором они вмещаются на соответствующих сторонах ленты, столь же правдоподобно, как любое другое. Пусть x – абсцисса начала записи гимна, y – сонаты. Пространство элементарных событий

$$\Omega = \{(x, y) : 0 \le x \le 10 - 2,\ 0 \le y \le 10 - 3\} -$$

прямоугольник с площадью $\mu(\Omega) = 8 \times 7 = 56$ м^2. На рис. 1 разными видами штриховки отмечены области, соответствующие повреждению гимна и сонаты, буквами $A,\ B,\ C,\ D$ – области, соответствующие каждому из событий $A,\ B,\ C,\ D$. Используя геометрическое определение вероятности, можно легко найти вероятности этих событий, например,

$$P(D) = \frac{\mu(D)}{\mu(\Omega)} = \frac{3 \cdot 4}{56} = \frac{3}{14}.$$

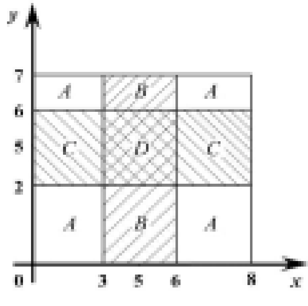


Рис. 1.

До сих пор мы рассматривали случаи, когда мера множества Ω была ограниченной и все элементарные события равновозможны. Рассмотрим теперь определение вероятности, свободное от этих ограничений. При этом в основу полагается идея определения вероятности как неотрицательной, нормированной и счетно-аддитивной функции множеств, являющихся событиями, т.е. $P(A)$, где $A \subseteq \Omega$. Совокупность подмножеств из Ω, на которых может быть определена вероятностная функция, должна быть построена специальным образом.

События, имеющие вероятность, образуют σ-алгебру, т.е. множество F подмножеств Ω, удовлетворяющее следующим свойствам:

а) $\text{Ù} \in \mathsf{F}$; б) $A \in \mathsf{F} \Rightarrow \overline{A} = \text{Ù} \setminus \text{Á} \in \mathsf{F}$;

б) $A_k \in \mathsf{F}$, $k = 1,2,... \Rightarrow \bigcup\limits_{k=1}^{\infty} A_k \in \mathsf{F} \left(\bigcap\limits_{k=1}^{\infty} \mathrm{A}k \in \mathsf{F} \right)$.

Простейшими примерами σ-алгебр являются: тривиальная σ-алгебра F, состоящая из двух событий $(\Omega, \varnothing)$, где $\varnothing$ – невозможное событие; σ-алгебра F, состоящая из всех событий.

Пример 1.4.

Пусть $\text{Ù} = \{\text{ù} : \text{ù} \in [0,1]\}$, $\text{Á} = \left\{\text{ù} : \text{ù} \in \left[0, \frac{2}{3}\right]\right\}$, $\text{Â} = \left\{\text{ù} : \text{ù} \in \left[\frac{1}{3}, 1\right]\right\}$.

Описать σ-алгебру событий F на Ù, порожденную событиями A и B.

Решение. Используя определение σ-алгебры получаем, что σ-алгебру событий F образуют следующие события (т.к. здесь $\text{Á} \cup \text{Â} = \text{Ù}$, $\overline{A} \cap \overline{\text{Â}} = \varnothing$):

$$\varnothing,\ \text{Ù} = \{\text{ù} : \text{ù} \in [0,1]\},\ \text{Á} = \left\{\text{ù} : \text{ù} \in \left[0, \frac{2}{3}\right]\right\},\ \text{Â} = \left\{\text{ù} : \text{ù} \in \left[\frac{1}{3}, 1\right]\right\},$$

$$\overline{\acute{A}} = \left\{ \grave{u} : \grave{u} \in \left(\frac{2}{3}, 1 \right] \right\}, \ \overline{\hat{A}} = \left\{ \grave{u} : \grave{u} \in \left[0, \frac{1}{3} \right) \right\}, \ \acute{A} \bigcap \hat{A} = \left\{ \grave{u} : \grave{u} \in \left[\frac{1}{3}, \frac{2}{3} \right] \right\},$$

$$\overline{\acute{A}} \bigcup \overline{\hat{A}} = \left\{ \grave{u} : \grave{u} \in \left\{ \left[0, \frac{1}{3} \right) \bigcup \left(\frac{2}{3}, 1 \right] \right\} \right\}.$$

Аксиоматическое определение вероятности. Вероятностью P, определенной на σ-алгебре событий F, называется числовая функция $P(A)$, $A \in \mathsf{F}$, удовлетворяющая следующим аксиомам:

1. $\forall \ A \in \mathsf{F} \ P(A) \geq 0$.
2. $P(\grave{U}) = 1$.
3. $\acute{A}_i \bigcap \acute{A}_k = \varnothing$, $i \neq k \Rightarrow P\left(\bigcup_{k=1}^{\infty} \acute{A}_k \right) = \sum_{k=1}^{\infty} P(\acute{A}_k)$.

Из данного определения следует, что $0 \leq P(A) \leq 1$, и вероятность является нормированной мерой, т.е. мерой, для которой выполняется условие нормировки 2.

Определение. Совокупность (Ω, F, P) называется вероятностным пространством.

Эквивалентным аксиоме 3 является требование аддитивности для конечного множества событий A_k и следующие аксиомы непрерывности (та или иная):

а) пусть случайные события $B_1, B_2, \ldots, B_n, \ldots \in \mathsf{F}$ таковы, что

$$B_{k+1} \subset B_k, \ \bigcap_{k=1}^{\infty} B_k = B, \text{ тогда } \lim_{n \to \infty} P(B_n) = P(B);$$

б) пусть $B_1, B_2, \ldots, B_n, \ldots \in \mathsf{F}$ и $B_k \subset B_{k+1}$, $\bigcup_{k=1}^{\infty} B_k = B$, тогда также $\lim_{n \to \infty} P(B_n) = P(B)$.

Можно показать, что из аксиомы 3 при выполнении вышеуказанного условия следует аксиома непрерывности б) и наоборот. Пусть, например, справедлива аксиома непрерывности б). Предположим, что $A_k \in$ F, $k = 1,2,\ldots,n,\ldots$ и $A_k \cap A_j = \varnothing,\ k \neq j$. Определим $B_n = \bigcup_{k=1}^{n} A_k$; ясно, что $B_n \subseteq B_{n+1}$. Определим также

$$B = \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty}\left(\bigcup_{k=1}^{n} A_k\right) = \bigcup_{k=1}^{\infty} A_k .$$

Из аксиомы непрерывности б) следует, что

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = P(B) = \lim_{n\to\infty} P(B_n) = \lim_{n\to\infty} P\left(\bigcup_{k=1}^{n} A_k\right) = \lim_{n\to\infty}\sum_{k=1}^{n} P(Á_n) = \sum_{k=1}^{\infty} P(Á_n),$$

откуда вытекает аксиома 3.

Приведем свойства вероятности:

а) $P(\varnothing) = 0$, следует из того, что $\varnothing \cup Ù = Ù$ и

$$P(\varnothing \cup Ù) = P(\varnothing) + P(Ù);$$

б) $P(\overline{A}) = 1 - P(Á)$, поскольку $Á \cup \overline{A} = Ù$ и

$$P(Á \cup \overline{A}) = P(Á) + P(\overline{A}) = P(Ù) = 1;$$

в) если $Á \subseteq Â$, то $P(Á) \leq P(B)$, следует из определения меры;

г) $P(Á \cup Â) = P(Á) + P(Â) - P(Á \cap B)$, поскольку

$$Á \cup Â = Á \cup (Â \setminus (Á \cap Â));$$

д) $P(Á \cup Â) \leq P(Á) + P(Â)$, следует из предыдущего свойства; равенство будет, если $Á \cap Â = \varnothing$.

§2. УСЛОВНАЯ ВЕРОЯТНОСТЬ И НЕЗАВИСИМОСТЬ СОБЫТИЙ

Определение. Пусть задано вероятностное пространство (Ω, F, P) и A, $B \in \mathsf{F}$. Предположим, что $P(B) > 0$. Условной вероятностью события A при условии, что произошло событие B, называется отношение вероятности $P(A \cap B)$ к $P(B)$ и обозначается $P(A / B)$, т.е.

$$P(A / B) = \frac{P(A \cap B)}{P(B)}.$$

Она обладает следующими свойствами:

а) если $B \subseteq A$, то $P(A/B) = 1$, т.к. в этом случае $A \cap B = B$;

б) если $A \cap B = \varnothing$, то $P((A \cup B)/ C) = P(A/C) + P(B / C)$, т.к.
$(\acute{A} \cap C) \cap (B \cap C) = \varnothing, P((\acute{A} \cup \hat{A}) \cap C) = P(A \cap C) + P(B \cap C)$;

в) $P(\overline{A}/B) = 1 - P(A/B)$, т.к.

$$1 = P(\Omega/B) = P((A \cup \overline{A})/B) = P(A/B) + P(\overline{A}/B).$$

Пример 1.5. Бросается игральный кубик. Какова вероятность того, что выпало число очков большее трех (событие A), если известно, что выпала четная грань (событие B)?

Решение. Событию B соответствует выпадение чисел 2, 4, 6; событию A – выпадение чисел 4, 5, 6; событию $A \cap B$ – 4, 6. Поэтому

$$P(A / B) = \frac{P(A \cap B)}{P(B)} = \frac{2}{6} : \frac{3}{6} = \frac{2}{3}.$$

Определение. Два события A и B называются независимыми, если $P(A \cap B) = P(A)P(B)$.

Определение. Два события $A_1, A_2, ..., A_n$ называются независимыми в совокупности, если для любых наборов индексов $\{i_1, i_2, ..., i_m\}$ таких, что $1 \le i_1 < i_2 < ... < i_m \le n$, $2 \le m \le n$, имеет место равенство

$$P\left(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_m}\right) = \prod_{k=1}^{m} P\left(A_{i_k}\right).$$

Если это равенство справедливо только для случая, когда $m = 2$, то события называются попарно независимыми.

Пример 1.6. Бросаются две монеты. Пусть событие A состоит в выпадении герба на первой монете, событие B состоит в выпадении цифры на второй монете; событие C – монеты выпадут разными сторонами. В этом случае:

$$\Omega = \{\text{ГЦ}, \text{ГГ}, \text{ЦГ}, \text{ЦЦ}\},\ A = \{\text{ГЦ}, \text{ГГ}\},\ B = \{\text{ГЦ}, \text{ЦЦ}\},\ C = \{\text{ГЦ}, \text{ЦГ}\},$$
$$AB = \{\text{ГЦ}\};$$

$$P(AB) = \frac{1}{4} = P(A)P(B);\ P(AC) = \frac{1}{4} = P(A)P(C);$$

$$P(BC) = \frac{1}{4} = P(B)P(C);\ P(ABC) = \frac{1}{4} \ne P(A)P(B)P(C) = \frac{1}{8},$$

т.е. A, B, C – попарно независимы (но независимости в совокупности здесь нет).

Следует отметить, что на практике независимость событий проверяется не из определения, а исходя из условий эксперимента: можно показать, что если события связаны с независимыми экспериментами, то и сами события будут независимыми.

Справедливо следующее утверждение, известное как **теорема умножения вероятностей**. *Пусть* $A_k \in \mathcal{F}$,

$$k = 1, 2, ..., n;\ P\left(\bigcap_{k=1}^{m} A_k\right) > 0$$ *для всех* $1 \le m \le n$, *тогда*

$$P\left(\bigcap_{k=1}^{n}\acute{A}_k\right)=P(\acute{A}_1)\prod_{m=1}^{n-1}P\left(\acute{A}_{m+1}/\bigcap_{k=1}^{m}\acute{A}_k\right).$$

Пример 1.7. Абонент забыл последнюю цифру номера телефона и поэтому набирает ее наугад. Определить вероятность того, что ему придется звонить не более чем три месяца.

Решение. Обозначим через A_i событие, заключающееся в том, что абонент звонит i-й раз и ему соответствует неудача, $i=\overline{1,3}$. Тогда имеем:

$$P(A_1)=1-\frac{1}{10}=\frac{9}{10},\ P(A_2/A_1)=\frac{8}{9},\ P(A_3/A_1A_2)=\frac{7}{8}.$$

Искомая вероятность равна

$$1-P(A_1A_2A_3)=1-P(A_1)P(A_2/A_1)P(A_3/A_1A_2)=0{,}3.$$

Определение. Пусть задано вероятностное пространство (Ω,F,P). Совокупность событий $A_1,\ A_2,\ldots,\ A_n\in\ \mathsf{F}$, называется полной группой событий, если выполнены следующие условия:

а) $\bigcup_{k=1}^{n}A_k=\Omega$; б) $A_i\cap A_k=\varnothing$, $i\neq k$; в) $P(A_k)>0,\ k=\overline{1,n}$.

Если $A_1,\ A_2,\ \ldots,\ A_n$ образуют полную группу событий, то для любого события

$$B=B\cap\Omega=B\cap\left(\bigcup_{k=1}^{n}A_k\right)=\bigcup_{k=1}^{n}(B\cap A_k),$$

и поэтому

$$P(B)=P\left(\bigcup_{k=1}^{n}(B\cap A_k)\right)=\sum_{k=1}^{n}P(B\cap A_k)=\sum_{k=1}^{n}P(A_k)P(B/A_k).$$

Формула

$$P(B)=\sum_{k=1}^{n}P(A_k)P(B/A_k)$$

называется **формулой полной вероятности для случайных событий.**

Если $P(B)>0$, то

$$P(A_k \cap B)=P(B)P(A_k/B)=P(A_k)P(B/A_k).$$

Отсюда следует, что

$$P(A_k/B)=\frac{P(A_k)P(B/A_k)}{\sum\limits_{i=1}^{n}P(A_i)P(B/A_i)}, \quad k=\overline{1,n},$$

т.е. имеет место **формула Байеса для случайных событий:**

$$P(A_k/B)=\frac{P(A_k)P(B/A_k)}{\sum\limits_{i=1}^{n}P(A_i)P(B/A_i)}, \quad k=\overline{1,n}.$$

При решении задач удобно применять следующую формулировку: если событие $B \in \mathrm{F}$ может происходить только с одним из событий A_1, $A_2, ..., A_n$ образующих полную группу событий, то при $P(B)>0$ справедливы формула полной вероятности и формула Байеса.

Пример 1.8. Для контроля продукции из трех партий деталей взята на проверку одна деталь. Какова вероятность выявления бракованной продукции, если в одной партии 2/3 деталей – бракованные, а в двух других все доброкачественные.

Решение. Пусть $B=$\{взятая деталь – бракованная\}, $A_k=$\{деталь берется из k -й партии\}, $k=\overline{1,3}$. Тогда

$$P(A_k)=\frac{1}{3}, k=\overline{1,3}; \; P(B/A_1)=\frac{2}{3}, \; P(B/A_2)=P(B/A_3)=0$$

и поэтому $P(B)=\sum\limits_{k=1}^{3}P(A_k)P(B/A_k)=\frac{2}{9}$.

Пример 1.9. Прибор состоит из двух узлов; работа каждого узла необходима для работы прибора в целом. Надежность (вероятность безотказной работы в течение времени t) первого узла равна p_1, второго p_2. Прибор испытывался в течение времени t, в результате чего обнаружено, что он вышел из строя (отказал). Найти вероятность того, что отказал только первый узел, а второй исправен.

Решение. Пусть $A_1=${оба узла исправны}, $A_2=${первый узел отказал, а второй исправен}, $A_3=${первый узел исправен, а второй отказал}, $A_4=${оба узла отказали}. Эти события образуют полную группу событий. Найдем их вероятности.

$$P(A_1)=p_1p_2,\ P(A_2)=(1-p_1)p_2,\ P(A_3)=p_1(1-p_2),$$
$$P(A_4)=(1-p_1)(1-p_2).$$

Поскольку наблюдалось событие $B=$ {прибор отказал}, то

$$P(B/A_1)=0,\ P(B/A_2)=P(B/A_3)=P(B/A_4)=1.$$

По формуле Байеса находим:

$$P(A_2/B)=\frac{(1-p_1)p_2}{(1-p_1)p_2+p_1(1-p_2)+(1-p_1)(1-p_2)}=\frac{(1-p_1)p_2}{1-p_1p_2}.$$

§3. СХЕМА НЕЗАВИСИМЫХ ИСПЫТАНИЙ БЕРНУЛЛИ

Определение. Испытанием (экспериментом, опытом) называется последовательность из двух актов: 1) создание комплекса условий, 2) наблюдение появившегося события. Испытания называются независимыми, если наблюдаемые события являются независимыми.

Определение. Независимыми испытаниями Бернулли называются такие испытания, для которых вероятности появления событий в каждом испытании одинаковы и не меняются от испытания к испытанию.

Нас будет интересовать следующая задача. Пусть производятся n испытаний Бернулли. В каждом испытании возможно появления события A с вероятностью p и невозможно с вероятностью $q=1-p$. Нужно определить $P_n(m)$ – вероятность того, что в n испытаниях событие A появляется ровно m раз.

Результат n испытаний удобно описать набором букв длиной n, который состоит из букв A и B: $\omega=(A\ A\ B...\ A\ B\ A)$, где буква A означает, что в испытании появилось событие A, а B – что в испытании появилось противоположное событие $\overline{A}$. Каждый набор интересующих нас исходов содержит m букв A и $n-m$ букв B, поэтому все такие исходы имеют одинаковую вероятность $p^m q^{n-m}$. Разные наборы отличаются только размещением букв A и B, поскольку число случаев, в которых появляется событие A, фиксировано. Размещение букв A и B однозначно определяется выбором m элементов из n, что можно сделать C_n^m способами. Поэтому

$$P_n(m)=\mathrm{C}_n^m p^m q^{n-m}=\mathrm{C}_n^m p^m (1-p)^{n-m},\ m=\overline{0,n}.$$

Эта формула называется **формулой Бернулли**. Очевидно, что

$$\sum_{m=0}^{n} P_n(m)=1\cdot$$

Пример 1.10. В течение смены, которая длится время t, эксплуатируется n ПЭВМ. Каждая ПЭВМ имеет надежность (вероятность безотказной работы) p и выходит из строя независимо от других. Найти вероятность $P(A)$ того, что инженер-электроник, вызванный по окончании времени t для ремонта неисправных ПЭВМ, справится со своей задачей за время τ, если на ремонт каждой неисправной ПЭВМ ему требуется время τ_0.

Решение. Событие A равносильно тому, что число вышедших из строя ПЭВМ меньше, чем $l=[\tau/\tau_0]$, где $[\tau/\tau_0]$ – наибольшее целое число, которое меньше либо равно τ/τ_0. Поэтому

$$P(A)=\sum_{m=0}^{l} C_n^m (1-p)^m p^{n-m} \cdot$$

Когда число испытаний велико, для вычисления $P_n(m)$ можно пользоваться приближенными формулами, которые вытекают из предельной теоремы Пуассона и локальной предельной теоремы Муавра–Лапласа.

В частности, имеет место **предельная теорема Пуассона:** *если* $n \to \infty,\ p \to 0$*, так, что* $np \to ё,\ 0 < ё < \infty$*, то*

$$P_n(m) \xrightarrow[n\to\infty]{} \frac{ё^m}{m!} e^{-\lambda} \cdot$$

Она выполняется потому, что если положить $np = ё_n$, то

$$P_n(m)=\frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} =$$

$$=\frac{n(n-1)\ldots(n-m+1)}{m!}\left(\frac{ё_n}{n}\right)^m \left(1-\frac{ё_n}{n}\right)^{n-m} =$$

$$=\frac{ё_n^m}{m!}\left(1-\frac{ё_n}{n}\right)^n \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{m-1}{n}\right)\left(1-\frac{ё_n}{n}\right)^{-m} .$$

Далее путем предельного перехода при $n \to \infty$ получим утверждение теоремы.

Приближенная формула, которая следует из этой теоремы, имеет вид (при больших n и малых p)

$$P_n(m) \approx \frac{ё^m}{m!} e^{-\lambda},\ m=\overline{0,n} \cdot$$

Она применяется при решении задач, в основном, когда

$$ё_n = np \le 10.$$

Пример 1.11. Вероятность появления события в каждом испытании равна 0,001. Проведено 5000 испытаний. Найти вероятность, что событие в них произойдет не менее двух раз.

Решение. $\text{ё}_n = np = 5000 \cdot 0{,}001 = 5 < 10.$ Искомая вероятность равна

$$P\{m \geq 2\} = \sum_{m=2}^{n} P_n(m) = 1 - P_n(0) - P_n(1) = 1 - P_{5000}(0) - P_{5000}(1) =$$

$$= 1 - e^{-5} - 5e^{-5} = 1 - 6e^{-5} \approx 0{,}9596 \cdot$$

Заметим, что в данном примере $p \approx 0$, а найденная вероятность $P\{m \geq 2\} \approx 1$.

При $n \to \infty$ имеет место также **локальная предельная теорема Муавра–Лапласа**:

$$\frac{P_n(m)}{\frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{x_m^2}{2}}} \xrightarrow[n \to \infty]{} 1,$$

$$\textit{где} \quad x_m = \frac{m - np}{\sqrt{np(1-p)}}, \ -\infty < a \leq x_m \leq b < +\infty.$$

Из нее при больших n вытекает следующая приближенная формула

$$P_n(m) \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{x_m^2}{2}} .$$

На практике ею обычно пользуются, когда $\text{ё}_n = np > 10$. Она даёт хорошие приближения при $p \approx \frac{1}{2}$ и часто используется, когда $n > 100$, $np(1-p) > 20$.

Интегральная предельная теорема Муавра–Лапласа имеет вид:

$$P\left(a \le \frac{m-np}{\sqrt{np(1-p)}} \le b\right) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{t^2}{2}}\,dt.$$

Для того чтобы показать, что она имеет место, можно воспользоваться предыдущей теоремой, из которой следует: $\forall\ \varepsilon > 0\ \exists\ n_\varepsilon$, такое, что $\forall\, n \ge n_\varepsilon$

$$\left|\frac{P_n(m)}{\frac{1}{\sqrt{2\pi np(1-p)}}e^{-\frac{x_m^2}{2}}}\right| < \varepsilon\,,$$

т.е.

$$(1-\varepsilon)\frac{1}{\sqrt{2\pi}}e^{-\frac{x_m^2}{2}}\frac{1}{\sqrt{np(1-p)}} < P_n(m) < (1+\varepsilon)\frac{1}{\sqrt{2\pi}}e^{-\frac{x_m^2}{2}}\frac{1}{\sqrt{np(1-p)}}.$$

Пусть $-\infty < a \le b < +\infty, \sum^* = \sum\limits_{m:a\le\frac{m-np}{\sqrt{np(1-p)}}\le b}$, тогда, учитывая, что

$$\sum{}^* P_n(m) = P\left(a \le \frac{m-np}{\sqrt{np(1-p)}} \le b\right),\ \Delta x_m = x_{m+1} - x_m = \frac{1}{\sqrt{np(1-p)}}\,,$$

можно записать

$$(1-\varepsilon)\frac{1}{\sqrt{2\pi}}\sum{}^* e^{-\frac{x_m^2}{2}}\Delta x_m < P\left(a \le \frac{m-np}{\sqrt{np(1-p)}} \le b\right) < (1+\varepsilon)\frac{1}{\sqrt{2\pi}}\sum{}^* e^{-\frac{x_m^2}{2}}\Delta x_m.$$

Если $n \to \infty$, то $\Delta x_m \to 0$, поэтому при $n \to \infty$:

$$(1-\varepsilon)\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{t^2}{2}}dt < P\left(a \le \frac{m-np}{\sqrt{np(1-p)}} \le b\right) < (1+\varepsilon)\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{t^2}{2}}dt \cdot$$

Из данной теоремы вытекает приближенная формула

$$P\left(a \le \frac{m-np}{\sqrt{np(1-p)}} \le b\right) \approx \Phi(b) - \Phi(a),$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$ – интеграл Лапласа;

$$\Phi(0) = 0; \ \Phi(-x) = -\Phi(x); \ \Phi(x) \approx 0{,}5, \ x \ge 5.$$

Таблица значений функции $\Phi(x)$ приведена в Приложении 2.

Пример 1.12. Театр, вмещающий 1000 зрителей, имеет два входа. У каждого входа имеется свой гардероб. Сколько мест должно быть в каждом из гардеробов, чтобы в среднем в 99 случаях из 100 все зрители могли раздеться в гардеробе того входа, через который они вошли? Предполагается, что зрители приходят парами и каждая пара независимо от других выбирает с вероятностью 0,5 любой из входов.

Решение. Пусть число мест в каждом гардеробе равно N, для его нахождения составим уравнение. Занумеруем гардеробы номерами 1 и 2. Выбор зрителями того или иного гардероба можно рассматривать как испытание Бернулли, в каждом из которых определенная пара с вероятностью 0,5 выбирает гардероб, например, №1.

По условию задачи $n = 500$, $p = 0{,}5$. Пусть событие A состоит в том, что зрители разденутся в гардеробе того входа, куда они зашли, m – число пар зрителей, выбравших гардероб №1. Событие A будет происходить, если $500 - \frac{N}{2} \le m \le \frac{N}{2}$. По условию $P(A) = 0{,}99$. Поэтому

$$0{,}99 = P(A) = P\left(500 - \frac{N}{2} \le m \le \frac{N}{2}\right) =$$

$$= P\left(\frac{500 - \frac{N}{2} - np}{\sqrt{np(1-p)}} \le \frac{m - np}{\sqrt{np(1-p)}} \le \frac{\frac{N}{2} - np}{\sqrt{np(1-p)}}\right) \approx$$

$$\approx \Phi\left(\frac{\frac{N}{2} - 250}{\sqrt{125}}\right) - \Phi\left(\frac{250 - \frac{N}{2}}{\sqrt{125}}\right) = 2\Phi\left(\frac{\frac{N}{2} - 250}{\sqrt{125}}\right), \text{ т.е. } \Phi\left(\frac{\frac{N}{2} - 250}{\sqrt{125}}\right) \approx 0{,}495.$$

С помощью таблицы для функции $\Phi(x)$ находим $\Phi(2{,}56) \approx 0{,}495,$ таким образом, $\dfrac{\frac{N}{2} - 250}{\sqrt{125}} \approx 2{,}56,$ откуда следует, что

$$N \approx 556.$$

Из интеграла предельной теоремы Муавра–Лапласа получаем

$$P\left(\left|\frac{m}{n} - p\right| \le \varepsilon\right) = P\left(-\varepsilon \le \frac{m - np}{n} \le \varepsilon\right) =$$

$$= P\left(\frac{-\varepsilon\sqrt{n}}{\sqrt{p(1-p)}} \le \frac{m - np}{\sqrt{np(1-p)}} \le \frac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}}\right) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}\, dt = 1,$$

т.е. $P\left(\left|\frac{m}{n} - p\right| \le \varepsilon\right) \xrightarrow[n\to\infty]{} 1 \quad \forall \varepsilon > 0.$ Последнее соотношение носит название ***закона больших чисел в форме Бернулли***. Из него следует, что при большом числе испытаний частота появления события почти не отличается от вероятности этого события.

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ И ИХ РАСПРЕДЕЛЕНИЯ

Пусть $\mathcal{F}$– σ-алгебра множеств пространства Ω. Пара $(\Omega, \mathcal{F})$ называется измеримым пространством. Отображение ξ одного измеримого пространства в другое называется измеримой функцией, т.е. если $\omega \in \Omega$ и $\xi(\omega) \in R$, то $\xi\,(.)\colon (\Omega, \mathcal{F}) \to (R, \mathcal{B}\,)$ Здесь $\mathcal{B}$ – σ-алгебра множеств пространства R. Другими словами, $\xi(\omega)$ – измеримая функция, если для любого борелевского множества B $\{\omega{:}\xi(\omega) \in B\} \subset \mathcal{F}$. Пусть теперь $(\Omega, \mathcal{F}, P)$ – вероятностное пространство, R – числовая ось, $\mathcal{B}$ – σ-алгебра, порожденная интервалами на числовой оси (эту σ-алгебру называют системой борелевских множеств или борелевским полем множеств).

Определение. Измеримая функция $\xi(\omega)$, т.е. такая, что $\forall\, B \in \mathcal{B}$ $\{\omega{:}\xi(\omega) \in B\} \subset \mathcal{F}$ называется случайной величиной (СВ).

Пример 1.13. Пусть Ω – множество студентов на факультете. Каждый отдельный студент – элемент $\omega \in \Omega$. Определим на элементах ω функцию $\xi(\omega)$, которая принимает значения, равные году рождения студента, который является элементом ω. Таким образом определенная функция является случайной величиной (имеется в виду, что, кроме Ω, задана вероятностная функция $P(\omega)$).

Пример 1.14. По промежуткам времени безотказной работы приборы делятся на несколько типов, например, первый, второй, третий. Пусть $\Omega(\omega)$ – множество значений, которые могут принимать промежутки времени безотказной работы прибора, а $\xi(\omega)$ – номер типа, который присваивается прибору с промежутком безотказной работы ω. Тогда $\xi(\omega)$ является СВ.

Определение. Действительная функция $f(x),\ x\in R$, называется борелевской, если прообраз $f^{-1}(Â)=\{x{:}f(x)\in Â\}$ любого борелевского множества B числовой прямой является борелевским множеством.

Пример 1.15. Покажем, что если $î(ù)$ – СВ, а $f(x)$ – борелевская функция, то $\eta(\omega)=f(\xi(\omega))$ – также СВ. Пусть (Ω,F,P) – вероятностное пространство, на котором задана $î(ù)$ и В – $ó$-алгебра борелевских множеств прямой, $B\in$ $â$. Т.к. $f(x)$ – борелевская функция, то $f^{-1}(B)=B_1\in$ $â$. Но $\eta^{-1}(B)=\xi^{-1}(B_1)\in\mathsf{F}$ и, следовательно, $ç(ù)$ – СВ.

Определение. Пусть $Â_x=(-\infty,x]$. Функция

$$F_î(x)=P\{ù:î(ù)\in Â_x\}=P\{ù:î(ù)\le x\}=P\{î\le x\}$$

называется функцией распределения (ф.р.) случайной величины $î(ù)$.

Теорема 1.1. *(свойства ф.р.). 1) $F_î(x)$ – монотонно неубывающая функция; 2) $\lim\limits_{x\to-\infty}F_î(x)=0$, $\lim\limits_{x\to+\infty}F_\xi(x)=1$; 3) $F_î(x)$ – непрерывная справа функция.*

Доказательство.

1) Поскольку $\{ù:î(ù)\le x_1\}\le\{ù:î(ù)\le x_2\}$ для $x_1\le x_2$, то

$$P\ \{ù:î(ù)\le x_1\}\le P\{ù:î(ù)\le x_2\}\Rightarrow F_î(x_1)\le F_î(x_2).$$

2) $\{ù:î(ù)\le x\}\xrightarrow[x\to-\infty]{}\varnothing\Rightarrow$ по аксиоме непрерывности $P\{ù:î(ù)\le x\}\xrightarrow[x\to-\infty]{}P(\varnothing)=0$; $\{ù:î(ù)\le x\}\xrightarrow[x\to+\infty]{}Ù\Rightarrow$ по аксиоме непрерывности $P\{ù:î(ù)\le x\}\xrightarrow[x\to+\infty]{}P(Ù)=1.$

3)Пусть
$x_1>x_2>\ldots,\ \lim\limits_{k\to\infty}x_k=x,\ \{ù:î(ù)\le x_k\}=Á_k,\ \{ù:î(ù)\le x\}=Á,$тогда $Á=\bigcap\limits_k Á_k\Rightarrow$

по аксиоме непрерывности $P(A) = F_\xi(x) = \lim\limits_{x_k \to x} P(A_k) = \lim\limits_{x_k \to x} F_\xi(x_k)$.

Теорема 1.2. *Если функция $F(x)$ обладает свойствами 1)–3), то существует вероятностное пространство $(\Omega, \mathcal{F}, P)$ и СВ $\xi(\omega)$, определенная на нем, такая, что $F_\xi(x) = F(x)$.*

Рассмотрим дискретное вероятностное пространство, в этом случае пространство элементарных событий состоит из дискретного множества элементарных событий (счетного или конечного)

$$\Omega = \{\omega_1, \omega_2, ..., \omega_k, ...\}.$$

Пусть $x_k = \xi(\omega_k)$, $\xi(\omega) \in \{x_1, x_2, ..., x_k, ...\}$, $A_k = \{\omega : \xi(\omega) = x_k\}$, $k = 1, 2, ...$. Случайные величины, которые могут принимать только конечное или счетное множество значений, называются дискретными. Для их описания удобно пользоваться набором вероятностей $\{p_1, p_2, ..., p_k, ...\}$, где $p_k = P(A_k) = P\{\omega : \xi(\omega) = x_k\}$, который называется распределением вероятностей дискретной СВ $\xi(\omega)$. Поскольку

$$\bigcup_{k=1}^{\infty} A_k = \Omega, \; A_i \cap A_j = \varnothing, \; i \neq j,$$

то $P\left(\bigcup\limits_{k=1}^{\infty} A_k\right) = \sum\limits_{k=1}^{\infty} P(A_k) = P(\Omega) = \sum\limits_{k=1}^{\infty} p_k = 1$ (условие нормировки).

Совокупность

$$\begin{matrix} x_1, x_2, ..., x_k, ... \\ p_1, p_2, ..., p_k, ... \end{matrix}$$

называется дискретным законом (рядом) распределения вероятностей.

Установим связь между распределением вероятностей и функцией распределения:

$$F_\xi(x) = P\{\omega : \xi(\omega) \le x\} = P\left(\bigcup_{\{k : x_k \le x\}} \{\omega : \xi(\omega) = x_k\}\right) =$$

$$= \sum_{\{k: x_k \le x\}} P\{\omega : \xi(\omega) = x_k\} = \sum_{\{k: x_k \le x\}} p_k,$$

$p_k = F_\xi(x_k) - F_\xi(x_{k-1}),$ если считать, что $F_\xi(x_0) = 0.$

В связи с тем, что свойства СВ полностью определяются свойствами их функций распределения, их принято классифицировать по характеру этих функций.

I. Дискретные СВ (ф.р.). В этом случае множество значений $\xi(\omega)$: $x_1, x_2, ..., x_k, ..$ – счетно либо конечно, $F_\xi(x) = \sum_{\{k: x_k \le x\}} p_k$; ф.р. обладает, кроме основных, следующими свойствами: 1) $F_\xi(x)$ имеет конечное или счетное множество точек разрыва первого рода, 2) если x – точка непрерывности $F_\xi(x)$, то

$$\exists \ \frac{dF_\xi(x)}{dx} \ \text{и} \ \frac{dF_\xi(x)}{dx} = 0.$$

Примеры дискретных распределений CB:

1) СВ имеет распределение Бернулли, если $\xi(\omega) \in \{0,1\}$,

$p_k = P\{\omega : \xi(\omega) = k\} = p^k (1-p)^{1-k}, k = 0, 1, \ 0 < p < 1,$ рис. 2,

$$F_\xi(x) = \begin{cases} 0, x < 0, \\ 1 - p, 0 \le x < 1, \\ 1, x \ge 1, \end{cases} \text{ рис. 3.}$$

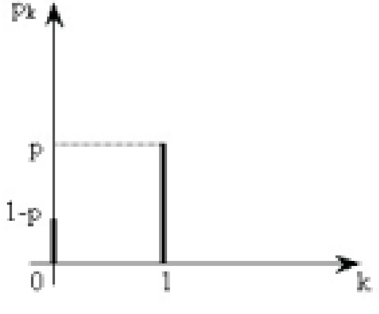


Рис. 2

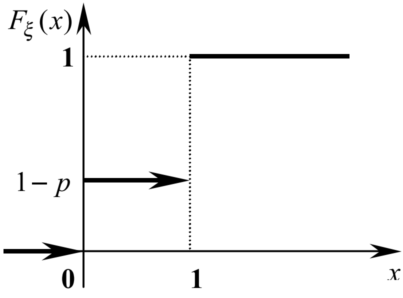


Рис. 3

2) СВ $\xi(\omega)$ имеет биномиальное распределение, если

$$\xi(\omega) \in \{0, 1, 2, \ldots, n\}$$

$$p_k = P\{\omega : \xi(\omega) = k\} = C_n^k p^k (1-p)^{n-k}, k = 0, 1, 2, \ldots, n, \; 0 < p < 1,$$

$$F_\xi(x) = \begin{cases} \sum\limits_{k=0}^{l} C_n^k p^k (1-p)^{n-k}, \; l \le x < l+1, \\ 1, x \ge n, \\ 0, x < 0, \end{cases}$$

отметим, что, как следует из формулы Бернулли, число появлений события в n независимых испытаниях Бернулли имеет биномиальное распределение;

3) СВ $\xi(\omega)$ имеет геометрическое распределение, если

$$\xi(\omega) \in \{0, 1, 2, \ldots k, \ldots\},$$

$$p_k = P\{\omega : \xi(\omega) = k\} = p(1-p)^k, k = 0, 1, 2, \ldots, 0 < p < 1,$$

$$F_\xi(x) = \begin{cases} 0, x < 0 \\ \sum\limits_{k=0}^{l} p(1-p)^k, l \le x < l+1 \end{cases} = \begin{cases} 0, x < 0, \\ 1-(1-p)^l, l \le x < l+1; \end{cases}$$

4) *СВ* $\xi(\omega)$ имеет распределение Пуассона с параметром $\lambda > 0$, если $\xi(\omega) \in \{0, 1, 2, \ldots k, \ldots\}$,

$$p_k = P\{\omega : \xi(\omega) = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, 2, \ldots, \text{ рис. 4,}$$

$$F_\xi(x) = \begin{cases} 0, x < 0, \\ e^{-\lambda} \sum_{k=0}^{l} \frac{\lambda^k}{k!}, l \le x < l+1, \end{cases} \text{ рис. 5;}$$

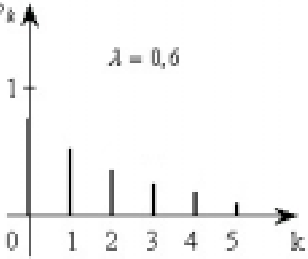


Рис. 4

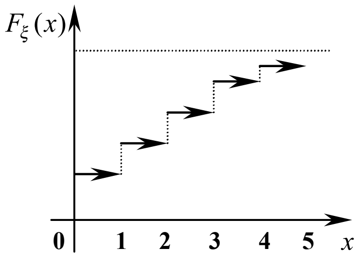


Рис. 5

Дадим интерпретацию некоторых из этих СВ. Предположим, что студент идет сдавать зачет. На некоторые вопросы по сдаваемому предмету он знает ответы, а на остальные – нет. Поэтому события, заключающиеся в том, что он получит зачет, являются случайными. Определим СВ следующим образом: если зачет сдан, то $\xi = 1$, если нет, то $\xi = 0$. Таким образом определенная СВ является бернуллиевой, параметр p в том случае соответствует относительному числу вопросов, на которые студент знает ответ. Пусть студенту необходимо сдать n зачетов и он делает по одной попытке получить каждый из этих зачетов. Определим СВ ξ как число зачетов, которые получит студент. Такая СВ будет биномиальной. Число студентов, которых успел выслушать преподаватель на зачете за фиксированный интервал времени, а также число заданий, которые выполняет ЭВМ за фиксированный промежуток времени, являются СВ, распределенными по закону Пуассона с соответственно определенными параметрами λ.

Пример 1.16. Производятся последовательные испытания трех приборов на надежность. Каждый следующий прибор испытывается только в том случае, если предыдущий оказался надежным. Составить таблицу распределения и найти ф.р. случайного числа испытанных приборов, если вероятность надежности каждого прибора равна $q = 1 - p = 0,9$.

Решение. СВ ξ , описывающая число испытанных приборов, имеет распределение вероятностей

$$p_k = P\{\xi = k\} = \begin{cases} p(1-p)^{k-1}, k = 1, 2, \\ (1-p)^{k-1}, k = 3, \end{cases}$$

поэтому таблица распределения имеет

k	1	2	3
p_k	0,1	$0,9 \cdot 0,1$	$(0,9)^2 \cdot 0,1$

$$\text{а ф.р. } F_\xi(x) = \begin{cases} 0, x < 1, \\ \sum_{i=1}^{k} p_i, k \le x < k+1, \\ 1, x \ge 3. \end{cases}$$

II. Непрерывные СВ (СВ с абсолютно непрерывными ф.р.).

В этом случае $F_\xi(x) -$ непрерывная функция и $F_\xi(x) = \int\limits_{-\infty}^{x} p_\xi(t)dt.$

Ясно, что $p_\xi(x) = \dfrac{dF_\xi(x)}{dx}$ в точках существования производной. Функция $p_\xi(x)$ называется плотностью распределения вероятностей СВ $\xi(\omega)$. Она обладает следующими свойствами:

1) $p_\xi(x) \ge 0$ (как производная неубывающей функции);

2) $\int\limits_{-\infty}^{\infty} p_\xi(x)dx = 1$ условие нормировки (следует из свойства ф.р. $F_\xi(+\infty) = 1)$;

3) кроме того,

$$P\{\omega : x_1 < \xi(\omega) \le x_2\} = F_\xi(x_2) - F_\xi(x_1) = \int\limits_{x_1}^{x_2} p_\xi(t)dt.$$

Рассмотрим примеры непрерывных СВ:

1) СВ $\xi(\omega)$ имеет равномерное распределение на отрезке $[a,b]$, если

$$p_\xi(x) = \begin{cases} \dfrac{1}{b-a}, x \in [a,b] \\ 0, x \notin [a,b] \end{cases}, \text{ рис. 6, } \quad F_\xi(x) = \begin{cases} 0, x < a, \\ \dfrac{x-a}{b-a}, a < x \le b, \\ 1, x > b; \end{cases} \text{ рис. 7,}$$

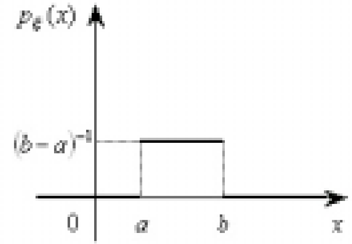


Рис. 6

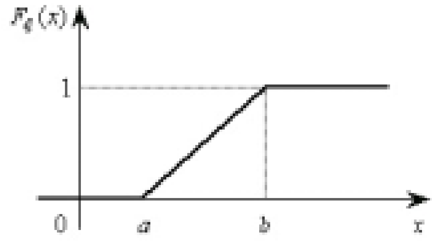


Рис. 7

2) СВ $\xi(\omega)$ имеет показательное (экспоненциальное) распределение с параметром $\lambda > 0$, если

$$p_\xi(x) = \begin{cases} 0, x < 0 \\ \lambda e^{-\lambda x}, x \ge 0 \end{cases}, \text{ рис. 8, } \quad F_\xi(x) = \begin{cases} 0, x < 0 \\ 1 - e^{-\lambda x}, x \ge 0 \end{cases}, \text{ рис. 9;}$$

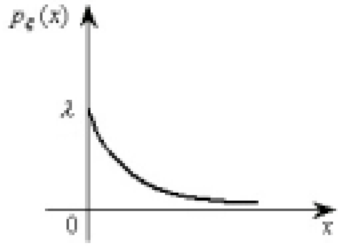


Рис. 8

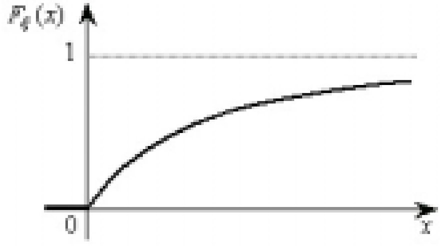


Рис. 9

3) СВ $\xi(\omega)$ имеет нормальное распределение с параметрами a, σ^2, $\xi(\omega) \sim N(a, \sigma^2)$, если

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \text{ рис. } 10,$$

$$F_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}} dt, \text{ рис. } 11;$$

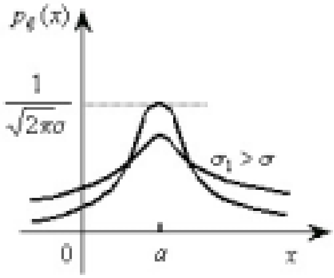


Рис. 10

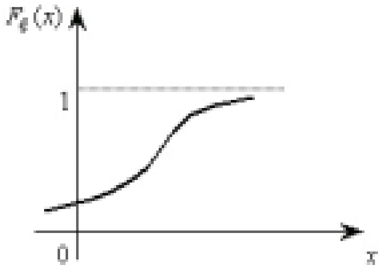


Рис. 11

4) СВ $\xi(\omega)$ имеет распределение Коши с параметром $a > 0$, если

$$p_\xi(x) = \frac{a}{\pi(a^2 + x^2)}, \; F_\xi(x) = \frac{1}{\pi} arctg \frac{x}{a}.$$

В частности, интервалы времени между соседними автомобилями на дорогах являются экспоненциально распределенными СВ с соответствующими параметрами $\lambda > 0$. Если $\xi(\omega)$ – СВ, имеющая экспоненциальное распределение, то $\forall ô \geq 0$

$$P(\xi > x+\tau / \xi > \tau) = \frac{P(\xi > x+\tau, \xi > \tau)}{P(\xi > \tau)} = \frac{P(\xi > x+\tau)}{P(\xi > \tau)} = \frac{1-P(\xi \leq x+\tau)}{1-P(\xi \leq \tau)} =$$

$$= \frac{1-1-e^{-\lambda(x+\tau)}}{1-1-e^{-\lambda\tau}} = e^{-\lambda x} = 1-F_{\xi}(x) = 1-P(\xi \leq x) = P(\xi > x) \Rightarrow$$

$$\Rightarrow P(\xi - \tau \leq x / \xi > \tau) = P(\xi \leq x),$$

следовательно, см. рис. 1.12, СВ $\xi - \tau$ имеет то же экспоненциальное распределение, что и СВ ξ .

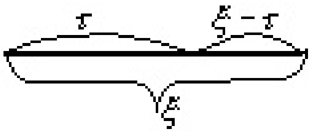


Рис. 12

Нормальное распределение является наиболее широко применяемым на практике непрерывным распределением. В частности, оно применяется при моделировании броуновского движения, которое мы рассмотрим в главах, посвященных случайным процессам.

Пример 1.17. Проверим, что функция $p(x) = \frac{1}{\sqrt{2ð}} e^{-\frac{x^2}{2}}$ является плотностью распределения вероятностей.

Решение. Ясно, что $f(x) > 0$. Нужно еще проверить условие нормировки. Это можно сделать несколькими способами: а) пользуясь равенством $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, где $\Gamma(k) = \int\limits_0^{\infty} t^{k-1} e^{-t} dt$ – гамма-функция;

б) возводя интеграл $\int\limits_{-\infty}^{\infty} f(x)dx$ в квадрат и переходя в двойном интеграле к полярным координатам.

Рассмотрим первый способ:

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = \frac{1}{\sqrt{\pi}}\int\limits_{0}^{\infty} e^{-\frac{x^2}{2}}\left(\frac{x^2}{2}\right)^{-\frac{1}{2}} d\left(\frac{x^2}{2}\right) =$$

$$= \left[\frac{x^2}{2} = t\right] = \frac{1}{\sqrt{\pi}}\int\limits_{0}^{\infty} t^{-\frac{1}{2}}e^{-t}dt = \frac{\Gamma\left(\frac{1}{2}\right)}{\sqrt{\pi}} = 1.$$

Второй способ дает:

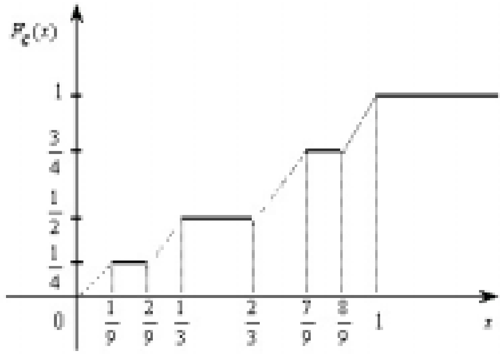
$$\left[\int\limits_{-\infty}^{\infty} f(x)dx\right]^2 = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}}dxdy =$$

$$= [x = r\cos\varphi,\ y = r\sin\varphi] = \frac{1}{2\pi}\int\limits_{0}^{2\pi} d\varphi \int\limits_{0}^{\infty} e^{-\frac{r^2}{2}} r dr = 1,$$

откуда следует тот же результат. Плотность $p(x)$ называют плотностью стандартного нормального распределения, у которого $a = 0,\ \sigma = 1$.

III. СВ с сингулярными ф.р. Кроме дискретных и непрерывных СВ, существуют другие СВ. В частности, кроме величин, которые на одних интервалах являются непрерывными, а на других дискретными, существуют величины, не являющиеся ни на одном интервале ни дискретными, ни непрерывными. К таким СВ относятся, например, те, ф.р. которых непрерывные, но при этом возрастают только на множестве лебеговской меры нуль. В качестве примера можно привести величину, ф.р. которой – известная кривая

Кантора. Кривая («лестница») Кантора строится следующим образом (рис. 13). В итоге ф.р. $F_\xi(x)$



$$F_\xi(x) = \frac{1}{2}, x \in \left(\frac{1}{3}, \frac{2}{3}\right],$$

$$F_\xi(x) = \frac{1}{4}, x \in \left(\frac{1}{9}, \frac{2}{9}\right],$$

$$F_\xi(x) = \frac{3}{4}, x \in \left(\frac{7}{9}, \frac{8}{9}\right],$$

• • •

и т.д. до бесконечности

Рис. 13

оказывается определенной на счетном множестве интервалов, которые являются интервалами смежности некоторого нигде не плотного совершенного множества меры нуль. На этом множестве доопределим функцию $F_\xi(x)$ по непрерывности. Величина ξ с таким образом определенной ф.р. не дискретна, поскольку ее ф.р. непрерывна, но в то же время не абсолютно непрерывна, т.к. ее ф.р. не является интегралом от своей производной.

Теорема Лебега. *Любую ф.р.* $F_\xi(x)$ *однозначно можно представить в виде (разложение Лебега)*

$$F_\xi(x) = a_1 F_1(x) + a_2 F_2(x) + a_3 F_3(x),$$

где $a_i \geq 0, i = \overline{1,3},\ a_1 + a_2 + a_3 = 1, F_1(x), F_2(x), F_3(x)$ – дискретная, абсолютно непрерывная и сингулярная ф.р. соответственно.

Покажем, что если $P\{\omega : \xi(\omega) = \overline{x}\} = 0,$ то $\overline{x}$ является точкой непрерывности функции $F_\xi(x)$. Действительно, $\forall \varepsilon > 0$

$$P\{\omega : \overline{x} - \varepsilon \leq \xi(\omega) \leq \overline{x} + \varepsilon\} \geq F_\xi(\overline{x} + \varepsilon) - F_\xi(\overline{x} - \varepsilon) \geq 0.$$

Отсюда имеем

$$P\{\omega:\xi(\omega)=\bar{x}\}=\lim_{\varepsilon\to 0}P\{\bar{x}-\varepsilon\le\xi(\omega)\le\bar{x}+\varepsilon\}=\lim_{\varepsilon\to 0}\left[F_\xi(\bar{x}+\varepsilon)-F_\xi(\bar{x}-\varepsilon)\right]=0.$$

Справедливо и обратное утверждение: если x – точка непрерывности функции $F_\xi(x)$, то $P\{\omega:\xi(\omega)=\bar{x}\}=0$.

Пусть $\xi(\omega)$ – СВ с абсолютно непрерывной ф.р. $F_\xi(x)$ и плотностью распределения $p_\xi(x)$, $y=f(x)$ – непрерывная возрастающая функция. Тогда для $CB\ \ \eta(\omega)=f(\xi(\omega))$

$$F_\eta(y)=P\{\eta\le y\}=P\{f(\xi)\le y\}=P\{\xi\le f^{-1}(y)\}=F_\xi(f^{-1}(y)),$$

и если, кроме того, $f(x)$ – дифференцируемая функция, то

$$p_\eta(y)=p_\xi\left(f^{-1}(y)\right)\frac{df^{-1}(y)}{dy}.$$

Если же функция $f(x)$ – убывающая, то

$$F_\eta(y)=1-F_\xi(f^{-1}(y)),\ \ p_\eta(y)=-p_\xi(f^{-1}(y))\frac{df^{-1}(y)}{dy}.$$

Пример 1.18. Случайная величина ξ имеет показательное распределение с параметром λ. Найти плотность распределения СВ $\eta=\sqrt{\xi}.$

Решение.

$$F_\eta(y)=P\{\sqrt{\xi}\le y\}=P\{\xi\le y^2\}=1-e^{-\lambda y^2},$$

$$p_\eta(y)=F_\eta^{'}(y)=2\lambda y e^{-\lambda y^2},\ y\ge 0.$$

Пример 1.19. Пусть СВ ξ равномерно распределена на отрезке [0,1]. Найти ф.р. и плотности распределения СВ: а) $\eta_1=\xi^2$, б) $\eta_2=\sqrt{\xi}$, в) $\eta_3=sin^2\left(\frac{\pi\xi}{2}\right)$.

Решение.

а) $F_{\eta_1}(x) = P\{\xi^2 \le x\} = P\{\xi \le \sqrt{x}\} = \sqrt{x},\ 0 \le x \le 1$,

$$p_{\eta_1}(x) = F'_{\eta_1}(x) = \frac{1}{2\sqrt{x}},\ 0 < x \le 1;$$

б) $F_{\eta_2}(x) = P\{\sqrt{\xi} \le x\} = P\{\xi \le x^2\} = x^2,\ p_{\eta_2}(x) = F'_{\eta_2}(x) = 2x,\ 0 \le x \le 1;$

в)

$$F_{\eta_3}(x) = P\left\{sin^2\left(\frac{\pi\xi}{2}\right) \le x\right\} = P\left\{sin\left(\frac{\pi\xi}{2}\right) \le \sqrt{x}\right\} = P\left\{\xi \le \frac{2}{\pi} arcsin\sqrt{x}\right\}, 0 \le x \le 1,$$

$$p_{\eta_3}(x) = F'_{\eta_3}(x) = \frac{1}{\pi\sqrt{x(1-x)}}, 0 < x < 1.$$

Пример 1.20. Найти плотности распределения СВ: а) $\eta_1 = \xi^3$, б) $\eta_2 = |\xi|$, если известна плотность распределения СВ ξ.

Решение.

а) $p_{\eta_1}(y) = p_{\hat{\imath}}(\sqrt[3]{y})(\sqrt[3]{y})' = \frac{1}{3}\sqrt[3]{y^2}\, p_{\hat{\imath}}(\sqrt[3]{y});$

б) $F_{\eta_2}(y) = P(\eta_2 \le y) = P(-y \le \hat{\imath} \le y) = \int_{-y}^{y} p_{\hat{\imath}}(x)dx = \int_{0}^{y}[p_{\hat{\imath}}(x) + p_{\hat{\imath}}(-x)]dx,\ y > 0,$

откуда следует, что $p_{\eta_2}(y) = p_{\hat{\imath}}(y) + p_{\hat{\imath}}(-y),\ y > 0,$ согласно определению плотности.

Пример 1.21. При проведении математических экспериментов на ЭВМ поведение построенной модели многократно наблюдают при различных случайных исходных условиях. Такой способ исследования называется методом статистических испытаний, или методом «Монте-Карло». При этом возникает задача получения случайных чисел, распределенных по любому какому угодно заданному закону. В ЭВМ эта задача решается при помощи функционального преобразования случайных чисел, распределенных равномерно в интервале [0,1], методы получения которых хорошо разработаны. Это делается следующим образом.

Пусть СВ $\xi(\omega)$ равномерно распределена на интервале [0,1]. Надо найти такое преобразование $f(x)$, чтобы СВ $\eta(\omega) = f(\hat{\imath}(\omega))$ имела заданную ф.р. $F(y)$. Т.к. $0 \le F(y) \le 1$, выберем $f(x)$ в виде $f(x) = F^{-1}(x)$. Рассмотрим СВ $\eta(\omega) = F^{-1}(\hat{\imath}(\omega))$. Для нее $p_\eta(y) = p_{\hat{\imath}}(F(y))\left|\dfrac{\partial F(y)}{\partial y}\right|$, но т.к. $\hat{\imath}(\omega)$ равномерно распределена на интервале [0,1], то $p_{\hat{\imath}}(F(y)) = 1$, и мы получаем $p_\eta(y) = \dfrac{\partial F(y)}{\partial y}$ (знак модуля здесь можно снять, т.к. $F(y)$ – неубывающая функция). Таким образом, $F_\eta(y) = F(y)$, что и требовалось доказать.

Модой дискретной СВ называется ее наиболее вероятное значение, **модой непрерывной СВ** ξ – значение аргумента x, при котором ее плотность распределения $p_\xi(x)$ максимальна. **Медианой СВ** ξ называется значение аргумента x, при котором $F_\xi(x) = 0{,}5$.

Рассмотрим теперь многомерные СВ. Пусть на некотором вероятностном пространстве (Ω, F, P) определены n измеримых функций СВ $\xi_1(\omega), \xi_2(\omega), ..., \xi_n(\omega)$. Совокупность этих функций $\xi(\omega)$ определяет отображение $(\Omega, \mathsf{F}) \to (R^n,\ \hat{a}^{\,n})$, где R^n – n-мерное действительное пространство, $\hat{a}^{\,n}$ – система борелевских множеств на R^n. Такая **совокупность** $\xi(\omega) = (\xi_1(\omega), \xi_2(\omega), ..., \xi_n(\omega))$ **называется многомерной СВ (или случайным вектором)**.

Функция n аргументов

$$F_\xi(x)=F_{\xi_1\xi_2\dots\xi_n}(x_1,x_2,\dots,x_n)=P\{\omega:\xi_1(\omega)\le x_1,\xi_2(\omega)\le x_2,\dots,\xi_n(\omega)\le x_n\}$$

называется n-мерной ф.р. n-мерной СВ.

Пример 1.22. Пусть $\xi_1(\omega)$ – СВ, равномерно распределенная на отрезке $[0,1]$, $\xi_2(\omega)=\xi_1^2(\omega)$. Тогда $\xi(\omega)=(\xi_1(\omega),\xi_2(\omega))$ – двумерная СВ. Найдем ее ф.р.

$$F_{\xi_1\xi_2}(x_1,x_2)=P\{\omega:\xi_1(\omega)\le x_1,\xi_2(\omega)\le x_2\}=P\{\omega:\xi_1(\omega)\le x_1,\xi_1^2(\omega)\le x_2\}$$

При $x_1<0,x_2<0$ это выражение равно 0, а при $x_1\ge 0,x_2\ge 0$

$$F_{\xi_1\xi_2}(x_1,x_2)=P\{\omega:\xi_1(\omega)\le x_1,\xi_1(\omega)\le\sqrt{x_2}\}=P\{\omega:\xi_1(\omega)\le\min(x_1,\sqrt{x_2})\}=$$

$$=\begin{cases}\min(x_1,\sqrt{x_2}), & \textit{если}\ \min(x_1,\sqrt{x_2})\le 1,\\ 1, & \textit{если}\ \min(x_1,\sqrt{x_2})>1,\end{cases}$$

Ф.р. многомерной СВ имеет следующие свойства:

1) $F_\xi(x)$ является неубывающей по всем аргументам;

2) $\lim\limits_{x_k\to-\infty}F_{\xi_1\dots\xi_n}(x_1,\dots,x_n)=0$;

3) $\lim\limits_{x_k\to+\infty,\,k=\overline{1,n}}F_{\xi_1\dots\xi_n}(x_1,\dots,x_n)=1$;

4) $F_{\xi_1\dots\xi_n}(x_1,\dots,x_n)$ непрерывна справа по всем аргументам;

5) условие согласованности:

$$\lim_{x_k\to+\infty}F_{\xi_1\,\xi_{k-1}\,\xi_k\,\xi_{k+1}\dots\xi_n}(x_1,\dots,x_{k-1},x_k,x_{k+1},\dots,x_n)=$$

$$=F_{\xi_1\,\xi_{k-1}\,\xi_{k+1}\dots\xi_n}(x_1,\dots,x_{k-1},x_{k+1},\dots,x_n)\cdot$$

Ф.р. меньшей размерности, которая получается из ф.р. большей размерности, если применить для нее условие согласованности, называется маргинальной;

6) отметим, что, для того чтобы некоторая функция $F(x_1, ..., x_n)$ была ф.р. n-мерной CB, недостаточно, чтобы для нее были выполнены условия 1) – 5). Необходимо также выполнение еще одного условия. Пусть $a_k < b_k$, $A_k = \{x : a_k < x \le b_k\}$, $k = \overline{1, n}$. Нетрудно видеть, что

$$P\{\omega : \xi_1(\omega) \le x_1, ..., \xi_{k-1}(\omega) \le x_{k-1}, a_k < \xi_k(\omega) \le b_k, \xi_{k+1}(\omega) \le x_{k+1}, ..., \xi_n(\omega) \le x_n\} =$$
$$= F_{\xi_1 ... \xi_n}(x_1, ..., x_{k-1}, b_k, x_{k+1}, ..., x_n) - F_{\xi_1 ... \xi_n}(x_1, ..., x_{k-1}, a_k, x_{k+1}, ..., x_n) =$$
$$= \Delta_k F_{\xi_1 ... \xi_n}(x_1, ..., x_n),$$

а также $P\{\omega : \xi_k(\omega) \in \acute{A}_k, k = \overline{1, n}\} = \Delta_1 \Delta_2 ... \Delta_n F_{\xi_1 ... \xi_n}(x_1, ..., x_n)$.

Отсюда ясно, что для многомерной ф.р. должно выполняться условие

$$\Delta_1 \Delta_2 ... \Delta_n F_{\xi_1 ... \xi_n}(x_1, ..., x_n) \ge 0 \cdot$$

Это условие не следует из свойств 1) – 5). Покажем это на примере.

Пример 1.23. Пусть $n = 2$,

$$F_{\xi_1\xi_2}(x_1, x_2) = \begin{cases} 0, x_1 < 0, x_2 < 0, x_1 + x_2 < 1, \\ 1, \text{в остальных случаях.} \end{cases}$$

Для такой функции, что легко проверить, выполняются свойства 1) – 5). Пусть $A_k = \left\{\omega : \frac{1}{3} < \xi_k(\omega) \le 1\right\}$, тогда

$$P\{\omega : \xi_k(\omega) \in \acute{A}_k, k = 1, 2\} = \Delta_1 \Delta_2 F_{\xi_1 \xi_2}(x_1, x_2) = \Delta_1 \left[F_{\xi_1 \xi_2}(x_1, 1) - F_{\xi_1 \xi_2}\left(x_1, \frac{1}{3}\right)\right] =$$
$$= F_{\xi_1 \xi_2}(1, 1) - F_{\xi_1 \xi_2}\left(\frac{1}{3}, 1\right) - F_{\xi_1 \xi_2}\left(1, \frac{1}{3}\right) + F_{\xi_1\xi_2}\left(\frac{1}{3}, \frac{1}{3}\right) = 1 - 1 - 1 + 0 = -1$$

Таким образом, получили, что если $F_{\xi_1 \xi_2}(x_1, x_2)$ – ф.р., то найденная вероятность –отрицательная. Это невозможно, значит, выполнение условий 1) – 5) является недостаточным для того, чтобы $F_{\xi_1\xi_2}(x_1, x_2)$ была ф.р.

Так же, как и в одномерном случае, $F_\xi(x)$ **относится к дискретному типу**, если каждая из СВ $\xi_k(\omega), k=\overline{1,n}$, принимает значения из счетного или конечного множества. Дискретную СВ $\xi(\omega)=(\xi_1(\omega),\xi_2(\omega),...,\xi_n(\omega))$ удобно описывать распределением вероятностей

$$P_\xi(x)=P\{\omega:\xi_1(\omega)=x_1,\xi_2(\omega)=x_2,...,\xi_n(\omega)=x_n\},\quad \sum_x P_\xi(x)=1,$$

где $x=(x_1,x_2,...,x_n)$.

$F_\xi(x)$ относится к абсолютно непрерывному типу распределения, если ее можно представить в виде

$$F_\xi(x)=F_{\xi_1...\xi_n}(x_1,...,x_n)=\int_{-\infty}^{x_1}dt_1...\int_{-\infty}^{x_n}p_{\xi_1...\xi_n}(t_1,...,t_n)dt_n\,.$$

Здесь $\dfrac{\partial^n F_{\xi_1...\xi_n}(x_1,...,x_n)}{\partial x_1...\partial x_n}=p_{\xi_1...\xi_n}(x_1,...,x_n)$,

$$p_{\xi_1...\xi_n}(x_1,...,x_n)\geq 0,\ \int_{-\infty}^{+\infty}dx_1...\int_{-\infty}^{+\infty}p_{\xi_1...\xi_n}(x_1,...,x_n)dx_n=1\,.$$

Функция $p_\xi(x)=p_{\xi_1...\xi_n}(x_1,...,x_n)$, обладающая перечисленными свойствами, называется **плотностью распределения вероятностей многомерной СВ** $\xi(\omega)=(\xi_1(\omega),\xi_2(\omega),...,\xi_n(\omega))$ (совместной плотностью вероятностей величин $\xi_1(\omega),\xi_2(\omega),...,\xi_n(\omega)$. Из условия согласованности для $F_\xi(x)$ вытекает следующее свойство для совместной плотности вероятностей

$$\int_{-\infty}^{+\infty}p_{\xi_1...\xi_{k-1}\xi_k\xi_{k+1}...\xi_n}(x_1,...x_{k-1},x_k,x_{k+1},...,x_n)dx_k=$$

$$=p_{\xi_1...\xi_{k-1}\xi_{k+1}...\xi_n}(x_1,...,x_{k-1},x_{k+1},...,x_n).$$

Плотность распределения, стоящая справа, называется маргинальной по отношению к исходной. Справедлива также следующая важная формула:

$$P\{\omega:(\xi_1(\omega),\dots,\xi_n(\omega))\in G\}=\underset{G}{\int\dots\int} p_{\xi_1\dots\xi_n}(x_1,\dots x_n)dx_1\dots dx_n,$$

например,

$$P\{\omega:\xi_1^2(\omega)+\xi_2^2(\omega)\le Z\}=\iint\limits_{x_1^2+x_2^2\le Z} p_{\xi_1\xi_2}(x_1,x_2)dx_1dx_2.$$

Рассмотрим примеры многомерных распределений.

1. Полиномиальное распределение имеет дискретная СВ

$$\xi(\omega)=(\xi_1(\omega),\xi_2(\omega),\dots,\xi_n(\omega)),\ \text{где}$$

$$\xi_k(\omega)\in\{0,1,2,\dots,N\},\ \sum_{k=1}^{n}\xi_k(\omega)=N,$$

$$P_\xi(x)=P_{\xi_1\dots\xi_n}(x_1,\dots,x_n)=N!\prod_{k=1}^{n}\frac{p_k^{x_k}}{x_k!},\ 0\le x_k\le N,\ k=\overline{1,n},\ \sum_{k=1}^{n}x_k=N,$$

$$0<p_k<1,\ k=\overline{1,n},\ \sum_{k=1}^{n}p_k=1.$$

2. Многомерное нормальное распределение имеет СВ $\xi(\omega)=(\xi_1(\omega),\xi_2(\omega),\dots,\xi_n(\omega))$, непрерывного типа, для которой

$$p_\xi(x)=p_{\xi_1\dots\xi_n}(x_1,\dots,x_n)=(2\pi)^{-\frac{n}{2}}\sqrt{|G|}\exp\left\{-\frac{1}{2}\sum_{i,j=1}^{n}(x_i-a_i)g_{ij}(x_j-a_j)\right\},$$

где $G=\|g_{ij}\|_{n\times n}$ – положительно определенная матрица, $|G|$ – ее определитель.

Пример 1.24. Пусть $(\xi(\omega),\eta(\omega))$ – двумерная СВ, имеющая нормальное распределение,

$$p_{\xi\eta}(x,y)=\frac{\sqrt{|G|}}{2\pi}exp\left\{-\frac{1}{2}\begin{pmatrix}x-a\\y-b\end{pmatrix}^T G\begin{pmatrix}x-a\\y-b\end{pmatrix}\right\},\ G=\|g_{ik}\|_{2\times2}.$$

Найдем плотности распределения СВ $\xi(\omega)$ и $\eta(\omega)$:

$$p_{\xi}(x)=\int_{-\infty}^{\infty} p_{\xi\eta}(x,y)dy=\sqrt{\frac{1}{2\pi}\left|g_{11}-\frac{g_{12}g_{21}}{g_{22}}\right|}exp\left\{-\frac{1}{2}\left(g_{11}-\frac{(g_{12}+g_{21})^2}{4g_{22}}\right)(x-a)^2\right\},$$

$$p_{\eta}(x)=\int_{-\infty}^{\infty} p_{\xi\eta}(x,y)dy=\sqrt{\frac{1}{\sqrt{2\pi}}\left|g_{22}-\frac{g_{12}g_{21}}{g_{11}}\right|}exp\left\{-\frac{1}{2}\left(g_{22}-\frac{(g_{12}+g_{21})^2}{4g_{11}}\right)(y-b)^2\right\}.$$

Из условий нормировки $\int_{-\infty}^{\infty} p_{\xi}(x)dx=\int_{-\infty}^{\infty} p_{\eta}(y)dy=1$ для элементов матрицы G вытекает следующее требование:

$$g_{12}=g_{21},\ g_{11}g_{22}-g_{12}g_{21}=|G|>0,$$

т.е. матрица G должна быть симметричной и положительно определенной.

Пусть (Ω, F, P) – вероятностное пространство, $A, B\in \mathrm{F}$. Если $P(B)>0$, то условная вероятность события A

$$P(A/B)=\frac{P(A\cap B)}{P(B)}$$

существует и, как мы видели раньше, удовлетворяет свойствам вероятностной функции P. Если зафиксировать событие B, то эту условную вероятностную функцию событий $A\in \mathrm{F}$ можно рассматривать как элемент для построения нового вероятностного пространства (Ω, F, P). Ранее была определена ф.р. $F_{\xi}(x)$ CB $\xi(\omega)=(\xi_1(\omega),\xi_2(\omega),...,\xi_n(\omega))$, рассматриваемой на вероятностном пространстве (Ω, F, P). Аналогично можно определить ф.р. СВ $\xi(\omega)$ на (Ω, F, P). Эта функция называется **условной ф.р. СВ** $\xi(\omega)$ при условии B и обозначается $F_{\xi}(x/B)$. Ее определяют следующим образом: пусть

$A = \{\omega : \xi_1(\omega) \le x_1, \xi_2(\omega) \le x_2, ..., \xi_n(\omega) \le x_n\}$, тогда $F_\xi(x/B) = F_{\xi_1 ... \xi_n}(x_1, ..., x_n / B) = P(A/B)$. Когда $F_\xi(x/B)$ – абсолютно непрерывная функция, то она имеет плотность распределения вероятностей

$$p_\xi(x/B) = \frac{\partial^n F_\xi(x/B)}{\partial x_1 ... \partial x_n},$$

которая называется **условной плотностью вероятностей СВ** $\xi(\omega)$.

Часто в качестве условия B используется событие, связанное с тем, что некоторая СВ $\xi(\omega)$ приняла определенное значение. Допустим, что на $(\Omega, \mathcal{F}, P)$ определена многомерная СВ $\eta(\omega) = (\eta_1(\omega), \eta_2(\omega), ..., \eta_m(\omega))$.
Пусть

$$B_\varepsilon = \{\omega : y \le \eta(\omega) \le y + \varepsilon\} = \{\omega : y_k \le \eta_k(\omega) \le y_k + \varepsilon, k = \overline{1, m}\}.$$

Если существует $\lim\limits_{\varepsilon \to 0} F_\xi(x/B_\varepsilon) = F_\xi(x/y)$, то он называется условной ф.р. СВ $\xi(\omega)$ при условии $\eta(\omega) = y$.

$$F_\xi(x/y) = \lim_{\varepsilon \to 0} F_\xi(x/B_\varepsilon) = \lim_{\varepsilon \to 0} F_\xi(x / y \le \eta(\omega) \le y + \varepsilon) =$$

$$= \lim_{\varepsilon \to 0} P(\xi(\omega) \le x / y \le \eta(\omega) \le y + \varepsilon) =$$

$$= \lim_{\varepsilon \to 0} \frac{P(\xi(\omega) \le x, y \le \eta(\omega) \le y + \varepsilon)}{P(y \le \eta(\omega) \le y + \varepsilon)} = \lim_{\varepsilon \to 0} \frac{F_{\xi\eta}(x, y + \varepsilon) - F_{\xi\eta}(x, y)}{F_{\xi\eta}(\infty, y + \varepsilon) - F_{\xi\eta}(\infty, y)},$$

где $F_{\xi\eta}(x, y)$ – известная $(n + m)$-мерная ф.р. СВ $\xi(\omega)$ и $\eta(\omega)$. Предположим, что $F_{\xi\eta}(x, y)$ – абсолютно непрерывная ф.р., тогда, пользуясь ее определением и правилом Лопиталя, будем иметь:

$$\lim_{\varepsilon\to 0}\frac{F_{\xi\eta}(x, y+\varepsilon)-F_{\xi\eta}(x, y)}{F_{\xi\eta}(\infty, y+\varepsilon)-F_{\xi\eta}(\infty, y)}=\lim_{\varepsilon\to 0}\frac{\int\limits_{-\infty}^{x}\int\limits_{y}^{y+\varepsilon}p_{\xi\eta}(t,\tau)d\tau dt}{\int\limits_{y}^{y+\varepsilon}p_{\eta}(\tau)d\tau}=$$

$$=\int\limits_{-\infty}^{x}\lim_{\varepsilon\to 0}\frac{\int\limits_{y}^{y+\varepsilon}p_{\xi\eta}(t,\tau)d\tau}{\int\limits_{y}^{y+\varepsilon}p_{\eta}(\tau)d\tau}dt=\int\limits_{-\infty}^{x}\frac{p_{\xi\eta}(t, y)}{p_{\eta}(y)}dt=\int\limits_{-\infty}^{x}p_{\xi}(t/y)dt,$$

если предел под интегралом существует, т.к.

$$\left(\int\limits_{y}^{y+\varepsilon}p_{\xi\eta}(t,\tau)d\tau\right)_{\varepsilon}^{'}=\int\limits_{y}^{y+\varepsilon}p_{\xi\eta}^{'}(t,\tau)d\tau+(y+\varepsilon)_{\varepsilon}^{'}\,p_{\xi\eta}(t, y+\varepsilon)-$$

$$-0\cdot p_{\xi\eta}(t, y)\xrightarrow[\varepsilon\to 0]{}p_{\xi\eta}(t, y).$$

В этих формулах $\int\limits_{-\infty}^{x}dt$ означает n-мерный интеграл $\int\limits_{-\infty}^{x_1}dt_1\ldots\int\limits_{-\infty}^{x_n}dt_n$. Таким образом,

$$F_{\xi}(x/y)=\int\limits_{-\infty}^{x}p_{\xi}(t/y)dt,$$

где $p_{\xi}(x/y)$ – условная плотность вероятностей, которая определяется равенством

$$p_{\xi}(x/y)=\frac{p_{\xi\eta}(x/y)}{p_{\eta}(y)}.$$

Из условия согласованности следует:

$$p_{\xi}(x)=\int_{-\infty}^{\infty}p_{\xi\eta}(x;y)dy=\int_{-\infty}^{\infty}p_{\xi}(x/y)p_{\eta}(y)dy,$$

$$p_{\xi}(x/y)=\frac{p_{\xi\eta}(x,y)}{p_{\eta}(y)}=\frac{p_{\xi}(x)p_{\eta}(y/x)}{\int_{-\infty}^{\infty}p_{\xi}(x)p_{\eta}(y/x)dx}.$$

Формула $p_{\xi}(x)=\int_{-\infty}^{\infty}p_{\xi}(x/y)p_{\eta}(y)dy$ называется **формулой полной вероятности для плотностей распределения**, а формула

$$p_{\xi}(x/y)=\frac{p_{\xi}(x)p_{\eta}(y/x)}{\int_{-\infty}^{\infty}p_{\xi}(x)p_{\eta}(y/x)dx}\ -$$

формулой Байеса для плотностей распределения.

Пример 1.25. Найдем условие плотности распределения вероятностей для СВ, рассматриваемых в предыдущем примере.

$$p_{\xi}(x/y)=\frac{p_{\xi\eta}(x,y)}{p_{\eta}(y)}=\sqrt{\frac{g_{11}}{2\pi}}exp\left\{-\frac{1}{2}\left(\sqrt{g_{11}}(x-a)+\frac{g_{12}}{\sqrt{g_{11}}}(y-b)\right)^2\right\},$$

$$p_{\eta}(y/x)=\frac{p_{\xi\eta}(x,y)}{p_{\eta}(x)}=\sqrt{\frac{q_{22}}{2\pi}}exp\left\{-\frac{1}{2}\left(\sqrt{g_{22}}\,y-b+\frac{g_{21}}{\sqrt{g_{22}}}(x-a)\right)^2\right\}.$$

Рассмотрим вопросы, связанные с **независимостью и функциональными преобразованиями СВ**. Пусть $\xi_k(\omega), k=\overline{1,n},$ – СВ, определенные на вероятностном пространстве $(\Omega, \mathrm{F}\ , P)$ $A_k=\{\omega:\xi_k(\omega)\in B_k\}$, где B_k – борелевское множество.

Определение. СВ $\xi_k(\omega)$, $k=\overline{1,n}$, называются независимыми в совокупности, если, как бы ни выбирались борелевские множества B_k, $k=\overline{1,n}$, случайные события A_k, $k=\overline{1,n}$, являются независимыми в совокупности, т.е.

$$P\left(\bigcap_{i=1}^{m} \mathrm{A}_{j_i}\right)=\prod_{i=1}^{m} P(\mathrm{A}_{j_i}) \ \forall \ m\le n,\ 1\le j_1<j_2<\ldots<j_m\le n.$$

Определение. СВ $\xi_k(\omega)$, $k=\overline{1,\, n}$, называется независимыми парами, если $\forall\ 1\le i<k\le n$ и для любых борелевских множеств B_i и B_k события A_i и A_k являются независимыми, т.е. $P(A_i\cap A_k)=P(A_i)P(A_k)$.

