Лазакович Н.В., Сташуленок С.П., Яблонский О.Л.

КУРС ТЕОРИИ ВЕРОЯТНОСТЕЙ

Учебное пособие

Лазакович Н.В., Сташуленок С.П., Яблонский О.Л. Курс теории вероятностей [Электронный ресурс]: Учебное пособие. — Электрон. текст. дан. (6.3 Мб) и анимации в системе Mathematica (102 Мб). — Мн.: “Электронная книга БГУ”, 2003, www.elbook.bsu.by. — 1 электрон. опт. диск CD-ROM.

МИНСК

«Электронная книга БГУ»

2003

elbook@bsu.by

Краткая авторская аннотация

В основу учебного пособия положен годовой курс лекций, которые авторы в течение ряда лет читали для студентов механико-математического факультета Белорусского государственного университета. В книге содержатся следующие разделы: вероятностные пространства, независимость, случайные величины, числовые характеристики случайных величин, характеристические функции, предельные теоремы, основы теории случайных процессов, элементы математической статистики и приложения, в которых приведены таблицы основных вероятностных распределений и значения некоторых из них. Большинство глав включает в себя дополнения, куда вынесены вспомогательный материал и темы для самостоятельного изучения. Изложение сопровождается большим количеством примеров, упражнений и задач, иллюстрирующих основные понятия и поясняющих возможные применения доказанных утверждений.

Для студентов математических специальностей университетов.

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ 5
ВВЕДЕНИЕ 6
Глава 1. ВЕРОЯТНОСТНЫЕ ПРОСТРАНСТВА 9
§ 1. Терминология теории вероятностей 9
§ 2. Аксиоматика Колмогорова 15
§ 3. Примеры вероятностных пространств 20
ДОПОЛНЕНИЯ К ГЛАВЕ 1 26
§ 1. Основные понятия комбинаторики 26
§ 2. Классические теоретико-вероятностные модели 30
§ 3. Частотное (статистическое) определение вероятности 33
§ 4. Борелевские сигма-алгебры 34
Задачи 36
Глава 2. НЕЗАВИСИМОСТЬ 39
§ 1. Условные вероятности 39
§ 2. Независимость событий 43
§ 3. Независимые испытания 46
ДОПОЛНЕНИЯ К ГЛАВЕ 2 51
§ 1. Предельные теоремы в схеме Бернулли 51
§ 2. Цепи Маркова 54
Задачи 58
Глава 3. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ 60
§ 1. Случайные величины и их распределения 60
§ 2. Классификация случайных величин 67
§ 3. Многомерные случайные величины 79
§ 4. Независимость случайных величин 86
ДОПОЛНЕНИЯ К ГЛАВЕ 3 91
§ 1. Функциональные преобразования случайных величин 91
Задачи 93
Глава 4. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН 95
§ 1. Математическое ожидание и его свойства 95
§ 2. Моменты случайных величин 103
§ 3. Неравенства. Коэффициент корреляции 109
ДОПОЛНЕНИЯ К ГЛАВЕ 4 116
§ 1. Интеграл Лебега 116
§ 2. Интегралы Римана – Стилтьеса и Лебега – Стилтьеса 117
§ 3. Условные математические ожидания 120
Задачи 124

Глава 5. ХАРАКТЕРИСТИЧЕСКИЕ ФУНКЦИИ 127
§ 1. Определение и простейшие свойства 127
§ 2. Формулы обращения для характеристических функций 136
§ 3. Непрерывность соответствия между множествами функций распределения и характеристических функций 141
ДОПОЛНЕНИЯ К ГЛАВЕ 5 149
§ 1. Производящие функции 149
§ 2. Решетчатые распределения 151
§ 3. Многомерные характеристические функции 152
§ 4. Многомерное нормальное распределение и связанные с ним распределения 154
Задачи 158
Глава 6. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ 161
§ 1. Центральная предельная теорема 161
§ 2. Сходимость случайных величин 172
§ 3. Законы больших чисел 176
ДОПОЛНЕНИЯ К ГЛАВЕ 6 189
§ 1. Сходимость рядов 189
Задачи 195
Глава 7. ОСНОВЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ 198
§ 1. Определение случайного процесса 198
§ 2. Случайные процессы с независимыми приращениями 204
§ 3. Корреляционная теория случайных процессов 216
§ 4. Марковские случайные процессы 228
ДОПОЛНЕНИЯ К ГЛАВЕ 7 235
§ 1. Обобщенные случайные процессы 235
Задачи 249
Глава 8. ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ 252
§ 1. Основные понятия и элементы выборочной теории 253
§ 2. Оценивание неизвестных параметров 260
§ 3. Проверка статистических гипотез 281
§ 4. Параметрические гипотезы 287
§ 5. Линейная регрессия и метод наименьших квадратов 294
Задачи 299
ПРИЛОЖЕНИЯ 301
ЛИТЕРАТУРА 307
ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ 309
УКАЗАТЕЛЬ ОБОЗНАЧЕНИЙ 314

ПРЕДИСЛОВИЕ

В основу настоящего учебного пособия положен расширенный курс лекций, которые в течение ряда лет читают авторы на механико-математическом факультете Белорусского государственного университета. Его содержание полностью соответствует программе годового курса теории вероятностей и математической статистики для студентов математических специальностей университетов.

Материал пособия разбит на введение, восемь глав (вероятностные пространства, независимость, случайные величины, числовые характеристики случайных величин, характеристические функции, предельные теоремы, основы теории случайных процессов, элементы математической статистики) и приложения, в которых приводятся таблицы основных вероятностных распределений и значений некоторых из них.

Первые семь глав имеют дополнения, которые, с одной стороны, должны сделать пособие самодостаточным, с другой – вынести часть материала данного курса на самостоятельное изучение.

В конце каждой главы приводятся задачи, значимость которых различна: одни носят характер простых упражнений, в других предлагается доказать утверждения, сформулированные, но не доказанные в основном тексте, третьи должны дать дополнительные сведения к рассматриваемому кругу вопросов.

Каждый параграф пособия имеет свою нумерацию определений, теорем, лемм, утверждений, упражнений, примеров и соотношений.

Пособие снабжено иллюстрациями, выполненными в Mathematica 4.1, которые можно посмотреть после нажатия на ссылку Иллюстрация в Mathematica. Кроме этого некоторые теоремы снабжены схемами доказательства, для просмотра которых достаточно нажать Схема доказательства.

Авторы считают своим долгом выразить искреннюю благодарность рецензентам, чьи критические замечания и полезные советы, несомненно, способствовали улучшению содержания пособия. Мы также признательны Н. В. Чесалину и всем студентам, оказавшим помощь в наборе текста пособия.

Н. В. Лазакович,
С. П. Сташулёнок,
О. Л. Яблонский

ВВЕДЕНИЕ

Теория вероятностей – это большой, интенсивно развивающийся раздел математики, изучающий случайные явления. Казалось бы, науке, в частности, математике, нет дела до случайных явлений, ведь наука открывает законы, которые помогают предсказывать течение того или иного процесса или явления, а случайное явление – как раз такое явление, исход которого предсказать невозможно. Однако и случайные явления подчиняются некоторым закономерностям, которые называют *вероятностными*. Прежде всего условимся, что мы будем иметь дело не со всеми случайными явлениями, а с *массовыми*, т. е. с такими, исходы которых в принципе возможно наблюдать в одних и тех же условиях много раз. Пусть при N-кратном осуществлении некоторого комплекса условий случайное событие A происходит $N(A)$ раз. Отношение $N(A)/N$ называется *относительной частотой события* A. Оказывается, для больших N относительная частота события A в массовых явлениях обладает так называемым свойством *устойчивости*, которое состоит в том, что отношение $N(A)/N$ колеблется относительно одного и того же числа. Это число $\mathrm{P}(A)$ в соответствующей математической модели называют *вероятностью события* A. Устойчивость относительных частот событий – объективное свойство массовых случайных явлений реального мира. Отсутствие устойчивости относительных частот в сериях испытаний свидетельствует о том, что условия, в которых проводятся испытания, претерпевают значительные изменения.

Теория вероятностей – это наука, которая изучает математические модели массовых случайных явлений. Если говорить конкретнее, то теория вероятностей устанавливает такие связи между вероятностями случайных событий в математических моделях, которые позволяют вычислять вероятности сложных событий по вероятностям более простых событий.

Призванная изучать количественные характеристики случайных событий, теория вероятностей, как и всякая точная наука, стала таковой лишь тогда, когда было четко сформулировано понятие вероятностной модели, и была создана ее аксиоматика.

В этой связи естественно хотя бы кратко остановиться на основных этапах развития теории вероятностей.

Возникновение теории вероятностей как науки относится к середине XVII в., и связано с именами Б. Паскаля, П. Ферма, Х. Гюйгенса. Хотя отдельные задачи, касающиеся подсчета шансов в азартных играх, рассматривались и ранее – в XV–XVI в. итальянскими математиками. Первые общие методы решения таких задач были, по-видимому, даны в знаменитой переписке Паскаля и Ферма, начавшейся в 1654 г. и в первой книге по теории вероятностей «О расчетах в азартной игре», опубликованной Гюйгенсом в 1657 г.

Истинная теория вероятностей начинается с работы Я. Бернулли «Искусство предположения» (1713), в которой была доказана первая предельная теорема – закон больших чисел и работы А. Муавра «Аналитические методы» (1730), в которой сформулирована и доказана так называемая центральная предельная теорема.

В 1812 г. вышел большой трактат П. Лапласа «Аналитическая теория вероятностей», в котором обобщалась теорема Муавра на несимметричный случай схемы Бернулли, и давались применения вероятностных методов к теории ошибок наблюдений.

К этому же периоду в развитии теории вероятностей, когда центральное место в исследованиях занимали предельные теоремы, относятся работы С. Пуассона и К. Ф. Гаусса. С именем Пуассона в современной теории вероятностей связаны понятия распределения и процесса, носящих его имя. Гауссу принадлежит заслуга создания теории ошибок.

Следующий важный период в становлении теории вероятностей связан с именами П. Л. Чебышева, А. А. Маркова, А. М. Ляпунова, создавших в начале XIX в. эффективные методы доказательства предельных теорем для сумм независимых случайных величин.

Современный период в развитии теории вероятностей начинается с установления аксиоматики. Первые работы в этом направлении принадлежат С. Н. Бернштейну, Р. Мизесу и Э. Борелю. В 1933 г. вышла книга А. Н. Колмогорова «Основные понятия теории вероятностей», в которой была предложена аксиоматика, получившая всеобщее признание и позволившая охватить не только все классические разделы теории вероятностей, но и дать строгую основу для развития ее новых разделов – теории случайных процессов и математической статистики. Изложение теории вероятностей в настоящей книге основано на аксиоматическом подходе А. Н. Колмогорова.

ЛАТИНСКИЙ АЛФАВИТ

A a	а	*J j*	йот(жи)	*S s*	эс
B b	бэ	*K k*	ка	*T t*	тэ
C c	цэ	*L l*	эль	*U u*	у
D d	дэ	*M m*	эм	*V v*	вэ
E e	е	*N n*	эн	*W w*	дубль-вэ
F f	эф	*O o*	О	*X x*	икс
G g	же	*P p*	пэ	*Y y*	игрек
H h	аш	*Q q*	ку	*Z z*	зэт
I i	и	*R r*	эр		

ГРЕЧЕСКИЙ АЛФАВИТ

Α α	альфа	Ι ι	йота	Ρ ρ	ро
Β β	бета	Κ κ	каппа	Σ σ	сигма
Γ γ	гамма	Λ λ	ламбда	Τ τ	тау
Δ δ	дельта	Μ μ	мю	Υ υ	ипсилон
Ε ε	эпсилон	Ν ν	ню	Φ φ	фи
Ζ ζ	дзета	Ξ ξ	кси	Χ χ	хи
Η η	эта	Ο ο	омикрон	Ψ ψ	пси
Θ θ	тэта	Π π	пи	Ω ω	омега

Глава 1. ВЕРОЯТНОСТНЫЕ ПРОСТРАНСТВА

В настоящей главе на основании аксиоматики Колмогорова строится математическая модель произвольного случайного явления – вероятностное пространство и приводятся наиболее распространенные примеры таких моделей.

Исторически в теории вероятностей сложилась своя терминология. Изложение материала начнем с обсуждения этой терминологии и ее интерпретации в теории множеств.

§ 1. Терминология теории вероятностей

Для математического описания экспериментов со случайными исходами, прежде всего, потребуется понятие *пространства элементарных событий (исходов)*. Таким пространством будем называть любое множество элементарных событий, обладающих следующими свойствами. Во-первых, все они взаимно исключают друг друга – являются непересекающимися, т. е. в результате эксперимента происходит одно и только одно элементарное событие. Во-вторых, каждый интересующий нас результат эксперимента может быть однозначно описан с помощью элементарных событий, которые будем обозначать $\omega, \omega \in \Omega$.

Пространство элементарных событий можно трактовать как множество всех исходов исследуемого случайного явления.

Отметим, что само понятие пространства элементарных событий математически является неопределяемым – оно исходно, так же как понятие точки в геометрии. Конкретная природа Ω нас, как правило, интересовать не будет.

Пример 1. Рассмотрим опыт с бросанием игральной кости. Пространство элементарных событий имеет вид $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, где ω_k – элементарное событие, состоящее в том, что на верхней грани кубика выпала цифра k, $k = 1, \ldots, 6$.