Ясно, что СВ, независимые в совокупности, являются и независимыми парами.

Теорема 1.3. *Пусть* $\xi_k(\omega)$, $k=\overline{1,\, n}$, *– независимые в совокупности СВ, а* $f_k(x)$, $k=\overline{1,\, n}$, *– борелевские функции. Тогда СВ* $\eta_k(\omega)=f_k(\xi_k(\omega))$, $k=\overline{1,\, n}$, *также независимы в совокупности.*

Доказательство. Пусть $B_{j_1}, B_{j_2},\ldots, B_{j_m}$ – борелевские множества $\forall\ m\le n$ и $1\le j_1<j_2<\ldots<j_m\le n$. Тогда

$$P\{\omega{:}f_{j_k}(\xi_{jk}(\omega))\in B_{j_k},\ k=\overline{1,m}\}=P\{\omega{:}\xi_{j_k}(\omega)\in f_{j_k}^{-1}(B_{j_k}),\ k=\overline{1,\, m}\}=$$
$$=\prod_{k=1}^{m} P\{\omega{:}\xi_{j_k}(\omega)\in f_{j_k}^{-1}(B_{j_k})\}=\prod_{k=1}^{m} P\{\omega : f_{j_k}(\xi_{j_k}(\omega))\in B_{j_k}\}.$$

Приведем два критерия независимости СВ в совокупности.

Теорема 1.4. *Для того чтобы СВ* $\xi_k(\omega)$, $k=\overline{1,\, n}$, *были независимыми в совокупности, необходимо и достаточно, чтобы* $\forall\ x_k, k=\overline{1,n}$,

$$F_{\xi_1\ldots\xi_n}(x_1,\ldots, x_n)=\prod_{k=1}^{n} F_{\xi_k}(x_k)$$

Если $F_{\xi_1\ldots\xi_n}(x_1, \ldots, x_n)$ – абсолютно непрерывная функция, то имеет место следующее утверждение.

Теорема 1.5. *Для того чтобы СВ* $\xi_k(\omega)$, $k=\overline{1,\,n}$, *были независимыми в совокупности, необходимо и достаточно, чтобы*

$$p_{\xi_1\ldots\xi_n}(x_1, \ldots, x_n) = \prod_{k=1}^{n} p_{\xi_k}(x_k).$$

Доказательство следует из предыдущей теоремы и определения абсолютно непрерывной ф.р. Докажем, например, необходимость

$$F_{\xi_1\ldots\xi_n}(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} dt_1 \ldots \int_{-\infty}^{x_n} p_{\xi_1\ldots\xi_n}(t_1, \ldots, t_n)dt_n = \prod_{k=1}^{n} F_{\xi_k}(x_k) =$$

$$= \prod_{k=1}^{n} \int_{-\infty}^{x_k} p_{\xi_k}(t_k)dt_k = \int_{-\infty}^{x_1} dt_1 \ldots \int_{-\infty}^{x_n} \prod_{k=1}^{n} p_{\xi_k}(t_k)dt_k.$$

Поскольку это выполняется $\forall x_1, x_2, \ldots, x_n$, то отсюда следует, что $p_{\xi_1\ldots\xi_n}(t_1, \ldots, t_n) = \prod_{k=1}^{n} p_{\xi_k}(t_k)$

Отметим, что из приведенных выше теорем следует, что если $\xi(\omega)$ и $\eta(\omega)$ – независимых СВ, то $F_\xi(x/y) = F_\xi(x)$, $p_\xi(x/y) = p_\xi(x)$, т.е. условная ф.р. и плотность распределения совпадают с безусловными.

Пример 1.26. Из примеров 1.24 и 1.25 следует, что, для того, чтобы нормально распределенные СВ $\xi(\omega)$ и $\eta(\omega)$ были независимыми, необходимо и достаточно, чтобы диагональные элементы матрицы G были равны нулю, т.е. $q_{12} = q_{21} = 0$. Приведем аналогичное утверждение для дискретных СВ.

Теорема 1.6. *Пусть* $\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)$ *– СВ, каждая из которых может принимать не более, чем счетное число значений. Они являются независимыми тогда и только тогда, когда* $\forall\ x_1, x_2, \ldots, x_n$

$$P\{\omega{:}\,\xi_1(\omega)=x_1,\ \xi_2(\omega)=x_2,\ \ldots,\ \xi_n(\omega)=x_n\}=\prod_{k=1}^{n}P\{\omega{:}\,\xi_k(\omega)=x_k\}$$

Доказательство основано на следующих соотношениях. Пусть $B_1, \ldots, B_n$ – произвольные борелевские множества на прямой, тогда

$$P\{\omega{:}\,\xi_1(\omega)\in B_1,\ldots,\xi_n(\omega)\in B_n\}=P\left\{\omega{:}\ \bigcup_{j_1}\{\xi_1(\omega)=x_1\},\ldots,\bigcup_{j_n}\{\xi_n(\omega)=x_n\}\right\},$$

объединение здесь берется по всем элементам множеств $B_1, \ldots, B_n$ положительной вероятности. Значит,

$$P\{\omega{:}\,\xi_1(\omega)\in B_1,\ldots,\ \xi_n(\omega)\in B_n\}=\sum_{j_1}\sum_{j_2}\ldots\sum_{j_n}P\{\omega{:}\,\xi_1(\omega)=x_{j_1},\ldots,\xi_n(\omega)=x_{j_n}\}=$$

$$=\sum_{j_1}\sum_{j_2}\ldots\sum_{j_n}P\{\omega{:}\,\xi_1(\omega)=x_{j_1}\}\cdot\ldots\cdot P\{\omega{:}\,\xi_n(\omega)=x_{j_n}\}=$$

$$=\sum_{j_1}P\{\omega{:}\,\xi_1(\omega)=x_{j_1}\}\cdot\ldots\cdot\sum_{j_n}P\{\omega{:}\,\xi_n(\omega)=x_{j_n}\}=$$

$$=P\{\omega{:}\,\xi_1(\omega)\in B_1,\ldots,\ \xi_n(\omega)\in B_n\}.$$

Пример 1.27. Пусть $(\Omega, \mathcal{F}, P)$ – вероятностное пространство, причем Ω состоит ровно из n точек, каждая из которых имеет положительную вероятность. Покажем, что на этом вероятностном пространстве не существует двух независимых СВ, каждая из которых принимает n различных значений.

Возьмем произвольное из n элементарных событий ω_1. Пусть ξ и η – произвольные СВ, определенные на этом вероятностном пространстве и принимающие n различных значений каждая. Положим $\xi(\omega_1)=a, \eta(\omega_1)=b.$ Тогда $P(\xi=a)=P(\{\omega_1\})=P(\eta=b),$ поскольку не существует других элементарных событий, на которых СВ ξ принимала бы значение a или СВ η принимала бы значение b. Следовательно,

$$P(\xi = a, \eta = b) = P(\{\omega_1\}), P(\xi = a)P(\eta = b) = P^2(\{\omega_1\}),$$ и т.к.

$P(\{\omega\}) \neq P^2(\{\omega\})$, то СВ ξ и η – зависимы.

Определение. Пусть $\hat{a}^{\,n}$ – система борелевских множеств в R^n, $\hat{a}$ – система борелевских множеств на прямой R. Функция n аргументов $f(x) \in R,\ x = (x_1, ..., x_n) \in R^n,$ называется **борелевской (измеримой по Борелю)**, если $f^{-1}(B) = \{x : f(x) \in B\} \in \hat{a}^{\,n}\ \forall\ B \in \hat{a}$, т.е. прообраз борелевского множества из R является борелевским множеством в R^n.

Теорема 1.7 ***(о суперпозиции измеримых функций).*** *Пусть* $\xi(\omega)$ – n*-мерная СВ, определенная на вероятностном пространстве* (Ω, F, P), $f(x)$ – *борелевская функция на* R^n. *Тогда* $\eta(\omega) = f(\xi(\omega)) = f(\xi_1(\omega), ..., \xi_n(\omega))$ *также является СВ на* (Ω, F, P).

Доказательство. Пусть $B \in \hat{a}$. Т.к. $\{\omega : f(\xi(\omega)) \in B\} = \{\omega : \xi(\omega) \in f^{-1}(B)\}$ и $f^{-1}(B)$ – борелевское множество из R^n, то

$$\{\omega : \xi(\omega) \in f^{-1}(B)\} = \xi^{-1}(f^{-1}(B)) \in \mathrm{F},$$

т.е. $\{\omega : \eta(\omega) \in B\} \in \mathrm{F}$, что и нужно было доказать.

Рассмотрим следующую задачу. Пусть $f_k(x_1, x_2, ..., x_n),\ k = \overline{1, m}$ – борелевские функции на R^n. Определим СВ

$$\eta_k(\omega) = f_k(\xi_1(\omega), \xi_2(\omega), ..., \xi_n(\omega)),\ k = \overline{1, m}.$$

Если $\xi(\omega)=(\xi_1(\omega),\xi_2(\omega),...,\xi_n(\omega))-n$-мерная СВ, то $\eta(\omega)=(\eta_1(\omega),\eta_2(\omega),...,\eta_m(\omega))$ будет m-мерной СВ. Пусть $F_\xi(x)$ и $F_\eta(y)-$ ф.р. соответственно СВ $\xi(\omega)$ и $\eta(\omega)$. Нужно выразить $F_\eta(y)$ через $F_\xi(x)$ и систему функций $f_k(x)$, $k=\overline{1,m}$.

Допустим, что ф.р. $F_\xi(x)$ – абсолютно непрерывная. Рассмотрим следующие случаи:

1) пусть $m=n$ и все функции $f_k(x)$, $k=\overline{1,n}$, являются дифференцируемыми и функционально независимыми, для последнего достаточно, чтобы

$$\det A=\det||a_{ik}||=\det\left\|\frac{\partial f_i(x)}{\partial x_k}\right\|\neq 0\ \ \forall\ x.$$

В данном случае будем иметь:

$$F_\eta(y)=P\{\eta_1\le y_1,...,\ \eta_n\le y_n\}=P\left\{f_k(\xi_1,...,\ \xi_n)\le y_k,\ k=\overline{1,\ n}\right\}=$$

$$=\int dx_1\ .\ \ .\ \ .\int\limits_{\left\{f_k(x_1,\ ...,\ x_n)\le y_k,k=\overline{1,\ n}\right\}} p_\xi(x_1,...,\ x_n)dx_n\ .$$

Сделаем замену переменных $f_k(x_1,...,x_n)=z_k$, $k=\overline{1,n}$, тогда

$$x_k=q_k(z_1,\ ...,\ z_n),\ f_k[q_1(z_1,\ ...,\ z_n),...,q_n(z_1,\ ...,\ z_n)]=$$
$$=z_k,\quad q(f_1(x_1,\ ...,\ x_n),...,f_k(x_1,\ ...,\ x_n))=x_k,\ k=\overline{1,\ n},$$

и поэтому

$$F_\eta(y)=\int\limits_{\left\{z_k\le y_k,k=\overline{1,n}\right\}} dq_1(z_1,\ ...,\ z_n)..\int p_\xi(q_1(z_1,\ ...,\ z_n),...,q_n(z_1,\ ...,\ z_n))dq_n(z_1,\ ...,\ z_n)=$$
$$=\int\limits_{-\infty}^{y_1}dz_1...\int\limits_{-\infty}^{y_n}p_\xi(q_1(z_1,\ ...,z_n),...,q_n(z_1,\ ...,z_n))\left|\frac{\partial(q_1,\ ...,\ q_n)}{\partial(z_1,\ ...,\ z_n)}\right|dz_n,$$

где $\left|\dfrac{\partial(q_1,...,q_n)}{\partial(z_1,...,z_n)}\right|$ – якобиан невырожденного преобразования

$x_k=q_k(z_1,...,z_n)$, $k=\overline{1,n}$, т.е. определителем ($n\times n$) – матрицы G

с элементами $q_{ik} = \dfrac{\partial q_i(z_1, ..., z_n)}{\partial z_k}$. Таким образом, $F_\eta(y)$ также абсолютно непрерывная ф.р., и плотность распределения СВ $\eta(\omega)$ равна

$$p_\eta(y_1, ..., y_n) = p_\xi(q_1(y_1, ..., y_n), ..., q_n(y_1, ..., y_n))\left|\frac{\partial(q_1, ..., q_n)}{\partial(y_1, ..., y_n)}\right|.$$

Данное выражение можно записать в более краткой форме:

$$p_\eta(y) = p_\xi(f^{-1}(y))\left|\frac{\partial f^{-1}(y)}{\partial(y)}\right|.$$

где $f(y) = (f_1(y), ..., f_n(y)) = (f_1(y_1, ..., y_n), ..., f_n(y_1, ..., y_n))$;

2) пусть теперь $m < n$ и $rang\left\|\dfrac{\partial f_i(x)}{\partial x_k}\right\| = m \ \ \forall x$. В этом случае систему функций $f_k(x), k = \overline{1, m},$ можно дополнить $(n - m)$ функциями $f_{m+j}(x), 1 \le j \le n - m,$ так, чтобы они были непрерывно дифференцируемы и все вместе функционально независимыми. Новая система функций определит n-мерную СВ $\overline{\eta}(\omega)$. Тогда из условия согласованности и предыдущего случая будем иметь:

$$p_{\varsigma}(y_1, ..., y_m) = \int_{-\infty}^{\infty} dy_{m+1} \cdots \int_{-\infty}^{\infty} p_{\overline{\varsigma}}(y_1, ..., y_n) dy_n =$$

$$= \int_{-\infty}^{\infty} dy_{m+1} \cdots \int_{-\infty}^{\infty} p_{\hat{\imath}}(q_1(y_1, ..., y_n), ..., q_n(y_1, ..., y_n))\left\|\frac{\partial(q_1, ..., q_n)}{\partial(y_1, ..., y_n)}\right\| dy_n .$$

Пример 1.28. Пусть у нас есть двумерная СВ $\xi(\omega) = (\xi_1(\omega), \xi_2(\omega))$. Образуем СВ $\eta(\omega) = \xi_1(\omega) + \xi_2(\omega)$ и найдем ее плотность распределения. В данном случае $n = 2$, $m = 1$, $f_1(x_1, x_2) = x_1 + x_2 = y_1$. Т.к. $m < n$, $n - m = 1$, то нужно ввести еще одну функцию, выберем $f_2(x_1, x_2) = x_2 = y_2$. Тогда обратное преобразование определяется функциями

$$x_1 = q_1(y_1, y_2) = y_1 - y_2, \ \ x_2 = q_2(y_1, y_2) = y_2.$$

Поэтому

$$\left\| \frac{\partial(q_1, q_2)}{\partial(y_1, y_2)} \right\| = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1 \text{ и } p_\eta(y) = \int_{-\infty}^{\infty} p_\xi(y - y_2, y_2) dy_2.$$

Если $\xi_1(\omega)$ и $\xi_2(\omega)$ – независимые СВ, то

$$p_\eta(y) = \int_{-\infty}^{\infty} p_{\xi_1}(y - y_2) p_{\xi_2}(y_2) dy_2.$$

Эта формула известна в литературе как свертка для плотностей распределения вероятностей и обозначается $p_\eta = p_{\xi_1} * p_{\xi_2}$.

Пример 1.29. Пусть СВ $\xi_1(\omega)$, $\xi_2(\omega)$ независимы и имеют стандартные гамма–распределения с параметрами p_1 и p_2 соответственно:

$$p_{\xi_i}(x) = \Gamma^{-1}(p_i) x^{p_i - 1} e^{-x}, \ \ x \ge 0, i = 1, 2,$$

где $\Gamma(p) = \int_0^{+\infty} x^{p-1} e^{-x} dx$ – гамма-функция. Покажем, что СВ $\eta(\omega) = \xi_1(\omega) + \xi_2(\omega)$ имеет гамма–распределение с параметром $p_1 + p_2$.

Согласно формуле для плотности суммы независимых СВ имеем

$$p_\eta(y) = \Gamma^{-1}(p_1)\Gamma^{-1}(p_2) \int_0^y (y - y_2)^{p_1 - 1} e^{-(y - y_2)} y_2^{p_2 - 1} e^{-y_2} dy_2 =$$

$$= \Gamma^{-1}(p_1)\Gamma^{-1}(p_2) e^{-y} \int_0^y (y - y_2)^{p_1 - 1} y_2^{p_2 - 1} dy_2.$$

Верхний предел в интеграле равен y, поскольку в данном соотношении должно быть $y - y_2 > 0$. Сделав замену переменной $y = y_2 u$, получим

$$p_\eta(y) = \Gamma^{-1}(p_1)\Gamma^{-1}(p_2) y^{p_1+p_2-1} e^{-y} \int_0^1 u^{p_2-1}(1-u)^{p_1-1} du.$$

Т.к. должно выполняться условие нормировки $\int_0^\infty p_\xi(y)dy = 1$, то величина $c = \Gamma^{-1}(p_1)\Gamma^{-1}(p_2)\int_0^1 u^{p_2-1}(1-u)^{p_1-1} du$ должна быть равна $\Gamma^{-1}(p_1 + p_2)$. Отметим также попутно доказанное равенство

$$\int_0^1 u^{p_2-1}(1-u)^{p_1-1} du = \frac{\Gamma(p_1)\Gamma(p_2)}{\Gamma(p_1+p_2)}.$$

Пример 1.30. Пусть $\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)$ – независимые СВ, имеющие экспоненциальное распределение с параметром λ, т.е.

$$p_{\xi_i}(x) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0, \\ 0, & x < 0. \end{cases}$$

Найдем распределение СВ $\eta_n(\omega) = \xi_1(\omega) + \xi_2(\omega) + \ldots + \xi_n(\omega)$.

Пусть $n = 2$. Имеем

$$p_{\eta_2}(y) = \int_0^y \lambda e^{-\lambda(y-y_2)} \lambda e^{-\lambda y_2} dy_2 = \lambda^2 e^{-\lambda y} \int_0^y dy_2 = \lambda^2 y e^{-\lambda y},\ y \ge 0.$$

Предположим, что для некоторого $n \ge 1$ справедлива формула

$$p_{\eta_n}(y) = \lambda \frac{(\lambda y)^{n-1}}{(n-1)!} e^{-\lambda y},\ y \ge 0.$$

Докажем, что она справедлива также и для $n+1$. По той же формуле для плотности суммы двух независимых СВ $\eta_{n+1}(\omega)=\eta_n(\omega)+\xi_{n+1}(\omega)$ получаем

$$p_{\eta_{n+1}}(y)=\int_0^y \lambda\frac{[\lambda(y-y_2)]^{n-1}}{(n-1)!}e^{-\lambda(y-y_2)}\lambda e^{-\lambda y_2}dy_2=\frac{\lambda^{n+1}}{(n-1)!}e^{-\lambda y}\int_0^y (y-y_2)^{n-1}dy_2=\lambda\frac{(\lambda y)^n}{n!}e^{-\lambda y},\ y\geq 0.$$

Распределение с плотностью $p_{\eta_n}(y)$ называется **распределением Эрланга n-го порядка**.

Если СВ $\xi_1(\omega),\ \xi_2(\omega)$ независимы и дискретны, то формула для распределения вероятностей их суммы $\eta(\omega)=\xi_1(\omega)+\xi_2(\omega)$ имеет вид

$$P_\eta(y)=P\{\omega{:}\xi_1(\omega)+\xi_2(\omega)=y\}=\sum_{\{y_1,\,y_2\,:\,y_1+y_2=y\}}P\{\omega{:}\xi_1(\omega)=y_1,\ \xi_2(\omega)=y_2\}=$$
$$=\sum_{\{y_1,y_2:y_1+y_2=y\}}P\{\omega{:}\xi_1(\omega)=y_1\}P\{\omega{:}\xi_2(\omega)=y_2\}=\sum_{\{y_1,y_2:y_1+y_2=y\}}P_{\xi_1}(y_1)P_{\xi_2}(y_2)=$$
$$=\sum_{\{y_2\}}P_{\xi_1}(y-y_2)P_{\xi_2}(y_2).$$

Пример 1.31. Пусть СВ $\xi_1(\omega),\ \xi_2(\omega)$ независимы и имеют биномиальные распределения:

$$P_{\xi_1}(\ell)=C_m^\ell p^\ell(1-p)^{n-\ell},\ \ P_{\xi_2}(k)=C_n^k p^k(1-p)^{n-k},\ \ \ell,k=\overline{0,n}.$$

Найдем распределение вероятностей СВ $\eta(\omega)=\xi_1(\omega)+\xi_2(\omega)$:

$$P_\eta(k)=\sum_{i=0}^n C_m^{k-i}p^{k-i}(1-p)^{m-k+i}C_n^i p^i(1-p)^{n-i}=p^k(1-p)^{n+m-k}\sum_{i=0}^n C_n^i C_m^{k-i}=$$
$$=C_{n+m}^k p^k(1-p)^{n+m-k},k=\overline{0,n}.$$

Здесь на последнем шаге использовано тождество

$$\sum_{i=0}^n C_n^i C_m^{k-i}=C_{n+m}^k\cdot$$

Пример 1.32. Пусть СВ $\xi_1(\omega)$, $\xi_2(\omega)$ независимы и имеют пуассоновские распределения с параметрами $\lambda,\ \mu$:

$$P_{\xi_1}(k)=\frac{\lambda^k}{k!}e^{-\lambda},\ \ P_{\xi_2}(\ell)=\frac{\mu^\ell}{\ell!}e^{-\mu},\ \ k,\ell=0,1,2,\ldots$$

Покажем, что их сумма $\xi_1(\omega)+\xi_2(\omega)=\eta(\omega)$ имеет пуассоновское распределение с параметром $\lambda+\mu$:

$$P_\eta(k)=\sum_{i=0}^{k}\frac{\lambda^{k-i}}{(k-i)!}e^{-\lambda}\,\frac{\mu^i}{i!}e^{-\mu}=\frac{e^{-\lambda-\mu}}{k!}\sum_{i=0}^{k}\frac{k!}{i!(k-i)!}\mu^i\lambda^{k-i}=\frac{(\lambda+\mu)^k}{k!}e^{-(\lambda+\mu)},$$

поскольку справедлива формула бинома Ньютона

$$(\mu+\lambda)=\sum_{i=0}^{k}C_k^i\mu^i\lambda^{k-i}.$$

Пример 1.33. Пусть имеем двумерную СВ $\xi(\omega)=(\xi_1(\omega),\ \xi_2(\omega))$ непрерывного типа, для которой известна плотность распределения $p_\xi(x_1,x_2)$. Необходимо найти $p_\eta(y)$, где

$\eta(\omega)=\dfrac{\xi_1(\omega)}{\xi_2(\omega)}$. В данном случае опять $n=2,\ m=1$. При этом

$f_1(x_1,x_2)=\dfrac{x_1}{x_2}=y_1$ и введем функцию $f_2(x_1,x_2)=x_2=y_2$. Обратное преобразование имеет вид:

$$x_1=q_1(y_1,y_2)=y_1y_2,\ \ x_2=q_2(y_1,y_2)=y_2\cdot$$

Якобиан этого преобразования равен

$$\left|\frac{\partial(q_1,q_2)}{\partial(y_1,y_2)}\right|=\begin{vmatrix}y_2 & y_1\\ 0 & 1\end{vmatrix}=y_2\cdot$$

Поэтому $p_{\bar{\eta}}(y_1,y_2)=p_\xi(y_1y_2,y_2)|y_2|$, и, таким образом

$$p_\eta(y)=\int_{-\infty}^{\infty}p_\xi(yy_2,y_2)|y_2|dy_2\cdot$$

Если СВ $\xi_1(\omega)$ и $\xi_2(\omega)$ независимы, то

$$p_{ç}(y)=\int\limits_{0}^{\infty}|y_2|\,p_{\hat{\imath}_1}(yy_2)\,p_{\hat{\imath}_2}(y_2)\,dy_2 =$$

$$=\int\limits_{0}^{\infty}y_2\,p_{\hat{\imath}_1}(yy_2)\,p_{\hat{\imath}_2}(y_2)\,dy_2-\int\limits_{-\infty}^{0}y_2\,p_{\hat{\imath}_1}(yy_2)\,p_{\hat{\imath}_2}(y_2)\,dy_2\,.$$

Используя этот результат, можно найти также плотность распределения произведения СВ $\eta(\omega)=\xi_1(\omega)\xi_2(\omega)$:

$$p_{\eta}(y)=\int\limits_{-\infty}^{\infty}p_{\xi}\left(y_2,\frac{y}{y_2}\right)\frac{1}{|y_2|}\,dy_2\,,$$

а если они независимы, то

$$p_{\eta}(y)=\int\limits_{-\infty}^{\infty}\frac{1}{|y_2|}\,p_{\xi_1}(y_2)\,p_{\xi_2}\left(\frac{y}{y_2}\right)dy_2\,.$$

§5• . ПРОСТРАНСТВА С МЕРОЙ, ИНТЕГРАЛ ЛЕБЕГА

Пусть × – произвольное множество элементов, назовем его пространством. Пространство × называется измеримым и обозначается $(X,\ \mathrm{F})$, если на нем выбрана некоторая σ-алгебра подмножеств F. Говорят, что в измеримом пространстве $(X,\ \mathsf{F})$ введена мера, если на элементах σ-алгебры F (т.е. на подмножествах ×, которые принадлежат F) задана неотрицательная функция $\mu=\mu(\mathrm{A})$, которая является счетно-аддитивной, т.е. для любого числа непересекающихся множеств $\mathrm{A}_1,\mathrm{A}_2,\ldots\in\mathsf{F}$

$$\mu\left(\bigcup_k \mathrm{A}_k\right)=\sum_k\mu(\mathrm{A}_k).$$

• Данный параграф является вспомогательным для изучения дальнейших параграфов

Значение $\mu(\mathrm{A})$, $\mathrm{A} \in \mathsf{F}$, называется мерой множества A, **совокупность** $(X,\ \mathsf{F},\ \mu)$ **называется пространством с мерой**. Ранее введенное вероятностное пространство $(X,\ \mathsf{F},\ P)$ – это пространство с нормированной мерой, $P(\Omega)=1$.

Рассмотрим, как строится мера Лебега в пространстве R^n. Для n-мерного параллелепипеда $\Delta = \{x = (x_1, x_2, \dots x_n): \alpha_i < x_i < b_i, i = \overline{1, n}\}$ мера Лебега равна $\mu(\Delta)=(b_1 - a_1)(b_2 - a_2)\dots(b_n - a_n)$. Пусть теперь $\mathrm{A} \in R^n$ – произвольное множество в R^n. Покроем это множество конечным либо счетным числом параллелепипедов $\Delta_1, \Delta_2, \dots, \Delta_k, \dots$. Нижняя грань чисел $\sum_k \mu(\Delta_k)$, которая берется по всевозможным покрытиям множества A, называется верхней мерой $m^*(\mathrm{A})$ множества A. Нижняя мера $m_*(\mathrm{A})$ определяется как $m(\Delta)-m^*(\overline{\mathrm{A}})$, где Δ – некоторый параллелепипед, содержащий множество A, $\overline{\mathrm{A}}$ – множество всех точек этого параллелепипеда, которые не принадлежат множеству A. Множества, для которых верхняя мера равна нижней, называются измеримыми по Лебегу, а **общее значение** $m(\mathrm{A})$ **верхней и нижней мер – мерой Лебега множества** A.

Допустим, что $\mu(X)<+\infty$. Напомним некоторые определения. Функция $f(x)$, заданная на $(X,\ \mathsf{F},\ \mu)$, называется измеримой, если $f^{-1}(B)=\{x : x \in X, f(x) \in B\} \subset \mathsf{F}$, где B – любое борелевское множество на числовой прямой. Измеримая действительная функция $f(x)$ называется простой (дискретной), если множество её значений не более чем счетно. Например, индикатор множества A

$$I_{\mathrm{A}}(x)=\begin{cases}1, x \in \mathrm{A},\\ 0, x \notin \mathrm{A}\end{cases}$$

является простой функцией.

Любую простую функцию $f(x)$ можно представить в виде

$$f(x)=\sum_i c_i I_{\mathrm{A}_i}(x),$$

где $c_1, c_2, \dots$ – значения, которые принимает $f(x)$,

$A_i = \{x : f(x) = c_i\}$, $i = 1, 2\ldots$. **Интегралом Лебега от простой функции по мере** μ называется число

$$\int\limits_X f(x)\mu(dx) = \sum_i c_i \mu(A_i),$$

если ряд в правой части сходится абсолютно.

Пример 1.34. Пусть $f(x)$ – конечная неотрицательная простая функция. Вывести неравенство

$$\int\limits_X f(x)\mu(dx) \le \max_x \{f(x)\}\mu\{x : f(x) > \varepsilon\} + \varepsilon\mu\{x : 0 < f(x) \le \varepsilon\}\cdot$$

Решение. Интеграл в левой части неравенства можно записать в виде

$$\int\limits_X f(x)\mu(dx) = \sum_{i=1}^{3} \int\limits_{X_i} f(x)\mu(dx),$$

где

$$X_1 = \{x : f(x) > \varepsilon\},\ X_2 = \{x : 0 < f(x) \le \varepsilon\},\ X_3 = \{x : f(x) = 0\}.$$

На множестве X_1 воспользуемся неравенством $f(x) \le \max\limits_x\{f(x)\}$, на множестве X_2 – неравенством $f(x) \le \varepsilon$. Тогда, применив свойство монотонности интеграла от простой функции ($\int\limits_X f_1(x)\mu(dx) \le \int\limits_X f_2(x)\mu(dx)$ при $f_1(x) \le f_2(x)$, что легко проверить, используя определения интеграла для простых функций), получаем требуемое неравенство.

Пусть $f(x)$ – измеримая функция, заданная на пространстве с мерой $(X,\ F,\ \mu)$.

Теорема 1.8. *Для любой измеримой функции $f(x)$ существует последовательность $\{f_n(x)\}$ простых функций, сходящаяся при $n \to \infty$ к $f(x)$ равномерно.*

Для доказательства достаточно положить $f_n(x)=\sum\limits_{k=-\infty}^{+\infty}\frac{k}{n}\chi_{A_k^n}(x)$, где

$$A_k^n=\left\{x:x\in X,\ \frac{k}{n}<f(x)\le\frac{k+1}{n}\right\}=f^{-1}\left\{\left(\frac{k}{n},\frac{k+1}{n}\right]\right\}.$$

Тогда, если $x_0\in X$ – произвольная точка, такая, что $\frac{k_0}{n}<f(x_0)\le\frac{k_0+1}{n}$, то $f_n(x_0)=\frac{k_0}{n}$, откуда следует, что

$$\left|f_n(x_0)-f(x_0)\right|\le\frac{1}{n}.$$

Интеграл Лебега от произвольной измеримой функции $f(x)$ определяется следующим образом:

$$\int\limits_X f(x)\mu(dx)=\lim_{n\to\infty}\int\limits_X f_n(x)\mu(dx)=\lim_{n\to\infty}\sum_{k=-\infty}^{\infty}\frac{k}{n}\mu\left\{x:\frac{k}{n}<f(x)\le\frac{k+1}{n}\right\}.$$

Основные свойства этого интеграла аналогичны свойствам интеграла Римана. Приведем некоторые из них:

1) $\int\limits_X c\mu(dx)=c\mu(X)$, где c – константа;

2) $\int\limits_X cf(x)\mu(dx)=c\int\limits_X f(x)\mu(dx)$;

3) $\int\limits_X [f_1(x)+f_2(x)]\mu(dx)=\int\limits_X f_1(x)\mu(dx)+\int\limits_X f_2(x)\mu(dx)$;

4) если $f_1(x)\ge f_2(x)$, то $\int\limits_X f_1(x)\mu(dx)\ge\int\limits_X f_2(x)\mu(dx)$;

5) если $X=\bigcup\limits_{i=1}^{\infty}A_i,\ A_i\in F,\ A_i\cap A_{\mathrm{j}}=\varnothing,\ i=j$, то

$$\int\limits_{X} f(x)\mu(dx)=\sum_{i=1}^{\infty}\int\limits_{A_i} f(x)\mu(dx).$$

Пусть теперь $f=(f_1, f_2, \ldots, f_n)$ – вектор-функция, заданная на $(X,\ \mathsf{F}\ \mu)$, компоненты которой являются измеримыми функциями. Она индуцирует на борелевских множествах пространства R^n меру μ_f, заданную равенством

$$\mu_f(B)=\mu\{x : x\in X, f(x)=(f_1(x), f_2(x), \ldots, f_n(x))\in B\}=\mu\{f^{-1}(B)\}.$$

Теорема 1.9 ***(формула замены переменных в интеграле Лебега).*** *Пусть* $g(x)=g(x_1, x_2, \ldots, x_n)$ *– борелевская функция, заданная на* R^n*. Тогда имеет место следующая формула:*

$$\int\limits_{X} g(f(x))\mu(dx)=\int\limits_{R^n} g(y)\mu_f(dy).$$

Для доказательства левую и правую часть этого равенства приведем к одному и тому же выражению. Тогда будем иметь:

$$\int\limits_{X} g(f(x))\mu(dx)=\lim_{n\to\infty}\sum_{k=-\infty}^{\infty}\frac{k}{n}\mu\left\{x : x\in X, \frac{k}{n}<g(f(x))\le\frac{k+1}{n}\right\}=$$

$$=\lim_{n\to\infty}\sum_{k=-\infty}^{\infty}\frac{k}{n}\mu\left\{x : x\in X, f(x)\in g^{-1}\left\{\left(\frac{k}{n},\frac{k+1}{n}\right]\right\}\right\}=\lim_{n\to\infty}\sum_{k=-\infty}^{\infty}\frac{k}{n}\mu_f\left\{g^{-1}\left(\frac{k}{n},\frac{k+1}{n}\right]\right\},$$

$$\int\limits_{R^n} g(y)\mu_f(dy)=\lim_{n\to\infty}\sum_{k=-\infty}^{\infty}\frac{k}{n}\mu_f\left\{y : y\in R^n, \frac{k}{n}<g(y)\le\frac{k+1}{n}\right\}=$$

$$=\lim_{n\to\infty}\sum_{k=-\infty}^{\infty}\frac{k}{n}\mu_f\left\{g^{-1}\left(\frac{k}{n},\frac{k+1}{n}\right]\right\}.$$

Перейдем от $(X,\ \mathsf{F},\ \mu)$ к вероятностному пространству $(\Omega,\ \mathsf{F},\ P)$. Заметим, что измеримая функция $\xi(\omega)$, которую мы назвали случайной величиной и которая задается на $(\Omega,\ \mathsf{F},\ P)$, индуцирует на борелевских множествах $\hat{A} \subset R$ меру $\tilde{N}_{\hat{\imath}} = \grave{\imath}_{\hat{\imath}}$, для которой

$$\tilde{N}_{\hat{\imath}}(\hat{A}) = \tilde{N}\{\grave{u} : \hat{\imath}(\grave{u}) \in \hat{A}\} = \tilde{N}\left(\hat{\imath}^{-1}(\hat{A})\right).$$

Если $\hat{A} = (x, x + \Delta x]$, то

$$\tilde{N}(\Delta\grave{u}) = \tilde{N}\{\grave{u} : x < \hat{\imath}(\grave{u}) \le x + \Delta x\} = \tilde{N}_{\hat{\imath}}((x, x + \Delta x]) =$$

$$= F_{\hat{\imath}}(x + \Delta x) - F_{\hat{\imath}}(x) = \Delta F_{\hat{\imath}}(x)\cdot$$

Из предыдущей теоремы следует

$$\int_{\grave{U}} g(\hat{\imath}(\grave{u}))\tilde{N}(d\grave{u}) = \int_R g(x) dF_{\hat{\imath}}(x),$$

что означает переход от интеграла Лебега по вероятностной мере к интегралу Лебега–Стилтьеса, в котором интегрирующей функцией является функция распределения.

Интегралом Лебега–Стилтьеса называют интеграл Лебега по мере Лебега–Стилтьеса. Мерой Лебега–Стилтьеса называется неотрицательная счетно-аддитивная функция $\grave{\imath}(\mathrm{A})$ множеств $\mathrm{A} \subset R$, для которой выполняется дополнительное условие:

$$\grave{\imath}((a, b]) = F(b) - F(a),$$

где $F(x)$ – неубывающая, непрерывная справа функция. Вероятность является частным случаем меры Лебега–Стилтьеса, при этом $F_{\hat{\imath}}(-\infty) = 0,\ F_{\hat{\imath}}(+\infty) = 1.$

Рассмотрим определение **интеграла Римана–Стилтьеса** от функции $g(x)$ с интегрирующей функцией $F(x)$ по ограниченному интервалу $[a, b]$. Пусть

$$a = x_0 < x_1 < \ldots < x_n = b,\ \Delta = \max_{1 \le k \le n}(x_k - x_{k-1})\cdot$$

Предел

$$\lim_{\substack{n \to \infty \\ \Delta \to 0}} \sum_{k=1}^{n} g(x_k)[F(x_k) - F(x_{k-1})] = \int_a^b g(x) dF(x),$$

если он существует, называется интегралом Римана–Стилтьеса. Заметим, что если функция $g(x)$ непрерывна на интервале $[a,b]$, то интеграл Римана–Стилтьеса совпадает с интегралом Лебега–Стилтьеса по этому интервалу. Разобьем интервал $[a,b]$ на малые части $\mathrm{A}_i^n = \left(x_i^{(n)}, x_{i+1}^{(n)}\right]$ точками $a = x_1^{(n)} < x_2^{(n)} < \ldots < x_{n+1}^{(n)} = b.$ Когда $g(x)$ – непрерывная функция, то последовательность

$$g_n(x) = \sum_{i=1}^{n} g\left(x_{i=1}^{(n)}\right)\text{É}_{\mathrm{A}_i^n}(x)$$

равномерно к ней сходится, если $\left|x_{i+1}^{(n)} - x_i^{(n)}\right| \xrightarrow[n\to\infty]{} 0$ для всех i. Значит,

$$(\textit{интеграл Л.-С-}) \int_a^b g(x)dF(x) = \lim_{n\to\infty} \sum_{i=1}^{n} g\left(x_{i+1}^{(n)}\right)\mu\left(A_i^n\right) =$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} g\left(x_{i+1}^{(n)}\right)\left[F\left(x_{i+1}^{(n)}\right) - F\left(x_i^{(n)}\right)\right] = \int_a^b g(x)dF(x)\ (\textit{интеграл Р.-С-}).$$

Слева в данной формуле стоит интеграл Лебега–Стилтьеса, а справа – Римана–Стилтьеса.

При определении несобственных интегралов Римана–Стилтьеса поступают таким же образом, как и в случае определения несобственных интегралов в обычном смысле Римана. В этом случае для того, чтобы интегралы Лебега–Стилтьеса и Римана–Стилтьеса совпадали, достаточно, чтобы последний существовал в смысле абсолютной сходимости.

Если $F_{\text{î}}(x)$ абсолютно непрерывная функция, то

$$\int_{\text{Ù}} g(\text{î}(\text{ù}))\text{Ñ}(d\text{ù}) = \int_R g(x)dF_{\text{î}}(x) = \int_R g(x)p_{\text{î}}(x)dx,$$

и вычисление рассматриваемых интегралов сводится к вычислению обычных интегралов Римана.

Следует отметить, что в основе построения интегралов Лебега и Римана лежат разные идеи. При построении интегралов Лебега исходят из разбиения значений интегрируемой функции и определения слагаемых в интегральной сумме через меру множества точек, которым соответствуют близкие значения интегрируемой функции. В интегральной сумме Римана исходят из разбиения значений переменной и определения слагаемых через меру множества точек, которым соответствуют близкие значения переменной. Результатом этих разных подходов является то, что соответствующие интегральные суммы Римана имеют предел лишь для не слишком разрывных функций, а интегральные суммы Лебега сходятся к предельным значениям для более широкого класса функций. Если ф.р. $F_{î}(x)$ не имеет сингулярной компоненты, как это обычно бывает в приложениях, то интегралы Лебега–Стилтьеса и Римана–Стилтьеса по этой интегрирующей функции совпадают, если они существуют.

Используя интегралы, рассмотренные выше, некоторые результаты, сформулированные ранее для абсолютно непрерывных ф.р. можно записать в более общей форме:

1) для любой СВ $î(ù)$ (если она и не имеет абсолютно непрерывной ф.р.)

$$Ñ\{ù : î(ù) \in G\} = \int_G dF_{î}(x);$$

2) для любых СВ $î(ù)$, $ç(ù)$ (они могут не иметь абсолютно непрерывные ф.р.)

$$F_{ç}(y) = \int_{-\infty}^{+\infty} F_{ç}(y/x)\, dF_{î}(x);$$

3) если $ç(ù) = î_1(ù) + î_2(ù)$, то

$$F_{ç}(y) = \int_{-\infty}^{+\infty} F_{î_1}(y - x)\, dF_{î_2}(x), \text{ т.е. } F_{ç} = F_{î_1} * F_{î_2},$$

здесь также не обязательно, чтобы СВ $î_1(ù), î_2(ù)$ имели абсолютно непрерывные ф.р.

§6. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

Перед тем как дать формальное определение первой числовой характеристики математического ожидания, рассмотрим несколько примеров.

Пример 1.35 *(доход телефонной станции)*. На телефонной станции в любой момент x в течение промежутка времени É фиксируется число $f(x)$ телефонных разговоров. Цена ì одного разговора за промежуток времени является функцией множества, которая ставит в соответствие каждому интервалу $É_s \subset É$ цену одного разговора за время $É_s$, $\mu_i : I_s \to \mu(I_s)$. Необходимо найти ожидаемый доход телефонной станции за промежуток времени É.

Для этого заметим, что $f(x)$ принимает конечное число значений. Пусть f_i – возможные значения $f(x)$, а $\{x : f(x) = f_i\}$ – множество моментов времени, в которых велось ровно f_i разговоров, $i = \overline{1, n}$. Тогда ожидаемый доход будет равен

$$\sum_{i=1}^{n} f_i \mu\{x : f(x) = f_i\}.$$

Таким образом, эта величина является интегралом Лебега от простой функции $f(x)$, принимающей конечное число значений, по мере ì. Его принято обозначать $\int_{É} f(x) ì(dx)$.

Пример 1.36 *(урожай с поля)*. Пусть $f(x), x \in Q$, – урожайность пшеницы на поле Q, например, Q – участок посева семян района, области под пшеницей; ì – площадь, т.е. функция множеств (участка), которая ставит в соответствие каждому участку $Q_i \subset Q$ число, равное его площади $\mu(Q_i)$, $\mu : Q_i \to \mu(Q_i)$. Необходимо найти величину урожая с поля Q. Здесь в отличие от предыдущего примера значения функции $f(x)$ непрерывно заполняют некоторый промежуток.

В качестве единицы измерения урожайности выберем число ε, например, $\varepsilon = 1$ *ц*, 10 *ц* и т.д. Пусть $\mathrm{A}_{k,\varepsilon} = \{x : k\varepsilon < f(x) \le (k+1)\varepsilon\}$ – участок поля, на котором урожайность меняется от $k\varepsilon$ до $(k+1)$å. Под урожайностью на $\mathrm{A}_{k,\varepsilon}$ будем понимать величину kå. Рассмотрим на Q функцию $f_\varepsilon(x)$, принимающую конечное число значений, а именно, $f_\varepsilon(x)=k\varepsilon$ для $x\in A_{k,\varepsilon}$, т.е. $f_\varepsilon(x)$ – величина урожайности, измеренная с точностью до ε. В качестве величины урожая с поля Q естественно рассматривать величину

$$\sum_k k\varepsilon\mu\{x : k\varepsilon < f(x)\le (k+1)\varepsilon\} -$$

интеграл Лебега от функции $f(x)$ по мере μ, т.е. $\int\limits_Q f(x)\mu(dx)$.

***Пример 1.37** (среднее значение СВ).* Пусть $\xi(\omega)$ – СВ со значениями на числовой оси R, заданная на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Нужно найти среднее значение СВ – значение, которое СВ принимает в среднем.

Как и в предыдущем примере, сначала выберем единицу измерения ε значений СВ $\xi(\omega)$ и рассмотрим множество точек $A_{k,\varepsilon}=\{\omega : k\varepsilon<\xi(\omega)\le(k+1)\varepsilon\}$, т.е. $A_{k,\varepsilon}$ – множество точек пространства элементарных событий Ω, в которых $\xi(\omega)$ принимает значения от $k\varepsilon$ до $(k+1)\varepsilon,\ k=1,2,\ldots$. На пространстве Ω рассмотрим последовательность СВ $\xi_\varepsilon(\omega)$, которые зависят от параметра ε, положив $\xi_\varepsilon(\omega)=k\varepsilon$ для $\omega\in A_{k,\varepsilon},\ k=1,2,\ldots$. В качестве среднего значения СВ $\xi(\omega)$ естественно рассматривать величину

$$\lim_{\varepsilon\to 0}\sum_k k\varepsilon P\{\omega : k\varepsilon<\xi(\omega)\le(k+1)\varepsilon\}=\int\limits_\Omega \xi(\omega)P(d\omega),$$

являющуюся интегралом Лебега от СВ $\xi(\omega)$ по вероятностной мере P, которую принято называть математическим ожиданием СВ $\xi(\omega)$ (средним значением СВ $\xi(\omega)$)

Определение. Возьмем в качестве ε величину $\frac{1}{n}$. Таким образом, математическим ожиданием СВ $\xi=\xi(\omega)$, заданной на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, называется число, равное

$$M\xi=\lim_{n\to\infty}\sum_{k=-\infty}^{\infty}\frac{k}{n}P\left\{\omega : \frac{k}{n}<\xi(\omega)\le\frac{k+1}{n}\right\}=\int\limits_\Omega \xi(\omega)P(d\omega)=\int\limits_X x\,dF_\xi(x),$$

где $F_\xi(x)$ – ф.р. СВ ξ.

Из этого определения вытекают следствия:

а) если $\xi(\omega)$ – дискретная СВ, которая принимает значения из множества $\{x_1, x_2, ..., x_k, ...\}$, то

$$M\xi = \sum_{k=1} x_k P\{\omega : \xi(\omega) = x_k\} = \sum_{k=1} x_k p_k \, ;$$

б) если $\xi(\omega)$ имеет абсолютно непрерывное распределение (плотность $p_\xi(x)$), то

$$M\xi = \int_{-\infty}^{+\infty} x p_\xi(x) dx \, ;$$

в) математическое ожидание $M\xi$ существует, если $M\,|\xi| < +\infty$.

Пример 1.38. Найдем математическое ожидание СВ ξ, распределенной по закону Пуассона.

Решение. Для такой СВ

$$\xi(\omega) \in \{0, 1, 2, ..., k, ...\}, \; p_k = \frac{\lambda^k}{k!} e^{-\lambda}, \; k = 0, 1, 2, ... \cdot$$

Поэтому

$$M\xi = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda} \lambda \sum_{r=0}^{\infty} \frac{\lambda^r}{r!} = e^{-\lambda} \lambda e^{\lambda} = \lambda \cdot$$

Пример 1.39. Пусть ξ принимает значения $x_k = (-1)^k 2^k$ с вероятностями $p_k = \dfrac{1}{2^k}, k = 1, 2, ...$. Попарно просуммировав последовательные значения $x_k p_k$, получим

$x_{2k-1} p_{2k-1} + x_{2k} p_{2k} = 0, k = 1, 2, ...$. На первый взгляд может показаться, что $M\xi = 0$. Однако

$$\sum_{\{k \,:\, x_k > 0\}} x_k p_k = \sum_{k=1}^{\infty} 2^{2k} 2^{-2k} = \infty, \; \sum_{\{k \,:\, x_k < 0\}} x_k p_k = -\sum_{n=1}^{\infty} 2^{2k-1} 2^{1-2k} = -\infty,$$

поэтому $M\xi$ не существует.

Пример 1.40. Найдем математическое ожидание СВ ξ, равномерно распределенной на отрезке $[a,b]$.

Решение. Данная СВ непрерывного типа. Учитывая для нее вид плотности распределения, будем иметь

$$\text{Ìî} = \int_a^b \frac{x}{b-a}\,dx = \frac{a+b}{2},$$

т.е. математическое ожидание равно середине отрезка $[a,b]$.

Пример 1.41. СВ ξ имеет гамма-распределение с параметром p (см. пример 1.28.). Найдем Ìî.

Решение.

$$M\xi = \int_0^\infty \frac{x}{\Gamma(p)} x^{p-1} e^{-x}\,dx = \frac{1}{\Gamma(p)} \int_0^\infty x^p e^{-x}\,dx = \frac{\Gamma(p+1)}{\Gamma(p)} = \frac{p\Gamma(p)}{\Gamma(p)} = p$$

для целых p.

Поскольку математическое ожидание определяется в общем случае в виде интеграла Лебега, то его свойства в основном следуют из свойств этого интеграла. Рассмотрим основные из них (будем предполагать, что соответствующие математические ожидания, входящие в приведенные в данных свойствах соотношения, существуют).

1. Если $P(\omega : \xi(\omega) \geq 0) = 1$, то $\text{Ìî} \geq 0$.
2. $\text{Ì}\,(c\text{î}) = c\text{Ì}\,\xi$, где c – константа, $\text{Ìc} = c$.
3. Если $P(\omega : \xi(\omega) \geq \eta(\omega)) = 1$, то $\text{Ìî} \geq \text{Ìç}$.
4. $\text{Ì}\,(\text{î} + \text{ç}) = \text{Ìî} + \text{Ìç}$.
5. $|\text{Ìî}| \leq \text{Ì}\,|\text{î}|$.

6. Пусть $I_A(\omega)=\begin{cases}1, & \omega\in A\\ 0, & \omega\notin A\end{cases}$, $A\in$ F. Такая СВ называется индикатором случайного события A. Тогда

$$MI_A(\omega)=\int_\Omega I_A(\omega)P(d\omega)=\int_A P(d\omega)=P(A),$$

т.е. математическое ожидание такой СВ равно вероятности события A.

7. Пусть ξ и η – независимые СВ, тогда $M(\xi\eta)=M\xi M\eta$.

Докажем это свойство. Рассмотрим два случая.

а) Пусть ξ и η – дискретные СВ, $\xi\in\{a_1,a_2,...\}$, $\eta\in\{b_1,b_2,...\}$.

Тогда, введя случайные события

$A_i=\{\omega:\xi(\omega)=a_i\}$, $B_j=\{\omega:\eta(\omega)=b_j\}$, можно с помощью их записать наши СВ

$$\xi(\omega)=\sum_i a_i I_{A_i}(\omega),\ \eta(\omega)=\sum_j b_j I_{B_j}(\omega),$$

и поэтому в этом случае будем иметь

$$M\xi M\eta=\left[\sum_i a_i P(A_i)\right]\left[\sum_j b_j P(B_j)\right]=$$

$$=\sum_{i,j} a_i b_j P(A_i)P(B_j)=\sum_{i,j} a_i b_j P(A_i B_j)=M(\xi,\eta).$$

б) Пусть теперь ξ и η – произвольные СВ. Введем дискретные СВ

$$\xi_n(\omega)=\frac{k}{n},\text{ если } \omega\in\left\{\omega:\frac{k}{n}<\xi(\omega)\le\frac{k+1}{n}\right\},$$

$$\eta_n(\omega)=\frac{k}{n},\text{ если } \omega\in\left\{\omega:\frac{k}{n}<\eta(\omega)\le\frac{k+1}{n}\right\},\ k\in\{...,-2,-1,0,1,2,...\}.$$

Они независимы, как функции от независимых СВ, поэтому $M(\xi_n\eta_n)=M\xi_n M\eta_n$. Кроме того, как было показано в предыдущем параграфе, $\xi_n \xrightarrow[n\to\infty]{} \xi$, $\eta_n \xrightarrow[n\to\infty]{} \eta$ равномерно. Рассмотрим разность

$$M(\xi\eta) - M(\xi_n\eta_n) = M(\xi\eta - \xi_n\eta_n) =$$
$$= M(\xi\eta - \xi_n\eta + \xi_n\eta - \xi_n\eta_n) = M[(\xi - \xi_n)\eta] + M[\xi_n(\eta - \eta_n)].$$

При $n \to \infty$ $|\xi - \xi_n| \le \frac{1}{n}$, т.е. $|\xi_n| \le |\xi| + \frac{1}{n}$, $|\eta - \eta_n| \le \frac{1}{n}$, следовательно,

$$|M(\xi\eta) - M(\xi_n\eta_n)| \le \frac{1}{n} M|\eta| + \frac{1}{n}\left(M|\xi| + \frac{1}{n}\right) \xrightarrow[n\to\infty]{} 0,$$

что и доказывает свойство в данном случае.

8. Пусть ξ – СВ, $g(x)$ – измеримая функция. Тогда $g(\xi)$ также является СВ и, если $M|g(\xi)| < +\infty$, то

$$Mg(\xi) = \int\limits_R g(x)dF_\xi(x),$$

в частности,

$$Mg(\xi) = \sum_{k=1}^{\infty} g(x_k)p_k, \text{ если } \xi \text{ – дискретная СВ},$$

$$Mg(\xi) = \int\limits_{-\infty}^{+\infty} g(x)p_\xi(x)dx, \text{ если } \xi \text{ – непрерывная СВ}.$$

Пример 1.42. Правомерно ли следующее рассуждение: «От общежития до университета расстояние равно 1 км, студент ходит в среднем со скоростью 5 км/ч, следовательно, в среднем на дорогу у него будет уходить 12 мин.»?

Решение. Прежде всего уточним формулировку задачи: имеется в виду, что скорость – случайная величина, математическое ожидание которой равно 5 км/ч. Рассуждение является неправомерным, т.к. из свойства 8 следует, что, вообще говоря, $M\frac{1}{\xi} \ne \frac{1}{M\xi}$.

Для многомерной СВ $\xi = (\xi_1, \xi_2, ..., \xi_n)$ математическое ожидание определяется в виде вектора

$$M\xi = (m_1, m_2, ..., m_n), \text{ где } m_k = M\xi_k = \int_{\Omega} \xi_k(\omega) P(d\omega), k = \overline{1, n}.$$

Если, например, СВ ξ – дискретного типа и

$$\xi_k \in \{0, 1, 2, ..., N\}, k = \overline{1, n}, \sum_{k=1}^{n} \xi_k(\omega) = N \ \forall \omega \in \Omega, \text{ то}$$

$$m_k = \sum_{x_1=0}^{N} \sum_{x_2=0}^{N-x_1} ... \sum_{x_n=0}^{N-x_1-x_2-...-x_{n-1}} x_k P\{\xi_1 = x_1, \xi_2 = x_2, ..., \xi_n = x_n\}, k = \overline{1, n}.$$

Если СВ $\xi = (\xi_1, \xi_2, ..., \xi_n)$ непрерывного типа и $g(x_1, x_2, ..., x_n)$ – измеримая функция, то

$$Mg(\xi_1, \xi_2, ..., \xi_n) = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} g(x_1, x_2, ..., x_n) p_{\xi_1 \xi_2 ... \xi_n}(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n,$$

где $p_{\xi_1 \xi_2 ... \xi_n}(x_1, x_2, ..., x_n)$ – плотность распределения n-мерной СВ ξ. В частности,

$$m_k = M\xi_k = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} x_k p_{\xi_1 \xi_2 ... \xi_n}(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n.$$

Рассмотрим понятие **условного математического ожидания**. Ранее было введено определение условной ф.р. (см. п.7):

$$F_{\xi}(x/B) = \frac{P(\omega : \xi(\omega) \le x, \omega \in B)}{P(\omega : \omega \in B)},$$

причем в качестве Â можно взять событие $B = \{\omega : \eta(\omega) = y\}$, где ç(ù) – некоторая СВ. Условным математическим ожиданием СВ $\xi(\omega)$ при условии события Â называется

$$M(\xi/B) = \int_{R} x dF_{\xi}(x/B) = \int_{\Omega} \xi(\omega) P(d\omega / B) = (P(B))^{-1} \int_{B} \xi(\omega) P(d\omega).$$

Если случайные события $B_1, B_2, \ldots, B_k, \ldots$ образуют полную группу событий, то

$$M\xi = \int_{\Omega} \xi(\omega) P(d\omega) = \sum_{k=1}^{\infty} \int_{B_k} \xi(\omega) P(d\omega) = \sum_{k=1}^{\infty} P(B_k) M(\xi / B_k).$$

Формула $\grave{I}\hat{\imath} = \sum_{k=1}^{\infty} \tilde{N}(\hat{A}_k)\grave{I}\,(\hat{\imath}/\hat{A}_k)$ называется формулой полной вероятности для математического ожидания.

Пусть $\hat{\imath}(\grave{u})$ – неотрицательная СВ, определенная на вероятностном пространстве $(\Omega,\ \mathsf{F}, P)$, $\hat{a}$ – σ-алгебра, содержащаяся в F, и пусть $\grave{I}\hat{\imath} < +\infty$. Условное математическое ожидание СВ ξ относительно σ-алгебры $\hat{a}$ определяется как СВ $M(\xi / \mathrm{B})$, удовлетворяющая условиям:

а) $M(\xi /\ \hat{a}\ )$ измерима относительно $\hat{a}$;

б) $\forall B \in\ \hat{a}\quad \int_B \xi(\omega) P(d\omega) = \int_B M(\xi / \mathrm{B}) P(d\omega)$.

Условным математическим ожиданием СВ относительно СВ η называется условное математическое ожидание ξ относительно σ-алгебры, порожденной $\c{c}$.

Пример 1.43. Пусть СВ ξ и $\c{c}$ независимы и имеют биномиальные распределения с параметрами n, p и m, p соответственно. Найдем условное математическое ожидание $\grave{I}\,(\hat{\imath}/\hat{\imath}+\c{c})$.

Решение. СВ $\xi + \eta$ имеет биномиальное распределение с параметрами $n+m, p$ (см. пример 8.6.).Найдем условное распределение $\hat{\imath}$ при условии $\hat{\imath}+\c{c}$:

$$\tilde{N}\{\hat{\imath} = k / \hat{\imath}+\c{c} = l\} = \frac{\tilde{N}\{\hat{\imath} = k, \hat{\imath}+\c{c} = l\}}{\tilde{N}\{\hat{\imath}+\c{c} = l\}} =$$

$$= \frac{\tilde{N}\{\hat{\imath} = k, \c{c} = l-k\}}{\tilde{N}\{\hat{\imath}+\c{c} = l\}} = \frac{\tilde{N}\{\hat{\imath} = k\}\tilde{N}\{\c{c} = l-k\}}{\tilde{N}\{\hat{\imath}+\c{c} = l\}} =$$

$$=\frac{C_n^k p^k (1-p)^{n-k} C_m^{l-k} p^{l-k} (1-p)^{m-l-k}}{C_{n+m}^l p^l (1-p)^{n+m-l}}=\frac{C_n^k C_m^{l-k}}{C_{n+m}^l}.$$

Поэтому, используя тождество из примера 1.31, получаем

$$M(\xi/\xi+\eta=l)=\sum_k k\,\frac{C_n^k C_m^{l-k}}{C_{n+m}^l}=n\sum_k\frac{C_{n-1}^{k-1}C_{(n+m-1)-(n-1)}^{(l-1)-(k-1)}}{\frac{(n+m)}{l}C_{n+m-1}^{l-1}}=\frac{l\cdot n}{n+m},$$

$$\text{Ì}\ (\text{î}/\text{î}\!+ n)=\frac{(\text{î}+\text{ç})n}{n+m}.$$

Рассмотрим неравенства, справедливые для математических ожиданий.

1) **Неравенство Чебышева.** Пусть $f(x)$ – неотрицательная, монотонно неубывающая борелевская функция, определенная на интервале $[0,+\infty)$ со значениями в R. Тогда для любой СВ ξ и $\forall\ \varepsilon>0$

$$\text{Ñ}\langle|\text{î}|\geq \text{å}\rangle\leq\frac{\text{Ì}f(|\text{î}|)}{f(\text{å})},$$

в частности,

$$\text{Ñ}\{|\text{î}|\geq \text{å}\}\leq\frac{\text{Ì}\ |\text{î}|^k}{\text{å}^k}\ \forall k\geq 0\ –$$

это неравенство называется **неравенством Маркова**. При $k=1$ имеем:

$$\text{Ñ}\{|\text{î}|\geq \text{å}\}\leq\frac{\text{Ì}\ |\text{î}|}{\text{å}}.$$

Неравенство Чебышева справедливо, поскольку

$$Mf(|\xi|)=\int_\Omega f(|\xi(\omega)|)P(d\omega)=\int_{\{\omega:|\xi(\omega)|\geq\varepsilon\}\cup\{\omega:|\xi(\omega)|<\varepsilon\}} f(|\xi(\omega)|)P(d\omega)=$$

$$= \int\limits_{\{\omega : |\xi(\omega)| \ge \varepsilon\}} f(|\xi(\omega)|)P(d\omega) + \int\limits_{\{\omega : |\xi(\omega)| < \varepsilon\}} f(|\xi(\omega)|)P(d\omega) \ge$$

$$\ge \int\limits_{\{\omega : |\xi(\omega)| \ge \varepsilon\}} f(|\xi(\omega)|)P(d\omega) \ge \int\limits_{\{\omega : |\xi(\omega)| \ge \varepsilon\}} f(\varepsilon)P(d\omega) = f(\varepsilon)P\{|\xi| \ge \varepsilon\}.$$

2) **Неравенство Иенсена.** Напомним, что функция $g(x)$ со значениями в R, заданная на интервале $I \subset R$, называется выпуклой на нем вниз (вверх), если для любого $x_0 \in I$ найдется число $\lambda(x_0)$ такое, что для всех $x \in I$:

$$g(x_0) + \lambda(x_0)(x - x_0) \le g(x) \quad (g(x_0) + \lambda(x_0)(x - x_0) \ge g(x)).$$

Т.к. $y = g(x_0) + \lambda(x_0)(x - x_0)$ является уравнением прямой, которая проходит через точку $(x_0, g(x_0))$ с угловым коэффициентом $\lambda(x_0)$, то выпуклость вниз (вверх) означает, что в любой точке $(x_0, g(x_0))$ можно провести касательную прямую так, что график функции $g(x)$ лежит выше (ниже) этой прямой.

Пусть $g(x)$ – выпуклая вниз (вверх) борелевская функция на R. Тогда, если $M|\xi| < +\infty$, выполняется неравенство Иенсена

$$g(M\xi) \le Mg(\xi) \quad (g(M\xi) \ge Mg(\xi)),$$

в частности,

$$(M\xi)^2 \le M\xi^2, \quad |M\xi|^r \le M|\xi|^r, \; r \ge 1.$$

Поясним, почему неравенство Иенсена имеет место, например, для выпуклой вниз функции $g(x)$. Для такой функции

$$g(x_0) + \lambda(x_0)(x - x_0) \le g(x), x \in R.$$

Положим $x_0 = M\xi$, $x = \xi$, тогда

$$g(M\xi) + \lambda(M\xi)(\xi - M\xi) \le g(\xi).$$

Взяв в этом неравенстве математическое ожидание в левой и правой части, получим требуемое неравенство.

3) **Неравенство Гёльдера.** Пусть $\frac{1}{p} + \frac{1}{q} = 1, p > 0, q > 0$, тогда

$$M|\xi\eta| \le \left(M|\xi|^p\right)^{\frac{1}{p}} \left(M|\eta|^q\right)^{\frac{1}{q}}.$$

Если в неравенстве Гёльдера положить $\xi = |\xi|^s$, $\eta = 1$, $p = \frac{t}{s}$, $\frac{1}{q} = 1 - \frac{s}{t}$, где $0 < s < t$, то получим **неравенство Ляпунова**:

$$\left(M|\xi|^s\right)^{\frac{1}{s}} \le \left(M|\xi|^t\right)^{\frac{1}{t}}.$$

При $p = q = 2$ из неравенства Гёльдера получаем **неравенство Коши–Буняковского (Шварца)**:

$$M|\xi\eta| \le \sqrt{M\xi^2}\sqrt{M\eta^2}.$$

Докажем неравенство Гёльдера. Введем CB

$$\alpha = \frac{\xi}{\left(M\xi^p\right)^{\frac{1}{p}}}, \quad \beta = \frac{\eta}{\left(M\eta^q\right)^{\frac{1}{q}}}.$$

Тогда нам требуется доказать, что $M\alpha\beta \le 1$. Рассмотрим функцию $\ln x$. Она является выпуклой кверху функцией, поэтому $\forall\ \lambda \in [0, 1]$

$$\ln[(1-\lambda)x_1 + \lambda x_2] \ge (1-\lambda)\ln x_1 + \lambda \ln x_2 = \ln\left(x_1^{1-\lambda} x_2^{\lambda}\right),\ x_1, x_2 > 0,$$

т.е. справедливо неравенство

$$(1-\lambda)x_1 + \lambda x_2 \ge x_1^{1-\lambda} x_2^{\lambda}.$$

Положим $x_1 = \alpha^p,\ x_2 = \beta^q,\ \lambda = \frac{1}{q},\ (1-\lambda) = \frac{1}{p}$. Тогда из этого неравенства будем иметь

$$\frac{1}{p}\alpha^p + \frac{1}{q}\beta^q \ge \alpha\beta,$$

откуда следует, что

$$M(\alpha\beta) \le \frac{1}{p}M\alpha^p + \frac{1}{q}M\beta^q = \frac{1}{p} + \frac{1}{q} = 1,$$

что и требовалось доказать.

Пример 1.44. Показать, что

$$\tilde{N}\left\{\left|\sum_{i=1}^{n}\hat{\imath}_i\right|\geq\mathring{a}\right\}\leq\frac{1}{\mathring{a}}\sum_{i=1}^{n}\grave{I}\;\left|\hat{\imath}_i\right|.$$

Решение. Применив неравенство Чебышева и свойство 5 математического ожидания, получим

$$\tilde{N}\left\{\left|\sum_{i=1}^{n}\hat{\imath}_i\right|\geq\mathring{a}\right\}\leq\frac{1}{\mathring{a}}\grave{I}\;\left|\sum_{i=1}^{n}\hat{\imath}_i\right|\leq\frac{1}{\mathring{a}}\sum_{i=1}^{n}\grave{I}\;\left|\hat{\imath}_i\right|\cdot$$

Приведем практический пример, иллюстрирующий, как можно использовать понятие математического ожидания.

Пример 1.45. Продавец получает N единиц товара в день и стремится заказать число N таким образом, чтобы максимизировать ожидаемую прибыль. Найти оптимальный параметр N^*, если число покупателей товара в данный день следует закону Пуассона с параметром $\ddot{e}>0$. Прибыль, получаемая от единицы проданного товара, равна a, убыток от нереализованной единицы товара – b, c – убыток, если покупатель желает приобрести товар, но запас товаровисчерпан.

Решение. Пусть $\hat{\imath}$ – число покупателей в данный день, тогда чистая прибыль продавца имеет вид

$$g_n(\xi)=\begin{cases}a\xi-b(N-\xi),\text{если }\ \xi\leq N,\\ aN-c(\xi-N),\text{если }\ \xi>N.\end{cases}$$

Ожидаемая прибыль является математическим ожиданием $G_N=Mg_N(\xi)$. Изменение ожидаемой прибыли при добавлении еще одной единицы товара равно

$$G_{N+1}-G_N=M\left[g_{N+1}(\xi)-g_N(\xi)\right]=M\left[-b+(a+b+c)u(\xi-N)\right],$$

где $u(x)=\begin{cases}1, x>0\\ 0, x\leq 0\end{cases}$ – функция Хевисайда. Поскольку

$$\grave{I}u\,(\c{c})=1\cdot\tilde{N}\{\c{c}>0\}+0\cdot\tilde{N}\{\c{c}\leq 0\}=\tilde{N}\{\c{c}>0\},$$

то приращение ожидаемой прибыли равно

$$G_{N+1} - G_N = -b + (a+b+c)\text{Ìu}(\text{î} - N) = -b + (a+b+c)\text{Ñ}\{\text{î} > N\}.$$

Для нахождения $\max\limits_N G_N$ достаточно решить уравнение $G_N = G_{N+1}$, откуда получаем

$$\text{Ñ}\{\text{î} > N\} = \frac{b}{a+b+c}, \text{ т.е. } 1 - F_{\text{î}}(N) = \frac{b}{a+b+c}.$$

Т.к. для закона Пуассона

$$F_{\text{î}}(N) = \sum_{k=0}^{N} \frac{\text{ë}^k}{k!} e^{-\text{ë}},$$

то уравнение для нахождения оптимального N^* имеет вид:

$$\sum_{k=0}^{N^*} \frac{\lambda^k}{k!} e^{-\lambda} = \frac{b}{a+b+c} + 1,$$

или, в неявной форме,

$$N^* = F_{\xi}^{-1}\left(\frac{b}{a+b+c} + 1\right).$$

Числовое значение данного выражения можно найти, воспользовавшись таблицей распределения Пуассона.

Рассмотрим теперь другие различные числовые характеристики СВ, которые являются математическими ожиданиями определенных функций от СВ.

Моменты СВ.

Определение. Начальным моментом порядка N СВ ξ называется $m_N = \text{Ìî}^N$, а центральным моментом порядка

$$N - \text{ì}_N = \text{Ì}(\text{î} - \text{Ìî})^N.$$

Определение. В случае многомерных СВ смешанным начальным моментом порядка $NCB\text{î} = (\text{î}_1, \text{î}_2, ..., \text{î}_n)$ называется

$\text{Ì}\left[\text{î}_1^{k_1} \text{î}_2^{k_2} \cdot ... \cdot \text{î}_n^{k_n}\right]$, где $\sum\limits_{i=1}^{n} k_i = N,\ k_i$ – целые неотрицательные числа. Смешанным центральным моментом порядка N называется

$$\text{Ì}\left[(\text{î}_1 - \text{Ìî}_1)^{k_1} (\text{î}_2 - \text{Ìî}_2)^{k_2} ... (\text{î}_n - \text{Ìî}_n)^{k_n}\right].$$

Пример 1.46. Найдем моменты СВ, имеющей нормальное распределение с параметрами $a,\ \sigma^2$.

Сделаем это вначале для случая $a = 0,\ \sigma = 1$, плотность распределения СВ при этом имеет вид

$$p_{\text{î}}(x) = \frac{1}{\sqrt{2\text{ð}}} e^{-\frac{x^2}{2}}, -\infty < x < +\infty,$$

поэтому

$$m_n = \text{Ìî}^{\,n} = \frac{1}{\sqrt{2\text{ð}}} \int_{-\infty}^{\infty} x^n e^{-\frac{x^2}{2}} dx \cdot$$

При нечетных n этот интеграл равен нулю как интеграл нечетной функции по симметричному промежутку интегрирования. Таким образом, остается найти m_n при четных n. Пусть $n = 2$, тогда

$$\text{Ìî}^{\,2} = \frac{1}{\sqrt{2\text{ð}}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx = \sqrt{\frac{2}{\text{ð}}} \int_{0}^{\infty} x^2 e^{-\frac{x^2}{2}} dx = \sqrt{\frac{2}{\text{ð}}} \int_{0}^{\infty} x d\left(-e^{-\frac{x^2}{2}}\right) =$$

$$-\sqrt{\frac{2}{\text{ð}}} x e^{-\frac{x^2}{2}} \Bigg|_0^{\infty} + \sqrt{\frac{2}{\text{ð}}} \int_{0}^{\infty} e^{-\frac{x^2}{2}} dx =$$

$$= \sqrt{\frac{2}{\text{ð}}} \int_{0}^{\infty} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\text{ð}}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \int_{-\infty}^{\infty} p_{\text{î}}(x) dx = 1;$$

при этом мы использовали предел $\lim\limits_{\text{÷}\to\infty} x\, e^{-\frac{x^2}{2}} dx = 0$, который легко проверяется по правилу Лопиталя. Аналогично, проинтегрировав интеграл, выражающий $\text{Ìî}^{\,n}$, по частям, получим

$$\text{Ìî}^{n}=\frac{1}{\sqrt{2\text{ð}}}\int_{-\infty}^{\infty}x^{n}e^{-\frac{x^{2}}{2}}dx=\sqrt{\frac{2}{\text{ð}}}\int_{0}^{\infty}x^{n}e^{-\frac{x^{2}}{2}}dx=\sqrt{\frac{2}{\text{ð}}}\int_{0}^{\infty}x^{n-1}d\left(-e^{-\frac{x^{2}}{2}}\right)=$$

$$=-\sqrt{\frac{2}{\text{ð}}}x^{n-1}e^{-\frac{x^{2}}{2}}\Bigg|_{0}^{\infty}+(n-1)\sqrt{\frac{2}{\text{ð}}}\int_{0}^{\infty}x^{n-2}e^{-\frac{x^{2}}{2}}dx=$$

$$=(n-1)\frac{1}{\sqrt{2\text{ð}}}\int_{-\infty}^{\infty}x^{n-2}e^{-\frac{x^{2}}{2}}dx=(n-1)\text{Ìî}^{n-2}\cdot$$

Итак, доказана рекуррентная формула

$$M\xi^{n}=(n-1)M\xi^{n-2},\ M\xi=0.$$

Отсюда

$$\text{Ìî}^{n}=(n-1)\text{Ìî}^{n-2}=(n-1)(n-3)\text{Ìî}^{n-4}=\ldots=$$

$$=\ldots=(n-1)(n-3)\cdot\ldots\cdot 3\,\text{Ìî}^{2}=(n-1)(n-3)\ldots 3\cdot 1.$$

Положим $n=2k$. Тогда

$$(n-1)(n-3)\cdot\ldots\cdot 1=(2k-1)(2k-3)\ldots 1=$$

$$=\frac{(2k-1)!}{[2(k-1)][2(k-2)]\cdot\ldots\cdot 2}=\frac{(2k-1)!}{2^{k-1}(k-1)!}.$$

Таким образом,

$$m_{2k}=\text{Ìî}^{2k}=\frac{(2k-1)!}{2^{k-1}(k-1)!},\ k=1,2,\ldots\ \cdot$$

Найдем теперь центральные моменты распределения СВ ç с плотностью

$$p_{\text{ç}}(x)=\frac{1}{\sqrt{2\text{ð}}\sigma}e^{-\frac{(x-a)^{2}}{2\sigma^{2}}},\ -\infty<x<+\infty\cdot$$

Если ввести СВ ξ по формуле $\xi=\dfrac{\eta-a}{\sigma}$, то имеем

$p_{\xi}(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, а из формулы $\eta=a+\sigma\xi$ находим

$$M\eta=a+\sigma M\xi=a.$$

Рассмотрим центральные моменты СВ η любых порядков. Т.к. $\eta-a=\sigma\xi$, то

$$M(\eta-M\eta)^n=M(\sigma\xi)^n=\sigma^n M\xi^n=\begin{cases}0,\ n=2k-1,\\ \dfrac{\sigma^{2k}(2k-1)!}{2^{k-1}(k-1)!},\ n=2k,\end{cases}\quad k=1,2,\ldots$$

Дисперсия и ее свойства.

Определение. Дисперсией СВ ξ называется ее центральный момент второго порядка:

$$D\xi=M(\xi-M\xi)^2,$$

$\sqrt{D\xi}$ называется средним квадратичным отклонением СВ ξ.

Дисперсия служит характеристикой (хотя и не полной) рассеяния значений СВ от ее среднего значения.

Из определения следует, что

$$D\xi=M\left[\xi^2-2\xi M\xi+(M\xi)^2\right]=M\xi^2-2(M\xi)^2+(M\xi)^2=M\xi^2-(M\xi)^2.$$

Рассмотрим свойства дисперсии.

1. $D\xi\geq 0$.

2. $D\xi=0$ только в том случае, когда $P(\xi=const)=1$.

3. $D(c\xi)=M(c\xi-cM\xi)^2=c^2M(\xi-M\xi)^2=c^2D\xi$.

4. $D(c+\xi)=M[c+\xi-M(c+\xi)]^2=M(\xi-M\xi)^2=D\xi$.

5. Пусть ξ и η – независимые СВ, тогда

$$D(\xi+\eta)=M[\xi+\eta-M(\xi+\eta)]^2=M[(\xi-M\xi)+(\eta-M\eta)]^2=$$

$$= \text{Ì}\,(\text{î}-\text{Ìî})^2+\text{Ì}\,(\text{ç}-\text{Ìç})^2+2\text{Ì}\,[(\text{î}-\text{Ìî})(\text{ç}-\text{Ìç})]=$$

$$= D\text{î}+D\text{ç}+2[\text{Ì}\,(\text{îç})+\text{ÌîÌç}-2\text{ÌîÌç}]=D\text{î}+D\text{ç},$$

поскольку для независимых СВ $\text{Ì}\,(\text{îç})=\text{ÌîÌç}$.

6. Если в неравенстве Чебышева для математического ожидания в качестве $f(x)$ взять x^2, а в качестве СВ взять $(\text{î}-\text{Ìî})$ (или, что то же самое, в неравенство Маркова вместо î подставить $\text{î}-\text{Ìî}$), то получим **неравенство Чебышева для дисперсии**:

$$P\{|\text{î}-\text{Ìî}|\geq \text{å}\}\leq \frac{D\text{î}}{\text{å}^2}\;\forall \text{å}>0.$$

Пример 1.47. Из предыдущего примера следует, что дисперсия СВ ç, имеющей нормальное распределение с параметрами a,σ^2, равна σ^2, т.е. $D\text{ç}=\sigma^2$.

Пример 1.48. Найдем дисперсию СВ î, равномерно распределенной на отрезке $[a,b]$.

Решение. Для такой СВ $\text{Ìî}=\dfrac{a+b}{2}$ (см. пример 1.40). Найдем начальный момент второго порядка:

$$m_2=\text{Ìî}^2=\int_{-\infty}^{\infty}x^2p_{\text{î}}(x)dx=\frac{1}{b-a}\int_a^b x^2dx=\frac{b^3-a^3}{3(b-a)}=\frac{a^3+ab+b^3}{3}.$$

Поэтому

$$D\text{î}=\text{Ìî}^2-(\text{Ìî})^2=\frac{(b-a)^2}{12}.$$

Пример 1.49 *(правило «трех сигм»).* Используя неравенство Чебышева, оценить вероятность того, что любая СВ î отклонится от своего среднего значения менее чем на три средних квадратичных отклонения этой величины.