Пример 2. Внутрь единичного квадрата бросается частица. В этом случае пространство элементарных событий может быть описано следующим образом:

$$\Omega = \{\omega = (x, y) : 0 \le x, y \le 1\},$$

т. е. здесь элементарное событие – произвольная точка квадрата.

В реальном мире случайное событие – это исход какого-либо эксперимента, который может как произойти, так и не произойти. В том случае, когда число элементарных событий не более чем счетно, *случайным событием* будем называть любое подмножество $A \subseteq \Omega$ элементов из Ω (событие A произошло, если произошло какое-либо из элементарных событий $\omega \in A$). Понятно, что элементарные события являются случайными событиями. В случае, когда пространство Ω несчетно бесконечно, *случайными событиями* будем называть элементы некоторой σ-алгебры $\mathcal{A}$ подмножеств из Ω (необходимость этого будет выяснена позже).

Определение 1. *Класс $\mathcal{A}$ подмножеств пространства Ω назовем алгеброй событий, если*

1) $\varnothing \in \mathcal{A}$, $\Omega \in \mathcal{A}$;

2) *из того, что* $A \in \mathcal{A}$ *следует* $\overline{A} = \Omega \setminus A \in \mathcal{A}$;

3) *из соотношений* $A \in \mathcal{A}$, $B \in \mathcal{A}$ *следует, что* $A \bigcup B \in \mathcal{A}$.

Определение 2. *Алгебра $\mathcal{A}$ называется σ-алгеброй (сигма-алгеброй), если из того, что $A_n \in \mathcal{A}$ для всех $n \in N$ следует, что*

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}.$$

Упражнение 1. Докажите, что σ-алгебра замкнута относительно всех теоретико-множественных операций.

Упражнение 2. Установите, совпадают ли понятия алгебры и σ-алгебры?

Достоверным событием будем называть случайное событие, которое всегда происходит при осуществлении данного комплекса условий. Несложно видеть, что оно совпадает со всем пространством элементарных событий, поэтому будем его обозначать Ω. *Невозможным событием* назовем случайное событие, которое никогда не происходит при реализации данного комплекса условий. Обозначать невозможное событие будем $\varnothing$. Так, в примере 1, событие, состоящее в выпадении не менее одного очка, является достоверным, а событие, состоящее в том, что выпало семь очков – невозможным.

Упражнение 3. Укажите невозможное и достоверное события для примера 2.

Пусть дано некоторое пространство элементарных событий Ω. Определим теперь операции между двумя событиями A и B. *Произведением* или *пересечением* событий назовем событие, обозначаемое $A \bigcap B$ или AB, которое происходит тогда и только тогда, когда одновременно происходят события A и B. События A и B назовем *несовместными*, если $A \bigcap B = \varnothing$ (т. е. вместе произойти они не могут). *Суммой* или *объединением* событий A и B назовем событие, обозначаемое $A \bigcup B$ или $A + B$ (в случае, когда они несовместны), которое происходит тогда и только тогда, когда происходят A или B, или оба вместе. *Разностью* событий A и B назовем событие $A \setminus B$, происходящее тогда и только тогда, когда произошло A, но не произошло B. *Событием, противоположным событию A,* будем называть событие $\overline{A} = \Omega \setminus A$. *Симметрической разностью* назовем событие $A \Delta B$, которое происходит тогда и только тогда, когда происходит хотя бы одно из событий $A \setminus B$ или $B \setminus A$.

Иллюстрация в Mathematica

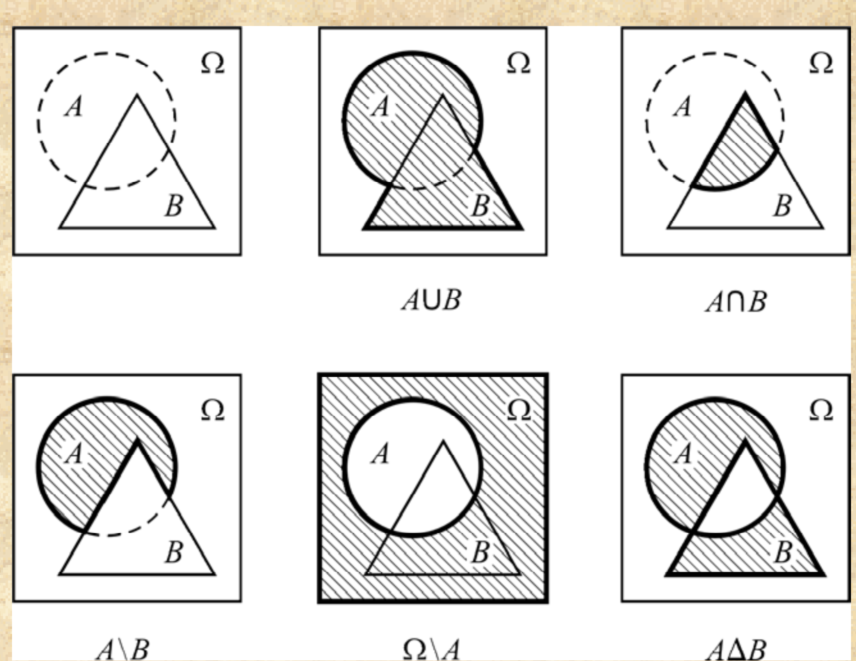

Рис. 1. Сумма, произведение, разность событий A и B; событие, противоположное A и симметрическая разность B и A.

Пример 3. Бросается игральная кость. Тогда $\Omega = \{1, 2, 3, 4, 5, 6\}$ (см. пример 1). Рассмотрим следующие случайные события:

A = {выпало четное число очков}, т. е. $A = \{2, 4, 6\}$,

B = {выпавшее число очков больше трех}, т. е. $B = \{4, 5, 6\}$.

Найдите $\overline{A}$, $\overline{B}$, $A \bigcup B$, $A \bigcap B$, $A \setminus B$. Являются ли события A и B несовместными?

Решение:

$\overline{A} = \{1, 3, 5\}$, $\overline{B} = \{1, 2, 3\}$, $A \bigcup B = \{2, 4, 5, 6\}$, $A \bigcap B = \{4, 6\}$, $A \setminus B = \{2\}$. События A и B не являются несовместными, так как $A \bigcap B \neq \varnothing$.

Упражнение 4. В рамках примера 2 задайте случайные события A и B и найдите $\overline{A}$, $\overline{B}$, $A \bigcup B$, $A \bigcap B$, $A \setminus B$. Выясните, являются ли A и B несовместными событиями?

Пусть $\{A_n\}_{n=1}^{\infty}$ – бесконечная последовательность случайных событий. Обозначим через A^* множество всех тех и только тех элементарных событий, которые принадлежат бесконечному числу множеств A_n. Тогда

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

Действительно, если $\omega \in A^*$, т. е. ω принадлежит бесконечному числу множеств A_n, то

$$\omega \in \bigcup_{m=n}^{\infty} A_m$$

для каждого n, и, следовательно,

$$\omega \in \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m,$$

т. е.

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

Если

$$\omega \in \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m,$$

то $\omega \in A^*$.

Пусть A_* – множество тех и только тех $\omega \in \Omega$, которые принадлежат всем A_n, за исключением конечного числа. Тогда, проводя рассуждения, аналогичные приведенным выше, получаем

$$A_* = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m.$$

Множество A^* можно интерпретировать как событие, состоящее в том, что произойдет бесконечно много событий из последовательности $\{A_n\}_{n=1}^{\infty}$. Множество A_* есть событие, состоящее в том, что произойдут все события A_n, начиная с некоторого номера, или, что то же самое, не произойдет только конечное число событий из последовательности $\{A_n\}_{n=1}^{\infty}$. Очевидно, что $A_* \subseteq A^*$. Событие A^* называется *верхним пределом* последовательности событий $\{A_n\}_{n=1}^{\infty}$,

$$A^* = \overline{\lim_{n\to\infty}} A_n.$$

Событие A_* называется *нижним пределом* последовательности событий $\{A_n\}_{n=1}^{\infty}$,

$$A_* = \varliminf_{n\to\infty} A_n .$$

Если $A^* = A_*$, *то говорят, что последовательность событий* $\{A_n\}_{n=1}^{\infty}$ *имеет предел*

$$\lim_{n\to\infty} A_n = \varlimsup_{n\to\infty} A_n = \varliminf_{n\to\infty} A_n .$$

Упражнение 5. Пусть $\{A_n\}$ – монотонно убывающая последовательность событий, т. е. $A_1 \supset A_2 \supset ... \supset A_n \supset ...$. Докажите, что

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n .$$

Упражнение 6. Пусть $\{A_n\}$ – монотонно возрастающая последовательность событий, т. е. $A_1 \subset A_2 \subset ... \subset A_n \subset ...$. Покажите, что

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n .$$

Упражнение 7. Пусть I_A – индикатор события A (характеристическая функция множества A). Докажите справедливость следующих соотношений:

1) $I_{A^*} = \varlimsup_{n\to\infty} I_{A_n} \Leftrightarrow A^* = \varlimsup_{n\to\infty} A_n$;

2) $I_{A_*} = \varliminf_{n\to\infty} I_{A_n} \Leftrightarrow A_* = \varliminf_{n\to\infty} A_n$;

3) $I_A = \lim_{n\to\infty} I_{A_n} \Leftrightarrow A = \lim_{n\to\infty} A_n$.

Упражнение 8. Докажите правило де Моргана: если имеются некоторые соотношения между событиями, выраженные через символы $\bigcap$, $\bigcup$, $\subset$, $\supset$, то при переходе к противоположным событиям все символы необходимо повернуть на 180°, т. е. заменить соответственно на $\bigcup$, $\bigcap$, $\supset$, $\subset$.

Подытоживая все, сказанное выше, приведем таблицу, показывающую, как некоторые понятия теории множеств интерпретируются в теории вероятностей.

Таблица 1

Обозначения	Язык теории множеств	Язык теории вероятностей
Ω	Универсальное множество, множество всех ω	Пространство элементарных событий (элементарных исходов эксперимента), достоверное событие
ω	Элемент множества Ω	Элементарное событие
A	Некоторое множество элементов ω	Событие *A* (если $\omega \in A$, то говорят, что наступило событие *A*)
$\varnothing$	Пустое множество	Невозможное событие
$\bar{A}$	Дополнение множества *A*	Событие, противоположное событию *A*
$A \subset B$	*A* – подмножество *B*	Из наступления события *A* необходимо следует событие *B*
$A \bigcup B$	Объединение множеств *A* и *B*; множество точек, входящих хотя бы в одно из множеств *A* или *B*	Событие, состоящее в том, что произошло или *A*, или *B*, или оба вместе
$A \bigcap B$	Пересечение множеств *A* и *B*; множество точек, входящих и в *A*, и в *B*	Событие, состоящее в том, что произошло и *A*, и *B*
$A \bigcap B = \varnothing$	*A* и *B* – непересекающиеся множества	*A* и *B* – несовместные события
$A \setminus B$	Разность множеств *A* и *B*	Событие, состоящее в том, что произойдет *A*, но не произойдет *B*
$\overline{\lim_{n\to\infty}} A_n$	Множество всех тех ω, которые принадлежат бесконечному числу множеств из последовательности $\{A_n\}$	Событие, состоящее в том, что произойдет бесконечное число событий из последовательности $\{A_n\}$
$\underline{\lim_{n\to\infty}} A_n$	Множество всех тех ω, которые принадлежат всем A_n, за исключением конечного числа (множество всех тех ω, которые не принадлежат только конечному числу множеств A_n)	Событие, состоящее в том, что произойдут все события A_n, за исключением конечного числа (событие, состоящее в том, что не произойдет только конечное число из событий последовательности $\{A_n\}$)
$\lim A_n$	Если $\overline{\lim_{n\to\infty}} A_n = \underline{\lim_{n\to\infty}} A_n$, то последовательность множеств $\{A_n\}$ имеет предел $\lim_{n\to\infty} A_n = \overline{\lim_{n\to\infty}} A_n = \underline{\lim_{n\to\infty}} A_n$	Если $\overline{\lim_{n\to\infty}} A_n = \underline{\lim_{n\to\infty}} A_n$, то последовательность событий $\{A_n\}$ имеет предел $\lim_{n\to\infty} A_n = \overline{\lim_{n\to\infty}} A_n = \underline{\lim_{n\to\infty}} A_n$

§ 2. Аксиоматика Колмогорова

В этом параграфе будет построена математическая модель произвольного случайного явления, т. е. аксиоматика теории вероятностей, предложенная в 1933 г. А. Н. Колмогоровым.

Определение 1. *Вероятностным пространством называется тройка* $(\Omega, \mathcal{A}, \mathrm{P})$, *где* Ω *– пространство элементарных событий,* $\mathcal{A}$ *– некоторая* σ*-алгебра подмножеств из* Ω, *функция* P: $\mathcal{A} \to R$ *и обладает следующими свойствами:*

1°. $\mathrm{P}(A) \geq 0$ *для всех* $A \in \mathcal{A}$ *(неотрицательность* P*);*

2°. $\mathrm{P}(\Omega) = 1$ *(нормированность* P*);*

3°. $\mathrm{P}(A+B) = \mathrm{P}(A) + \mathrm{P}(B)$ *для любых* $A, B \in \mathcal{A}$, $A \bigcap B = \varnothing$ *(аддитивность* P*);*

4°. *Для любых событий* $B_1, B_2, \ldots, B_n, \ldots \in \mathcal{A}$, *таких, что* $B_n \downarrow \varnothing$, *т. е.* $B_1 \supset B_2 \supset \ldots \supset B_n \supset \ldots$ *и*

$$\bigcap_{n=1}^{\infty} B_n = \varnothing,$$

выполняется

$$\mathrm{P}(B_n) \xrightarrow[n \to \infty]{} 0 \text{ (непрерывность } \mathrm{P}\text{)}.$$

Функция P *называется вероятностью.*

Замечание 1. Вероятностное пространство является математической моделью произвольного случайного явления. Действительно, учитывая, что исходы такого явления случайны, для его описания необходимо рассматривать как все исходы (это заложено в Ω), так и вероятности, с которыми они происходят (за это отвечает P). В простейших случаях, т. е. в случаях когда число элементов не более чем счетное, для полного описания случайного явления можно было бы ограничиться двойкой (Ω, P) (см. гл. 1, § 3). В общем случае, т. е. когда число элементов несчетное, в Ω могут найтись подмножества, для которых вероятность P определить нельзя. Причина этого заключается в существовании так называемых неизмеримых множеств, что в свою очередь кроется в топологической структуре классических пространств (прямой, плоскости и т. д.). Поэтому приходится отказаться от того, чтобы назвать событием любое подмножество Ω. Следует выделить среди всех подмножеств класс $\mathscr{A}$ измеримых подмножеств. Из курса теории меры (см., напр., [1], [20], [29]) известно, что $\mathscr{A}$ является σ-алгеброй. Элементы $\mathscr{A}$ называются *случайными событиями*. Отметим также, что с точки зрения теории меры, функция P – конечная мера, а $(\Omega, \mathscr{A}, \mathrm{P})$ – пространство с мерой.

Покажем, что P – это σ-аддитивная мера. Для этого введем *аксиому σ-аддитивности* P:

3^*. Пусть $A_1,\ A_2,\ ...,\ A_n,\ ... \in \mathscr{A}$ и $A_i \bigcap A_j = \varnothing$ для $i \neq j$. Тогда

$$\mathrm{P}\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathrm{P}(A_n).$$

Утверждение 1. *Система аксиом* 1°– 4° *эквивалентна системе аксиом* 1°, 2°, 3^*.

Доказательство. Рассмотрим сначала необходимость. Пусть выполняются аксиомы 1°– 4°. Докажем справедливость аксиомы 3^*. Несложно видеть, что

$$\sum_{k=1}^{\infty} A_k = \left(\sum_{k=1}^{n-1} A_k\right) + B_n,$$

где

$$B_n = \sum_{k=n}^{\infty} A_k,$$

поэтому в силу справедливости аксиомы 3°

$$\mathrm{P}\left(\sum_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{n-1} \mathrm{P}(A_k) + \mathrm{P}(B_n).$$

Переходя к пределу при $n \to \infty$ в последнем равенстве и учитывая аксиому 4°, получаем

$$\mathrm{P}\left(\sum_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathrm{P}(A_k).$$

Докажем теперь достаточность. Покажем, что из системы аксиом 1°, 2°, 3^* вытекает аксиома 4°. Пусть случайные события $B_1, B_2, ...,$ $B_n, ... \in \mathscr{A}$ таковы, что $B_1 \supset B_2 \supset ... \supset B_n \supset ...$ и

$$\bigcap_{n=1}^{\infty} B_n = \varnothing.$$

Введем в рассмотрение случайные события $A_1, A_2, \ldots, A_n, \ldots$, положив по определению $A_1 = B_1 \setminus B_2,\ A_2 = B_2 \setminus B_3, \ldots,\ A_n = B_n \setminus B_{n+1}, \ldots$, тогда

$$B_1 = \sum_{k=1}^{\infty} A_k\ ,\ B_n = \sum_{k=n}^{\infty} A_k\ .$$

В силу справедливости аксиомы 3^* имеем

$$\mathrm{P}(B_1) = \sum_{k=1}^{\infty} \mathrm{P}(A_k),\ \mathrm{P}(B_n) = \sum_{k=n}^{\infty} \mathrm{P}(A_k),$$

поэтому

$$\mathrm{P}(B_n) \underset{n\to\infty}{\to} 0,$$

как остаток сходящегося ряда.

З а м е ч а н и е 2. Происхождение аксиом 1°, 2°, 3° можно объяснить, исходя из свойства статистической устойчивости частот (см. также § 3 из дополнений к гл. 1). Пусть A и B – несовместные события, $N(A)/N$ и $N(B)/N$ – их относительные частоты в какой-либо длинной серии из N наблюдений. Так как $N(A) \geq 0$, то $N(A)/N \geq 0$, следовательно, число $\mathrm{P}(A)$, к которому близко отношение $N(A)/N$, должно быть неотрицательным. Для достоверного события $N(\Omega) = N$, поэтому надо потребовать, чтобы $\mathrm{P}(\Omega) = 1$. Для несовместных событий $N(A+B) = N(A) + N(B)$, откуда

$$\frac{N(A+B)}{N} = \frac{N(A)}{N} + \frac{N(B)}{N},$$

что и приводит к аксиоме 3°.

Аксиома 4° (или 3^*) имеет несколько иное происхождение, связанное не с реальным свойством устойчивости частот, а с требованиями развиваемой на основе аксиоматики математической теории. Поясним сказанное на примере. Пусть на единичный квадрат случайно бросается частица, причем вероятность ее попадания в любой внутренний квадрат со сторонами, параллельными сторонам основного квадрата, равна площади внутреннего квадрата. С помощью аксиомы 3° можно получить вероятность попадания частицы в любую фигуру, составленную из объединения конечного числа квадратов. Если мы хотим иметь также возможность находить вероятности попадания частицы в более сложные фигуры, например, в круг, то это можно сделать с помощью аксиомы 3^*, приближая круг фигурами, составленными из конечных сумм таких квадратов.

Свойства вероятностей

1. *Если* $A \subset B$, *то* $\mathrm{P}(A) \le \mathrm{P}(B)$.

Доказательство. Так как $B = A + (B \setminus A)$, то из аксиомы 3^* получаем $\mathrm{P}(B) = \mathrm{P}(A) + \mathrm{P}(B \setminus A)$, поэтому, учитывая, что $\mathrm{P}(B \setminus A) \ge 0$, получаем требуемое.

2. *Если* $A \subset B$, *то* $\mathrm{P}(B \setminus A) = \mathrm{P}(B) - \mathrm{P}(A)$.

Доказательство вытекает из свойства 1.

3. *Для любого случайного события A справедливо* $\mathrm{P}(\overline{A}) = 1 - \mathrm{P}(A)$.

Доказательство следует из свойства 2, если $B = \Omega$, и аксиомы 2°.

4. $\mathrm{P}(\varnothing) = 0$.

Доказательство вытекает из аксиомы 2° и свойства 3 при $A = \Omega$.

Замечание 3. Иногда ошибочно считают, что событие нулевой вероятности есть невозможное событие. Это не так. Например, пусть выбирается наугад число из отрезка $[0;1]$. Положим $\Omega = [0;1]$, $\mathscr{A}$ – борелевская σ-алгебра на $[0;1]$ (т. е. наименьшая σ-алгебра, содержащая все интервалы $(a;b) \subset [0;1]$), P – мера Лебега на $[0;1]$, а M – множество рациональных точек отрезка $[0;1]$. Тогда вероятность того, что событие M произойдет, равна нулю. Однако это событие может произойти, т. е. оно не является невозможным. Отметим также, что событие, вероятность которого равна единице, не обязательно является достоверным событием.

5. *Для каждого случайного события A выполняется* $0 \le \mathrm{P}(A) \le 1$.

Доказательство следует из включений $\varnothing \subset A \subset \Omega$, свойств 1 и 4 и аксиомы 2°.

6. *Пусть A и B – случайные события. Тогда*

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B).$$

Доказательство. Очевидно, что $A \cup B = A + (B \setminus (A \cap B))$, а поэтому в силу аксиомы 3° и 2-го свойства вероятностей получаем требуемое.

7. *Пусть* A_1, ..., A_n *– случайные события. Тогда*

$$P\left(\bigcup_{k=1}^{n} A_k\right)=\sum_{i=1}^{n} P(A_i)-\sum_{1\le i_1<i_2\le n} P\left(A_{i_1}\cap A_{i_2}\right)+\sum_{1\le i_1<i_2<i_3\le n} P\left(A_{i_1}\cap A_{i_2}\cap A_{i_3}\right)-\ldots+$$

$$+(-1)^{n-1} P\left(A_1\cap\ldots\cap A_n\right).$$

Д о к а з а т е л ь с т в о вытекает из свойства 6 после применения метода математической индукции.

П р и м е р 1. На отдельных карточках написаны числа 1, 2, ..., n. Карточки разложены в ряд случайным образом. Какова вероятность того, что хотя бы одно из чисел окажется на месте с таким же номером?

Пусть A_i – событие, состоящее в том, что карточка с числом i окажется на месте с номером i. Надо вычислить $P\left(\bigcup_{i=1}^{n} A_i\right)$. Имеем

$$P(A_i)=\frac{(n-1)!}{n!}=\frac{1}{n},\quad P\left(A_{i_1}\cap A_{i_2}\right)=\frac{(n-2)!}{n!},\quad \ldots,\quad P\left(A_{i_1}\cap\ldots\cap A_{i_k}\right)=\frac{(n-k)!}{n!}.$$

Используя 7-е свойство вероятностей, получим

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\frac{C_n^1(n-1)!}{n!}-\frac{C_n^2(n-2)!}{n!}+\frac{C_n^3(n-3)!}{n!}-\ldots+(-1)^{n-1}C_n^n\frac{1}{n!}=$$

$$=1-\frac{1}{2!}+\frac{1}{3!}-\ldots+(-1)^{n-1}\frac{1}{n!}.$$

Заметим, что при $n\to\infty$ искомая вероятность стремится к числу $1-e^{-1}\approx 0,63$.

8. *Для любого конечного или счетного числа случайных событий* $\{A_n\}$ *имеют место неравенства*

$$P\left(\bigcup_n A_n\right)\le\sum_n P(A_n), \tag{1}$$

$$P\left(\bigcap_n A_n\right)\ge 1-\sum_n P\left(\overline{A}_n\right). \tag{2}$$

Для д о к а з а т е л ь с т в а неравенства (1) воспользуемся тем, что

$$\bigcup_n A_n=\sum_n B_n,$$

где $B_1=A_1,\ B_2=\overline{A}_1\cap A_2,\ \ldots,\ B_n=\overline{A}_1\cap\overline{A}_2\cap\ldots\cap\overline{A}_{n-1}\cap A_n,\ \ldots$. События B_n попарно несовместны и $B_n\subset A_n$, поэтому

$$P\left(\bigcup_n A_n\right)=P\left(\sum_n B_n\right)=\sum_n P(B_n)\le\sum_n P(A_n)$$

Неравенство (2) вытекает из следующих рассуждений:

$$P\left(\bigcap_n A_n\right)=1-P\left(\bigcup_n \overline{A}_n\right)\ge 1-\sum_n P\left(\overline{A}_n\right).$$

9. *Вероятность* P *непрерывная функция на σ-алгебре* $\mathscr{A}$, *т. е. если*

$$A_n \underset{n\to\infty}{\longrightarrow} A,$$

то

$$\mathrm{P}(A_n) \underset{n\to\infty}{\longrightarrow} \mathrm{P}(A).$$

Доказательство. Несложно видеть, что для $n = 1, 2, \ldots$ справедливы включения

$$\bigcap_{m=n}^{\infty} A_m \subset A_n \subset \bigcup_{m=n}^{\infty} A_m,$$

а поэтому в силу свойства 1 вероятностей

$$\mathrm{P}\left(\bigcap_{m=n}^{\infty} A_m\right) \le \mathrm{P}(A_n) \le \mathrm{P}\left(\bigcup_{m=n}^{\infty} A_m\right).$$

В последнем соотношении перейдем к пределам при $n \to \infty$ и воспользуемся результатами упражнений 4 и 5 (гл. 1, § 1)

$$\lim_{n\to\infty} \mathrm{P}\left(\bigcap_{m=n}^{\infty} A_m\right) = \mathrm{P}\left(\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m\right) \le \underline{\lim}_{n\to\infty} \mathrm{P}(A_n) \le$$

$$\le \overline{\lim_{n\to\infty}} \mathrm{P}(A_n) \le \lim_{n\to\infty} \mathrm{P}\left(\bigcup_{m=n}^{\infty} A_m\right) = \mathrm{P}\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m\right).$$

Из того, что $A_n \underset{n\to\infty}{\longrightarrow} A$, т. е.