Решение. Из неравенства Чебышева следует

$$P\{|\xi - M\xi| < \varepsilon\} \geq 1 - \frac{D\xi}{\varepsilon^2}.$$

В нашем случае $\varepsilon = 3\sigma$, где $\sigma = \sqrt{D\xi}$ – среднее квадратичное отклонение. Поэтому $P\{|\xi - M\xi| < 3\sigma\} \geq 1 - \frac{\sigma^2}{9\sigma^2}$, т.е.

$$P\{|\xi - M\xi| < 3\sigma\} \geq \frac{8}{9}.$$

Таким образом, полученная вероятность не меньше, чем $\frac{8}{9}$.

Ковариация и ее свойства.

Определение. Ковариацией СВ ξ и η называется

$$\operatorname{cov}(\xi,\eta) = M\left(\overset{0}{\xi} \cdot \overset{0}{\eta}\right) = M\left[(\xi - M\xi)(\eta - M\eta)\right],$$

здесь $\overset{0}{\xi} = \xi - M\xi$, $\overset{0}{\eta} = \eta - M\eta$ – центрированные СВ по отношению к СВ ξ и η.

Рассмотрим свойства ковариации.

1. $\operatorname{cov}(\xi,\xi) = M\,(\xi - M\xi)^2 = D\xi$.

2. Если СВ ξ и η независимы, то $\operatorname{cov}(\xi,\eta) = 0$, это следует из определения ковариации и свойства 6 математического ожидания.

3. Для произвольных СВ ξ и η

$$D(\xi + \eta) = D\xi + D\eta + 2\operatorname{cov}(\xi,\eta),$$

что следует из определения дисперсии и ковариации.

4. Пусть $\xi_1 = x\xi - \eta$, где x принимает действительные значения. Рассмотрим дисперсию СВ ξ_1

$$D\xi_1 = M\left[x\xi - \eta - M(x\xi - \eta)\right]^2 = M\left[(x\xi - M(x\xi)) - (\eta - M\eta)\right]^2 =$$

$$= x^2 M(\xi - M\xi)^2 - 2xM\left[(\xi - M\xi)(\eta - M\eta)\right] + M(\eta - M\eta)^2 =$$

$$= x^2 D\xi - 2x\operatorname{cov}(\xi,\eta) + D\eta \cdot$$

Это выражение является квадратным трехчленом относительно x. Поскольку $D\xi_1 \geq 0$, то дискриминант $\left[2\operatorname{cov}(\xi,\eta)\right]^2 - 4D\xi\, D\eta$ должен быть меньше либо равен нулю, т.е. должно выполняться неравенство

$$\left[\operatorname{cov}(\xi,\eta)\right]^2 \leq D\xi D\eta,$$

т.е.

$$\left|\operatorname{cov}(\xi,\eta)\right| \leq \sqrt{D\xi\, D\eta},$$

или

$$-\sqrt{D\xi\, D\eta} \leq \operatorname{cov}(\xi,\eta) \leq \sqrt{D\xi\, D\eta}\cdot$$

Коэффициент корреляции и его свойства. Нормированной СВ по отношению к СВ ξ называется СВ $\xi^0 = \dfrac{\xi - M\xi}{\sqrt{D\xi}}$; очевидно, что $M\xi^0 = 0,\ D\xi^0 = 1$.

Определение. Коэффициентом корреляции СВ ξ и η называется

$$r(\xi,\eta) = M\left(\xi^0\eta^0\right) = M\left[\frac{(\xi - M\xi)(\eta - M\eta)}{\sqrt{D\xi}\,\sqrt{D\eta}}\right].$$

Он характеризует меру степени зависимости между СВ ξ и η. СВ ξ и η называются некоррелированными, если $r(\xi,\eta) = 0$ $\left(\operatorname{cov}(\xi,\eta) = 0\right)$.

Пример 1.50. Приведем два примера, показывающие, что из равенства нулю коэффициента корреляции двух случайных величин не следует их независимость.

1. Пусть вероятностное пространство $(\Omega,\ \mathrm{F}, P)$ представляет собой отрезок $[0,1]$ с σ-алгеброй F борелевских подмножеств и мерой Лебега P.

Рассмотрим СВ ξ и η, определенные следующим образом:

$$\xi=\xi(\omega)=\begin{cases}1, & 0\le\omega\le\dfrac{1}{4},\\ -1, & \dfrac{1}{4}<\omega<\dfrac{1}{2},\\ 0, & \dfrac{1}{2}<\omega\le 1,\end{cases}\quad \eta=\eta(\omega)=\begin{cases}0, & 0\le\omega\le\dfrac{1}{2},\\ 1, & \dfrac{1}{2}<\omega\le\dfrac{3}{4},\\ -1, & \dfrac{3}{4}<\omega\le 1.\end{cases}$$

Найдем их математические ожидания:

$$M\xi=1\cdot P\left\{\omega: 0\le\omega\le\frac{1}{4}\right\}-1\cdot P\left\{\omega:\frac{1}{4}<\omega<\frac{1}{2}\right\}=\frac{1}{4}-\frac{1}{4}=0,$$

$$M\eta=1\cdot P\left\{\omega:\frac{1}{2}<\omega\le\frac{3}{4}\right\}-1\cdot P\left\{\omega:\frac{3}{4}<\omega\le 1\right\}=\frac{1}{4}-\frac{1}{4}=0.$$

Очевидно также, что $\xi\eta=0$ при всех $\omega\in[0,1]$. Поэтому $\operatorname{cov}(\xi,\eta)=M(\xi\eta)-M\xi M\eta=0$, т.е. СВ ξ и η некоррелированы. Но эти СВ зависимы, т.к. $P\{\omega:\xi(\omega)=1,\ \eta(\omega)=1\}=0,$ а

$$P\{\omega:\xi(\omega)=1\}P\{\omega:\eta(\omega)=1\}=\frac{1}{4}\cdot\frac{1}{4}=\frac{1}{16}, \text{ т.е.}$$

$$P\{\omega:\xi(\omega)=1,\ \eta(\omega)=1\}\ne P\{\omega:\xi(\omega)=1\}P\{\omega:\eta(\omega)=1\}\cdot$$

2. Рассмотрим теперь две непрерывные СВ ξ и η. Пусть СВ ξ имеет плотность распределения $p_\xi(x)=\mathrm{A}e^{-x^2},\ -\infty<x<+\infty$, а $\eta=\xi^2$. Тогда

$$\operatorname{cov}(\xi,\eta)=M(\xi\eta)-M\xi M\eta=M\xi^3-M\xi M\xi^2=$$

$$= Á\int_{-\infty}^{\infty} x^3 e^{-x^2}\,dx - Á\int_{-\infty}^{\infty} x e^{-x^2}\,dx\, Ì\,\xi^2 = 0 - 0Ì\,î^{\,2} = 0,$$

т.к. подынтегральные функции в интегралах – нечетные, и, следовательно, $\operatorname{cov}(î,ç) = r(î,ç) = 0$, т.е. СВ $î$ и $ç$ некоррелированы, в то время как они связаны функциональной зависимостью $ç = î^2$.

Рассмотрим свойства коэффициента корреляции.

1. $|r(î,ç)| \le 1$. Оно выполняется поскольку

$$0 \le D\left(î^0 \pm ç^0\right) = Ì\left(î^0 \pm ç^0\right)^2 = Ì\left(î^0\right)^2 + Ì\left(ç^0\right)^2 \pm 2Ì\left(î^0 ç^0\right) =$$
$$= 2\left[1 \pm Ì\left(î^0 ç^0\right)\right] \Rightarrow 1 \pm Ì\left(î^0 ç^0\right) \ge 0 \Rightarrow |r(î,ç)| \le 1.$$

2. $r(î,ç) = \pm 1$ тогда и только тогда, когда $î$ и $ç$ связаны линейной зависимостью.

Докажем вначале необходимость. Пусть $r(î,ç) = \pm 1$. Рассмотрим выражение

$$0 \le Ì\left(\frac{î - Ìî}{\sqrt{Dî}} \pm \frac{ç - Ìç}{\sqrt{Dç}}\right) = 1 \pm 2r(î,ç) + 1 = 2\left[1 \pm r(î,ç)\right].$$

Отсюда следует, что если $r(î,ç) = \pm 1$, то $\dfrac{î - Ìî}{\sqrt{Dî}} \pm \dfrac{ç - Ìç}{\sqrt{Dç}} = 0$, т.е.

$$î = \mp\sqrt{\frac{Dî}{Dç}}\,ç \pm Ì\sqrt{\frac{Dî}{Dç}} + Ìî,$$

значит, $î$ и $ç$ линейно зависимы.

Достаточность следует из того, что, если $ç = aî + b,\ a \ne 0$, то

$$r(î,ç) = \frac{Ì\left[(î - Ìî)(ç - Ìç)\right]}{\sqrt{Dî}\sqrt{Dç}} = \frac{aÌ\left(î - Ìî\right)^2}{|a|\sqrt{Dî}} = \frac{a}{|a|} = signa = \begin{cases} 1, & a > 0, \\ -1, & a < 0. \end{cases}$$

3. Если $î$ и $ç$ – независимые СВ, то $r(î,ç)=0$. Обратное утверждение не обязательно, но оно будет справедливо, когда $î$ и $ç$ – нормально распределенные СВ.

Важной характеристикой n-мерной СВ является **ковариационная матрица** $Ê=\left\|Ê_{ij}\right\|_{n\times n}$, где $Ê_{ij}=Ì\left[(î_i-Ìî_i)(î_j-Ìî_j)\right]$ – ковариационный момент СВ $î_i$ и $î_j$; ясно, что $Ê_{ij}=Ê_{ji},\ Ê_{ii}=Dî_i,\ i,j=\overline{1,n}$. **Нормированной ковариационной (корреляционной) матрицей** R называется матрица, элементами которой являются коэффициенты корреляции СВ $î_i$ и $î_j$:

$$R=\left\|r_{ij}\right\|_{n\times n},\ r_{ij}=\frac{Ê_{ij}}{\sqrt{Dî_i}\sqrt{Dî_j}},\ r_{ij}=r_{ji},\ r_{ii}=1,\ i,j=\overline{1,n}\cdot$$

Энтропия. Количество информации. Часто числовыми характеристиками СВ являются функционалы, определяющие различие между двумя распределениями вероятностей. Такое различие необходимо знать, например, если, кроме факта зависимости между двумя СВ, нужно знать, насколько велика эта зависимость. Рассмотрим случай абсолютно непрерывных распределений. В качестве величины, измеряющей степень зависимости двух СВ, можно использовать расстояние между распределениями $p_{îç}(t,ô)$ и $p_î(t)p_ç(ô)$. Расстояние между двумя распределениями измеряется различными способами. Одним из них является так называемое количество информации Шеннона.

Определение. Энтропией называется функционал

$$ç(î)=Ì\left[\ln p_î(î)\right]=\int_{-\infty}^{\infty}\ln\left(p_î(x)\right)p_î(x)dx\cdot$$

В физике это мера беспорядка (неопределенности); чем больше неопределенность, тем больше энтропия.

Следует отметить, что ξ может быть и многомерной СВ, т.е. может быть, что $H(\xi) = H(\xi_1, \xi_2, \ldots, \xi_n)$.

Определение. Количеством информации Шеннона называется величина

$$I(\xi,\eta) = -H(\xi) - H(\eta) + H(\xi,\eta).$$

В теории информации эта величина означает, что СВ ξ содержит $I(\xi,\eta)$ информации о СВ η.

Т.к.

$$p_{\xi}(x) = \int_{-\infty}^{\infty} p_{\xi\eta}(x,y)dy,\ H(\xi) = \int_{-\infty}^{\infty} \ln\big(p_{\xi}(x)\big)p_{\xi}(x)dx =$$

$$= \int_{-\infty}^{\infty} \ln p_{\xi}(x) \int_{-\infty}^{\infty} p_{\xi\eta}(x,y)dydx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \ln\big(p_{\xi}(x)\big)p_{\xi\eta}(x,y)dy\,dx,$$

$$H(\eta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \ln\big(p_{\eta}(y)\big)p_{\xi\eta}(x,y)dydx, \text{ то}$$

$$I(\xi,\eta) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} \ln\left[\frac{p_{\xi\eta}(x,y)}{p_{\xi}(x)p_{\eta}(y)}\right]dy.$$

Если ξ и η – независимые СВ, то $I(\xi,\eta) = 0$, т.е. никакой информации СВ ξ не содержит о СВ η. В этом случае $H(\xi,\eta) = H(\xi) + H(\eta)$.

§7. ФУНКЦИОНАЛЬНЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

Определение. Характеристической функцией СВ $\xi(\omega)$ называется комплекснозначная функция

$$\varphi_{\xi}(t)=Me^{it\xi(\omega)}=\int\limits_{\Omega}e^{it(\omega)}P(d\omega)=\int\limits_{X}e^{itx}dF_{\xi}(x),$$

где $t-$ действительная переменная.

Таким образом, характеристическая функция является преобразованием Фурье–Стилтьеса ф.р. $F_{\xi}(x)$. Если $F_{\xi}(x)$ – абсолютно непрерывная ф.р., то $\varphi_{\xi}(t)$ является обычным преобразованием Фурье плотности распределения $p_{\xi}(x)$:

$$\varphi_{\xi}(t)=\int\limits_{-\infty}^{\infty}e^{itx}p_{\xi}(x)dx;$$

если ξ – СВ, принимающая значения $x_0, x_1, ..., x_k, ...$ соответственно с вероятностями $p_0, p_1, ..., p_k, ...$, то

$$\varphi_{\xi}(t)=\sum_{k=0}^{\infty}e^{itx_k}p_k.$$

Характеристическая функция существует всегда, т.к.

$$\left|\varphi_{\xi}(t)\right|=\left|\int\limits_{-\infty}^{+\infty}e^{itx}dF_{\xi}(x)\right|\le\int\limits_{-\infty}^{+\infty}\left|e^{itx}\right|dF_{\xi}(x)=\int\limits_{-\infty}^{+\infty}dF_{\xi}(x)=1,$$

очевидно, что ее можно записать в виде

$$\varphi_{\xi}(t)=Me^{it\xi}=M\cos(t\xi)+iM\sin(t\xi).$$

Рассмотрим свойства характеристических функций.

1. $\varphi_{\xi}(0)=1,\ \left|\varphi_{\xi}(t)\right|\le 1$.

2. $\varphi_{a\xi+b}(t)=Me^{it(a\xi+b)}=e^{itb}\varphi_{\xi}(at)$, где a и b – константы.

3. Пусть $\xi_1, \xi_2, ..., \xi_n$ – независимые CB. Тогда

$$\varphi_{\xi_1+\xi_2+...+\xi_n}(t)=Me^{it(\xi_1+\xi_2+...+\xi_n)}=Me^{it\xi_1}Me^{it\xi_2}\cdot...\cdot Me^{it\xi_n}=$$

$$=\varphi_{\xi_1}(t)\varphi_{\xi_2}(t)...\varphi_{\xi_n}(t).$$

4. $\varphi_{\xi}(t)$ – равномерно непрерывная функция, т.е. $\forall \varepsilon > 0\ \exists \delta > 0$, такое, что $\left|\varphi_{\xi}(t+\delta) - \varphi_{\xi}(t)\right| < \varepsilon$ для всех t.

5. Если $M\ |\xi|^{k} < +\infty,\ k \geq 1$, то существует непрерывная k-я производная характеристической функции $\varphi_{\xi}(t)$ и

$$\varphi_{\xi}^{(k)}(0) = i^{k} M\xi^{k}.$$

Поясним это свойство. Т.к.

$$\left|\int_{-\infty}^{+\infty} ix\, e^{itx}\, dF_{\xi}(x)\right| \leq \int_{-\infty}^{+\infty} |x|\, dF_{\xi}(x) = M\ |\xi| < +\infty,$$

то интеграл $\int_{-\infty}^{+\infty} ix\, e^{itx}\, dF_{\xi}(x)$ сходится равномерно относительно t, поэтому возможно дифференцирование под знаком интеграла и

$$\varphi_{\xi}'(t) = i\int_{-\infty}^{+\infty} x e^{itx}\, dF_{\xi}(x),\ \text{т.е.}\ \varphi_{\xi}'(0) = iM\,\xi;$$

далее доказательство можно провести по индукции.

Из этого свойства вытекает, в частности, что, если $M\ |\xi|^{k} < +\infty$, в окрестности точки $t = 0$ справедливо разложение функции $\varphi_{\xi}(t)$ в ряд:

$$\varphi_{\xi}(t) = 1 + \sum_{j=1}^{k} \frac{i^{j} M\xi^{j}}{j!} t^{j} + 0\left(t^{k}\right).$$

6. $\overline{\varphi_{\xi}(t)} = \varphi_{\xi}(-t) = \varphi_{-\xi}(t)$, где $\overline{\varphi_{\xi}(t)}$ – функция, комплексно сопряженная функции $\varphi_{\xi}(t)$. Отсюда следует, что, если $F_{\xi}(x) = 1 - F_{\xi}(-x)$ (или, что для непрерывной СВ эквивалентно тому, что плотность распределения $p_{\xi}(x)$ симметрична относительно нуля), характеристическая функция является действительной.

Пример 1.51. Найдем характеристическую функцию СВ, распределенной по нормальному закону. Рассмотрим вначале случай стандартного нормального распределения, когда $a = 0,\ \sigma^2 = 1$. В этом случае

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},\ \varphi_\xi(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{itx - \frac{x^2}{2}} dx \cdot$$

Продифференцировав функцию $\varphi_{\hat{\imath}}(t)$ по t (см. свойство 5), получим:

$$\varphi'_{\hat{\imath}}(t) = \frac{1}{\sqrt{2\eth}} \int\limits_{-\infty}^{+\infty} ix\, e^{itx - \frac{x^2}{2}} dx = -\frac{i}{\sqrt{2\eth}} \int\limits_{-\infty}^{+\infty} e^{itx} d\left(e^{-\frac{x^2}{2}} \right) =$$

$$= -\frac{i}{\sqrt{2\eth}} e^{itx} e^{-\frac{x^2}{2}} \Bigg|_{-\infty}^{+\infty} - \frac{t}{\sqrt{2\eth}} \int\limits_{-\infty}^{+\infty} e^{itx - \frac{x^2}{2}} dx = -t\varphi_{\hat{\imath}}(t),$$

откуда следует, что $d(\ln \varphi_{\hat{\imath}}(t)) = -t,\ \ln \varphi_{\hat{\imath}}(t) = -c_1 \dfrac{t^2}{2}$, т.е. $\varphi_{\hat{\imath}}(t) = ce^{-\frac{t^2}{2}}$. Поскольку $\varphi_{\hat{\imath}}(0) = 1$, то $c = 1$ и $\varphi_{\hat{\imath}}(t) = e^{-\frac{t^2}{2}}$.

Если СВ $\hat{\imath}$ имеет нормальное распределение с произвольными параметрами $a,\ \sigma^2$, то СВ $\eta = \dfrac{\xi - a}{\sigma}$ является нормированной по отношению к ней, $\grave{I}ç = 0,\ Dç = 1$, и $\xi = \sigma\,\eta + a$. Тогда, поскольку $\varphi_{\sigma\eta} = e^{-\frac{(\sigma t)^2}{2}}$, то

$$\varphi_\xi(t) = Me^{it\xi} = e^{ita} Me^{it\sigma\eta} = e^{ita} e^{-\frac{\sigma^2 t^2}{2}} = e^{ita - \frac{\sigma^2 t^2}{2}}.$$

Пример 1.52. Найдем характеристическую функцию дискретной СВ, распределенной по закону Пуассона. Используя свойство 8 для математического ожидания (см. п.1.6), имеем

$$\varphi_{\hat{\imath}}(t)=\sum_{k=0}^{\infty} e^{itk}\,\frac{\ddot{e}^{k}}{k\,!}\,e^{-\ddot{e}}\,,$$

поэтому

$$\varphi'_{\hat{\imath}}(t)=i\sum_{k=0}^{\infty} ke^{itk}\,\frac{\ddot{e}^{k}}{k\,!}\,e^{-\ddot{e}}=i\ddot{e}e^{it}\sum_{k=1}^{\infty} e^{it(k-1)}\,\frac{\ddot{e}^{k-1}}{(k-1)!}\,e^{-\ddot{e}}=i\ddot{e}e^{it}\varphi_{\hat{\imath}}(t),$$

откуда следует, что

$$\varphi_{\hat{\imath}}(t)=c\ \exp\left\{i\ddot{e}\int_{0}^{t} e^{ix}dx\right\}=c\ \exp\left[\ddot{e}\left(e^{it}-1\right)\right].$$

Т.к. $\varphi_{\hat{\imath}}(0)=1$, то $c=1$, таким образом,

$$\varphi_{\hat{\imath}}(t)=\exp\left[\ddot{e}\left(e^{it}-1\right)\right].$$

Пример 1.53. Выяснить, является ли функция $f(t)=e^{-t^4}$ характеристической функцией.

Решение. Найдем первую и вторую производные функции $f(t)$

$$\varphi'_{\hat{\imath}}(t)=-4t^{3}e^{-t^{4}},\ \varphi''_{\hat{\imath}}(t)=4t^{2}e^{-t^{4}}\left(4t^{4}-3\right).$$

Поскольку из свойства 5 дисперсии следует, что $D\hat{\imath}=\grave{I}\hat{\imath}^{\,2}-\left(\grave{I}\hat{\imath}\right)^{2}=-\varphi''_{\hat{\imath}}(0)+\left[\varphi'_{\hat{\imath}}(0)\right]^{2}$, то дисперсия нашей СВ должна была бы равняться нулю. Но тогда $\tilde{N}\left(\hat{\imath}=const\right)=1$, т.е. $f(t)$ – характеристическая функция вырожденного распределения, и, согласно определению, по модулю должна равняться единице, что противоречит условию задачи.

Характеристическая функция многомерной СВ определяется следующим образом

$$\varphi_{\xi_1\xi_2\ldots\xi_n}(t_1, t_2, \ldots, t_n) = Me^{i\sum\limits_{k=1}^{n} t_k\xi_k} = \int\limits_{\mathbb{R}^n} e^{i\sum\limits_{k=1}^{n} t_k x_k} dF_{\xi_1\xi_2\ldots\xi_n}(x_1, x_2, \ldots, x_n).$$

Приведем для нее свойство, аналогичное свойству 5 характеристических функций одномерных СВ:

$$\left.\frac{\partial\, \varphi^m_{\xi_1\xi_2\ldots\xi_n}(t_1, t_2, \ldots, t_n)}{\partial\, t_1^{k_1}\, \partial\, t_2^{k_2} \ldots \partial\, t_n^{k_n}}\right|_{t_1 = t_2 = \ldots = t_n = 0} = i^m M\left(\xi_1^{k_1}\xi_2^{k_2}\ldots\xi_n^{k_n}\right),$$

где $m = k_1 + k_2 + \ldots + k_n$.

Рассмотрим одну важную теорему.

Теорема 1.10 (об обращении характеристической функции). *Если $F_\xi(x)$ – ф.р. СВ ξ, $\varphi_\xi(t)$ – ее характеристическая функция, то для любых точек непрерывности x и y функции $F_\xi(x)$:*

$$F_\xi(y) - F_\xi(x) = \lim_{T \to \infty} \frac{1}{2\pi} \int\limits_{-T}^{T} \left(e^{-itx} - e^{-ity}\right)\varphi_\xi(t)\frac{1}{it}\, dt.$$

Если $F_\xi(x)$ – абсолютно непрерывная функция, то это соотношение легко получить из обычной формулы обратного преобразования Фурье

$$p_\xi(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} e^{-itx}\varphi_\xi(t)dt,$$

поскольку

$$F_\xi(y) - F_\xi(x) = \int\limits_x^y p_\xi(z)dz = \frac{1}{2\pi}\int\limits_x^y\int\limits_{-\infty}^{+\infty} e^{-itx}\varphi_\xi(t)dtdz =$$

$$= \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}\varphi_\xi(t)\int\limits_x^y e^{-itz}dzdt = \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}\varphi_\xi(t)\frac{1}{it}\left(e^{-itx} - e^{-ity}\right)dt.$$

Из этой теоремы вытекает **теорема единственности**: *характеристическая функция СВ* $\hat{\imath}$ *однозначно определяет ее ф.р.*

Когда СВ $\hat{\imath}$ является целочисленной, т.е. множество ее значений – числа натурального ряда, то характеристическая функция является функцией $z = e^{it}$ и, согласно свойству 8 для математического ожидания (см. п.1.6), имеем вид

$$\varphi_{\hat{\imath}}(t) = \o_{\hat{\imath}}(z) = \grave{I} z^{\hat{\imath}} = \sum_{k=0}^{\infty} z^k p_k \,,$$

где $p_k = \tilde{N}(\hat{\imath} = k)$, $k = 0, 1, 2, \dots$.

Определение. Функция

$$\o_{\hat{\imath}}(z) = p_0 + p_1 z + p_2 z^2 + \dots + p_k z^k + \dots, \ |z| \le 1 \,,$$

называется **производящей функцией СВ** $\hat{\imath}$.

Пример 1.54. Пусть СВ $\hat{\imath}$ имеет геометрическое распределение, в этом случае $p_k = p(1-p)^k$, $k = 0, 1, 2, \dots,\ 0 < p < 1$, поэтому

$$\o_{\hat{\imath}}(z) = \sum_{k=0}^{\infty} z^k p(1-p)^k = p \sum_{k=0}^{\infty} [(1-p)z]^k = \frac{p}{1-(1-p)z} \,.$$

Она имеет следующие свойства.

1. Функция $\o_{\hat{\imath}}(z) = \varphi_{\hat{\imath}}(t)$, рассматриваемая как функция аргумента t, является периодичной с периодом $2\eth$, поэтому достаточно знать поведение $\o_{\hat{\imath}}(z)$ на окружности $|z| = 1$, чтобы полностью описать все свойства $\varphi_{\hat{\imath}}(t)$.

2. Из определения следует, что

$$\o_{\hat{\imath}}(0) = p_0, \ \o_{\hat{\imath}}(1) = p_0 + p_1 + \dots + p_k + \dots = 1 \,.$$

3. По значениям $\o_{\hat{\imath}}(z)$ однозначно определяются вероятности $p_0, p_1, \dots, p_k, \dots$. Например,

$$p_1 = \left.\frac{d\text{ø}_{\text{î}}(z)}{dz}\right|_{z=0} = \left[p_1 + 2p_2 z + \ldots + kp_k z^{k-1} + \ldots\right]_{z=0},$$

$$p_2 = \frac{1}{2}\left[\frac{d^2\text{ø}_{\text{î}}(z)}{dz^2}\right]_{z=0} = \frac{1}{2}\left[p_2 + 6p_3 z + \ldots + k(k-1)p_k z^{k-2} + \ldots\right]_{z=0},$$

общее выражение имеет вид

$$p_k = \frac{1}{k!}\left[\frac{d^k}{dz^k}\text{ø}_{\text{î}}(z)\right]_{z=0},\ k = 0, 1, 2, \ldots$$

4. Если $\text{Ì}\ |\text{î}| < +\infty$, то

$$\left.\frac{d\text{ø}_{\text{î}}(z)}{dz}\right|_{z=1} = \left[\sum_{k=0}^{\infty} kz^{k-1}p_k\right]_{z=1} = \sum_{k=0}^{\infty} kp_k = \text{Ìî},$$

$$\left[\frac{d^2\text{ø}_{\text{î}}(z)}{dz^2}\right]_{z=1} = \left[\sum_{k=1}^{\infty} k(k-1)z^{k-2}p_k\right]_{z=1} = \sum_{k=1}^{\infty} k(k-1)p_k = \text{Ì}\ [\text{î}(\text{î}-1)],$$

в общем случае

$$\left[\frac{d^k}{dz^k}\text{ø}_{\text{î}}(z)\right]_{z=1} = \text{Ì}\ [\text{î}(\text{î}-1)\ldots(\text{î}-k+1)], \text{ если } \text{Ì}\ |\text{î}|^k < +\infty.$$

5. Пусть $\text{î}_1, \text{î}_2, \ldots, \text{î}_n$ – целочисленные независимые одинаково распределенные СВ, ç – независимая от них целочисленная СВ, $\text{Ì}\ |\text{î}| < +\infty$, $\text{Ì}\ |\eta| < +\infty$, $\text{ô} = \text{î}_1 + \text{î}_2 + \ldots + \text{î}_n$ (т.е. ô является суммой случайного числа СВ). Тогда

$$\text{ø}_{\text{ô}}(z) = \text{ø}_{\text{ç}}\left[\text{ø}_{\text{î}_1}(z)\right],$$

поскольку

$$\text{ø}_{\text{ô}}(z) = \text{Ì}z^{\text{ô}} = \sum_{k=0}^{\infty} \text{Ì}\ \left(z^{\text{ô}}/\text{ç} = k\right) =$$

$$=\sum_{k=0}^{\infty} \text{Ì}z^{\text{î}_1+\text{î}_2+\ldots+\text{î}_k}\,\text{Ñ}(\text{ç}=k)\text{Ì}z^{\text{î}_1+\text{î}_2+\ldots+\text{î}_k}\,\text{Ñ}(\text{ç}=k)=\sum_{k=0}^{\infty}\text{Ì}z^{\text{î}_1}\,\text{Ì}z^{\text{î}_2}\cdot\ldots\cdot\text{Ì}z^{\text{î}_k}\,\text{Ñ}(\text{ç}=k)=$$

$$=\sum_{k=0}^{\infty}\left[\text{ø}_{\text{î}_1}(z)\right]^k\text{Ñ}(\text{ç}=k)=\text{ø}_{\text{ç}}\left[\text{ø}_{\text{î}_1}(z)\right]\cdot$$

6. Пусть î и ç – независимые СВ. Тогда

$$\text{ø}_{\text{î}+\text{ç}}(z)=\text{Ì}z^{\text{î}+\text{ç}}=\text{Ì}z^{\text{î}}\,\text{Ì}z^{\text{ç}}=\text{ø}_{\text{î}}(z)\text{ø}_{\text{ç}}(z)\cdot$$

Пример 1.55. Найти распределение вероятностей, которому соответствует производящая функция

$$\text{ø}_{\text{î}}(z)=1+\frac{1-z}{z}\ln(1-z)\cdot$$

Решение. Разложим производящую функцию в ряд по степеням z:

$$\text{ø}_{\text{î}}(z)=1+\frac{1}{z}\ln(1-z)-\ln(1-z)=$$

$$=1-\frac{1}{z}\sum_{k=1}^{\infty}\frac{z^k}{k}+\sum_{k=1}^{\infty}\frac{z^k}{k}=1-\sum_{k=0}^{\infty}\frac{z^k}{k+1}+\sum_{k=1}^{\infty}\frac{z^k}{k}=\sum_{k=1}^{\infty}\left(\frac{1}{k}-\frac{1}{k+1}\right)z^k=\sum_{k=1}^{\infty}\frac{z^k}{k(k+1)}\cdot$$

Таким образом, функция $\text{ø}_{\text{î}}(z)$ является производящей функцией СВ, которая может принимать целые положительные значения $1,2,\ldots,k,\ldots$ и

$$p_k=\text{Ñ}(\text{î}=k)=\frac{1}{k(k+1)},\ k=1,2,\ldots\cdot$$

§8. СХОДИМОСТЬ СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Со сходимостью случайных последовательностей в теории вероятностей связан ряд классических результатов, рассматриваемых в данном параграфе.

Определение. Последовательность СВ $\{\xi_k(\omega),\ k = 1, 2, \ldots\}$ сходится к СВ $\xi(\omega)$ **по вероятности** (обозначается $\xi_k(\omega) \xrightarrow[k\to\infty]{P} \xi(\omega)$), если

$$\forall \varepsilon > 0 \ \lim_{k\to\infty} P\{\omega : |\xi_k(\omega) - \xi(\omega)| \geq \varepsilon\} = 0 \cdot$$

Существует еще несколько типов сходимости случайных последовательностей:

– **сходимость почти наверное** (п.н.) (с вероятностью 1):

$$\xi_k(\omega) \xrightarrow[k\to\infty]{п.\ н.} \xi(\omega) \left(\xi_k(\omega) \xrightarrow[k\to\infty]{P=1} \xi(\omega) \right), \text{ если}$$

$$P\{\omega : \lim_{k\to\infty} \xi_k(\omega) = \xi(\omega)\} = 1;$$

можно показать, что это соотношение эквивалентно следующему

$$P\left\{ \bigcap_{k\geq 1} \bigcup_{n\geq k} \{\omega : |\xi_n(k) - \xi(\omega)| \geq \varepsilon\} \right\} = 0;$$

– **сходимость в среднем порядка** r (предполагается, что $M\,|\xi_k(\omega)|^r < +\infty$):

$$\xi_k(\omega) \xrightarrow[k\to\infty]{ср.\ п.\ r} \xi(\omega), \text{ если } \lim_{k\to\infty} M\left\{|\xi_k(\omega) - \xi(\omega)|^r\right\} = 0;$$

если $r = 1$, то имеем **сходимость в среднем**; если $r = 2$ – то **сходимость в среднем квадратичном**, которая записывается в виде $\underset{k\to\infty}{l.i.m.}\, \xi_k(\omega) = \xi(k)$ (limit in the mean);

– **сходимость по распределению** (слабая сходимость):

$\xi_k(\omega) \xrightarrow[k\to\infty]{D} \xi(\omega)$, если $\lim_{k\to\infty} F_{\xi_k}(x) = F_{\xi}(x)$ во всех точках непрерывности ф.р. $F_{\xi}(x)$.

Пример 1.56. Пусть $\xi(\omega)$ – СВ, $\xi_k(\omega)=\min\{\xi(\omega), k\}$. Докажем, что последовательность СВ $\{\xi_k(\omega)\}$, $k=1, 2, \dots$, сходится к СВ $\xi(\omega)$ почти наверное.

Имеем:

$$P\{\omega : \xi_k(\omega)=\xi(\omega), k\ge N\}=P\{\omega : \min\{\xi(\omega), k\}=\xi(\omega), k\ge N\}=$$

$$=P\{\omega : \xi(\omega)\le N\}.$$

Вместе с тем при любом конкретном ω, если

$\xi_k(\omega)=\xi(\omega)$, $k\ge N$, то $\xi_k(\omega)\xrightarrow[k\to\infty]{}\xi(\omega)$, т.е.

$$\{\omega : \xi_k(\omega)=\xi(\omega),\ k\ge N\}\subseteq\{\omega : \xi_k(\omega)\xrightarrow[k\to\infty]{}\xi(\omega)\}.$$

Поэтому

$$P\{\omega : \xi_k(\omega)\xrightarrow[k\to\infty]{}\xi(\omega)\}\ge P\{\omega : \xi_k(\omega)=\xi(\omega),\ k\ge N\}=$$

$$=P\{\omega : \xi(\omega)\le N\}=F_\xi(N).$$

Далее, т.к. $\lim\limits_{N\to\infty}F_\xi(N)=1$ и вероятность случайного события не может быть больше единицы, то, устремив N к ∞, получаем

$$P\{\omega : \xi_k(\omega)\xrightarrow[k\to\infty]{}\xi(\omega)\}=1.$$

Пример 1.57. Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, где $\Omega=[0,1]$, $\mathcal{F}$ – σ-алгебра борелевских множеств на Ω, P – мера Лебега, задана СВ

$$\xi_k(\omega)=\begin{cases} k^{\frac{1}{\alpha}}, \ \omega\in\left(0,\dfrac{1}{k}\right],\\ 0, \ \omega\notin\left(0,\dfrac{1}{k}\right].\end{cases}$$

Поскольку $Ì\,\left|î_k(ù)\right|^{á}=\left(k^{\frac{1}{\alpha}}\right)^{á}P\left\{ù:ù\in\left(0,\frac{1}{k}\right]\right]\right\}=k\cdot\frac{1}{k}=1$, то последовательность $\{î_k(ù),\ k=1,2,...\}$ не сходится к нулю в среднем порядка α.

Для любых видов сходимости в математике существуют **критерии типа Коши**. Приведем такие критерии для некоторых типов сходимости случайных последовательностей. Например, для того чтобы $î_k(ù)\xrightarrow[k\to\infty]{ср.п.\ r}î(ù)$, необходимо и достаточно, чтобы последовательность $\{î_k(ù),\ k=1,2,...\}$ была фундаментальной в среднем порядка r, т.е.

$$Ì\,\left\{\left|î_k(ù)-î_n(ù)\right|^r\right\}\xrightarrow[k,n\to\infty]{}0\cdot$$

Аналогично:

$$î_k(ù)\xrightarrow[k\to\infty]{P}î(ù)\Leftrightarrow\forall å>0\quad\lim_{k,n\to\infty}P\left\{ù:\left|î_k(ù)-î_n(ù)\right|>å\right\}=0,$$

$$î_k(ù)\xrightarrow[k\leftarrow\infty]{п.н.}î(ù)\Leftrightarrow\forall å>0\ \ P\left[\bigcap_{k\ge1}\left\{ù:\sup_{n\ge0}\left|î_{k+n}(ù)-î_k(ù)\right|>å\right]=0\cdot\right.$$

Теорема 1.11 ***(достаточный признак сходимости почти наверное).*** *Если* $\forall å>0$ *с вероятностью 1 происходит лишь конечное число событий*

$$Á_k(å)=\left\{ù:\left|î_k(ù)-î(ù)\right|\ge å\right\},\ k=1,2,...,$$

то $î_k(ù)\xrightarrow[л\to\infty]{п.н.}î(ù)\cdot$

Приведем краткое доказательство теоремы. Пусть $\text{å} = \dfrac{1}{n}$ и $\hat{A}_n$ – событие, заключающееся в том, что число событий $\acute{A}_k\left(\dfrac{1}{n}\right)$ является конечным. Очевидно, что $\hat{A}_1 \supseteq \hat{A}_2 \supseteq \ldots \supseteq \hat{A}_n \supseteq \ldots$, и по аксиоме непрерывности для $\hat{A} = \bigcap\limits_{n=1}^{\infty} \hat{A}_n$ имеем $\tilde{N}(\hat{A}) = \lim\limits_{n\to\infty} \tilde{N}(\hat{A}_n)$. Но $P(\hat{A}_n) = 1 \ \forall n > 0$, поэтому $P(\hat{A}) = 1$ и

P {число событий $\acute{A}_k(0) = \left\{\grave{u} : \left|\hat{\imath}_k(\grave{u}) - \hat{\imath}(\grave{u})\right| > 0\right\}$ – конечно$\} = 1$,

следовательно, $P\left\{\omega : \lim\limits_{k\to\infty} \xi_k(\omega) = \xi(\omega)\right\} = 1$.

При исследовании сходимости случайных последовательностей часто используется следующая лемма.

Лемма Бореля–Кантелли ***(закон нуля и единицы).*** *Пусть* $\left\{\acute{A}_k,\ k = 1, 2, \ldots\right\}$ *– последовательность случайных событий, заданных на вероятностном пространстве* (Ω, F, P), $\acute{A}$ *– событие, заключающееся в том, что произойдет бесчисленное множество событий* $\acute{A}_k$, *т.е.* $A = \bigcap\limits_{n=1}^{\infty} \bigcup\limits_{k=n}^{\infty} A_k$. *Тогда:*

а) если $\sum\limits_{k=1}^{\infty} P(\acute{A}_k) < +\infty$, *то* $P(\acute{A}) = 0$,

б) если $\sum\limits_{k=1}^{\infty} P(\acute{A}_k) = +\infty$ *и события* $\acute{A}_k$, $k = 1, 2, \ldots$, *независимы в совокупности, то* $P(\acute{A}) = 1$.

Часть а) данной леммы, например, вытекает из того, что

$$Á \subseteq \bigcup_{k=n}^{\infty} Á_k \forall n \geq 1, \; P(Á) \leq P\left(\bigcup_{k=n}^{\infty} Á_k\right) \leq \sum_{k=n}^{\infty} P(Á_k) \xrightarrow[n\to\infty]{} 0 \cdot$$

Теорема 1.12 ***(о связи между сходимостью в среднем квадратичном и почти наверное).*** *Если* $î_k(ù) \xrightarrow[k \to \infty]{ср.кв.} î(ù)$ *достаточно быстро, в том смысле, что*

$$\sum_{k=1}^{\infty} Ì\left\{\left|î_k(ù) - î(ù)\right|^2\right\} < +\infty,$$

то $î_k(ù) \xrightarrow[k \to \infty]{п.н.} î(ù)$.

Для доказательства теоремы можно воспользоваться неравенством Чебышева, из которого следует, что

$$P(Á_k(å)) \leq \frac{1}{å^2} Ì\left\{\left|î_k(ù) - î(ù)\right|^2\right\},$$

поэтому

$$\sum_{k=1}^{\infty} P(Á_k(å)) \leq \frac{1}{å^2} \sum_{k=1}^{\infty} Ì\left\{\left|î_k(ù) - î(ù)\right|^2\right\} < +\infty \cdot$$

Таким образом, по лемме Бореля–Кантелли, при этом с вероятностью 1 происходит лишь конечное число событий $Á_k(å)$, откуда следует, что

$$î_k(ù) \xrightarrow[k\to\infty]{п.н.} î(ù) \cdot$$

Для рассмотрения очень важного критерия сходимости в среднем квадратичном нам понадобится следующая лемма.

Лемма. *Пусть*

$$î_k(ù) \xrightarrow[k\to\infty]{ср.кв.} î(ù), \; ç_n(ù) \xrightarrow[n\to\infty]{ср.кв.} ç(ù) \cdot$$

Тогда

$$M\{\xi_k(\omega)\xi_n(\omega)\}\xrightarrow[k,n\to\infty]{} M\{\xi(\omega)\eta(\omega)\}\cdot$$

Доказательство ее следует из соотношения

$$M\{\xi_k(\omega)\eta_n(\omega)-\xi(\omega)\eta(\omega)\}=M\{[\xi_k(\omega)-\xi(\omega)][\eta_n(\omega)-\eta(\omega)]\}+$$
$$+M\{[\xi_k(\omega)-\xi(\omega)]\eta(\omega)\}+M\{[\eta_n(\omega)-\eta(\omega)]\xi(\omega)\}$$

и неравенства Шварца для математических ожиданий, которое используется для каждого слагаемого в правой части этого соотношения, например,

$$\left|M[\xi_k(\omega)-\xi(\omega)][\eta_n(\omega)-\eta(\omega)]\right|^2\le$$
$$\le M[\xi_k(\omega)-\xi(\omega)]^2\,M[\eta_n(\omega)-\eta(\omega)]^2\xrightarrow[k,n\to\infty]{}0\Rightarrow$$
$$\Rightarrow M\{[\xi_k(\omega)-\xi(\omega)][\eta_n(\omega)-\eta(\omega)]\}\xrightarrow[k,n\to\infty]{}0.$$

Теорема 1.13 ***(критерий сходимости в среднем квадратичном).*** *Для того чтобы последовательность* $\{\xi_k(\omega),\ k=1,2,\ldots\}$ *сходилась в среднем квадратичном к некоторой СВ, необходимо и достаточно, чтобы существовал предел*

$$\lim_{k,n\to\infty} M\{\xi_k(\omega)\xi_n(\omega)\}=A<+\infty\cdot$$

Докажем *необходимость*. Пусть $\xi_k(\omega)\xrightarrow[k\to\infty]{ср.кв.}\xi(\omega)$. Положим $\eta_n(\omega)=\xi_n(\omega)$, тогда, по лемме, будем иметь

$$M\{\xi_k(\omega)\eta_n(\omega)\}=M\{\xi_k(\omega)\xi_n(\omega)\}\xrightarrow[k,n\to\infty]{}M\xi^2(\omega)$$

и в качестве A можно взять $M\xi^2(\omega)$.

Достаточность. Пусть $\exists\lim\limits_{k,n\to\infty} M\{\xi_k(\omega)\xi_n(\omega)\}=A$. Тогда

$$M\{\xi_k(\omega)-\xi_n(\omega)\}^2=,$$
$$=M\xi_k^2(\omega)+M\eta_n^2(\omega)-2M\{\xi_k(\omega)\eta_n(\omega)\}\xrightarrow[k,n\to\infty]{}A+A-2A=0,$$

следовательно, последовательность $\{\xi_k(\omega),\ k=1,2,\ldots\}$ является фундаментальной в среднем квадратичном и поэтому сходится в этом смысле.

Рассмотрим непосредственные соотношения между различными типами сходимости случайных последовательностей.

Очевидно, что

$$P\left[\bigcup_{n\ge k}\{\omega:|\xi_n(\omega)-\xi(\omega)|\ge\varepsilon\}\right]\ge P\{\omega:|\xi_k(\omega)-\xi(\omega)|\ge\varepsilon\},$$

откуда вытекает, что из сходимости почти наверное следует сходимость по вероятности.

Используя неравенства Чебышева и Ляпунова для математических ожиданий, получаем, что

$$P\{\omega:|\xi_k(\omega)-\xi(\omega)|\ge\varepsilon\}\le\frac{1}{\varepsilon}M|\xi_k(\omega)-\xi(\omega)|\le$$

$$\le\frac{1}{\varepsilon}\left(M|\xi_k(\omega)-\xi(\omega)|^r\right)^{\frac{1}{r}}\le\frac{1}{\varepsilon}\left(M|\xi_k(\omega)-\xi(\omega)|^q\right)^{\frac{1}{q}}\ \ \forall\varepsilon>0,\ q>r>1.$$

Поэтому из сходимости в среднем порядка q следует сходимость в среднем порядка $r<q$, в частности, сходимость в среднем; из сходимости в среднем следует сходимость по вероятности.

Пример 1.58. Покажем, что из сходимости по вероятности следует сходимость по распределению.

Пусть $\xi_k(\omega)\xrightarrow[k\to\infty]{P}\xi(\omega)$ и x – точка непрерывности ф.р. $F_\xi(x)$. Тогда $\forall\ \varepsilon>0$ имеем

$$F_{\xi_k}(x)=P\{\omega:\xi_k(\omega)\le x\}=P\{\omega:\xi_k(\omega)\le x,\ |\xi_k(\omega)-\xi(\omega)|\le\varepsilon\}+$$
$$+P\{\omega:\xi_k(\omega)\le x,\ |\xi_k(\omega)-\xi(\omega)|>\varepsilon\}.$$

Если $\xi_k(\omega)\le x,\ |\xi_k(\omega)-\xi(\omega)|\le\varepsilon$, то $\xi(\omega)\le x+\varepsilon$, и поэтому

$$F_{\xi_k}(x)\le P\{\omega:\xi(\omega)\le x+\varepsilon\}+P\{\omega:|\xi_k(\omega)-\xi(\omega)|>\varepsilon\}\cdot$$

Аналогично

$$F_{\xi_k}(x) \geq P\{\omega : \xi_k(\omega) \leq x,\ |\xi_k(\omega) - \xi(\omega)| \leq \varepsilon\} \geq$$
$$\geq P\{\omega : \xi(\omega) + \varepsilon \leq x,\ |\xi_k(\omega) - \xi(\omega)| \leq \varepsilon\} \geq$$
$$\geq P\{\omega : \xi(\omega) \leq x - \varepsilon\} - P\{\omega : |\xi_k(\omega) - \xi(\omega)| > \varepsilon\}.$$

Переходя в полученных неравенствах к пределу при $k \to \infty$ с учетом того, что $\lim\limits_{k\to\infty} P\{\omega : |\xi_k(\omega) - \xi(\omega)| > \varepsilon\} = 0$, будем иметь

$$F_{\xi}(x - \varepsilon) \leq \lim_{k\to\infty} F_{\xi_k}(x) \leq F_{\xi}(x + \varepsilon)\cdot$$

Взяв предел при $\varepsilon \to 0$, получим требуемый результат.

Соотношение между различными типами сходимости случайных последовательностей можно отобразить на следующей схеме (рис. 14).

$$\boxed{\xi_k(\omega) \xrightarrow[k\to\infty]{ср.п.q} \xi(\omega)} \overset{q>r}{\Rightarrow} \boxed{\xi_k(\omega) \xrightarrow[k\to\infty]{ср.п.r} \xi(\omega)}$$

$$\Downarrow$$

$$\boxed{\xi_k(\omega) \xrightarrow[k\to\infty]{п.н.} \xi(\omega)} \Rightarrow \boxed{\xi_k(\omega) \xrightarrow[k\to\infty]{P} \xi(\omega)} \Rightarrow \boxed{\xi_k(\omega) \xrightarrow[k\to\infty]{ср.п.r} \xi(\omega)}$$

Рис. 14

§ 9. ЗАКОН БОЛЬШИХ ЧИСЕЛ

С законом больших чисел в форме Бернулли мы познакомились в §3. Из приведенного там утверждения следует, что доля успешных испытаний из n независимых испытаний Бернулли при $n \to \infty$ приближается к вероятности одного успешного испытания (частота появления события стремится к вероятности данного события).

Определение. Если $\frac{1}{n}\sum\limits_{k=1}^{n}(\xi_k(\omega) - a_k) \xrightarrow[n\to\infty]{P} 0$, где $a_k = M\xi_k(\omega)$, то говорят, что для последовательности СВ $\{\xi_k(\omega), k = 1, 2, \ldots\}$ выполняется **закон больших чисел**.

Используя определение сходимости по вероятности, получаем, что в этом случае

$$\lim_{n\to\infty} P\left\{\omega : \left|\frac{1}{n}\sum_{k=1}^{n}\left(\xi_k(\omega) - a_k\right)\right| \geq \varepsilon\right\} = 0,$$

или, что то же самое,

$$\lim_{n\to\infty} P\left\{\omega : \left|\frac{1}{n}\sum_{k=1}^{n}\left(\xi_k(\omega) - a_k\right)\right| < \varepsilon\right\} = 1.$$

Пример 1.59 **(***закон больших чисел в форме Чебышева***).** Пусть СВ $\xi_1(\omega), \xi_2(\omega), \ldots$ независимы и имеют одинаковые математические ожидания $M\xi_k(\omega) = a$ и дисперсии $D\xi_k(\omega) = \sigma^2$. Тогда для них имеет место закон больших чисел, т.е.

$$\lim_{n\to\infty} P\left\{\omega : \left|\frac{1}{n}\sum_{k=1}^{n}\xi_k(\omega) - a\right| < \varepsilon\right\} = 1.$$

Это вытекает из неравенства Чебышева для дисперсий, из которого следует, что

$$P\left\{\omega : \left|\frac{1}{n}\sum_{k=1}^{n}\xi_k(\omega) - a\right| \geq \varepsilon\right\} \leq \frac{D\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k(\omega)\right)}{\varepsilon^2} = \frac{1}{n^2}\cdot\frac{n\sigma^2}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

Отметим, что аналогичный результат имеет место и в случае, когда СВ $\xi_1(\omega), \xi_2(\omega), \ldots$ независимы, одинаково распределены и имеют конечное математическое ожидание (*теорема Хинчина*).

Смысл закона больших чисел состоит в том, что среднее арифметическое СВ $\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)$ при достаточно большом n будет с большой вероятностью почти не отличаться от $\frac{1}{n}\sum_{k=1}^{n} M\xi_k(\omega)$, или, в частности, от a.

В экспериментальных науках среднее арифметическое $\overline{x} = \frac{1}{n}\sum_{k=1}^{n} x_k$ результатов $x_1, x_2, \ldots, x_n$ измерений некоторой величины a рассматривают как более точное приближение к истинному значению a этой величины по сравнению с отдельным измерением. Вероятностная модель измерений дается последовательностью СВ $\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)$. Если измерения не содержат систематической ошибки, т.е. отклонение $x_i - a$ объясняется чисто случайными погрешностями, то следует предположить, что $\grave{I}\,(\hat{\imath}_k(\grave{u}) - a) = 0$, т.е. $M\xi_k(\omega) = a,\ k = \overline{1, n}$. Если к тому же измерения независимы и одинаково точны, т.е. $D\xi_k(\omega) = \sigma^2,\ k = \overline{1, n}$, то закон больших чисел позволяет объяснить экспериментально наблюдаемую закономерность о стабилизации с ростом n средних арифметических $\overline{x}$ вблизи a. Качество приближения к a оценки $S(x_1, x_2, \ldots, x_n) = \overline{x}$ при конечных n можно характеризовать неравенством Чебышева

$$P\{\omega : |S(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)) - a| \geq \varepsilon\} \leq \frac{DS}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2},$$

оценивающим степень конкретизации распределения вероятностей СВ $S(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))$ вокруг точки a. Таким образом, точность приближения $S(x_1, x_2, \ldots, x_n) = \overline{x} \approx a$ можно оценивать одним числом – дисперсией DS.

С помощью закона больших чисел в форме Чебышева можно легко доказать утверждение для схемы Бернулли, приведенное в начале параграфа. Именно, пусть

$$\xi_k(\omega) = \begin{cases} 1, \text{если } k-e \text{ испытание успешно,} \\ 0, \text{если } k-e \text{ испытание неудачно,} \end{cases}$$

p – вероятность успеха в отдельном испытании. Тогда $\sum_{k=1}^{n} \hat{\imath}_k(\grave{u})$ – число успехов в n испытаниях, а $\frac{1}{n}\sum_{k=1}^{n} \xi_k(\omega)$ – доля успехов в n испытаниях. Поскольку $M\xi_k(\omega) = p$ и $\hat{\imath}_k^2(\grave{u}) = \hat{\imath}_k(\grave{u})$, то $\grave{I}\hat{\imath}_k^{\,2}(\grave{u}) = p$, и поэтому $D\hat{\imath}_k(\grave{u}) = p - p^2$. Легко проверить, что $p - p^2 \le \frac{1}{4}$, таким образом,

$$P\left\{\omega : \left|\frac{1}{n}\sum_{k=1}^{n} \xi_k(\omega) - p\right| \ge \varepsilon\right\} \xrightarrow[n\to\infty]{} 0 \cdot$$

Приведем результаты, связанные с выполнением закона больших чисел в более общих ситуациях.

Теорема 1.14. *Для того чтобы для последовательности СВ* $\{\xi_k(\omega), k = 1, 2, \ldots\}$ *(они могут быть и зависимыми) выполнялся закон больших чисел, необходимо и достаточно, чтобы*

$$\lim_{n\to\infty} \grave{I}\left\{\frac{\left[\sum_{k=1}^{n}\left(\hat{\imath}_k(\grave{u}) - a_k\right)\right]^2}{n^2 + \left[\sum_{k=1}^{n}\left(\hat{\imath}_k(\grave{u}) - a_k\right)\right]^2}\right\} = 0 \cdot$$

Доказательство. Необходимость. Пусть $\c{c}_n(\grave{u}) = \frac{1}{n}\sum_{k=1}^{n}\left(\hat{\imath}_k(\grave{u}) - a_k\right)$. Тогда

$$P\{\omega : |\eta_n(\omega)| \ge \varepsilon\} = \int\limits_{|y|\ge\varepsilon} dF_{\eta_n}(y) \ge \int\limits_{|y|\ge\varepsilon} \frac{y^2}{1+y^2} dF_{\eta_n}(y) = \int\limits_{-\infty}^{\infty} \frac{y^2}{1+y^2} dF_{\eta_n}(y) -$$

$$- \int\limits_{|y|<\varepsilon} \frac{y^2}{1+y^2} dF_{\eta_n}(y) \ge M\left(\frac{\eta_n^2(\omega)}{1+\eta_n^2(\omega)}\right) - \int\limits_{|y|<\varepsilon} y^2 dF_{\eta_n}(y) \ge M\left(\frac{\eta_n^2(\omega)}{1+\eta_n^2(\omega)}\right) - \varepsilon^2.$$

Отсюда следует, что

$$0 \le M\left(\frac{\eta_n^2(\omega)}{1+\eta_n^2(\omega)}\right) \le \varepsilon^2 + P\{\omega : |\eta_n(\omega)| \ge \varepsilon\}.$$

В необходимости нам дано, что $\lim\limits_{n\to\infty} P\{\omega : |\eta_n(\omega)| \ge \varepsilon\} = 0$. Поэтому, если выбрать ε достаточно малым, а n достаточно большим, мы можем сделать правую часть неравенства сколь угодно малой.

Достаточность. Нам дано, что $\lim\limits_{n\to\infty} M\left(\frac{\eta_n^2(\omega)}{1+\eta_n^2(\omega)}\right) = 0$. Поскольку

$$P\{\omega : |\eta_n(\omega)| \ge \varepsilon\} = \int\limits_{|y|\ge\varepsilon} dF_{\eta_n}(y) \le \frac{1+\varepsilon^2}{\varepsilon^2} \int\limits_{|y|\ge\varepsilon} \frac{y^2}{1+y^2} dF_{\eta_n}(y) \le$$

$$\le \frac{1+\varepsilon^2}{\varepsilon^2} \int\limits_{-\infty}^{\infty} \frac{y^2}{1+y^2} dF_{\eta_n}(y) = \frac{1+\varepsilon^2}{\varepsilon^2} M\left(\frac{\eta_n^2(\omega)}{1+\eta_n^2(\omega)}\right),$$

то $\lim\limits_{n\to\infty} P\{\omega : |\eta_n(\omega)| \ge \varepsilon\} = 0$.

Рассмотрим следствия из этой теоремы.

Теорема Маркова. *Если*

$$\lim_{n\to\infty} \frac{D\left(\sum\limits_{k=1}^{n} \xi_k(\omega)\right)}{n^2} = 0,$$

то имеет место закон больших чисел.

Это вытекает из соотношения

$$\frac{1}{n^2} D\left(\sum_{k=1}^{n} \xi_k(\omega)\right) = M\eta_n^2(\omega) \ge M\left(\frac{\eta_n^2(\omega)}{1+\eta_n^2(\omega)}\right).$$

Теорема Чебышева. *Если последовательность независимых СВ* $\{\xi_k(\omega), k = 1, 2, \ldots\}$ *(не обязательно одинаково распределенных) такова, что* $D\xi_k(\omega) \le c, k = 1, 2, \ldots$, *где* c *– некоторая константа, то имеет место закон больших чисел.*

Это следует из теоремы Маркова и независимости СВ, поскольку

$$\frac{1}{n^2} D\left(\sum_{k=1}^{n} \xi_k(\omega)\right) = \frac{1}{n^2} \sum_{k=1}^{n} D\xi_k(\omega) \le \frac{c}{n}.$$

Пример 1.60. Пусть $\{\xi_k(\omega), k = 1, 2, \ldots\}$ – последовательность независимых СВ и

$$P\{\omega : \xi_k(\omega) = \pm 2^k\} = 2^{-(2k+1)},\ P\{\omega : \xi_k(\omega) = 0\} = 1 - 2^{-2k}.$$

В данном случае

$$M\xi_k(\omega) = \frac{2^k}{2^{2k+1}} - \frac{2^k}{2^{2k+1}} = \frac{1}{2} - \frac{1}{2} = 0,\ M\xi_k^2(\omega) = \frac{2^{2k}}{2^{2k+1}} = \frac{1}{2},$$

поэтому $D\xi_k(\omega) = \frac{1}{2}$, и из теоремы Чебышева следует, что эта последовательность удовлетворяет закону больших чисел.

Приведем вспомогательное неравенство, известное как **неравенство Гаека–Реньи**: пусть $\{\xi_k(\omega), k = 1, 2, \ldots\}$ последовательность независимых СВ, $M\xi_k(\omega) = a_k$, $D\xi_k(\omega) = \sigma_k^2$, $k = 1, 2, \ldots$. Если $c_1, c_2, \ldots$ – невозрастающая последовательность неотрицательных чисел, то для любых целых m и n, таких, что $m < n$, и $\forall\ \varepsilon > 0$,

$$P\left\{\omega : \max_{m \le k \le n} c_k \left|\sum_{i=1}^{k} (\xi_i(\omega) - a_i)\right| \ge \varepsilon\right\} \le \frac{1}{\varepsilon^2}\left(c_m^2 \sum_{i=1}^{m} \sigma_i^2 + \sum_{i=m+1}^{n} c_i^2 \sigma_i^2\right).$$

Теорема 1.15 ***(усиленный закон больших чисел в форме Колмогорова).*** *Пусть* $\{\xi_k(\omega), k = 1, 2, \ldots\}$ *– последовательность независимых СВ* $\sum\limits_{k=1}^{\infty} \dfrac{\sigma_k^2}{k^2} < \infty$. *Тогда для нее выполняется усиленный закон больших чисел, т.е.*

$$\frac{1}{n}\sum_{k=1}^{n}\left(\xi_k(\omega) - a_k\right) \xrightarrow[n\to\infty]{п.н.} 0$$

(в отличие от закона больших чисел здесь вместо сходимости по вероятности рассматривается сходимость почти наверное).

Доказательство. В неравенстве Гаека–Реньи положим $c_k = \dfrac{1}{k}$ и, т.к. оно справедливо $\forall\ n$, то

$$P\left\{\omega : \max_{m \le k} \frac{1}{k}\left|\sum_{i=1}^{k}\left(\xi_i(\omega) - a_i\right)\right| \ge \varepsilon\right\} \le \frac{1}{\varepsilon^2}\left(\frac{1}{m^2}\sum_{i=1}^{m}\sigma_i^2 + \sum_{i=m+1}^{\infty}\frac{\sigma_i^2}{i^2}\right).$$

Поскольку $\sum\limits_{i=1}^{\infty}\dfrac{\sigma_i^2}{i^2} < \infty$, то $\lim\limits_{m\to\infty}\sum\limits_{i=m+1}^{\infty}\dfrac{\sigma_i^2}{i^2} = 0$. В связи с этим, устремляя m к ∞, получаем

$$\lim_{m\to\infty} P\left\{\omega : \max_{m \le k} \frac{1}{k}\left|\sum_{i=1}^{k}\left(\xi_i(\omega) - a_i\right)\right| \ge \varepsilon\right\} = 0,$$

что и доказывает теорему.

Аналогом теоремы Хинчина для закона больших чисел является следующее утверждение.

Теорема 1.16 ***(Колмогорова).*** *Пусть* $\{\xi_k(\omega), k = 1, 2, \ldots\}$ *– последовательность независимых одинаково распределенных СВ. Для выполнения усиленного закона больших чисел необходимо и достаточно существования у СВ конечного математического ожидания.*

Пример 1.61. Вернемся к схеме независимых испытаний Бернулли. Пусть $v_n(\omega)$ – число успехов в n испытаниях Бернулли с вероятностью успеха p. Тогда, как рассмотрено выше, $ί_n(ù)=\sum_{k=1}^{n} î_k(ù)$, где $\xi_k(\omega), k=\overline{1,n}$ – независимые одинаково распределенные СВ, $P\{\omega:\xi_k(\omega)=1\}=p,\ P\{\omega:\xi_k(\omega)=0\}=1-p$. Поскольку $M\xi(\omega)=p, k=\overline{1,n}$, то

$$\frac{ί_n(ù)}{n}\xrightarrow[n\to\infty]{п.н.} p\,.$$

§ 10. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

В данном параграфе рассмотрим группу предельных теорем теории вероятностей, устанавливающих сходимость законов распределения для последовательностей сумм случайных величин. Эта группа теорем ввиду ее особой важности, как для теории, так и для приложений, носит название **центральной предельной теоремы**. Две из таких теорем, а именно, теоремы Муавра–Лапласа, мы рассматривали в §3. Они являются частными случаями более общих теорем, которые мы ниже рассмотрим.

Теорема 1.17 ***(Линдеберга).*** *Если последовательность независимых СВ* $\xi_1(\omega), \xi_2(\omega), \ldots$ *при любых* $ô>0$ *удовлетворяет условию Линдеберга*

$$\lim_{n\to\infty}\frac{1}{Â_n^2}\sum_{k=1}^{n}\int\limits_{|x-a_k|>ôÂ_n}(x-a_k)^2\,dF_{î_k}(x)=0,$$

где $F_{î_k}(x)$ – ф.р. СВ $\xi_k(\omega)$, $a_k=M\xi_k(\omega)$, $B_n^2=D\left(\sum_{k=1}^{n}\xi_k(\omega)\right)$, то при $n\to\infty$ равномерно относительно x

$$P\left\{\omega : \frac{1}{B_n}\sum_{k=1}^{n}\left(\xi_k(\omega) - a_k\right) < x\right\} \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt \cdot$$

Теорему Линдеберга используют в основном в теоретических работах, требующих большой общности рассмотрения. В приложениях чаще применяют теорему Ляпунова, условия которой проверяются более эффективно.

Теорема 1.18 ***(Ляпунова).*** *Если для последовательности независимых СВ* $\xi_1(\omega), \xi_2(\omega), \ldots$ *можно подобрать такое* $\delta > 0$, *что*

$$\frac{1}{B_n^{2+\delta}}\sum_{k=1}^{n} M\left|\xi_k(\omega) - a_k\right|^{2+\delta} \xrightarrow[n\to\infty]{} 0,$$

то справедлива центральная предельная теорема.

Для ее доказательства достаточно показать, что из условия Ляпунова вытекает условие Линдеберга. Если $\left|x - a_k\right| > \tau B_n$, то

$$\frac{\left|x - a_k\right|^{\delta}}{\tau^{\delta} B_n^{\delta}} > 1,$$

поэтому

$$\frac{1}{B_n^2}\sum_{k=1}^{n}\int\limits_{\left|x-a_k\right|>\tau B_n}(x - a_k)^2\, dF_{\xi_k}(x) \le \frac{1}{\tau^{\delta} B_n^{2+\delta}}\sum_{k=1}^{n}\int\limits_{\left|x-a_k\right|>\tau B_n}\left|x - a_k\right|^{2+\delta} dF_{\xi_k}(x) \le$$

$$\le \frac{1}{\tau^{\delta} B_n^{2+\delta}}\sum_{k=1}^{n}\int_{-\infty}^{\infty}\left|x - a_k\right|^{2+\delta} dF_{\xi_k}(x) = \frac{1}{\tau^{\delta} B_n^{2+\delta}}\sum_{k=1}^{n} M\left|\xi_k(\omega) - a_k\right|^{2+\delta} \xrightarrow[n\to\infty]{} 0$$

т.е. действительно из условия Ляпунова следует условие Линдеберга.

Пример 1.62. Пусть СВ $\xi_1(\omega), \xi_2(\omega), \ldots$ независимы и

$$M\xi_k(\omega) = 0,\ M\,\xi_k^2 = \sigma_k^2,\ 0 < \sigma_k^2 < \infty,\ M\left|\xi_k(\omega)\right|^3 < \infty,\ k = 1, 2, \ldots \cdot$$

Тогда $B_n = \left(\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2\right)^{\frac{1}{2}}$. Предположим, что выполнено условие Ляпунова

$$\frac{1}{\hat{A}_n^3}\sum_{k=1}^{n} \grave{I} \left|\hat{\imath}_k(\grave{u})\right|^3 \xrightarrow[n\to\infty]{} 0 \cdot$$

Воспользуемся неравенством Ляпунова для математического ожидания

$$\left(M\left|\xi_k(\omega)\right|^s\right)^{\frac{1}{s}} \le \left(M\left|\xi_k(\omega)\right|^t\right)^{\frac{1}{t}},\ 0 < s < t\,,$$

из которого вытекает, что

$$\sigma_k^3 = \left(M\xi_k^2(\omega)\right)^{\frac{3}{2}} \le M\left|\xi_k(\omega)\right|^3,$$

и поэтому

$$\frac{1}{B_n^3}\max_{k\le n}\sigma_k^3 \le \frac{1}{B_n^3}\sum_{k=1}^{n}\sigma_k^3 \xrightarrow[n\to\infty]{} 0 \cdot$$

Но т.к. $\sigma_k < B_n,\ k = \overline{1,\ n}$, то

$$\frac{1}{B_n^2}\max_{k\le n}\sigma_k^2 \xrightarrow[n\to\infty]{} 0\,,$$

а из этого соотношения следует условие Линдеберга, т.е. в данном случае справедлива центральная предельная теорема.

Нужно отметить, что, если все $\xi_k(\omega),\ k = 1,\ 2,\ \ldots$, независимы и одинаково распределены, $D\xi_k(\omega) = \sigma^2 < \infty,\ k = 1,\ 2,\ \ldots$, условие Линдеберга также выполняется и

$$P\left\{\grave{u} : \frac{\sum_{k=1}^{n}\hat{\imath}_k(\grave{u}) - na}{\sigma\sqrt{n}} < x\right\} \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\eth}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt \cdot$$

Практический смысл центральной предельной теоремы состоит в следующем: если некоторый процесс происходит под воздействием большого числа независимо действующих случайных факторов, каждый из которых лишь очень мало изменяет ход процесса, то распределение суммарного действия этих случайных факторов можно очень близко аппроксимировать нормальным законом.

Пример 1.63. Нагрузка потребительской сети (телефонной, информационной, электрической и т.п.) в данный момент времени является результатом суммирования большого числа элементарных нагрузок, вносимых индивидуальными потребителями. Поскольку случайности, определяющие поведение потребителей, естественно считать независимыми, то имеет место суммирование большого числа независимых СВ, пренебрежимо мало влияющих на сумму. Различные статистические исследования подтверждают нормальный закон нагрузки.

Пример 1.64. Всякое измерение неизбежно сопряжено с погрешностями; систематические погрешности можно устранить, случайные же ошибки измерения полностью никогда не могут быть устранены. Случайная погрешность измерения вызывается многими причинами, каждая из которых лишь незначительно влияет на результат. Каждая из причин порождает свою, так называемую элементарную, погрешность измерения. Например, при взвешивании некоторого тела на точных весах случайная ошибка в определении его веса складывается из элементарных погрешностей, вызываемых атмосферными причинами (колебаниями температуры, влажности воздуха, воздушными потоками); попаданием на чаши весов пылинок; неточностями, допущенными измерителем при снятии показаний со шкалы весов; незначительными вибрациями основания весов (которые в свою очередь могут вызываться многими причинами) и т.п. Реально наблюдаемая случайная погрешность измерения есть *сумма элементарных* погрешностей. Так как количество элементарных погрешностей велико и роль каждой из них в образовании случайной погрешности измерения мала, то в силу теоремы Ляпунова случайная погрешность измерения должна быть распределена приближенно *по нормальному закону.*

Опыт показывает, что наблюдающиеся распределения вероятностей случайных ошибок измерения очень хорошо согласуются с нормальным законом. Итак, при *прямых измерениях случайная погрешность измерения распределена по закону, близкому к нормальному.*

Пример 1.65. Дано 5000 независимых одинаково распределенных СВ с дисперсией 50. Найти вероятность того, что среднее арифметическое этих СВ отклонится от своего математического ожидания не более, чем на $\Delta = 0{,}12$.

Решение. Пусть $M\xi_k(\omega)=a,\ D\xi_k(\omega)=\sigma^2$. Тогда искомая вероятность равна

$$P=P\left\{\omega:\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k(\omega)-a\right|\le\Delta\right\}=P\left\{\omega:\frac{\sigma}{\sqrt{n}}\left|\frac{\sum\limits_{k=1}^{n}\xi_k(\omega)-a}{\sigma\sqrt{n}}\right|\le\Delta\right\}=$$

$$=P\left\{\omega:\left|\frac{\sum\limits_{k=1}^{n}\xi_k(\omega)-a}{\sigma\sqrt{n}}\right|\le\frac{\sqrt{n}}{\sigma}\Delta\right\}\approx\Phi\left(\frac{\sqrt{n}}{\sigma}\Delta\right)-\Phi\left(-\frac{\sqrt{n}}{\sigma}\Delta\right)=2\Phi\left(\frac{\sqrt{n}}{\sigma}\Delta\right),$$

где $\Phi(x)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_0^x e^{-\frac{t^2}{2}}dt$.

Подставляя значения $n=5000,\ \sigma=\sqrt{50},\ \Delta=0{,}12$, получаем $P\approx 2\Phi(1{,}2)=0.7698\cdot$

Пример 1.66. Для вычисления интеграла $I=\displaystyle\int_a^b f(x)dx$ часто используют метод статистических испытаний (Монте-Карло), при котором

$$I\approx I_n=\frac{1}{n}\sum_{k=1}^{n}f(\xi_k(\omega)),$$

где $\xi_k(\omega), k=\overline{1,n}$, – независимые, равномерно распределенные на отрезке $[a,b]$ СВ. Сколько нужно провести статистических испытаний при вычислении интеграла $\displaystyle\int_0^{\frac{\pi}{2}}\cos x dx$, чтобы с вероятностью

0,9 можно было считать относительную погрешность в вычисленном значении интеграла меньшей, чем 5 %?

Решение. Интеграл

$$\frac{2}{\pi} I = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \cos x dx$$

можно рассматривать как математическое ожидание функции $\cos(\xi(\omega))$, где $\xi(\omega)$ – равномерно распределенная на интервале $\left[0, \frac{\pi}{2}\right]$ СВ. Тогда приближенное значение интеграла равно

$$I_n = \frac{\pi}{2n}\sum_{k=1}^{n} \cos x_k ,$$

где x_k – случайные числа из интервала $\left[0, \frac{\pi}{2}\right]$. Т.к. $\cos x_k$ независимы, одинаково распределены и имеют конечную, отличную от нуля, дисперсию, то

$$DI_n = \frac{\pi^2}{4n} D\cos(\xi(\omega)) = \frac{\pi^2 - 8}{8n}$$

и

$$P\left\{\frac{I_n - I}{\sqrt{DI_n}} < \varepsilon\right\} = P\left\{\sqrt{\frac{8n}{\pi^2 - 8}}\left|I_n - I\right| < \varepsilon\right\} \approx$$

$$\approx \Phi_1(\varepsilon) - \Phi_1(-\varepsilon) = 2\Phi_1(\varepsilon) - 1 = 0,9 ,$$

где $\Phi_1(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$. Отсюда с помощью таблицы для функции $\Phi_1(x)$ находим $\varepsilon = 1,645$. Для того чтобы относительная погрешность $\frac{|\acute{E}_n - \acute{E}|}{\acute{E}}$ была меньше $0{,}05$, учитывая, что $\acute{E} = 1$, необходимо провести столько опытов, чтобы

$$\sqrt{\frac{8n}{\pi^2 - 8}}\, 0{,}05 > 1{,}645,$$

откуда получаем $n > 252$.

ГЛАВА 2. ОСНОВНЫЕ ПОНЯТИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

§11. ОПРЕДЕЛЕНИЕ СЛУЧАЙНОГО ПРОЦЕССА И ПРИМЕРЫ

При изучении различных реальных явлений рассматривают всевозможные величины, зависящие от времени: $x = x(t)$. Такие величины принято называть процессами. Говоря о случайных процессах, как правило, имеют в виду некоторую СВ $\xi(t)$, меняющуюся с течением времени t. Будем называть случайным процессом $\xi(t)$ функцию от действительного параметра $t \in T$ (времени), значения которой при каждом t являются СВ.

Строгое определение случайного процесса состоит в следующем.

Определение. Набор $\xi(t) = \{\xi(t,\omega),\ t \in T,\ \omega \in \Omega\}$ СВ, определенных на одном и том же вероятностном пространстве (Ù ,F,P), называется случайной функцией.

Определение. Пусть (Ù ,F,P) – вероятностное пространство, T – некоторое множество моментов времени. И пусть каждому $\omega \in \Omega$ поставлена в соответствие функция $\xi_t = \xi(t) = \xi(t,\omega)$, $t \in T$, со значениями в n-мерном пространстве $(n \geq 1)$ такая, что при каждом фиксированном $t \in T$ функция $\xi(t,\omega)$ является СВ. Эта функция называется случайным процессом.

Следовательно, случайный процесс – это совокупность СВ, зависящих от времени. Естественно, его можно рассматривать как функцию двух переменных: t и ω. Как следует из определения, $\xi(t,\omega)$ является n–мерной СВ при любом фиксированном t.

Определение. При фиксированном $\omega \in \Omega$ функция $\xi(t,\omega)$ называется реализацией (траекторией или выборочной функцией) случайного процесса. Совокупность всех реализаций случайного процесса называется ансамблем (семейством) реализаций.

Пример 2.1. На метеорологической станции фиксируется температура воздуха $\xi(t, \omega)$, которая является функцией двух переменных: времени и элементарного события ù. Пусть t^* – некоторый настоящий момент времени. На основании проведенных измерений мы знаем, как вел себя процесс $\xi(t)$ в прошлом при $t < t^*$, и не знаем, как он будет вести себя в будущем. Мы можем только знать прогнозы, как он поведет себя при $t > t^*$, но их точность уменьшается с ростом разности $t - t^*$. При каждом фиксированном $t > t^*$ температура $\xi(t)$ является случайной величиной. Обычно мы не имеем возможности точно предсказать, какой будет температура в будущем; она зависит от многих случайных факторов, совместное влияние которых можно записать как элементарное событие ω. Поэтому температура воздуха как функция данных двух переменных является случайным процессом. То же самое можно сказать о количестве атмосферных осадков, уровне воды в реке в зависимости от времени и места.

Пример 2.2. Частица находится в начальный момент времени $t = 0$ в точке ζ на оси и затем движется вдоль этой оси со скоростью η. Причем (η, ζ) является двумерной СВ, имеющей нормальное распределение с параметрами $N(a_\eta, a_\zeta, \sigma_\eta^2, \sigma_\zeta^2)$. Положение частицы на оси можно выразить соотношением $\xi(t, \omega) = \eta(\omega)t + \zeta(\omega)$. Реализациями случайного процесса $\xi(t, \omega)$ являются полупрямые вида $\xi(t) = at + b$, $t \geq 0$, где (a, b) – значение двумерной СВ $(\eta(\omega) + \zeta(\omega))$ при некотором фиксированном $\omega = \overline{\omega}$.

Данный случайный процесс можно трактовать также как семейство реализаций вида $\eta(\omega)t+\zeta(\omega)$, где $\omega\in\Omega$, а также как семейство СВ $\xi(t,\omega)$, таких, что в каждый фиксированный момент времени $t=t^*\geq 0$ СВ $\xi(t^*,\omega)$ имеет нормальное распределение с параметрами $N(a_\eta t+a_\zeta,\ \sigma_\eta^{\ 2}t^2+2r\sigma_\eta\sigma_\zeta t+\sigma_\zeta^{\ 2})$, где r – коэффициент корреляции СВ $\eta(\omega)$ и $\zeta(\omega)$. Это можно показать, используя функциональные преобразования СВ.