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

приходим к выводу, что

$$\lim_{n\to\infty} \mathrm{P}(A_n) = \mathrm{P}(A) = \mathrm{P}\left(\lim_{n\to\infty} A_n\right).$$

§ 3. Примеры вероятностных пространств

1. *Классическое вероятностное пространство* – математическая модель случайного явления, число исходов которого конечно и все они равновозможны (т. е. все элементарные события равновероятны). В этом случае тройка $(\Omega, \mathscr{A}, \mathrm{P})$ имеет следующий вид:

$$\Omega = \{\omega_1, \ \omega_2, \ldots, \ \omega_n\}, \ \mathscr{A} = \mathscr{P}(\Omega), \ \mathrm{P}(\omega_i) = \frac{1}{n}, \ i = 1, 2, \ldots, n.$$

Здесь и в дальнейшем $\mathscr{P}(\Omega)$ будет обозначать множество всех подмножеств множества Ω. На любом подмножестве из Ω определим вероятность P следующим образом:

$$\mathrm{P}(A)=\sum_{i:\,\omega_i\in A}\mathrm{P}(\omega_i)=\frac{1}{n}\sum_{i:\,\omega_i\in A}1=\frac{|A|}{|\Omega|},$$

где $|\Omega|$ – число всех элементарных событий, $|A|$.– число элементарных событий, принадлежащих A. Несложно убедиться, что функция P удовлетворяет всем аксиомам вероятности. Она носит название *классической вероятности*.

Замечание 1. Модель классического вероятностного пространства используется в тех случаях, когда элементарные события обладают свойством симметрии в том смысле, что все они находятся в одинаковом отношении к тем условиям, которые определяют характер испытания. Например, результат бросания игральной кости или монеты «симметричен» по отношению к выпадению того или иного числа очков на кости или той или иной стороны монеты. Таким же свойством обладают правильно организованная жеребьевка и тираж лотереи.

Для нахождении вероятностей по схеме классического определения широко используются комбинаторные методы (см. § 1 дополнений к гл. 1).

Пример 1. Построим вероятностную модель эксперимента, состоящего в однократном подбрасывании симметричной монеты. В этом случае $\Omega=\{\omega:\ \omega=\mathit{Г}\vee P\}$ или просто $\Omega=\{\mathit{Г},P\}$, где элементарные события $\mathit{Г}$ и P заключаются соответственно в выпадении герба и решки, $\mathscr{A}=\mathscr{P}(\Omega)=$ $=\{\varnothing,\ \Omega,\ \{\mathit{Г}\},\ \{P\}\}$, и $\mathrm{P}(\omega=\mathit{Г})=\mathrm{P}(\omega=P)=1/2$.

Пример 2. Пусть бросается симметричная игральная кость. Тогда $\Omega=$ $=\{\omega:\ \omega=i,\ i=1,\ 2,\ 3,\ 4,\ 5,\ 6\}$, $\mathscr{A}=\mathscr{P}(\Omega)$ (σ-алгебра $\mathscr{A}$ состоит из 64 случайных событий), $\mathrm{P}(\omega=i)=1/6$.

Рассмотрим событие $B=\{$выпавшее число очков кратно трем$\}$, т. е. $B=$ $=\{3,\ 6\}$. В этом случае $\mathrm{P}(B)=\mathrm{P}(\omega=3)+\mathrm{P}(\omega=6)=1/3$.

2. *Конечное вероятностное пространство* – математическая модель случайного явления с конечным числом, вообще говоря, не равновозможных исходов (т. е. вероятности элементарных событий могут быть различными). В этом случае тройка $(\Omega,\ \mathscr{A},\ \mathrm{P})$ имеет следующий вид:

$$\Omega=\{\omega_1,\ \omega_2,\ ...,\ \omega_n\},\ \mathscr{A}=\mathscr{P}(\Omega),$$

$$\mathrm{P}(\omega_i)=p_i,\ \ p_i\ge 0,\ \ i=1,\ 2,\ ...,\ n,\ \ \sum\nolimits_{i=1}^{n}p_i=1.$$

Для любого подмножества A из Ω вероятность P определяется так:

$$\mathrm{P}(A)=\sum_{i:\,\omega_i\in A}p_i.$$

Несложно убедиться, что P удовлетворяет всем аксиомам вероятности.

П р и м е р 3. Бросается симметричная игральная кость. В качестве пространства элементарных событий возьмем множество

$$\Omega=\{\omega_1=\{1,\ 2,\ 3\},\ \omega_2=\{4,\ 5\},\ \omega_3=\{6\}\},\ \mathscr{A}=\mathscr{P}(\Omega)$$

(σ-алгебра $\mathscr{A}$ состоит из 8 случайных событий), $\mathrm{P}(\omega_1)=\dfrac{1}{2}$, $\mathrm{P}(\omega_2)=\dfrac{1}{3}$, $\mathrm{P}(\omega_3)=\dfrac{1}{6}$.

В рамках данной модели случайное событие B из прим. 2 не является элементом σ-алгебры, а, значит, для него вероятность не определена.

3. *Дискретное вероятностное пространство* – математическая модель случайного явления, число исходов которого не более чем счетное, поэтому

$$\Omega=\{\omega_1,\ \omega_2,\ \ldots,\ \omega_n,\ \ldots\},\ \mathscr{A}=\mathscr{P}(\Omega),$$

$$\mathrm{P}(\omega_i)=p_i,\ \ p_i\ge 0,\ \ i=1,\ 2,\ \ldots,\ \sum\nolimits_{i=1}^{\infty}p_i=1.$$

Для любого события A вероятность определим как сумму вероятностей всех элементарных событий, содержащихся в A:

$$\mathrm{P}(A)=\sum_{k:\ \omega_k\in A}p_k.$$

З а м е ч а н и е 2. В рассмотренных примерах вероятностных пространств (см. прим. 1 – 3) в качестве σ-алгебр случайных событий брались множества всех подмножеств пространства элементарных событий, т. е. фактически, под вероятностными пространствами понимались совокупности $(\Omega,\ \mathrm{P})$.

П р и м е р 4. Двое по очереди бросают одну и ту же симметричную монету. Выигрывает тот, у кого раньше выпадет герб. Найдем вероятность того, что выиграет бросающий первым.

Пространство элементарных событий рассматриваемого эксперимента имеет следующий вид:

$$\Omega=\{\omega_1,\ \omega_2,\ \ldots,\ \omega_n,\ \ldots\},$$

где событие ω_n состоит в том, что впервые герб выпал при n-м подбрасывании,

$$\mathrm{P}(\omega_n)=\frac{1}{2^n}.$$

Событие A, состоящее в выигрыше первого игрока, имеет следующий вид:

$$A=\bigcup_{k=1}^{\infty}\omega_{2k-1},$$

поэтому

$$\mathrm{P}(A)=\sum_{k=1}^{\infty}\frac{1}{2^{2k-1}}=\frac{0,5}{1-0,25}=\frac{2}{3}.$$

Таким образом, у первого игрока шансы выиграть больше в 2 раза.

4. *Геометрическое вероятностное пространство.* Теперь рассмотрим непрерывную вероятностную схему, т. е. случай, когда пространство элементарных событий представляет собой область *n*-мерного евклидова пространства (отрезок, многоугольник, многогранник, шар и т. п.) с конечным *n*-мерным объемом. Под *n-мерным объемом* будем понимать соответствующую меру Лебега. Естественно обобщить принцип равновозможности элементарных исходов классического вероятностного пространства и на эту схему. Однако в непрерывном случае число элементарных событий несчетное, и поэтому невозможно приписать каждому элементарному событию иной вероятности, кроме нуля. К построению математической модели (вероятностного пространства) в этом случае подойдем по-другому (сравните с построением общего вероятностного пространства). *Геометрическим вероятностным пространством* назовем тройку $(\Omega, \mathcal{A}, \mathrm{P})$, где $\Omega \subset R^n$ – измеримое по Лебегу множество конечной меры, $\mathcal{A}$ – некоторая σ-алгебра подмножеств Ω (σ-алгебра всех измеримых по Лебегу подмножеств Ω). В качестве $\mathcal{A}$ можно выбрать также борелевскую σ-алгебру, т. е. σ-алгебру, порожденную всеми открытыми подмножествами Ω (см. дополнения к гл. 1, § 4). Функцию P зададим на элементах $A \in \mathcal{A}$ следующим образом:

$$\mathrm{P} : \mathcal{A} \to R,$$

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

где λ – мера Лебега.

Определенная таким образом вероятность называется *геометрической.*

З а д а ч а о в с т р е ч е. Двое договорились встретиться в определенном месте между 19 и 20 часами. Пришедший первым ждет второго в течение 20 мин., затем уходит. Какова вероятность того, что встреча состоится?

Р е ш е н и е. Введем независимые переменные, зная значения которых мы можем ответить на вопрос задачи «да» или «нет»: x – момент прихода первого, y – момент прихода второго. Для того, чтобы встреча состоялась, необходимо и достаточно, чтобы $|x-y| \le 20$. Тогда

$$\Omega = \{\omega :\ \omega = (x, y),\ 0 \le x, y \le 60\},$$

$$A = \{\omega :\ \omega = (x, y),\ 0 \le x, y \le 60,\ |x-y| \le 20\}.$$

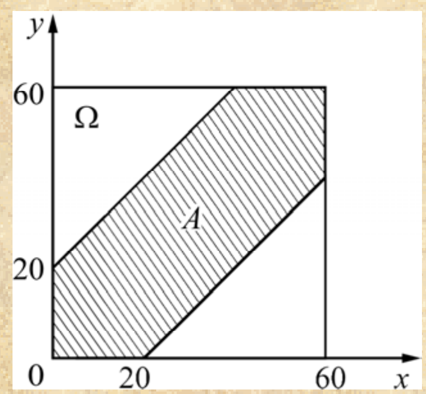

Рис. 2

Все возможные исходы на изображены на рис. 2 точками квадрата со стороной 60, исходы, при которых встреча состоится, расположены в заштрихованном шестиугольнике A. Несложный подсчет показывает, что вероятность в этом случае равна $5/9$.

П а р а д о к с Б е р т р а н а. Геометрическое определение вероятности часто подвергалось критике за некорректность определения вероятности события. В качестве наиболее яркого примера можно привести задачу французского математика, жившего в XIX в., Жозефа Бертрана.

З а д а ч а. Наудачу берется хорда в круге. Чему равна вероятность того, что ее длина превосходит длину стороны вписанного равностороннего треугольника?

Р е ш е н и е 1. Из соображений симметрии можно заранее задать направление хорды. Проведем диаметр, перпендикулярный хорде. Получим, что только хорды, которые пересекают диаметр в промежутке от $1/4$ до $3/4$ его длины, будут превосходить сторону правильного треугольника. Значит, вероятность равна $1/2$.

Р е ш е н и е 2. Не ограничивая общности, можно закрепить один из концов хорды на окружности. Касательная к окружности в этой точке и две стороны правильного треугольника с вершиной в этой точке образуют три угла по 60°. Условию задачи удовлетворяют только хорды, попадающие в средний угол. Таким образом, при этом способе вычисления, искомая вероятность оказывается равной $1/3$.

Р е ш е н и е 3. Для определения положения хорды, достаточно задать ее середину. Хорда удовлетворяет условию задачи тогда и только тогда, когда ее середина находится внутри круга, концентрического данному, но половинного радиуса. Площадь этого круга равна $1/4$ площади данного. Таким образом, искомая вероятность равна $1/4$.

Дело в том, что в условии задачи не определено проведение хорды наудачу. В самом деле, в решении 1 вдоль одного из диаметров AB окружности катится стержень (рис. 3, a). Множество всех возможных мест остановки этого стержня есть множество точек отрезка AB длины, равной диаметру. Равновероятными считаются события, состоящие в том, что остановка произойдет в интервале длины Δh, где бы внутри диаметра ни располагался этот отрезок. В решении 2 стержень, закрепленный на шарнире, расположенном в одной из точек окружности, совершает колебания величиной 180° (рис. 3, *б*). При этом предполагается, что остановка стержня внутри дуги окружности длины h зависит только от длины дуги, но не от ее положения, т. е. равновероятными событиями считаются остановки стержня в любых дугах окружности одинаковой длины. В решении 3 (рис. 3, *в*) точка бросается внутрь круга радиуса R наудачу, и вычисляется вероятность попадания ее внутрь концентрического круга радиуса $R/2$. Таким образом, разные ответы объясняются различными постановками задач.

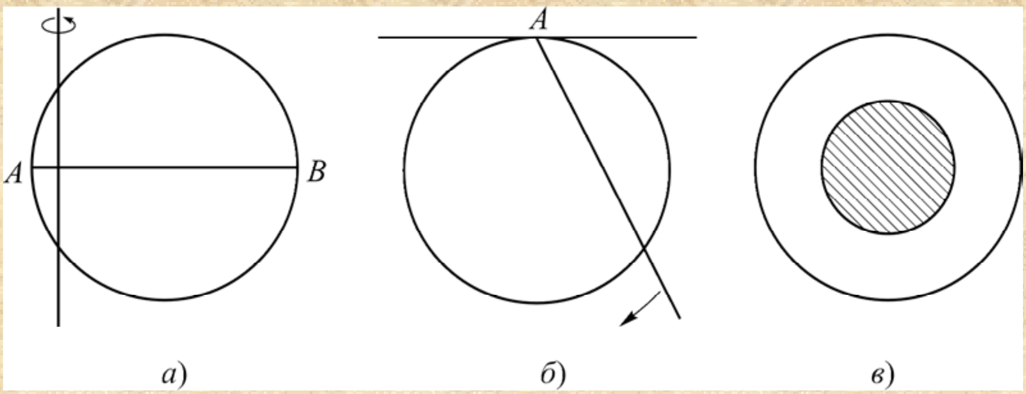

Рис. 3

ДОПОЛНЕНИЯ К ГЛАВЕ 1

§ 1. Основные понятия комбинаторики

Комбинаторика – это раздел математики, в котором изучаются расположения объектов в соответствии со специальными правилами и методы подсчета числа всех возможных способов, которыми эти расположения могут быть сделаны.