Определение. Случайный процесс $\xi(t)$, $t\in T$, у которого множество T не более чем счетно, называется процессом с дискретным временем или случайной последовательностью. Если множество T совпадает с промежутком прямой, например, $T=[0;+\infty)$, то $\xi(t)$ называется процессом с непрерывным временем.

Пример 2.3. Важным примером случайного процесса с непрерывным временем, $T=[0;+\infty)$, является пуассоновский процесс [4], который обозначим $\nu(t)=\nu(t,\omega)$. Его реализация представляет собой число наступлений некоторого события за период $[0,t)$. Очевидно, всякая возможная реализация является неубывающей ступенчатой функцией, которая возрастает единичными скачками в некоторые случайные моменты времени $t_1,t_2,\ldots$ (рис 15). Очевидно также $\nu(0)=0$. Конкретными примерами наблюдаемых величин, образующих подобного рода процессы, являются: число телефонных вызовов на коммутатор телефонной станции; число фотонов, испускаемых веществом при радиоактивном распаде; число происшествий на данном перекрестке; число ошибок в тексте, набираемом на компьютере; число выходов из строя некоторого устройства; число вызовов, поступающих на станцию скорой помощи; число запросов, поступающих на сервер компьютерной сети. Рассмотрение этих процессов как пуассоновских основывается на законе редких событий. Представим себе большое число испытаний Бернулли с малой вероятностью наступления события и постоянным средним числом таких наступлений. При этих условиях локальная предельная теорема Пуассона утверждает, что число наступлений события подчиняется закону Пуассона.

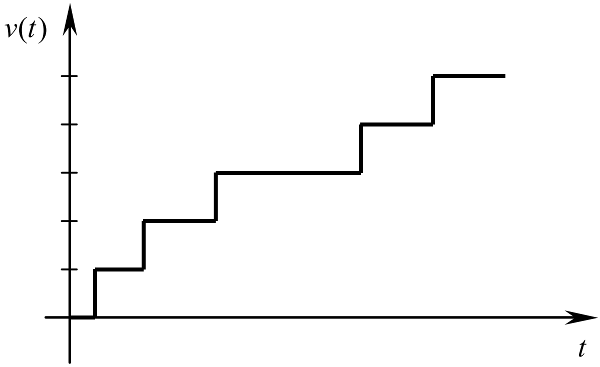


Рис. 15

Предположим, что число наступлений события в некотором интервале не зависит от числа наступлений этого события в любом другом не пересекающемся с ним интервале. Предположим также, что СВ $\nu(t_0+t)-\nu(t_0)$, зависит только от t и не зависит от t_0 или значения $\nu(t_0)$. Сформулируем еще два предположения, согласующихся с интуитивным описанием процесса, данным выше:

а) вероятность того, что за период времени Δt произойдет по меньшей мере одно событие, равна

$$p(\Delta t)=\lambda\Delta t+o(\Delta t),\ \Delta t\to 0,\ \lambda>0,$$

напомним, что $\lim\limits_{\Delta t\to 0}\dfrac{o(\Delta t)}{\Delta t}=0$;

б) вероятность того, что за время Δt произойдут два или более события, есть $o(\Delta t)$; это предположение равнозначно исключению возможности более чем однократного одновременного наступления события; в приведенных нами физических процессах это условие обычно выполняется.

Пусть $P_k(t)$ означает вероятность того, что за время t произойдет ровно k событий, т.е.

$$P_k(t) = P\{\nu(t) = k\},\ k = 0,\ 1,\ 2, \ldots \cdot \tag{2.1}$$

Условие б) можно записать в виде $\sum_{k=2}^{\infty} P_k(\Delta t) = o(\Delta t)$ и, очевидно, $p(\Delta t) = \sum_{k=1}^{\infty} P_k(\Delta t)$. В силу предположения о независимости имеем

$$P_0(t + \Delta t) = P_0(t)\, P(\Delta t) = P_0(t)(1 - p(\Delta t)),$$

и поэтому

$$\frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = -\frac{p(\Delta t)}{\Delta t} P_0(t) \cdot$$

Из предположения а) следует, что $\lim\limits_{\Delta t \to 0} \dfrac{p(\Delta t)}{\Delta t} = \lambda$. Поэтому вероятность $P_0(t)$ того, что событие не наступит на интервале $[0, t)$, удовлетворяет дифференциальному уравнению $P_0^{'}(t) = -\lambda P_0(t)$, общее решение которого имеет вид $P_0(t) = c\, e^{-\lambda t}$. Константа c определяется из начального условия $P_0(t) = 1$, из которого следует, что $c = 1$. Поэтому

$$P_0(t) = e^{-\lambda t} \cdot \tag{2.2}$$

Найдем теперь $P_k(t)$ $\forall k \geq 1$. Легко видеть, что

$$P_k(t + \Delta t) = P_k(t)\, P_0(\Delta t) + P_{k-1}(t) P_1(\Delta t) + \sum_{i=2}^{k} P_{k-i}\, P_i(\Delta t). \tag{2.3}$$

Поскольку по определению $P_0(\Delta t) = 1 - p(\Delta t)$ и из предположения а) следует, что

$$P_1(\Delta t) = p(\Delta t) + o(\Delta t),$$

а также справедливо очевидное неравенство

$$\sum_{i=2}^{k} P_{k-i} P_i(\Delta t) \le \sum_{i=2}^{k} P_i(\Delta t) = o(\Delta t),$$

так как $P_{k-i}(t) \le 1, \; i = \overline{2,k}$, то (2.3) мы можем переписать в виде

$$P_k(t+\Delta t) - P_k(t) = -P_k(t)\, p(\Delta t) + P_{k-1}(t) P_1(\Delta t) + \sum_{i=2}^{k} P_{k-i} P_i(\Delta t) =$$

$$= -\lambda P_k(t)\Delta t + \lambda P_{k-1}(t)\Delta t + o(\Delta t).$$

Отсюда следует, что

$$\frac{P_k(t+\Delta t) - P_k(t)}{\Delta t} = -\lambda P_k(t) + \lambda P_{k-1}(t) + \frac{o(\Delta t)}{\Delta t},$$

и при $\Delta t \to 0$, получаем

$$P_k'(t) = -\lambda P_k(t) + \lambda P_{k-1}(t), \; k = 1, 2, \ldots \tag{2.4}$$

при начальных условиях $P_k(0) = 0, \; k = 1, 2, \ldots$.

Для решения разностно–дифференциальных уравнений (2.4) введем функции

$$Q_k(t) = P_k(t)\, e^{-\lambda t}, \; k = 0, 1, 2, \ldots.$$

Подставляя $Q_k(t)$ в (2.3), будем иметь

$$Q_k(t) = \lambda Q_{k-1}(t), \; k = 1, \; 2, \ldots, \tag{2.5}$$

где $Q_0(t) = 1$, поскольку имеет место соотношение (2.2); начальные условия остаются теми же: $Q_k(0) = 0, \; k = 1, \; 2, \ldots$. Последовательно решая уравнения (2.5), получаем:

$$Q_1'(t) = \lambda \text{ или } Q_1(t) = \lambda t + c_1, \text{ поэтому } Q_1(t) = \lambda t;$$

$$Q_2'(t)=\frac{\lambda^2 t}{2} \text{ или } Q_2(t)=\frac{(\lambda t)^2}{2}+c_2, \text{ поэтому } Q_2(t)=\frac{(\lambda t)^2}{2!};$$

. . .

$$Q_k'(t)=\frac{\lambda^k t^{k-1}}{(k-1)!} \text{ или } Q_k(t)=\frac{(\lambda t)^k}{k!}+c_k, \text{ поэтому } Q_k(t)=\frac{(\lambda t)^k}{k!}.$$

Следовательно, имеем

$$P_k(t)=\frac{(\lambda t)^k}{k!}e^{-\lambda t},\ k=0,\ 1,\ 2,... \tag{2.6}$$

Другими словами, при каждом t СВ $\nu(t)$ подчиняется распределению Пуассона с параметром λt. В частности, среднее число наступлений события за время t равно λt. Для нахождения значений функции $P_k(t)$ можно воспользоваться таблицей из Приложения 3.

Пуассоновский процесс является одним из важных случайных процессов, широко используемых в математическом моделировании, поэтому позднее мы еще вернемся к этому процессу.

Определение. При фиксированном t для случайного процесса $\xi(t,\omega)$ получаем СВ, которую назовем сечением случайного процесса в момент времени t.

В дальнейшем случайный процесс $\xi(t,\omega)$ для удобства будем обозначать $\xi(t)$. Рассмотрим сечение случайного процесса $\xi(t_1)$ в момент t_1. Функция $F(x_1,t_1)=P(\xi(t_1)\le x_1)$ носит название одномерной ф.р. случайного процесса в момент времени t_1. Если зафиксировать два момента времени t_1 и t_2, то вероятность совместного выполнения случайных событий $\{\xi(t_1)\le x_1\}$ и $\{\xi(t_2)\le x_2\}$ задает двумерную ф.р. случайного процесса

$$F(x_1,t_1; x_2,t_2) = P(\xi(t_1) \le x_1, \xi(t_2) \le x_2).$$

Определение. Пусть $t_1 < t_2 < ... < t_n$ – некоторые фиксированные моменты времени. Рассмотрим n-мерную СВ $(\xi(t_1), \xi(t_2),..., \xi(t_n))$. Распределение вероятностей этой СВ называется n-мерным распределением вероятностей случайного процесса, а ф.р. этой СВ

$$\begin{aligned} &F(x_1,t_1; x_2,t_2; ...; x_n,t_n) = \\ &= P\big(\xi(t_1) \le x_1,\ \xi(t_2) \le x_2, ...,\ \xi(t_n) \le x_n\big) \end{aligned} \tag{2.7}$$

соответственно n-мерной ф.р. случайного процесса.

Заметим, что n-мерная ф.р. случайного процесса является функцией $2n$ переменных $x_1, x_2, ..., x_n, t_1, t_2, ..., t_n$, в то время как ф.р. n-мерной СВ является функцией n переменных $x_1, x_2, ..., x_n$.

Отметим также, что для функции $F(x_1,t_1; x_2,t_2; ...; x_n,t_n)$ должны выполняться условия согласованности:

$$\begin{aligned} &F(x_1,t_1; x_2,t_2; ...; x_m,t_m; +\infty, t_{m+1}; ...; +\infty, t_n) = \\ &= F(x_1,t_1; x_2,t_2; ...; x_m,t_m) \quad \forall m < n \end{aligned} \tag{2.8}$$

$$F(x_{i_1},t_{i_1}; x_{i_2},t_{i_2}; ...; x_{i_n},t_{i_n}) = F(x_1,t_1; x_2,t_2; ...; x_n,t_n), \tag{2.9}$$

где $i_1, i_2, ..., i_n$ – любая перестановка индексов $1, 2, ..., n$.

Теперь можно сформулировать еще одно определение случайного процесса.

Определение. Случайным процессом $\xi(t)$, заданным на множестве T, $t \in T$, называется семейство распределений (2.7), удовлетворяющее условиям согласованности (2.8), (2.9). Набор функций $F(x_1,t_1; x_2,t_2; ...; x_n,t_n)$ для $n = 1, 2, ...$ называют конечномерным распределением случайного процесса.

Приведем один из простейших примеров случайного процесса – так называемый веерный процесс и его одномерную ф.р. [6].

Пример 2.4. Пусть η – любая одномерная СВ. Определим случайный процесс формулой $\xi(t) = \eta(t - t_0) + b,\ -\infty < t < +\infty$, где t_0, b – фиксированные числа. При фиксированном t функция $\xi(t)$ является СВ, при этом

$$F(x,t) = P(\eta(t - t_0) + b \le x) = \begin{cases} F_\eta\left(\dfrac{x-b}{t-t_0}\right),\ t > t_0, \\ \int\limits_{-\infty}^{x} \delta(t-b)\,dt,\ t = t_0, \\ 1 - F_\eta\left(\dfrac{x-b}{t-t_0} + 0\right),\ t < t_0. \end{cases}$$

Здесь $\delta(\cdot)$ – δ-функция Дирака, которая определяется заданием интеграла $\int\limits_a^b f(x)\,\delta(x)\,dx$ для любой непрерывной функции $f(x)$ следующим образом:

$$\int\limits_a^b f(x)\,\delta(x)\,dx = \begin{cases} f(0),\ 0 \in [a, b], \\ 0,\ 0 \notin [a, b], \end{cases}$$

аналогично

$$\int\limits_a^b f(x)\,\delta(x-\tau)\,dx = \begin{cases} f(\tau),\ \tau \in [a, b], \\ 0,\ \tau \notin [a, b], \end{cases}$$

поэтому в нашем случае

$$\int\limits_{-\infty}^{x} \delta(t-b)\,dt = \begin{cases} 1,\ b \in [-\infty, x], \\ 0,\ b \notin [-\infty, x]. \end{cases}$$

Если зафиксировать η, то реализация случайного процесса $\xi(t)$ представляет собой линейную функцию, график которой проходит через точку (t_0, b). Различные траектории процесса в данном случае представляют собой пучок прямых («веер») (рис. 16), откуда и название процесса.

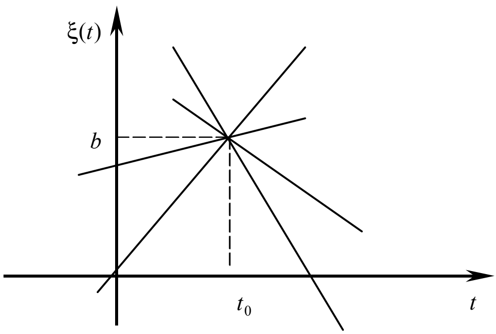


Рис. 16

Определение. Если существует неотрицательная функция $p(x_1,t_1;x_2,t_2;...;x_n,t_n)$, такая, что $\forall\, x_1, x_2, ..., x_n$

$$F(x_1,t_1;x_2,t_2;...;x_n,t_n) =$$

$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}...\int_{-\infty}^{x_n} p(y_1,t_1;y_2,t_2;...;y_n,t_n)\,dy_1\,dy_2...dy_n,$$

то она называется n-мерной плотностью распределения случайного процесса.

Очевидно, что в точках непрерывности плотности распределения имеем

$$p(x_1,t_1;x_2,t_2;...;x_n,t_n) = \frac{\partial^n F(x_1,t_1;x_2,t_2;...;x_n,t_n)}{\partial x_1\,\partial x_2\,...\partial x_n}.$$

При этом условие согласованности (2.8) принимает вид

$$p(x_1,t_1;x_2,t_2;...;x_m,t_m)=$$
$$=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}...\int\limits_{-\infty}^{\infty}p(x_1,t_1;x_2,t_2;...;x_n,t_n)\,dx_{m+1}\,dx_{m+2}...dx_n\ .$$

Характеристическая функция конечномерного распределения вводится аналогично, как для многомерных случайных величин:

$$\varphi(u_1,t_1;u_2,t_2;...;u_n,t_n)=Me^{i\sum\limits_{k=1}^{n}u_k\xi(t_k)}=$$
$$=\int\limits_{-\infty}^{\infty}...\int\limits_{-\infty}^{\infty}e^{i\sum\limits_{k=1}^{n}u_k x_k}p(x_1,t_1;x_2,t_2;...;x_n,t_n)\,dx_1\,dx_2\,...dx_n \qquad (2.10)$$

До сих пор мы рассматривали описание одного случайного процесса. При решении многих задач приходится иметь дело с несколькими случайными процессами. Для задания, например, двух случайных процессов $\eta(t)$ и $\xi(t)$ вводится конечномерная ф.р. размерности $n+m$:

$$F(x_1,t_1;x_2,t_2;...;x_n,t_n;\ y_1,\tau_1;y_2,\tau_2;...;y_m,\tau_m)=$$
$$=P(\xi(t_1)\le x_1,\xi(t_2)\le x_2,...,\ \xi(t_n)\le x_n;$$
$$\eta(\tau_1)\le y_1,\eta(\tau_2)\le y_2,...,\ \eta(\tau_m)\le y_m).$$

Она в общем случае не обладает свойством симметрии относительно всех перестановок аргументов.

Реально протекающие явления, как правило, описываются случайными процессами, принимающими действительные значения. Однако иногда бывает полезным воспользоваться понятием комплексного случайного процесса

$$\zeta(t)=\xi(t)+i\,\eta(t),$$

который можно рассматривать как двумерный случайный процесс, каждая компонента которого является действительным процессом. Конечномерное распределение порядка n случайного процесса $\zeta(t)$ задается $2n$-мерным совместным распределением процессов $\xi(t)$ и $\eta(t)$, т.е.

$$F_\zeta(x_1,t_1;x_2,t_2;...;x_n,t_n;y_1,t_1;y_2,t_2;...;y_n,t_n)\cdot$$

§12. СТАТИСТИЧЕСКИЕ СРЕДНИЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ПРОЦЕССОВ

Конечномерные распределения дают исчерпывающую характеристику случайного процесса. Однако во многих случаях представляют интерес более сжатые характеристики, отражающие основные свойства случайных процессов. Роль таких характеристик играют моментные функции или статистические средние.

Определение. Средним значением случайного процесса $\xi(t)$ (статистическим средним) $m_\xi(t)$ называется математическое ожидание сечения случайного процесса в момент времени t:

$$m_\xi(t) = M\,\xi(t) = \int_{-\infty}^{\infty} x\,dF(x,\, t). \tag{2.12}$$

Оно определяется одномерной функцией распределения $F(x,\, t)$.

Определение. Дисперсией случайного процесса $\xi(t)$ называется дисперсия сечения случайного процесса, которая также определяется одномерным распределением:

$$D_\xi(t) = M[\xi(t) - m_\xi(t)]^2 = \int_{-\infty}^{\infty} [x - m_\xi(t)]\,dF(x,\, t). \tag{2.13}$$

Определение. Корреляционной функцией случайного процесса $\xi(t)$ называется математическое ожидание произведения двух сечений случайного процесса в моменты времени $t_1,\, t_2$:

$$R_\xi(t_1, t_2) = M[\xi(t_1)\xi(t_2)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1\, x_2\, dF(x_1, t_1; x_2, t_2). \tag{2.14}$$

Она определяется двумерной функцией распределения $F(x_1, t_1; x_2, t_2)$. Эту функцию также называют функцией автокорреляции.

Определение. Ковариационная функция случайного процесса – это математическое ожидание произведения центрированных сечений процесса в моменты времени t_1, t_2:

$$K_\xi(t_1,t_2)=\operatorname{cov}\big(\xi(t_1)\xi(t_2)\big)=M\big[\big(\xi(t_1)-m_\xi(t_1)\big)\big(\xi(t_2)-m_\xi(t_2)\big)\big] \qquad (2.15)$$

Величину

$$r_\xi(t_1,t_2)=\frac{K_\xi(t_1,t_2)}{\sqrt{D_\xi(t_1)D_\xi(t_2)}}=\frac{K_\xi(t_1,t_2)}{\sqrt{K_\xi(t_1,t_1)K_\xi(t_2,t_2)}}$$

называют нормированной ковариационной функцией, или коэффициентом корреляции случайного процесса.

Очевидно, что при $t_1=t_2=t$ ковариационная функция совпадает с дисперсией случайного процесса:

$$K_\xi(t,t)=D_\xi(t)=\sigma_\xi^{\,2}(t).$$

Для двух случайных процессов $\xi(t)$ и $\eta(t)$ вводится понятие взаимной функции корреляции, или функции кросс-корреляции:

$$R_{\xi\eta}(t_1,t_2)=M[\xi(t_1)\eta(t_2)], \qquad (2.16)$$

а совместная корреляционная функция двух случайных процессов определяется как матричная функция

$$\begin{pmatrix} R_\xi(t_1,t_2) & R_{\xi\eta}(t_1,t_2) \\ R_{\eta\xi}(t_1,t_2) & R_\eta(t_1,t_2) \end{pmatrix}.$$

Пример 2.5. Пусть случайный процесс $\xi(t)$ является семейством трех реализаций $\{\xi_1(t),\xi_2(t),\xi_3(t)\}$ и определяется следующим образом (рис. 17):

$$\xi_k(t)=kt,\ P\big(\xi(t)=\xi_k(t)\big)=\frac{k}{6},\ k=\overline{1,3},\ T=[0,+\infty).$$

Тогда из формул (2.12) – (2.15) следует

$$m_\xi(t) = M\,\xi(t) = \frac{1}{6}t + \frac{2}{6}\cdot 2t + \frac{3}{6}\cdot 3t = \frac{14}{6}t = \frac{7}{3}t,$$

$$K_\xi(t_1,t_2) = M\left[\left(\xi(t_1) - \frac{7}{3}t_1\right)\left(\xi(t_2) - \frac{7}{3}t_2\right)\right] = \frac{5}{9}t_1 t_2,$$

$$D_\xi(t) = \frac{5}{9}t_2, \qquad r_\xi(t_1,t_2) = 1.$$

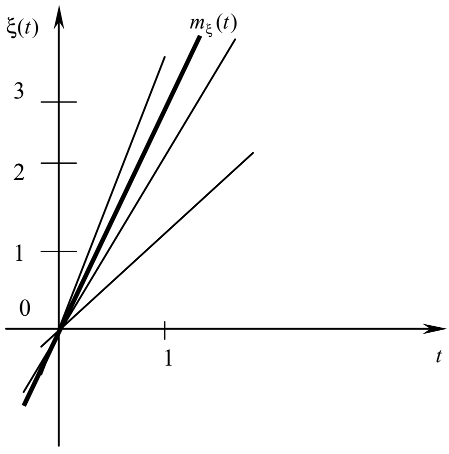


Рис. 17

Пример 2.5. Пусть случайный процесс $\xi(t,\omega)$ определяется выражением

$$\xi(t,\omega) = \alpha(\omega)\cos\lambda t + \beta(\omega)\sin\lambda t, \qquad (2.17)$$

где $\lambda > 0$, $T = [0, +\infty)$, $\alpha(\omega), \beta(\omega) \sim N(0,\sigma^2)$, и эти СВ независимы. Его можно переписать в виде

$$\xi(t,\omega) = \gamma(\omega)\cos\big(\lambda t + \varphi(\omega)\big),$$

то есть
$$\xi(t,\omega)=\gamma(\omega)\left[\cos\lambda t\ \cos\varphi(\omega)-\sin\lambda t\ \sin\varphi(\omega)\right]=$$
$$=\left[\gamma(\omega)\ \cos\varphi(\omega)\right]\cos\lambda t-\left[\gamma(\omega)\sin\varphi(\omega)\ \right]\sin\lambda t. \qquad (2.18)$$

Приравнивая выражения (2.17), (2.18), имеем
$$\alpha(\omega)=\gamma(\omega)\ \cos\varphi(\omega), \qquad \beta(\omega)=-\gamma(\omega)\sin\varphi(\omega),$$
откуда находим
$$\gamma(\omega)=\sqrt{\alpha^2(\omega)+\beta^2(\omega)}, \qquad \cos\varphi(\omega)=\frac{\alpha(\omega)}{\sqrt{\alpha^2(\omega)+\beta^2(\omega)}}\ .$$

Процесс (2.17) характеризует случайные колебания, амплитуда которых $\gamma(\omega)$ и фаза $\varphi(\omega)$ являются СВ. Реализация процесса – периодические функции с периодом $\frac{2\pi}{\lambda}$ (рис. 18).

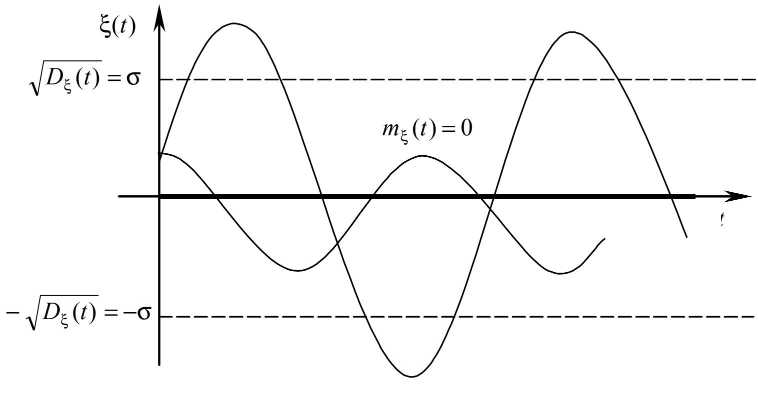


Рис. 18

Найдем среднее значение процесса и его ковариационную функцию :
$$m_\xi(t)=M\alpha(\omega)\cos\lambda t+M\beta(\omega)\sin\lambda t=0,$$
$$K_\xi(t_1,t_2)=M\left[\left(\alpha(\omega)\cos\lambda t_1+\beta(\omega)\sin\lambda t_1\right),\left(\alpha(\omega)\cos\lambda t_2+\beta(\omega)\sin\lambda t_2\right)\right]=$$

$$= M\left(\alpha^2(\omega)\cos\lambda t_1\cos\lambda t_2\right) + M\left(\beta^2(\omega)\sin\lambda t_1\sin\lambda t_2\right) = \sigma^2\cos\lambda(t_2 - t_1), \quad (2.19)$$

поскольку $M\left(\alpha(\omega)\beta(\omega)\right) = M\alpha(\omega)\,M\beta(\omega) = 0$. Используя (2.19), находим дисперсию процесса и нормированную ковариационную функцию:

$$D_\xi(t) = K_\xi(t,t) = \sigma^2,\ r_\xi(t_1,t_2) = \cos\lambda(t_2 - t_1). \quad (2.20)$$

Пример 2.7. Обобщим случайный процесс (2.17) на сумму случайных колебаний. Рассмотрим процесс

$$\xi(t,\omega) = \sum_{k=1}^{n}\left(\alpha_k(\omega)\cos\lambda_k t + \beta_k(\omega)\sin\lambda_k t\right), \quad (2.21)$$

где $\lambda_k > 0,\ k = \overline{1,n},\ T = [0,+\infty),\ \alpha_k(\omega),\beta_k(\omega) \sim N(0,\sigma_k{}^2)$, причем СВ $\alpha_k(\omega),\beta_k(\omega),\ k = \overline{1,n}$, независимы. Для такого процесса будем иметь

$$m_\xi(t) = 0, \qquad K_\xi(t_1,t_2) = \sum_{k=1}^{n}\sigma_k{}^2\cos\lambda_k(t_2 - t_1).$$

Перечислим основные свойства корреляционной функции случайного процесса $\xi(t)$.

1. Корреляционная функция является симметричной функцией своих аргументов:

$$R_\xi(t_1,t_2) = R_\xi(t_2,t_1), \quad (2.22)$$

что следует из ее определения.

2. $\left|R_\xi(t_1,t_2)\right| \le \sqrt{R_\xi(t_1,t_1)R_\xi(t_2,t_2)}$, это свойство следует из неравенства Коши–Буняковского для математических ожиданий $M^2(uv) \le Mu^2\,Mv^2$. Обозначив в нем $u = \xi(t_1), v = \xi(t_2)$, имеем

$$M^2\left[\xi(t_1)\xi(t_2)\right] = M\,\xi^2(t_1)\;M\,\xi^2(t_2).$$

3. Корреляционная функция случайного процесса является положительно определенной, т.е.

$\forall\, n$ и действительных $\lambda_1,\lambda_2,\ldots,\lambda_n$

$$\sum_{i,j=1}^{n} R_\xi(t_i, t_j)\lambda_i\lambda_j \geq 0.$$

Это следует из того, что

$$\sum_{i,j=1}^{n} R_\xi(t_i, t_j)\lambda_i\lambda_j = \sum_{i,j=1}^{n} M[\xi(t_i)\xi(t_j)]\lambda_i\lambda_j = M\left[\sum_{i,j=1}^{n}\xi(t_i)\xi(t_j)\lambda_i\lambda_j\right] =$$

$$= M\left[\sum_{i=1}^{n}\xi(t_i)\lambda_i \sum_{j=1}^{n}\xi(t_j)\lambda_j\right] = M\left[\sum_{i=1}^{n}\xi(t_i)\lambda_i\right]^2 \geq 0\,.$$

4. Пусть $\xi(t) = \sum_{i=1}^{n}\xi_i(t)$. Тогда

$$R_\xi(t_1, t_2) = \sum_{i=1}^{n} R_{\xi_i}(t_1, t_2) + \sum_{i \neq j}^{n} R_{\xi_i\xi_j}(t_1, t_2), \tag{2.23}$$

т.е., корреляционная функция суммы случайных процессов равна сумме корреляционных функций плюс сумма всех взаимных корреляционных функций. Очевидно, что для некоррелированных случайных процессов

$$R_\xi(t_1, t_2) = \sum_{i=1}^{n} R_{\xi_i}(t_1, t_2)\,. \tag{2.24}$$

ГЛАВА 3 ПРОЦЕССЫ С КОНЕЧНЫМИ МОМЕНТАМИ ВТОРОГО ПОРЯДКА. КОРРЕЛЯЦИОННАЯ ТЕОРИЯ

§13. СХОДИМОСТЬ В СРЕДНЕМ КВАДРАТИЧНОМ ДЛЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

Понятие сходимости, а следовательно, непрерывности, дифференцируемости и интегрируемости для случайных процессов в определенной мере отличается от соответствующих понятий в математическом анализе. Наиболее удобно и просто эти понятия вводятся для процессов, удовлетворяющих условию $M|\xi(t)|^2 < +\infty$, которые называются процессами с конечными моментами второго порядка, или гильбертовыми процессами. Физически это означает, что процесс имеет конечную мощность, что выполнимо для большинства реальных процессов.

Сформулируем основные типы сходимости для случайных процессов [10]. Пусть случайный процесс $\xi_h(t)$ зависит от некоторого параметра h.

Определение. Последовательность $\xi_h(t)$ сходится к $\xi(t)$ по вероятности при $h \to h_0$, если

$$\lim_{h \to h_0} P\{|\xi_h(t) - \xi(t)| > \varepsilon\} = 0,$$

что обозначается аналогично, как для сходимости случайных последовательностей, то есть

$$\xi_h(t) \xrightarrow[h \to h_0]{P} \xi(t).$$

Определение. Последовательность $\xi_h(t)$ сходится к $\xi(t)$ в среднем квадратичном при $h \to h_0$, если

$$\lim_{h\to h_0} M\left\{\left|\xi_h(t)-\xi(t)\right|^2\right\}=0,$$
что также обозначим аналогично, как для такого же типа сходимости случайных последовательностей, то есть

$$\xi_h(t)\xrightarrow[h\to h_0]{ср.кв.}\xi(t) \text{ или } \underset{h\to h_0}{l.i.m.}\,\xi_h(t)=\xi(t) .$$

Напомним, что сходимость в среднем квадратичном – более сильная сходимость, чем по вероятности. Именно такой тип сходимости будем использовать при определении непрерывности, дифференцируемости и интегрируемости случайных процессов в данной главе.

Лемма. *Если* $\xi_h(t)\xrightarrow[h\to h_0]{ср.кв}\xi(t)$ *и* $\eta_h(t)\xrightarrow[h\to h_0]{ср.кв}\eta(t)$, *то*

$$M\{\xi_h(t_1)\eta_l(t_2)\}\xrightarrow[h,l\to h_0]{}M\{\xi(t_1)\eta(t_2)\}. \qquad (3.1)$$

Доказательство. Рассмотрим математическое ожидание разности

$$M\{\xi_h(t_1)\eta_l(t_2)-\xi(t_1)\eta(t_2)\}=M\{(\xi_h(t_1)-\xi(t_1)(\eta_l(t_2)-\eta(t_2))\}+$$
$$+M\{(\xi_h(t_1)-\xi(t_1))\eta(t_2)\}+M\{(\eta_l(t_2)-\eta(t_2))\xi(t_1)\}.$$

Далее во всех слагаемых воспользуемся неравенством Коши–Буняковского:

$$M^2\{(\xi_h(t_1)-\xi(t_1))(\eta_l(t_2)-\eta(t_2))\}\le$$
$$\le M\left\{(\xi_h(t_1)-\xi(t_1))^2\right\}M\left\{(\eta_l(t_2)-\eta(t_2))^2\right\},$$

$$M^2\{(\xi_h(t_1)-\xi(t_1))\eta(t_2)\}\le M\left\{(\xi_h(t_1)-\xi(t_1))^2\right\}M\eta^2(t_2),$$

$$M^2\{(\eta_l(t_2)-\eta(t_2))\xi(t_1)\}\le M\left\{(\eta_l(t_2)-\eta(t_2))^2\right\}M\xi^2(t_1).$$

Так как мы рассматриваем процессы с ограниченными моментами второго порядка, то

$M\xi^2(t_1)<+\infty$, $M\eta^2(t_2)<+\infty$. По условию леммы и определению сходимости в среднем квадратичном, остальные сомножители в правой части стремятся к нулю при $h,l\to h_0$. Поэтому

$$M\{\xi_h(t_1)\eta_l(t_2)-\xi(t_1)\eta(t_2)\}\xrightarrow[h,l\to h_0]{}0\cdot$$

На основании этой леммы можно доказать критерий сходимости в среднем квадратичном.

Теорема 3.1. *Для того чтобы* $\xi_h(t)\xrightarrow[h\to h_0]{ср.кв.}\xi(t)$*, необходимо и достаточно, чтобы существовал предел*

$$\lim_{h,l\to h_0} M\{\xi_h(t)\xi_l(t)\}=A<+\infty\cdot \tag{3.2}$$

Доказательство теоремы во многом аналогично доказательству теоремы 1.13, хотя оно было приведено сокращенно.

Необходимость. Пусть $\xi_h(t)\xrightarrow[h\to h_0]{ср.кв.}\xi(t)$. В качестве $\eta_l(t)$ возьмем $\xi_l(t)$, тогда $\eta_h(t)\xrightarrow[h\to h_0]{ср.кв.}\eta(t)$. Тогда, по только что доказанной лемме,

$$M\{\xi_h(t)\eta_l(t)\}=M\{\xi_h(t)\xi_l(t)\}\xrightarrow[h,l\to h_0]{}M\xi^2(t)<+\infty\ \cdot$$

Достаточность. Пусть существует конечный предел

$$\lim_{h,l\to h_0} M\{\xi_h(t)\xi_l(t)\}=A\cdot$$

Тогда

$$M\{(\xi_h(t)-\xi_l(t))^2\}=M\xi_h^2(t)+M\xi_l^2(t)-2M\{\xi_h(t)\xi_l(t)\}\xrightarrow[h,l\to h_0]{}A+A-2A=0, \tag{3.3}$$

то есть последовательность $\xi_h(t)$ сходится. Тогда по лемме Больцано–Вейерштрасса из нее можно выделить подпоследовательность, сходящуюся к $\xi(t)$:

$$\xi_{h_1}(t),\xi_{h_2}(t),\ldots,\xi_{h_n}(t)\xrightarrow[h_i\to h_0]{ср.кв.}\xi(t)\cdot \tag{3.4}$$

Осталось показать, что наша подпоследовательность $\xi_h(t)$ сходится к этому же пределу $\xi(t)$. Используя неравенство $(a+b)^2 \le 2a^2 + 2b^2$, имеем

$$M\left\{(\xi_h(t) - \xi(t))^2\right\} = M\left\{\left[(\xi_h(t) - \xi_{h_i}(t)) + (\xi_{h_i}(t) - \xi(t))\right]^2\right\} \le$$

$$\le 2M\left\{(\xi_h(t) - \xi_{h_i}(t))^2\right\} + 2M\left\{(\xi_{h_i}(t) - \xi(t))^2\right\}.$$

Тогда с учетом (3.3), (3.4) получаем, что

$$M\left\{(\xi_h(t) - \xi(t))^2\right\} \xrightarrow[h \to h_0]{} 0,$$

что и доказывает теорему.

Следствие. *Для того чтобы* $\underset{h \to h_0}{l.i.m.}\,\xi_h(t) = \eta$*, где* η *– СВ с конечным вторым моментом, необходимо и достаточно, чтобы*

$$\lim_{h,l \to h_0} M\{\xi_h(t)\xi_l(t)\} = M\eta^2. \qquad (3.5)$$

Доказательство следует из предыдущей леммы и теоремы.

§12. НЕПРЕРЫВНОСТЬ СЛУЧАЙНЫХ ПРОЦЕССОВ

Непрерывность случайных процессов определяется по-разному, в зависимости от выбранного типа сходимости. Сходимость по вероятности определяет стохастическую непрерывность процесса. Это самый слабый вид непрерывности.

Определение. Случайный процесс $\xi(t)$ называется стохастически непрерывным в точке t_0, если

$$\lim_{h,l \to h_0} P\{|\xi(t) - \xi(t_0)| > \varepsilon\} = 0.$$

Стохастическая непрерывность процесса относится свойствам, однозначно определяемым ее двумерным распределением.

Определение. Случайный процесс $\xi(t)$ называется непрерывным в среднем квадратичном в точке t_0, если

$$\xi_h(t) \xrightarrow[t \to t_0]{ср.кв.} \xi(t_0) \cdot$$

Из непрерывности в среднем квадратичном в точке t следует стохастическая непрерывность $\xi(t)$ в той же точке. Это следует из неравенства Чебышева:

$$P\{|\xi(t+\Delta t) - \xi(t)| > \varepsilon\} \leq \frac{M\{(\xi(t+\Delta t) - \xi(t))^2\}}{\varepsilon^2}.$$

Обратное утверждение в общем случае неверно.

Пример 3.1. Случайный процесс $\xi(t)$ называется процессом Коши, если его приращения $\xi(t+\Delta t) - \xi(t)$ независимы и подчиняются распределению Коши, т.е. имеют плотность распределения

$$p(x) = \frac{(\Delta t)^2}{\pi((\Delta t)^2 + x^2)} \cdot$$

Этот процесс стохастически непрерывен, но не является непрерывным в среднем квадратичном, так как для распределения Коши

$$M\{(\xi(t+\Delta t) - \xi(t))^2\} = \infty \cdot$$

Стохастическая непрерывность и непрерывность в среднем квадратичном не означает, что реализации случайного процесса непрерывны с вероятностью 1. Примером является процесс Пуассона, который принимает целочисленные значения и является непрерывным в среднем квадратичном.

Теорема 3.2 ***(критерий непрерывности случайного процесса в среднем квадратичном).*** *Для того чтобы процесс $\xi(t)$ был непрерывным в среднем квадратичном в точке t_0, необходимо и достаточно, чтобы его корреляционная функция $R(t_1, t_2)$ была непрерывной в точке $t_0 = t_1 = t_2$.*

Доказательство. Необходимость. Пусть случайный процесс $\xi(t)$ непрерывен в среднем квадратичном в точке t_0, т.е.

$$M(\xi(t)-\xi(t_0))^2 \xrightarrow[t\to t_0]{} 0 \cdot \qquad (3.6)$$

Рассмотрим модуль разности, применяя для него неравенство Коши–Буняковского,

$$\left|R(t_0+h,t_0+l)-R(t_0,t_0)\right| = \left|M\left[\xi(t_0+h)\xi(t_0+l)\right]-M\xi^2(t_0)\right| =$$

$$=\left|M\left[(\xi(t_0+h)-\xi(t_0))\xi(t_0+l)\right]+M\left[(\xi(t_0+l)-\xi(t_0))\xi(t_0)\right]\right| \le$$

$$\le \sqrt{M(\xi(t_0+h)-\xi(t_0))^2}\sqrt{M\xi^2(t_0+l)}+$$

$$+\sqrt{M(\xi(t_0+l)-\xi(t_0))^2}\sqrt{M\xi^2(t_0)}.$$

Используя (3.6), получим, что при $h, l \to 0$ правая часть неравенства стремится к нулю, а это означает, что функция $R(t_1,t_2)$ непрерывна в точке $t_0=t_1=t_2$.

Достаточность. Имеем:

$$M(\xi(t)-\xi(t_0))^2 = M\xi^2(t)-2M\left[\xi(t)\xi(t_0)\right]+$$

$$+M\xi^2(t_0)=R(t,t)-2R(t,t_0)+R(t_0,t_0)=$$

$$=\left[R(t,t)-R(t,t_0)\right]+\left[R(t_0,t_0)-R(t,t_0)\right].$$

Поскольку ковариационная функция $R(t_1,t_2)$ непрерывна в точке $t_0=t_1=t_2$, то обе последние разности стремятся к нулю при $t\to t_0$, а это означает, что соотношение (3.6) выполняется.

Приведем сокращенный вариант другого доказательства данной теоремы, используя критерий сходимости в среднем квадратичном (теорема 3.1). Пусть $\xi_h(t_0)=\xi(t_0+h)$. Тогда, используя этот критерий, имеем:

$$M\left[\xi(t_0+h)\xi(t_0+l)\right]=R(t_0+h,\ t_0+l)\xrightarrow[h,l\to 0]{}R(t_0,t_0), \qquad (3.7)$$

а это в свою очередь есть не что иное как условие непрерывности корреляционной функции $R(t_1,t_2)$ в точке $t_0=t_1=t_2$.

Теорема 3.2. *Если* $R(t,t)$ *непрерывна* $\forall\ t\in T$, *то* $R(t_1,t_2)$ *непрерывна* $\forall\ t_1,t_2\in T$, *т.е. из непрерывности корреляционной функции при совпадающих аргументах следует ее непрерывность при несовпадающих аргументах.*

Доказательство. Так как $R(t_1,t_1)$ и $R(t_2,t_2)$ непрерывны, то из предыдущей теоремы следует, что

$$\xi(t_1+h)\xrightarrow[h\to 0]{ср.кв.}\xi(t_1),\ \xi(t_2+h)\xrightarrow[h\to 0]{ср.кв.}\xi(t_2).$$

Тогда из леммы имеем

$$M\left[\xi(t_1+h)\xi(t_2+l)\right]=R(t_1+h,t_2+l)\xrightarrow[h,l\to 0]{}M\left[\xi(t_1)\xi(t_2)\right]=R(t_1,t_2),$$

§15. ДИФФЕРЕНЦИРУЕМОСТЬ СЛУЧАЙНЫХ ПРОЦЕССОВ

Определение. Случайный процесс $\xi(t)$ называется дифференцируемым в среднем квадратичном точке t_0, если существует предел

$$\underset{h\to 0}{l.i.m.}\frac{\xi(t_0+h)-\xi(t_0)}{h}=\xi'(t_0). \qquad (3.8)$$

Случайная функция $\xi'(t_0)$ называется среднеквадратичной производной процесса в точке t_0.

Теорема 3.4 ***(критерий дифференцируемости случайного процесса в среднем квадратичном).*** *Для того чтобы процесс* $\xi(t)$ *имел производную в среднем квадратичном в точке* t_0, *необходимо и достаточно, чтобы существовала смешанная частная производная второго порядка корреляционной функции* $\frac{\partial^2 R(t_1,t_2)}{\partial t_1 \partial t_2}$ *в точке* $t_0=t_1=t_2$.

Доказательство. Обозначим

$$\xi_h(t_0)=\frac{\xi(t_0+h)-\xi(t_0)}{h},\ \xi_l(t_0)=\frac{\xi(t_0+l)-\xi(t_0)}{l}.$$

Тогда по критерию сходимости в среднем квадратичном (3.2) для существования производной $\xi'(t_0)$ необходимо и достаточно существование предела

$$\lim_{h,l\to 0} M\left\{\frac{\xi(t_0+h)-\xi(t_0)}{h}\cdot\frac{\xi(t_0+l)-\xi(t_0)}{l}\right\}.$$

Имеем:

$$M\left\{\frac{\xi(t_0+h)-\xi(t_0)}{h}\cdot\frac{\xi(t_0+l)-\xi(t_0)}{l}\right\}=$$

$$=\frac{1}{hl}\left[R(t_0+h,t_0+l)-R(t_0,t_0+l)-R(t_0+h,t_0)+R(t_0,t_0)\right]=$$

$$=\frac{1}{l}\left[\frac{R(t_0+h,t_0+l)-R(t_0,t_0+l)}{h}\right]-\frac{1}{l}\left[\frac{R(t_0+h,t_0)-R(t_0,t_0)}{h}\right]\xrightarrow[h\to 0]{}$$

$$\xrightarrow[h\to 0]{}\frac{1}{l}\frac{\partial R(t_1,t_0+l)}{\partial t_1}\bigg|_{t_1=t_0}-\frac{1}{l}\frac{\partial R(t_1,t_0)}{\partial t_1}\bigg|_{t_1=t_0}\xrightarrow[l\to 0]{}\frac{\partial^2 R(t_1,t_2)}{\partial t_1\partial t_2}\bigg|_{t_0=t_1=t_2},$$

а последняя частная производная, по условию теоремы, существует. Тогда, используя критерий (3.2), получаем, что существует предел (3.8).

Теорема 3.5. *Если для всех* $t\in T$ *существует вторая смешанная производная от корреляционной функции при совпадающих аргументах* $\frac{\partial^2 R(t_1,t_2)}{\partial t_1\partial t_2}\bigg|_{t_1=t_2}$*, то* $\forall\ t_1,t_2\in T$ *существуют первые и вторая смешанная производные при несовпадающих аргументах*

$$\frac{\partial R(t_1,t_2)}{\partial t_1},\ \frac{\partial R(t_1,t_2)}{\partial t_2},\ \frac{\partial^2 R(t_1,t_2)}{\partial t_1\partial t_2}.$$

Доказательство. Действительно, исходя из предыдущей теоремы, существование $\left.\dfrac{\partial^2 R(t_1,t_2)}{\partial t_1 \partial t_2}\right|_{t_1=t_2}$ означает существование $\xi'(t_1)$. Обозначив

$$\xi_h(t_1) = \frac{\xi(t_1+h)-\xi(t_1)}{h}, \quad \eta_h(t_2) = \xi(t_2),$$

из леммы получаем, что существует предел

$$\lim_{h\to 0} M\left\{\frac{\xi(t_1+h)-\xi(t_1)}{h}\cdot \xi(t_2)\right\} = \\ = \lim_{h\to 0}\frac{R(t_1+h,t_2)-R(t_1,t_2)}{h} = \frac{\partial R(t_1,t_2)}{\partial t_1}.$$

Аналогично доказывается существование остальных производных.

Пример 3.2. Покажем, что случайный процесс с корреляционной функцией $R(t_1,t_2)=\min(t_1,t_2)$ не дифференцируем в среднем квадратичном.

Пусть $0 < h < \ell$. Легко проверить, что функция $R(t_1,t_2)$ обладает своим свойством корреляционной функции. Имеем (см. доказательство теоремы 3.4):

$$\frac{\partial^2 R(t,t)}{\partial t^2} = \lim_{h\to 0}\left[R(t+h,t+h)-R(t,t+h)-R(t+h,t)+R(t,t)\right] =$$

$$= \lim_{h\to 0}\frac{1}{h^2}\left[\min(t+h,t+h)-\min(t,t+h)-\min(t+h,t)+\min(t,t)\right] =$$

$$= \lim_{h\to 0}\frac{1}{h^2}\left[t+h-t-t+t\right] = \lim_{h\to 0}\frac{1}{h} = \infty,$$

поэтому на основании критерия из теоремы 3.4 получаем, что процесс с такой корреляционной функцией не дифференцируем в среднем квадратичном.

§16. ИНТЕГРИРОВАНИЕ СЛУЧАЙНЫХ ПРОЦЕССОВ

Пусть случайный процесс $\xi(t)$ и не случайная функция $g(t)$ заданы на интервале $[a,b]$. Разобьем этот интервал точками $a = t_1 < t_2 < ... < t_n = b$

Определение. Случайный процесс $\xi(t)$ интегрируем в среднем квадратичном на отрезке $[a,b]$, если существует конечный предел в среднем интегральных сумм:

$$\underset{\substack{\max(t_{i+1}-t_i)\to 0 \\ \bar{t}_i\in[t_i,\, t_{i+1})}}{l.i.m.} \sum_{i=1}^{n-1} g(\bar{t}_i)\xi(\bar{t}_i)(t_{i+1}-t_i) = \int_b^a g(t)\xi(t)dt \cdot \qquad (3.9)$$

Теорема 3.6 ***(критерий интегрируемости случайного процесса в среднем квадратичном)***. *Для того чтобы процесс* $\xi(t)$ *был интегрируем в среднем квадратичном, необходимо и достаточно, чтобы существовал двойной интеграл от его корреляционной функции*

$$\int_a^b \int_a^b g(t)g(n)R(t,u)dtdu < +\infty \cdot \qquad (3.10)$$

Доказательство. По критерию сходимости (3.2), для существования интеграла (3.9) необходимо и достаточно существование предела

$$\lim_{\substack{\max\limits_i(t_{i+1}-t_i)\to 0, \\ \max\limits_j(u_{j+1}-u_j)\to 0, \\ \bar{t}_i\in[t_i,t_{i+1}), \\ \bar{u}_j\in[u_j,u_{j+1})}} M\left\{\sum_i g(\bar{t}_i)\xi(\bar{t}_i)(t_{i+1}-t_i)\sum_j g(\bar{u}_j)\xi(\bar{u}_j)(t_{j+1}-t_j)\right\} = A.$$

Но поскольку

$$M\left\{\sum_i g(\bar{t}_i)\xi(\bar{t}_i)(t_{i+1}-t_i)\sum_j g(\bar{u}_j)\xi(\bar{u}_j)(u_{j+1}-u_j)\right\}=$$
$$=M\sum_i g(\bar{t}_i)g(\bar{u}_j)\xi(\bar{t}_i)\xi(\bar{u}_j)(t_{i+1}-t_i)(u_{j+1}-u_j)=$$
$$\sum_j g(\bar{t}_i)g(\bar{u}_j)R(\bar{t}_i,\bar{u}_j)(t_{i+1}-t_i)(u_{j+1}-u_j)\xrightarrow[\substack{\max\limits_i(t_{i+1}-t_i)\to 0\\ \max\limits_j(u_{j+1}-u_j)\to 0}]{}$$
$$\xrightarrow[\substack{\max\limits_i(t_{i+1}-t_i)\to 0\\ \max\limits_j(u_{j+1}-u_j)\to 0}]{}\int_a^b\int_a^b g(t)g(u)R(t,u)dtdu\ ,$$

то как раз в качестве A можно взять двойной интеграл (3.10).

Под несобственными интегралами в среднем квадратичном $\int_{-\infty}^{\infty} g(t)\xi(t)dt$ или $\int_a^{\infty} g(t)\xi(t)dt$ будем понимать соответственно пределы $\lim\limits_{\substack{a\to-\infty\\ b\to+\infty}}\int_a^b g(t)\xi(t)dt$, $\lim\limits_{b\to+\infty}\int_a^b g(t)\xi(t)dt$. Для того чтобы они существовали, необходимо и достаточно существование интегралов

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(t)g(u)R(t,u)dtdu \text{ и } \int_a^{\infty}\int_a^{\infty} g(t)g(u)R(t,u)dtdu\,.$$

Пример 3.3. Вернемся к случайному процессу, рассмотренному в примере 2.6. Поскольку среднее значение процесса равно нулю, то его корреляционная и ковариационные функции совпадают и, на основании (2.19),

$$R(t_1,t_2)=\sigma^2\cos[\lambda(t_2-t_1)]\cdot$$

Тогда из теорем 3.3, 3.5, 3.6 следует, что такой процесс непрерывен и дифференцируем в среднем квадратичном $\forall\ t\in[0,+\infty)$.

Покажем, что его среднеквадратичная производная является случайным процессом $\xi'(t,w)$, определяемым соотношением

$$\xi'(t, w) = -\lambda\alpha(w)\sin\lambda t + \lambda\beta(w)\cos\lambda t \cdot \qquad (3.11)$$

Для этого воспользуемся определением (3.8) и определением сходимости в среднем квадратичном:

$$\lim_{h\to 0} M\left\{\left|\frac{\xi(t+h)-\xi(t)}{h} - \xi'(t)\right|^2\right\} =$$

$$= \lim_{h\to 0} M\left\{\left|\alpha(w)\frac{\cos\lambda(t+h) - \cos\lambda t)}{h} + \beta(w)\frac{\sin\lambda(t+h) - \sin\lambda t}{h} + \right.\right.$$

$$\left.\left. - + \lambda\alpha(w)\sin\lambda t - \lambda\beta(w)\cos\lambda t\right|^2\right\} =$$

$$= \lim_{h\to 0} M\left\{\left|\alpha(w)\left(\frac{\cos\lambda(t+h) - \cos\lambda t}{h} + \lambda\sin\lambda t\right) + \right.\right.$$

$$\left.\left. + \beta(w)\left(\frac{\sin\lambda(t+h) - \sin\lambda t}{h} - \lambda\cos\lambda t\right)\right|^2\right\} =$$

$$= M\left\{\left|\alpha(w)\left((\cos\lambda t)' + \lambda\sin\lambda t\right) + \beta(w)\left((\sin\lambda t)' - \lambda\cos\lambda t\right)\right|^2\right\} = 0,$$

поэтому формула (3.11) справедлива.

Покажем теперь, что

$$\int_a^b \xi(t)dt = \frac{1}{\lambda}\alpha(w)(\sin\lambda b - \sin\lambda a) - \frac{1}{\lambda}\beta(w)(\cos\lambda b - \cos\lambda a). \qquad (3.12)$$

Для этого также воспользуемся определением сходимости в среднем квадратичном и определением интеграла (3.9). Поскольку

$$\lim_{n\to\infty} M\left\{\left|\sum_{i=1}^{n-1}(\alpha(w)\cos\lambda\bar{t}_i + \beta(w)\sin\lambda\bar{t}_i)(t_{i+1} - t_i) - \right.\right.$$

$$\left.\left. - \frac{1}{\lambda}\alpha(w)(\sin\lambda b - \sin\lambda a) + \frac{1}{\lambda}\beta(w)(\cos\lambda b - \cos\lambda a)\right|^2\right\} =$$

$$=\lim_{n\to\infty} M\left\{\left|\alpha(w)\sum_{i=1}^{n-1}\left(\cos\lambda\bar{t}_i\right)\left(t_{i+1}-t_i\right)-\frac{1}{\lambda}\left(\sin\lambda b-\sin\lambda a\right)\right|^2-\right.$$

$$\left.-\left|\beta(w)\sum_{i=1}^{n-1}\left(\sin\lambda\bar{t}_i\right)\left(t_{i+1}-t_i\right)+\frac{1}{\lambda}\left(\cos\lambda b-\cos\lambda a\right)\right|^2\right\}=0,$$

то соотношение (3.12) имеет место.

§17. СТОХАСТИЧЕСКИЕ ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

Под стохастическим обыкновенным дифференциальным уравнением обычно понимается обыкновенное дифференциальное уравнение относительно среднеквадратичной производной случайного процесса.

Рассмотрим линейное дифференциальное уравнение первого порядка

$$\xi'(t)+a(t)\xi(t)=\eta(t), \tag{3.13}$$

удовлетворяющее начальному условию $\xi(0)=x_0$, где $a(t)$ действительная функция, $n(t), t\geq 0$, – известный случайный процесс, x_0 – заданное число. Кроме моментных функций, определяемых соотношениями (2.12), (2.14), рассмотрим взаимную функцию корреляции процессов $\xi(t)$ и $\eta(t)$ $R_{\xi\eta}(t_1,t_2)$, имеющую вид (2.16). Определяя среднее значение от обеих сторон уравнения (3.13), получим

$$m'_\xi(t)+a(t)m_\xi(t)=m_\eta(t),\;\; m(t_0)=x_0.$$

Найдем теперь дифференциальное уравнение для корреляционной функции $R_\xi(t_1,t_2)$. Сначала умножим обе стороны уравнения (3.13) при $t=t_1$, на $\eta(t_2)$ и определим средние значения для обеих сторон. Тогда будем иметь

$$\frac{\partial}{\partial t_1}R_{\xi\eta}(t_1,t_2)+a(t_1)R_{\xi\eta}(t_1,t_2)=R_\eta(t_1,t_2)\cdot \tag{3.14}$$

Можно найти решение этого уравнения при начальном условии $R_{\xi\eta}(0, t_2) = x_0 m_\eta(t_2)$. Далее, умножив обе стороны уравнения (3.13) при $t = t_2$ на $\xi(t_1)$ и определив средние значения, получим

$$\frac{\partial}{\partial t_2} R_\xi(t_1, t_2) + a(t_2) R_\xi(t_1, t_2) = R_{\xi\eta}(t_1, t_2), \qquad (3.15)$$

для которого можно также найти решение при начальном условии $R_\xi(t_1, 0) = x_0 m_\xi(t_1)$.

Рассмотренный метод можно также применить в случае, когда в качестве x_0 берется некоторая СВ.

Пример 3.4. Найдем среднее значение, корреляционную функцию и дисперсию процесса $\xi(t)$, удовлетворяющего дифференциальному уравнению

$$\xi'(t) + \xi(t) = \eta(t) \qquad (3.16)$$

с начальным условием $\xi(0) = 0$, где $\eta(t)$ – случайный процесс, определенный в примере 2.6, причем $\sigma^2 = 1$.

Естественно, что

$$m'_\xi(t) + m_\xi(t) = 0,$$

и поэтому из начального условия для уравнения (3.16) вытекает условие $m_\xi(0) = 0$, откуда следует, что $m_\xi(t) = 0$.

Из уравнения (3.14) и (2.19) получаем

$$\frac{\partial}{\partial t_1} R_{\xi\eta}(t_1, t_2) + R_{\xi\eta}(t_1, t_2) = \cos\lambda(t_2 - t_1)$$

и начальное условие $R_{\xi\eta}(0, t_2) = 0$. Решением этого уравнения является

$$R_{\xi\eta}(t_1, t_2) = \frac{1}{2\lambda}\left[\cos\lambda(t_2 - t_1) - \sin\lambda(t_2 - t_1)\right] +$$

$$+\frac{e^{-\lambda t_1}}{2\lambda}\left[\sin\lambda t_2-\cos\lambda t_2\right]=f(t_1,t_2)$$

Тогда из (3.15) будем иметь уравнение

$$\frac{\partial}{\partial t_2}R_\xi(t_1,t_2)+R_\xi(t_1,t_2)=f(t_1,t_2)$$

с начальным условием $R_\xi(t_1,0)=0$. Его решением является функция

$$R_\xi(t_1,t_2)=\frac{1}{2\lambda^2}\left[\cos\lambda(t_2-t_1)+e^{-\lambda(t_1+t_2)}-\cos\lambda t_1 e^{-\lambda t_2}-\cos\lambda t_2 e^{-\lambda t_1}\right].$$

Поскольку $m_\xi(t)=0$, то $K_\xi(t_1,t_2)=R_\xi(t_1,t_2)$. Поэтому для дисперсии процесса $\xi(t)$ получаем

$$D_\xi(t)=\frac{1}{2\lambda^2}\left[\sin^2\lambda t+\left(e^{-\lambda t}-\cos\lambda t\right)^2\right].$$

§18. РАЗЛОЖЕНИЕ СЛУЧАЙНЫХ ПРОЦЕССОВ ПО ОРТОГОНАЛЬНЫМ ФУНКЦИЯМ

Пусть случайный процесс $\xi(t)$ задан на интервале $[0,\mathrm{T}]$ с корреляционной функцией $R(t_1,t_2)$ и нулевым математическим ожиданием, $m_\xi(t)=0$. Рассмотрим интегральное уравнение

$$\int_0^{\mathrm{T}} R(t,u)\varphi(u)du=\lambda\varphi(t), \tag{3.17}$$

которое называется однородным интегральным уравнением Фредгольма II рода. Из теории интегральных уравнений известно, что его решение существует при счетном наборе постоянных $\lambda_1\ge\lambda_2\ge\ldots\ge\lambda_n$, называемых собственными числами уравнения Фредгольма. Решение $\varphi_i(t)$, соответствующее λ_i, называется собственной функцией ядра $R(t,u)$. Собственные функции могут быть сделаны ортогональными, то есть

$$\int_0^T \varphi_i(t)\varphi_j(t)dt = \delta_{ij}, \tag{3.18}$$

где δ_{ij} – символ Кронекера. Так как ядро уравнения (3.17) является корреляционной функцией $R(t,u)$, а значит, положительно определенной, то все $\lambda_i > 0$, и по теореме Мерсера [10] ядро может быть представлено в виде ряда по собственным функциям

$$R(t,u) = \sum_{i=1}^{\infty} \lambda_i \varphi_i(t)\varphi_i(u)\cdot \tag{3.19}$$

Теорема 3.7. *Случайный процесс* $\xi(t)$ *можно представить в виде ряда*

$$\xi(t) = \underset{n\to\infty}{l.i.m.} \sum_{k=1}^{n} \xi_k(w)\varphi_k(t), \tag{3.20}$$

где $\xi_k(w) = \int_0^T \xi(t)\varphi_k(t)dt$ – некоррелированные СВ с $M\xi_k(w) = 0$, $D\xi_k(w) = \lambda_k$.

Доказательство. Рассмотрим свойства СВ $\xi_k(w), k = 1,2,\ldots$. Поскольку $M\xi(t) = 0$, то

$$M\xi_k(w) = M\left\{\int_0^T \xi(t)\varphi_k(t)dt\right\} = \int_0^T M\xi(t)\varphi_k(t)dt = 0\cdot$$

Найдем корреляционный момент $M\{\xi_k(w)\xi_l(w)\}$:

$$M\{\xi_k(w)\xi_l(w)\} = M\left\{\int_0^T \int_0^T \xi(t)\xi(u)\varphi_k(t)\varphi_l(u)dtdu\right\} =$$

$$=\int_0^T \varphi_k(t)dt\int_0^T R(t,u)\varphi_l(u)du$$

С учетом (3.17) и (3.18) имеем

$$M\{\xi_k(w)\xi_l(w)\}=\lambda_l\int_0^T \varphi_k(t)\varphi_l(t)dt=\lambda_l\delta_{kl}\cdot$$

Отсюда видно, что СВ $\xi_k(w), k=1, 2,...$, некоррелированы и $D\xi_k(w)=\lambda_k$. Найдем также вспомогательное выражение

$$M\{\xi_k(w)\xi(t)\}=M\left\{\xi(t)\int_0^T \xi(u)\varphi_k(u)du\right\}=\int_0^T R(t,u)\varphi_k(u)du=\lambda_k\varphi_k(t)\cdot$$

Докажем теперь основное утверждение теоремы. Рассмотрим предел

$$\lim_{n\to\infty} M\left[\xi(t)-\sum_{k=1}^{n}\xi_k(w)\varphi_k(t)\right]^2=$$

$$=\lim_{n\to\infty}\left\{R(t,t)-2\sum_{k=1}^{n}\{\xi(t)\xi_k(w)\}\varphi_k(t)+\sum_{k=1}^{n}\lambda_k\varphi_k^2(t)\right\}=$$

$$=\lim_{n\to\infty}\left\{R(t,t)-\sum_{k=1}^{n}\lambda_k\varphi_k^2(t)\right\}=0,$$

согласно (3.19). Используя определение сходимости в среднем квадратичном, получаем представление (3.20).

Разложение (3.20) является единственным разложением случайного процесса в ряд по ортогональным функциям с некоррелированными коэффициентами [10]. Основное достоинство этого разложения заключается в том, что оно позволяет вместо случайного процесса $\xi(t)$ рассматривать ряд (3.20) с некоррелированными СВ $\xi_1(w),\xi_2(w),...,\xi_k(w),...\cdot$

ГЛАВА 4 ПРОЦЕССЫ С НЕЗАВИСИМЫМИ ПРИРАЩЕНИЯМИ. ГАУССОВСКИЙ И ВИНЕРОВСКИЙ СЛУЧАЙНЫЕ ПРОЦЕССЫ

§19. ПРОЦЕССЫ С НЕЗАВИСИМЫМИ ПРИРАЩЕНИЯМИ

Рассмотренное в §4 понятие независимости СВ играет важную роль в теории случайных процессов. На основе этого понятия был введен класс процессов с независимыми приращениями.

Определение. Случайный процесс $\xi(t)$ называется процессом с независимыми приращениями, если для любого n и для любых моментов времени

$t_1 < t_2 < \ldots < t_n$ СВ $\xi(t_1)$, $\xi(t_2)-\xi(t_1),\ldots,\xi(t_n)-\xi(t_{n-1})$ независимы.

Определение. Процесс с независимыми приращениями $\xi(t)$ называется однородным, если $\xi(0)=0$ и для любых моментов времени $t_1 < t_2$ распределение СВ $\xi(t_2)-\xi(t_1)$ зависит только от разности $t_2 - t_1$ и не зависит от t_1.

Пример 4.1. Одним из важных примеров процесса с независимыми приращениями является пуассоновский процесс (см. пример 2.3). Введем понятие данного процесса несколько по другому.

Назовём пуассоновским процессом $v(t)$ процесс с независимыми приращениями, для которого при любых $t_1 < t_2$ СВ $v(t_2)-v(t_1)$ принимает только неотрицательные целые значения, $v(0)=0$, а также

$$P\{v(t_2)-v(t_1)=k\}=\frac{[\lambda(t_2-t_1)]^k}{k!}e^{-\lambda(t_2-t_1)}, k=0,1,2,\ldots \tag{4.1}$$

Очевидно, что пуассоновский процесс является однородным, так как распределение СВ $v(t_2)-v(t_1)$ зависит только от разности

$t_2 - t_1$. Вероятность того, что на интервале времени $[t_1, t_2)$ не наступит изменение состояния процесса (вероятность того, что реализация процесса на этом интервале является константой), равна

$$P_0(t_2 - t_1) = e^{-\lambda(t_2 - t_1)}.$$

Вероятность того, что значение процесса увеличится на 1 (произойдет единичный скачок) на интервале времени $[t_1, t_2)$, равна

$$P_1(t_2 - t_1) = \lambda(t_2 - t_1)e^{-\lambda(t_2 - t_1)},$$

и так далее. Реализации процесса являются ступенчатыми неубывающими функциями, поскольку с вероятностью 1 СВ $\xi(t_2) - \xi(t_1)$ принимает только неотрицательные значения, и вероятность того, что на интервале времени $[t_1, t_2)$ не произойдет изменение значения процесса положительна.

Остановимся ещё на одном свойстве пуассоновского процесса. Поскольку $v(0) = 0$, то очевидно, что $v(t)$ имеет такое же распределение, как $v(t) - v(0)$, т.е распределение пуассоновского процесса в момент времени t совпадает с распределением приращения $v(t) - v(0)$ в этот момент. Поэтому знание распределения вероятностей данной разности определяет распределение процесса. Если $v(t)$ означает число событий, наступивших на интервале времени $[0, t)$, то $v(t_2) - v(t_1)$ является числом событий, наступивших на интервале времени $[t_1, t_2)$. Обозначим через t', t'' две соседние точки скачков реализации процесса, т.е.

$$v(t_2) - v(t_1) = 0 \text{ для } t' < t_1 \le t_2 < t'',$$

$$\lim_{t \to t'-0} v(t) < \lim_{t \to t'+0} v(t), \lim_{t \to t''-0} v(t) < \lim_{t \to t''+0} v(t).$$

Рассмотрим ф.р. СВ $t'' - t'$:

$$P(t'' - t' \le t) = 1 - P(t'' - t' > t) = 1 - P(t'' > t' + t) =$$
$$= 1 - P(v(t' + t) - v(t') = 0) = 1 - e^{-\lambda t},$$

согласно (4.1). Отсюда следует, что время пребывания процесса в каждом из состояний имеет экспоненциальное распределение.

Найдем средние характеристики пуассоновского процесса. Из соотношения (4.1) следует (2.6), откуда (см. §6, пример) имеем

$$m_j(t) = \sum_{k=0}^{\infty} k \frac{(\lambda t)^k}{k!} e^{-\lambda t} = \lambda t. \tag{4.2}$$

Используя (2.6), (4.1), (4.2) можно найти ковариационную функцию

$$K_v(t_1, t_2) = M[v(t_1)v(t_2)] - m_v(t_1)m_v(t_2),$$

которая при $t_1 < t_2$ равна

$$K_v(t_1, t_2) = \sum_{\substack{k_1, k_2 = 0 \\ k_1 \le k_2}}^{\infty} k_1 k_2 P(v(t_1) = k_1, v(t_2) = k_2) - \lambda^2 t_1 t_2 =$$

$$= \sum_{\substack{k_1, k_2 = 0 \\ k_1 \le k}}^{\infty} \left[k_1(k_2 - k_1) + k_1^2\right] P(v(t_1) = k_1) P(v(t_2) - v(t_1) = k_2 - k_1) - \lambda^2 t_1 t_2 =$$

$$= \lambda t_1 \lambda(t_2 - t_1) + \lambda t_1 + \lambda^2 t_1^2 - \lambda^2 t_1 t_2 = \lambda t_1.$$

Отсюда следует, что дисперсия процесса имеет вид

$$D_v(t) = K_v(t, t) = \lambda t,$$

а корреляционная функция

$$R_v(t_1, t_2) = \lambda t_1(\lambda t_2 + 1).$$

Корреляционная функция является непрерывной и поэтому из теоремы 3.2 вытекает, что пуассоновский процесс является непрерывным в среднем квадратичном. Из (4.3) следует, что

$$\frac{\partial^2 R_v(t, t)}{\partial t^2} = \frac{\partial^2 K_v(t, t)}{\partial t^2} + \lambda^2,$$

поэтому $\dfrac{\partial^2 R_v(t, t)}{\partial t^2}$ существует тогда и только тогда, когда существует $\dfrac{\partial^2 K_v(t, t)}{\partial t^2}$.

Но (см. пример 3.2)

$$\frac{\partial^2 K_v(t,t)}{\partial t^2} = \lim_{h^2\to 0}\frac{1}{h^2}\left[K_v(t+h,t+h) - K_v(t,t+h) - K_v(t+h,t) + K_v(t,t)\right] =$$

$$= \lim_{h^2\to 0}\frac{1}{h^2}\left[K_v(t+h,t+h) - K_v(t,t+h) - K_v(t,t+h) + K_v(t,t)\right] =$$

$$= \lim_{h^2\to 0}\frac{1}{h^2}\left[\lambda(t+h) - \lambda t - \lambda t + \lambda t\right] = \lim_{h\to 0}\frac{\lambda}{h} = \infty,$$

следовательно, на основании теоремы 3.5, пуассоновский процесс не дифференцируем в среднем квадратичном.

§20. ОБОБЩЕННЫЙ ПУАССОНОВСКИЙ ПРОЦЕСС

Пусть $t_1, t_2, ..., t_n, ...$ – моменты изменения состояний пуассоновского процесса. Таким образом, $t_1, t_2 - t_1, ..., t_n - t_{n-1}, ...$ – независимые СВ с плотностью распределения $p(x) = \lambda e^{-\lambda x}, x \geq 0$. Пусть $\eta_1, \eta_2, ..., \eta_n, ...$ – СВ с общей ф.р. $H(x)$. Допустим, что СВ $(t_1, t_2, ..., t_n, ...; \eta_1, \eta_2, ..., \eta_n, ...)$ независимы в совокупности. Обозначим

$$\xi(t) = \sum_{t_n < t} \eta_n = \sum_{n=1}^{v(t)} \eta_n,$$

где $v(t)$ – пуассоновский процесс с параметром λ. Случайный процесс $\xi(t)$, определенный этим равенством, называется обобщенным пуассоновским процессом.

Основное свойство обобщенного пуассоновского процесса – независимость его приращений в непересекающихся интервалах времени: если $n \geq 2$ – любое число, $0 \leq \tau_0 < \tau_1 < ... < \tau_n$, то СВ

$$\zeta_1 = \xi(\tau_1) - \xi(\tau_0),\ ...,\ \zeta_n = \xi(\tau_n) - \xi(\tau_{n-1})$$

независимы в совокупности. Действительно,

$$\zeta_i = \sum_{\tau_{i-1} < t_n \leq \tau_n} \eta_n .$$

По формуле полной вероятности имеем

$$P\left(\zeta_i \le x_i,\, i=\overline{1,n}\right)=$$
$$=\sum_{k_0,k_1,\ldots,k_n} P\left(v_i = k_i,\, i=\overline{1,n}\right) P\left(\zeta_i \le x_i,\, i=\overline{1,n} \,\middle|\, v_i = k_i,\, i=\overline{1,n}\right),$$

где $v_i = v(\tau_i) - v(\tau_{i-1}), i \ge 1, v_0 = v(\tau_0)$.

На основании свойства пуассоновского процесса и формулы (4.1) можно записать

$$P\left(v_i = k_i,\, i=\overline{1,n}\right)=\prod_{i=0}^{n} \frac{\left[\lambda(\tau_i - \tau_{i-1})\right]^{k_i}}{k_i!} e^{-\lambda(\tau_i - \tau_{i-1})}, \qquad (4.6)$$

где предполагается, что $a_{-1} = 0$. Если теперь зафиксировать $v_0 = k_0, v_1 = k_1, \ldots, v_n = k_n$, то

$$\xi(\tau_0) = \eta_1 + \eta_2 + \ldots + \eta_{k_0},$$
$$\xi(\tau_1) = \eta_1 + \eta_2 + \ldots + \eta_{k_0+k_1},$$
$$\ldots$$
$$\xi(\tau_n) = \eta_1 + \eta_2 + \ldots + \eta_{k_0+k_1+\ldots+k_n}.$$

Отсюда

$$\begin{gathered} \zeta_1 = \eta_{k_0+1} + \eta_{k_0+2} + \ldots + \eta_{k_0+k_1}, \\ \zeta_2 = \eta_{k_0+k_1+1} + \eta_{k_0+k_1+2} + \ldots + \eta_{k_0+k_1+k_2}, \\ \ldots \\ \zeta_n = \eta_{k_0+\ldots+k_{n-1}+1} + \eta_{k_0+\ldots+k_{n-1}+2} + \ldots + \eta_{k_0+k_n}. \end{gathered} \qquad (4.7)$$

Таким образом, ξ_i при фиксированных $k_0, k_1, \ldots, k_n$ представляет собой суммы независимых СВ (разных для различных i).

Пусть $H^{(m)}(x) = P(\eta_1 + \eta_2 + \ldots + \eta_m \le x)$, тогда

$$\begin{gathered} P\left(\zeta_i \le x_i,\, i=\overline{1,n} \,\middle|\, v_0 = k_0, v_1 = k_1, \ldots, v_n = k_n\right)= \\ = H^{(k_1)}(x_1) H^{(k_2)}(x_2) \cdot \ldots \cdot H^{(k_n)}(x_n). \end{gathered} \qquad (4.8)$$

Подставив соотношения (4.8), (4.6) в формулу (4.5), найдем

$$P\left(\zeta_i \le x_i, i=\overline{1,n}\right)=\sum_{k_0=0}^{\infty}\frac{(\lambda\tau_0)^{k_0}}{k_0!}e^{-\lambda\tau_0}\prod_{i=1}^{n}\sum_{k_i=0}^{\infty}\frac{[\lambda(\tau_i-\tau_{i-1})]^{k_i}}{k_i!}e^{-\lambda(\tau_i-\tau_{i-1})}H^{(k_i)}(x_i)\cdot$$

Просуммировав по k_0 от 0 до ∞, приходим к формуле

$$P\left(\xi_i \le x_i, i=\overline{1,n}\right)=\prod_{i=1}^{n}\frac{[\lambda(\tau_i-\tau_{i-1})]^{k_i}}{k_i!}e^{-\lambda(\tau_i-\tau_{i-1})}H^{(k_i)}(x_i), \tag{4.9}$$

откуда следует, что ζ_i независимы в совокупности. Если $(\tau_i-\tau_{i-1})=\tau, i=\overline{1,n}$, то ζ_i одинаково распределены.

Положим $n=1, \tau_0=0, \tau_1=t, k_1=k$. Тогда из формулы (4.9) находим

$$F_{\xi(t)}(x)=P(\xi(t)\le x)=\sum_{k=0}^{\infty}\frac{(\lambda t)^k}{k!}e^{-\lambda t}H^{(k)}(x)\cdot \tag{4.10}$$

Пусть

$$\psi(\theta)=\int_{-\infty}^{\infty}e^{i\theta x}dH(x),\ \varphi_t(\theta)=\int_{-\infty}^{\infty}e^{i\theta x}dF_{\xi(t)}(x)\cdot$$

Тогда

$$\psi^k(\theta)=\int_{-\infty}^{\infty}e^{i\theta x}dH^{(k)}(x) \tag{4.11}$$

как характеристическая функция суммы независимых одинаково распределенных СВ. Умножая левую и правую части (4.10) на $e^{i\theta x}$, проинтегрировав от $-\infty$ до ∞ и учитывая равенство (4.11), получим

$$\varphi_t(\theta)=\sum_{k=0}^{\infty}\frac{[\lambda t\psi(\theta)]^k}{k!}e^{-\lambda\theta}=e^{-\lambda t[1-\psi(\theta)]}. \tag{4.12}$$

Пусть, например, интервалы между поступлениями запросов на центральный сервер компьютерной сети имеют экспоненциальное распределение с параметром λ. Длительности обработки запросов – независимые СВ с ф.р. $H(x)$. Тогда характеристическая функция суммарного времени обработки запросов, которые поступили на сервер за время t, определяется формулой (4.12). Найдем среднее значение этой СВ. Предположив, что $M\eta_1 < \infty$, имеем

$$M_{\zeta} = -i\varphi_t'(0) = -i\lambda t\psi'(0) = -i\lambda tiM\eta_1 = \lambda tM\eta_1 \cdot$$

§21. ГАУССОВСКИЙ СЛУЧАЙНЫЙ ПРОЦЕСС

Важную роль во многих прикладных задачах играют случайные процессы, конечномерное распределение которых является нормальным (гауссовским). Процессы такого рода возникают тогда, когда складывается большое число независимых или слабозависимых случайных процессов примерно одинаковой мощности. В этом случае, как показано ниже, с увеличением числа слагаемых сумма сходится к гауссовскому (нормальному) случайному процессу независимо от того, как распределены отдельные слагаемые.