В этом параграфе изложены основные понятия комбинаторики.

Основной принцип комбинаторики (правило умножения). *Пусть требуется выполнить одно за другим k действий. Если первое действие можно выполнить n_1 способами, второе – n_2, и так до k-го действия, которое можно выполнить n_k способами, то все k действий вместе могут быть выполнены*

$$n_1 \times n_2 \times \cdots \times n_k$$

способами.

Это утверждение может быть легко доказано методом математической индукции, проведите его самостоятельно.

Пример 1. Из города *A* в город *B* ведет m различных путей, а из города *B* в город *C* – n путей. Несложно видеть, что число различных путей из города *A* в город *C* равно $m \times n$.

Упорядоченные множества, перестановки. *Множество называется упорядоченным, если каждому элементу этого множества поставлено в соответствие некоторое число (номер элемента) от 1 до n, где n – число элементов множества, так, что различным элементам соответствуют различные числа.*

Упорядоченные множества считаются различными, если они отличаются либо своими элементами, либо их порядком. Различные упорядоченные множества, которые отличаются лишь порядком элементов (т. е. могут быть получены из того же самого множества) называются *перестановками* этого множества.

Пусть P_n – число перестановок множества, содержащего n различных элементов. Тогда

$$P_n = 1 \times 2 \times 3 \times \cdots \times n = n!.$$

Действительно, если последовательно выбирать элементы исходного множества и размещать их в определенном порядке на n местах, то на первое место можно поставить любой из n элементов. После того, как заполнено первое место, на второе место можно поставить любой из оставшихся $n-1$ элементов и т. д. По правилу умножения получаем требуемое.

Размещения из n по k. Рассмотрим теперь упорядоченные подмножества основного множества Ω. Упорядоченные k-элементные подмножества (содержащие k элементов) множества из n элементов называются *размещениями из n элементов по k.*

Число размещений из n элементов по k равно

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

В самом деле, на первое место упорядоченного k-элементного подмножества можно поставить любой из n элементов основного множества, на второе – любой из $n-1$ оставшихся и т. д. На k-е место можно поставить любой из $n-k+1$ оставшихся элементов основного множества, и по правилу умножения получаем требуемое.

Пример 2. Найдем число телефонных номеров, состоящих из шести различных цифр.

$$A_{10}^6 = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 = 151\ 200.$$

Сочетания из n по k. Пусть Ω – основное множество, состоящее из n элементов. Произвольное (неупорядоченное) k-элементное подмножество Ω называется *сочетанием* из n элементов по k.

Число всех сочетаний из n элементов по k равно

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

Докажите это самостоятельно.

Пример 3. Рассмотрим прямоугольную сетку размера $m \times n$, состоящую из квадратов размера 1×1. Найдем число кратчайших путей на этой сетке, ведущих из левого нижнего угла – точки $(0;0)$ в правый верхний угол – точку $(m;n)$.

Каждый кратчайший путь из точки $(0;0)$ в точку $(m;n)$ состоит из $m+n$ отрезков, среди которых m горизонтальных и n вертикальных. Разные пути отличаются лишь порядком чередования горизонтальных и вертикальных отрезков. Поэтому общее число путей равно числу способов, которыми из $m+n$ отрезков можно выбрать m горизонтальных отрезков, т. е. C_{m+n}^m. Можно было бы рассматривать число способов выбора не m горизонтальных, а n вертикальных отрезков, и тогда получили бы ответ C_{m+n}^n. Итак, геометрически установлено равенство

$$C_{m+n}^m = C_{m+n}^n,$$

в справедливости которого нетрудно убедиться, используя представление C_n^k.

Разбиения на группы; перестановки с повторениями. Пусть Ω – множество из n элементов. Поставим следующий вопрос: каким числом способов множество Ω можно представить в виде объединения m попарно непересекающихся множеств $B_1, B_2, ..., B_m$, содержащих соответственно $k_1, ..., k_m$ элементов $k_1 + ... + k_m = n$?

Для ответа на этот вопрос поступим следующим образом: возьмем произвольное k_1-элементное подмножество B_1 множества Ω (это можно сделать $C_n^{k_1}$ способами), среди $n-k_1$ оставшихся элементов возьмем k_2-элементное подмножество B_2 (это можно сделать $C_{n-k_1}^{k_2}$ способами) и т. д. Общее число способов выбора $B_1, B_2, ..., B_m$ по правилу умножения равно

$$C_n^{k_1}\ C_{n-k_1}^{k_2}\ C_{n-k_1-k_2}^{k_3}\ \dots\ C_{n-k_1-\dots-k_{m-1}}^{k_m}=\frac{n!}{k_1!(n-k_1)!}\cdot\frac{(n-k_1)!}{k_2!(n-k_1-k_2)!}\cdot\frac{(n-k_1-k_2)!}{k_3!(n-k_1-k_2-k_3)!}\times$$

$$\times\ \dots\ \frac{(n-k_1-\dots-k_{m-1})!}{k_m!(n-k_1-\dots-k_m)!}=\frac{n!}{k_1!k_2!\ \dots\ k_m!}.$$

Таким образом, доказано следующее утверждение.

Пусть $k_1,\ ...,\ k_m$ *– целые неотрицательные числа, причем* $k_1+\dots+k_m=n.$ *Число способов, которыми можно представить множество* Ω *из n элементов в виде суммы m множеств* $B_1,\ B_2,\ ...,\ B_m$*, число элементов которых составляет соответственно* $k_1,\ ...,\ k_m$*, равно*

$$P_n(k_1,\ k_2,\ ...,\ k_m)=\frac{n!}{k_1!k_2!\ \dots\ k_m!}.$$

Числа $P_n(k_1,\ k_2,\ ...,\ k_m)$ называются *полиномиальными коэффициентами*. Они имеют еще одну очень важную комбинаторную интерпретацию.

Пусть имеется n букв: k_1 букв a_1, k_2 букв a_2, …, k_m букв a_m, $k_1+\dots+$ $+k_m=n.$ Определим, сколько различных слов можно составить из этих букв. Занумеруем места, на которых стоят буквы, числами 1, …, n. Каждое слово определяется множествами B_1 (номера мест, где стоит буква a_1), B_2 (номера мест, где стоит буква a_2), …, B_m (номера мест, где стоит буква a_m). Следовательно, число различных слов равно числу способов, которыми можно представить множество $\Omega=\{1,\ 2,\ ...,\ n\}$ в виде суммы (объединения) множеств $B_1,\ B_2,\ \dots,\ B_m$, т. е. это число равно $P_n(k_1,\ k_2,\ ...,\ k_m)$.

Слова длины n, которые можно получить из k_1 букв a_1, k_2 букв a_2, …, k_m букв a_m, называются еще *перестановками с повторениями*. Приведенное выше рассуждение показывает, что справедлив следующий факт.

Число различных перестановок, которые можно составить из n элементов, среди которых имеется k_1 *элементов первого типа,* k_2 *элементов второго типа, ...,* k_m *элементов m-го типа, равно* $P_n(k_1,\ k_2,\ ...,\ k_m)$.

П р и м е р 4. Найдем число различных слов, которые можно получить, переставляя буквы в слове математика. Оно равно

$$P_{10}(3,\ 2,\ 2,\ 1,\ 1,\ 1)=\frac{10!}{3!\cdot 2!\cdot 2!}=151\ 200.$$

П о л и н о м и а л ь н а я ф о р м у л а. Рассмотрим задачу о том, как раскрывать скобки при вычислении выражения

$$(a_1+a_2+\dots+a_k)^n.$$

Справедливо утверждение

$$(a_1+a_2+\dots+a_k)^n=\sum_{r_1\ge 0,\ ...,\ r_k\ge 0,\ r_1+\dots+r_k=n}\frac{n!}{r_1!\ \dots\ r_k!}a_1^{r_1}\ \dots\ a_1^{r_k}.$$

Д о к а з а т е л ь с т в о. Перемножим последовательно n раз $(a_1 + a_2 + ... + a_k)$. Тогда получим k^n слагаемых вида $d_1\, d_2\, \ldots\, d_n$, где каждый множитель d_i, $i = \overline{1, n}$, равен или a_1, или a_2, …, или a_k. Обозначим через $B(r_1, ..., r_k)$ совокупность всех тех слагаемых, где a_1 встречается множителем r_1 раз, a_2 – r_2 раз, …, a_k – r_k раз. Число таких слагаемых равно $P_n(r_1, r_2, ..., r_k)$. Утверждение доказано.

С о ч е т а н и я с п о в т о р е н и я м и. *Сочетаниями из m элементов по n элементов с* *п о в т о р е н и я м и* *называются группы, содержащие n элементов, причем каждый элемент принадлежит одному из m типов.*

Например, из трех элементов *a*, *b*, *c* можно составить такие сочетания по два элемента с повторениями

$$aa,\ bb,\ cc,\ ab,\ ac,\ bc.$$

Убедимся в том, что справедливо следующее утверждение: *число различных сочетаний из m элементов по n с повторениями равно*

$$f_m^n = C_{m+n-1}^{m-1} = C_{m+n-1}^n.$$

Действительно, каждое сочетание полностью определяется, если указать сколько элементов каждого из *m* типов в него входит. Поставим в соответствие каждому сочетанию последовательность из нулей и единиц, составленную по следующему правилу: напишем подряд столько единиц, сколько элементов первого типа входит в сочетание, далее поставим нуль и после него напишем подряд столько единиц, сколько элементов второго типа содержит это испытание и т. д. Например, написанным выше сочетаниям из трех букв по две будут соответствовать такие последовательности 1100, 1001, 0101, 1010, 0110, 0011.

Таким образом, каждому сочетанию из *m* по *n* соответствует последовательность из *n* единиц и $m-1$ нулей и наоборот. Утверждение доказано.

П р и м е р 5. Найдем число способов, которыми можно выбрать три из двенадцати букв А, А, А, Т, Т, Т, Г, Г, Г, Ц, Ц, Ц. В рассматриваемом случае $n = 4$ (имеем четыре типа А, Т, Г, Ц), а $n = 3$. Поэтому искомое число равно

$$f_4^3 = C_{3+4-1}^3 = C_6^3 = 20.$$

П р и м е р 6. Кости домино можно рассматривать как сочетание с повторениями по два из семи чисел: 0, 1, 2, 3, 4, 5, 6. Число таких сочетаний равно

$$f_7^2 = C_{7+2-1}^2 = C_8^2 = 28.$$

П р и м е р 7. Найдем число целых неотрицательных решений уравнения

$$x_1 + x_2 + ... + x_m = n.$$

Существует взаимно однозначное соответствие между решениями указанного уравнения и сочетаниями из *m* элементов по *n*. Если имеем целые неотрицательные числа $x_1, x_2, ..., x_m$, такие, что $x_1 + x_2 + ... + x_m = n$, то можем составить сочетание из *m* элементов, взяв x_1 элементов первого типа, …, x_m элементов *m*-го типа. Наоборот, имея сочетание из *m* элементов по *n*, получим некоторое решение исходного уравнения в целых неотрицательных числах. Поэтому число целых неотрицательных решений равно

$$f_m^n = C_{m+n-1}^n.$$

П р и м е р 8. Найдем число различных частных производных порядка n у бесконечно дифференцируемой функции m переменных. Частные производные порядка n не зависят от порядка дифференцирования, а зависят лишь от того, сколько раз мы дифференцируем по каждой переменной. Поэтому число различных частных производных равно числу различных целых неотрицательных решений уравнения $x_1 + x_2 + ... + x_m = n$, т. е.

$$f_m^n = C_{m+n-1}^n.$$

У т в е р ж д е н и е 1. *Пусть $n > m$. Тогда число таких сочетаний из m по n с повторениями, в которых каждый предмет представлен хотя бы один раз, равно C_{n-1}^{m-1}.*

В самом деле, поставим в соответствие, как и раньше, каждому сочетанию из m по n набор из $m-1$ нулей и n единиц. Тогда тем сочетаниям, которые каждый из m элементов содержит хотя бы один раз, будут соответствовать наборы, в которых нет нулей, стоящих рядом (в таком наборе каждый нуль стоит между двумя единицами). Каждый такой набор заканчивается единицей. Все такие наборы можно указать, отметив те из $n-1$ единиц, справа от которых стоит нуль. Это можно сделать C_{n-1}^{m-1} способами.

§ 2. Классические теоретико-вероятностные модели

С л у ч а й н ы й   в ы б о р   с   в о з в р а щ е н и е м. Пусть $M = \{1,\ 2,\ \ldots,\ N\}$, $\omega = (i_1,\ i_2,\ \ldots,\ i_n)$ – упорядоченный набор из n элементов множества M. Вероятностную схему, в которой пространство $\Omega = \{\omega = (i_1,\ i_2,\ \ldots,\ i_n):\ i_k \in M,\ k = 1,\ 2, \ldots,\ n\}$ и все элементарные события ω равновероятны, называют *схемой случайного выбора с возвращением.* Эту схему можно интерпретировать следующим образом. Пусть в урне содержатся N одинаковых шаров, пронумерованных числами множества M. Из урны выбирается один шар, записывается его номер, шар возвращается обратно в урну, после чего шары в урне тщательно перемешивают. По этой схеме выбирают n шаров и записываются их номера: $i_1,\ i_2,\ \ldots,\ i_n$. Рассмотрим случай, когда в урне K белых шаров и $N-K$ черных, причем номера белых шаров $1,\ 2,\ \ldots,\ K$, а черных $K+1,\ \ldots,\ N$. Какова вероятность события A_k, заключающегося в том, что среди выбранных n шаров ровно k белых?