Пусть случайный процесс $\xi(t)$ задан n сечениями $\xi(t_1), \xi(t_2), ..., \xi(t_n)$, для которых известны математические ожидания $M\xi(t_i) = a_i, i = \overline{1,n}$, и матрица ковариаций $K = \left\| K(t_i, t_j) \right\|$, $i, j = \overline{1,n}$, где

$$K(t_i, t_j) = M\left[(\xi(t_i) - a_i)(\xi(t_j) - a_j) \right] \cdot$$

Определение. Случайный процесс $\xi(t)$ называется гауссовским, или нормальным, если его n-мерная плотность распределения является нормальной

$$p(x_1, t_1; x_2, t_2; ...; x_n, t_n) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{|K|}} \exp\left\{ -\frac{1}{2|K|} \sum_{i,j=1}^{n} \left[(x_i - a_i) K_{ij} (x_j - a_j) \right] \right\}$$

для любого $n > 0$. Здесь $|K|$ – определитель матрицы K, $K_{i,j}$ – алгебраическое дополнение элемента $K(t_i, t_j)$ матрицы K.

Почти все свойства гауссовских процессов вытекают из свойств многомерных нормальных СВ. Например, коэффициенты разложения гауссовского случайного процесса в ряд по ортогональным функциям являются независимыми нормальными СВ. Это следует из того, что линейное преобразование многомерной нормальной СВ является также многомерной нормально распределенной СВ.

Теорема 4.1 ***(центральная предельная теорема для случайных процессов).*** *Пусть дана последовательность сумм случайных процессов*

$$\eta_n(t) = \sum_{k=1}^{m_n} \xi_{n_k}(t), \quad n = 1, 2, \ldots,$$

и выполняются следующие условия:

а) при фиксированном n *и любых моментах времени*

$$t_1, t_2, \ldots, t_{m_n}$$

СВ $\xi_{n_1}(t_1), \xi_{n_2}(t_2), \ldots, \xi_{n_{m_n}}(t_{m_n})$ *независимы и имеют конечные моменты* $M\xi_{n_k}(t) = 0$, $M\left[\xi_{n_k}^2(t)\right] = b_{n_k}^2(t)$*;*

б) существует предел

$$\lim_{n \to \infty} R_n(t_1, t_2) = R(t_1, t_2),$$

где $R_n(t_1, t_2) = M\left[\eta_n(t_1)\eta_n(t_2)\right]$ *–- корреляционная функция процесса* $\eta_n(t)$*;*

в) суммы $\eta_n(t)$ *при любом* t *удовлетворяют условию Линдеберга*

$$\frac{1}{B_n^2(t)} \sum_{k=1}^{m_n} \int_{|x| \ge \tau B_n^2} x^2 dF_{n_k}(x, t) \xrightarrow[m_{n \to \infty}]{} 0 \quad \forall \tau > 0,$$

где

$$B_n^2(t) = \sum_{k=1}^{m_n} b_{n_k}(t) = R_n(t, t), \; F_{n_k}(t) \text{- ф.р. СВ } \xi_{n_k}(t).$$

Тогда случайный процесс $\eta_n(t)$ при $n \to \infty$ сходится к гауссовскому случайному процессу с нулевым математическим ожиданием и дисперсией $R(t_1, t_2)$.

Эта теорема вытекает из центральной предельной теоремы для сумм независимых СВ.

§22. ВИНЕРОВСКИЙ СЛУЧАЙНЫЙ ПРОЦЕСС

История возникновения этого случайного процесса началась с наблюдений английского ботаника Р.Броуна в 1827 г., заметившего, что маленькие частицы, помещенные в жидкость, совершают непрерывное беспорядочное движение. В 1905 г. А.Эйнштейн объяснил это явление тем, что наблюдаемые молекулы подвержены непрерывным соударениям с молекулами окружающей среды. Выведенные А.Эйнштейном аналитические результаты были позднее проверены экспериментально и обобщены другими физиками и математиками. Первое математически чёткое построение теории таких процессов было дано Н.Винером в 1918 г. и в его более поздних работах.

Пусть $\xi(t)$ – расстояние броуновской частицы от начальной точки в момент времени t. Смещение $\xi(t_2)-\xi(t_1)$ за интервал времени $[t_1, t_2)$ можно рассматривать как сумму большого числа малых смещений. В этой ситуации применима центральная предельная теорема, поэтому естественно ожидать, что $\xi(t_2)-\xi(t_1)$ имеет нормальное распределение. Аналогично естественно предположить, что распределения величин $\xi(t_2)-\xi(t_1)$ и $\xi(t_2+\tau)-\xi(t_1+\tau)$ совпадают при любом $\tau > 0$, если среда находится в равновесии. Наконец, интуитивно ясно, что смещение $\xi(t_2)-\xi(t_1)$ должно зависеть только от $t_2 - t_1$, а не от момента начала наблюдения.

Винеровский процесс, называемый также процессом броуновского движения, играет фундаментальную роль при изучении многих случайных процессов других типов и широко применяется в различных областях, в частности, в финансовой математике и экономике [9] (см. главу 11).

Определение. Однородный гауссовский процесс с независимыми приращениями $\xi(t)$, для которого

$$m_\xi(t)=0, \; D_\xi(t)=\sigma^2 t,$$

называется винеровским процессом (или процессом броуновского движения).

Ковариационная функция такого процесса при $t_1 < t_2$ имеет вид

$$K(t_1,t_2)=M[\xi(t_1)\xi(t_2)]=M[\xi(t_1)(\xi(t_2)-\xi(t_1))+\xi^2(t_1)]=$$
$$=M\xi(t_1)M[\xi(t_2)-\xi(t_1)]+M\xi^2(t_1)=M\xi^2(t_1)=\sigma^2 t_1 . \quad (4.13)$$

Отсюда следует, что винеровский процесс является непрерывным, но не является дифференцируемым в среднем квадратичном; это можно показать аналогично, как для пуассоновского процесса.

Ковариационная матрица винеровского процесса записывается в виде

$$K=\sigma^2\begin{pmatrix} t_1 & t_1 & . & . & . & t_1 \\ t_1 & t_2 & . & . & . & t_2 \\ & . & . & . & & \\ t_1 & t_2 & . & . & . & t_n \end{pmatrix}.$$

Определение. Винеровский процесс, у которого $\sigma=1, \xi(0)=0$, называется стандартным винеровским процессом (стандартным броуновским движением) и обычно обозначается $W(t)$.

Найдем n-мерную плотность распределения такого процесса. Введем обозначения

$$\eta_1=W(t_1), \eta_i=W(t_i)-W(t_{i-1}), \quad (4.14)$$

где $i>1, t_1<t_2<\ldots<t_n$. Из (4.13), (4.14) следует, что

$$M\eta_1^2 = t_1,$$

$$M\eta_i^2 = M[\xi(t_i) - \xi(t_{i-1})]^2 = (t_i - 2t_{i-1} + t_{i+1}) = t_i - t_{i-1},\ i = 2, 3, ..., n. \tag{4.15}$$

$$M(\eta_i \eta_j) = M[(W(t_i) - W(t_{i-1}))(W(t_j) - W(t_{j-1}))] =$$

$$= (t_i - t_{i-1} - t_i + t_{i-1}) = 0,\ 1 < i < j < n.$$

СВ $(\eta_1, \eta_2, ..., \eta_n)$ является n-мерной СВ, имеющей нормальное распределение. Её плотность распределения записывается в виде

$$\tilde{p}(y_1, t_1; y_2, t_2; ...; y_n, t_n) =$$

$$= \frac{1}{\sqrt{2\pi t_1}} \exp\left(-\frac{y_1^2}{2t_1}\right) \prod_{i=2}^{n} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left[-\frac{y_i^2}{2(t_i - t_{i-1})}\right],$$

поэтому n-мерная плотность распределения винеровского процесса имеет вид

$$p(x_1, t_1; x_2, t_2; ...; x_n, t_n) =$$

$$= \frac{1}{\sqrt{2\pi t_1}} \exp\left(-\frac{x_1^2}{2t_1}\right) \prod_{i=2}^{n} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left[-\frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}\right].$$

ГЛАВА 5. МАРКОВСКИЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ

§23. ОПРЕДЕЛЕНИЯ И ПРИМЕРЫ

Марковские случайные процессы, или процессы без последствия, являются удобной математической моделью для многих реальных процессов. Рассмотрим систему, которая может находиться в различных состояниях, и пусть ее эволюция во времени носит стохастический характер, то есть состояние системы в момент времени t в общем случае не определяется однозначно состояниями системы в предыдущие моменты $s < t$. Тогда состояние этой системы можно описать некоторым случайным процессом $\xi(t)$, заданным на интервале времени T и принимающим значения из множества X.

Пусть заданы n сечений этого процесса $\xi(t_1), \xi(t_2), ..., \xi(t_{n+1})$ в моменты $t_1 < t_2 < ... < t_{n+1}$. Рассмотрим условную плотность вероятностей

$$p(x_{n+1}, t_{n+1} / x_1, t_1; x_2, t_2; ...; x_n, t_n) = \frac{p(x_1, t_1; x_2, t_2; ...; x_{n+1}, t_{n+1})}{p(x_1, t_1; x_2, t_2; ...; x_n, t_n)}. \quad (5.1)$$

Случайный процесс $\xi(t)$ называется марковским, если его условная плотность распределения (5.1) не зависит от значений процесса в моменты $t_1, t_2, ..., t_{n-1}$, а определяется лишь значением $\xi(t_n) = x_n$, т.е.

$$p(x_{n+1}, t_{n+1} / x_1, t_1; x_2, t_2; ...; x_n, t_n) = p(x_{n+1}, t_{n+1} / x_n, t_n) \cdot \quad (5.2)$$

Эту условную плотность вероятностей называют вероятностью перехода системы из состояния x_n, в котором она находилась в момент времени t_n, в состояние x_{n+1} в момент $t_{n+1} > t_n$ и обозначают

$$p(x_{n+1}, t_{n+1} / x_n, t_n) = p(x_{n+1}, t_{n+1}; x_n, t_n) \cdot \quad (5.3)$$

Рассмотрим многомерную плотность распределения процесса $\xi(t)$. С учетом условия марковости (5.2) ее можно записать в виде

$$p(x_1,t_1;x_2,t_2;...;x_{n+1},t_{n+1}) = p(x_{n+1},t_{n+1}/x_n,t_n)\,p(x_n,t_n/x_{n-1},t_{n-1})\times$$
$$\times ...\cdot p(x_2,t_2/x_1,t_1)\cdot p(x_1,t_1)\cdot \tag{5.4}$$

Отсюда видно, что начальное одномерное распределение $p(x_1,t_1)$ и вероятности перехода $p(t,x;y,s)$, $t>s$, полностью задают марковский процесс. Вероятности перехода удовлетворяют двум основным соотношениям:

а) $\int p(x,t;y,s)dx = 1$ (условие нормировки);

б) проинтегрируем (5.4) по некоторому промежуточному значению x_i, например, по x_2:

$$p(x_1,t_1;x_3,t_3;...;x_{n+1},t_{n+1}) = p(x_{n+1},t_{n+1};x_n,t_n)\cdot...\cdot p(x_1,t_1)\times$$
$$\times\int p(x_3,t_3;x_2,t_2)p(x_2,t_2;x_1,t_1)dx_2\,;$$

с другой стороны, по формуле (5.4) имеем

$$p(x_1,t_1;x_3,t_3;...;x_{n+1},t_{n+1}) = p(x_{n+1},t_{n+1};x_n,t_n)\cdot...\cdot p(x_4,t_4;x_3,t_3)\times$$
$$\times\, p(x_3,t_3;x_1,t_1)\,p(x_1,t_1)\,;$$

из сопоставлений этих выражений видно, что

$$p(x_3,t_3;x_1,t_1) = \int p(x_3,t_3;x_2,t_2)p(x_2,t_2;x_1,t_1)dx_2. \tag{5.5}$$

В общем случае определение марковского процесса может быть дано следующим образом. Пусть $X \subseteq R$.

Определение. Случайный процесс $\xi(t)$, $t\in T$, со значениями в X называется марковским, если $\forall t_1 < t_2 < ... < t_n < t_{n+1} \in T$ и любых борелевских множеств $B_1, B_2, ..., B_{n-1}, B_{n+1}$ из R при фиксированном борелевском множестве B_n выполняется следующее соотношение

$$P\big(\xi(t_{n+1})\in B_{n+1}/\xi(t_1)\in B_1, \xi(t_2)\in B_2,...,\xi(t_n)\in B_n\big) =$$
$$= P\big(\xi(t_{n+1})\in B_{n+1}/\xi(t_n)\in B_n\big)\cdot$$

При этом очевидно (по определению условной вероятности), что

$$P\big(\xi(t_1)\in B_1,\xi(t_2)\in B_2,...,\xi(t_n)\in B_n\big)>0\cdot$$

Если t_n – настоящий момент времени, t_{n+1} – некоторый момент в будущем, а $t_1,t_2,...,t_{n-1}$ – некоторые моменты в прошлом. Поэтому можно сказать, что марковский случайный процесс – это процесс, из которого «при фиксированном настоящем будущее не зависит от прошлого».

Марковские процессы можно разбить на классы в зависимости от структуры множества значений случайного процесса X и интервала наблюдения T.

Определение. Если множество $X=\{i_1, i_2, ..., i_n, ...\}$ счетно или конечно, то марковский процесс называется цепью Маркова. Множество значений цепи Маркова X называют еще фазовым пространством или пространством состояний цепи Маркова.

Определение. Цепь Маркова, у которой множество T дискретно, например, $T=\{0, 1, 2, ...\}$, называется цепью с дискретным временем, а цепь Маркова, у которой T – некоторый интервал положительной длины (конечный или бесконечный), называется цепью с непрерывным временем.

Для цепи Маркова с дискретным временем можно записать

$$P(\xi_{n+1}=i_{n+1}/\xi_1=i_1,\xi_2=i_2,...,\xi_n=i_n)=P(\xi_{n+1}=i_{n+1}/\xi_n=i_n),$$

где $i_n=\xi(t_n)$ – состояние цепи Маркова в некоторый дискретный момент времени t_n, $t_1<t_2<...<t_n<...$. То есть ξ_{n+1} полностью определяется значением ξ_n и $(n+1)$-м элементарным событием ω_{n+1} из последовательности независимых элементарных событий $\{\omega_n\}$.

Рассмотрим примеры цепей Маркова с дискретным временем.

Пример 5.1. Производится последовательность независимых испытаний, каждое из которых может быть «успешным» с вероятностью p и «неудачным» с вероятностью $q = 1 - p$. Обозначим через ν_n – число успехов в первых n испытаниях, $n = 1,2,\ldots$. Введем СВ

$$\eta_{n+1} = \eta_{n+1}(\omega_{n+1}) = \begin{cases} 1, & \text{если}\,(n+1)-\text{е испытание успешно}, \\ 0, & \text{в противном случае.} \end{cases}$$

Очевидно, что $\nu_{\text{n+1}} = \nu_{\text{n}} + \eta_{n+1}$, поэтому $\{\nu_n\}$ – цепь Маркова.

Пример 5.2. Имеется водохранилище, которое в любой момент времени $t \geq 0$ характеризуется некоторым уровнем наполнения $x(t)$. Минимальный уровень равен 0. Приток воды происходит в моменты $\tau, 2\tau, \ldots, n\tau, \ldots$, причем порции воды, поступающие в эти моменты – независимые СВ $\eta_1^{(\omega_1)}, \eta_2^{(\omega_2)}, \ldots, \eta_n^{(\omega_n)}$. Сток воды равномерен, т.е. за время dt в интервале между $n\tau$ и $(n+1)\tau$ уровень наполнения водохранилища уменьшается на величину αdt, $\alpha > 0$, если этот уровень еще не достиг нуля. Пусть ξ_n – уровень наполнения водохранилища в момент, непосредственно следующий за моментом поступления n-й порции воды:

$$\xi_n = x(n\tau + o) = \lim_{\varepsilon \downarrow 0} x(n\tau + \varepsilon)\cdot$$

Тогда

$$\xi_{n+1} = \begin{cases} \xi_n - \alpha\tau + \eta_{n+1}, & \xi_n - \alpha\tau > 0, \\ \eta_{n+1}, & \xi_n - \alpha\tau \leq 0, \end{cases}$$

и поэтому $\{\xi_n\}$ является цепью Маркова.

Опишем похожую модель из теории запасов.

Пример 5.3. Рассмотрим систему, в которой запасается некоторый товар с целью удовлетворения постоянного спроса. Предположим, что восполнение запаса осуществляется в моменты времени $t_1, t_2, \ldots, t_n, \ldots$, а суммарный спрос $\eta_n(\omega_{n+1})$ на товар в интервале (t_{n-1}, t_n) является СВ с распределением

$$P\{\eta_n = k\} = p_k,\ k = 0,1,2,...,$$

одинаковым для всех интервалов, $p_k \geq 0$, $\sum_{k=0}^{\infty} p_k = 1$. Уровень запаса фиксируется в начале каждого периода. Стратегия запаса такова: если имеющееся количество товара не превышает некоторого критического уровня u, то производится немедленное пополнение запаса до $U > u$. Если же имеющееся количество товара больше u, то пополнение не производится. Пусть ξ_n означает уровень наличного запаса непосредственно перед моментом t_n. Множество состояний процесса $\{\xi_n\}$ складывается из возможных значений уровня запаса

$$U, U-1, ..., 1, 0, -1, -2, ...,$$

где отрицательные значения интерпретируются как неудовлетворенный спрос (эти заказы подлежат немедленному исполнению при пополнении запаса). Согласно описанной стратегии уровни запаса двух последовательных периодов связаны соотношением

$$\xi_{n+1} = \begin{cases} \xi_n - \eta_{n+1},\ u < \xi_n \leq U, \\ U - \eta_{+1},\ \xi_n \leq u. \end{cases}$$

Если предположить, что СВ $\eta_1, \eta_2, ..., \eta_n, ...$ независимы, то уровни запаса $\{\xi_n\}$, $n = 0,1,2,...$, образуют марковскую цепь.

Пример 5.4. Предположим, что изменение величины заряда $\xi \geq 0$ в электронной лампе описывается следующим законом. В интервалах между попаданием на сетку заряженных частиц заряд убывает по экспоненциальному закону с параметром α; в момент попадания n-й частицы заряд увеличивается на СВ $\eta_n(\omega_n)$. Пусть τ_n – время между попаданием на сетку $(n-1)$-й и n-й частиц, ξ_n – величина заряда после попадания на сетку n-й частицы. Тогда

$$\xi_{n+1} = \xi_n e^{-\alpha\tau_{n+1}} + \eta_{n+1}.$$

Если двумерные СВ (τ_n, η_n) независимы в совокупности, то последовательность $\{\xi_n\}$ является цепью Маркова.

Пример 5.5. Пусть на стоянку такси в единичные моменты времени прибывают (по одной в каждый момент) машины. Если на стоянке нет ожидающих, то машина немедленно уезжает. Обозначим через $\eta_n(\omega_n)$ число пассажиров такси, приходящих в момент n на стоянку, и будем предполагать, что $\eta_1, \eta_2, ..., \eta_n, ...$ – независимые СВ. Пусть ξ_n – длина очереди в момент n, $\xi_0 = 0$. Тогда, если $\xi_n = i$, то в следующий момент $n+1$ длина очереди ξ_{n+1} станет равной

$$\begin{cases} \eta_{n+1}, & i = 0, \\ i - 1 + \eta_{n+1}, & i \geq 1. \end{cases}$$

Иначе говоря,

$$\xi_{n+1} = (\xi_n - 1)^+ + \eta_{n+1},$$

где $a^+ = \max(a, 0)$, и значит, последовательность $\{\xi_n\}$ образует цепь Маркова.

Определение. Если множество состояний X непрерывно и система переходит из состояния в состояние в произвольные моменты времени из множества T, то соответствующий процесс называется непрерывным марковским процессом.

Например, координата броуновской частицы принимает значения из непрерывного множества X и переходы происходят в произвольные моменты времени.

§24. ОДНОРОДНЫЕ ЦЕПИ МАРКОВА

Рассмотрим цепь Маркова с дискретным временем и, для простоты, с конечным числом состояний $X = \{1, 2, ..., N\}$. Вероятности

$$p_{ij}^{(n)} = P(\xi_{m+n} = j \mid \xi_m = i),\ n \ge 1,\ m \ge 1,\ i, j \in X \quad (5.6)$$

называются вероятностями перехода цепи Маркова за n шагов, а $p_{ij}^{(1)} = p_{ij}$ – просто вероятностями перехода (за один шаг). Матрица вида

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ . & . & . & . \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{pmatrix}$$

называется матрицей вероятностей перехода цепи Маркова. Очевидно, что сумма элементов в каждой строке этой матрицы равна 1, т.е. $\sum_{j=1}^{\infty} p_{ij} = 1$, $i = \overline{1, N}$, т.к. за один шаг цепь Маркова либо остается в своем состоянии, либо переходит в какое-то иное состояние. Пусть $p_i^{(0)}$ вероятность состояния i на нулевом шаге; набор $\{p_i^{(0)}, i \in X\}$ называется начальным распределением цепи Маркова.

Определение. Если вероятности перехода (5.6) не зависят от m, то цепь Маркова называется однородной.

Понятно, что свойства однородных марковских цепей полностью определяются начальными распределениями и вероятностями перехода p_{ij}. В конкретных случаях для описания эволюции цепи вместо явного выписывания матрицы $P = \|p_{ij}\|$ используют ориентированный граф, вершинами которого являются состояния из множества X, а стрелка

$$i \xrightarrow{p_{ij}} j$$

означает, что из состояния i возможен переход в состояние j с вероятностью p_{ij}; в том случае, когда $p_{ij}=0$, соответствующая стрелка не проводится.

Пример 5.6. **Одномерное случайное блуждание.** При рассмотрении системы случайных блужданий состояние системы для наглядности интерпретируют как движение движущейся частицы.

Одномерное случайное блуждание представляет собой марковскую цепь, пространство состояний которой состоит из конечного или бесконечного множества целых чисел. Если частица находится в состоянии i, то за один шаг она может либо перейти в одно из своих соседних состояний $(i-1$ или $i+1)$ соответственно с вероятностями q_i и p_i, либо остаться в состоянии i с вероятностью r_i. Граф переходов для случая, когда пространством состояний служит множество неотрицательных целых чисел, представлен на рис. 19.

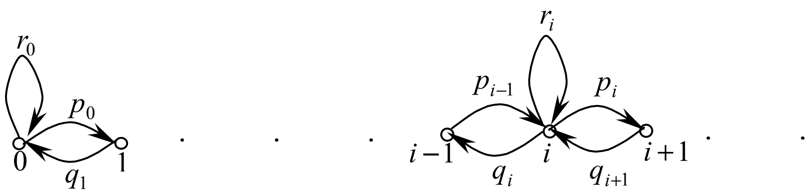


Рис. 19

Матрица вероятностей переходов имеет вид

$$P=\begin{pmatrix} r_0 & p_0 & 0 & 0 & & & \\ q_1 & r_1 & p_1 & 0 & & & \\ 0 & q_2 & r_2 & p_2 & & & \\ & . & . & . & . & & \\ & & q_i & r_i & p_i & 0 & \\ & & . & . & . & . & . \end{pmatrix},$$

где $p_0, r_0 \ge 0$, $p_0+r_0=1$; $p_i, q_i>0$, $r_i \ge 0$, $p_i+q_i+r_i=1$, $i\ge 1$.

Пусть ξ_n – состояние частицы через n шагов. Тогда

$$p_{i\,i+1} = P(\xi_{n+1} = i+1 / \xi_n = i) = p_i, \; p_{i\,i-1} = P(\xi_{n+1} = i-1 / \xi_n = i) = q_i,$$

$$p_{ii} = P(\xi_{n+1} = i / \xi_n = i) = r_i, \; i \ge 1, \; q_0 = 0,$$

$$p_{ij} = 0 \text{ для } |i - j| > 1.$$

В пользу названия «случайное блуждание» для процесса такого типа говорит тот факт, что его реализация описывает путь «абсолютно пьяного» человека, делающего случайным образом шаг назад или шаг вперед. Процессом случайного блуждания описывают также капитал игрока, участвующего в серии партий азартной игры. Предположим, что игрок A, имеющий капитал i, играет с бесконечно богатым партнером; при этом вероятность того, что он выиграет партию и увеличит свой капитал на единицу, равна p_i, а вероятность того, что он проиграет и тем самым уменьшит свой капитал на единицу, равна $q_i = 1 - p_i$, $i \ge 1$. Зависимость вероятностей выигрыша и проигрыша от i отражает возможную зависимость условий игры от капитала. Так, можно условиться, что, оказавшись в состоянии 0, соответствующем разорению игрока A, процесс остается в этом состоянии, т.е. $r_0 = 1$. Процесс $\{\xi_n\}$, где ξ_n – размер капитала игрока A после n партий, является процессом случайного блуждания. Этот процесс известен под названием «задачи о разорении игрока».

Пусть теперь $X = \{0, \pm 1, \pm 2, ..., \pm N\}$, $p_0^{(0)} = 1$, $p_{NN} = p_{(-N)(-N)} = 1$, а для $|i| < N$:

$$p_{ij} = \begin{cases} p, j = i+1, \\ q, j = i-1, \\ 0, \text{в остальных случаях}. \end{cases} \quad (5.7)$$

Граф переходов для случая $N = 3$ изображен на рис. 20.

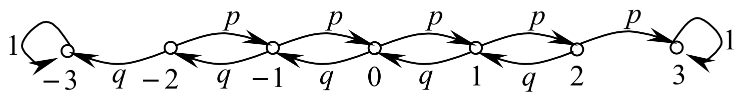


Рис. 20

Эта цепь отвечает игре двух игроков A и B, когда капитал каждого равен N и на каждом шаге игрок A с вероятностью p выигрывает у игрока B $+1$ и проигрывает -1 с вероятностью q. Если трактовать состояние i как величину выигрыша игрока A у игрока B, то достижение состояний N и $-N$ означает разорение игрока B и игрока A соответственно.

В самом деле, если $\eta_1, \eta_2, ...\eta_n$ – независимые бернуллиевские СВ, для которых $P(\eta_i = +1) = p$, $P(\eta_i = -1) = q$, $S_0 = 0$, а $S_n = \eta_1 + \eta_2 + ... + \eta_n$ – величина выигрыша игрока A у игрока B, то последовательность $\{S_n\}$ образует цепь Маркова с $p_0 = 1$ и вероятностями перехода (5.7), поскольку $S_{n+1} = S_n + \eta_{n+1}$ и

$$P(S_{n+1} = j / S_n = i_n, S_{n-1} = i_{n-1}, ...) =$$
$$= P(S_n + \eta_{n+1} = j / S_n = i_n, S_{n-1} = i_{n-1}, ...) =$$
$$= P(S_n + \eta_{n+1} = j / S_n = i_n) = P(\eta_{n+1} = j - i_n).$$

Рассмотрим теперь цепь Маркова с непрерывным временем. Вероятности

$$p_{ij}(t) = P\big(\xi(s+t) = j / \xi(s) = i\big),\ s \le t,\ i, j \in X. \qquad (5.8)$$

называются вероятностями перехода цепи Маркова за время t. Введем также следующее обозначение для вероятностей состояний

$$p_i(t) = P\big(\xi(t) = i\big),\ t > 0,\ i \in X.$$

Набор $\{p_i(0), i \in X\}$ называется начальным распределением цепи Маркова с непрерывным временем.

Определение. Если вероятности перехода (5.8) не зависят от s, то цепь Маркова с непрерывным временем называется однородной.

Вернемся опять к пуассоновскому случайному процессу (см. примеры 2.3, 4.1) и покажем в качестве примера, что он является однородной цепью Маркова с непрерывным временем. При этом само определение процесса мы введем опять несколько по-другому.

Пример 5.7. **Простейший поток событий. Пуассоновский процесс.**

Пусть $\{z_k\}$, $k = 1,2,...$ – последовательность независимых положительных СВ, имеющих одно и то же экспоненциальное распределение с ф.р.

$$F_{z_K}(x) = P(z_k \le x) = 1 - e^{-\lambda x},\ x \ge 0 \cdot$$

Положим $t_0 = 0$, $t_k = z_1 + z_2 + ... + z_k$, $k \ge 1$; очевидно, что $t_{k+1} > t_k$. Множество точек $\{t_k\}$, $k = 0,1,2,...$, называется **простейшим пуассоновским потоком однородных событий.** Число тех k, для которых $a \le t_k < b$, называется числом событий потока в интервале $[a,b]$. Обозначим через $\nu(t)$, $t \ge 0$, число событий простейшего потока в интервале $[0,t)$. Определенный таким образом случайный процесс $\nu(t)$ назовем пуассоновским процессом с параметром λ.

Покажем, что выполняется следующее соотношение

$$P(\nu(s+t) = j \mid \nu(s) = i) = \frac{(\lambda t)^{j-i}}{(j-i)!} e^{-\lambda t},\ j \ge i \cdot \qquad (5.9)$$

Введем следующие обозначения (см. рис 5.3):

$$
\begin{aligned}
&z_1' = z_{i+1} - (s - t_i), && t_0' = 0, \\
&z_2' = z_{i+2}, && t_k' = z_1' + z_2' + \ldots + z_k' \\
&z_3' = z_{i+3}, && t_{i+k}' = s + t_k'. \\
&\ldots
\end{aligned}
\tag{5.10}
$$

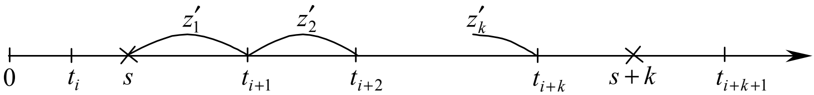


Рис. 21

Ясно, что

$$P(\nu(t) = k) = P(t_k < t,\, t_{k+1} \geq t) \cdot \tag{5.11}$$

Кроме того (см. рис. 21),

$$P(\nu(s+t) = i + k \mid \nu(s) = i) = P(t_k' < t, t_{k+1}' \geq t \mid \nu(s) = i)\,. \tag{5.12}$$

Из свойства экспоненциального распределения (см. §4) следует, что СВ $z_1' = z_{i+1} - \tau$, где $\tau = s - t_i$, имеет такое же экспоненциальное распределение, как и СВ z_{i+1}, если $\tau > 0$ и $z_1' \geq 0$, то есть, если $t_i < s$ и $t_{i+1} \geq s$, что эквивалентно $\nu(s) = i$. Отсюда вытекает, что при выполнении условия $(\nu(s) = i)$ СВ $\{z_k'\}$, $k = 1,2,\ldots$, независимы и одинаково распределены по экспоненциальному закону с параметром λ. Тогда из (5.10) – (5.12) получаем

$$P\bigl(\nu(s+t) = i + k \mid \nu(s) = i\bigr) = P\bigl(\nu(t) = k\bigr) \cdot \tag{5.13}$$

Теперь по индукции докажем, что выполняются соотношения (2.1), (2.6). При $k = 0$ имеем $P(\nu(t) = 0) = P(t_1 > t) = e^{-\lambda t}$, то есть (2.6) выполнено. Предположим, что (2.6) выполняется при некотором k и покажем, что оно будет иметь место для $k+1$. Используя определение $\nu(t)$ и формулу полной вероятности, имеем

$$P\big(\nu(t)=k+1\big)=\int_0^t P\big(\nu(t-x)=k\big)d(1-e^{-\lambda x})=$$

$$=\int_0^t \frac{[\lambda(t-x)]^k}{k!}e^{-\lambda(t-x)}\lambda e^{-\lambda x}dx=\frac{\lambda^{k+1}e^{-\lambda t}}{k!}\int_0^t (t-x)^k\,dx=$$

$$=-\frac{\lambda^{k+1}e^{-\lambda t}}{(k+1)!}(t-x)^{k+1}\Bigg|_0^t=\frac{(\lambda t)^{k+1}}{(k+1)!}e^{-\lambda t},$$

т.е. формула (2.6) справедлива и для $k+1$, а значит, справедлива $\forall\ k\geq 0.$. Из соотношений (5.13) и (2.6) следует (5.9).

Рассмотрим основные свойства пуассоновского процесса и соответственно простейшего потока событий.

1. Стационарность означает, что для любых интервалов времени $[\tau_0,\tau_1),[\tau_1,\tau_2),...,[\tau_{n-1},\tau_n)$, где $0\leq\tau_0<\tau_1<\tau_2<...<\tau_n$, любого $\tau>0$ и любых неотрицательных чисел $k_1,k_2,...,k_n$ вероятность того, что на интервале $[\tau_0,\tau_1)$ произойдет k_1 событий потока, …, на интервале $[\tau_1,\tau_2)$ произойдет k_2 событий потока, …, на интервале $[\tau_{n-1},\tau_n)$ произойдет k_n событий потока равна вероятности того, что на интервале $[\tau_0+\tau,\tau_1+\tau)$ произойдет k_1 событий потока, на интервале $[\tau_1+\tau,\tau_2+\tau)$ произойдет k_2 событий потока, на интервале $[\tau_{n-1}+\tau,\tau_n+\tau)$ – k_n событий потока.

Это свойство можно также сформулировать по-другому: для любых фиксированных $0\leq\tau_0<\tau_1<\tau_2<...<\tau_n$, любого $\tau>0$ многомерные СВ

$$(\nu(\tau_1)-\nu(\tau_0),\nu(\tau_2)-\nu(\tau_1),...,\nu(\tau_n)-\nu(\tau_{n-1}))$$

и

$$(\nu(\tau_1+\tau)-\nu(\tau_0+\tau),\nu(\tau_2+\tau)-\nu(\tau_1+\tau),...,\nu(\tau_n+\tau)-\nu(\tau_{n-1}+\tau))$$

имеют одинаковое распределение.

2. Ординарность означает, что вероятность того, что на интервале $[t, t+\Delta t)$ произойдет два или более событий, равна $o(\Delta t)$ при $\Delta t \to 0$, т.е.

$$P\big(\nu(t+\Delta t)-\nu(t)\geq 2\big)=o(\Delta t), \ \Delta t \to 0.$$

3. Отсутствие последействия означает, что при $0 \leq \tau_0 < \tau_1 < \tau_2 < ... < \tau_n$,

СВ $(\nu(\tau_1)-\nu(\tau_0), \nu(\tau_2)-\nu(\tau_1), ..., \nu(\tau_n)-\nu(\tau_{n-1})$ независимы.

Можно показать, что выполняется следующее соотношение

$$P\big(\nu(\tau_1)-\nu(\tau_0)=k_1,\ \nu(\tau_2)-\nu(\tau_1)=k_2,\ ...,\ \nu(\tau_n)-\nu(\tau_{n-1})=k_n\big)=$$

$$=\prod_{i=1}^{n}\left\{\frac{[\lambda(\tau_i-\tau_{i-1})]^{k_i}}{k_i!}e^{-\lambda(\tau_i-\tau_{i-1})}\right\}, \tag{5.14}$$

откуда и следует свойство отсутствия последействия простейшего потока событий.

Свойство стационарности вытекает из того, что в последнюю формулу входят лишь разности $\tau_i - \tau_{i-1}$. Таким образом, если ко всем τ_i добавить некоторое τ, то вид вероятности (5.14) не изменится.

Величина $\nu(t+\Delta t)-\nu(t)$, согласно (5.9), имеет распределение Пуассона с параметром $\lambda\Delta t$. Отсюда следует свойство ординарности, поскольку

$$P\big(\nu(t+\Delta t)-\nu(t)\geq 2\big)=1-P\big(\nu(t+\Delta t)-\nu(t)=0\big)-$$

$$-P(\nu(t+\Delta t)-\nu(t)=1)=1-e^{-\lambda\Delta t}-\lambda\Delta t e^{-\lambda\Delta t}=$$

$$=e^{-\lambda\Delta t}(e^{-\lambda\Delta t}-1-\lambda\Delta t)=\frac{(\lambda\Delta t)^2}{2}+O(\Delta t^3).$$

Простейший поток событий и пуассоновский процесс (как уже упоминалось выше) имеют огромное число приложений. Этими математическими схемами можно описать самые различные модели многих явлений. Об их применении речь будет идти в последующих главах.

Покажем, что пуассоновский процесс является цепью Маркова. В силу отсутствия последствия

$$P(\nu(\tau_1)=k_1,\nu(\tau_2)=k_2,...,\nu(\tau_n)=k_n)=$$

$$P(\nu(\tau_1)=k_1,\nu(\tau_2)-\nu(\tau_1)=k_2-k_1,...,\nu(\tau_n)-\nu(\tau_{n-1})=k_n-k_{n-1})=$$

$$=P(\nu(\tau_1)=k_1)\prod_{l=2}^{n}P(\nu(\tau_l)-\nu(\tau_{l-1})=k_l-k_{l-1}),$$

$$P(\nu(\tau_1)=k_1,\nu(\tau_2)=k_2,\ ...,\ \nu(\tau_{n+1})=k_{n+1})=\cdot$$

$$=P(\nu(\tau_1)=k_1)\prod_{l=2}^{n+1}P(\nu(\tau_l)-\nu(\tau_{l-1})=k_l-k_{l-1})\cdot$$

Поэтому, используя определение условной вероятности и формулу (5.14), имеем

$$P(\nu(\tau_{n+1})=k_{n+1}\mid\nu(\tau_1)=k_1,\nu(\tau_2)=k_2,...,\nu(\tau_n)=k_n)=$$

$$=P(\nu(\tau_{n+1})-\nu(\tau_n)=k_{n+1}-k_n)=$$

$$=\frac{[\lambda(\tau_{n+1}-\tau_n)]^{k_{n+1}-k_n}}{(k_{n+1}-k_n)!}e^{-\lambda(\tau_{n+1}-\tau_n)}.$$

Кроме того, используя (5.13) и (2.6), получаем

$$P(\nu(\tau_{n+1})=k_{n+1}/\nu(\tau_n)=k_n)=P(\nu(\tau_{n+1}-\tau_n)=k_{n+1}-k_n)=$$

$$=\frac{[\lambda(\tau_{n+1}-\tau_n)]^{k_{n+1}-k_n}}{(k_{n+1}-k_n)!}e^{-\lambda(\tau_{n+1}-\tau_n)}.$$

Таким образом,

$$P(\nu(\tau_{n+1})=k_{n+1}/\nu(\tau_1)=k_1,\nu(\tau_2)=k_2,...,\nu(\tau_n)=k_n)=,$$

$$P=P(\nu(\tau_{n+1})=k_{n+1}/\nu(\tau_n)=k_n),$$

значит, $\nu(t)$ – цепь Маркова. Ее ординарность следует из соотношения (5.9) (выражение справа в нем не зависит от s).

ГЛАВА 6. ЦЕПИ МАРКОВА С ДИСКРЕТНЫМ ВРЕМЕНЕМ

§23. УРАВНЕНИЯ ЧЕПМЕНА–КОЛМОГОРОВА

Рассмотрим однородную цепь Маркова с дискретным временем. Ее вероятности перехода (5.6) не зависят от m.

Теорема 6.1. *Переходные вероятности* $p_{ij}^{(n)}$ *однородной цепи Маркова удовлетворяют* ***уравнению Чепмена–Колмогорова***:

$$p_{ij}^{(n+l)} = \sum_k p_{ik}^{(n)} p_{kj}^{(l)}. \tag{6.1}$$

Доказательство. Используя определение условной вероятности, легко проверить, что для любых случайных событий A, B, C справедливо равенство

$$P(A \cap B / C) = P(A / C)\, P(B / A \cap C), \tag{6.2}$$

если $P(C) \neq 0$, $P(A \cap C) \neq 0$. Случайные события $(\xi_n = k)$, $k = 1, 2, \ldots$, образуют полную группу, поэтому $\bigcup_k (\xi_n = k) = \Omega$. Для вероятностей переходов за $n+l$ шагов можно записать

$$p_{ij}^{(n+l)} = P\big(\xi_{n+l} = j / \xi_0 = i\big) = P\big((\xi_{n+l} = j) \cap \Omega / \xi_0 = i\big) =$$

$$= P\left((\xi_{n+l} = j) \cap \left(\bigcup_k (\xi_n = k) \right) / \xi_0 = i \right) = \sum_k P\big((\xi_{n+l} = j) \cap (\xi_n = k) / \xi_0 = i\big)$$

Пусть $A = (\xi_n = k)$, $B = (\xi_{n+l} = j)$, $C = (\xi_0 = i)$. Тогда, используя марковское свойство и равенство (6.2), получаем

$$p_{ij}^{(n+l)} = \sum_k P\big(\xi_n = k / \xi_0 = i\big) P\big(\xi_{n+l} = j / (\xi_n = k) \cap (\xi_0 = i)\big) =$$

$$= \sum_k p_{ik}^{(n)} P\big(\xi_{n+l} = j / \xi_n = k\big) = \sum_k p_{ik}^{(n)} p_{kj}^{(l)},$$

что и требовалось доказать.

Из соотношения (6.1) следует, что переход цепи Маркова из состояния i в состояние j за $n+l$ шагов может осуществиться путем перехода ее в некоторое промежуточное состояние k за первые n шагов и далее перехода из состояния k в состояние j за оставшиеся l шагов.

Особо важны следующие два частных случая уравнения (6.1):

обратное уравнение (когда $n=1$)

$$p_{ij}^{(l+1)} = \sum_k p_{ik} p_{kj}^{(l)} \tag{6.3}$$

и **прямое уравнение** (когда $l=1$)

$$p_{ij}^{(n+1)} = \sum_k p_{ik}^{(n)} p_{kj}, \tag{6.4}$$

см. рис. 22, 23.

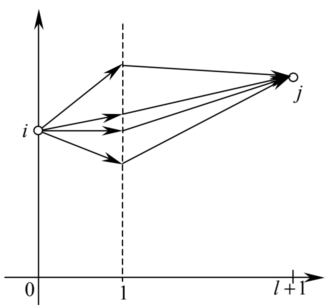


Рис. 22. К обраттному уравнению

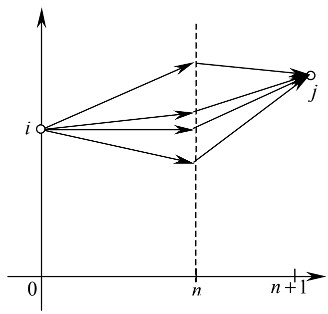


Рис. 23. К прямому уравнени.

Из обратного уравнения (6.3) следует

$$P^{(l)} = PP^{(l-1)} = PPP^{(l-2)} = \ldots = P^l,$$

где P – матрица вероятностей переходов за один шаг, $P^{(l)} = \left\| p_{ij}^{(l)} \right\|$ – матрица вероятностей переходов за l шагов, т.е. для того чтобы получить матрицу вероятностей переходов за l шагов, нужно матрицу вероятностей переходов за один шаг возвести в степень l.

Пример 6.1. Некоторая совокупность семей поделена на три группы: i_1 – семьи, не имеющие квартиры и не намеревающиеся ее купить; i_2 – семьи, не имеющие квартиры, но собирающиеся ее купить; i_3 – семьи, имеющие квартиру. Статистические обследования дали возможность оценить вероятность перехода семей из одной группы в другую на протяжении года. При этом матрица вероятностей переходов оказалась следующей

$$P = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0 & 0.7 & 0.3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Например,

$p_{33} = 1 = P$ {семьи, которые в предыдущем году имели квартиру, и в следующем году будут ее иметь},

$p_{23} = 0.3 = P$ {семья, не имевшая в предыдущем году квартиры, будет ее иметь следующем году},

$p_{21} = 0 = P$ {семья, которая хотела купить квартиру в предыдущем году, в следующем году от этого намерения отказалась}.

Нужно найти следующие вероятности:

$P_1 = P$ {семья, не имевшая в предыдущем году квартиры и не собирающаяся ее купить, будет находиться в такой же ситуации через два года},

$P_2 = P$ {семья, не имевшая квартиры и не намеревающаяся ее приобрести, будет иметь квартиру через два года}.

В данном случае имеем цепь Маркова с тремя состояниями i_1, i_2, i_3. Найдем матрицу вероятностей переходов за два шага (матрицу вероятностей перехода семей из одного состояния в другое через два года):

$$P^{(2)} = P^2 = \begin{pmatrix} 0{,}8 & 0{,}1 & 0{,}1 \\ 0 & 0{,}7 & 0{,}3 \\ 0 & 0 & 1 \end{pmatrix}^2 = \begin{pmatrix} 0{,}64 & 0{,}15 & 0{,}21 \\ 0 & 0{,}49 & 0{,}51 \\ 0 & 0 & 1 \end{pmatrix}.$$

Поэтому искомые вероятности равны

$$P_1 = p_{11}^{(2)} = 0{,}64;\ P_2 = p_{23}^{(2)} = 0{,}51.$$

Пример 6.2. Рассмотрим однородную марковскую цепь с двумя состояниями, 0 и 1, и матрицей вероятностей переходов

$$P = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}.$$

Нетрудно найти матрицу переходных вероятностей за два шага:

$$P^{(2)} = \begin{pmatrix} p_{00}^2 + p_{01}p_{10} & p_{01}(p_{00} + p_{11}) \\ p_{10}(p_{00} + p_{11}) & p_{11}^2 + p_{01}p_{10} \end{pmatrix}.$$

Можно также по индукции показать, что

$$P^{(n)} = P^n = \frac{1}{2 - p_{00} - p_{11}} \begin{pmatrix} 1 - p_{11} & 1 - p_{00} \\ 1 - p_{11} & 1 - p_{00} \end{pmatrix} +$$

$$+ \frac{(p_{00} + p_{11} - 1)^n}{2 - p_{00} - p_{11}} \begin{pmatrix} 1 - p_{00} & -(1 - p_{00}) \\ -(1 - p_{11}) & 1 - p_{11} \end{pmatrix},$$

в предположении, что

$$|p_{00} + p_{11} - 1| < 1. \tag{6.5}$$

Отсюда видно, что если выполняется неравенство (6.5), то

$$P^{(n)} \xrightarrow[n \to \infty]{} \frac{1}{2 - p_{00} - p_{11}} \begin{pmatrix} 1 - p_{11} & 1 - p_{00} \\ 1 - p_{11} & 1 - p_{00} \end{pmatrix},$$

и значит,

$$\lim_{n \to \infty} p_{i0}^{(n)} = \frac{1 - p_{11}}{2 - p_{00} - p_{11}},\ \lim_{n \to \infty} p_{i1}^{(n)} = \frac{1 - p_{00}}{2 - p_{00} - p_{11}}.$$

Таким образом, при выполнении условия (6.5) поведение рассматриваемой марковской цепи подчиняется следующей закономерности: влияние начального состояния на вероятность нахождения цепи в том или ином состоянии исчезает с ростом времени (числа шагов); $p_{ij}^{(n)}$ сходятся к предельным значениям π_j, не зависящим от i и образующим распределение вероятностей: $\pi_0, \pi_1 \geq 0$, $\pi_0 + \pi_1 = 1$. Если к тому же все элементы $p_{ij} > 0$, то предельные значения $\pi_0, \pi_1 > 0$; к этому свойству мы вернемся в §29.

§26. НАХОЖДЕНИЕ ВЕРОЯТНОСТЕЙ ПЕРЕХОДОВ С ПОМОЩЬЮ ПРОИЗВОДЯЩИХ ФУНКЦИЙ

Рассмотрим однородную цепь Маркова с дискретным временем и конечным числом состояний, $X = \{1, 2, ..., N\}$. Прямое уравнение Чепмена–Колмогорова (6.4) для нее можно переписать в виде

$$p_{ij}^{(n)} = \sum_{k=1}^{N} p_{ik}^{(n-1)} p_{kj}, \quad i, j = \overline{1, N}, \quad n \geq 2 \cdot \tag{6.6}$$

Данное соотношение обычно используют для вычисления $p_{ij}^{(n)}$ при небольших n. При больших n используют следующий метод.

Обозначим

$$p_{ij}^{(0)} = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \tag{6.7}$$

Тогда уравнение (6.6.) выполняется при $n = 1$, что можно проверить непосредственной подстановкой. Введем в рассмотрение производящие функции

$$\psi_{ij}(z) = \sum_{n=0}^{\infty} p_{ij}^{(n)} z^n \cdot \tag{6.8}$$

Ряд в правой части сходится по крайней мере при $|z|<1$, так как $0 \le p_{ij}^{(n)} \le 1$. Умножив обе части уравнения (6.6) на z^n и просуммировав по n от 1 до ∞, получим

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} z^n = \sum_{n=1}^{\infty}\sum_{k=1}^{N} p_{ik}^{(n-1)} z^n p_{kj}$$

или

$$\sum_{n=0}^{\infty} p_{ij}^{(n)} z^n - p_{ij}^{(0)} = z\sum_{n=0}^{\infty}\sum_{k=1}^{N} p_{ik}^{(n)} z^n p_{kj} \cdot$$

Отсюда следует, что

$$\psi_{ij}(z) - p_{ij}^{(0)} = z\sum_{k=1}^{N} \psi_{ik}(z) p_{kj} \cdot$$

Подставив в это равенство (6.7), находим

$$\psi_{ij}(z) - z\sum_{k=1}^{N} \psi_{ik}(z) p_{kj} = \delta_{ij}, \; i,j=\overline{1,N} \cdot \tag{6.9}$$

Получим систему N^2 уравнений с N^2 неизвестными. Однако как в левую, так и в правую часть равенства (6.9) входит одно и то же i, поэтому можно отдельно решить N уравнений при фиксированном i. Обозначим через $\Delta(z)$ определитель этой системы (один и тот же для всех i). Имеем

$$\Delta(z) = \begin{vmatrix} 1-zp_{11} & -zp_{21} & \ldots & -zp_{N1} \\ -zp_{12} & 1-zp_{22} & \ldots & -zp_{N2} \\ . & . & . & . \\ -zp_{1N} & -zp_{2N} & \ldots & 1-zp_{NN} \end{vmatrix}.$$

При малом $|z|$ диагональные элементы близки к 1, а недиагональные – к 0, т.е. определитель $\Delta(z)$ близок к 1. Значит, $\Delta(z) \ne 0$

в некоторой окрестности точки $z=0$ и система уравнений (6.9) имеет единственное решение. Так как коэффициенты этих уравнений – линейные функции z, то

$$\psi_{ij}(z)=\frac{R_{ij}(z)}{S(z)}, \tag{6.10}$$

где $R_{ij}(z)$, $S(z)$ – некоторые полиномы. Выражение (6.10) после выделения целой части можно разложить на простейшие дроби вида

$\dfrac{\beta_r}{(1-\alpha_r z)^{r+1}}$, где α_r, β_r – некоторые комплексные постоянные, $r=0, 1, 2, ...\cdot$

Далее будем исходить из тождества

$$\sum_{n=0}^{\infty}\alpha^n z^n=\frac{1}{1-\alpha z}\cdot$$

Продифференцировав его r раз, получим

$$\sum_{n=r}^{\infty}n(n-1)\cdot...\cdot(n-r+1)\alpha^n z^{n-r}=\frac{\alpha^r r!}{(1-\alpha z)^{r+1}}$$

или, что то же самое,

$$\frac{1}{r!}\sum_{n=r}^{\infty}(n+r)(n+r-1)\cdot...\cdot(n+1)\alpha^n z^n=\frac{1}{(1-\alpha z)^{r+1}}\cdot$$

Таким образом, каждой элементарной дроби вида $\dfrac{\beta}{(1-\alpha z)^{r+1}}$, $r=1, 2, ...$, соответствует составляющая $p_{ij}^{(n)}$ вида $\dfrac{\beta}{r!}(n+r)(n+r-1)\cdot...\cdot(n+1)\alpha^n$; дроби $\dfrac{\beta}{1-\alpha z}$ соответствует составляющая $p_{ij}^{(n)}$ вида $\beta\alpha^n$. Следовательно, $p_{ij}^{(n)}$ равно конечной сумме всех указанных составляющих.

Описанным способом можно найти пределы $\lim\limits_{n\to\infty} p_{ij}^{(n)}$, если они существуют.

Пример 6.3. Рассмотрим конечную однородную цепь Маркова с двумя состояниями, матрица вероятностей переходов которой имеет вид

$$P=\begin{pmatrix}1-a & a\\ b & 1-b\end{pmatrix}.$$

Для определенности положим $i=1$. Тогда уравнения (6.9) принимают вид

$$\psi_{11}(z)-z\psi_{11}(z)p_{11}-z\psi_{12}(z)p_{21}=1,$$

$$\psi_{12}(z)-z\psi_{11}(z)p_{12}-z\psi_{12}(z)p_{22}=0\cdot$$

Определитель этой системы равен

$$\Delta(z)=\begin{vmatrix}(1-zp_{11}) & -zp_{21}\\ -zp_{12} & (1-zp_{22})\end{vmatrix}=(1-zp_{11})(1-zp_{22})-z^2p_{12}p_{21}=$$

$$=z^2(1-a-b)-z(2-a-b)+1\cdot$$

Один из корней уравнения $\Delta(z)=0$ равен 1, другой равен $\dfrac{1}{1-a-b}$.

Поэтому вероятности $p_{1j}^{(n)}$ должны иметь следующий вид:

$$p_{1j}^{(n)}=c_j1^n+\beta_j(1-a-b)^n=c_j+\beta_j(1-a-b)^n,\ j=1,2\cdot$$

Далее, по правилу Крамера, находим

$$\psi_{11}(z)=\frac{\Delta_1(z)}{\Delta(z)},\ \psi_{12}(z)=\frac{\Delta_2(z)}{\Delta(z)},$$

где

$$\Delta_1(z)=\begin{vmatrix}1 & -zp_{21}\\ 0 & 1-zp_{22}\end{vmatrix},$$

$$\Delta_2(z) = \begin{vmatrix} 1 - zp_{22} & 1 \\ -zp_{12} & 0 \end{vmatrix} = zp_{12} = az \cdot$$

Откуда следует, что

$$\psi_{11}(z) = \frac{1-(1-b)^2 z}{z^2(1-a-b) - z(2-a-b)+1} =$$

$$= \frac{a}{a+b} \cdot \frac{1}{1-z(1-a-b)} + \frac{b}{(a+b)(1-z)},$$

$$\psi_{12}(z) = \frac{az}{z^2(1-a-b) - z(2-a-b)+1} =$$

$$= -\frac{a}{a+b} \cdot \frac{1}{1-z(1-a-b)} + \frac{a}{a+b} \cdot \frac{1}{1-z},$$

таким образом,

$$p_{11}^{(n)} = \frac{b}{a+b} + \frac{a}{a+b}(1-a-b)^n, \ n \geq 1,$$

$$p_{12}^{(n)} = \frac{a}{a+b} - \frac{a}{a+b}(1-a-b)^n, \ n \geq 1 \cdot$$

Легко убедиться, что $p_{11}^{(n)} + p_{12}^{(n)} = 1$ при любом $n \geq 1$. Кроме того,

$$\lim_{n\to\infty} p_{11}^{(n)} = \frac{b}{a+b}, \ \lim_{n\to\infty} p_{12}^{(n)} = \frac{a}{a+b},$$

если только исключить тривиальные случаи, когда $a = b = 0$ либо $a = b - 1 \cdot$

§27. КЛАССИФИКАЦИЯ СОСТОЯНИЙ ЦЕПИ МАРКОВА ПО АРИФМЕТИЧЕСКИМ СВОЙСТВАМ ВЕРОЯТНОСТЕЙ ПЕРЕХОДА

Будем рассматривать цепь Маркова с дискретным временем и счетным числом состояний, $X = \{1, 2, ...\}$.

Определение. Состояние $i \in X$ называется **несущественным**, если из него с положительной вероятностью можно за конечное число шагов выйти, но нельзя в него вернуться, т.е. $\exists\, m,\ j$, что

$$p_{ij}^{(m)} > 0, \text{ но } \forall\, n,\ j \quad p_{ji}^{(n)} = 0 \cdot$$

Если из множества X выделить все несущественные состояния, то оставшееся множество **существенных** состояний обладает тем свойством, что, попав в него, цепь Маркова никогда из него не выйдет (рис. 24). Как станет ясно из дальнейшего, основной интерес представляют именно существенные состояния.

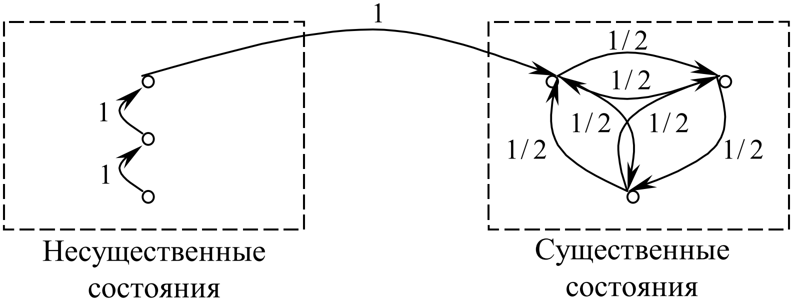


Рис. 24

Рассмотрим множество существенных состояний.

Определение. Состояние j называется **достижимым** из состояния i (обозначается $i \to j$), если $\exists\, n \geq 0$, что $p_{ij}^{(n)} > 0$ ($p_{ij}^{(0)} = \delta_{ij}$). Состояния i и j называются **сообщающимися** (обозначается $i \leftrightarrow j$), если j достижимо из i и i достижимо из j.

По определению отношение «$\leftrightarrow$» является симметричным ($i \leftrightarrow j \Rightarrow j \leftrightarrow i$), рефлексивным ($i \leftrightarrow i$) и нетрудно убедиться, что оно транзитивно ($i \leftrightarrow j,\ j \leftrightarrow k \Rightarrow i \leftrightarrow k$). Поэтому множество существенных состояний разбивается на конечное или счетное число непересекающихся множеств $X_1, X_2, \ldots$, состоящих из сообщающихся состояний и характеризующихся тем, что переходы между различными множествами невозможны.

Определение. Множества $X_1, X_2, \ldots$ называются **классами,** или **неразложимыми классами,** существенных сообщающихся состояний. Цепь Маркова, состояния которой образуют один неразложимый класс, называется **неразложимой**.

Пример 6.4. Рассмотрим цепь Маркова с множеством состояний $X = \{1, 2, 3, 4, 5\}$ и матрицей вероятностей переходов

$$P = \left(\begin{array}{cc|ccc} 1/3 & 2/3 & 0 & 0 & 0 \\ 1/4 & 3/4 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right) = \left(\begin{array}{c|c} P_1 & 0 \\ \hline 0 & P_2 \end{array}\right).$$

Граф переходов для этой цепи имеет вид (рис. 25).

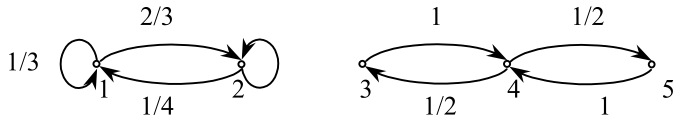


Рис. 25

Ясно, что у рассматриваемой цепи есть два неразложимых класса $X_1 = \{1, 2\}$ и $X_2 = \{3, 4, 5\}$, и исследование ее свойств сводится к исследованию свойств каждой из двух цепей, множествами состояний которых являются X_1 и X_2, а матрицы вероятностей переходов равны соответственно P_1 и P_2.

Проведенная классификация позволяет провести матрицу вероятностей переходов к каноническому виду. Для этого выделим неразложимые классы и перенумеруем их, а также отдельно выделим несущественные состояния. Тогда матрица P примет вид

$$P = \begin{array}{cc}
 & \begin{array}{ccccc} \text{I} & \text{II} & & k & \begin{array}{c}\text{несущ.}\\ \text{сост.}\end{array} \end{array} \\
\begin{array}{c} \text{I} \\ \text{II} \\ \\ k \\ \begin{array}{c}\text{несущ.}\\ \text{сост.}\end{array} \end{array} & \begin{pmatrix} P_1 & 0 & \cdots & 0 & 0 \\ 0 & P_2 & \cdots & 0 & 0 \\ . & & . & . & \\ 0 & 0 & \cdots & P_k & 0 \\ B_1 & B_2 & \cdots & B_k & R \end{pmatrix}
\end{array},$$

где P_s – матрица вероятностей переходов s -го неразложимого класса, B_s – матрица вероятностей переходов из несущественных состояний в s -й класс, $s = \overline{1, k}$; R – матрица вероятностей переходов по несущественным состояниям.

Рассмотрим теперь какой-либо неразложимый класс, изображенный на рис. 26.

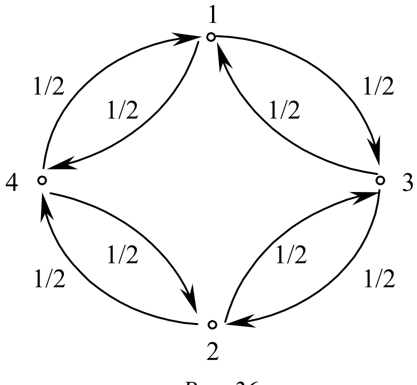


Рис. 26

Заметим, что здесь возвращение в каждое состояние возможно лишь за четное число шагов, переход в соседнее состояние – за нечетное число шагов, а матрица вероятностей переходов имеет блочную структуру:

$$P = \left(\begin{array}{cc|cc} 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \\ \hline 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{array}\right).$$

Отсюда видно, что класс $X_s = \{1, 2, 3, 4\}$ разбивается на два подкласса $C_0 = \{1, 2\}$ и $C_1 = \{3, 4\}$, обладающих следующим свойством цикличности: за один шаг цепь Маркова из C_0 непременно переходит в C_1, а из C_1 – в C_0. Этот пример показывает возможность разбиения неразложимых классов на **циклические подклассы**.

Определение. Будем говорить, что состояние j имеет период $d(j)$, если выполнены следующие два условия:

1) $p_{jj}^{(n)} > 0$ только для тех n, которые имеют вид $n = d(j)m$;

2) $d(j)$ есть наибольшее из чисел, обладающих свойством 1).

Иначе говоря, $d(j)$ есть наибольший общий делитель чисел n таких, что $p_{jj}^{(n)} > 0$ (если $p_{jj}^{(n)} = 0$ для всех $n \ge 1$, то полагаем $d(j) = 0$).

Теорема 6.2 ***(свойство периода состояния).*** *Если состояния i и j сообщающиеся ($i \leftrightarrow j$), то их периоды равны ($d(i) = d(j)$). Иными словами, все состояния одного неразложимого класса X имеют один и тот же период $d = d(X)$.*

Доказательство. Пусть $i, j \in X$, т.е. $i \leftrightarrow j$. Тогда $\exists\, n, l$, что $p_{ij}^{(n)} > 0$, $p_{ji}^{(l)} > 0$. Отсюда, используя уравнение Чепмена–Колмогорова, получаем

$$p_{ii}^{(n+l)} = \sum_k p_{ik}^{(n)} p_{ki}^{(l)} \ge p_{ij}^{(n)} p_{ji}^{(l)} > 0,$$

откуда следует, что $n+l$ делится на $d(i)$, т.е.

$$n+l=r_1 d(i)\cdot$$

Аналогично

$$p_{jj}^{(n+l)}=\sum_k p_{jk}^{(n)}p_{kj}^{(l)}\geq p_{ji}^{(n)}p_{ij}^{(l)}>0,$$

поэтому $n+l$ делится на $d(j)$, т.е.

$$n+l=r_2 d(j)\cdot$$

Таким образом, $n+l$ делится и на $d(j)$, и на $d(i)$.

Далее $\forall\ r>0$ имеем

$$p_{ii}^{(n+rd(j)+l)}=\sum_{k,m} p_{ik}^{(n)}p_{km}^{(rd(j))}p_{mi}^{(l)}\geq p_{ij}^{(n)}p_{jj}^{(rd(j))}p_{ji}^{(l)}>0,$$

и, так как $n+l$ делится на $d(i)$, то $rd(j)$ также должно делиться на $d(i)$. Аналогично, в силу симметрии, можно показать, что $\forall\ r'>0\ \ r'd(i)$ делится на $d(j)$. Поскольку r, r' – произвольные положительные целые числа, а $d(i)$ и $d(j)$ – наибольшие общие делители соответствующих чисел, то $d(i)=d(j)$.

Определение. Если $d(j)=1$ ($d(X)=1$), то состояние j (класс X) называется апериодическим (эргодическим).

Пусть $d=d(X)$ – период неразложимого класса X. Несмотря на сложность переходов внутри класса, можно обнаружить некоторую цикличность в переходах из одной группы состояний в другую. Чтобы показать это, выберем некоторое начальное состояние i_0 и введем для $d\geq 1$ следующие подклассы:

$$C_0=\left\{j\in X: p_{i_0 j}^{(n)}>0\Rightarrow n\equiv 0(\mathrm{mod}\ d)\right\},$$

$$C_1=\left\{j\in X: p_{i_0 j}^{(n)}>0\Rightarrow n\equiv 1(\mathrm{mod}\ d)\right\},$$

$$\ldots$$

$$C_{d-1}=\left\{j\in X: p_{i_0 j}^{(n)}>0\Rightarrow n\equiv d-1(\mathrm{mod}\ d)\right\}.$$

Ясно, что $X = C_0 + C_1 + ... + C_{d-1}$. Покажем, что за один шаг движение из подкласса в подклассе осуществляется так, как показано на рис. 27.

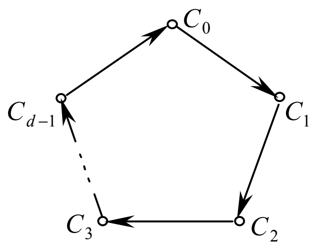
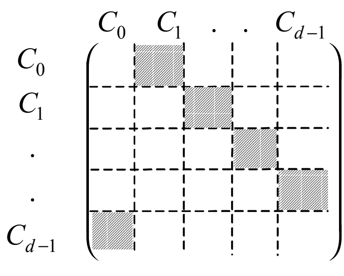


Рис. 27

В самом деле, пусть состояние $i \in C_p$ и $p_{ij} > 0$. Покажем, что тогда $j \in C_{p+1(\mathrm{mod}\, d)}$. Пусть n таково, что $p_{i_0 i}^{(n)} > 0$. Тогда $n = ad + p$, т.е. $n \equiv p(\mathrm{mod}\, d)$. Значит, $n+1 \equiv p+1(\mathrm{mod}\, d)$, откуда следует, что $p_{i_0 j}^{(n+1)} > 0$ и $j \in C_{p+1(\mathrm{mod}\, d)}$. Из-за переходов между подклассами, изображенными на рис. 27, подклассы C_p, $p = \overline{1, d-1}$, называют **циклическими подклассами.**

Заметим, что из приведенных рассуждений следует, что матрица вероятностей переходов неразложимого класса имеет блочную структуру



Рассмотрим некоторый подкласс C_p. Если в начальный момент времени система находилась в C_0, то в дискретные моменты времени $k = p + rd$, $r = 0, 1, \ldots$, она будет находиться в подклассе C_p. Поэтому с каждым подклассом C_p можно связать новую цепь Маркова с матрицей вероятностей переходов $\left\| p_{ij}^{(d)} \right\|$, $i, j \in C_p$, которая будет неразложимой и апериодической. Следовательно, в дальнейшем при рассмотрении предельных свойств вероятностей переходов $p_{ij}^{(n)}$ можно ограничиться рассмотрением только эргодических классов.

С учетом проведенной в данном параграфе классификации все состояния цепи Маркова можно расположить по следующей схеме (рис.28).

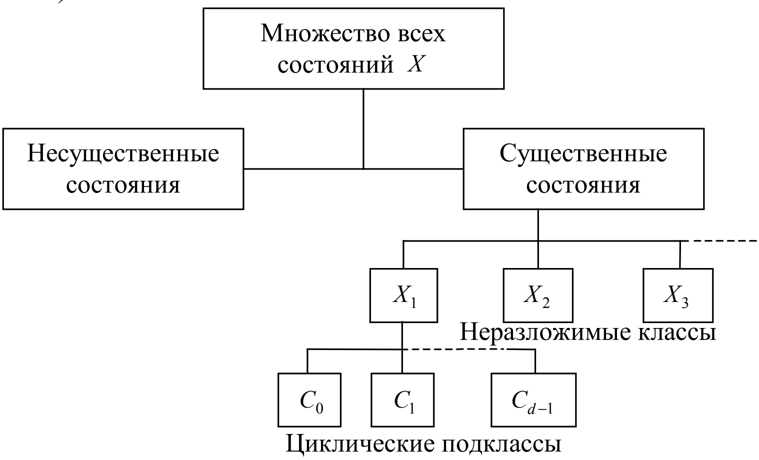


Рис.28

§28. КЛАССИФИКАЦИЯ СОСТОЯНИЙ ПО АСИМПТОТИЧЕСКИМ СВОЙСТВАМ ПЕРЕХОДНЫХ ВЕРОЯТНОСТЕЙ

Пусть $\{\xi_n\}$, $n = 0, 1, 2, \ldots$, – цепь Маркова с дискретным временем. Введем в рассмотрение следующие вероятности:

$f_{ii}^{(k)} = P\{\xi_k = i;\ \xi_l \neq i,\ 1 \le l \le k-1 / \xi_0 = i\}$ – вероятность первого возвращения цепи Маркова в состояние i в дискретный момент времени k (на k-м шаге), когда $\xi_0 = i$;

$f_{ij}^{(k)} = P\{\xi_k = j;\ \xi_l \neq j,\ 1 \le l \le k-1 / \xi_0 = i\}$ – вероятность первого попадания в состояние j в момент времени k, когда $\xi_0 = i$.

Используя марковское свойство, аналогично как при выводе уравнения Чепмена–Колмогорова, можно показать, что

$$p_{ij}^{(n)} = \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)} . \qquad (6.11)$$

Введем также для каждого состояния $i \in X$, $X = \{1, 2, \ldots\}$, величину

$$f_{ii} = \sum_{n=1}^{\infty} f_{ii}^{(n)} ,$$

которая является вероятностью того, что цепь Маркова, выходящая из состояния i, рано или поздно вернется в это состояние.

Определение. Состояние $i \in X$ называется **возвратным**, если $f_{ii} = 1$, и невозвратным, если $f_{ii} < 1$.

Каждое возвратное состояние можно в свою очередь отнести к одному из двух типов в зависимости от величины **среднего времени возвращения** (от его конечности или бесконечности). Величина

$$\mu_i = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$$

по определению математического ожидания равна среднему числу шагов, за которые цепь Маркова возвращается в состояние i, т.е. для цепи с дискретным временем характеризует среднее время возвращения в состояние i. Величина μ_i^{-1}, очевидно, характеризует интенсивность возвращения в состояние i.

Определение. Возвратное состояние i называется **положительным**, если $\mu_i^{-1} > 0$, и нулевым, если $\mu_i^{-1} = 0$.

Итак, в зависимости от свойств вероятностей $p_{ii}^{(n)}$ получаем классификацию состояний цепи, изображенную на рис. 6.8.

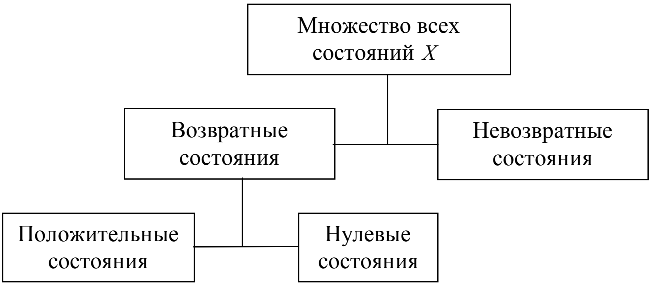


Рис. 29

Поскольку отыскание функций $f_{ii}^{(n)}$ довольно сложно, то для определения возвратности состояний полезен следующий критерий.

Теорема 6.3 *(критерий возвратности состояний).* *Состояние $i \in X$ возвратно тогда и только тогда, когда*

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty \cdot \tag{6.12}$$

Если состояние $i \in X$ возвратно и $i \leftrightarrow j$, то состояние $i \in X$ также возвратно.

Доказательство. Из соотношения (6.11) следует

$$p_{ii}^{(n)} = \sum_{l=1}^{n} f_{ii}^{(l)} p_{ii}^{(n-l)},$$

и значит,

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} = \sum_{n=1}^{\infty}\sum_{l=1}^{n} f_{ii}^{(l)} p_{ii}^{(n-l)}.$$

Поменять местами индексы суммирования нам поможет рис. 30, на котором линиями указано направление движения по целым точкам.

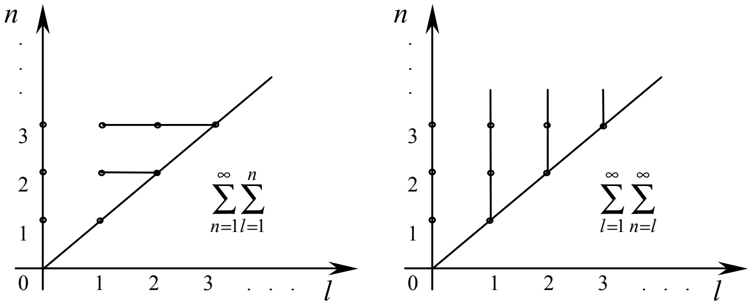


Рис. 30

Поменяв местами индексы суммирования, далее будем иметь

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} = \sum_{l=1}^{\infty} f_{ii}^{(l)} \sum_{n=l}^{\infty} p_{ii}^{(n-l)} = f_{ii} \sum_{n=0}^{\infty} p_{ii}^{(n)} =$$

$$= f_{ii}\left(p_{ii}^{(0)} + \sum_{n=1}^{\infty} p_{ii}^{(n)} \right) = f_{ii}\left(1 + \sum_{n=1}^{\infty} p_{ii}^{(n)} \right).$$

Отсюда следует, что если $\sum_{n=1}^{\infty} p_{ii}^{(n)} = A < \infty$, то $f_{ii} = \dfrac{A}{1+A} < 1$, и поэтому состояние i невозвратно.

Пусть выполняется соотношение (6.12). Тогда

$$\sum_{n=1}^{N} p_{ii}^{(n)} = \sum_{n=1}^{N}\sum_{l=1}^{n} f_{ii}^{(l)} p_{ii}^{(n-l)} =$$

$$= \sum_{l=1}^{n} f_{ii}^{(l)} \sum_{n=1}^{N} p_{ii}^{(n-l)} \le \sum_{l=1}^{N} f_{ii}^{(l)} \sum_{m=0}^{N} p_{ii}^{(m)} \text{ для } N \ge n.$$

Поэтому

$$\sum_{l=1}^{n} f_{ii}^{(l)} \ge \frac{\sum_{n=1}^{N} p_{ii}^{(n)}}{\sum_{m=0}^{N} p_{ii}^{(m)}},$$

откуда получаем

$$f_{ii}=\sum_{l=1}^{\infty} f_{ii}^{(l)} \geq \sum_{l=1}^{N} f_{ii}^{(l)} \geq \frac{\sum\limits_{n=1}^{N} p_{ii}^{(n)}}{\sum\limits_{m=0}^{N} p_{ii}^{(m)}} \xrightarrow[N\to\infty]{} 1 \cdot$$

Итак, если соотношение (6.12) выполнено, то $f_{ii}=1$, т.е. состояние i возвратно.

Докажем вторую часть теоремы. Пусть $p_{ij}^{(r)}>0$, $p_{ji}^{(l)}>0$. Тогда

$$p_{ii}^{(n+r+l)}=\sum_{k,\,m} p_{ik}^{(n)} p_{km}^{(r)} p_{mi}^{(l)} \geq p_{ij}^{(r)} p_{jj}^{(n)} p_{ji}^{(l)},$$

и если $\sum\limits_{n=1}^{\infty} p_{jj}^{(n)}=\infty$, то и соотношение (6.12) выполняется, т.е. состояние i возвратно. Теорема доказана.

Пример 6.5. Рассмотрим одномерное случайное блуждание частицы по целочисленным точкам действительной прямой (см. пример 5.6). За каждый переход частица перемещается на единицу вправо с вероятностью p и на единицу влево с вероятностью q, $p+q=1$. Следовательно, используя формулу Бернулли, получаем

$$p_{ii}^{(2n+1)}=0, \quad p_{ii}^{(2n)}=C_{2n}^{n} p^{n} q^{n}=\frac{(2n)!}{n!\,n!}(pq)^{n}, \quad n=0,\,1,\,2,\,\ldots$$

Воспользуемся формулой Стирлинга

$$n! \sim \sqrt{2\pi}\, n^{n+1/2} e^{-n} \cdot$$

Тогда

$$p_{ii}^{(2n)} \sim \frac{2^{2n}(pq)^{n}}{\sqrt{\pi n}}=\frac{(4pq)^{n}}{\sqrt{\pi n}}.$$

Легко проверить, что $pq = p(1-p) \le \frac{1}{4}$, причем равенство имеет место только тогда, когда $p = q = \frac{1}{2}$. В этом случае $p_{ii}^{(2n)} \sim \frac{1}{\sqrt{\pi n}}$. Поэтому ряд $\sum_{n=0}^{\infty} p_{ii}^{(2n)}$ расходится тогда и только тогда, когда $p = \frac{1}{2}$, и в данном случае все состояния являются возвратными. При $p \ne q$, когда $4pq < 1$ и $\sum_{n=0}^{\infty} p_{ii}^{(2n)} < \infty$, все состояния являются невозвратными. Интуитивно ясно, что в этом случае вероятность того, что частица, отправляясь из состояния i, будет смещаться к $+\infty$, если $p > q$, и к $-\infty$, если $p < q$, ни разу не возвращаясь в исходное состояние.

Пример 6.6. Обратимся теперь к двумерному случайному блужданию по двумерной целочисленной решетке. Пусть вероятности смещения частицы на единицу вправо, влево, вверх, вниз являются одинаковыми и равны $\frac{1}{4}$. Рассмотрим все траектории, состоящие из j перемещений вправо, j перемещений влево, k перемещений вверх, k перемещений вниз, $2j + 2k = 2n$, от состояния i. Легко убедиться, воспользовавшись полиномиальным распределением, что

$$p_{ii}^{(2n+1)} = 0, \quad n = 0, 1, 2, \ldots,$$

$$p_{ii}^{(2n)} = \sum_{\substack{j,k \\ j+k=n}} \frac{(2n)!}{j!\,j!\,k!\,k!} \left(\frac{1}{4}\right)^{2n}, \quad n = 1, 2, \ldots. \tag{6.13}$$

Умножая числитель и знаменатель в правой части выражения (6.13) на $(n!)^2$, получаем

$$p_{ii}^{(2n)} = \left(\frac{1}{4}\right)^{2n} C_{2n}^{n} \sum_{j=0}^{n} C_n^j \, C_n^{n-j} \; ,$$

но

$$\sum_{j=0}^{n} C_n^j \, C_n^{n-j} = C_{2n}^{n} \cdot$$

Следовательно,

$$p_{ii}^{(2n)} = \left(\frac{1}{4}\right)^{2n} \left(C_{2n}^{n}\right)^2 \cdot$$

Формула Стирлинга дает

$$p_{ii}^{(2n)} \sim \frac{1}{\pi n} \, .$$

Таким образом, соотношение (6.12), и в этом случае все состояния являются возвратными.

Теорема 6.4. *Если состояние* $j \in X$ *невозвратно, то для любого* $i \in X$

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty \, ,$$

и, значит, $\lim\limits_{n \to \infty} p_{ij}^{(n)} = 0$.

Доказательство. Из (6.11) и теоремы 6.3 следует

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} = \sum_{n=1}^{\infty} \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)} = \sum_{l=1}^{\infty} f_{ij}^{(l)} \sum_{n=0}^{\infty} p_{jj}^{(n)} =$$

$$= f_{ij} \sum_{n=0}^{\infty} p_{jj}^{(n)} \le \sum_{n=0}^{\infty} p_{jj}^{(n)} < \infty,$$

поскольку $f_{ij} = \sum_{l=1}^{\infty} f_{ij}^{(l)} \le 1$ – вероятность того, что цепь Маркова, вышедшая из состояния i, рано или поздно попадет в состояние j. Итак, ряд $\sum_{n=1}^{\infty} p_{ij}^{(n)}$ сходится и, следовательно, $p_{ij}^{(n)} \to 0$ при $n \to \infty$.

Покажем, что свойство возвратности состояния является свойством всего класса. Рассмотрим марковскую цепь с конечным числом состояний $X = \{1, 2, ..., N\}$.

Теорема 6.5. *Все состояния апериодической неразложимой цепи Маркова возвратны.*

Доказательство. Предположим, что все состояния невозвратны. Тогда в силу предыдущей теоремы и конечности множества состояний имеем

$$1 = \lim_{n\to\infty} \sum_{j=1}^{N} p_{ij}^{(n)} = \sum_{j=1}^{N} \lim_{n\to\infty} p_{ij}^{(n)} = 0 \cdot$$

Из этого противоречия следует, что все состояния не могут быть невозвратными.

Пусть состояние i возвратно, а j – произвольное состояние. Так как i и j сообщающиеся, то, по теореме 6.3, состояние j также возвратно, что и завершает доказательство.

Таким образом, все состояния конечного эргодического класса возвратны. Невозвратные состояния возможны только при бесконечном числе состояний.

§29. ЭРГОДИЧЕСКИЕ ЦЕПИ МАРКОВА

Мы уже рассматривали примеры (пример 6.2), в которых вероятности переходов $p_{ij}^{(n)}$ сходятся к предельным значениям π_j, не зависящим от i и образующим распределение вероятностей. Следующая теорема описывает широкий класс марковских цепей, обладающих так называемым свойством **эргодичности**: пределы $\lim\limits_{n\to\infty} p_{ij}^{(n)} = \pi_j$ не только существуют, не зависят от i, образуют распределение вероятностей, т.е. $\pi_j \geq 0$, $j \in X$, $\sum\limits_{j\in X} \pi_j = 1$, но и таковы, что $\pi_j > 0$ $\forall$ $j \in X$. Такие распределения $\{\pi_j\}$, $j \in X$, называются эргодическими. Сами вероятности называются финальными вероятностями состояний.

Определение. Цепь Маркова с дискретным временем называется **эргодической**, если

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j, \quad j \in X, \quad \sum_{j\in X} \pi_j = 1.$$

Теорема 6.6 ***(эргодическая теорема для цепей Маркова с дискретным временем и конечным числом состояний).*** *Пусть* $X = \{1, 2, ..., N\}$. *Тогда:*

1) если $\exists n_0$ такое, что

$$\min_{i,j} p_{ij}^{(n_0)} > 0, \tag{6.14}$$

то $\exists$ числа $\pi_1, \pi_2, ..., \pi_N$ такие, что

$$\pi_j > 0, \quad j = \overline{1, N}, \quad \sum_{j=1}^{N} \pi_j = 1 \tag{6.15}$$

и $\forall i \in X$

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j; \tag{6.16}$$

2) числа $\pi_1, \pi_2, ..., \pi_N$ удовлетворяют системе уравнений

$$\pi_j = \sum_{k=1}^{N} \pi_k p_{kj} = 1, \quad j = \overline{1, N}. \tag{6.17}$$

Доказательство. Докажем первую часть теоремы. Введем обозначения

$$m_j^{(n)} = \min_i p_{ij}^{(n)},\ M_j^{(n)} = \max_i p_{ij}^{(n)},\ i, j = \overline{1,\ N}\cdot$$

Поскольку

$$p_{ij}^{(n+1)} = \sum_{k=1}^{N} p_{ik} p_{kj}^{(n)},$$

то

$$m_j^{(n+1)} = \min_i p_{ij}^{(n+1)} = \min_i \sum_k p_{ik} p_{kj}^{(n)} \geq$$

$$\geq \min_i \sum_k p_{ik} \min_k p_{kj}^{(n)} = m_j^{(n)} \min_i \sum_k p_{ik} = m_j^{(n)},$$

откуда следует, что $m_j^{(n)} \leq m_j^{(n+1)}$. Аналогично можно показать, что $M_j^{(n)} \geq M_j^{(n+1)}$, т.е. наибольшая из вероятностей $p_{ij}^{(n)}$ с ростом n не может возрастать, а наименьшая не может убывать. Далее покажем, что

$$\max_{i,k} \left| p_{ij}^{(n)} - p_{kj}^{(n)} \right| \xrightarrow[n\to\infty]{} 0,$$

тогда (6.17) выполняется. Для этого достаточно показать, что

$$M_j^{(n)} - m_j^{(n)} \xrightarrow[n\to\infty]{} 0,\ j = \overline{1,\ N}\cdot \tag{6.18}$$

Пусть $\varepsilon = \min_{i,j} p_{ij}^{(n_0)}$, $\varepsilon > 0$ согласно (6.14). Тогда

$$p_{ij}^{(n_0+n)} = \sum_k p_{ik}^{(n_0)} p_{kj}^{(n)} = \sum_k \left[p_{ik}^{(n_0)} - \varepsilon p_{jk}^{(n)} \right] p_{kj}^{(n)} + \varepsilon \sum_k p_{jk}^{(n)} p_{kj}^{(n)} =$$

$$= \sum_k \left[p_{ik}^{(n_0)} - \varepsilon p_{jk}^{(n)} \right] p_{kj}^{(n)} + \varepsilon p_{jj}^{(2n)} \geq$$

$$\geq m_j^{(n)} \sum_k \left[p_{ik}^{(n_0)} - \varepsilon p_{jk}^{(n)} \right] + \varepsilon p_{jj}^{(2n)} = m_j^{(n)} (1-\varepsilon) + \varepsilon p_{jj}^{(2n)}\cdot$$

Отсюда следует, что

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

Аналогичным образом можно показать, что

$$M_j^{(n_0+n)} \le M_j^{(n)}(1-\varepsilon) + \varepsilon p_{jj}^{(2n)}.$$

Объединяя эти неравенства, получаем

$$M_j^{(n_0+n)} - m_j^{(n_0+n)} \le \left(M_j^{(n)} - m_j^{(n)}\right)(1-\varepsilon)$$

и, следовательно,

$$M_j^{(l\,n_0+n)} - m_j^{(l\,n_0+n)} \le \left(M_j^{(n)} - m_j^{(n)}\right)(1-\varepsilon)^l \xrightarrow[l\to\infty]{} 0.$$

Итак, для некоторой подпоследовательности $\{n_l\}$, $l = 1, 2, \ldots$,

$$M_j^{(n_l)} - m_j^{(n_l)} \le \xrightarrow[n_l\to\infty]{} 0,$$

и поскольку разность $M_j^{(n)} - m_j^{(n)}$ монотонна по n, то соотношение (6.18) выполняется, а следовательно, имеет место и соотношение (6.16).

Ясно также, что при $n \ge n_0$

$$m_j^{(n)} \ge m_j^{(n_0)} \ge \varepsilon > 0,$$

откуда вытекает, что $\pi_j > 0$, $j = \overline{1, N}$.

Докажем вторую часть теоремы. Запишем прямое уравнение Чепмена–Колмогорова

$$p_{ij}^{(n+1)} = \sum_{k=1}^{N} p_{ik}^{(n)} p_{kj}, \; i, j = \overline{1, N}.$$

Тогда

$$\lim_{n\to\infty} p_{ij}^{(n+1)} = \sum_{k=1}^{N} \lim_{n\to\infty} p_{ik}^{(n)} p_{kj}$$

и, используя (6.16), получаем систему уравнений (6.17). Теорема доказана.

Следствие. Для апериодического неразложимого класса с конечным числом состояний справедливо соотношение (6.16).

Если обозначить $\pi = (\pi_1, \pi_2, \ldots, \pi_N)$, то (6.17) можно переписать в матричной форме

$$\pi = \pi P, \tag{6.19}$$

где P – матрица вероятностей переходов цепи Маркова.

Рассмотрим еще одно важное следствие, вытекающее из эргодической теоремы. Пусть A – некоторая группа состояний цепи Маркова $\{\xi_n\}$, $A \subseteq X$, и $I_A(x)$ – индикатор множества A:

$$I_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

Рассмотрим величину

$$v_A(n) = \frac{I_A(\xi_0) + I_A(\xi_1) + \ldots + I_A(\xi_n)}{n+1},$$

характеризующую долю времени, проводимого цепью Маркова во множестве A. Поскольку

$$M\left[I_A(\xi_k)/\xi_0 = i\right] = P(\xi_k \in A/\xi_0 = i) = \sum_{j\in A} p_{ij}^{(k)} = p_i^{(k)}(A),$$

то

$$M\left[v_A(n)/\xi_0 = i\right] = \frac{1}{n+1}\sum_{k=0}^{n} p_i^{(k)}(A),$$

и, в частности,

$$M\left[v_{\{j\}}(n)/\xi_0 = i\right] = \frac{1}{n+1}\sum_{k=0}^{n} p_{ij}^{(k)}.$$

Из математического анализа известно, что если последовательность $x_n \xrightarrow[n\to\infty]{} x$, то $\dfrac{x_0 + x_1 + \ldots + x_n}{n+1} \xrightarrow[n\to\infty]{} x$. Поэтому, если $p_{ij}^{(k)} \xrightarrow[k\to\infty]{} \pi_j$, то

$$M\, v_{\{j\}}(n) \xrightarrow[n\to\infty]{} \pi_j, \quad M\, v_A(n) \xrightarrow[n\to\infty]{} \pi_A, \text{ где } \pi_A = \sum_{j\in A} \pi_j.$$

Для эргодических цепей можно доказать большее – для величин $I_A(\xi_0), I_A(\xi_1), \ldots, I_A(\xi_n), \ldots$ справедлив закон больших чисел.

Следствие (закон больших чисел). Для цепи Маркова с конечным числом состояний $\forall\ \varepsilon > 0$ и произвольного начального распределения

$$P\{|v_A(n) - \pi_A| > \varepsilon\} \xrightarrow[n\to\infty]{} 0 \cdot$$

Пример 6.7. Две автомашины A и B сдаются в аренду по одной и той же цене. Каждая из них может находиться в одном из двух состояний: i_1 – машина работает хорошо, i_2 – машина требует ремонта, которые образуют цепь Маркова. Матрицы вероятностей переходов между состояниями за сутки для этих машин равны соответственно

$$P_A = \begin{pmatrix} 0{,}8 & 0{,}2 \\ 0{,}7 & 0{,}3 \end{pmatrix},\ P_B = \begin{pmatrix} 0{,}7 & 0{,}3 \\ 0{,}6 & 0{,}4 \end{pmatrix}.$$

Определить финальные вероятности состояний для обеих автомашин. Какую автомашину стоит арендовать?

Решение. Существование финальных вероятностей состояний следует из условия (6.14). Найдем их для первой $\pi^{(1)} = \left(\pi_1^{(1)}, \pi_2^{(1)}\right)$ и для второй $\pi^{(2)} = \left(\pi_1^{(2)}, \pi_2^{(2)}\right)$ автомашины, используя соотношение (6.19):

$$\pi^{(1)} = \pi^{(1)} P_A,\ \pi^{(2)} = \pi^{(2)} P_B \cdot$$

Из первой системы уравнений имеем

$$\begin{cases} 0{,}8\pi_1^{(1)} + 0{,}7\pi_2^{(1)} = \pi_1^{(1)}, \\ 0{,}2\pi_1^{(1)} + 0{,}3\pi_2^{(1)} = \pi_2^{(1)}, \\ \pi_1^{(1)} + \pi_2^{(1)} = 1, \end{cases}$$

откуда получаем (любое из первых двух уравнений системы можно исключить) $\pi_1^{(1)} = \dfrac{7}{9}$, $\pi_2^{(1)} = \dfrac{2}{9}$. Аналогично для второй автомашины $\pi_1^{(2)} = \frac{2}{3}$, $\pi_2^{(2)} = \frac{1}{3}$. Отсюда следует, что первая автомашина будет более часто находиться в исправном состоянии, чем вторая, т.е. лучше арендовать первую автомашину.

§30. О СРЕДНИХ ВРЕМЕНАХ ПЕРЕХОДОВ МЕЖДУ СОСТОЯНИЯМИ

Обозначим через E_k множество состояний k-го эргодического класса, Q – множество несущественных состояний, m_i – среднее время перехода из i-го несущественного состояния, $i \in Q$, в k-й эргодический класс (среднее число шагов, за которые может осуществиться такой переход). Найдем это время.

Цепь Маркова за один шаг может перейти из i-го несущественного состояния в некоторое состояние $j \in E_k$ с вероятностью $\sum\limits_{j \in E_k} p_{ij}$, тогда время перехода равно 1. Если же цепь за первый шаг перейдет в другое состояние $l \in Q$ с вероятностью p_{il}, то среднее время перехода в эргодический класс будет равно $1 + m_l$, и окончательно имеем

$$m_i = 1 \cdot \sum_{j \in E_k} p_{ij} + \sum_{l \in Q} p_{il}(1 + m_l) =$$

$$= \sum_{j \in E_k} p_{ij} + \sum_{l \in Q} p_{il} + \sum_{l \in Q} m_l p_{il} = \sum_{j \in Q \cup E_k} p_{ij} + \sum_{l \in Q} m_l p_{il} \cdot \qquad (6.20)$$

Решая эту систему уравнений относительно m_i, $i \in Q$, получим нужные нам средние времена.

Если нас интересуют средние времена перехода в каждый эргодический класс, то систему уравнений (6.20) удобно записать в матричном виде. Обозначим

$$m = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_k \end{pmatrix}, \; e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \; R = \left\| p_{ij} \right\|, \; i, j \in Q \cdot$$

Учитывая, что в этом случае $\sum_{j \in Q \cup E_k} p_{ij} = 1$, имеем

$$m = e + Rm \text{, т.е. } m = (I - R)^{-1} e,$$

где I – единичная матрица. Последнее соотношение и определяет вектор среднего времени перехода из несущественного состояния в эргодические классы.

Рассмотрим теперь среднее время перехода из состояния в состояние внутри класса. Пусть E – множество состояний одного класса, $i, j \in E$. Через m_{ij} обозначим среднее число шагов, необходимых для переходов из i-го состояния в j-е состояние. За один шаг цепь может из i-го состояния с вероятностью p_{ij} перейти в j-е состояние, тогда время перехода будет равно 1, и оно будет равно $1 + m_{kj}$, если за этот шаг цепь перейдет из i-го состояния в некоторое промежуточное k-е состояние. Тогда по формуле для условного математического ожидания имеем

$$m_{ij} = 1 \cdot p_{ij} + \sum_{k \neq j} p_{ik}(m_{kj} + 1)$$

или

$$m_{ij} = 1 + \sum_{k} p_{ik} m_{kj} - p_{ij} m_{jj}. \tag{6.21}$$

Введем матричные обозначения

$$M = \|m_{ij}\|, \; i, j \in E, \; S = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ . & . & . & . \\ 1 & 1 & \dots & 1 \end{pmatrix}, D = \begin{pmatrix} m_{11} & 0 & 0 & \dots & 0 \\ 0 & m_{22} & 0 & \dots & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & . & m_{rr} \end{pmatrix},$$

где r – число состояний данного класса. Тогда систему уравнений (6.21) можно переписать в матричном виде

$$M = S + PM - PD,$$

откуда получаем

$$M = (I - P)^{-1}(SPD) \cdot$$

Эта формула определяет искомые средние значения времени перехода из одного состояния в другое.

Выясним теперь, что собой представляют диагональные элементы матрицы D. Для этого умножим систему уравнений (6.21) на финальную вероятность i-го состояния π_i и просуммируем по i:

$$\sum_i \pi_i m_{ij} = 1 + \sum_k m_{kj} \sum_i \pi_i p_{ik} - m_{jj} \sum_i \pi_i p_{ij} \cdot$$

Тогда с учетом формулы (6.17) имеем

$$\sum_i \pi_i m_{ij} = 1 + \sum_k \pi_k m_{kj} - \pi_j m_{jj},$$

откуда находим

$$m_{jj} = \frac{1}{\pi_j} \cdot$$

Таким образом, среднее время возвращения в исходное состояние обратно пропорционально финальной вероятности этого состояния. Это в свою очередь означает, что среднее время возвращения в положительное возвратное состояние конечно, в то время как для нулевых возвратных состояний оно бесконечно.

§31. СТАЦИОНАРНЫЕ ЦЕПИ МАРКОВА

Система уравнений (6.17) играет большую роль в теории марковских цепей. Всякое ее неотрицательное решение, удовлетворяющее условию нормировки, принято называть **стационарным** или **инвариантным** распределением вероятностей марковской цепи с матрицей переходных вероятностей $P = \left\| p_{ij} \right\|,\ i, j = \overline{1, N}$.

Определение. Набор чисел $\{p_j,\ j \in X\}$ называется стационарным распределением цепи Маркова с дискретным временем, если

1) $\{p_j,\ j \in X\}$ является распределением вероятностей, т.е.

$$p_j \ge 0,\ j \in X,\ \sum_{j \in X} p_j = 1;$$

2) $\forall\ j \in X$, $\forall\ n \geq 0$ имеет место равенство

$$p_j = \sum_{i \in X} p_i p_{ij}^{(n)} .$$

Рассмотрим смысл этого определения. Предположим, что на n-м шаге стационарное распределение достиглось, т.е. $p_j^{(n)} = p_j$, $\forall j \in X$, где $p_j^{(n)}$ – вероятность того, что цепь Маркова находится в состоянии j на n-м шаге. Тогда, используя формулу полной вероятности, марковское свойство и приведенное выше определение, $\forall m > n$ будем иметь

$$p_j^{(m)} = P(\xi_m = j) = \sum_{i \in X} P(\xi_n = i) P(\xi_m = j / \xi_n = i) =$$

$$= \sum_{i \in X} p_i^{(n)} p_{ij}^{(m-n)} = \sum_{i \in X} p_i p_{ij}^{(m-n)} = p_j, \ \forall j \in X,$$

т.е. стационарное распределение будет сохраняться на любом шаге $m > n$. В частности, если начальное распределение $\{p_j^{(0)}, j \in X\}$ совпадает со стационарным, то на любом шаге цепь Маркова будет иметь стационарное распределение, поэтому такая цепь Маркова называется **стационарной.**

То же самое произойдет, если в качестве начального взять эргодическое распределение $(\pi_1, \pi_2, ..., \pi_N)$, $p_j^{(0)} = \pi_j$ $j = \overline{1, N}$. Тогда

$$p_j^{(1)} = \sum_{i=1}^{N} \pi_i p_{ij} = \pi_j \quad j = \overline{1, N},$$

и вообще $p_j^{(n)} = \pi_j$, $j = \overline{1, N}$. Условие (6.14) эргодической теоремы 6.6 гарантирует как существование пределов $\pi_j = \lim_{n \to \infty} p_{ij}^{(n)}$, не зависящих от i, так и существование эргодического распределения, т.е. распределения с $\pi_j > 0$, $j = \overline{1, N}$. Распределение $(\pi_1, \pi_2, ..., \pi_N)$ оказывается также и стационарным распределением. Покажем, что набор $(\pi_1, \pi_2, ..., \pi_N)$ является единственным стационарным распределением. Пусть $(\tilde{\pi}_1, \tilde{\pi}_2, ..., \tilde{\pi}_N)$ – еще одно стационарное распределение. Тогда

$$\tilde{\pi}_j = \sum_{k=1}^{N} \tilde{\pi}_k p_{kj}^{(n)} \text{ и так как } p_{kj}^{(n)} \xrightarrow[n \to \infty]{} \pi_j \text{, то}$$

$$\tilde{\pi}_j = \sum_{k=1}^{N} \left(\tilde{\pi}_k \pi_j\right) = \pi_j, \quad j = \overline{1,N}.$$

Отметим, что стационарное распределение вероятностей (и к тому же единственное) может существовать и для неэргодических цепей. Действительно, если

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

то

$$P^{2n} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad P^{2n+1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

и поэтому пределы $\lim\limits_{n\to\infty} p_{ij}^{(n)}$ не существуют. В то же время система

$$\pi_j = \sum_{k=1}^{2} \pi_k p_{kj}, \quad j = 1, 2,$$

превращается в систему уравнений

$$\begin{cases} \pi_1 = \pi_2, \\ \pi_2 = \pi_1, \end{cases}$$

единственное решение которой (π_1, π_2), удовлетворяющее условию нормировки $\pi_1 + \pi_2 = 1$, есть $\left(\frac{1}{2}, \frac{1}{2}\right)$.

Приведем без доказательства ряд теорем, касающихся эргодических и стационарных цепей Маркова с дискретным временем и часто используемых при решении различных задач. Для марковской цепи с конечным числом состояний справедливо следующее утверждение.

Теорема 6.7 ***(эргодическая теорема Маркова–Бернштейна).*** *Если существуют такие $j \in X$ и $n \ge 1$, что $\forall\ i \in X$ выполняется неравенство*

$$p_{ij}^{(n)} \ge \lambda > 0,$$

то цепь Маркова является эргодической; существует единственное стационарное распределение этой цепи, совпадающее с эргодическим.

Определение. Цепь Маркова с дискретным временем называется неприводимой, если $\forall i, j \in X$ $\exists n > 0$ такое, что $p_{ij}^{(n)} > 0$.

Для неприводимых марковских цепей с конечным или счетным числом состояний справедливы следующие три утверждения.

Теорема 6.8 ***(эргодическая теорема Феллера).*** *Неприводимая апериодическая цепь Маркова относится к одному из следующих двух классов:*

а) все состояния цепи невозвратны (либо нулевые), в этом случае $\forall\ i, j \in X$ $p_{ij}^{(n)} \xrightarrow[n\to\infty]{} 0$, *и не существует стационарного распределения цепи;*

б) все состояния цепи положительны, в этом случае $p_{ij}^{(n)} \xrightarrow[n\to\infty]{} p_j > 0$, при этом $\{p_j, j \in X\}$ является стационарным распределением цепи и не существует других ее стационарных распределений.

Из теоремы Феллера следуют два нижеприведенные утверждения.

Теорема 6.9 ***(эргодическая теорема Фостера).*** *Для того чтобы неприводимая апериодическая цепь Маркова была эргодична, необходимо и достаточно, чтобы система уравнений*

$$x_j = \sum_{j \in X} x_i p_{ij}, j \in X,$$

имела нетривиальное решение $(x_i, i \in X)$, такое, что $\sum_{j \in X} |x_i| < \infty$. При этом существует единственное стационарное распределение, которое совпадает с эргодическим.

Теорема 6.10 ***(эргодическая теорема Маркова).*** *Для тог, чтобы неприводимая апериодическая цепь Маркова была эргодична, достаточно существования* $\varepsilon > 0$, *натурального числа* i_0 *и набора неотрицательных чисел* $x_0, x_1, x_2, \ldots$ *таких, что*

$$\sum_{j \in X} p_{ij} x_j \le x_i - \varepsilon \text{ для } i > i_0,$$

$$\sum_{j\in X} p_{ij} x_j < \infty \text{ для } i \le i_0 .$$

При этом существует единственное стационарное распределение, совпадающее с эргодическим.

Пример 6.8. На письменном столе лежит стопка из m книг. Если обозначить каждую книгу соответствующим номером, то порядок их расположения сверху вниз можно описать перестановкой из m чисел $(i_1, i_2, ..., i_m)$, где i_1 – номер книги, лежащей сверху, i_2 – номер следующей книги, …, i_m – номер последней книги, лежащей в самом низу. Предположим, что книга с номером k берется с вероятностью α_k, $k = \overline{1, m}$, причем при возвращении она кладется сверху.

Рассмотрим произвольное состояние $(i_1, i_2, ..., i_m)$. На следующем шаге оно либо останется неизменным, что происходит с вероятностью α_{i_1} при выборе лежащей сверху книги с номером i_1, либо меняется на одно из $m-1$ состояний вида $(i_k, i_1, ...)$, что происходит с вероятностью α_{i_k} при выборе книги с номером $i_k = i_1$.

В данном примере мы имеем цепь Маркова с состояниями, каждое из которых описывается соответствующей перестановкой $(i_1, i_2, ..., i_m)$ и указанными переходными вероятностями. Если некоторая книга i никогда не выбирается $(\alpha_i = 0)$, то все состояния $(i_1, i_2, ..., i_m)$, где $i_1 = i$ (книга с номером i лежит сверху), являются невозвратными, поскольку после первого же шага выбирается некоторая j-я книга, $j \ne i$, и в дальнейшем i-я книга, никогда не вынимаемая из стопки, опускается ниже. Если каждая книга выбирается с положительной вероятностью $\alpha_i > 0$, то каждое состояние достижимо из любого другого состояния (всего различных состояний, т.е. различных перестановок $(i_1,\ i_2, ...,\ i_m)$ имеется $m!$), и все эти сообщающиеся состояния являются положительными, образуя один неразложимый класс. Из каждого состояния $(i_1,\ i_2, ...,\ i_m)$ можно перейти за m шагов в любое состояние $(j_1, j_2, ..., j_m)$ с положительной вероятностью, превосходящей произведение $\alpha_1 \cdot \alpha_2 \cdot ... \cdot \alpha_m$ (это произведение равно вероятности перехода из $(i_1, i_2, ..., i_m)$ в $(j_1, j_2, ..., j_m)$, когда на первом шаге выбирается книга с номером j_m, на втором шаге – книга с номером j_{m-1} и т.д., на последнем m-м шаге выбирается книга с номером j_1).

Следовательно, выполняется условие (6.14) эргодической теоремы 6.6 и с течением времени устанавливается стационарное распределение вероятностей.

Рассмотрим вначале случай $m = 2$. Тогда имеется лишь два состояния $(1,2)$ и $(2,1)$. Вероятности переходов имеют вид

$$p_{11} = p_{21} = \alpha_1, \; p_{12} = p_{22} = \alpha_2,$$

а матрица вероятностей переходов

$$P = \begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_1 & \alpha_2 \end{pmatrix}.$$

Найдем вероятности перехода за два шага, используя соотношение $P^{(2)} = P^2$:

$$p_{11}^{(2)} = p_{21}^{(2)} = \alpha_1^2 + \alpha_1\alpha_2 = \alpha_1(\alpha_1 + \alpha_2) = \alpha_1,$$

$$p_{12}^{(2)} = p_{22}^{(2)} = \alpha_1\alpha_2 + \alpha_2^2 = \alpha_2(\alpha_1 + \alpha_2) = \alpha_2.$$

Ясно, что $P^{(n)} = P^n$. При любом начальном распределении вероятностей $\{p_1^{(0)}, p_2^{(0)}\}$ имеем

$$p_1^{(n)} = p_1^{(0)} p_{11}^{(n)} + p_2^{(0)} p_{21}^{(n)} = \alpha_1 (p_1^{(0)} + p_1^{(0)}) = \alpha_1,$$

$$p_2^{(n)} = p_1^{(0)} p_2^{(n)} + p_2^{(0)} p_{22}^{(n)} = \alpha_2 (p_1^{(0)} + p_1^{(0)}) = \alpha_2,$$

т.е. стационарное распределение устанавливается уже на втором шаге.

Рассмотрим теперь случай произвольного m. Обозначим через

$$p_{(i_1, i_2, \dots, i_m)(j_1, j_2, \dots, j_m)}$$

вероятность перехода из состояния $(i_1, i_2, \dots, i_m)$ в состояние $(j_1, j_2, \dots, j_m)$. Как было показано,

$$p_{(i_1, i_2, \dots, i_m)(j_1, j_2, \dots, j_m)} = \begin{cases} p_{i_k} \text{ при } (j_1, j_2, \dots, j_m) = (i_k, \dots), \\ 0 \text{ при остальных } (j_1, j_2, \dots, j_m), \end{cases}$$

где перестановка $(i_k,\ldots)$ получается из $(i_1,i_2,\ldots,i_m)$ выбором некоторого i_k и перестановкой его на первое место. Стационарные вероятности $p_{(i_1,i_2,\ldots,i_m)}$, согласно (6.17), являются решением следующей системы линейных уравнений

$$p_{(j_1,j_2,\ldots,j_m)} = \alpha_{j_1} \sum_{k=1}^{m} p_{(j_2,\ldots,j_{k-1},j_1,j_k,\ldots)} .$$

Интересно выяснить, с какой вероятностью каждая из имеющихся книг оказывается лежащей сверху, когда через достаточно большое число шагов практически устанавливается стационарное распределение вероятностей (т.е. когда стопка книг будет с неизменными вероятностями $p_{(i_1,i_2,\ldots,i_m)}$ занимать соответствующие положения $(i_1,i_2,\ldots,i_m)$).

Вероятность того, что сверху лежит книга с номером i, очевидно, равна

$$p_i = \sum_{i_2,\ldots,i_m} p_{(i,i_2,\ldots,i_m)} ,$$

где сумма берётся по всем состояниям, в которых на первом месте стоит i. Из уравнения для стационарных вероятностей состояний получаем, что

$$p_i = \sum_{i_2,\ldots,i_m} \alpha_i \sum_k p_{(i_2,\ldots,i_{k-1},i,i_k,i_m)} =$$

$$= \alpha_i \sum_{i_2,\ldots,i_m} p_{(i_1,i_2,\ldots,i_m)} = \alpha_i, \; i = \overline{1,m},$$

т.е. вероятность p_i того, что сверху находится книга с номером i, равна вероятности α_i, с которой эта книга выбирается. Таким образом, чем чаще берётся та или иная книга, тем с большей вероятностью она будет находиться сверху.

§32. ОПТИМАЛЬНЫЕ СТРАТЕГИИ В МАРКОВСКИХ ЦЕПЯХ

Рассмотрим цепь Маркова с конечным множеством состояний $X = \{i_1, i_2, ..., i_N\}$ и матрицей вероятностей переходов $P = \left\| p_{ij} \right\|_{N \times N}$. Пусть при переходе из состояния i_j в состояние i_k мы получаем некоторое вознаграждение r_{jk}. Например, пусть i_j, $j = \overline{1, N}$ – различные места в городе; переход из i_j в i_k означает перевозку пассажиров такси по данному маршруту; тогда r_{jk} – это прибыль, входящая в плату за проезд из (i_j) в (i_k). Найдем общее вознаграждение, которое следует ожидать после переходов цепи Маркова за n шагов.

Предположим, что цепь Маркова находится в состоянии i_j и пусть $\nu_j(n)$ – суммарное среднее вознаграждение за n шагов, если вначале цепь находилась в состоянии i_j. Тогда

$$q_j = \sum_{k=1}^{N} r_{jk} p_{jk} \quad –$$

средний выигрыш после первого шага. Если после первого шага цепь Маркова находится в состоянии i_l, то средний выигрыш на оставшихся $n-1$ шагах составит $\nu_l(n-1)$. Поэтому полный средний выигрыш за n шагов равен

$$\nu_j(n) = q_j + \sum_{l=1}^{N} p_{jl} \nu_l(n-1) \cdot$$

Введем матричные обозначения

$$V(n) = \begin{pmatrix} \nu_1(n) \\ \nu_2(n) \\ \cdot \\ \cdot \\ \cdot \\ \nu_N(n) \end{pmatrix}, \quad Q = \begin{pmatrix} q_1 \\ q_2 \\ \cdot \\ \cdot \\ \cdot \\ q_N \end{pmatrix}.$$

Тогда в матричной форме последнее соотношение может быть переписано как

$$V(n) = Q + PV(n-1) .$$

В стационарном режиме среднее вознаграждение за один шаг может быть найдено в виде

$$g = \sum_{j=1}^{N} g_j \pi_j , \qquad (6.22)$$

где π_j, $j = \overline{1, N}$, – финальные вероятности состояний.

Пример 6.9. Автомашина может находиться в двух состояниях: i_1 – работает хорошо, i_2 – требует ремонта. На следующий день работы она меняет свое состояние в соответствии с матрицей вероятностей переходов

$$P = \begin{pmatrix} 0{,}7 & 0{,}3 \\ 0{,}6 & 0{,}4 \end{pmatrix}.$$

Пусть, если машина работает нормально, мы имеем прибыль $\$40$; когда она начинает работу в нормальном состоянии, а затем требует ремонта (либо наоборот), прибыль равна $\$20$; если машина требует ремонта, то потери составляют $\$20$. В данном случае

$$R = \left\| r_{jk} \right\|_{2\times 2} = \begin{pmatrix} 40 & 20 \\ 20 & -20 \end{pmatrix},$$

$q_1 = 40 \cdot 0{,}7 + 20 \cdot 0{,}3 = 34$, $q_2 = 20 \cdot 0{,}6 - 20 \cdot 0{,}4 = 4$.

Найдем ожидаемую прибыль за два перехода между состояниями (за два шага). Так как $V(1) = \begin{pmatrix} 34 \\ 4 \end{pmatrix}$, то

$$V(2) = \begin{pmatrix} 34 \\ 4 \end{pmatrix} + \begin{pmatrix} 0{,}7 & 0{,}3 \\ 0{,}6 & 0{,}4 \end{pmatrix} V(1) = \begin{pmatrix} 59 \\ 26 \end{pmatrix}.$$

Таким образом, ожидаемая прибыль за два перехода составит $\$59$, если процесс начал развиваться из состояния i_1, и $\$26$, если процесс начал развиваться из состояния i_2.

Найдем финальные вероятности состояний $\pi = (\pi_1, \pi_2)$ из соотношений $\pi = \pi P$, $\pi_1 + \pi_2 = 1$, откуда получаем $\pi_1 = \frac{2}{3}$, $\pi_2 = \frac{1}{3}$. Далее, используя (6.22), находим стационарное ожидаемое вознаграждение:

$$q = 34 \cdot \frac{2}{3} + 4 \cdot \frac{1}{3} = 24 \cdot$$

Следовательно, если автомашина работает в течение многих переходов между состояниями, и неизвестно ее текущее состояние, то ожидаемая прибыль за один шаг процесса равна $\$24$.

ГЛАВА 7. ЦЕПИ МАРКОВА С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

§31. НЕКОТОРЫЕ ОПРЕДЕЛЕНИЯ И СВОЙСТВА

Рассмотрим марковский процесс $\xi(t)$ с конечным или счетным множеством состояний X, который изменяет свои состояния в произвольные моменты времени непрерывного интервала T. Напомним, что такой процесс называется цепью Маркова с непрерывным временем. Очевидно, что для такой цепи Маркова

$$P\big(\xi(t_2) = j / \xi(s) = k, \xi(t_1) = i\big) = P\big(\xi(t_2) = j / \xi(t_1) = i\big) = p_{ij}(t_1, t_2), \quad (7.1)$$

где $i, j \in X,\ s, t_1, t_2 \in T,\ s < t_1 < t_2$.

Определение. Вероятность $p_{ij}(t_1, t_2)$ называется вероятностью перехода из состояния i в состояние j за промежуток времени $[t_1, t_2)$,

$$\sum_{j \in X} p_{ij}(t_1, t_2) = 1.$$

В дальнейшем в основном будем рассматривать случай, когда $T = [0, +\infty)$.

Теорема 7.1 ***(уравнение Чепмена–Колмогорова).*** *Переходные вероятности* $p_{ij}(t_1, t_2)$ *удовлетворяют уравнению*

$$p_{ij}(t_1, t_3) = \sum_k p_{ik}(t_1, t_2)\, p_{kj}(t_2, t_3), \quad t_1 < t_2 < t_3. \quad (7.2)$$

Доказательство. Используя определение условной вероятности и то, что события $\big(\xi(t_2) = k\big)$, $k = 1, 2, \ldots$, образуют полную группу, а значит, $\bigcup_k \big(\xi(t_2) = k\big) = \Omega$, будем иметь

$$p_{ij}(t_1, t_3) = P\big(\xi(t_3) = j / \xi(t_1) = i\big) = P\big(\big(\xi(t_3) = j\big) \cap \Omega / \xi(t_1) = i\big) =$$

$$= P\left((\xi(t_3) = j) \cap \left(\bigcup_k (\xi(t_2) = k) \right) / \xi(t_1) = i \right) =$$

$$= \sum_k P(\xi(t_3) = j, \xi(t_2) = k / \xi(t_1) = i) =$$

$$= \sum_k \frac{P(\xi(t_3) = j, \xi(t_2) = k, \xi(t_1) = i)}{P(\xi(t_1) = i)} =$$

$$= \sum_k \frac{P(\xi(t_3) = j / \xi(t_2) = k, \xi(t_1) = i) P(\xi(t_2) = k, \xi(t_1) = i)}{P(\xi(t_1) = i)} =$$

$$= \sum_k P(\xi(t_3) = j / \xi(t_2) = k, \xi(t_1) = i) P(\xi(t_2) = k / \xi(t_1) = i) \cdot$$

Далее, используя марковское свойство (7.1), получаем

$$p_{ij}(t_1, t_3) = \sum_k P(\xi(t_3) = j / \xi(t_2) = k) P(\xi(t_2) = k / \xi(t_1) = i) =$$

$$= \sum_k p_{ik}(t_1, t_2)\, p_{kj}(t_2, t_3),$$

что и требовалось доказать.

Напомним, что если вероятности переходов $p_{ij}(t_1, t_2)$ зависят только от разности моментов времени, т.е.

$$p_{ij}(t_1, t_2) = p_{ij}(t_2 - t_1),$$

то цепь Маркова называется однородной.

Для однородной цепи Маркова выражение (7.2) может быть переписано в виде

$$p_{ij}(t_3 - t_1) = \sum_k p_{ik}(t_2 - t_1)\, p_{kj}(t_3 - t_2),$$

или

$$p_{ij}(t + \tau) = \sum_k p_{ik}(t)\, p_{kj}(\tau) \cdot \qquad (7.3)$$

Таким образом, в этом случае переходные вероятности обладают следующими свойствами:

1) $p_{ij}(t) \geq 0$;

2) $\sum_{j\in X} p_{ij}(t) = 1$;

3) вероятности $p_{ij}(t)$ равномерно непрерывны по t (примем без доказательства, вернемся к нему в §34);

4) выполняется условие стохастической непрерывности процесса

$$\lim_{t\to 0} p_{ij}(t) = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Это условие означает, что с вероятностью 1 цепь Маркова не изменит своего состояния за бесконечно малый промежуток времени $t \to 0$. Следует отметить, что стохастическая непрерывность не означает непрерывность реализаций марковской цепи. Это особенно очевидно в данном случае, так как множество ее значений дискретно, т.е. реализации разрывны, тем не менее условие стохастической непрерывности выполнено. Это происходит потому, что разрывы каждой реализации цепи происходят в случайные моменты времени, и вероятность того, что разрыв произойдет именно в данный момент времени t, равна нулю.

Определение. Цепь с непрерывным временем называется **неприводимой**, если $\forall\ i, j \in X\ \ \exists\ \ 0 < t_{ij} < \infty$, такое, что $p_{ij}(t_{ij}) > 0$.

Сформулируем еще несколько утверждений относительно вероятностей $p_{ij}(t)$.

Теорема 7.2. *Для любого состояния $i \in X$ вероятность возвращения в исходное состояние строго больше нуля при всех $t > 0$, т.е.*

$$p_{ii}(t) > 0 \quad \forall\ i \in X \text{ и } \forall\ t > 0 \cdot$$

Доказательство. В силу стохастической непрерывности найдем такое $p > 0$, что $\forall\ t \le h$

$$p_{ii}(t) > 0.$$

Пусть t – произвольный момент времени. Всегда $\exists\ n$, такое, что $\dfrac{t}{n} < h$. Тогда, используя уравнение Чепмена–Колмогорова, получаем

$$p_{ii}(t) = \sum_{\{k_i,\, i=\overline{1,n}\}} p_{ik_1}\left(\frac{t}{n}\right) p_{k_1k_2}\left(\frac{t}{n}\right) \cdot \ldots \cdot p_{k_ni}\left(\frac{t}{n}\right) \ge \left[p_{ii}\left(\frac{t}{n}\right)\right]^n > 0 \cdot$$

Теорема 7.3. *Если* $\exists\ t_0$, *что* $p_{ij}(t_0) > 0$, *то* $p_{ij}(t) > 0\ \forall\ t > 0$.

Доказательство также вытекает из уравнения Чепмена–Колмогорова и предыдущей теоремы:

$$p_{ij}(t) = \sum_k p_{ik}(t - t_0)\, p_{kj}(t_0) \ge p_{ii}(t - t_0)\, p_{ij}(t_0) > 0 \cdot$$

§34. СИСТЕМЫ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ КОЛМОГОРОВА ДЛЯ ОДНОРОДНОЙ ЦЕПИ МАРКОВА С КОНЕЧНЫМ ЧИСЛОМ СОСТОЯНИЙ

В общем случае вероятности перехода $p_{ij}(t)$ однородной марковской цепи для произвольного $t \ge 0$ удовлетворяют системам дифференциальных уравнений, вывод которых мы опишем в данном параграфе.

Определение. Пусть

$$\lambda_i = \lim_{\Delta t \to 0} \frac{1 - p_{ii}(\Delta t)}{\Delta t}, \quad \lambda_{ij} = \lim_{\Delta t \to 0} \frac{p_{ij}(\Delta t)}{\Delta t},\ i \ne j, \quad (7.4)$$

причем $0 \le \lambda_i \le +\infty$, $0 \le \lambda_{ij} < +\infty$. λ_{ij} называется **интенсивностью перехода** цепи Маркова из состояния i в состояние j, λ_i называется **интенсивностью выхода** из состояния i, $i, j \in X$.

Определение. Если $\lambda_i = \infty$, то состояние i называется **мгновенным**; если $\lambda_i < \infty$, то состояние i называется **немгновенным,** или **задерживающим**. Если $\lambda_i = 0$, то состояние i называется **поглощающим**.

Можно показать, что всегда $\lambda_i \geq \sum_{j \neq i} \lambda_{ij}$, $i \in X$.

Определение. Немгновенное состояние i называется **регулярным**, если $\lambda_i = \sum_{j \neq i} \lambda_{ij}$ и **нерегулярным**, если $\lambda_i > \sum_{j \neq i} \lambda_{ij}$. Если все состояния цепи Маркова регулярные, то она называется **регулярной**.

Из соотношения (7.4) следует, что

$$p_{ii}(\Delta t) = 1 - \lambda_i \Delta t + o(\Delta t),$$

$$p_{ij}(\Delta t) = \lambda_{ij} \Delta t + o(\Delta t), \ i \neq j.$$

Теорема 7.4. *Для однородной цепи Маркова с конечным числом состояний переходные вероятности* $p_{ij}(t)$ *удовлетворяют системам дифференциальных уравнений*

$$p'_{ij}(t) = -\lambda_j p_{ij}(t) + \sum_{k \neq j} \lambda_{kj} p_{ik}(t), \ i, j \in X, \tag{7.6}$$

$$p'_{ij}(t) = -\lambda_i p_{ij}(t) + \sum_{k \neq i} \lambda_{ik} p_{kj}(t), \ i, j \in X \tag{7.7}$$

с начальными условиями

$$p_{ij}(0) = \delta_{ij}, \ i, j \in X. \tag{7.8}$$

Система (7.6) называется системой прямых дифференциальных уравнений, а система (7.7) – системой обратных дифференциальных уравнений Колмогорова.

Доказательство. Из уравнения Чепмена–Колмогорова (7.3) следует, что

$$p_{ij}(t + \Delta t) = \sum_k p_{ik}(t) p_{kj}(\Delta t).$$

Выделим в сумме с правой стороны слагаемое при $k = j$. Тогда, используя асимптотические выражения (7.5), получим

$$p_{ij}(t + \Delta t) = p_{ij}(t)\left[1 - \lambda_i \Delta t + o(\Delta t)\right] +$$

$$+\sum_{k\neq j} p_{ik}(t)\left[\lambda_{kj}\,\Delta t+o(\Delta t)\right].$$

Следовательно

$$\frac{p_{ij}(t+\Delta t)-p_{ij}(t)}{\Delta t}=p_{ij}(t)\left[-\lambda_i+\frac{o(\Delta t)}{\Delta t}\right]+$$

$$+\sum_{k\neq j} p_{ik}(t)\left[\lambda_{kj}+\frac{o(\Delta t)}{\Delta t}\right].$$

Устремив Δt к нулю, будем иметь систему уравнений (7.6).

Аналогично можно поступить при выводе системы уравнений (7.7):

$$p_{ij}(t+\Delta t)=\sum_k p_{ik}(\Delta t)\,p_{kj}(t)=$$

$$=\left[1-\lambda_i\,\Delta t+o(\Delta t)\right]p_{ij}(t)+\sum_{k\neq i}\left[\lambda_{ik}\,\Delta t+o(\Delta t)\right]p_{kj}(t),$$

$$\frac{p_{ij}(t+\Delta t)-p_{ij}(t)}{\Delta t}=\left[-\lambda_i+\frac{o(\Delta t)}{\Delta t}\right]p_{ij}(t)+$$

$$+\sum_{k\neq i}\left[\lambda_{ik}+\frac{o(\Delta t)}{\Delta t}\right]p_{kj}(t)\ .$$

Устремляя Δt к нулю, получаем систему уравнений (7.7).

Отметим, что решения систем прямых и обратных уравнений при одних и тех же начальных условиях (7.8) совпадают.

Отметим также, что система прямых дифференциальных уравнений имеет место и для безусловных вероятностей состояний $p_j(t)$:

$$p'_j(t)=-\lambda_j\,p_j(t)+\sum_{k\neq j}\lambda_{kj}\,p_k(t),\ \ j\in X \tag{7.9}$$

при любом начальном распределении $\{p_j(0),\ j\in X\}$. Это следует из формулы полной вероятности и соотношений (7.5):

$$p_j(t+\Delta t)=\sum_k p_k(t)\,p_{kj}(\Delta t)=$$

$$= p_j(t)\left[1 - \lambda_j \Delta t + o(\Delta t)\right] + \sum_{k \neq j} p_k(t)\left[\lambda_{kj} \Delta t + o(\Delta t)\right].$$

Следует также заметить, что в случае, когда множество состояний цепи Маркова X счетно, проведенные рассуждения не всегда правильны, поскольку сумма счетного числа величин порядка $o(\Delta t)$ не обязательно дает $o(\Delta t)$. В этом случае для справедливости систем уравнений Колмогорова приходится накладывать дополнительные условия. Например, использованный выше прием для вывода системы прямых уравнений (7.6) остается в силе, если в асимптотических выражениях (7.5) остаточные члены $o(\Delta t)$ таковы, что $\frac{o(\Delta t)}{\Delta t} \underset{\Delta t \to 0}{\to} 0$ равномерно по всем i, причем для фиксированного j интенсивности λ_{ij} равномерно ограничены:

$$\lambda_{ij} \leq \Lambda_j < \infty, \ i \in X.$$

Предположим, что существует стационарное распределение цепи Маркова $\{p_j,\ j \in X\}$, где $p_j = \lim_{t \to \infty} p_j(t)$. Поскольку стационарные вероятности не зависят от времени и являются константами, то при $t \to \infty$ производные $p'_j(t)$ становятся равными нулю, и тогда из системы уравнений (7.9) имеем

$$\lambda_j p_j = \sum_{k \neq j} \lambda_{kj} p_k, \ j \in X.$$

Полученные уравнения называются **уравнениями равновесия**. Им можно дать следующее толкование. Назовем произведение $\lambda_j p_j$ интенсивности выхода из состояния j и вероятности этого состояния потоком вероятности из состояния j. Произведение $\lambda_{kj} p_k$ назовем потоком вероятности из состояния k в состояние j. Тогда из уравнений равновесия следует, что в стационарном режиме поток вероятности из любого состояния равен сумме потоков вероятностей из всех других состояний в данное состояние.

Интенсивности переходов между состояниями можно использовать для следующего определения.

Определение. Состояние j стохастически непрерывной однородной цепи Маркова с непрерывным временем называется **достижимым из состояния** i, если либо $i = j$, либо $\lambda_{ij} > 0$, либо $\lambda_{ik_1}\lambda_{k_1k_2}\cdot\ldots\cdot\lambda_{k_{n-1}k_n}\lambda_{k_nj} > 0$.

Пример 7.1. Найдем решения уравнений (7.9), удовлетворяющие начальным условиям

$$p_j(0) = \begin{cases} 1, & j = 0, \\ 0, & j > 0, \end{cases}$$

если

$$\lambda_j = \lambda,\ j = 0,\ 1,\ 2, \ldots,$$

$$\lambda_{kj} = \begin{cases} \lambda, & j = k+1, \\ 0, & j \neq k+1, \end{cases} \quad k = 0, 1, 2, \ldots,\ j = 1, 2, \ldots .$$

Система уравнений в данном случае принимает вид

$$\begin{cases} p_0'(t) = -\lambda\, p_0(t), \\ p_j'(t) = -\lambda\, p_j(t) + \lambda\, p_{j-1}(t),\ j = 1, 2, \ldots. \end{cases}$$

Она совпадает с системой уравнений для вероятностей состояний пуассоновского процесса (пример 2.3), поэтому получаем

$$p_j(t) = \frac{(\lambda t)^j}{j!} e^{-\lambda t},\ j = 0, 1, 2, \ldots .$$

Пример 7.2. Пусть $\xi(t)$ является однородной цепью Маркова с двумя состояниями, $X = \{0, 1\}$. Например, $\xi(t) = 0$ может означать, что некоторое устройство находится в рабочем состоянии, $\xi(t) = 1$ – что оно находится в нерабочем состоянии в момент времени t. Пусть $\lambda_{01} = \lambda,\ \lambda_{10} = \mu$, т.е. $\lambda_0 = \lambda,\ \lambda_1 = \mu$, и пусть $\xi(0) = 0$, т.е. $p_0(0)=1,\ p_1(0)=0$. Система уравнений (7.9) в этом случае имеет вид

$$\begin{cases} p_0'(t) = -\lambda\, p_0(t) + \mu\, p_1(t), \\ p_1'(t) = -\mu\, p_1(t) + \lambda\, p_0(t). \end{cases} \tag{7.10}$$

Поскольку $p_1(t) = 1 - p_0(t)$, то первое уравнение (7.10) можно переписать в виде

$$p_0'(t) + (\lambda+\mu)p_0(t) = \mu \cdot \tag{7.11}$$

Общее решение однородного уравнения $p_0'(t) + (\lambda+\mu)p_0(t) = 0$ равно $c\,e^{-(\lambda+\mu)t}$; частное решение неоднородного уравнения (7.11) можно найти в виде константы, поскольку это уравнение первого порядка, и записать как $\dfrac{\mu}{\lambda+\mu}$. Следовательно, общее решение уравнения (7.11) примет вид

$$p_0(t) = c\,e^{-(\lambda+\mu)t} + \frac{\mu}{\lambda+\mu} \cdot$$

Учитывая начальное условие, получаем $c = \dfrac{\lambda}{\lambda+\mu}$, и, таким образом,

$$p_0(t) = \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t} + \frac{\mu}{\lambda+\mu},$$

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

Используя эти вероятности, можно записать выражение для среднего значения цепи Маркова

$$m_\xi(t) = 0 \cdot p_0(t) + 1 \cdot p_1(t) = p_1(t) = \frac{\lambda}{\lambda+\mu}\left(1 - e^{-(\lambda+\mu)t}\right) \cdot$$

Найдем теперь распределение времени пребывания нашей цепи в состоянии 0, которое обозначим через η_0. Пусть $G(t) = P(\eta_0 > t)$. Из определения λ_0 следует, что

$$p_{00}(t, t+\Delta t) = p_{00}(\Delta t) = 1 - \lambda_0\,\Delta t + o(\Delta t) = 1 - \lambda\,\Delta t + o(\Delta t)\cdot$$

Поскольку

$$G(t+\Delta t) = G(t)\,p_{00}(t, t+\Delta t) = G(t)\big[1 - \lambda\,\Delta t + o(\Delta t)\big],$$

то

$$G(t) = -\lambda\,G(t),\quad G(0) = 1\cdot$$

Поэтому
$$G(t) = e^{-\lambda t}\cdot$$
Таким образом

$$P(\eta_0 \le t) = 1 - G(t) = 1 - e^{-\lambda t},$$

т.е. СВ η_0 имеет экспоненциальное распределение с параметром λ.

Очевидно, что $M\eta_0 = \dfrac{1}{\lambda}$. Аналогично можно показать, что время пребывания η_1 цепи в состоянии 1 имеет также экспоненциальное распределение, но с параметром μ. Из полученных соотношений также следует, что

$$\lambda_0 = \frac{1}{M\,\eta_0},\qquad \lambda_1 = \frac{1}{M\,\eta_1}\cdot$$

Данные результаты можно обобщить: можно показать, что в случае однородной цепи Маркова с конечным числом состояний, имеющей конечные интенсивности λ_i, $i \in X$, время пребывания в фиксированном состоянии имеет экспоненциальное распределение.

Иногда удобно записывать решения систем уравнений Колмогорова в матричной форме. Пусть имеем однородную цепь Маркова с непрерывным временем и конечным числом состояний, $X = \{1, 2, \ldots, N\}$. Введем матрицу вероятностей переходов

$$P(t) = \left\| p_{ij}(t) \right\|_{N\times N} \cdot$$

Тогда уравнение Чепмена–Колмогорова (7.3) можно записать более компактно в виде

$$P(t+\tau) = P(t)P(\tau), \; t,\tau \ge 0 \cdot \qquad (7.12)$$

Условие 5) (§31) говорит о непрерывности $P(t)$ при $t = 0$, поскольку из него следует, что $P(0) = I$, где I – единичная матрица. Из (7.12) вытекает, что $P(t)$ непрерывна при всех $t > 0$. Действительно, если $\tau = \Delta t > 0$ в (7.12), то

$$\lim_{\Delta t\to 0} P(t+\Delta t) = P(t)\lim_{\Delta t\to 0} P(\Delta t) = P(t)I = P(t) \cdot \qquad (7.13)$$

С другой стороны, при $t > 0$, $0 < \Delta t < t$ можно записать (7.12) в виде

$$P(t) = P(t-\Delta t)P(t) \cdot$$

Но $P(\Delta t)$ при достаточно малых Δt близка к единичной матрице. Поэтому обратная матрица $P^{-1}(\Delta t)$ к матрице $P(\Delta t)$ существует и также близка к I. Следовательно

$$P(t) = P(t)\lim_{\Delta t\to 0} P^{-1}(\Delta t) = \lim_{\Delta t\to 0} P(t-\Delta t) \cdot \qquad (7.14)$$

Из (7.13), (7.14) вытекает непрерывность матрицы $P(t)$.

Обозначим через A матрицу

$$A = \begin{pmatrix} -\lambda_1 & \lambda_{12} & \lambda_{13} & \dots & \lambda_{1N} \\ \lambda_{21} & -\lambda_2 & \lambda_{23} & \dots & \lambda_{2N} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \lambda_{N1} & \lambda_{N2} & \dots & \lambda_{N\,N-1} & -\lambda_N \end{pmatrix},$$

которая называется инфинитезимальной. Предельные соотношения (7.4) запишем в матричном виде

$$\lim_{\Delta t\to 0} \frac{P(\Delta t) - I}{\Delta t} = A \cdot$$

С помощью этого соотношения и равенства (7.12) находим

$$\frac{P(t+\Delta t)-P(t)}{\Delta t}=\frac{P(t)\left[P(\Delta t)-I\right]}{\Delta t}=\frac{P(\Delta t)-I}{\Delta t}P(t),$$

Откуда при $\Delta t \to 0$ получаются матричные дифференциальные уравнения

$$P'(t)=P(t)\,A=A\,P(t), \tag{7.15}$$

где $P'(t)=\left\|p'_{ij}(t)\right\|_{N\times N}$. Эти уравнения при начальных условиях $P(0)=I$ можно решить стандартным методом теории обыкновенных дифференциальных уравнений [*]. Тогда получим

$$P(t)=e^{At}=I+\sum_{n=1}^{\infty}\frac{A^n t^n}{n!}.$$

Практически мы должны найти собственные значения $q_1, q_2, \dots, q_N$ матрицы A и полную систему соответствующих им правых собственных векторов $u^{(1)}, u^{(2)}, \dots, u^{(N)}$, если это возможно [4]. Затем мы должны воспользоваться представлением

$$P(t)=U\,Q(t)U^{-1},$$

где U – матрица, столбцами которой являются соответственно вектора $u^{(1)}, u^{(2)}, \dots, u^{(N)}$, а $Q(t)$ – диагональная матрица

$$Q(t)=\begin{pmatrix} e^{q_1 t} & 0 & 0 & \dots & 0 \\ 0 & e^{q_2 t} & 0 & \dots & 0 \\ . & . & . & . & . \\ 0 & 0 & \dots & 0 & e^{q_N t} \end{pmatrix}.$$

Систему прямых дифференциальных уравнений Колмогорова можно применить для получения соотношений для финальных вероятностей состояний эргодических марковских цепей с непрерывным временем.

---

[*] Гантмахер Ф.Р. Теория матриц. – М.: Наука, 1988. – 548 с.

Определение. Цепь Маркова с непрерывным временем называется **эргодической**, если

$$\lim_{t\to\infty} p_{ij}(t) = \pi_j, \ \ j \in X, \ \ \sum_{j\in X} \pi_j = 1 \cdot$$

Можно показать, что для цепи с конечным числом состояний $X = \{1, 2, ..., N\}$ финальные вероятности состояний π_j, $j = \overline{1, N}$, всегда существуют. Другими словами, при бесконечно длительной эволюции в цепи устанавливается стационарное распределение вероятностей, не зависящее от времени. Уравнения для вычисления финальных вероятностей можно получить, переходя к пределу при $t \to \infty$ в системе уравнений (7.6). Тогда с учетом того, что $p'_{ij}(t) = 0$, имеем

$$\begin{cases} \lambda_j \, \pi_j = \sum\limits_{k=1}^{N} \pi_k \, \lambda_{kj}, \ \ j = \overline{1, N}, \\ \sum\limits_{k=1}^{N} \pi_k = 1. \end{cases} \qquad (7.16)$$

Система уравнений (7.15) однозначно определяет финальные вероятности состояний.

Покажем теперь, как интенсивности переходов (7.4) используются для нахождения средних времен переходов между состояниями. Рассмотрим цепь Маркова с непрерывным временем и счетным числом состояний. Обозначим через τ_{ij} среднее время перехода из i-го состояния в j-е, $i, j \in X$. За время Δt цепь Маркова может перейти из состояния i в состояние j с вероятностью $p_{ij}(\Delta t)$ и в состояние, отличное от j-го, с вероятностью $p_{ik}(\Delta t)$, $k \neq j$. Тогда, используя формулу для условного математического ожидания, имеем

$$\tau_{ij} = \Delta t \, p_{ij}(\Delta t) + \sum_{k\neq j} (\Delta t + \tau_{kj}) \, p_{ik}(\Delta t) \, . \qquad (7.17)$$

Подставляя вероятности (7.5) в (7.16), получим

$$\tau_{ij} = \lambda_{ij}(\Delta t)^2 + (\tau_{ij} + \Delta t)(1 - \lambda_i \Delta t) + \sum_{k \neq j} (\tau_{kj} + \Delta t) \lambda_{ik} \Delta t + o(\Delta t)$$

или

$$-\lambda_i \tau_{ij} \Delta t + \sum_{k \neq i, j} \lambda_{ik} \tau_{kj} \Delta t + o(\Delta t) = -\Delta t \cdot$$

Поделив данное выражение на Δt и переходя к пределу при $\Delta t \to 0$, будем иметь

$$-\lambda_i \tau_{ij} + \sum_{k \neq j} \lambda_{ik} \tau_{kj} = -1, \quad i, j \in X \cdot \tag{7.18}$$

Из этой системы уравнений можно найти среднее время перехода из состояния i в состояние j, $i \neq j$.

§35. ПРОЦЕСС ГИБЕЛИ И РАЗМНОЖЕНИЯ

Так называется однородная марковская цепь с непрерывным временем и счетным множеством состояний $X = \{0, 1, 2, \ldots\}$, в которой из состояния n возможен лишь непосредственный переход в состояния $n-1$ и $n+1$, а из состояния 0 – в состояние 1. Такие процессы хорошо описывают задачи в области биологии, физики, социологии, массового обслуживания. Состояние процесса можно интерпретировать, например, как число особей некоторой популяции, переход из состояния n в $n+1$ – как рождение новой особи, а переход из состояния n в $n-1$ – как гибель некоторой особи.

Будем предполагать, что все состояния процесса регулярны. Для сокращения записей обозначим $\lambda_{n\,n+1} = \lambda_n$, $\lambda_{n\,n-1} = \mu_n$. Тогда из (7.5) получаем

$$p_{nn+1}(\Delta t) = \lambda_n \Delta t + o(\Delta t),$$

$$p_{nn-1}(\Delta t) = \mu_n \Delta t + o(\Delta t),$$

$$p_{nn}(\Delta t) = 1 - (\lambda_n + \mu_n)\Delta t + o(\Delta t) \cdot$$

Применяя вышеприведенную интерпретацию, можно сказать, что $\lambda_n \Delta t$ с точностью до $o(\Delta t)$ есть вероятность рождения новой особи в популяции из n особей за время Δt, а $\mu_n \Delta t$ с точностью до $o(\Delta t)$ – вероятность гибели особи в этой популяции за время Δt.

Системы прямых и обратных дифференциальных уравнений Колмогорова для вероятностей переходов процесса имеют вид:

$$p'_{ij}(t) = \lambda_{j-1} p_{ij-1}(t) - (\lambda_j + \mu_j) p_{ij}(t) + \mu_{j+1} p_{ij+1}(t), \; i, j \in X, \qquad (7.18)$$

$$p'_{ij}(t) = \mu_i p_{i-1j}(t) - (\lambda_i + \mu_i) p_{ij}(t) + \lambda_i p_{i+1j}(t), \; i, j \in X, \qquad (7.19)$$

где $\mu_0 = \lambda_{-1} = 0$, а система уравнений для безусловных вероятностей состояний:

$$\begin{cases} p'_0(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t), \\ p'_j(t) = -\lambda_{j-1} p_{j-1}(t) - (\lambda_j + \mu_j) p_j(t) + \mu_{j+1} p_{j+1}(t), \; j = 1, 2, \dots . \end{cases} \qquad (7.20)$$

Ее можно записать в матричной форме (7.15), при этом инфинитезимальная матрица процесса равна

$$A = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \dots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \dots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \dots \\ 0 & 0 & \mu_3 & -(\lambda_3 + \mu_3) & \dots \\ \dots & \dots & \dots & \dots & . \end{pmatrix}.$$

Оказывается, если выполняется условие

$$\sum_{j=1}^{\infty} \prod_{i=1}^{j} \frac{\lambda_{i-1}}{\mu_i} < \infty, \qquad (7.21)$$

то процесс гибели и размножения эргодичен, существуют вероятности $p_j = \lim_{t \to \infty} p_j(t)$, $j \in X$, совокупность которых образует единственное стационарное распределение, совпадающее с эргодическим, т.е. стационарные вероятности состояний равны финальным вероятностям состояний. Для выполнения неравенства (7.21) достаточно, чтобы существовало такое N, для которого $\frac{\lambda_{i-1}}{\mu_i} < 1$ при всех $i > N$. Найдем стационарные вероятности состояний. Система уравнений равновесия для них, как следует из (7.20), имеет вид

$$\begin{cases} 0 = -\lambda_0 p_0 + \mu_1 p_1, \\ 0 = -\lambda_{j-1} p_{j-1} - (\lambda_j + \mu_j) p_j + \mu_{j+1} p_{j+1}, \; j = 1, 2, \ldots . \end{cases} \tag{7.22}$$

Обозначим $z_j = -\lambda_j p_j + \mu_{j+1} p_{j+1}$, $j = 0, 1, 2, \ldots$. Тогда из (7.22) получаем

$$\begin{cases} z_0 = 0, \\ z_j - z_{j-1} = 0, \; j = 1, 2, \ldots, \end{cases}$$

откуда следует, что $z_j = 0$ для всех $j = 0,\ 1,\ 2, \ldots$. Следовательно,

$$p_j = \frac{\lambda_{j-1}}{\mu_j} p_{j-1} = \frac{\lambda_{j-1}}{\mu_j} \cdot \frac{\lambda_{j-2}}{\mu_{j-1}} p_{j-2} = \ldots =$$

$$= \frac{\lambda_{j-1}}{\mu_j} \cdot \frac{\lambda_{j-2}}{\mu_{j-1}} \cdot \ldots \cdot \frac{\lambda_0}{\mu_1} p_0 = \prod_{i=1}^{j} \frac{\lambda_{i-1}}{\mu_i} p_0 . \tag{7.23}$$

Вероятность p_0 найдем, используя условие нормировки $\sum_{j=0}^{\infty} p_j = 1$. Подставляя (7.23) в это условие, находим

$$p_0 = \left[1 + \sum_{j=1}^{\infty} \prod_{i-1}^{j} \frac{\lambda_{i-1}}{\mu_i} \right]^{-1} .$$

Ясно, что для существования p_0 нужно, чтобы выполнялось неравенство (7.21). Если

$$\sum_{j=1}^{\infty} \prod_{i=1}^{j} \frac{\lambda_{i-1}}{\mu_i} = \infty ,$$

то $p_0 = 0$ и стационарного (финального) распределения не существует. Это означает, что эволюция процесса протекает в одну сторону, при этом номер состояния все возрастает.

Рассмотрим еще одну особенность, которая возникает в процессе чистого размножения, связанную с временами переходов процесса между состояниями. Пусть $\mu_i = 0$, $i \in X$. Найдем τ_{in} – среднее время перехода из состояния i в состояние n. Система (7.17) в данном случае имеет вид

$$-\lambda_i \tau_{in} + \lambda_i \tau_{i+1n} = -1, \ i \in X,$$

откуда

$$\tau_{in} = \tau_{i+1n} + \frac{1}{\lambda_i}. \qquad (7.24)$$

Так как рассматриваемый процесс чистого размножения протекает в одном направлении, то, как показано выше, стационарные (финальные) вероятности не существуют. При этом время возвращения в исходное состояние, естественно, равно нулю, $\tau_{nn} = 0$. Тогда

$$\tau_{n-1n} = \frac{1}{\lambda_{n-1}}.$$

Далее, полагая $i = n-2, n-3, \ldots$, из (7.24) получим

$$\tau_{in} = \sum_{j=i}^{n-1} \frac{1}{\lambda_j}.$$

Рассмотрим среднее время перехода в состояние с бесконечно большим номером, т.е.

$$\lim_{n \to \infty} \tau_{in} = \tau_{i\infty} = \sum_{j=i}^{\infty} \frac{1}{\lambda_j}.$$

Здесь опять возможны два случая, связанных со сходимостью ряда.

1. Если ряд $\sum_{j=i}^{\infty}\frac{1}{\lambda_j}$ расходится, то в состояние с бесконечно большим номером процесс перейдет за бесконечно большое время, что вполне реально.

2. Если же этот ряд сходится, т.е. $\sum_{j=i}^{\infty}\frac{1}{\lambda_j}<\infty$, то это соответствует тому, что процесс за конечное время переходит в состояние с бесконечно большим номером. Физически это означает, что в системе, описываемой процессом чистого размножения, возникает лавинообразное явление, или «взрыв». Для биологических систем, например, это соответствует эпидемии.

Пример 7.3 (*линейный рост с иммиграцией [4]*). Процесс гибели и размножения называется процессом с линейным ростом, если

$$\lambda_n=\lambda n+a,\ \mu_n=\mu n,\ a>0.$$

Такие процессы также возникают при изучении биологического воспроизведения и роста популяций. Если состояние n описывает текущий размер популяции, то средняя мгновенная интенсивность роста равна $\lambda n+a$. Аналогично, вероятность того, что состояние процесса уменьшится на 1 за малый промежуток времени Δt, равна $\mu n\Delta t+o(\Delta t)$. Величина λn соответствует естественному приросту популяции размера n, в то время как второе слагаемое a можно интерпретировать как интенсивность роста популяции за счет внешнего источника, такого, например, как иммиграция.

Если подставить значения λ_n и μ_n в систему прямых уравнений Колмогорова (7.18), то получим

$$p'_{i0}(t)=-ap_{i0}(t)+\mu p_{i1}(t)$$

$$p'_{ij}(t)=\big(\lambda(j-1)+a\big)p_{ij-1}(t)-\big((\lambda+\mu)j+a\big)p_{ij}(t)+$$

$$+\mu(j+1)p_{ij+1}(t),\ j=1,2,\ldots$$

Если теперь умножить j-е уравнение на j и просуммировать уравнения, то получим среднее значение нашего процесса

$$m(t)=M\xi(t)=\sum_{j=1}^{\infty} jp_{ij}(t).$$

Можно показать, что оно удовлетворяет линейному неоднородному дифференциальному уравнению первого порядка

$$m'(t)-(\lambda-\mu)m(t)=a$$

с начальным условием $m(0)=i$, если $\xi(0)=i$. Решение этого уравнения имеет вид

$$m(t)=\begin{cases} at+i,\ \lambda=\mu, \\ \dfrac{a}{\lambda-\mu}\left[e^{(\lambda-\mu)t}-1\right]-ie^{(\lambda-\mu)t},\ \lambda\neq\mu. \end{cases}$$

Ясно, что если $\lambda\geq\mu$, то $m(t)\xrightarrow[t\to\infty]{}\infty$. Если же $\lambda<\mu$, то средний размер популяции при больших t примерно равен

$$m(t)\approx\frac{a}{\mu-\lambda}.$$

§36. АНАЛИЗ МАРКОВСКИХ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ

В этом параграфе мы рассмотрим применение процесса гибели и размножения к анализу систем массового обслуживания (СМО), которые являются адекватными математическими моделями многих объектов в технике, экономике, природе и обществе. Довольно эффективны подходы теории массового обслуживания для описания процессов в электронных вычислительных машинах, их отдельных устройств или, наоборот, в компьютерных системах и сетях, составленных из них.

Общая модель обслуживания имеет следующую структуру. На некоторое устройство (СМО) в случайные моменты времени поступают заявки и требуют какого-либо рода обслуживания. Под термином «заявки» можно понимать, например, запросы, поступающие на сервер компьютерной сети; корабли, входящие в порт, или самолеты, прилетающие в аэропорт; клиентов, требующих обслуживания в банках, страховых компаниях, магазинах; неисправные машины, ожидающие ремонта и т. д. Употребление того или иного термина часто определяется характером решаемых прикладных задач. Вместо термина «заявка» поэтому используют также термин «запрос», «вызов», «требование», «клиент», а вместо термина «обслуживающее устройство» – «линия», «канал», «прибор».

СМО классифицируются по следующим признакам.

1. Входящий поток заявок, для задания которого необходимо знать распределение интервалов времени между моментами поступления заявок. Мы будем рассматривать СМО с простейшим (пуассоновским) потоком заявок, т. е. когда моменты поступления заявок образуют простейший поток событий (пример 5.7), в котором интервалы между моментами поступления заявок распределены по экспоненциальному закону с параметром λ.

2. Распределение времен обслуживания заявок. Будем предполагать, что длительности обслуживания заявок распределены экспоненциально с параметром μ.

3. Количество линий обслуживания. Будем рассматривать однолинейные и многолинейные СМО.

4. Дисциплина обслуживания заявок. Дисциплиной обслуживания называется правило, в соответствии с которым на обслуживание выбирается заявка из числа ожидающих обслуживания. В большинстве моделей принята дисциплина «первый пришел – первый обслуживается» – прямой порядок обслуживания, когда заявки обслуживаются в той последовательности, в которой они поступают на вход системы. Все СМО, которые будут рассмотрены в этой главе, такого типа.

Для обозначения достаточно простых СМО, анализируемых в этом параграфе, обычно используют символику, введенную английским математиком Кендаллом: $A/B/n/N$. Буква A обозначает тип входящего потока заявок, например, если $A=M$, то входящий поток заявок является простейшим. Буква B обозначает тип распределения времен обслуживания заявок в линиях, которые мы будем считать идентичными. Например, $B=M$ означает одинаковое экспоненциальное распределение времен обслуживания для всех линий СМО. Число n обозначает количество линий обслуживания, а N – число мест ожидания в очереди. В случае $N=0$ заявки, поступающие в систему в моменты времени, когда все линии заняты обслуживанием других заявок, уходят из системы без ожидания (теряются). Поэтому такие СМО называются системами с потерями. Если же $N=\infty$, все заявки терпеливо ожидают начала обслуживания, соответствующее СМО называются системами с ожиданием. При этом последний символ в обозначениях Кенделла обычно опускается. Если $0<N<\infty$, СМО называется системой с ограниченным числом мест ожидания.

Пусть $\xi(t)$ – число заявок в системе (в очереди и на обслуживании) в момент времени t. Вероятностное распределение $\xi(t)$ после момента t определяется:

а) числом заявок в системе в момент времени t,

б) моментами поступления заявок после момента t,

в) моментами окончаний обслуживания заявок после момента t.

Рассмотрим СМО с простейшим входящим потоком заявок и экспоненциальным обслуживанием. В силу «отсутствия памяти» у экспоненциального распределения (см. §4), моменты поступления заявок после момента t не зависят от предыстории системы до момента t. Аналогично, моменты окончания обслуживания заявок после момента t не зависят от предыстории системы до момента t. Поэтому вероятностное поведение $\xi(t)$ после момента t зависит только от $\xi(t)$ и не зависит от поведения $\xi(t)$ до момента t.

Следовательно $\xi(t)$ – цепь Маркова с конечным или счетным числом состояний. Такие СМО называются марковскими. Для нахождения вероятностей $p_j(t) = P(\xi(t) = j),\ j \in X$, нужно решить систему дифференциальных уравнений Колмогорова для безусловных вероятностей состояний.

В большинстве рассматриваемых СМО $\xi(t)$ является процессом гибели и размножения со следующим графом переходом (рис. 31):

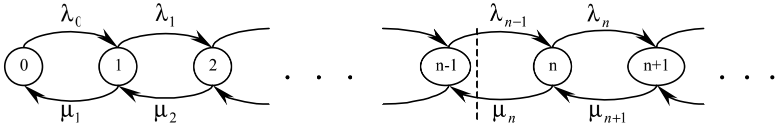


Рис.31

Вершины данного графа обозначают состояния СМО, а дуги возможные переходы из состояния в состояние. При этом каждой дуге приписывается некоторый вес, равный интенсивности соответствующего перехода. В случае, если существует стационарный

режим, то стационарные вероятности состояний $p_j,\ j \in X$, находятся из системы уравнений равновесия (7.22), которую можно переписать в виде

$$\lambda_0 p_0 = \mu_1 p_1, \tag{7.25}$$

$$(\lambda_j + \mu_j) p_j = \lambda_{j-1} p_{j-1} + \mu_{j+1} p_{j+1}, \quad j = 1,\ 2, \cdots. \tag{7.26}$$

Вычитая из уравнения (7.26) при $j = 1$ уравнение (7.25), получаем

$$\lambda_1 p_1 = \mu_2 p_2 \cdot$$

Далее, вычитая из уравнения (7.26) при $j = 2$ последнее, получим

$$\lambda_2 p_2 = \mu_3 p_3 \cdot$$

Продолжая этот процесс, будем иметь

$$\lambda_{n-1} p_{n-1} = \mu_n p_n, \quad n = 1, 2, \cdots. \tag{7.27}$$

Для составления уравнений (7.27) можно использовать следующую геометрическую интерпретацию и понятие потока вероятности (см. §32). Проведем в графе вертикальное сечение между двумя состояниями $n-1$ и n (рис. 31). Тогда поток вероятности через это сечение слева $\lambda_{n-1} p_{n-1}$ в стационарном режиме равен потоку вероятности через это сечение справа $\mu_n p_n$.