За элементарное событие примем любые подмножества из n элементов, выбранные из множества N шаров. Известно, что число таких подмножеств равно N^n. Каждый набор шаров, входящий в событие A_k, состоит из двух частей: k белых и $n-k$ черных шаров. Все такие наборы можно получить следующим образом: сначала выберем номера мест в наборе, где будут расположены белые шары. Это можно сделать C_n^k различными способами. Затем составим части наборов из белых шаров, их число равно K^k, и отдельно – части наборов из черных шаров, число таких частей $(N-K)^{n-k}$. Значит, число элементарных событий в A_k равно $C_n^k K^k (N-K)^{n-k}$. Используя классическое определение вероятности, получаем

$$\mathrm{P}\left(A_k\right)=\frac{C_n^k K^k\left(N-K\right)^{n-k}}{N^n}.$$

П р и м е р 1. С л у ч а й н ы е ч и с л а. Рассмотрим теперь один важный частный случай схемы выбора с возвращением. Пусть $M=\{0, 1, 2, \ldots, 9\}$, тогда пространством Ω для данной схемы выбора с возвращением является множество всех цифровых последовательностей $\omega=\left(i_1, i_2, \ldots, i_n\right)$ длины n. Значит, число элементов пространства Ω равно 10^n. Несложно видеть, что вероятность любой фиксированной последовательности ω равна 10^{-n}.

Случайными (равномерно распределенными) числами называют цифровую последовательность $\omega=\left(i_1, i_2, \ldots, i_n\right)$, полученную в результате проведения реального опыта, который хорошо описывается рассматриваемой схемой. Такого типа опыты проводят, например, при розыгрыше номеров в спортлото. Оценка близости реального опыта и его математической модели является одной из задач математической статистики. Последовательности случайных чисел можно использовать для получения реализаций случайных процессов. Например, если в последовательности случайных чисел $\omega=\left(i_1, i_2, \ldots, i_n\right)$ четные цифры заменить буквой Γ, а нечетные – буквой P, то полученную последовательность из букв Γ и P можно рассматривать как реализацию опыта, заключающегося в n-кратном подбрасывании симметричной монеты, где Γ – выпадение герба, P – выпадение решки. Действительно, каждая последовательность, состоящая из букв Γ и P, может быть получена из одинакового числа равновероятных цифровых последовательностей, а, значит, последовательности, состоящие из Γ и P, равновероятны. При расчетах методом Монте-Карло требуются длинные последовательности случайных чисел, которые в процессе вычислений можно непрерывно вводить в компьютер. Методы получения случайных чисел, основанные на реальном извлечении шаров из урны, для этой цели непригодны, поэтому для практических вычислений пользуются какими-либо алгоритмами получения последовательности псевдослучайных чисел.

С л у ч а й н ы й в ы б о р б е з в о з в р а щ е н и я. Пусть $M=\{1, 2, \ldots, N\}$, $\omega=\left(i_1, i_2, \ldots, i_n\right)$ – упорядоченный набор из n различных элементов множества M. Вероятностную схему, в которой пространство $\Omega=\left\{\omega=\left(i_1, i_2, \ldots, i_n\right): i_k \in M,\right.$ $\left.k=1, 2, \ldots, n, i_l \neq i_j, l \neq j\right\}$ и все элементарные события ω равновероятны, называют *схемой случайного выбора без возвращения.* Эту схему интерпретируют следующим образом. Пусть в урне находятся N пронумерованных шаров, причем шары с номерами $1, 2, \ldots, K$ белого цвета, а остальные $N-K$ черного. Выборка без возвращения состоит в том, что мы наудачу вынимаем из урны последовательно n шаров, не возвращая их обратно в урну. Пространство элементарных событий Ω в этом случае состоит из множества всех упорядоченных наборов $\omega=\left(i_1, i_2, \ldots, i_n\right)$, где числа $i_k \in M$. Здесь i_k – номер шара, который достали на k-м шаге. Мощность пространства Ω равна числу размещений из N элементов по

n, т. е. $|\Omega| = N(N-1)\ \ldots\ (N-n+1)$. Вычислим вероятность события S_k, состоящего в том, что среди выбранных n шаров ровно k белых. Несложно видеть, что

$$|S_k| = C_n^k \cdot \frac{K!}{(K-k)!} \cdot \frac{(N-K)!}{(N-K-n+k)!}.$$

Тогда

$$\mathrm{P}(S_k) = \frac{|S_k|}{|\Omega|} = \frac{C_n^k C_{N-n}^{K-k}}{C_N^K} = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n}.$$

Примеры классических теоретико-вероятностных моделей

Модель Максвелла – Больцмана

Имеется система, состоящая из r различимых (пронумерованных) шаров (частиц), каждый из которых может находиться в одном из n ящиков (состояний) вне зависимости от того, где при этом находятся остальные шары. В этой модели возможно всего n^r различных размещений r шаров по n ящикам. Если при этом все такие размещения (состояния системы) считаются равновероятными, то говорят *о статистике Максвелла – Больцмана.* Вероятность каждого состояния равна n^{-r}.

Модель Бозе – Эйнштейна

Рассмотрим совокупность r неразличимых шаров (частиц), каждый из которых независимо от остальных может находиться в одном из n ящиков (состояний). Поскольку шары неразличимы, каждое состояние такой системы задается упорядоченным набором $(r_1, r_2, \ldots, r_n)$, где r_k – число шаров в k-м ящике, $\sum_{k=1}^{n} r_k = r$. Теперь подсчитаем число различных состояний системы (т. е. число различных упорядоченных наборов).

Рис. 4

Состояние системы удобно представить так, как это показано на рис. 4., где изображена конфигурация из $r = 10$ точек (частиц) и $n-1=4$ отрезков (границ ящиков). Каждая такая конфигурация задает размещение неразличимых шаров по ящикам и наоборот, если задано состояние системы через упорядоченный набор n чисел, то ему соответствует одна конфигурация. Нарисованной конфигурации соответствует упорядоченный набор чисел $(4,\ 2,\ 0,\ 1,\ 3)$. Очевидно, что каждая конфигурация определяется положениями внутренних $n-1$ отрезков, которые могут находиться в $n+r-1$ позициях. Значит, имеется ровно C_{n+r-1}^{n-1} различных конфигураций и столько же различных состояний рассматриваемой системы. Если потребовать, чтобы все состояния системы были равновероятными, то получится вероятностная схема, которая называется *моделью Бозе – Эйнштейна.* При этом вероятность каждого состояния системы равна $\dfrac{1}{C_{n+r-1}^{n-1}}$.

Модель Ферми – Дирака

Данная модель определяется аналогично модели Бозе – Эйнштейна, в которой дополнительно действует принцип запрета, требующий, чтобы в каждой ячейке находилось не более одного шара (частицы). Поскольку и в этом случае шары неразличимы, то состояние системы характеризуется упорядоченным набором $(r_1,\ r_2,\ \ldots,\ r_n)$, где $r_k = 0,\ 1,$ при этом $r \le n$. Задать состояние такой системы можно, указав заполненные ячейки, а их можно выбрать C_n^r различными способами. Столько же состояний будет и у системы Ферми – Дирака. Если все состояния равновероятны, то говорят *о статистике Ферми – Дирака.* Вероятность каждого состояния в таком случае равна $\dfrac{1}{C_n^r}$.

§ 3. Частотное (статистическое) определение вероятности

Пусть проводят некоторый эксперимент (например, бросаю игральной кости), Ω – множество всех исходов данного эксперимента, A – некоторое случайное событие, N – количество экспериментов (количество подбрасываний игральной кости), $N(A)$ – количество экспериментов, в которых произошло событие A. Величина $\dfrac{N(A)}{N}$ называется *частотой появления события A.* Во многих случаях на практике частота $\dfrac{N(A)}{N}$ ведет себя таким образом, будто существует предел $\dfrac{N(A)}{N}$ при $N \to \infty$. Предположим, что этот предел действительно существует. Обозначим его $\mathrm{P}(A)$. Таким образом, $\mathrm{P}(A) = \lim\limits_{N\to\infty} \dfrac{N(A)}{N}$ и называется *частотной, или статистической, вероятностью события A.*

Упражнение 1. Докажите справедливость следующих свойств частотной вероятности P:

1) для любого события A

$$\mathrm{P}(A) \ge 0;$$

2) для достоверного события Ω

$$\mathrm{P}(\Omega) = 1;$$

3) если события $A_1,\ A_2,\ \ldots,\ A_n$ попарно несовместны, то

$$\mathrm{P}(A_1 + \ldots + A_n) = \mathrm{P}(A_1) + \ldots + \mathrm{P}(A_n).$$

В заключение этого параграфа остановимся на одной интересной и весьма распространенной, особенно среди естествоиспытателей, концепции вероятности, предложенной Р. Мизесом. Согласно Мизесу, если частота по мере увеличения числа экспериментов все меньше и меньше уклоняется от вероятности, то в пределе по определению и должна быть вероятность (при этом на последовательность испытаний дополнительно накладываются достаточно громоздкие условия). По мнению Мизеса, любое априорное определение вероятности обречено на неудачу, и лишь данное им эмпирическое – способно удовлетворить интересы естествознания, математики и философии.

В концепции Мизеса вероятность теряет характер объективной числовой характеристики некоторых реальных явлений. Действительно, до осуществления бесконечного числа испытаний нельзя даже говорить о вероятности того или иного события и, следовательно, использовать результаты теории вероятностей. Подробнее об этой теории можно узнать из книги Р. Мизеса «Вероятность и статистика».

Отметим, что впервые статистическое определение вероятности было дано Я. Бернулли.

§ 4. Борелевские сигма-алгебры

Пусть Ω – пространство элементарных событий, $\mathcal{A}$ – некоторое множество (система) событий.

О п р е д е л е н и е 1. *Сигма-алгеброй (σ-алгеброй) $\sigma(\mathcal{A})$, порожденной множеством событий $\mathcal{A}$, называется наименьшая σ-алгебра, содержащая $\mathcal{A}$, т. е.*

$$\sigma(\mathcal{A}) = \bigcap_{\alpha} \mathcal{B}_\alpha, \tag{1}$$

где $\mathcal{B}_\alpha$ – σ-алгебра, содержащая $\mathcal{A}$, а пересечение берется по всем таким σ-алгебрам.

У т в е р ж д е н и е 1. *Если события $A_1, \ldots, A_n \in \mathcal{A}$, то*

$$\bigcup_{i=1}^{n} A_i \in \sigma(\mathcal{A}), \bigcap_{i=1}^{n} A_i \in \sigma(\mathcal{A})$$

для $i = \overline{1,n}$.

Д о к а з а т е л ь с т в о. Если $A_1, \ldots, A_n \in \mathcal{A}$, то при любом α из соотношения (1) следует

$$\bigcup_{i=1}^{n} A_i \in B_\alpha, \bigcap_{i=1}^{n} A_i \in B_\alpha, A_i \in B_\alpha$$

для $i = \overline{1,n}$.

Отсюда

$$\bigcup_{i=1}^{n} A_i \in \bigcap_{\alpha} B_\alpha = \sigma(\mathcal{A}), \bigcap_{i=1}^{n} A_i \in \bigcap_{\alpha} B_\alpha = \sigma(\mathcal{A}), A_i \in \bigcap_{\alpha} B_\alpha = \sigma(\mathcal{A}).$$

О п р е д е л е н и е 2. *Борелевской σ-алгеброй на R называется σ-алгебра $\mathcal{B}(R)$, порожденная системой множеств $\mathcal{A} = \{(-\infty; x]: x \in R\}$, т. е.*

$$\mathcal{B}(R) = \sigma\{(-\infty; x]: x \in R\}.$$

Множества, которые принадлежат $\mathcal{B}(R)$, называются борелевскими мно-

жествами на прямой R.

Утверждение 2. *Для любых* a, $b \in R$, *множества* $(a;b]$, $[a;b]$, $[a;b)$, $(a;b)$, $\{a\}$ ($\{a\}$ – *множество, состоящее из точки* a) *являются борелевскими множествами.*

Доказательство вытекает из утверждения 1 и равенств

$$(a;b] = (-\infty;b] \setminus (-\infty;a], \ \{a\} = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}; a\right],$$

$$[a;b] = \{a\} \cup (a;b], \ [a;b) = [a;b] \setminus \{b\}, \ (a;b) = (a;b] \setminus \{b\}.$$

Утверждение 3. *Борелевская* σ-*алгебра* $\mathscr{B}(R)$ *совпадает со следующими* σ-*алгебрами*

$$\sigma\{(a;b]:\ a,\ b \in R\},\ \sigma\{[a;b]: a,\ b \in R\},\ \sigma\{[a;b):\ a,\ b \in R\},\ \sigma\{(a;b):\ a,\ b \in R\}.$$

Докажем, что $\mathscr{B}(R) = \sigma\{(a;b]:\ a,\ b \in R\}$. Для остальных случаев доказательства аналогичны.

Из утверждения 2 следует, что

$$\mathscr{B}(R) = \sigma\{(-\infty;x] \cup (a;b]:\ x,\ a,\ b \in R\}, \tag{2}$$

так как $\mathscr{B}(R)$ включает множества $(-\infty;x]$ и $(a;b]$.