Пусть в систему поступает простейший поток заявок с параметром λ. Как следует из примера 2.3, λ является средним числом заявок, поступающих в единицу времени, поэтому λ характеризует интенсивность простейшего потока заявок. Аналогично, если μ – параметр экспоненциального распределения времен обслуживания заявок, то μ^{-1} является средним временем обслуживания заявок, а μ характеризует интенсивность обслуживания заявок, если линия работает без перерыва. Введем следующие обозначения: пусть $a_k(\Delta t)$ – вероятность поступления k заявок за время Δt, $a_{\geq 2}(\Delta t)$ – вероятность поступления двух или более заявок за время Δt. Из формулы (2.6) следует, что

$$a_k(\Delta t) = \frac{(\lambda \Delta t)^k}{k!} e^{-\lambda \Delta t}, k = 0, 1, 2, \ldots \cdot$$

Поэтому на основании этой формулы и свойства ординарности простейшего потока:

$$a_0(\Delta t) = e^{-\lambda \Delta t} = 1 - \lambda \Delta t + o(\Delta t), \tag{7.28}$$

$$a_1(\Delta t) = \lambda \Delta t e^{-\lambda \Delta t} = \lambda \Delta t + o(\Delta t), \tag{7.29}$$

$$a_{\geq 2}(\Delta t) = o(\Delta t). \tag{7.30}$$

Рассмотрим однолинейную СМО с экспоненциальным обслуживания заявок с интенсивностью μ. Пусть $b_k(\Delta t)$ – условная вероятность того, что, начиная с определенного момента, за время

Δt будет обслужено k заявок при условии, что в указанный момент линией обслуживалась некоторая заявка; $b_{\geq k}(t)$ – аналогичная вероятность обслуживания k и более заявок. Тогда

$$b_{\geq 1}(\Delta t) = 1 - e^{-\mu\Delta t} = 1 - (1 - \mu\Delta t + o(\Delta t)) =$$
$$= \mu\Delta t + o(\Delta t),$$
$$b_0(\Delta t) = 1 - b_{\geq 1}(\Delta t) = 1 - \mu\Delta t + o(\Delta t). \qquad (7.31)$$

Если бы линия работала непрерывно, то моменты окончаний обслуживания заявок образовывали бы простейший поток с параметром μ и вероятность $b_{\geq 2}(\Delta t)$, которую при данном предположении обозначим через $\widetilde{b}_{\geq 2}(\Delta t)$, была бы в силу ординарности равна $o(\Delta t)$. Следовательно

$$b_{\geq 2}(\Delta t) \leq \widetilde{b}_{\geq 2}(\Delta t) = o(\Delta t), \qquad (7.32)$$
$$b_1(\Delta t) = b_{\geq 1}(\Delta t) - b_{\geq 2}(\Delta t) = \mu\Delta t + o(\Delta t). \qquad (7.33)$$

Рассмотрим теперь n-линейную СМО, у которой n идентичных линий работают независимо друг от друга. Обозначим через $b_k^{(n)}(\Delta t)$ вероятность того, что в такой системе за время Δt будет обслужено k заявок. Поскольку линии работают независимо, то

$$b_0^{(n)}(\Delta t) = b_0(\Delta t) b_0(\Delta t) \cdot \ldots \cdot b_0(\Delta t) =$$
$$= (1 - \mu\Delta t + o(\Delta t))^n = 1 - n\mu\Delta t + o(\Delta t). \qquad (7.34)$$

Событие {в n-линейную СМО за время Δt обслужилась ровно одна заявка} равносильно событию {за это время заявка обслужилась в некоторой одной линии данной СМО, а в остальных линиях завершения обслуживания заявок не произошло}. Поэтому, используя формулу полной вероятности, будем иметь

$$b_1^{(n)}(\Delta t) = b_1(\Delta t) b_0(\Delta t) \cdot \ldots \cdot b_0(\Delta t) + b_0(\Delta t) b_1(\Delta t) \cdot \ldots \cdot b_0(\Delta t) + \ldots +$$
$$+ b_0(\Delta t) \cdot \ldots \cdot b_0(\Delta t) b_1(\Delta t) = n \cdot b_1(\Delta t) [b_0(\Delta t)]^{n-1} = n\mu\Delta t + o(\Delta t), \qquad (7.35)$$

и

$$b_{\geq 2}{}^{(n)}(\Delta t) = 1 - b_0{}^{(n)}(\Delta t) - b_1{}^{(n)}(\Delta t) = o(\Delta t), \tag{7.36}$$

где $b_{\geq 2}{}^{(n)}(\Delta t)$ – вероятность того, что в n-линейную СМО за время Δt будет обслужено более одной заявки.

Рассмотрим некоторые простейшие СМО.

Система M / M / 1. Определим вероятности переходов для процесса $\xi(t)$. Переход процесса $\xi(t)$ из состояния i в состояние $i-1$ за время Δt означает, что в системе за это время должно обслужиться на одну заявку больше, чем поступило. Следовательно, из формулы полной вероятности имеем

$$p_{ii-1}(\Delta t) = a_0(\Delta t) b_1(\Delta t) + \sum_{k=1}^{\infty} a_k(\Delta t) b_{k+1}(\Delta t).$$

Но, так как $\sum_{k=1}^{\infty} a_k(\Delta t) b_{k+1}(\Delta t) \leq b_{\geq 2}(\Delta t)$, поскольку $a_k(\Delta t) \leq 1, k = 1,2,...$, то из (7.28), (7.32), (7.33) следует, что

$$p_{ii-1}(\Delta t) = \mu \Delta t + o(\Delta t),\ i \geq 1. \tag{7.37}$$

Аналогично, переход процесса из состояния i в состояние $i+1$ за время Δt означает, что в систему должно за это время поступить на одну заявку больше, чем обслужиться. Поэтому

$$p_{ii+1}(\Delta t) = a_1(\Delta t) b_0(\Delta t) + \sum_{k=2}^{\infty} a_k(\Delta t) b_{k-1}(\Delta t) \leq a_1(\Delta t) b_0(\Delta t) + a_{\geq 2}(\Delta t),$$

а из (7.29) – (7.31) вытекает, что

$$p_{ii+1}(\Delta t) = \lambda \Delta t + o(\Delta t),\ i \geq 0. \tag{7.38}$$

Кроме того,

$$p_{ii}(\Delta t) = a_0(\Delta t) b_0(\Delta t) + \sum_{k=1}^{\infty} a_k(\Delta t) b_k(\Delta t) \leq$$

$$\leq a_0(\Delta t) b_0(\Delta t) + a_{\geq 1}(\Delta t) b_{\geq 1}(\Delta t) + a_0(\Delta t) b_0(\Delta t) + o(\Delta t),$$

и учитывая (7.28), (7.31), получаем

$$p_{ii}(\Delta t) = 1 - (\lambda + \mu)\Delta t + o(\Delta t),\ i \geq 1. \tag{7.39}$$

Аналогично можно найти

$$p_{00}(\Delta t) = 1 - \lambda\Delta t + o(\Delta t), \tag{7.40}$$

$$p_{ik}(\Delta t) = o(\Delta t), \ |i-k| \geq 2.$$

Таким образом, из (7.37) – (7.40) следует, что $\xi(t)$ является процессом гибели и размножения со следующим графом переходов (рис. 32).

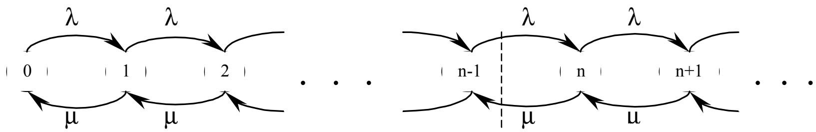


Рис. 32

Из условия эргодичности для процесса гибели и размножения следует, что если

$$\rho = \frac{\lambda}{\mu} < 1,$$

то существует единственное стационарное распределение, совпадающее с эргодическим, ρ называется коэффициентом загрузки системы. Уравнение равновесия (7.27) имеют вид

$$\lambda p_{n-1} = \mu p_n, \ n = 1, 2, ...,$$

откуда находим

$$p_n = \rho p_{n-1} = \rho^2 p_{n-2} = ... = \rho^n p_0.$$

Вероятность p_0 можно найти, используя условие нормировки $\sum_{n=0}^{\infty} p_n = 1$, откуда следует, что $p_0 = 1 - \rho$, и поэтому

$$p_n = \rho^n (1-\rho), \ n = 0, 1, 2, ...,$$

т.е. число заявок в такой СМО в стационарном имеет геометрическое распределение.

Легко найти производящую функцию такого распределения:

$$\psi(z)=\sum_{n=0}^{\infty}p_n z^n=\frac{1-\rho}{1-\rho z},\ |z|<1.$$

Отсюда получаем выражение для среднего числа заявок в системе (в очереди и на обслуживании) в стационарном режиме:

$$M\xi=\sum_{n=0}^{\infty}np_n=\psi'(z)\big|_{z=1}=\frac{\rho}{1-\rho}.$$

Ясно, что $M\xi\xrightarrow[\rho\to 1]{}\infty$, т.е. при стремлении ρ к 1 очередь в СМО неограниченно растет.

При рассмотрении последующих СМО выражения для переходных вероятностей находятся подобным образом, как в предыдущем случае, с использованием соотношений (7.28) – (7.36). Поэтому в дальнейшем мы сразу будем изображать граф переходов для соответствующего процесса гибели и размножения. Нужно только учесть, что если обслуживанием заявок в многолинейной СМО заняты k линий, то интенсивность обслуживания равна $k\mu$.

Система M / M / n / 0. Это система с потерями без ожидания. Если заявка поступает в систему в момент, тогда обслуживанием заняты все n линий, то она теряется. Такая система была введена датским инженером Эрлангом в начале прошлого столетия и применена в качестве модели обработки вызовов, поступающих на телефонную станцию. Граф переходов для такой СМО имеет вид (рис. 33)

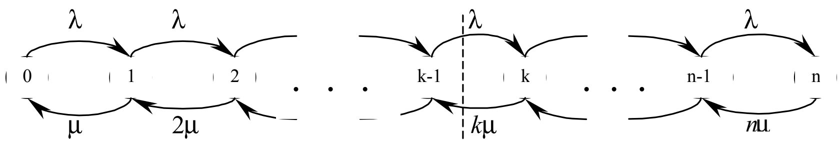


Рис. 33

Поскольку число состояний системы конечно, то единственное стационарное распределение, совпадающее с эргодическим, существует всегда при любых параметрах λ,μ (см. §32). Запишем уравнения равновесия для стационарных вероятностей состояний:

$$\lambda p_{k-1}=k\mu p_k\ ,k=\overline{1,n}\cdot$$

Отсюда получаем

$$p_k = \frac{\rho}{k} p_{k-1} = \frac{\rho^2}{k(k-1)} p_{k-2} = \ldots = \frac{\rho^k}{k!} p_0 \ , k = \overline{1,n} \cdot \qquad (7.41)$$

Вероятность p_0, как всегда, можем найти из условия нормировки, которое в данном случае имеет вид $\sum_{k=0}^{n} p_k = 1$, откуда

$$p_0 = \left(\sum_{k=0}^{n} \frac{\rho^k}{k!} \right)^{-1} .$$

Таким образом,

$$p_k = \frac{\rho^k}{k!} \left(\sum_{i=0}^{n} \frac{\rho^i}{i!} \right)^{-1} .$$

В частности, вероятность потери заявки равна (формула потерь Эрланга)

$$p_n = \frac{\rho^n}{n!} \left(\sum_{i=0}^{n} \frac{\rho^i}{i!} \right)^{-1} .$$

Система M / M / n. Это многолинейная система с ожиданием. Если обслуживанием заявок заняты все n линий, то интенсивность обслуживания равна $n\mu$. Гsраф переходов для этой СМО имеет вид (рис. 34)

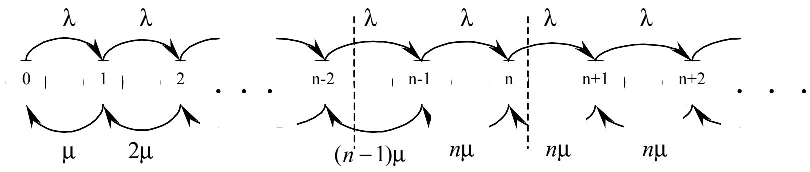


Рис. 34

Стационарное распределение существует, если

$$\rho_n = \frac{\rho}{n} < 1 \cdot$$

Уравнение равновесия имеет вид

$$\begin{cases} \lambda p_{k-1} = k\mu p_k \ , k = \overline{1, n} \,, \\ \lambda p_{k-1} = n\mu p_k \ , k \geq n+1, \end{cases}$$

откуда, аналогично предыдущему случаю (7.41), получаем

$$p_k = \frac{\rho^k}{k!} p_0 \ , k \leq n \ ,$$

$$p_k = \frac{\rho}{n} p_{k-1} = \left(\frac{\rho}{n}\right)^2 p_{k-2} = \ldots = \left(\frac{\rho}{n}\right)^{k-n} p_n = \left(\frac{\rho}{n}\right)^{k-n} \frac{\rho^n}{n!} p_0 =$$

$$= \frac{\rho^k}{n^{k-n} n!} p_0 \ , k \geq n+1 \, .$$

Условие нормировки в этом случае принимает вид

$$\sum_{k=0}^{\infty} p_k = p_0 \left[\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{n^n}{n!} \sum_{k=n+1}^{\infty} \left(\frac{\rho}{n}\right)^k \right] = 1 \,,$$

откуда следует, что

$$p_0 = \left[\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{n^n \rho_n^{\ n+1}}{n!(1-\rho_n)} \right]^{-1} .$$

Среднее число заявок в системе в стационарном режиме равно

$$M\xi = \sum_{k=0}^{\infty} k p_k = p_0 \left\{ \sum_{i=1}^{n} \frac{\rho^i}{(i-1)!} + \frac{n^n \rho_n^{\ n+1} [1 - n(1-\rho_n)]}{n!(1-\rho_n)^2} \right\} .$$

Система M / M / n / N. Это многолинейная система с ограниченным числом мест для ожидания. Она отличается от предыдущей СМО тем, что в ней имеется только N мест для ожидания. Поэтому граф переходов в этом случае имеет вид (рис. 35)

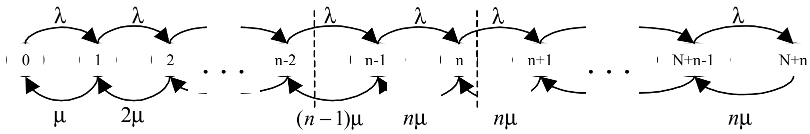


Рис. 35

Стационарное распределение существует при любых λ, μ, поскольку число состояний системы конечно. Уравнения равновесия принимают вид

$$\begin{cases} \lambda p_{k-1} = k\mu p_k \ , k = \overline{1, n}\,, \\ \lambda p_{k-1} = n\mu p_k \ , k = \overline{n+1, N+n}\,. \end{cases}$$

Отсюда следует, что стационарные вероятности p_k , $k \in X$, имеют такой же вид, как и для предыдущей СМО, с той лишь разницей, что они определены для $k = \overline{0\,, N+n}$. Таким образом,

$$p_k = \begin{cases} \dfrac{\rho^k}{k!}\, p_0 \ , k = \overline{0\,, n}\,, \\ \dfrac{n^n {\rho_n}^k}{n!}\, p_0 \ , k = \overline{n+1\,, N+n}\,. \end{cases}$$

Вероятность p_0 определяется из условия нормировки $\sum\limits_{k=0}^{n+N} p_k = 1$:

$$p_0 \left[\sum_{k=0}^{n} \frac{\rho^k}{k!} + \sum_{k=n+1}^{N+n} \frac{n^n {\rho_n}^k}{n!} \right] = 1\,,$$

откуда получаем

$$p_0 = \left[\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{\rho^n \rho_n (1 - {\rho_n}^N)}{(1-\rho_n)} \right]^{-1}.$$

Среднее число заявок в системе определяется соотношением

$$M\xi=\sum_{k=0}^{n+N}kp_k=p_0\left\{\sum_{i=1}^{n}\frac{\rho^i}{(i-1)!}+\frac{\rho^n\rho_n}{n!(1-\rho_n)}\left[n-(n+N){\rho_n}^N+\frac{1-{\rho_n}^N}{1-\rho_n}\right]\right\},$$

а вероятность потери заявки

$$p_{n+N}=p_0\,\frac{\rho^{n+N}}{n!\,n^N}.$$

§35. ВЕТВЯЩИЕСЯ ПРОЦЕССЫ С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

Предположим, что имеется некоторая совокупность частиц, которые с течением времени превращаются в частицы такого же типа, причем этот процесс «размножения» обладает следующим свойством: каждая из исходных частиц через время t независимо от других частиц и от обстоятельств, предшествующих исходному моменту, с вероятностью $p_k(t)$ порождает группу из k частиц. Обозначим через $\xi(t)$ число частиц, имеющихся к моменту времени t. Очевидно, эволюция величины $\xi(t)$ представляет собой цепь Маркова с непрерывным временем. Процесс $\xi(t)$ такого вида называется **ветвящимся процессом**. Такие процессы служат математическими моделями многих реальных процессов (химические реакции, ядерные процессы, выживание фамилий, генетические и экологические процессы и др.) [4, 11].

Пусть в некоторый исходный момент времени t_0 имеется ровно i частиц и пусть $\xi_j(t)$ – число частиц, порожденных j-й частицей через время t, $j=1,2,...,i$. Предположим, что СВ $\xi_1(t),\xi_2(t),...,\xi_i(t)$ независимы между собой и имеют одно и то же распределение вероятностей

$$p_k(t)=P\big(\xi_j(t)=k\big),\ k=0,1,2,...,\ j=1,2,...,i.$$

Общее число частиц через время t будет

$$\xi(t)=\xi_1(t)+\xi_2(t)+...+\xi_i(t). \tag{7.42}$$

Будем рассматривать однородные ветвящиеся процессы. Предположим, что отдельная частица за малый промежуток времени Δt с вероятностью

$$p_k(\Delta t) = \lambda_k \Delta t + o(\Delta t),\ k \neq 1,$$

превращается в k новых частиц, и с вероятностью

$$p_1(\Delta t) = 1 - \lambda \Delta t + o(\Delta t)$$

остается неизменной. Пусть также $\lambda_1 = -\lambda$, $\sum_i \lambda_i = 0$ и переходные вероятности $p_k(t) = p_{1k}(t)$ удовлетворяют обратным дифференциальным уравнениям Колмогорова (7.7):

$$p_k'(t) = \sum_i \lambda_i p_{ik}(t),\ k = 0,1,2,..., \tag{7.43}$$

где $p_{ik}(t)$ – переходные вероятности ветвящегося процесса, $p_{ik}(t)$ есть вероятность того, что i частиц за время t порождают k частиц.

Введем производящие функции

$$\psi(t,z) = \sum_{k=0}^{\infty} p_k(t) z^k, \tag{7.44}$$

$$\psi_i(t,z) = \sum_{k=0}^{\infty} p_{ik}(t) z^k.$$

При каждом z, для которого $|z| < 1$, из (7.43) имеем

$$\frac{d}{dt}\sum_{k=0}^{\infty} p_k(t) z^k = \sum_{k=0}^{\infty} z^n \frac{d}{dt} p_k(t) = \sum_i \lambda_i \sum_{k=0}^{\infty} p_{ik}(t) z^k,$$

что дает следующее дифференциальное уравнение для производящей функции $\psi(t,z)$:

$$\frac{d}{dt}\psi(t,z) = \sum_i \lambda_i \psi_i(t,z). \tag{7.45}$$

Но, по определению математического ожидания для дискретной СВ,

$$\psi(t,z) = Mz^{\xi_j(t)},\ j = 1,2,...,i,$$

$$\psi_i(t,z) = Mz^{\xi(t)},$$

где $\xi(t) = \sum_{j=1}^{i} \xi_j(t)$. Поскольку все величины $\xi_j(t)$, $j = \overline{1,i}$, независимы, то

$$Mz^{\xi(t)} = Mz^{\xi_1(t)} \cdot Mz^{\xi_2(t)} \cdot \ldots \cdot Mz^{\xi_i(t)}.$$

Следовательно,

$$\psi_i(t,z) = [\psi(t,z)]^i, \; i = 1, 2, \ldots,$$

и дифференциальное уравнение для производящей функции $\psi(t,z)$ можно записать в виде

$$\frac{d}{dt}\psi(t,z) = \sum_i \lambda_i \psi^k(t,z). \tag{7.46}$$

Будем считать, что для ветвящегося процесса $\xi(t)$ параметры λ_i, $i = 0, 1, 2, \ldots$, заданы. Введем функцию.

$$f(x) = \sum_{i=0}^{\infty} \lambda_i x^i. \tag{7.47}$$

Согласно (7.46), производящая функция $\psi(t,z)$ является решением дифференциального уравнения вида

$$\frac{dx}{dt} = f(x). \tag{7.48}$$

Поскольку $\psi(0,z) = z$, производящая функция $\psi(t,z)$ при каждом z, $0 \le z \le 1$, совпадает с решением $x = x(t)$ этого уравнения, удовлетворяющим начальному условию $x(0) = z$. Вместо уравнения (7.48) удобно рассмотреть эквивалентное ему дифференциальное уравнение для обратной к $x = x(t)$ функции $t = t(x)$:

$$\frac{dt}{dx} = \frac{1}{f(x)}.$$

Решение этого уравнения, функция $t = t(x)$, имеет вид

$$t = \int_z^x \frac{du}{f(u)}, \quad 0 \le x \le 1 \cdot \qquad (7.49)$$

Пример 7.4. Пусть

$$\lambda_0 = \lambda, \; \lambda_1 = -\lambda, \; \lambda_i = 0, \; i = 2,3,...$$

В этом случае $f(x) = \lambda - \lambda x = \lambda(1 - x)$ и

$$t = \int_z^x \frac{du}{f(u)} = -\frac{1}{\lambda}\left[\ln(1-x) - \ln(1-z)\right] \cdot$$

Из этого соотношения легко находится функция $\psi = \psi(t,z)$, а именно,

$$\ln(1-\psi) = -\lambda t + \ln(1-z),$$

т.е.

$$\psi(t,z) = 1 - e^{-\lambda t}(1-z) \cdot$$

Вероятности $p_k(t)$, определяемые из разложения (7.44), имеют вид

$$p_0(t) = 1 - e^{-\lambda t}, \; p_1(t) = e^{-\lambda t}, \; p_k(t) = 0, \; k = 2,3,...$$

Пример 7.5. Предположим, что некоторые частицы размножаются делением пополам. В соответствующем этому явлению ветвящемся процессе нужно взять

$$\lambda_0 = 0, \; \lambda_1 = -\lambda, \; \lambda_2 = \lambda, \; \lambda_k = 0, \; i = 3,4,...$$

В данном случае $f(x) = -\lambda x + \lambda x^2 = \lambda x(x-1)$ и

$$t = \int_z^x \frac{du}{f(u)} = \frac{1}{\lambda}\int_z^x \left[\frac{1}{u-1} - \frac{1}{u}\right] du =$$

$$= \frac{1}{\lambda}\left[\ln\left(1 - \frac{1}{x}\right) - \ln\left(1 - \frac{1}{z}\right)\right] \cdot$$

Таким образом, функция $\psi = \psi(t, z)$ может быть определена из соотношения

$$t = \frac{1}{\lambda}\left[\ln\left(1 - \frac{1}{\psi}\right) - \ln\left(1 - \frac{1}{z}\right)\right],$$

откуда следует, что

$$\psi(t, z) = e^{-\lambda t}\frac{z}{1 - z(1 - e^{-\lambda t})} = e^{-\lambda t} z \sum_{k=0}^{\infty}(1 - e^{-\lambda t})^k z^k \cdot$$

В итоге получаем, что

$$p_0(t) = 0,\ p_k(t) = e^{-\lambda t}(1 - e^{-\lambda t})^{k-1},\ k = 1, 2, \ldots \cdot$$

В заключение данного параграфа проанализируем эффекты вырождения и взрыва для ветвящихся процессов. Определим вероятность того, что через некоторое время t не останется ни одной частицы. Естественно, этого не может произойти, если $\lambda_0 = 0$, т.е. если частицы не могут исчезать, а могут лишь размножаться. Если в исходный момент $t = 0$ имеется одна частица, то эта вероятность равна $p_0(t) = \psi(t, 0)$. Если же вначале имеется i частиц, то эта вероятность имеет вид

$$p_{i0}(t) = \psi^i(t, 0) = p_0^i(t) \cdot$$

Как функция от t вероятность $p_0(t)$, как следует из вышесказанного и (7.43), является решением дифференциального уравнения (7.48), отвечающим параметру $z = 0$:

$$p_0'(t) = f\left[p_0(t)\right],\ p_0(0) = 0\,.$$

Проведя анализ дифференциального уравнения (7.48), где $f(x)$ имеет вид (7.47), можно показать [11], что это решение при $t \to \infty$ асимптотически приближается к некоторому значению $p_0 = \alpha$, являющемуся наименьшим корнем уравнения $f(x) = 0$, т.е.

$$\lim_{t \to \infty} p_0(t) = \alpha \cdot$$

Таким образом, $p_0 = \alpha$ есть **вероятность вырождения ветвящегося процесса** $\xi(t)$ – вероятность того, что к некоторому моменту времени не останется ни одной частицы. Если функция $f(x)$ является положительной на интервале $0 \le x < 1$, то вероятность вырождения ветвящегося процесса $\xi(t)$ равна 1.

Рассмотрим теперь так называемое **явление взрыва**, когда образуется бесконечно много частиц. Вероятность того, что взрыв произойдет до момента t, если вначале была одна частица, равна

$$p_\infty(t) = P\big(\xi(t) \ge \infty\big) = 1 - P\big(\xi(t) < \infty\big) =$$

$$= 1 - \sum_{k=0}^{\infty} P\big(\xi(t) = k\big) = 1 - \sum_{k=0}^{\infty} p_k(t) = 1 - \lim_{z \to 1} \psi(t, z) \cdot \tag{7.50}$$

В случае, когда $x(t) = 1$ является единственной интегральной кривой дифференциального уравнения (7.48), проходящей через точку $(0,1)$, или при выполнении условия

$$\int_{x_0}^{1} \frac{dx}{f(x)} = -\infty \tag{7.51}$$

с помощью (7.44) получаем

$$\lim_{z \to 1} \psi(t, z) = 1,$$

и тогда, согласно (7.50), $p_\infty(t) = 0 \quad \forall\, t > 0$, так что возможность взрыва исключена. В [11] показано, что при условии

$$\int_{x_0}^{1} \frac{dx}{f(x)} > -\infty \tag{7.52}$$

интегральные кривые $x(t, z)$ при $z \to 1$ монотонно сходятся к функции $x_0(t)$, так что **вероятность взрыва** есть

$$p_\infty(t) = 1 - x_0(t) > 0 \cdot$$

Ясно, что если вначале было i частиц, то вероятность взрыва через время t при условии (7.50) также равна нулю, а при условии (7.52) она равняется

$$p_{i\infty}(t) = 1 - x_0^i(t) \cdot$$

ГЛАВА 8. НЕПРЕРЫВНЫЕ МАРКОВСКИЕ ПРОЦЕССЫ

§38. ОБОБЩЕННОЕ УРАВНЕНИЕ МАРКОВА

Рассмотрим теперь класс марковских процессов, который характеризуется тем, что множество его состояний X является непрерывным и время перехода из одного состояния в другое также изменяется непрерывно. Такие процессы образуют класс непрерывных марковских процессов, широко используемых в качестве моделей различных реальных процессов.

Докажем вспомогательную формулу.

Лемма 8.1. *Пусть* η, ζ *– некоторые СВ,* B *– случайное событие. Тогда*

$$P(B/\eta = x) = \int_{-\infty}^{+\infty} P(B/\eta = x, \zeta = y)dF_{\zeta}(y/\eta = x), \qquad (8.1)$$

где $F_{\zeta}(y/\eta = x)$ *– условная ф.р. СВ* ζ *при условии, что* $\eta = x$.

Доказательство. Введем случайные события

$$A_i = (y_i < \zeta \le y_{i+1}),\, i = 1,2\, ...,$$

где y_i, $i = 1,\ 2 \ldots$ – точки разбиения переменной

$$y,\ y_1 < y_2 < \ldots < y_i < y_{i+1} < \ldots \, .$$

Тогда

$$A_i \bigcap A_j = \varnothing,\, i \ne j,\, \bigcup_i A_i = \Omega,\, i, j = 1,2,\ldots.$$

Далее, используя определение условной вероятности и условной ф.р., будем иметь

$$P(B/\eta = x) = P(B \bigcap \Omega / \eta = x) = P(B \bigcap (\bigcup_i A_i)/\eta = x) =$$

$$= \sum_i P(B \bigcap A_i / \eta = x) =$$

$$= \sum_i P(B/\eta = x, A_i)P(A_i/\eta = x) =$$

$$\sum_i P(B/\eta = x,\, y_i < \zeta \le y_{i+1})\big[F_\zeta(y_{i+1}/\eta = x) - F_\zeta(y_i/\eta = x)\big] \xrightarrow[\Delta y_i \to 0]{}$$

$$\xrightarrow[\Delta y_i \to 0]{} \int_{-\infty}^{+\infty} P(B/\eta = x,\, \zeta = y)\, dF_\xi(y/\eta = x).$$

Лемма доказана.

Говорят, что случайный процесс $\xi(t)$ **непрерывен**, если за малые промежутки времени лишь с малой вероятностью он может получить заметные по величине приращения [3].

Пусть $\xi(t)$ – марковский процесс. Полная вероятностная характеристика этого процесса определяется функцией

$$F(t, x; \tau, y) = P(\xi(\tau) \le y / \xi(t) = x),\ \ \tau > t,$$

которая называется **переходной функцией марковского процесса** (марковской переходной функцией).

Определение [10]. Случайный процесс $\xi(t)$ называется **непрерывным марковским**, если для любых моментов времени $s, t, \tau \in T$, таких, что $0 \le s < t < \tau$, и любых действительных y выполняется равенство

$$P(\xi(\tau) \le y / \xi(t) = x, \xi(s) = z) = P(\xi(\tau) \le y / \xi(t) = x) = F(t, x; \tau, y). \qquad (8.2)$$

Теорема 8.1. *Переходная функция непрерывного марковского процесса удовлетворяет* ***обобщенному уравнению Маркова*** *(уравнению Чепмена–Колмогорова)*

$$F(t, x; \tau, y) = \int_{-\infty}^{+\infty} F(s, \mathrm{z}; \tau, y)\, d_z F(t, \mathrm{x}; s, \mathrm{z}) \quad \forall\, 0 \le t < s < \tau. \qquad (8.3)$$

Доказательство. В соотношении (8.1) положим

$$B = \{\xi(\tau) \le y\},\ \ \eta = \xi(t),\ \ \zeta = \xi(s),$$

а y заменим на *z*. Тогда, используя (8.2), имеем

$$P(\xi(\tau) \le y / \xi(t) = x) =$$

$$= \int_{-\infty}^{+\infty} P(\xi(\tau) \le y / \xi(t) = x, \xi(s) = z)\, dP(\xi(s) \le z / \xi(t) = x) =$$

$$= \int_{-\infty}^{+\infty} P(\xi(\tau) \le y / \xi(s) = z) dP(\xi(s) \le z / \xi(t) = x),$$

откуда и следует (8.3).

Отметим некоторые свойства, которым должна удовлетворять функция $F(t, x; \tau, y)$. Прежде всего для нее, как для ф.р., должны быть выполнены соотношения:

1) $0 \le F(t, x; \tau, y) \le 1$,
2) $\lim\limits_{y \to -\infty} F(t, x; \tau, y) = 0$,
3) $\lim\limits_{y \to +\infty} F(t, x; \tau, y) = 1$.

Кроме того, функция $F(t, x; \tau, y)$ непрерывна справа относительно аргумента y и

$$\lim_{\tau \to t+0} F(t, x; \tau, y) = \begin{cases} 1, & y \ge x, \\ 0, & y < x. \end{cases}$$

Если существует плотность

$$f(t, x; \tau, y) = \frac{\partial}{\partial y} F(t, x; \tau, y),$$

то для нее выполняются следующие очевидные равенства

$$\int_{-\infty}^{y} f(t, x; \tau, z) dz = F(t, x; \tau, y),$$

$$\int_{-\infty}^{+\infty} f(t, x; \tau, z) dz = 1,$$

а обобщенное уравнение Маркова имеет вид

$$f(t, x; \tau, y) = \int_{-\infty}^{+\infty} f(s, z; \tau, y) f(t, x; s, z) dz, \tag{8.4}$$

оно вытекает из (8.3).

§39. ДИФФУЗИОННЫЕ ПРОЦЕССЫ

Среди непрерывных марковских процессов особую роль играют диффузионные процессы, которые являются математической моделью движения частицы в процессе диффузии, а также предельной моделью для дискретных процессов, описывающих различные биологические явления. В настоящее время они имеют широкое применение в экономико-математическом и финансово-математическом моделировании.

Определение. Марковский процесс $\xi(t)$ называется диффузионным, если его переходная функция $F(t, x; \tau, y)$ удовлетворяет следующим условиям.

1. $$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|y-x|>\delta} dF(t-\Delta t, x; t, y) = 0 \ \forall\ \delta > 0. \qquad (8.5)$$

Это более точное условие непрерывности случайного процесса. Оно требует, чтобы вероятность того, что $|\xi(t) - \xi(t-\Delta t)| > \varepsilon$, была величиной более высокого порядка малости, чем Δt.

2. Существует функция $a(t, x)$ такая, что

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|y-x|>\delta} (y-x) d_y F(t-\Delta t, x; t, y) = a(t, x) \ \forall\ \delta > 0. \qquad (8.6)$$

3. Существует предел

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|y-x|>\delta} (y-x)^2 d_y F(t-\Delta t, x; t, y) = b(t, x) \ \forall\ \delta > 0. \qquad (8.7)$$

Заметим, что левые части равенств (8.6), (8.7) зависят от δ, но, в силу определения непрерывности процесса, т.е. в силу (8.5), эта зависимость является лишь кажущейся.

Для того чтобы выяснить физический смысл функций $a(t, x)$ и $b(t, x)$, мы несколько усилим требование непрерывности процесса. А именно, предположим, что вместо (8.5) при любом $\delta > 0$ имеет место соотношение

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|y-x|>\delta} (y-x)^2 \, d_y F(t-\Delta t, x; t, y) = 0 . \tag{8.8}$$

Легко видеть, что из (8.8) следует (8.5). Условия (8.6), (8.7) могут быть записаны теперь в виде

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{-\infty}^{+\infty} (y-x) d_y F(t-\Delta t, x; t, y) = a(t, x) , \tag{8.9}$$

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{-\infty}^{+\infty} (y-x)^2 \, d_y F(t-\Delta t, x; t, y) = b(t, x) . \tag{8.10}$$

Но

$$\int_{-\infty}^{+\infty} (y-x) d_y F(t-\Delta t, x; t, y) = M\left[\xi(t) - \xi(t-\Delta t)\right]$$

является математическим ожиданием изменения процесса $\xi(t)$ за время Δt, а

$$\int_{-\infty}^{+\infty} (y-x)^2 \, d_y F(t-\Delta t, x; t, y) = M\left[\xi(t) - \xi(t-\Delta t)\right]^2$$

есть математическое ожидание квадрата изменения процесса $\xi(t)$ и, следовательно, пропорционально кинетической энергии в предположении, что $\xi(t)$ является координатой движущейся под влиянием случайных воздействий точки. Поэтому из (8.9), (8.10) ясно, что функция $a(t, x)$ есть средняя скорость изменения процесса $\xi(t)$ (и называется **коэффициентом сноса,** или переноса), а функция $b(t, x)$ пропорциональна средней кинетической энергии изучаемой нами системы (называется **коэффициентом диффузии**).

Отметим, что при некоторых ограничениях на $F(t, x; \tau, y)$ задание коэффициентов $a(t, x)$, $b(t, x)$ полностью определяет марковский процесс.

§40. ОБРАТНОЕ УРАВНЕНИЕ КОЛМОГОРОВА

Далее мы введем дифференциальные уравнения в частных производных, которым при выполнении некоторых условий удовлетворяет переходная функция марковского процесса $F(t, x; \tau, y)$ и ее производных. Эти уравнения впервые строго были доказаны А.Н-.Колмогоровым, хотя второе из них встречалось в работах физиков М.Планка и А.Фоккера в связи с развитием теории диффузии, и носят название уравнений Колмогорова.

Теорема 8.2. *Пусть существуют непрерывные частные производные первого и второго порядков*

$$\frac{\partial F(t, x; \tau, y)}{\partial x} \quad \text{и} \quad \frac{\partial^2 F(t, x; \tau, y)}{\partial x^2}$$

при любых значениях t, x и ô$> t$. *И пусть выполнены условия (8.5)-(8.7). Тогда функция* $F(t, x; \tau, y)$ *удовлетворяет уравнению*

$$\frac{\partial F(t, x; \tau, y)}{\partial t} = -a(t, x)\frac{\partial F(t, x; \tau, y)}{\partial x} - \frac{1}{2}b(t, x)\frac{\partial^2 F(t, x; \tau, y)}{\partial x^2}, \quad (8.11)$$

*которое называется **обратным уравнением Колмогорова**.*

Доказательство. Заменив в обобщенном уравнении Маркова (8.3) t на $t - \Delta t, s$ на t и используя условие нормировки для ф.р., получим

$$F(t - \Delta t, x; \tau, y) = \int_{-\infty}^{+\infty} F(t, z; \tau, z) d_z F(t - \Delta t, x; t, z),$$

$$\int_{-\infty}^{+\infty} d_z F(t - \Delta t, x; t, z) = 1.$$

Тогда

$$\frac{\partial F(t,x;\tau,y)}{\partial t}=-\lim_{\Delta t\to 0}\frac{1}{\Delta t}\left[F(t-\Delta t,x;\hat{o},y)-F(t,x;\hat{o},y)\right]=$$

$$=-\lim_{\Delta t\to 0}\frac{1}{\Delta t}\left[\int_{-\infty}^{+\infty}F(t,z;\hat{o},y)d_zF(t-\Delta t,x;t,z)-\int_{-\infty}^{+\infty}F(t,x;\hat{o},y)d_zF(t-\Delta t,x;t,z)\right]=$$

$$=-\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{-\infty}^{+\infty}\left[F(t,z;\hat{o},y)-F(t,x;\hat{o},y)\right]d_zF(t-\Delta t,x;t,z)=$$

$$=-\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|z-x|>\delta}\left[F(t,z;\hat{o},y)-F(t,x;\hat{o},y)\right]d_zF(t-\Delta t,x;t,z)+$$

$$-\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|z-x|\le\delta}\left[F(t,z;\tau,y)-F(t,x;\tau,y)\right]d_z\,F(t-\Delta t,x;t,z)\cdot\qquad(8.12)$$

При сделанных нами предположениях по формуле Тейлора имеет место равенство

$$F(t,z;\tau,y)=F(t,x;\tau,y)+(z-x)\frac{\partial F(t,x;\tau,y)}{\partial x}+$$

$$+\frac{1}{2}(z-x)^2\,\frac{\partial^2F(t,x;\tau,y)}{\partial x^2}+o\left((z-x)^2\right)\cdot$$

Тогда, подставляя это выражение в (8.12) и используя (8.5) – (8.7), далее будем иметь

$$-\lim_{\Delta t\to 0}\frac{1}{\Delta t}\left[o(\Delta t)+\int_{|z-x|\le\delta}\left[(z-x)\frac{\partial F(t,x;\tau,y)}{\partial x}+\right.\right.$$

$$\left.\left.+\frac{1}{2}(z-x)^2\,\frac{\partial^2F(t,x;\tau,y)}{\partial x^2}+o\left((z-x)^2\right)\right]d_zF(t-\Delta t,x;t,z)\right]=$$

$$=-\frac{\partial F(t,x;\tau,y)}{\partial x}\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|z-x|\le\delta}(z-x)\,d_zF(t-\Delta t,x;t,z)-$$

$$-\frac{1}{2}\frac{\partial^2 F(t,x;\tau,y)}{\partial x^2}\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|z-x|\le\delta}(z-x)^2\,d_zF(t-\Delta t,x;t,z)-$$

$$-\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|z-x|\le\delta}o\big((z-x)^2\big)d_zF(t-\Delta t,x;t,z)\cdot \qquad (8.13)$$

Последнее слагаемое стремится к нулю при $\delta\to 0$. Но так как левая часть равенства (8.12) от δ не зависит и пределы (8.6) и (8.7) также от δ не зависят, выражение (8.13) существует и равно

$$-a(t,x)\frac{\partial F(t,x;\tau,y)}{\partial x}-\frac{1}{2}b(t,x)\frac{\partial^2 F(t,x;\tau,y)}{\partial x^2},$$

что и приводит нас к уравнению (8.11).

Если предположить, что существует плотность распределения

$$f(t,x;\tau,y)=\frac{\partial}{\partial y}F(t,x;\tau,y), \qquad (8.14)$$

то простое дифференцирование (8.11) показывает, что она удовлетворяет уравнению

$$\frac{\partial f(t,x;\tau,y)}{\partial t}+a(t,x)\frac{\partial f(t,x;\tau,y)}{\partial x}+\frac{1}{2}b(t,x)\frac{\partial^2 f(t,x;\tau,y)}{\partial x^2}=0\cdot$$

§41. УРАВНЕНИЕ КОЛМОГОРОВА–ФОККЕРА–ПЛАНКА

Перейдем теперь к выводу прямого уравнения Колмогорова, которое является сопряженным к обратному.

Теорема 8.3. *Пусть имеется диффузионный процесс, т.е. выполняются соотношения (8.5) – (8.7), существует плотность распределения (8.14) и существуют производные*

$$\frac{\partial f(t,x;\tau,y)}{\partial\tau},\ \frac{\partial}{\partial y}\big[a(\tau,y)f(t,x;\tau,y)\big],\ \frac{\partial^2}{\partial y^2}\big[b(\tau,y)f(t,x;\tau,y)\big]\cdot$$

Тогда плотность $f(t,x;\tau,y)$ *удовлетворяет уравнению*

$$\frac{\partial f(t,x;\tau,y)}{\partial \tau} = -\frac{\partial}{\partial y}\left[a(\tau,y) f(t,x;\tau,y)\right] + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[b(\tau,y) f(t,x;\tau,y)\right], \qquad (8.15)$$

которое называется ***прямым уравнением Колмогорова (уравнением Колмогорова–Фоккера–Планка).***

Доказательство. Пусть $R(y)$ – неотрицательная непрерывная функция такая, что

$$R(y) = 0, \text{ если } y < a \text{ и } y > b$$

$$R(a) = R(b) = R'(a) = R'(b) = R''(a) = R''(b) = 0\,, \qquad (8.16)$$

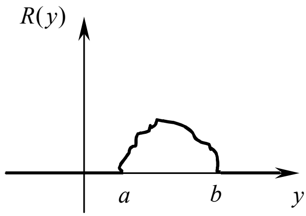


Рис. 36

Рассмотрим интеграл

$$I = \int_a^b \frac{\partial f(t,x;\tau,y)}{\partial \tau} R(y)dy = \lim_{\Delta\tau \to 0} \int_{-\infty}^{\infty} \frac{f(t,x;\tau+\Delta\tau,y) - f(t,x;\tau,y)}{\partial \tau} R(y)dy.$$

Согласно обобщенному уравнению Маркова (8.4)

$$f(t,x;\tau+\Delta\tau,y) = \int_{-\infty}^{\infty} f(s,z;\tau+\Delta\tau,y) f(t,x;s,z)dz$$

и, если заменить s на τ, то

$$f(t,x;\tau+\Delta\tau,y) = \int_{-\infty}^{\infty} f(\tau,z;\tau+\Delta\tau,y) f(t,x;\tau,z)dz\,.$$

Поэтому

$$I = \lim_{\Delta\tau \to 0} \frac{1}{\Delta\tau} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(t,z;\tau+\Delta\tau,y) f(t,x;\tau,z)dz - f(t,x;\tau,y) \right] R(y)dy =$$

$$= \lim_{\Delta\tau \to 0} \frac{1}{\Delta\tau}\left[\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(t,x;\tau,z) f(t,z;\tau+\Delta\tau,y) R(y)dz\,dy - \int\limits_{-\infty}^{\infty} f(t,x;\tau,y)R(y)dy\right].$$

После замены в двойном интеграле y на z, а z на y, получим

$$I = \lim_{\Delta\tau \to 0} \frac{1}{\Delta\tau}\int\limits_{-\infty}^{\infty} f(t,x;\tau,y)\left[\int\limits_{-\infty}^{\infty} f(\tau,y;\tau+\Delta\tau,z)R(z)dz - R(y)\right]dy. \quad (8.17)$$

Разложим функцию $R(z)$ в ряд Тейлора в окрестности точки $z = y$:

$$R(z) = R(y) + (z-y)R'(y) + \frac{1}{2}(z-y)^2 R''(y) + o\left((z-y)^2\right). \qquad (8.18)$$

Так как в силу ограниченности функции $R(z)$ и условия (8.5)

$$\int\limits_{|y-z|>\delta} f(\tau,y;\tau+\Delta\tau,z)R(z)dz = o(\Delta\tau)$$

и

$$\int\limits_{|y-z|\le\delta} f(\tau,y;\tau+\Delta\tau,z)dz = 1 + o(\Delta\tau),$$

то выражение в квадратной скобке в (8.17) с учетом разложения (8.18) принимает вид

$$\int\limits_{-\infty}^{\infty} f(\tau,y;\tau+\Delta\tau,z)R(z)dz - R(y) =$$

$$= R'(y)\int\limits_{|y-z|\le\delta}(z-y) f(\tau,y;\tau+\Delta\tau,z)dz +$$

$$+ \frac{1}{2}R''(y)\int\limits_{|y-z|\le\delta}\left[(z-y)^2 + o\left((z-y)^2\right)\right] f(\tau,y;\tau+\Delta\tau,z)dz + o(\Delta\tau).$$

Таким образом,

$$I = \lim_{\Delta\tau \to 0} \frac{1}{\Delta\tau}\int\limits_{-\infty}^{\infty} f(t,x;\tau,y)\left\{ R'(y)\int\limits_{|y-z|\le\delta}(z-y) f(\tau,y;\tau+\Delta\tau,z)dz + \right.$$

$$\left. + \frac{1}{2}R''(y)\int\limits_{|y-z|\le\delta}\left[(z-y)^2 + o\left((z-y)^2\right)\right] f(\tau,y;\tau+\Delta\tau,z)dz + o(\Delta\tau)\right\}dy$$

Далее, с учетом (8.6), (8.7), (8.16), имеем

$$I = \int_{-\infty}^{\infty} f(t,x;\tau,y)\left[a(\tau,y)R'(y) + \frac{1}{2}b(\tau,y)R''(y)\right]dy =$$

$$= \int_{a}^{b} f(t,x;\tau,y)\left[a(\tau,y)R'(y) + \frac{1}{2}b(\tau,y)R''(y)\right]dy \cdot \qquad (8.19)$$

Воспользовавшись формулой интегрирования по частям и равенствами (8.16), находим, что

$$\int_a^b f(t,x;\tau,y)\,a(\tau,y)R'(y)dy = -\int_a^b R(y)\frac{\partial}{\partial y}\left[a(\tau,y)f(t,x;\tau,y)\right]dy,$$

$$\int_a^b f(t,x;\tau,y)\,b(\tau,y)R''(y)dy = \int_a^b R(y)\frac{\partial^2}{\partial y^2}\left[b(\tau,y)f(t,x;\tau,y)\right]dy \cdot$$

В результате подстановки полученных выражений в (8.19) получаем

$$\int_a^b \frac{\partial f(t,x;\tau,y)}{\partial \tau}R(y)dy = \int_a^b \left\{-\frac{\partial}{\partial y}\left[a(\tau,y)f(t,x;\tau,y)\right] + \right.$$

$$\left. + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[b(\tau,y)f(t,x;\tau,y)\right]\right\}R(y)dy$$

или

$$\int_a^b \left\{\frac{\partial f(t,x;\tau,y)}{\partial \tau} + \frac{\partial}{\partial y}\left[a(\tau,y)f(t,x;\tau,y)\right] - \right.$$

$$\left. - \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[b(\tau,y)f(t,x;\tau,y)\right]\right\}R(y)dy = 0 \cdot$$

Так как на промежутке $[a,b]$ функция $R(y)$ не меняет знак, последнее соотношение может быть справедливо только тогда, когда

$$\frac{\partial f(t,x;\tau,y)}{\partial \tau} + \frac{\partial}{\partial y}\left[a(\tau,y)f(t,x;\tau,y)\right] - \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[b(\tau,y)f(t,x;\tau,y)\right] = 0 \cdot$$

Таким образом, получили уравнение Колмогорова–Фоккера–Планка, которое вместе с начальным условием

$$f(t,x;t,y)=\delta(y-x)$$

однозначно определяет функцию $f(t,x;\tau,y)\ \ \forall\ \tau>t$.

В заключение рассмотрим частный случай уравнений Колмогорова, когда функция $f(t,x;\tau,y)$ зависит от t, τ и $y-x$, но не зависит от самих x и y. Физически это означает, что процесс $\xi(t)$ протекает однородно в пространстве: вероятность получить приращение $\Delta=y-x$ не зависит от того, в каком положении x находится процесс в момент времени t. Очевидно, что в этом случае функции $a(t,x)$ и $b(t,x)$ не зависят от x, а являются функциями только от одного аргумента t:

$$a(t,x)=a(t),\ b(t,x)=b(t)\cdot$$

Обратное и прямое уравнения Колмогорова в этом случае имеют вид

$$\begin{cases}\dfrac{\partial f}{\partial t}=-a(t)\dfrac{\partial f}{\partial x}-\dfrac{1}{2}b(t)\dfrac{\partial^2 f}{\partial x^2},\\ \dfrac{\partial f}{\partial \tau}=-a(\tau)\dfrac{\partial f}{\partial y}+\dfrac{1}{2}b(\tau)\dfrac{\partial^2 f}{\partial y^2}.\end{cases}\tag{8.20}$$

Пусть $a(t)=0,\ b(t)=1$, уравнения (8.20) при этом превращаются в уравнение теплопроводности

$$\frac{\partial f}{\partial \tau}=\frac{1}{2}\frac{\partial^2 f}{\partial y^2}\tag{8.21}$$

и ему сопряженное

$$\frac{\partial f}{\partial t}=-\frac{1}{2}\frac{\partial^2 f}{\partial x^2}\cdot\tag{8.22}$$

Из общей теории теплопроводности известно, что единственное решение этих уравнений, удовлетворяющее условию

$$\lim_{\tau\to t}\int\limits_{|y-x|>\delta} f(t,x;\tau,y)dy=0,$$

которое следует из (8.5), имеет вид

$$f(t,x;\tau,y)=\frac{1}{\sqrt{2\pi(\tau-t)}}e^{-\frac{(y-x)^2}{2(\tau-t)}}.$$

Если сделать замену переменных

$$x'=x-\int_a^t a(z)\,dz,\qquad y'=y-\int_a^\tau b(z)\,dz,$$

$$t'=\int_a^t b(z)\,dz,\qquad \tau'=\int_a^\tau b(z)\,dz,$$

то уравнения (8.20) сводятся к уравнениям (8.21), (8.22). Отсюда следует, что решения уравнений (8.20) можно записать в виде

$$f(t,x;\tau,y)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-x-\alpha)^2}{2\sigma^2}},$$

где

$$\alpha=\int_t^\tau a(z)\,dz,\qquad \sigma^2=\int_t^\tau b(z)\,dz,$$

т.е. плотность $f(t,x;\tau,y)$ является нормальной.

§42. ДОПРЕДЕЛЬНЫЕ МОДЕЛИ ДИФФУЗИОННЫХ ПРОЦЕССОВ

Рассмотрим еще одно применение диффузионных марковских процессов. Представим частицу, совершающую скачкообразные движения по оси Ox. За время Δt частица совершает скачок на величину $\pm\delta$ с вероятностью $\frac{1}{2}$, причем скачки совершаются независимо друг от друга. Пусть $x(T)$ – координата частицы в момент времени T, а Δx_k – скачок в k-й момент времени. Очевидно, что

$$x(T)=\sum_{k=1}^{n}\Delta x_k\text{, где } n=\frac{T}{\Delta t}. \tag{8.23}$$

Найдем среднее значение и дисперсию положения частицы в момент T:

$$M\{x(T)\} = M\left\{\sum_{k=1}^{n} \Delta x_k\right\} = \sum_{k=1}^{n} M\{\Delta x_k\} = 0,$$

$$D\{x(T)\} = D\left\{\sum_{k=1}^{n} \Delta x_k\right\} = \sum_{k=1}^{n} D\{\Delta x_k\} = \sum_{k=1}^{n} M\left\{(\Delta x_k)^2\right\} = \delta^2 n,$$

или, согласно (8.23),

$$D\{x(T)\} = \delta^2 \frac{T}{\Delta t} = T \frac{\delta^2}{\Delta t}.$$

Случаи, когда $\lim\limits_{\Delta t \to 0} \dfrac{\delta^2}{\Delta t} = \infty \vee 0$ для практики малоинтересны.

Рассмотрим случай, когда

$$\lim_{\Delta t \to 0} \frac{\delta^2}{\Delta t} = const,$$

и пусть $\delta^2 = b\Delta t$, тогда $D\{x(T)\} = bT$ – конечная величина. Для разностей

$$x(t_2) - x(t_1) = \sum_{t_1 \le k\Delta t \le t_2} \Delta x_k, \quad t_2 > t_1,$$

имеем

$$M\{x(t_2) - x(t_1)\} = 0,$$

$$D\{x(t_2) - x(t_1)\} = b|t_2 - t_1|.$$

Тогда в силу центральной предельной теоремы

$$p(x_2, t_2 / x_1, t_1) = \frac{1}{\sqrt{2\pi b|t_2 - t_1|}} e^{-\frac{[x_2 - x_1]^2}{2b|t_2 - t_1|}},$$

и все разности по предположению независимы. Поэтому в пределе получим, что $x(T)$ – винеровский процесс. Остановимся на особенностях траекторий этого процесса. Как определено выше, приращение процесса за время Δt равно $\delta = \sqrt{b\Delta t}$ и $\delta \to 0$ при $\Delta t \to 0$, т.е. траектории винеровского процесса непрерывны. С другой стороны,

$$\lim_{\Delta t\to 0}\frac{\Delta x}{\Delta t}=\lim_{\Delta t\to 0}\frac{\pm\delta}{\Delta t}=\lim_{\Delta t\to 0}\frac{\pm\sqrt{b\Delta t}}{\Delta t}=\pm\infty,$$

поэтому производная не существует. Можно сказать, что траектории винеровского процесса ни в одной точке не имеют производной. Это утверждение верно для всех диффузионных процессов.

В данном случае случайное блуждание частицы описывается диффузионным процессом с коэффициентом сноса $a(t,x)=0$ и коэффициентом диффузии $b(t,x)=b$.

В заключение этого параграфа докажем лемму, которая понадобится нам в следующей главе.

Лемма 8.2. *Пусть* $\xi(t)$ *– винеровский процесс с нулевым коэффициентом сноса и единичным коэффициентом диффузии. Тогда для* $0\le t_1<t_2<...<t_n=T$

$$S_n=\sum_{i=1}^{n-1}\left[\xi(t_{i+1})-\xi(t_i)\right]^2\xrightarrow[\max_i|t_{i+1}-t_i|\to 0]{\text{ср.кв.}}T.$$

Доказательство. По определению сходимости в среднем квадратичном нужно доказать, что

$$M\left\{(S_n-T)^2\right\}\xrightarrow[n\to\infty]{}0.$$

Обозначим $\eta_i=\left[\xi(t_{i+1})-\xi(t_i)\right]^2$, тогда $S_n=\sum\limits_{i=1}^{n-1}\eta_i$. Так как $\xi(t)$ – винеровский процесс, то $\left[\xi(t_{i+1})-\xi(t_i)\right]$ – нормально распределенные СВ и

$$M\left[\xi(t_{i+1})-\xi(t_i)\right]=0,\quad D\left[\xi(t_{i+1})-\xi(t_i)\right]=|t_{i+1}-t_i|.$$

Тогда

$$M\{\eta_i\}=D\left[\xi(t_{i+1})-\xi(t_i)\right]=t_{i+1}-t_i,$$

$$D\{\eta_i\}=M\left\{\eta_i^{\,2}\right\}-M^2\{\eta_i\}.$$

Величины $M\left\{\eta_i^{\,2}\right\}$ можно найти, используя выражение для четных моментов нормальных СВ:

$$M\left\{\xi^{2k}\right\}=(2k-1)!!\sigma^{2k}=1\cdot 3\cdot 5\cdot\ldots\cdot(2k-1)\sigma^{2k},$$

значит,

$$M\left\{\eta_i^{\ 2}\right\} = M\left\{\left[\xi(t_{i+1}) - \xi(t_i)\right]^4\right\} = 3(t_{i+1} - t_i)^2,$$

и

$$D\left\{\eta_i\right\} = 2(t_{i+1} - t_i)^2 \cdot$$

Поэтому

$$M\left\{S_n\right\} = T$$

$$D\left\{S_n\right\} = 2\sum_{i=1}^{n-1}(t_{i+1} - t_i)^2 \cdot$$

Рассмотрим теперь выражение

$$M\left\{(S_n - T)^2\right\} = M\left\{S_n^{\ 2}\right\} - 2M\left\{S_n\right\}T + T^2 =$$

$$= D\left\{S_n\right\} + T^2 - 2TM\left\{S_n\right\} + T^2 = D\left\{S_n\right\} =$$

$$= 2\sum_{i=1}^{n-1}(t_{i+1} - t_i)^2 \le 2T \max_i \left|t_{i+1} - t_i\right| \cdot$$

Так как T конечно, то при $\max\limits_i \left|t_{i+1} - t_i\right| \to 0$ $M\left\{(S_n - T)^2\right\} \to 0$, что и доказывает сходимость $\left(S_n - T\right)$ в среднем квадратичном.

ГЛАВА 9. СТОХАСТИЧЕСКИЕ ИНТЕГРАЛЫ И ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

В данной главе мы будем рассматривать интегралы вида

$$\int_a^b \varphi(\xi(t),t)d\xi(t), \tag{9.1}$$

где $\varphi(\cdot)$ – неслучайная функция, $\xi(t)$ – случайный процесс. Реализации процесса $\xi(t)$ в общем случае являются функциями неограниченной вариации, следовательно, этот интеграл нельзя понимать как интеграл Стилтьеса или Лебега–Стилтьеса.

Определение. Интеграл вида (9.1) называют **стохастическим интегралом,** если $\xi(t)$ – диффузионный процесс.

Как следует из §42, не совсем понятно, что понимать в этом случае под дифференциалом $d\xi(t)$. Выясним это, рассмотрев два более подробных определения стохастического интеграла.

§43. СТОХАСТИЧЕСКИЙ ИНТЕГРАЛ В ФОРМЕ ИТО

Разобьем интервал $[a,b]$ точками $a = t_1 < t_2 < ... < t_{n-1} = b$, и рассмотрим интегральную сумму

$$S_n = \sum_{i=1}^{n-1} \varphi(\xi(t_i),t_i)[\xi(t_{i+1}) - \xi(t_i)].$$

Если существует предел в среднем квадратичном

$$\underset{\substack{n\to\infty \\ \max_i (t_{i+1}-t_i)\to 0}}{l.i.m.} \sum_{i=1}^{n-1} \varphi(\xi(t_i),t_i)[\xi(t_{i+1}) - \xi(t_i)] = \int_a^b \varphi(\xi(t),t)d\xi(t),$$

то он называется **стохастическим интегралом в форме Ито.**

Так как диффузионный процесс $\xi(t)$ недифференцируем ни в одной точке, то интеграл Ито обладает рядом свойств, отличных от интеграла Римана. Покажем это отличие на частном примере.

Пример 9.1. Пусть надо вычислить интеграл

$$I = \int_{t_0}^{t} [\xi(\tau) - \xi(t_0)] d\xi(\tau),$$

где $\xi(\tau)$ – винеровский процесс с нулевым математическим ожиданием и единичной дисперсией, $M\xi(\tau) = 0$, $D\xi(\tau) = 1$.

Если этот интеграл понимать как интеграл Римана, то

$$I = \int_{t_0}^{t} [\xi(\tau) - \xi(t_0)] d[\xi(\tau) - \xi(t_0)] =$$

$$= \int_{t_0}^{t} d\left\{ \frac{[\xi(\tau) - \xi(t_0)]^2}{2} \right\} = \frac{[\xi(\tau) - \xi(t_0)]^2}{2}. \tag{9.2}$$

Но так как $\xi(t)$ недифференцируем, то интеграл I вычислим как предел интегральной суммы. Для этого разобьем область интегрирования на $n-1$ частей точками $t_0 = t_1 < t_2 < \ldots < t_{n-1} = t$. Пусть

$$\delta_{i+1} = \xi(t_{i+1}) - \xi(t_i).$$

Тогда

$$\xi(t_i) - \xi(t_0) = \sum_{k=1}^{i-1} \delta_k,$$

и по определению интеграла Ито имеем

$$I = \underset{n\to\infty}{l.i.m.} \sum_{i=1}^{n-1} [\xi(t_i) - \xi(t_0)][\xi(t_{i+1}) - \xi(t_i)] =$$

$$= \underset{n\to\infty}{l.i.m.} \sum_{i=1}^{n-1} \left(\sum_{k=1}^{i-1} \delta_k \right) \delta_i .$$

Чтобы вычислить двойную сумму в I, воспользуемся соотношением

$$\left(\sum_{k=1}^{n-1} \delta_k \right)^2 = \sum_{k=1}^{n-1} \delta_k^{\ 2} + 2 \sum_{i=1}^{n-1} \delta_i \left(\sum_{k=1}^{i-1} \delta_k \right).$$

Отсюда следует

$$\sum_{i=1}^{n-1}\delta_i\left(\sum_{k=1}^{i-1}\delta_k\right)=\frac{1}{2}\left[\left(\sum_{k=1}^{n-1}\delta_k\right)^2-\sum_{k=1}^{n-1}\delta_k^2\right].$$

Тогда

$$I=\underset{n\to\infty}{l.i.m.}\ \frac{1}{2}\left[\left(\sum_{k=1}^{n-1}\delta_k\right)^2-\sum_{k=1}^{n-1}\delta_k{}^2\right]=$$

$$=\frac{1}{2}\left[\xi(t)-\xi(t_0)\right]^2-\underset{n\to\infty}{l.i.m.}\frac{1}{2}\sum_{i=1}^{n-1}\left[\xi(t_{i+1})-\xi(t_i)\right]^2.$$

По лемме 8.2 второе слагаемое равно $t-t_0$, поэтому

$$I=\frac{\left[\xi(t)-\xi(t_0)\right]^2}{2}-\frac{t-t_0}{2}.$$

Сравнивая это выражение с интегралом Римана (9.2), видим, что они отличаются на величину $\frac{t-t_0}{2}$ за счет недифференцируемости траекторий винеровского процесса $\xi(t)$.

§44. СТОХАСТИЧЕСКИЙ ИНТЕГРАЛ В ФОРМЕ СТРАТАНОВИЧА

При решении прикладных задач обычно имеют дело с допредельными моделями броуновского движения, траектории которых непрерывны. Поэтому желательно ввести такое определение стохастического интеграла, в котором его значение было бы устойчивым по отношению к предельному переходу от процессов с дифференцируемыми траекториями к винеровскому процессу. Таким определением является определение в форме Стратановича.

Определение. Пусть имеем разбиение интервала интегрирования $[a,b]$ точками $a=t_1<t_2<\ldots<t_{n-1}=b$. Рассмотрим интегральную сумму

$$S_n = \sum_{i=1}^{n-1} \varphi\left(\frac{\xi(t_{i+1}) + \xi(t_i)}{2}, t_i\right)[\xi(t_{i+1}) - \xi(t_i)] \cdot \qquad (9.3)$$

Если при $\max_i (t_{i+1} - t_i) \to 0$ существует прием в среднем квадратичном, то он называется **стохастическим интегралом в форме Стратановича,** или симметризованным стохастическим интегралом, и обозначается

$$\oint_a^b \varphi(\xi(t), t) d\xi(t) = \underset{n \to \infty}{l.i.m.} \sum_{i=1}^{n-1} \varphi\left(\frac{\xi(t_{i+1}) - \xi(t_i)}{2}, t_i\right)[\xi(t_{i+1}) - \xi(t_i)] \cdot$$

Отличие в аргументе подынтегральной функции обеспечивает необходимую устойчивость стохастического интеграла к предельному переходу.

Рассмотрим связь интеграла в форме Ито и в форме Стратановича. Предположим, что функция $\varphi(x, y)$ дифференцируема по первому аргументу. Используя разложение в ряд Тэйлора в окрестности точки $\xi(t_i)$, имеем

$$\varphi\left(\frac{\xi(t_{i+1}) + \xi(t_i)}{2}, t_i\right) = \varphi\left(\xi(t_i) + \frac{\xi(t_{i+1}) - \xi(t_i)}{2}, t_i\right) =$$

$$= (\xi(t_i), t_i) + \frac{1}{2}[\xi(t_{i+1}) - \xi(t_i)]\frac{\partial \varphi}{\partial \xi} + \frac{1}{4}[\xi(t_{i+1}) - \xi(t_i)]^2 \frac{\partial^2 \varphi}{\partial \xi^2} + \ldots \cdot$$

Подставляя это разложение в сумму (9.3) и переходя к пределу при $n \to \infty$, получаем

$$\oint_a^b \varphi(\xi(t), t) d\xi(t) = \int_a^b \varphi(\xi(t), t) d\xi(t) +$$

$$+ \frac{1}{2} \underset{n \to \infty}{l.i.m.} \sum_{i=1}^{n-1} \frac{\partial \varphi(\xi(t_i), t_i)}{\partial \xi}[\xi(t_{i+1}) - \xi(t_i)] + \ldots = \int_a^b \varphi(\xi(t), t) d\xi(t) +$$

$$+ \frac{1}{2}\int_a^b \frac{\partial \varphi(\xi(t), t)}{\partial \xi} d\xi(t) \cdot \qquad (9.4)$$

Так как интеграл Стратановича больше отвечает прикладным задачам, а интеграл в форме Ито более удобен в теоретических преобразованиях, то решение конкретной задачи можно начинать с использования представления Ито, а затем по формуле (9.4) перейти к интегралу в форме Стратановича. Если $\varphi(\xi(t),t)$ не зависит от $\xi(t)$, то данные интегралы совпадают.

Пример 9.2. Найдем среднее значение и корреляционную функцию случайного процесса, заданного выражением

$$\eta(t)=\int_0^1 e^{i\tau t}d_\tau\xi(\tau),$$

где $\xi(\tau)$ – винеровский процесс, $\tau\in T=[0,+\infty)$.

Мы знаем (п. 4.4), что для винеровского процесса

$$M\xi(\tau)=0,\ D\xi(\tau)=\sigma^2\tau,$$

$$D(\xi(\tau_2)-\xi(\tau_1))=M\left[(\xi(\tau_2)-\xi(\tau_1))^2\right]=M\xi^2(\tau_2)-2M[\xi(\tau_2)\xi(\tau_1)]+$$
$$+M\xi^2(\tau_1)=\sigma^2(\tau_2-2\tau_1+\tau_1)=\sigma^2(\tau_2-\tau_1).$$

Заметим, что последнее выражение можно представить в виде

$$D(\xi(\tau_2)-\xi(\tau_1))=\int_{\tau_1}^{\tau_2}s(t)dt,$$

где $s(t)=\sigma^2$. Найдем среднее значение процесса $\eta(t)$:

$$M\eta(t)=M\left(\int_0^1 e^{i\tau t}d_\tau\xi(\tau)\right)=M\left(\underset{n\to\infty}{l.i.m.}\sum_{k=1}^{n-1}e^{i\tau_k t}[\xi(\tau_{k+1})-\xi(\tau_k)]\right)=$$
$$=\underset{n\to\infty}{l.i.m.}\sum_{k=1}^{n-1}e^{i\tau_k t}M[\xi(\tau_{k+1})-\xi(\tau_k)]=0.$$

Найдем корреляционную функцию комплекснозначного процесса $\eta(t)$ при $t_1\neq t_2$:

$$R_\eta(t_1,t_2)=M\left[\eta(t_1)\overline{\eta(t_2)}\right]=M\left(\int_0^1 e^{it_1\tau}d_\tau\xi(t)\overline{\int_0^1 e^{it_2\tau}d_\tau\xi(t)}\right)=$$

$$=M\,\underset{n\to\infty}{l.i.m.}\sum_{k=1}^{n-1}e^{it_1\tau_k}\left[\xi(\tau_{k+1})-\xi(\tau_k)\right]\overline{\sum_{k=1}^{n-1}e^{it_2\tau_k}\left[\xi(\tau_{k+1})-\xi(\tau_k)\right]}=$$

$$=\underset{n\to\infty}{l.i.m.}\sum_{k=1}^{n-1}e^{it_1\tau_k}\,\overline{e^{it_2\tau_k}}\,D\left[\xi(\tau_{k+1})-\xi(\tau_k)\right]=$$

$$=\underset{n\to\infty}{l.i.m.}\sum_{k=1}^{n-1}e^{it_1\tau_k}\,\overline{e^{it_2\tau_k}}\int_{\tau_k}^{\tau_{k+1}}s(t)dt=\int_0^1 e^{it_1\tau}e^{-it_2\tau}s(\tau)d\tau=$$

$$=\sigma^2\int_0^1 e^{i(t_1-t_2)\tau}d\tau=\frac{\sigma^2}{i(t_1-t_2)}\left[e^{i(t_1-t_2)}-1\right].$$

Для $t_1=t_2$ имеем

$$R_\eta(t,t)=M\left[\eta(t)\overline{\eta(t)}\right]=\sigma^2\int_0^1 dt=\sigma^2\,.$$

Таким образом,

$$R_\eta(t_1,t_2)=\begin{cases}\dfrac{\sigma^2}{i(t_1-t_2)}\left[e^{i(t_1-t_2)}-1\right], & t_1\neq t_2,\\ \sigma^2, & t_1=t_2.\end{cases}$$

§45. СТОХАСТИЧЕСКИЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

Многие реальные объекты описываются дифференциальными уравнениями вида

$$\frac{dX(t)}{dt}=a\big(X(t),t\big)+\sigma\big(X(t),t\big)\frac{d\xi(t)}{dt} \tag{9.5}$$

или в форме дифференциалов

$$dX(t)=a\big(X(t),t\big)+\sigma\big(X(t),t\big)d\xi(t)\,. \tag{9.6}$$

Если $\xi(t)$ – винеровский случайный процесс, то дифференциальное уравнение называется **стохастическим.** Стохастические дифференциальные уравнения принято обычно писать в форме дифференциалов, поскольку винеровский процесс является недифференцируемым.

Решение уравнений (9.5), (9.6) определяется интегральным равенством

$$X(t) = X(t_0) + \int_{t_0}^{t} a(X(\tau),\tau)d\tau + \int_{t_0}^{t} \sigma(X(\tau),\tau)d\xi(\tau), \tag{9.7}$$

где первый интеграл является стохастическим в среднеквадратичном смысле (см. §16), второй интеграл является стохастическим в форме Ито или в форме Стратановича, $X(t_0)$ – значение процесса $X(t)$ в начальный момент времени. (9.7) называется **стохастическим уравнением Ито**.

Справедливо следующее утверждение, которое мы примем без доказательства.

Теорема 9.1. *Пусть* $a(x,t)$, $\sigma(x,t)$ *– непрерывные по* t *функции и существует константа* L, *такая, что*

$$|a(x,t) - a(y,t)| + |\sigma(x,t)| \le L|x-y|,$$

$$a^2(x,t) + \sigma^2(x,t) \le L^2(1+x^2).$$

Тогда существует «единственное» решение уравнения (9.7) *в том смысле, что*

$$P\left(\sup_{t_0 \le t \le T} |X_1(t) - X_2(t)| = 0\right) = 1,$$

где $X_1(t)$, $X_2(t)$ *– некоторые решения уравнения* (9.7).

Пусть $a(x,t) = \beta(t)x + \alpha(t)$, $\sigma(x,t) = b(t)$. Тогда решение уравнения (9.7), как легко проверить, имеет вид

$$X(t)=e^{\int\limits_{t_0}^{t}\beta(\tau)d\tau}\left[X(t_0)+\int\limits_{t_0}^{t}\alpha(\tau)e^{-\int\limits_{t_0}^{\tau}\beta(u)du}d\tau+\int\limits_{t_0}^{t}b(\tau)e^{-\int\limits_{t_0}^{\tau}\beta(u)du}d\xi(\tau)\right]. \quad (9.8)$$

Отсюда следует, что среднее значение данного решения можно записать в виде

$$m_X(t)=MX(t)=e^{\int\limits_{t_0}^{t}\beta(\tau)d\tau}\left[MX(t_0)+\int\limits_{t_0}^{t}\alpha(\tau)e^{-\int\limits_{t_0}^{\tau}\beta(u)du}d\tau\right].$$

Покажем, что процесс $X(t)$, определяемый уравнением (9.7), является марковским. Действительно, распределение вероятностей $X(t)$ при $t>t_0$ при заданном значении $X(t_0)$ зависит лишь от $X(t_0)$ и не зависит от прошлых значений $X(s)$, $s<t_0$, как видно из (9.7). А это является основным свойством марковского процесса. Поэтому процесс $X(t)$ полностью описывается коэффициентами сноса и диффузии. Пусть $\xi(t)$ – винеровский процесс с нулевым средним и единичным коэффициентом диффузии, второй интеграл в (9.7) будем понимать в форме Ито. Тогда

$$X(t+\Delta t)-X(t)=a(X(t),t)\Delta t+\sigma(X(t),t)[\xi(t+\Delta t)-\xi(t)]+o(\Delta t),$$

откуда после усреднения имеем

$$M[X(t+\Delta t)-X(t)/X(t)]=a(X(t),t)\Delta t,$$

т.е. $a(x,t)$ является коэффициентом сноса процесса $X(t)$. Найдем коэффициент диффузии. Далее, используя (9.8), получаем

$$[X(t+\Delta t)-X(t)]^2=a^2(X(t),t)(\Delta t)^2+\sigma^2(X(t),t)[\xi(t+\Delta t)-\xi(t)]^2+$$
$$+2a(X(t),t)\sigma(X(t),t)\Delta t[\xi(t+\Delta t)-\xi(t)]+o(\Delta t).$$

Отсюда, после усреднения будем иметь

$$M\left\{[X(t+\Delta t)-X(t)]^2/X(t)\right\}=\sigma^2(X(t),t)\Delta t,$$

таким образом, коэффициент диффузии процесса $X(t)$ равен $\sigma^2(x,t)$.

Знание коэффициентов сноса и диффузии позволяет записать уравнение Колмогорова–Фоккера–Планка для переходной плотности вероятностей, которая однозначно определяет поведение процесса $X(t)$.

ГЛАВА 10. СТАЦИОНАРНЫЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ

§46. СТАЦИОНАРНЫЕ В УЗКОМ И ШИРОКОМ СМЫСЛЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ

Определение. Случайный процесс $\xi(t)$ называется **стационарным в узком смысле**, если для конечномерных ф.р. выполняется равенство

$$F(x_1,t_1;x_2,t_2;\ldots;x_n,t_n)=F(x_1,t_1+\tau;x_2,t_2+\tau;\ldots;x_n,t_n+\tau) \quad (10.1)$$

$\forall\,\tau$ и $\forall\, n=1,2,\ldots$, т.е. его конечномерное распределение инвариантно относительно сдвига всех моментов времени t_i, $i=\overline{1,n}$, на одну и ту же величину τ.

При $n=1$ условие стационарности (10.1) дает

$$F(x_1,t_1)=F(x_1,t_1+\tau).$$

Полагая $\tau=-t_1$, имеем

$$F(x_1,t_1)=F(x_1,0),$$

т.е. одномерное распределение стационарного процесса не зависит от времени и совпадает с распределением в момент времени $t=0$. А одномерное распределение, согласно (2.12), (2.13), определяет среднее значение и дисперсию случайного процесса, следовательно, для стационарного в узком смысле случайного процесса среднее значение и дисперсия не зависят от времени:

$$m_\xi(t)=m,\ D_\xi(t)=\sigma^2.$$

При $n=2$ из условия стационарности следует

$$F(x_1,t_1;x_2,t_2)=F(x_1,t_1+\tau;x_2,t_2+\tau).$$

Полагая $\tau=-t_1$, имеем

$$F(x_1,t_1;x_2,t_2)=F(x_1,x_2,t_2-t_1),$$

т.е. двумерное распределение зависит лишь от разности моментов времени t_2 и t_1. Поэтому корреляционная функция стационарного в узком смысле случайного процесса, согласно (2.14), зависит только от одного аргумента $t = t_2 - t_1$:

$$R_\xi(t_1, t_2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_1 x_2 dF(x_1, t_1; x_2, t_2) =$$
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_1 x_2 dF(x_1, t_1, t_2 - t_1) = R_\xi(t_2 - t_1) = R_\xi(t). \tag{10.2}$$

С практической точки зрения стационарность в узком смысле означает, что все статистические характеристики процесса не подвергаются изменению при переносе оси времени, в частности, одномерное распределение не зависит от времени вообще. Например, стационарность в узком смысле некоторого производственного процесса означает, что не подвергаются систематическим изменениям ни условия производства, ни поставляемое сырье, ни другие факторы; существует возможность появления только некоторых случайных изменений и то только таких, распределения вероятностей которых не подвергаются изменениям при переносе времени. Стационарным в узком смысле случайным процессом не может быть измеряемая температура воздуха в течение года, поскольку известно, что распределения температуры в разных местах и в разные поры года являются различными. Аналогично, стационарным процессом не является курс доллара, выраженный в рублях.

Определение. Случайный процесс $\xi(t)$ называется **стационарным в широком смысле**, если его среднее значение не зависит от времени

$$m_\xi(t) = m, \tag{10.3}$$

а корреляционная функция зависит лишь от разности аргументов,

$$R_\xi(t_1, t_2) = R_\xi(t_2 - t_1), \tag{10.4}$$

Очевидно, что из стационарности в узком смысле следует стационарность в широком смысле. Наоборот верно не всегда. Но для гауссовских случайных процессов верно и обратное утверждение.