Обозначим $\mathscr{B}_1(R) = \sigma\{(a;b]:\ a,\ b \in R\}$. Покажем, что $\mathscr{B}_1(R) = \mathscr{B}(R)$. Из утверждения 1 следует, что

$$(-\infty;x] = \bigcup_{n \geq x}^{\infty} (-n;x] \in \mathscr{B}_1(R).$$

Отсюда

$$\mathscr{B}_1(R) = \sigma\{(a;b] \cup (-\infty;x]:\ x,\ a,\ b \in R\}.$$

Поэтому из соотношений (1) и (2) получаем

$$\mathscr{B}_1(R) = \mathscr{B}(R).$$

Пусть $\Omega = R^n = \{(x_1,\ ...,\ x_n):\ x_i \in R,\ i = \overline{1,n}\}$, $B_i \in R$, $i = \overline{1,n}$. Обозначим $B_1 \times \times B_2 \times ... \times B_n = \{(x_1, ..., x_n):\ x_i \in B_i,\ i = \overline{1,n}\}$. Множество $B_1 \times B_2 \times ... \times B_n$ называется *декартовым произведением множеств* B_1, B_2, ..., B_n.

Определение 3. *Борелевской* σ-*алгеброй* $\mathscr{B}(R^n)$ *называется* σ-*алгебра, порожденная системой множеств*

$$\mathscr{A} = \{(-\infty;x_1] \times \ldots \times (-\infty;x_n]:\ (x_1,\ ...,\ x_n) \in R^n\},$$

т. е.

$$\mathscr{B}(R^n) = \sigma(\mathscr{A}) = \sigma((-\infty;x_1] \times \ldots \times (-\infty;x_n]:\ (x_1,\ ...,\ x_n) \in R^n).$$

Определение 4. *Множества*

$$\{(a_1;b_1] \times ... \times (a_n;b_n]:\ a_i,\ b_i \in R,\ a_i < b_i,\ i = \overline{1,n}\}$$

называются прямоугольниками (*параллелепипедами*) *в* R^n.

У т в е р ж д е н и е 4. *Борелевская σ-алгебра* $\mathscr{B}\left(R^n\right)$ *совпадает с σ-алгеброй, порожденной прямоугольниками в* R^n.

Докажите это утверждение самостоятельно.

У п р а ж н е н и е 1. Введем на числовой прямой R метрику

$$\rho_1(x,y)=\frac{|x-y|}{1+|x-y|},$$

эквивалентную обычной евклидовой метрике $d(x,y)=|x-y|$, и пусть

$$\mathscr{B}_0(R)=\sigma\left(\left\{x\in R:\ \rho_1(x,y)<\rho\right\},\ \rho>0,\ y\in R\right).$$

Докажите, что $\mathscr{B}_0(R)=\mathscr{B}(R)$.

У п р а ж н е н и е 2. Пусть $\mathscr{B}_0(R)$ – алгебра, порожденная открытыми множествами

$$\left\{x\in R^n:\ \rho_n(x,y)<\rho\right\},\ \rho>0,\ y\in R^n,$$

где

$$\rho_n(x,y)=\sum_{k=1}^{n}2^{-k+1}\rho_1(x_k,y_k),\ x=(x_1,\ ...,\ x_n),\ y=(y_1,\ ...,\ y_n).$$

Докажите, что $\mathscr{B}_0\left(R^n\right)=\mathscr{B}\left(R^n\right)$.

Задачи

1. Установите справедливость следующих свойств операций $\bigcup$ и $\bigcap$:

1) $A\bigcup B=B\bigcup A,\ A\bigcap B=B\bigcap A$ (*коммутативность*);
2) $A\bigcup(B\bigcup C)=(A\bigcup B)\bigcup C,\ A\bigcap(B\bigcap C)=(A\bigcap B)\bigcap C$ (*ассоциативность*);
3) $A\bigcap(B\bigcup C)=(A\bigcap B)\bigcup(A\bigcap C),\ A\bigcup(B\bigcap C)=(A\bigcup B)\bigcap(A\bigcup C)$ (*дистрибутивность*);
4) $A\bigcup A=A,\ A\bigcap A=A$ (*идемпотентность*).

Покажите также, что

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

2. Докажите, что

$$\mathrm{P}(A\triangle B)=\mathrm{P}(A)+\mathrm{P}(B)-2\mathrm{P}(A\bigcap B)=\mathrm{P}(A\bigcup B)-\mathrm{P}(A\bigcap B).$$

3. Пусть множество Ω состоит из N элементов. Покажите, что общее число $d(N)$ различных разбиений множества Ω определяется формулой

$$d(N)=e^{-1}\sum_{k=0}^{\infty}\frac{k^N}{k!}.$$

У к а з а н и е. Сначала убедитесь в том, что

$$d(N)=\sum_{k=0}^{N-1}C_{N-1}^{k}d(k),\ d(0)=1.$$

4. Пусть $A_1,\ ...,\ A_n$ – события и величины $S_0,\ S_1,\ ...,\ S_n$ определены следующим образом:

$$S_0 = 1,\ S_r = \sum_{I_r} \mathrm{P}\left(A_{k_1} \dots A_{k_r}\right),\ 1 \le r \le n,$$

где суммирование распространяется по неупорядоченным подмножествам $I_r = \{k_1, \ldots, k_r\}$ множества $\{1, \ldots, n\}$.

Пусть событие B_m состоит в том, что одновременно произойдет в точности m событий из $A_1, \ldots, A_n$. Покажите, что

$$\mathrm{P}(B_m) = \sum_{r=m}^{n} (-1)^{r-m} C_r^m S_r.$$

В частности, для $m = 0$

$$\mathrm{P}(B_0) = 1 - S_1 + S_2 - \ldots \pm S_n.$$

Покажите также, что вероятность того, что одновременно произойдут по крайней мере m событий из $A_1, \ldots, A_n$, равна

$$\mathrm{P}(B_1) + \ldots + \mathrm{P}(B_n) = \sum_{r=m}^{n} (-1)^{r-m} C_{r-1}^{m-1} S_r.$$

В частности, вероятность того, что произойдет по крайней мере одно из событий $A_1, \ldots, A_n$, равна

$$\mathrm{P}(B_1) + \ldots + \mathrm{P}(B_n) = S_1 - S_2 + S_3 - \ldots \pm S_n.$$

5. Используя вероятностные соображения, докажите справедливость следующих тождеств:

1) $\sum\limits_{k=0}^{n} C_n^k = 2^n;$

2) $\sum\limits_{k=0}^{n} \left(C_n^k\right)^2 = C_{2n}^n;$

3) $\sum\limits_{k=0}^{n} (-1)^{n-k} C_m^k = C_{m-1}^n,\ m \ge n-1;$

4) $\sum\limits_{k=0}^{m} k(k-1) C_m^k = m(m-1)2^{m-2},\ m \ge 2.$

6. З а д а ч а Б а н а х а. Некий математик носит с собой два коробка спичек, число которых в каждом из коробков равно N. Каждый раз, когда ему нужна спичка, он берет наудачу один из коробков. Когда-нибудь наступит такой момент, что вынутый коробок окажется пуст. Найдите вероятность того, что:

а) второй коробок содержит r спичек;

б) в момент, когда впервые один из коробков оказался пуст, в другом было r спичек;

в) отсутствие спичек будет впервые обнаружено не в том коробке, который опустел первым; $r = 0, 1, \ldots, N$.

7. З а д а ч а о б а л л о т и р о в к е. Пусть два кандидата R и S получили на выборах соответственно r и s, $r > s$, голосов. Какова вероятность того, что кандидат R в течение всех выборов по количеству полученных голосов был впереди S? При этом предполагается выполненной следующая несколько наивная процедура подсчета голосов: каждый избиратель отдает свой голос или кандидату R или кандидату S с вероятностью $\frac{1}{2}$, последовательно опрашиваются все избиратели и на каждом шаге подсчитывается разность голосов, поданных за R и за S.

8. З а д а ч а о п ь я н и ц е. Пьяница стоит на расстоянии одного шага от края утеса. Он шагает случайным образом либо к краю утеса, либо от него. На каждом шаге вероятность отойти от края равна p, а шаг к краю утеса имеет вероятность $q = 1 - p$. Чему равна вероятность того, что пьяница не избежит падения?

9. С е м е й н а я з а д а ч а. В семье четыре сестры, они моют посуду по очереди. Из четырех разбитых тарелок три разбиты младшей сестрой, и поэтому ее называют неуклюжей. Можно ли ее оправдать, приписывая эти неудачи случайности? Обсудите связь с размещением шаров по ящикам.

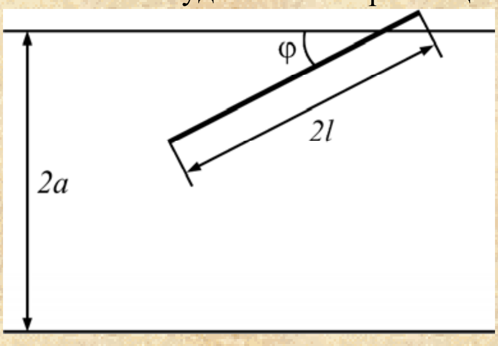

Рис. 5

10. З а д а ч а Б ю ф ф о н а. Плоскость разграфлена параллельными прямыми, расстояния между которыми $2a$ (рис. 5). На нее наудачу бросается игла длины $2l$, $l < a$. Найдите вероятность того, что игла пересечет какую-нибудь прямую. Слово «наудачу» подразумевает следующее: центр иглы случайно попадает на отрезок длины $2a$, перпендикулярный параллельным прямым, вероятность того, что угол φ, образованный иглой и проведенными прямыми, будет заключаться между углами α и β, пропорциональна $|\alpha - \beta|$, и положение центра иглы и величина φ независимы.

11. Т е о р е т и к о - в е р о я т н о с т н а я ф о р м у л и р о в к а в е л и к о й т е о р е м ы Ф е р м а. В двух урнах содержится одно и то же количество шаров, несколько черных и несколько белых в каждой. Из них вынимаются n $(n \geq 3)$ шаров с возвращением. Найдите число n и содержимое обеих урн, если вероятность того, что все белые шары взяты из первой урны, равна вероятности того, что из второй урны взяты либо все белые, либо все черные шары.

12. R различимых шаров раскладывают по n ящикам. Найдите вероятность того, что останется ровно m пустых ящиков.

13. R неразличимых шаров раскладывают по n ящикам. Определить вероятность того, что останется ровно m пустых ящиков.

14. Два игрока A и B играют в следующую игру. На первом шаге игрок A расставляет разные числа от 1 до 18 по своему усмотрению на шести гранях трех игральных костей. На втором шаге игрок B, внимательно изучив кости, выбирает одну. На третьем шаге A выбирает одну из оставшихся костей. После этого каждый бросает кость. Побеждает тот, на чьей кости выпало большее количество очков. Какому из игроков игра более выгодна?

Глава 2. НЕЗАВИСИМОСТЬ

В этой главе изложены понятия независимости случайных событий, классов случайных событий, испытаний со случайными исходами и некоторые вопросы, связанные с ними. Отметим, что независимость является одним из важнейших понятий теории вероятностей: именно это понятие определило то своеобразие, которое выделяет теорию вероятностей из общей теории, занимающейся исследованием измеримых пространств с мерой.

Изучение независимости случайных событий начнем с обсуждения понятия условной вероятности.

§ 1. Условные вероятности

Условная вероятность оценивает то изменение в степени уверенности о наступлении некоторого события, которое происходит после получения дополнительной информации.

Рассмотрим два случайных события A и B. Пусть известно, что событие B наступило, но неизвестно, какое конкретно из элементарных событий ω, составляющих событие B, произошло. Что можно сказать в этом случае о вероятности наступления события A?

П р и м е р 1. Пусть дважды брошена симметричная игральная кость, A – событие, состоящее в том, что при первом бросании выпала единица, а B – событие, состоящее в том, что сумма появившихся очков строго меньше четырех.

Вычислим условную вероятность события A, если известно, что произошло событие B.

Имеем $\Omega=\left\{(i,\, j):\ i=\overline{1,6},\, j=\overline{1,6}\right\}$, $A=\left\{(1,\ 1),\ (1,\ 2),\ (1,\ 3),\ (1,\ 4),\ (1,\ 5),\ (1,\ 6)\right\}$, $B=\left\{(1,\ 1),\ (1,\ 2),\ (2,\ 1)\right\}$, $A\cap B=\left\{(1,\ 1),\ (1,\ 2)\right\}$. Следовательно, по определению классической вероятности,

$$\mathrm{P}(A)=\frac{6}{36}=\frac{1}{6},\ \mathrm{P}(B)=\frac{3}{36}=\frac{1}{12},\ \mathrm{P}(A\cap B)=\frac{2}{36}=\frac{1}{18}.$$

Если событие B произошло, то событие A может произойти лишь тогда, когда произойдут элементарные события $(1,\ 1)$ или $(1,\ 2)$, поэтому естественно считать, что

$$\mathrm{P}(A\,|\,B)=\frac{n(A\cap B)}{n(B)}=\frac{2}{3}=\frac{\dfrac{n(A\cap B)}{n(\Omega)}}{\dfrac{n(B)}{n(\Omega)}}=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)},$$

где $n(X)$ – число элементарных событий, входящих в случайное событие X.