Теорема 10.1. *Гауссовский процесс стационарный в широком смысле является стационарным и в узком смысле.*

Доказательство. Согласно §19 n-мерная ф.р. процесса равна

$$F(x_1,t_1;x_2,t_2;\dots;x_n,t_n)=$$
$$=\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|K|}}\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\dots\int_{-\infty}^{x_n}\exp\left\{-\frac{1}{2|K|}\sum_{i,j=1}^{n}\left[(y_i-a_i)K_{ij}(y_j-a_j)\right]\right\}dy_1dy_2\dots dy_n, \tag{10.5}$$

где $a_i=M\xi(t_i)$, $K=\left\|K(t_i,t_j)\right\|=\left\|M\left[\left(\xi(t_i)-a_i\right)\left(\xi(t_j)-a_j\right)\right]\right\|_{n\times n}$,

$i,j=\overline{1,n}$, K_{ij} – алгебраическое дополнение элемента $K(t_i,t_j)$ матрицы ковариаций K. Для стационарного в широком смысле процесса, как следует из (2.14), (10.3), (10.4),

$$K(t_i,t_j)=K(t_i+\tau,t_j+\tau)=K(t_j-t_i),$$

поэтому многомерная ф.р. (10.5) зависит только от n и $K(t_j-t_i)$:

$$F(x_1,t_1;x_2,t_2;\dots;x_n,t_n)=f\left(x_i,n,K(t_j-t_i)\right),\ i,j=\overline{1,n}.$$

Отсюда видно, что при переходе к моментам времени $t_i+\tau$, $i=\overline{1,n}$, ф.р. не изменится, т.е. соотношение (10.1) выполнено.

Как указано выше, функцию корреляции стационарного в широком смысле случайного процесса можно рассматривать как функцию одного аргумента $R_\xi(t)$. Как следует из свойств этой функции для произвольного случайного процесса (см. §12), $R_\xi(t)$ обладает следующими свойствами:

а) $R_\xi(0)\geq 0$, (10.6)

б) $\left|R_\xi(t)\right|\leq R_\xi(0)$, (10.7) в) $R_\xi(t)$ является положительно определенной функцией, т.е. $\forall\ n$ и любых комплексных чисел $z_1, z_2, ..., z_n$

$$\sum_{i,j=1}^{n} R_\xi(t_i - t_j) z_i \bar{z}_j \geq 0 \cdot \tag{10.8}$$

Отметим, что частными случаями теорем 3.2, 3.4 являются следующие утверждения.

Теорема 10.2. *Для того чтобы стационарный в широком смысле случайный процесс $\xi(t)$ был непрерывен в среднем квадратичном, необходимо и достаточно, чтобы его корреляционная функция была непрерывной в точке $t = 0$.*

Теорема 10.3. *Для того чтобы стационарный в широком смысле процесс был дифференцируем в среднем квадратичном, необходимо и достаточно, чтобы $-R''_\xi(0) < +\infty$.*

Пример 10.1. Рассмотрим случайный процесс

$$\xi(t) = \xi \cos t + \eta \sin t,$$

где ξ, η – независимые СВ, причем $M\xi = M\eta = 0$, $D\xi = D\eta = 1$. Очевидно, что $m_\xi(t) = 0$. Поскольку $M\xi^2 = D\xi = 1$, $M\eta^2 = D\eta = 1$, $M(\xi\eta) = M\xi M\eta = 0$, то корреляционная функция процесса равна

$$R_\xi(t_1, t_2) = M[\xi(t_1)\xi(t_2)] = M[(\xi \cos t_1 + \eta \sin t_1)(\xi \cos t_2 + \eta \sin t_2)] =$$

$$= M(\xi^2 \cos t_1 \cos t_2 + \xi\eta \sin t_1 \cos t_2 + \xi\eta \cos t_1 \sin t_2 + \eta^2 \sin t_1 \sin t_2) =$$

$$= \cos t_1 \cos t_2 + \sin t_1 \sin t_2 = \cos(t_1 - t_2) = \cos(t_2 - t_1),$$

поэтому случайный процесс $\xi(t)$ является стационарным в широком смысле.

Пусть теперь у нас имеются два случайных процесса $\xi(t)$ и $\eta(t)$.

Определение. Процессы $\xi(t)$ и $\eta(t)$ называются **стационарно связными в узком смысле**, если их совместная конечномерная ф.р. любого порядка (2.11) не зависит от положения начала отсчета времени, т.е.

$$F(x_1,t_1;x_2,t_2;...;x_n\,,t_n;y_1,\tau_1;y_2,\tau_2;...;y_m\,,\tau_m)=$$
$$=F(x_1,t_1+\tau;x_2,t_2+\tau;...;x_n\,,t_n+\tau;y_1,\tau_1+\tau;y_2,\tau_2+\tau;...;y_m\,,\tau_m+\tau)$$
$$\forall\ \tau\ \text{и}\ \forall\ n,m=1,2,...\ .$$

Определение. Комплексный случайный процесс

$$\zeta(t)=\xi(t)+i\eta(t)$$

называется стационарным в узком смысле, если $\xi(t)$ и $\eta(t)$ являются стационарно связными в узком смысле, т.е. его конечномерное распределение любого порядка n удовлетворяет соотношению

$$F_\zeta(x_1,t_1;x_2,t_2;...;x_n\,,t_n;y_1,t_1;y_2,t_2;...;y_n\,,t_n)=$$
$$=F_\zeta(x_1,t_1+\tau;x_2,t_2+\tau;...;x_n\,,t_n+\tau;y_1,t_1+\tau;y_2,t_2+\tau;...;y_n\,,t_n+\tau)\cdot$$

Определение. Пусть $\zeta(t)$ – комплексный случайный процесс, такой, что $M\zeta^2(t)<\infty$. Он называется стационарным в широком смысле, если

$$M\zeta(t)=const\ (\text{не зависит от } t)\,, \tag{10.9}$$

$$R_\zeta(t_1,t_2)=M\left[\zeta(t_1)\overline{\zeta(t_2)}\right]=R_\zeta(t_2-t_1)\cdot \tag{10.10}$$

Пример 10.2. Рассмотрим случайный процесс

$$\zeta(t)=\sum_{k=-n}^{n}\alpha_k(\omega)e^{i\lambda_k t}\ ,$$

где $\lambda_{-n},\lambda_{-n+1},...,\lambda_n$ – действительные постоянные,

$\alpha_{-n}(\omega),\alpha_{-n+1}(\omega),...,\alpha_n(\omega)$ – СВ, такие, что

$M\alpha_k(\omega)=0$, $M\left|\alpha_k(\omega)\right|^2=\sigma_k{}^2<\infty$, $k=-n,-n+1,...,n$. Ясно, что

$$M\zeta(t)=0\cdot$$

Найдем корреляционную функцию процесса $\zeta(t)$:

$$R_\zeta(t_1,t_2)=M\left[\zeta(t_1)\overline{\zeta(t_2)}\right]=M\left(\sum_{k=-n}^{n}\alpha_k(\omega)e^{i\lambda_k t_1}\sum_{k=-n}^{n}\alpha_k(\omega)e^{-i\lambda_k t_2}\right)=$$

$$=\sum_{k=-n}^{n}M\left[\alpha_k{}^2(\omega)\right]e^{i\lambda_k(t_1-t_2)}=\sum_{k=-n}^{n}\sigma_k^2 e^{-i\lambda_k(t_2-t_1)}.$$

Таким образом, выполнены условия (10.9), (10.10), поэтому процесс $\zeta(t)$ является стационарным в широком смысле.

§47. СПЕКТРАЛЬНАЯ ПЛОТНОСТЬ СЛУЧАЙНОГО ПРОЦЕССА

Докажем сначала в общем виде одно утверждение.

Теорема 10.4 ***(Бохнера).*** *Для функции* $R(t)$*, обладающей свойствами (10.6) – (10.8), существует функция* $F(\lambda)$ *– неубывающая, ограниченной вариации, такая, что*

$$R(t)=\int_{-\infty}^{+\infty}e^{i\lambda t}dF(\lambda)\cdot \tag{10.11}$$

Доказательство. Пусть $z_i=e^{-i\lambda t_i}$, $i=1,2,\ldots$. Тогда из определения положительно определенной функции (10.6) следует, что

$$\sum_{i,j}R(t_i-t_j)e^{-i\lambda(t_i-t_j)}\geq 0,$$

поэтому

$$g(\lambda,A)=\frac{1}{\sqrt{2\pi}A}\int_0^A\int_0^A R(t-u)e^{-i\lambda(t-u)}dtdu\geq 0\ \ \forall\ A>0. \tag{10.12}$$

Сделаем в интеграле (10.12) замену переменных

$$t-u=x,$$

$$u=y,$$

тогда $x=t-y$. При этом область интегрирования B перейдет в область интегрирования C, как показано на рис. 37.

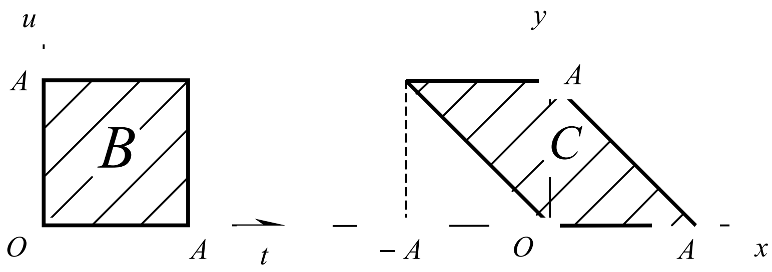


Рис. 37

Далее будем иметь

$$g(\lambda, A) = \frac{1}{\sqrt{2\pi}A}\left[\int_{-A}^{0}\int_{-x}^{A} dy R(x) e^{-i\lambda x} dx + \int_{0}^{A}\int_{0}^{A-x} dy R(x) e^{-i\lambda x} dx\right] =$$

$$= \frac{1}{\sqrt{2\pi}A}\left[\int_{-A}^{0}(A+x)R(x)e^{-i\lambda x}dx + \int_{0}^{A}(A-x)R(x)e^{-i\lambda x}dx\right] = \qquad (10.13)$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-A}^{A}\left(1-\frac{|x|}{A}\right)R(x)e^{-i\lambda x}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mu\left(\frac{x}{A}\right)R(x)e^{-i\lambda x}dx,$$

где

$$\mu(t) = \begin{cases} 1-|t|, \ |t| \le 1, \\ 0, \ |t| > 1. \end{cases}$$

Покажем, что функция $g(\lambda, A)$ интегрируема на всей прямой. Рассмотрим интеграл

$$\int_{-\infty}^{\infty}\mu\left(\frac{\lambda}{2M}\right)g(\lambda, A)d\lambda = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mu\left(\frac{x}{A}\right)R(x)\int_{-\infty}^{\infty}\mu\left(\frac{\lambda}{2M}\right)e^{-i\lambda x}d\lambda dx.$$

Поскольку

$$\int_{-\infty}^{\infty} \mu\left(\frac{\lambda}{2M}\right) e^{-i\lambda x} d\lambda = \int_{-2M}^{2M} \left(1 - \frac{|\lambda|}{2M}\right) e^{-i\lambda x} d\lambda =$$

$$= 2\int_{0}^{2M} \left(1 - \frac{\lambda}{2M}\right) \cos(\lambda x) d\lambda = 2\int_{0}^{2M} \left(1 - \frac{\lambda}{2M}\right) \frac{d\sin(\lambda x)}{x} =$$

$$= 2\left(1 - \frac{\lambda}{2M}\right) \frac{\sin(\lambda x)}{x}\Bigg|_{0}^{2M} + \frac{1}{Mx}\int_{0}^{2M} \sin(\lambda x) d\lambda =$$

$$= -\frac{\cos(\lambda x)}{Mx^2}\Bigg|_{0}^{2M} = \frac{1 - \cos(2Mx)}{Mx^2} = \frac{2\sin^2(Mx)}{Mx^2},$$

то с учетом (10.7) получаем

$$\left| \int_{-\infty}^{\infty} \mu\left(\frac{\lambda}{2M}\right) g(\lambda, A) d\lambda \right| = \left| M\sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} \mu\left(\frac{x}{A}\right) R(x) \frac{\sin^2(Mx)}{(Mx)^2} dx \right| \le$$

$$\le R(0)\sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} \frac{\sin^2(Mx)}{(Mx)^2} d(Mx) = R(0).$$

Таким образом, $g(\lambda, A)$ интегрируема на всей прямой и, как следует из (10.13), является обратным преобразованием Фурье для функции $\mu\left(\frac{x}{A}\right)R(x)$. Тогда из свойств преобразований Фурье получаем, что в этом случае существует прямое преобразование Фурье

$$\mu\left(\frac{x}{A}\right) R(x) = \int_{-\infty}^{+\infty} g(\lambda, A) e^{i\lambda x} d\lambda \cdot \tag{10.14}$$

При $x = 0$ имеем

$$R(0) = \int_{-\infty}^{+\infty} g(\lambda, A) d\lambda,$$

откуда, учитывая условие нормировки, вытекает, что функция $\dfrac{g(\lambda, A)}{R(0)}$ является плотностью распределения вероятностей некоторой СВ. В свою очередь тогда из (10.14) следует, что $\mu\left(\dfrac{x}{A}\right)\dfrac{R(x)}{R(0)}$ является характеристической функцией $\forall A$.

При $A \to \infty$

$$\mu\left(\frac{x}{A}\right)\frac{R(x)}{R(0)} \longrightarrow \frac{R(x)}{R(0)},$$

поэтому $\dfrac{R(x)}{R(0)}$ также является характеристической функцией. Тогда из теоремы об обращении характеристической функции вытекает, что существует неубывающая функция $F_1(\lambda)$ ограниченной вариации, такая, что

$$\frac{R(x)}{R(0)} = \int_{-\infty}^{+\infty} e^{i\lambda x} dF_1(\lambda),$$

откуда следует выражение (10.11). Теорема доказана.

Следствие. Если $R(t)$ – корреляционная функция стационарного случайного процесса $\xi(t)$, то справедливо выражение (10.11). При этом $F(\lambda)$ называется **спектральной функцией** процесса $\xi(t)$. Если функция $F(\lambda)$ дифференцируема, то $f(\lambda) = F'(\lambda)$ называется **спектральной плотностью** процесса $\xi(t)$,

$$f(\lambda) = \int_{-\infty}^{+\infty} e^{-i\lambda t} R(t) dt,$$

$$F(\lambda)=\int_{-\infty}^{\lambda} f(u)du \cdot$$

Если $R(t)$ – вещественная функция, то

$$R(t)=\int_{-\infty}^{+\infty}\cos(\lambda t)f(\lambda)d\lambda,$$

$$f(\lambda)=\int_{-\infty}^{+\infty}\cos(\lambda t)R(t)dt \cdot$$

Пример 10.3. Найдем спектральную плотность с корреляционной функцией

$$R(t)=\frac{1}{2\pi}\sigma^2 e^{-\alpha|t|},\ \alpha>0.$$

Используя формулу (10.15), получаем

$$f(\lambda)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-i\lambda t}\sigma^2 e^{-\alpha|t|}dt=$$

$$=\frac{\sigma^2}{2\pi}\int_{-\infty}^{0}e^{-i\lambda t+\alpha t}dt+\frac{\sigma^2}{2\pi}\int_{0}^{+\infty}e^{-i\lambda t-\alpha t}dt=$$

$$=\frac{\sigma^2}{2\pi}\left(\frac{1}{\alpha-i\lambda}+\frac{1}{\alpha+i\lambda}\right)=\frac{\sigma^2}{\pi}\cdot\frac{\alpha}{\alpha^2+\lambda^2}.$$

Сформулируем без доказательства нужное нам в дальнейшем утверждение.

Теорема 10.5. *Для любого стационарного в широком смысле и непрерывного в среднем квадратичном случайного процесса* $\xi(t)$ *существует процесс* $\eta(t)$ *с независимыми приращениями, такой, что справедливо спектральное представление*

$$\xi(t)=\int_{-\infty}^{+\infty}e^{i\lambda t}d_\lambda\eta(\lambda)\cdot \tag{10.16}$$

Рассмотрим случайный процесс вида

$$\zeta(t)=a_0\xi(t)+a_1\frac{d\xi(t)}{dt}+\ldots+a_k\frac{d^k\xi(t)}{dt^k}, \qquad (10.17)$$

где $a_0, a_1, ..., a_k$ – постоянные коэффициенты, $\xi(t)$ – стационарный в широком смысле, имеющий спектральную плотность $f_\xi(\lambda)$ процесс, имеющий также производную k-го порядка в среднем квадратичном. Согласно теореме 10.5 процесс $\xi(t)$ имеет спектральное представление (10.16). Кроме того, можно показать, что

$$\frac{d^r\xi(t)}{dt^r}=\frac{d^r}{dt^r}\int_{-\infty}^{+\infty}e^{i\lambda t}d_\lambda\eta(\lambda)=\int_{-\infty}^{+\infty}(i\lambda)^r e^{i\lambda t}d_\lambda\eta(\lambda),$$

т.е. можно дифференцировать под знаком интеграла. Учитывая это, процесс (10.17) можно переписать в виде

$$\zeta(t)=\int_{-\infty}^{+\infty}\left[a_0+a_1(i\lambda)+\ldots+a_k(i\lambda)^k\right]e^{i\lambda t}d_\lambda\eta(\lambda).$$

Можно также показать, что корреляционная функция этого процесса равна

$$R_\zeta(t_1,t_2)=M\left[\zeta(t_1)\overline{\zeta(t_2)}\right]=$$

$$=M\left(\int_{-\infty}^{+\infty}\left[a_0+a_1(i\lambda)+\ldots+a_k(i\lambda)^k\right]e^{i\lambda t_1}d_\lambda\eta(\lambda)\times\right.$$

$$\left.\times\overline{\int_{-\infty}^{+\infty}\left[a_0+a_1(i\lambda)+\ldots+a_k(i\lambda)^k\right]e^{i\lambda t_2}d_\lambda\eta(\lambda)}\right)=$$

$$=\int_{-\infty}^{+\infty}\left|a_0+a_1(i\lambda)+\ldots+a_k(i\lambda)^k\right|^2 e^{i\lambda(t_1-t_2)}f_\xi(\lambda)d\lambda.$$

Тогда спектральная плотность имеет вид

$$f_\zeta(\lambda)=\left|a_0+a_1(i\lambda)+\ldots+a_k(i\lambda)^k\right|^2 f_\xi(\lambda). \qquad (10.18)$$

Рассмотрим теперь выражение (10.17) как стохастическое дифференциальное уравнение. Если предположить, что его решение

$\xi(t)$ является стационарным в широком смысле случайным процессом, то спектральная плотность этого решения равна

$$f_\xi(\lambda) = \frac{f_\zeta(\lambda)}{\left|a_0 + a_1(i\lambda) + \ldots + a_k(i\lambda)^k\right|^2}.$$

Пример 10.4. Найдем среднее значение $m_\xi(t)$ и функцию корреляции $R_\xi(t)$ стационарного в широком смысле решения стохастического дифференциального уравнения.

$$\xi'(t) + \alpha\xi(t) = \zeta(t), \tag{10.19}$$

где $\zeta(t)$ является стационарным в широком смысле случайным процессом с нулевым средним значением, $M\zeta(t) = 0$, и спектральной плотностью

$$f_\zeta(\lambda) = e^{-\frac{1}{2}\lambda^2}, \ \alpha - \text{действительная постоянная.}$$

Предположим, что решение уравнения (10.19) является стационарным процессом, поэтому $m_\xi(t) = const$. Кроме того, беря математическое ожидание от обеих сторон в (10.19), получаем

$$m_\xi'(t) + \alpha m_\xi(t) = 0,$$

откуда следует, что $m_\xi(t) = 0$.

Из (10.18) находим выражение, которому удовлетворяет спектральная плотность процесса $\xi(t)$,

$$|\alpha + i\lambda|^2 f_\xi(\lambda) = (\alpha^2 + \lambda^2) f_\xi(\lambda) = f_\zeta(\lambda) = e^{-\frac{1}{2}\lambda^2},$$

откуда

$$f_\xi(\lambda) = \frac{1}{\alpha^2 + \lambda^2} e^{-\frac{1}{2}\lambda^2}.$$

Поэтому из (10.11) получаем

$$R_\xi(t) = \int_{-\infty}^{+\infty} e^{i\lambda t} f_\xi(\lambda) d\lambda = \int_{-\infty}^{+\infty} e^{i\lambda t} \frac{1}{\alpha^2 + \lambda^2} e^{-\frac{1}{2}\lambda^2} d\lambda. \tag{10.20}$$

Далее заметим, что функция $f_1(\lambda)=e^{-\frac{1}{2}\lambda^2}$ является характеристической функцией СВ ξ_1, имеющей плотность распределения

$$p_{\xi_1}(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},$$

а функция $f_2(\lambda)=\dfrac{\alpha^2}{\alpha^2+\lambda^2}$ является характеристической функцией СВ ξ_2, имеющей плотность распределения

$$p_{\xi_2}(x)=\frac{\alpha}{2}e^{-\alpha|x|}.$$

Поэтому функцию

$$f_{\xi_1+\xi_2}(\lambda)=f_1(\lambda)f_2(\lambda)=\frac{\alpha^2}{\alpha^2+\lambda^2}e^{-\frac{1}{2}\lambda^2} \qquad (10.21)$$

можно трактовать как характеристическую функцию суммы двух независимых СВ $\xi_1+\xi_2$. Плотность распределение этой суммы имеет вид

$$p_{\xi_1+\xi_2}(x)=\int_{-\infty}^{+\infty}p_{\xi_1}(u)p_{\xi_2}(x-u)du=$$

$$=\frac{\alpha}{2\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{u^2}{2}-\alpha|x-u|}du\cdot \qquad (10.22)$$

Используя (10.21), получаем

$$\frac{1}{\alpha^2}\int_{-\infty}^{+\infty}e^{-i\lambda t}\frac{\alpha^2}{\alpha^2+\lambda^2}e^{-\frac{1}{2}\lambda^2}d\lambda=\frac{1}{2\pi\alpha^2}\int_{-\infty}^{+\infty}e^{-i\lambda t}2\pi f_{\xi_1+\xi_2}(\lambda)d\lambda\cdot \qquad (10.23)$$

Как известно (§7), для характеристической функции справедливы соотношения

$$f_{\xi_1+\xi_2}(\lambda) = \int_{-\infty}^{+\infty} e^{i\lambda t} p_{\xi_1+\xi_2}(t)dt,$$

$$p_{\xi_1+\xi_2}(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-i\lambda t} f_{\xi_1+\xi_2}(\lambda)d\lambda. \qquad (10.24)$$

Тогда из (10.20), (10.22) – (10.24) вытекает, что

$$R_\xi(t) = \frac{2\pi}{\alpha^2} p_{\xi_1+\xi_2}(t) = \frac{\sqrt{2\pi}}{2\alpha}\int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}-\alpha|t-u|}du.$$

§48. ЭРГОДИЧЕСКОЕ СВОЙСТВО СЛУЧАЙНЫХ ПРОЦЕССОВ

Пусть $\xi(t),\ t \geq 0$ – случайный процесс с непрерывным временем.

Определение. Временным средним некоторой функции $g\left(\xi(t,\overline{\omega})\right)$ реализации случайного процесса $\xi(t,\overline{\omega})$ называется величина

$$\widehat{M}g\left(\xi(t,\overline{\omega})\right) = \lim_{t\to\infty}\frac{1}{t}\int_0^t g\left(\xi(\tau,\overline{\omega})\right)d\tau. \qquad (10.25)$$

Например, временное среднее случайного процесса равно

$$\widehat{m} = \widehat{M}\xi(t,\overline{\omega}) = \lim_{t\to\infty}\frac{1}{t}\int_0^t \xi(\tau,\overline{\omega})d\tau,$$

а временная дисперсия

$$\widehat{D} = \widehat{M}\left[\xi(t,\overline{\omega}) - \widehat{m}\right]^2 = \lim_{t\to\infty}\frac{1}{t}\int_0^t \left[\xi(\tau,\overline{\omega}) - \widehat{m}\right]^2 d\tau.$$

Определение. Говорят, что случайный процесс $\xi(t,\omega)$ обладает **эргодическим свойством**, если

$$Mg\left(\xi(t,\omega)\right) = \int_\Omega g\left(\xi(t,\omega)\right)P(d\omega) =$$

$$= \widehat{M} g\left(\xi(t,\overline{\omega})\right) = \lim_{t\to\infty} \frac{1}{t}\int_0^t g\left(\xi(\tau,\overline{\omega})\right) d\tau \cdot \qquad (10.26)$$

Свойство эргодичности обычно рассматривается применительно к стационарным процессам, для которых $Mg\left(\xi(t,\omega)\right)$ не зависит от t. В этом случае при выполнении (10.26) вместо среднего по вероятностной мере на множестве Ω

$$\int_\Omega g\left(\xi(t,\omega)\right) P(d\omega)$$

можно взять одну реализацию процесса $\xi(t,\overline{\omega})$ и вычислить среднее по времени (10.25).

Сформулируем сейчас утверждения, являющиеся теоретической основой для оценки среднего значения процесса с помощью его одной реализации.

Теорема 10.6. *Пусть* $\xi(t)$ *– случайный процесс с непрерывной корреляционной функцией* $R(t_1,t_2)$. *Тогда необходимым и достаточным условием для того, чтобы*

$$\underset{t\to\infty}{l.i.m.} \frac{1}{t}\int_0^t \xi(\tau) d\tau = 0 \qquad (10.27)$$

является

$$\lim_{t\to\infty} \frac{1}{t^2}\int_0^t\int_0^t R(t_1,t_2)\, dt_1\, dt_2 = 0 \cdot \qquad (10.28)$$

Доказательство. Справедливо следующее соотношение

$$M\left(\frac{1}{t}\int_0^t \xi(\tau)d\tau\right)^2 = \frac{1}{t^2} M\left(\int_0^t \xi(\tau)d\tau \int_0^t \xi(\tau)d\tau\right) =$$

$$= \frac{1}{t^2}\int_0^t\int_0^t M\left[\xi(t_1)\xi(t_2)\right] dt_1\, dt_2 = \frac{1}{t^2}\int_0^t\int_0^t R(t_1,t_2)\, dt_1\, dt_2 \cdot \qquad (10.29)$$

Из определения сходимости в среднем квадратичном следует, что (10.27) эквивалентно соотношению

$$\lim_{t\to\infty} M\left(\frac{1}{t}\int_0^t \xi(\tau)d\tau\right)^2 = 0\cdot$$

Поэтому из (10.29) следует, что выражения (10.27) и (10.28) эквивалентны.

Теорема 10.7. *Пусть* $\xi(t)$ *– действительный стационарный в широком смысле случайный процесс, имеющий непрерывную корреляционную функцию* $R(t)$. *Необходимым и достаточным условием выполнения равенства (10.27) служит*

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t\left(1-\frac{\tau}{t}\right)R(\tau)d\tau = 0\cdot \tag{10.30}$$

Доказательство. Нужно показать, что для стационарного в широком смысле случайного процесса условие (10.28) принимает вид (10.30). Заметим, что

$$\frac{1}{t^2}\int_0^t\int_0^t R(t_1,t_2)dt_1\,dt_2 = \frac{1}{t^2}\int_0^t\int_0^t R(t_2-t_1)dt_1\,dt_2 =$$

$$=\frac{1}{t^2}\int_{-t}^0 R(\tau)d\tau\int_{-\tau}^t dt_1 + \frac{1}{t^2}\int_0^t R(\tau)d\tau\int_0^{t-\tau}dt_1 =$$

$$=\frac{1}{t^2}\int_{-t}^0 (t+\tau)R(\tau)d\tau + \frac{1}{t^2}\int_0^t (t-\tau)R(\tau)d\tau =$$

$$=\frac{1}{t^2}\int_0^t (t-\tau)R(-\tau)d\tau + \frac{1}{t^2}\int_0^t (t-\tau)R(\tau)d\tau =$$

$$=\frac{2}{t^2}\int_0^t (t-\tau)R(\tau)d\tau = \frac{2}{t}\int_0^t\left(1-\frac{\tau}{t}\right)R(\tau)d\tau,$$

откуда следует эквивалентность выражений (10.28) и (10.30).

Заметим также, что достаточным условием для того, чтобы имело место равенство (10.30), является

$$\lim_{t\to\infty} R(t) = 0\cdot$$

Это следует из того, что

$$\frac{1}{t}\int_0^t (1-\frac{\tau}{t})R(\tau)d\tau \le \frac{1}{t}\int_0^t R(\tau)d\tau =$$

$$=\frac{1}{t}\int_0^{t_0} R(\tau)d\tau + \frac{1}{t}\int_{t_0}^t R(\tau)d\tau \to 0 \text{ при } t\to\infty$$

для достаточно больших t_0.

Пример 10.5. Покажем, что для винеровского процесса равенство (10.28) не имеет места.

Из соотношения (4.13) с учетом того, что $M\xi(t)=0$, получаем

$$\frac{1}{t^2}\int_0^t\int_0^t R(t_1,t_2)dt_1\,dt_2 = \frac{2\sigma^2}{t^2}\int_0^t dt_1\int_0^t t_1\,dt_2 =$$

$$=\frac{2\sigma^2}{t^2}\int_0^t t_1(t-t_1)dt_1 = \frac{2\sigma^2}{t^2}\left(t\frac{t^2}{2}-\frac{t^3}{3}\right) = \frac{\sigma t}{3}\xrightarrow[t\to\infty]{}\infty,$$

поэтому в данном случае (10.28) не выполняется.

Пример 10.6. Проверим, выполняется ли условие (10.27) для процесса (2.17).

Данный процесс, как следует из примера 2.6, является стационарным в широком смысле. Из (2.19) вытекает, что имеет место равенство (10.30), поскольку

$$\frac{\sigma^2}{t}\int_0^t\left(1-\frac{\tau}{t}\right)\cos(\lambda\tau)d\tau = \frac{\sigma^2}{\lambda^2 t^2}\left[1-\cos(\lambda\tau)\right]\xrightarrow[t\to\infty]{}0,$$

поэтому, на основании теоремы 10.7, условие (10.27) выполняется.

§49. МАРТИНГАЛЫ

Мартингалы являются классом случайных процессов с очень полезными свойствами, которые в последнее время все шире находят применение, например, в финансовой математике и финансовой экономике [9].

Исследование зависимости между сечениями случайного процесса можно осуществлять различными способами. При исследовании стационарных случайных процессов, например, основным показателем этой зависимости была корреляционная функция, и все выводы делались на основании свойств непрерывности, дифференцируемости и интегрируемости этой функции. В теории непрерывных марковских процессов основной характеристикой зависимости служит переходная функция, которая полностью определяет эволюцию этих процессов; в однородных цепях Маркова такой характеристикой зависимости является матрица вероятностей переходов. В данном параграфе остановимся на достаточно обширном классе случайных процессов, для которых изучение зависимости проводится методами, основанными на исследовании свойств условных математических ожиданий.

Пусть имеем два сечения случайного процесса $\xi(t_1)$ и $\xi(t_2)$, $t_2 > t_1$, и пусть существует одномерная $p_\xi(x,t) > 0$ и двумерная $p_\xi(x,t_1;y,t_2)$ плотности распределения этого процесса. Тогда условное математическое ожидание можно ввести следующим образом

$$M\{\xi(t_2)/\xi(t_1)=x\}=\frac{1}{p_\xi(x,t_1)}\int_{-\infty}^{\infty} y p_\xi(x,t_1;y,t_2)dy,$$

если этот интеграл существует. Очевидно, что $M\{\xi(t_2)/\xi(t_1)=x\}$ есть функция от x, то есть от значения процесса в момент $t_1 < t_2$.

Рассмотрим теперь процесс $\xi(t)$, заданный многомерной плотностью распределения $p(x_1,t_1;x_2,t_2;\ldots;x_n,t_n)$, и предположим, что существуют условные математические ожидания

$$M\{\xi(t_n)/\xi(t_1)=x_1,\xi(t_2)=x_2,\ldots,\xi(t_{n-1})=x_{n-1}\}.$$

Определение. Случайный процесс $\xi(t)$ называется **мартингалом**, если существует условное математическое ожидание

$$M\{\xi(t_i)/\xi(t_1)=x_1,\xi(t_2)=x_2,\ldots,\xi(t_{i-1})=x_{i-1}\}=x_{i-1} \quad \forall\ i,$$

которое не зависит от предшествующих значений процесса $\xi(t_k)$, $t_k < t_{i-1}$, $k = \overline{1, i-2}$.

Мартингалы можно рассматривать как модели безобидных игр. Если $\xi(t_{i-1})$ описывает состояние капитала игрока в момент времени t_{i-1}, тогда по определению мартингала средняя величина его капитала в момент t_i при условии, что в момент t_{i-1} он располагал капиталом x_{i-1}, равна x_{i-1} независимо от того, каков был его капитал до момента t_{i-1}.

Рассмотрим примеры мартингалов.

Пример 10.7. Винеровский процесс с нулевым коэффициентом сноса и единичным коэффициентом диффузии является мартингалом. Действительно, учитывая независимость приращений винеровского процесса, найдем условное математическое ожидание значения процесса в момент t_i, если в предыдущий момент времени $t_{i-1} < t_i$ значение процесса фиксировано $\xi(t_{i-1}) = x_{i-1}$:

$$M\{\xi(t_i)/x_{i-1}\} = M\{[(\xi(t_i) - x_{i-1}) + x_{i-1}]/x_{i-1}\} =$$
$$= M\{(\xi(t_i) - x_{i-1})/x_{i-1}\} + x_{i-1} = x_{i-1}.$$

Пример 10.8. Рассмотрим пуассоновский процесс $\xi(t)$ с началом в нуле, то есть $\xi(0) = 0$. Тогда центрированный процесс

$$\eta(t) = \xi(t) - M\xi(t) = \xi(t) - \lambda t,$$

как нетрудно показать, также является мартингалом.

Пример 10.9. Рассмотрим пуассоновский процесс $\xi(t)$ с началом в нуле, $\xi(0) = 0$, с нулевым средним значением, $M\xi(t) = 0$, и независимыми приращениями. Пусть также

$$M[\xi(t) - \xi(\tau)]^2 = F(t) - F(\tau),\ 0 \le \tau < t.$$

Тогда процесс $\xi^2(t) - F(t)$ является мартингалом. Для того чтобы это показать, найдем условное математическое ожидание

$$M\left\{\left[\xi^2(t) - F(t)\right] / \xi(\tau) = x\right\} = M\left\{\left[\xi(t) - \xi(\tau) + \xi(\tau)\right]^2 - F(t) / \xi(\tau) = x\right\} =$$

$$= \xi^2(\tau) + 2\xi(\tau) M\left\{\left[\xi(t) - \xi(\tau)\right] / \xi(\tau) = x\right\} + M\left\{\left[\xi(t) - \xi(\tau)\right]^2 / \xi(\tau) = x\right\} - F(t) =$$

$$= \xi^2(\tau) + M\left\{\left[\xi(t) - \xi(\tau)\right] / \xi(\tau) = x\right\} - F(t) = \xi^2(\tau) - F(\tau),$$

то есть определение мартингала выполняется.

Определение. Случайный процесс $\xi(t)$ называется **субмартингалом**, если

$$M\left\{\xi(t_i) / \xi(t_{i-1}) = x_{i-1}\right\} \geq x_{i-1},$$

и **супермартингалом**, если

$$M\left\{\xi(t_i) / \xi(t_{i-1}) = x_{i-1}\right\} \leq x_{i-1} \quad \forall\ i$$

Очевидно, что если $\xi(t)$ – субмартингал, то $(-\xi(t))$ – супермартингал.

Теорема 10.8. *Пусть $g(x)$ – выпуклая вниз функция, $\xi(t)$ – мартингал. Тогда, если существует математическое ожидание $Mg(\xi(t))$, то $\eta(t) = g(\xi(t))$ – субмартингал.*

Доказательство. Данное утверждение вытекает из неравенства Иенсена для математического ожидания СВ

$$M\{g(\xi)\} \geq g(M\xi),$$

если $g(x)$ – выпуклая вниз функция. В нашем случае

$$M\left\{g(\xi(t_i)) / \xi(t_{i-1}) = x_{i-1}\right\} \geq g\left(M\left\{\xi(t_i) / \xi(t_{i-1}) = x_{i-1}\right\}\right).$$

По определению мартингала,

$$M\left\{\xi(t_i) / \xi(t_{i-1}) = x_{i-1}\right\} = x_{i-1},$$

поэтому

$$M\left\{g(\xi(t_i)) / \xi(t_{i-1}) = x_{i-1}\right\} \geq g(x_{i-1}),$$

то есть

$$M\left\{\eta(t_i) / \eta(t_{i-1}) = g(x_{i-1})\right\} \geq g(x_{i-1})$$

или

$$M\{\eta(t_i)/\eta(t_{i-1}) = y_{i-1}\} \ge y_{i-1},$$

что и требовалось доказать.

Основные соотношения и свойства мартингалов исследуются для случайных моментов времени, связанных с моментами достижения ими некоторых заранее заданных значений. Их называют моментами остановки. Основные утверждения, касающиеся моментов остановки, приведены, например, в [13].

ГЛАВА 11. О ПРИМЕНЕНИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

В данной главе речь пойдет о применении цепи Маркова с непрерывным временем и конечным числом состояний при исследовании сети массового обслуживания, являющейся вероятностной моделью локальной вычислительной сети, а также о применении винеровских процессов в финансовой математике.

§50. ИССЛЕДОВАНИЕ МАРКОВСКОЙ СЕТИ МАССОВОГО ОБСЛУЖИВАНИЯ В СТАЦИОНАРНОМ И ПЕРЕХОДНОМ РЕЖИМЕ

В §34 был проведен анализ марковских систем массового обслуживания (СМО). Под сетью массового обслуживания (МО) принято понимать совокупность СМО, между которыми циркулируют заявки, переходящие из одной системы в другую. Сети МО часто применяются в качестве математических моделей различных информационно-компьютерных систем и сетей и других объектов [14, 15].

Рассмотрим замкнутую сеть МО, состоящую из однолинейных систем обслуживания $S_1, S_2, ..., S_n$ (рис. 38). Система S_i обслуживает в общем случае K_i заявок типа i, $i = \overline{1, n-1}$. После обслуживания в системе S_i заявка типа i поступает в систему S_n, обслуживается там и вновь поступает в систему S_i, $i = \overline{1, n-1}$. Таким образом, в сети постоянно обслуживаются $K_1 + K_2 + ... + K_{n-1} = K$ заявок. Систему обслуживания S_n будем называть центральной, а $S_1, S_2, ..., S_{n-1}$ – периферийными СМО.

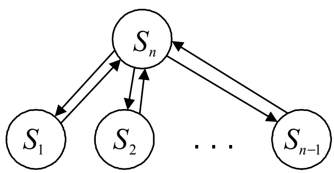


Рис. 38

Времена обслуживания заявок в системе S_i имеют экспоненциальное распределение с параметром μ_i, $i = \overline{1, n-1}$. Пусть $B(l, t)$

– ф.р. длительностей обслуживания заявок в центральной СМО и пусть $K_i = 1$, $i = \overline{1, n-1}$. Состояние сети в момент времени t характеризуется вектором

$$k(t) = (k, t) = (k_1, k_2, ..., k_{n-1}, n-1-k_1-k_2-...-k_{n-1}, t),$$

где $k_i = 0 \vee 1$ – число заявок в системе S_i, $i = \overline{1, n-1}$. Такая сеть служит моделью локальной сети со случайным множественным доступом, если под заявками понимать передаваемые в локальной сети пакеты; при этом каждая из $(n-1)$ первых СМО будет моделью сетевой станции, а центральная СМО – моделью канала связи.

Рассмотрим множество натуральных чисел $\Omega = \{1, 2, ..., n-1\}$ и для каждого набора $k = (k_1, k_2, ..., k_{n-1})$ – множества $\Omega_0(k) = \{i \mid k_i = 0,\ 1 \le i \le n-1\}$, $\Omega_1(k) = \{i \mid k_i = 1,\ 1 \le i \le n-1\}$, очевидно, что для любого k $\Omega_0(k) \cup \Omega_1(k) = \Omega$. Случайный процесс $\{k(t)\}$ является марковским с дискретным множеством состояний и непрерывным временем. Сети МО, состояния которых образуют марковский процесс, называются марковскими. Обозначим через p_i – вероятность перехода заявок в i-ю периферийную СМО после обслуживания в центральной СМО, $i = \overline{1, n-1}$, $\sum_{i=1}^{n-1} p_i = 1$. Возможны следующие переходы в состояние (k, t) за время Δt для этого процесса: из состояния $(k + I_i - I_n, t)$, $i \in \Omega_0(k)$, с вероятностью $\mu_i \Delta t + o(\Delta t)$; из состояния $(k - I_i + I_n, t)$, $i \in \Omega_1(k)$, с вероятностью $p_i \mu_n \left(n - |\Omega_1(k)|\right) \Delta t + o(\Delta t)$; из состояния (k, t) с вероятностью $1 - \sum_{i \in \Omega_1(k)} \mu_i \Delta t - \sum_{i \in \Omega_0(k)} p_i \mu_n \left(|\Omega_0(k)|\right) \Delta t + o(\Delta t)$; из остальных состояний – с вероятностью $o(\Delta t)$, где $|\Omega_0(k)|$ – мощность множества

$\Omega_0(k)$, I_i – вектор размерности n с нулевыми компонентами, за исключением i-й, которая равна 1. Тогда используя формулу полной вероятности и переходя к пределу при $\Delta t \to 0$, получим систему уравнений Колмогорова для вероятностей состояний:

$$\frac{dP(k,t)}{dt} = \left\{ \sum_{i\in\Omega_1(k)} \mu_i + \sum_{i\in\Omega_0(k)} p_i \mu_n \left(|\Omega_0(k)|\right) \right\} P(k,t) + \sum_{i\in\Omega_0(k)} \mu_i P(k + I_i - I_n, t) +$$

$$+ \sum_{i\in\Omega_1(k)} p_i \mu_n \left(|\Omega_0(k)|\right) P(k + I_i - I_n, t) \cdot \qquad (11.1)$$

Будем считать, что в центральной СМО S_n заявки на обслуживание выбираются с одинаковыми вероятностями, т.е. если в очереди на обслуживание находится l заявок, то интенсивность обслуживания заявок равна $\frac{1}{l}\mu_n(l)$. Кроме того, предположим, что функция $\mu_n(l)$ зависит от числа заявок линейно, т.е. $\mu_n(l) = \mu_n l$ и пусть $\mu_i = \mu$, $i = \overline{1, n-1}$. В этом случае система (11.1) имеет вид:

$$\frac{dP(k,t)}{dt} = -\left(\mu\Omega_1(k) + \mu_n |\Omega_0(k)|\right) P(k,t) +$$

$$+ \mu \sum_{i\in\Omega_0(k)} P(k + I_i - I_n, t) + \mu_n \sum_{i\in\Omega_1(k)} P(k - I_i + I_n, t) \cdot \qquad (11.2)$$

Исследуем более подробно эту систему.

Терема 11.1. *Стационарное решение системы (11.2) записывается в виде*

$$P(k) = G \prod_{i=1}^{n-1} \left(\frac{\mu_n}{\mu} \right)^{k_i}.$$

Доказательство. Стационарные вероятности состояний, согласно (11.2), удовлетворяют системе уравнений

$$\left(\mu|\Omega_1(k)| + \mu_n |\Omega_0(k)|\right) P(k) =$$

$$= \mu \sum_{i \in \Omega_0(k)} P(k + I_i - I_n, t) + \mu_n \sum_{i \in \Omega_1(k)} P(k - I_i + I_n, t) \cdot \quad (11.4)$$

Используя (11.3), получаем

$$P(k + I_i - I_n, t) = \frac{\mu_n}{\mu} P(k), \quad P(k - I_i + I_n, t) = \frac{\mu}{\mu_n} P(k), \quad k_i \neq 0,$$

таким образом,

$$\mu \sum_{i \in \Omega_0(k)} P(k + I_i - I_n, t) = \mu_n \left|\Omega_0(k)\right| P(k),$$

$$\mu_n \sum_{i \in \Omega_1(k)} P(k - I_i + I_n, t) = \mu \left|\Omega_1(k)\right| P(k) \cdot$$

То есть (11.4) выполняется.

В системе уравнений (11.2) C_{n-1}^{j} одинаковых уравнений для состояний с j нулями, $j = \overline{0, n-1}$, т.к. $P(k,t) = P(\overline{k},t)$, если число нулей в состояниях k и $\overline{k}$ совпадают, поскольку в этом случае $\left|\Omega_0(k)\right| = \left|\Omega_0(\overline{k})\right|$, $\left|\Omega_1(k)\right| = \left|\Omega_1(\overline{k})\right|$ и

$$\sum_{i \in \Omega_0(k)} P(k + I_i - I_n, t) = \sum_{i \in \Omega_0(\overline{k})} P(\overline{k} + I_i - I_n, t),$$

$$\sum_{i \in \Omega_1(k)} P(k - I_i + I_n, t) = \sum_{i \in \Omega_1(\overline{k})} P(\overline{k} - I_i + I_n, t) \cdot$$

Поэтому ее можно свести к системе $2^{n-1} - \sum_{j=0}^{n-1} \left(C_{n-1}^{j} - 1\right) = n$ различных линейных дифференциальных уравнений с постоянными коэффициентами, которую можно записать в матричной форме

$$\frac{dP(u,t)}{dt} = AP(u,t), u = \overline{1,n}, \quad (11.5)$$

где

$$A=\begin{pmatrix} -(n-1)\mu & (n-1)\mu_n & 0 & \dots & & 0 \\ \mu & -\left[(n-2)\mu+\mu_n\right] & (n-2)\mu_n & \dots & & 0 \\ 0 & 2\mu & -\left[(n-3)\mu+2\mu_n\right] & (n-3)\mu_n & 0\ \dots & 0 \\ & & \dots & & & \\ 0 & \dots & 0 & (n-2)\mu & -\left[\mu+(n-2)\mu_n\right] & 2\mu_n \\ 0 & \dots & 0 & 0 & (n-1) & \mu-(n-1)\mu_n \end{pmatrix}$$

Решение этой системы с матрицей Якоби может быть записано в виде [14]:

$$P(u,t)=\sum_{i=1}^{n}\sum_{\lambda}P(i,0)\frac{M_i(\lambda)M_u(\lambda)e^{\lambda t}}{(\mu\mu_n)^i\, i!(n-1)(n-2)\dots(n-i)L_n(\lambda)},u=\overline{1,n},$$

причем суммирование производится по всем корням λ (действительным и различным) многочлена $M_n(\lambda)$, где $M_1(\lambda)=1$, $M_2(\lambda)=\lambda+(n-1)\mu$,

$$M_{u+1}=[\lambda+n\mu-\mu_n+(\mu_n-\mu)u]M_u(\lambda)-(\mu\mu_u)^u u!(n-1)(n-2)\dots$$
$$\dots(n-u)M_{u-1}(\lambda),u=\overline{2,n-1},$$

$$L_n(\lambda)=\sum_{j=1}^{n}\frac{M_j^2}{(\mu\mu_n)^j\, j!(n-1)(n-2)\cdot\ldots\cdot(n-j)}.$$

Пример 11.1. Рассмотрим случай, когда $n=4$. Состояниями сети будут: $(1,1,1,0)$, $(1,1,0,1)$, $(1,0,1,1)$, $(0,1,1,1)$, $(1,0,0,2)$, $(0,1,0,2)$, $(0,0,1,2)$, $(0,0,0,3)$. Пусть $P(1,0)=1$, $P(i,0)=0$, $i=\overline{2,8}$. Для упрощения записей переобозначим их соответственно 1, 2, 3, …, 8. Можно показать [14], что в этом случае $P(2,t)=P(3,t)=P(4,t)$,

$P(5,t)=P(6,t)=P(7,t),$ и система уравнений для вероятностей состояний сводится к системе четырех дифференциальных уравнений для $P(1,t),$ $P(2,t),$ $P(5,t)$, $P(8,t)$ (11.5) с матрицей

$$A=\begin{pmatrix} -3\mu & 3\mu_4 & 0 & 0 \\ \mu & -(2\mu+\mu_4) & 2\mu_4 & 0 \\ 0 & 2\mu_4 & -(\mu+2\mu_4) & \mu_4 \\ 0 & 0 & 3\mu & -3\mu_4 \end{pmatrix},$$

которая в свою очередь, учитывая условие нормировки, сводится к линейному уравнению третьего порядка

$$\frac{d^3P(1,t)}{dt^3}+6\gamma\frac{d^2P(1,t)}{dt^2}+11\gamma^2\frac{dP(1,t)}{dt}+6\gamma^3P(1,t)=3\mu_4, \quad (11.6)$$

где $\gamma=\mu+\mu_4$. Частным решением его является $\dfrac{\mu_4}{2\gamma^3}$. Характеристическое уравнение однородного уравнения $y^3+6\gamma y^2+11\gamma^2 y+6\gamma^3=0$ имеет корни $y_1=-\gamma$, $y_2=-2\gamma$, $y_3=-3\gamma$. Таким образом, общее решение уравнения (11.6) имеет вид:

$$P(1,t)=c_1e^{-\gamma t}+c_2e^{-2\gamma t}+c_3e^{-3\gamma t}+\frac{\mu_4}{2\gamma^3}\cdot$$

Используя его, можно получить:

$$P(2,t)=\frac{c_1}{3\mu_4}(2\mu-\mu_4)e^{-\gamma t}+\frac{c_2}{3\mu_4}(\mu-2\mu_4)e^{-2\gamma t}+c_3e^{-3\gamma t}+\frac{\mu_4}{2\gamma},$$

$$P(8,t)=\frac{c_1\mu^2}{3\mu_4^2}e^{-\gamma t}+\frac{c_2\mu}{\mu_4}e^{-2\gamma t}+c_3e^{-3\gamma t}-\frac{3\gamma\mu+\mu_4^2}{2\gamma^2\mu_4}+1,$$

$$P(5,t)=\frac{1}{3}[1-P(1,t)-P(8,t)-3P(2,t)]\cdot$$

Константы c_1, c_2, c_3 находятся, используя начальные условия:

$$c_1 = \mu_4^2 \gamma^{-2} \left(1 - \gamma^{-2} - \frac{\gamma^2 + 3\mu^2}{2\mu_4 \gamma^3} \right),$$

$$c_2 = 3\mu_4 - 2\gamma^{-2} - 2c_1, \; c_3 = 1 - \frac{1}{2}\mu_4 \gamma^{-2} - c_1 - c_2 .$$

Знание этих вероятностей состояний позволяет найти такие характеристики локальной сети, как, например, среднее число станций, ожидающих передачи пакетов в момент времени t, $3P(2,t) + 6P(5,t) + 3P(8,t) - 1$, коэффициент загрузки канала связи $1 - P(1,t)$ и другие.

Отметим, что применение марковских сетей МО в качестве моделей обработки исков в страховых компаниях, а также использование уравнения Колмогорова–Фоккера–Планка при исследовании таких сетей рассмотрены в [15].

§51. ПРИМЕНЕНИЕ ВИНЕРОВСКОГО ПРОЦЕССА И СТОХАСТИЧЕСКИХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ В ФИНАНСОВОЙ МАТЕМАТИКЕ

В §22 было дано определение стандартного винеровского процесса $W(t)$. Он обладает следующими свойствами [9, 16]:

1) $W(t)$ имеет непрерывные траектории, т.е. является непрерывным в среднеквадратичном смысле;

2) $W(0) = 0$;

3) $W(t)$ имеет независимые приращения, т.е. для всяких $s < t \le u < v$ приращения $W(t) - W(s)$ и $W(v) - W(u)$ являются независимыми случайными величинами;

4) для всяких $s < t$ приращение $W(t) - W(s)$ является случайной величиной, имеющей нормальное распределение вероятностей с нулевым математическим ожиданием и дисперсией $t - s$;

5) с вероятностью единица траектории процесса $W(t)$ являются непрерывными функциями времени, которые не дифференцируемы в любой точке t;

6) условное распределение $W(t)$ при заданном $W(\tau)$ для $\tau \le t$ является нормальным со средним значением $W(\tau)$ и дисперсией $t-\tau$;

Заметим, что два последних свойства следуют из предыдущих.

Стандартный винеровский процесс иногда называют **процессом случайного блуждания,** или **стандартным броуновским движением.** В финансовой литературе чаще всего используется последний термин. Характер изменения процесса во времени представлен на рис. 39.

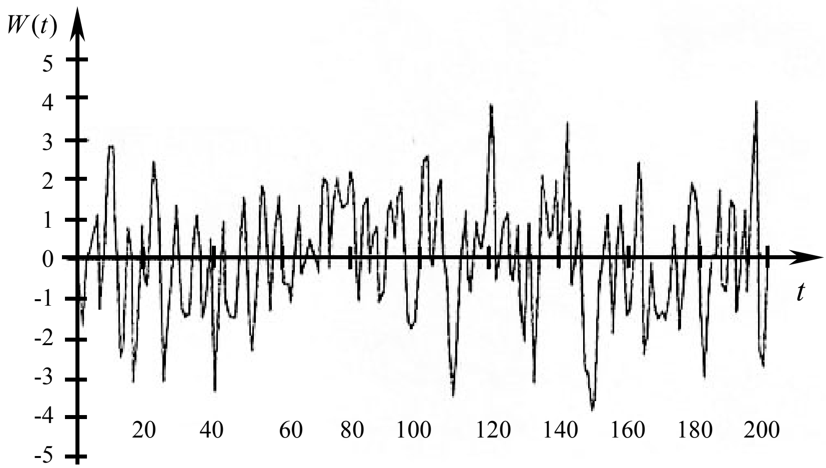


Рис. 39. Типичная реализация процесса броуновского движения

Рассмотрим более подробно некоторые свойства стандартного винеровского процесса. Зафиксируем два момента времени τ и t, $\tau < t$, и используем более удобные обозначения: $\Delta t = t - \tau$, $\Delta W(t) = W(t) - W(\tau)$. Из перечисленных выше свойств стандартного винеровского процесса легко получить следующие результаты:

$$E\left[\Delta W(t)\right]=0,\ E\left[\left(\Delta W(t)\right)^2\right]=\Delta t,$$

$$\mathrm{Var}\left[\Delta \mathrm{W(t)}\right]=\Delta t,\ \mathrm{Var}\left[\left(\Delta W(t)\right)^2\right]=2(\Delta t)^2 \cdot$$

Отсюда видно, что квадрат винеровских приращений $\left(\Delta W(t)\right)^2$ имеет математическое ожидание, которое равно приращению времени Δt. Однако существенным фактом является то, что при малых приращениях времени Δt дисперсия $\left(\Delta W(t)\right)^2$ пренебрежимо мала по сравнению со своим математическим ожиданием. Другими словами, когда Δt стремится к нулю, $\left(\Delta W(t)\right)^2$, конечно, тоже стремится к нулю, но дисперсия будет приближаться к 0 гораздо быстрее, чем математическое ожидание. Таким образом, при уменьшении Δt квадрат $\left(\Delta W(t)\right)^2$ будет выглядеть все более и более «детерменированным», и мы должны принять, что в пределе становится справедливым формальное равенство

$$\left(\Delta W(t)\right)^2=dt \cdot \qquad (11.7)$$

Обычным (не стандартным) процессом броуновского движения $X(t)$ является процесс, приращения которого связаны с приращениями стандартного процесса $W(t)$ соотношением $\frac{\left(\Delta X(t)-\mu\Delta t\right)}{\sigma}=\Delta W(t)$, или

$$\Delta X(t)=\mu\Delta t+\sigma\Delta W(t),$$

где параметры μ и σ определяют соответственно математическое ожидание и дисперсию процесса $X(t)$. Этот процесс является наиболее простым представителем процессов с независимыми приращениями.

В общем случае приращения процесса $X(t)$ можно определить соотношением

$$\Delta X(t)=\mu\left(X(t),t\right)\Delta t+\sigma\left(X(t),t\right)\Delta W(t),$$

откуда, переходя к пределу $\Delta t \to dt$, получаем стохастическое дифференциальное уравнение для процесса с независимыми приращениями:

$$dX(t) = \mu\big(X(t),t\big)dt + \sigma\big(X(t),t\big)dW(t), \; X(0) = X_0. \quad (11.8)$$

Рассмотрим применение уравнений такого типа в финансовой математике.

Пример 11.2 *(арифметическое броуновское движение).*

Пусть $\mu\big(X(t),t\big) = \mu$ и $\sigma\big(X(t),t\big) = \sigma$ – две константы, в этом случае имеем уравнение

$$dX(t) = \mu dt + \sigma dW(t).$$

Тогда говорят, что процесс $X(t)$ следует арифметическому броуновскому движению с **дрейфом** μ и **волатильностью** σ. Процесс описывает реальные экономические показатели, которые растут с линейной скоростью и проявляют увеличивающуюся неопределенность. Отметим следующие свойства такого процесса $X(t)$:

1) $X(t)$ может быть положительным или отрицательным;

2) если $u > t$, то $X(u)$ является будущим значением процесса по отношению к моменту t; распределение $X(u)$ при заданном $X(t)$ является нормальным со средним $X(t) + \alpha(u-t)$ и стандартным отклонением $\sigma\sqrt{(u-t)}$;

3) дисперсия предсказания (прогноза) процесса $X(u)$ стремится к бесконечности с ростом u (при заданных t, $X(t)$).

Рис. 40 демонстрирует выборочную траекторию арифметического броуновского движения с положительным дрейфом $(\mu > 1)$.

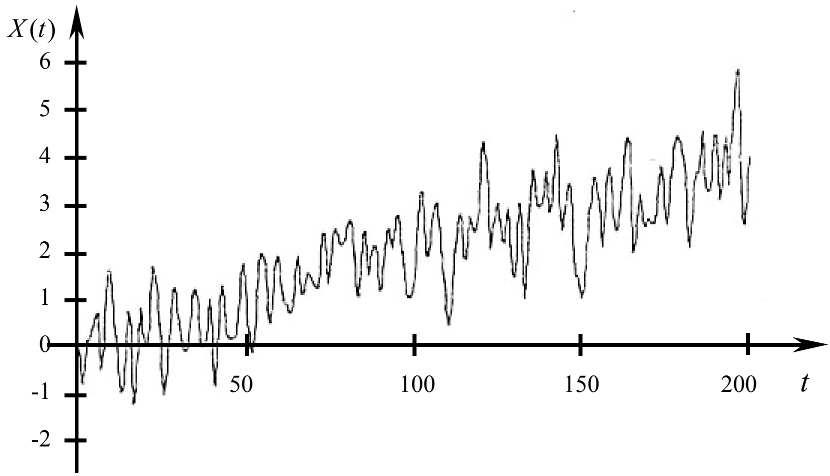


Рис. 40. Арифметическое броуновское движение.
Дрейф равен 0,02, волатильность равна 0,5.

Эти три свойства показывают, что арифметическое броуновское движение описывает показатели, которые могут становиться положительными или отрицательными, имеют нормальное распределение ошибки предсказания и имеют дисперсию предсказания, которая увеличивается линейно во времени. Например, этот процесс может служить математической моделью потока платежей.

Пример 11.3 *(геометрическое броуновское движение).*

Пусть $\mu\big(X(t),t\big)=\mu X(t)$ и $\sigma\big(X(t),t\big)=\sigma X(t)$. Тогда из (11.8) имеем

$$dX(t)=\alpha X(t)dt+\sigma X(t)dW(t)\cdot$$

В этом случае говорят, что процесс $X(t)$ следует геометрическому броуновскому движению с дрейфом μ и волатильностью σ. Процесс соответствует экономическим показателям, которые растут экспоненциально со средней скоростью μ и имеют волатильность, пропорциональную значению самого процесса. Процесс также проявляет увеличивающуюся неопределенность предсказания.

Приведем следующие свойства процесса $X(t)$, отметив при этом, что если $X(t)$ стартует с положительного значения, он будет оставаться положительным:

1) $X(t)$ имеет поглощающий барьер в 0; таким образом, если $X(t)$ достигает 0 (событие нулевой вероятности), то $X(t)$ будет оставаться нулевым;

2) условное распределение $X(u)$ при заданном $X(t)$ является логарифмически нормальным (логнормальным); условное среднее $\ln(X(t))$ для $u > t$ равно $\ln(X(t)) + \alpha(u-t) - 0.5\sigma^2(u-t)$ и условное стандартное отклонение $\ln(X(t))$ равно $\sigma\sqrt{(u-t)}$; $\ln(X(t))$ является нормально распределенным; условное ожидаемое значение $X(u)$ равно $X(t)\exp[\alpha(u-t)]$;

3) дисперсия предсказания $X(u)$ неограниченно увеличивается с ростом u.

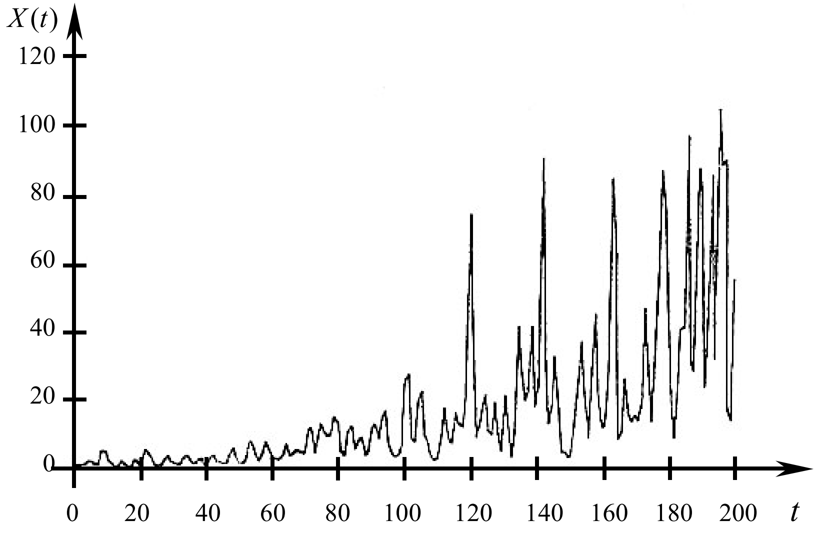


Рис. 41. Геометрическое броуновское движение. Коэффициенты дрейфа и волатильность такие же, как у процесса, представленного на рис. 40.

Геометрическое броуновское движение часто используется для моделирования стоимости активов, так как пропорциональные изменения цены активов являются независимыми и одинаково нормально распределенными. Оно может быть также использовано для моделирования процессов, которые являются положительными и увеличиваются (в среднем) с постоянной экспоненциальной скоростью. Например, можно применить геометрическое броуновское движение для моделирования номинальной цены товаров потребления или доходов от конкретной деятельности. В некоторых случаях также желательна отрицательная скорость изменения положительной переменной. Геометрическое броуновское движение также может быть использовано. Для моделирования процессов, изображенных на рис. 40 и рис. 41, использован один и тот же процесс $W(t)$, чтобы продемонстрировать, как могут выглядеть соответствующие ему арифметическое и геометрическое броуновские движения.

Пример 11.4 *(процесс, возвращающийся к среднему).*

Процесс, возвращающийся к среднему, называется процессом Орнштейна–Уленбека, когда $\gamma = 0$. Пусть $\mu(X(t), t) = k(\mu - X(t))$ и $\sigma(X(t), t) = \sigma X(t)$, где $k \geq 0$ и γ произвольные. В этом случае получаем уравнение

$$dX(t) = k(\mu - X(t))dt + \sigma X(t) dW(t),$$

и говорят, что процесс $X(t)$ следует процессу возвращения к среднему с параметром регулирования скорости k, средним установления μ и волатильностью σ. Выбор γ дает возможность изменять характер волатильности процесса. Этот процесс является подходящим для описания экономических показателей, которые имеют тенденцию устанавливаться к среднему значению, но могут быть подвержены краткосрочным возмущениям. Мы предположим, что k, μ и γ являются положительными. Процесс имеет следующие свойства:

1) $X(t)$ является положительным для положительных стартовых значений;

2) когда $X(t)$ достигает 0, дрейф становится положительным и волатильность исчезает;

3) когда u становится неограниченным, дисперсия предсказания $X(u)$ является конечной;

4) если $\gamma = 0$, распределение $X(u)$ при заданном $X(t)$ для $u > t$ является нормальным; при этом условное среднее распределение равно

$$\left(X(t-\mu)\right)\exp\left[-k(u-t)\right]+\mu,$$

а условная дисперсия имеет вид

$$\left[\sigma^2/(2k)\right]\left(1-\exp\left[-2k(u-t)\right]\right);$$

5) если $\gamma = 0{,}5$, распределение $X(u)$ при заданном $X(t)$ для $u > t$ является нецентральным χ^2; среднее распределения равно

$$\left(X(t)-\mu\right)\exp\left[-k(u-t)\right]+\mu,$$

а дисперсия распределения равна

$$X(t)\frac{\sigma^2}{k}\left(e^{-k(u-t)}-e^{-2k(u-t)}\right)+\mu\frac{\sigma^2}{2k}\left(1-e^{-k(u-t)}\right)^2.$$

Процесс, возвращающийся к среднему, является подходящим для описания процентных ставок или индексов инфляции, которые могут иметь устойчивые установившиеся значения, и не подходит для торгуемых активов. Можно также моделировать саму волатильность (если волатильность изменяется непредсказуемо) как процесс, возвращающийся к среднему. Процесс, возвращающийся к среднему при $\gamma = 0{,}5$, может выглядеть подобно представленному на рис. 42, где использован тот же процесс $W(t)$, что и на рис. 40 и 41.

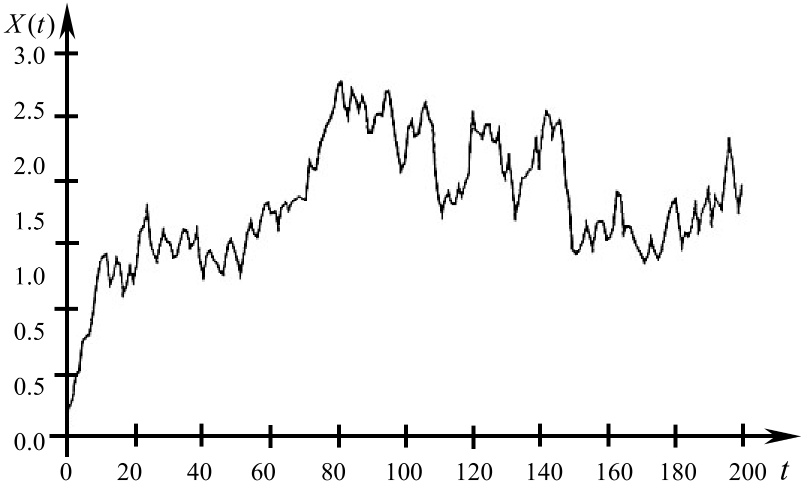


Рис. 42. Процесс, возвращающийся к среднему,

$\gamma = 0{,}5$; $k = 0{,}7$; $\mu = 2$; $\sigma = 0{,}1$

ПРИЛОЖЕНИЕ 1

ТАБЛИЦА ЗНАЧЕНИЙ ФУНКЦИИ $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

ПРИЛОЖЕНИЕ 2

ТАБЛИЦА ЗНАЧЕНИЙ ФУНКЦИИ $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{z^2}{2}}dz$

x	Φ*(x)*	*x*	Φ*(x)*	*x*	Φ*(x)*	*x*	Φ*(x)*
0,00	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,01	0,0040	0,33	0,1293	0,65	0,2422	0,97	0,3340
0,02	0,0080	0,34	0,1331	0,66	0,2454	0,98	0,3365
0,03	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,04	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,05	0,0199	0,37	0,1443	0,69	0,2549	1,01	0,3438
0,06	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,07	0,0279	0,39	0,1517	0,71	0,2611	1,03	0,3485
0,08	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,09	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,10	0,0398	0,42	0,1628	0,74	0,2703	1,06	0,3554
0,11	0,0438	0,43	0,1664	0,75	0,2734	1,07	0,3577
0,12	0,0478	0,44	0,1700	0,76	0,2764	1,08	0,3599
0,13	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,14	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,15	0,0596	0,47	,0,1808	0,79	0,2852	1,11	0,3665
0,16	0,0636	0,48	0,1844	0,80	0,2881	1,12	0,3686
0,17	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,18	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,19	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3749
0,20	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,21	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,22	0,0871	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,23	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,24	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,25	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869
0,26	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3883
0,27	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,28	0,1103	0,60	0,2257	0,92	0,3212	1,24	0,3925
0,29	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,30	0,1179	0,62	0,2324	0,94	0,3264		
0,31	0,1217	0,63	0,2357	0,95	0,3289		

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,26	0,396	1,59	0,4441	1,92	0,4726	2,50	0,4938
1,27	0,398	1,60	0,4452	1,93	0,4732	2,52	0,4941
1,28	0,399	1,61	0,4463	1,94	0,4738	2,54	0,4945
1,29	0,401	1,62	0,4474	1,95	0,4744	2,56	0,4948
1,30	0,403	1,63	0,4484	1,96	0,4750	2,58	0,4951
1,31	0,404	1,64	0,4495	1,97	0,4756	2,60	0,4953
1,32	0,406	1,65	0,4505	1,98	0,4761	2,62	0,4956
1,33	0,408	1,66	0,4515	1,99	0,4767	2,64	0,4959
1,34	0,409	1,67	0,4525	2,00	0,4772	2,66	0,4961
1,35	0,411	1,68	0,4535	2,02	0,4783	2,68	0,4963
1,36	0,413	1,69	0,4545	2,04	0,4793	2,70	0,4965
1,37	0,414	1,70	0,4554	2,06	0,4803	2,72	0,4967
1,38	0,416	1,71	0,4564	2,08	0,4812	2,74	0,4969
1,39	0,417	1,72	0,4573	2,10	0,4821	2,76	0,4971
1,40	0,419	1,73	0,4582	2,12	0,4830	2,78	0,4973
1,41	0,420	1,74	0,4591	2,14	0,4838	2,80	0,4974
1,42	0,422	1,75	0,4599	2,16	0,4846	2,82	0,4976
1,43	0,423	1,76	0,4608	2,18	0,4854	2,84	0,4977
1,44	0,425	1,77	0,4616	2,20	0,4861	2,86	0,4979
1,45	0,426	1,78	0,4625	2,22	0,4868	2,88	0,4980
1,46	0,427	1,79	0,4633	2,24	0,4875	2,90	0,4981
1,47	0,429	1,80	0,4641	2,26	0,4881	2,92	0,4982
1,48	0,430	1,81	0,4649	2,28	0,4887	2,94	0,4984
1,49	0,431	1,82	0,4656	2,30	0,4893	2,96	0,4985
1,50	0,433	1,83	0,4664	2,32	0,4898	2,98	0,4986
1,51	0,434	1,84	0,4671	2,34	0,4904	3,00	0,49865
1,52	0,435	1,85	0,4678	2,36,	0,4909	3,20	0,49931
1,53	0,437	1,86	0,4686	2,38	0,4913	,3,40	0,49966
1,54	0,438	1,87	0,4693	2,40	0,4918	3,60	0,49984
1,55	0,439	1,88	0,4699	2,42	0,4922	3,80	0,49992
1,56	0,440	1,89	0,4706	2,44	0,4927	4,00	0,49996
1,57	0,441	1,90	0,4713	2,46	0,4931	4,50	0,49999
1,58	0,442	1,91	0,4719	2,48	0,4934	5,00	0,49999

ПРИЛОЖЕНИЕ 3

РАСПРЕДЕЛЕНИЕ ПУАССОНА $P(\xi = k) = \dfrac{\lambda^k}{k!} e^{-\lambda}$

k	0	1	2	3	4	5	6	7	8	9	10	11
0,1	0,9048	0,0905	0,0045	0,0002	0,0000							
0,2	0,8187	0,1637	0,0164	0,0011	0,0001	0,0000						
0,3	0,7408	0,2222	0,0333	0,0033	0,0003	0,0000						
0,4	0,6703	0,2681	0,0536	0,0072	0,0007	0,0001	0,0000					
0,5	0,6065	0,3033	0,0758	0,0126	0,0016	0,0002	0,0000					
0,6	0,5488	0,3293	0,0988	0,0198	0,0030	0,0004	0,0000					
0,7	0,4966	0,3476	0,1217	0,0284	0,0050	0,0007	0,0001	0,0000				
0,8	0,4493	0,3595	0,1438	0,0383	0,0077	0,0012	0,0002	0,0000				
0,9	0,4066	0,3659	0,1647	0,0494	0,0111	0,0020	0,0003	0,0000				
1,0	0,3679	0,3679	0,1839	0,0613	0,0153	0,0031	0,0005	0,0001	0,0000			
1,5	0,2231	0,3347	0,2510	0,1255	0,0471	0,0141	0,0035	0,0008	0,0001	0,0000		
2,0	0,1353	0,2707	0,2707	0,1804	0,0902	0,0361	0,0120	0,0034	0,0009	0,0002	0,0000	
2,5	0,0821	0,2052	0,2565	0,2138	0,1336	0,0668	0,0278	0,0099	0,0031	0,0009	0,0002	0,0000
3,0	0,0498	0,1494	0,2240	0,2240	0,1680	0,1008	0,0504	0,0216	0,0081	0,0027	0,0008	0,0002
3,5	0,0302	0,1057	0,1850	0,2158	0,1888	0,1322	0,0771	0,0385	0,0169	0,0066	0,0023	0,0007
4,0	0,0183	0,0733	0,1465	0,1954	0,1954	0,1563	0,1042	0,0595	0,0298	0,0132	0,0053	0,0019
4,5	0,0111	0,0500	0,1125	0,1687	0,1898	0,1708	0,1281	0,0824	0,0463	0,0232	0,0104	0,0043
5,0	0,0067	0,0337	0,0842	0,1404	0,1755	0,1755	0,1462	0,1044	0,0653	0,0363	0,0181	0,0082
6,0	0,0025	0,0149	0,0446	0,0892	0,1339	0,1606	0,1606	0,1377	0,1033	0,0688	0,0413	0,0225
7,0	0,0009	0,0064	0,0223	0,0521	0,0912	0,1277	0,1490	0,1490	0,1304	0,1014	0,0710	0,0452
8,0	0,0003	0,0027	0,0107	0,0286	0,0573	0,0916	0,1221	0,1396	0,1396	0,1241	0,0993	0,0722
9,0	0,0001	0,0011	0,0050	0,0150	0,0337	0,0607	0,0911	0,1171	0,1318	0,1318	0,1186	0,0970
10,0	0,0000	0,0005	0,0023	0,0076	0,0189	0,0378	0,0631	0,0901	0,1126	0,1251	0,1251	0,1137

k	12	13	14	15	16	17	18	19	20	21	22	23	24
3,0	0,0001	0,0000											
3,5	0,0002	0,0001	0,0000										
4,0	0,0006	0,0002	0,0001	0,0000									
4,5	0,0016	0,0006	0,0002	0,0001	0,0000								
5,0	0,0034	0,0013	0,0005	0,0002	0,0000								
6,0	0,0113	0,0052	0,0022	0,0009	0,0003	0,0001	0,0000						
7,0	0,0263	0,0142	0,0071	0,0033	0,0014	0,0006	0,0002	0,0001	0,0000				
8,0	0,0481	0,0296	0,0169	0,0090	0,0045	0,0021	0,0009	0,0004	0,0002	0,0001	0,0000		
9,0	0,0728	0,0504	0,0324	0,0194	0,0109	0,0058	0,0029	0,0014	0,0006	0,0003	0,0001	0,0000	
10,0	0,0948	0,0729	0,0521	0,0347	0,0217	0,0128	0,0071	0,0037	0,0019	0,0009	0,0004	0,0002	0,0001

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. Вентцель А.Д. Курс теории случайных процессов. – М.: Наука, 1975. – 319 с.

2. Гихман И.И., Скороход А.В. Введение в теорию случайных процессов. – М.: Наука, 1977. – 567 с.

3. Гнеденко Б.В. Курс теории вероятностей. – М.: Наука, 1988. – 447 с.

4. Карлин С. Основы теории случайных процессов. – М.: Наука, 1971. – 536 с.

5. Кемени Д., Снелл Дж., Кнепп А. Счетные цепи Маркова. – М.: Наука, 1987. – 414 с.

6. Коваленко И.Н. Филлипова А.А. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1982. – 256 с.

7. Лапо П.М., Маталыцкі М.А. Задачы па тэорыі імавернасцей. – Мінск: Універсітэцкае, 1995. – 87 с.

8. Маталыцкий М.А., Романюк Т.В. Теория вероятностей в примерах и задачах. Учебное пособие для студентов математических специальностей вузов. – Гродно: ГрГУ, 2002. – 247 с.

9. Медведев Г.А. Математические основы финансовой экономики. Часть 1. Мартингальный подход. – Минск: БГУ, 2003. – 287 с.

10. Радюк Л.Е., Терпугов А.Ф. Теория вероятностей и случайных процессов. – Томск: ТГУ, 1988. – 175 с.

11. Розанов Ю.А. Теория вероятностей, случайные процессы и математическая статистика. – М.: Наука, 1985. – 320 с.

12. Феллер В. Введение в теорию вероятностей и ее приложения. Т.1. – М.: Мира, 1984. – 526 с.

13. Ширяев А.Н. Вероятность. – М.: Наука, 1980. – 573 с.

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

14. Маталыцкий М.А., Сети массового обслуживания в стационарном и переходном режимах: Монография. – Гродно: ГрГУ, 2001. – 211 с.

15. Маталыцкий М.А., Романюк Т.В. Приближенные методы анализа сетей с центральной системой обслуживания и их применения: Монография. – Гродно: ГрГУ, 2003. – 200 с.

16. Медведев Г.А. Математические модели финансовых рисков. Часть 1. – Минск: БГУ, 1999. – 239 с.

Содержание

Учебное издание

Маталыцкий Михаил Алексеевич

**ЭЛЕМЕНТЫ ТЕОРИИ
СЛУЧАЙНЫХ ПРОЦЕССОВ**

Учебное пособие

Редактор Н.Н. Красницкая
Компьютерная верстка: Р.Н. Баранчик
Дизайн обложки: О.В. Канчуга

Сдано в набор 09.04.2004. Подписано в печать 05.07.2004.
Формат 60×84/16. Бумага офсетная. Печать RISO. Гарнитура Таймс.
Усл.печ. л. 18,91. Уч.-изд. л. 17,93. Тираж 200 экз. Заказ .

Учреждение образования «Гродненский государственный
университет имени Янки Купалы».
ЛИ №02330/0133257 от 30.04.2004. Ул. Пушкина, 39, 230012, Гродно.

Отпечатано на технике издательского центра Учреждения образования
«Гродненский государственный университет имени Янки Купалы».
ЛП №02330/0056882 от 30.04.2004. Ул. Пушкина, 39, 230012, Гродно.