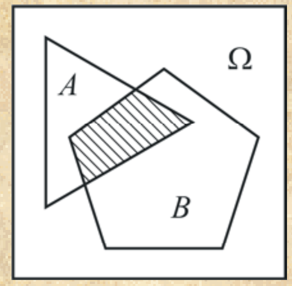

Рис. 1

П р и м е р 2. Внутрь области Ω случайным образом бросается точка. Пусть событие A заключается в том, что точка попала в треугольную область A, а событие B – в том, что точка попала в пятиугольную область B (рис. 1).

Определим вероятность события A при условии, что произошло событие B. Обозначим $S(X)$ – площадь области X, тогда

$$\mathrm{P}(A \mid B) = \frac{S(A \cap B)}{S(B)} = \frac{\dfrac{S(A \cap B)}{S(\Omega)}}{\dfrac{S(B)}{S(\Omega)}} = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}.$$

Дадим теперь определение условной вероятности, согласующееся с интуитивным представлением о ней.

О п р е д е л е н и е 1. *Пусть вероятность события* B *– положительная величина, т. е.* $\mathrm{P}(B) > 0$. *Условной вероятностью события* A *при условии, что произошло событие* B, *называется число*

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

У п р а ж н е н и е 1. Пусть задано вероятностное пространство $(\Omega, \mathcal{A}, \mathrm{P})$, $A \in \mathcal{A}$, $B \in \mathcal{A}$, $\mathrm{P}(B) > 0$. Обозначим $\mathrm{P}_B(A) = \mathrm{P}(A \mid B)$. Докажите, что функция $\mathrm{P}_B(A) \colon \mathcal{A} \to R$ является вероятностью.

У п р а ж н е н и е 2. Обозначим $\mathcal{A}_B = \mathcal{A} \cap B = \{A \cap B : A \in \mathcal{A}\}$. Докажите, что $(B, \mathcal{A}_B, \mathrm{P}_B)$ – вероятностное пространство.

На практике часто бывает так, что известны или достаточно просто определяются именно условные вероятности и с их помощью необходимо вычислить безусловную вероятность некоторого случайного события. Простейшей формулой для решения задач такого типа является формула умножения вероятностей.

Теорема 1 (**Теорема умножения**). *Пусть* $\mathrm{P}(A) > 0$, $\mathrm{P}(B) > 0$. *Тогда*

$$\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B \mid A) = \mathrm{P}(B) \cdot \mathrm{P}(A \mid B).$$

Д о к а з а т е л ь с т в о теоремы следует из определения условной вероятности.

Распространим теорему умножения на случай, когда число сомножителей равно n.

Теорема 2 (**Теорема умножения**). *Пусть*

$$\mathrm{P}\left(A_1 \cap A_2 \cap \ldots \cap A_{n-1}\right) > 0.$$

Тогда

$$\mathrm{P}\left(A_1 \cap A_2 \cap \ldots \cap A_n\right) =$$

$$= \mathrm{P}\left(A_1\right) \cdot \mathrm{P}\left(A_2 \mid A_1\right) \cdot \mathrm{P}\left(A_3 \mid A_1 \cap A_2\right) \cdot \ldots \cdot \mathrm{P}\left(A_n \mid A_1 \cap \ldots \cap A_{n-1}\right).$$

Доказательство теоремы проводится методом математической индукции с использованием результатов теоремы 1.

Пример 3. Студент знает 20 вопросов из 25. Экзаменатор задал ему три вопроса. Какова вероятность того, что студент ответит на все вопросы?

Решение. Введем следующие события:

$$A_1 = \{\text{студент ответил на первый вопрос}\},$$

$$A_2 = \{\text{студент ответил на второй вопрос}\},$$

$$A_3 = \{\text{студент ответил на третий вопрос}\}.$$

Найдем числовые значения вероятностей $\mathrm{P}\left(A_1\right)$, $\mathrm{P}\left(A_2 \mid A_1\right)$, $\mathrm{P}\left(A_3 \mid A_1 \cap A_2\right)$:

$\mathrm{P}\left(A_1\right) = \dfrac{20}{25}$, так как всего вопросов – 25, а «хороших» – 20; $\mathrm{P}\left(A_2 \mid A_1\right) = \dfrac{19}{24}$, так как при подсчете этой вероятности необходимо учесть, что на первый вопрос уже дан ответ, $\mathrm{P}\left(A_3 \mid A_1 \cap A_2\right) = \dfrac{18}{23}$.

Искомую вероятность $\mathrm{P}\left(A_1 \cap A_2 \cap A_3\right)$ находим по теореме 2

$$\mathrm{P}\left(A_1 \cap A_2 \cap A_3\right) = \mathrm{P}\left(A_1\right) \cdot \mathrm{P}\left(A_2 \mid A_1\right) \cdot \mathrm{P}\left(A_3 \mid A_1 \cap A_2\right) = \frac{20}{25} \cdot \frac{19}{24} \cdot \frac{18}{23} \approx 0{,}496.$$

Следующая простая, но важная формула, которая называется формулой полной вероятности, является основным средством при подсчете вероятностей сложных событий с использованием условных вероятностей. Приведем сначала следующее определение.

Определение 2. *Конечное или счетное число случайных событий* $A_1, A_2, \ldots, A_n, \ldots$ *образует полную группу событий (разбиение), если:*

1. $\mathrm{P}\left(A_i\right) > 0,\ i = 1,\ 2,\ \ldots;$

2. $A_i \cap A_j = \varnothing,\ i \neq j;$

3. $\sum\limits_k A_k = \Omega.$

Теорема 3 (**Формула полной вероятности**). *Пусть случайные события* $A_1, \ldots, A_n$ *образуют полную группу событий. Тогда для про-*

извольного события B, *рассматриваемого на том же вероятностном пространстве,*

$$\mathrm{P}(B)=\sum_{k=1}^{n}\mathrm{P}(A_k)\cdot\mathrm{P}(B\mid A_k).$$

Доказательство. Представим событие B в виде

$$B=B\bigcap\Omega=B\bigcap\left(\sum_{k=1}^{n}A_k\right)=\sum_{k=1}^{n}(B\bigcap A_k).$$

Тогда, пользуясь свойством аддитивности вероятности и теоремой 2 умножения, получим

$$\mathrm{P}(B)=\sum_{k=1}^{n}\mathrm{P}(A_k\bigcap B)=\sum_{k=1}^{n}\mathrm{P}(A_k)\cdot\mathrm{P}(B\mid A_k).$$

Замечание 1. Формула полной вероятности верна и при выполнении следующих условий на события A_i:

1. $\mathrm{P}(A_i)>0$,
2. $A_i\bigcap A_j=\varnothing,\ i\neq j$,
3. $B\subset\sum_{i}A_i$.

Пример 4. Студент выучил 20 билетов из 25 и идет отвечать вторым. Какова вероятность, что он вытянет «счастливый» билет?

Решение. Введем случайные события

$A_1=\{$первый отвечающий выбрал «счастливый» билет$\}$, $A_2=\overline{A_1}$. Тогда

$$\mathrm{P}(A_1)=\frac{20}{25},\quad \mathrm{P}(A_2)=\frac{5}{25},\quad \mathrm{P}(B\mid A_1)=\frac{19}{24},\quad \mathrm{P}(B\mid A_2)=\frac{20}{24},$$

где событие $B=\{$второй отвечающий выбрал «счастливый» билет$\}$,

$$\mathrm{P}(B)=\frac{20}{25}\cdot\frac{19}{24}+\frac{5}{25}\cdot\frac{20}{24}=\frac{4}{5}.$$

Таким образом, вероятность выбрать «счастливый» билет такая же, как, если бы студент отвечал первым.

Во многих приложениях теории вероятностей встречается следующая задача. Пусть до опыта об исследуемом случайном явлении имеются гипотезы $A_1, A_2, \ldots, A_n$. После опыта становится известной информация о результатах этого явления, но не полная: результаты наблюдений показывают, не какой конкретно элементарный исход $\omega\in\Omega$ произошел, а что наступило некоторое событие B. Считая, что до опыта были известны *(априорные)* вероятности $\mathrm{P}(A_1), \ldots, \mathrm{P}(A_n)$ и условные вероятности $\mathrm{P}(B\mid A_1), \ldots, \mathrm{P}(B\mid A_n)$, необходимо определить послеопытные *(апостериорные)* вероятности $\mathrm{P}(A_1 \mid B)$, …, $\mathrm{P}(A_n \mid B)$. Решение поставленной задачи дают формулы Байеса.

Теорема 4 (**Формулы Байеса**). *Пусть выполнены условия теоремы 3 и* $\mathrm{P}(B) > 0$. *Тогда для любых значений* $k = 1, 2, \ldots, n$ *имеют место формулы :*

$$\mathrm{P}(A_k \mid B) = \frac{\mathrm{P}(A_k) \cdot \mathrm{P}(B \mid A_k)}{\sum_{j=1}^{n} \mathrm{P}(A_k) \cdot \mathrm{P}(B \mid A_k)}.$$

Доказательство. По определению условной вероятности,

$$\mathrm{P}(A_k \mid B) = \frac{\mathrm{P}(A_k \cap B)}{\mathrm{P}(B)}.$$

Применив к числителю теорему умножения, а к знаменателю формулу полной вероятности, получим требуемое.

Пример 5. В урне находятся 2 шара (оба белые, или оба черные, или один черный и один белый). Из урны достали один шар, он оказался белым. Какова вероятность того, что и второй шар белый?

Решение. Введем следующие события: A_1 = {в урне находятся два белых шара}, A_2 = {в урне находятся два черных шара}, A_3 = {в урне находятся белый и черный шары}, B = {вынутый из урны шар оказался белым}. Тогда $\mathrm{P}(A_1) = \mathrm{P}(A_2) = \mathrm{P}(A_3) = 1/3$ и

$$\mathrm{P}(A_1 \mid B) = \frac{\mathrm{P}(A_1) \cdot \mathrm{P}(B \mid A_1)}{\mathrm{P}(A_1) \cdot \mathrm{P}(B \mid A_1) + \mathrm{P}(A_2) \cdot \mathrm{P}(B \mid A_2) + \mathrm{P}(A_3) \cdot \mathrm{P}(B \mid A_3)} =$$

$$= \frac{\frac{1}{3} \cdot 1}{\frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot \frac{1}{2}} = \frac{2}{3}.$$

§ 2. Независимость событий

В этом параграфе введем понятие независимости событий, опираясь на аппарат условных вероятностей. Пусть A и B – такие случайные события, что $\mathrm{P}(A) > 0$, $\mathrm{P}(B) > 0$ и A не зависит от B. Из последнего естественно предположить, что $\mathrm{P}(A \mid B) = \mathrm{P}(A)$. Но из теоремы умножения $\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B \mid A) = \mathrm{P}(B) \cdot \mathrm{P}(A \mid B)$, а поэтому, во-первых, $\mathrm{P}(B \mid A) = \mathrm{P}(B)$ (т. е. событие B не зависит от события A), а во-вторых, $\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B)$. Эти соображения обосновывают естественность следующего определения.

О п р е д е л е н и е 1. *События* A *и* B *называются независимыми, если*

$$\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B). \qquad (1)$$

З а м е ч а н и е 1. В приведенном определении вероятности событий A и B могут быть равными нулю.

У п р а ж н е н и е 1. Пусть $\mathrm{P}(A) = 0$, а B – произвольное случайное событие. Докажите, что события A и B независимы.

У п р а ж н е н и е 2. Пусть $\mathrm{P}(A) = 1$, а B – произвольное случайное событие. Докажите, что события A и B независимы.

У п р а ж н е н и е 3. Пусть события A и B независимы. Являются ли независимыми следующие пары событий: A и $\overline{B}$, $\overline{A}$ и B, $\overline{A}$ и $\overline{B}$?

У п р а ж н е н и е 4. Пусть события A и B несовместны. Будут ли они независимыми?

Обычно независимость A и B, которую иногда называют *теоретико-вероятностной, или статистической, независимостью* (в отличие от причинной независимости реальных явлений), не устанавливается с помощью равенства (1), а постулируется на основе каких-либо внешних соображений. Равенство (1) позволяет вычислить вероятность $\mathrm{P}(A \cap B)$, зная вероятности $\mathrm{P}(A)$ и $\mathrm{P}(B)$ двух независимых событий A и B. При установлении независимости событий A и B часто используется следующий принцип: *события* A *и* B, *реальные прообразы которых* $\tilde{A}$ *и* $\tilde{B}$ *причинно независимы, независимы в теоретико-вероятностном смысле.* И, конечно, из теоретико-вероятностной независимости событий A и B не следует причинная независимость их реальных прообразов $\tilde{A}$ и $\tilde{B}$. Следующий пример показывает, что независимость может исчезнуть, если незначительно изменить вероятностную модель (см. также гл. 2, задача 9).

П р и м е р 1. Из колоды в пятьдесят две карты случайно вынимается одна. Рассмотрим следующие случайные события: A – вынута карта бубновой масти, B – вынута дама. Являются ли события A и B независимыми?

Найдем вероятности событий A, B и $A \cap B$. Всего в колоде 13 карт бубновой масти, поэтому $\mathrm{P}(A) = \frac{13}{52} = \frac{1}{4}$. В колоде четыре дамы, значит, вероятность вынуть даму $\mathrm{P}(B) = \frac{4}{52} = \frac{1}{13}$. Произведением событий A и B является случайное событие – $\{$вынута бубновая дама$\}$, поэтому $\mathrm{P}(A \cap B) = \frac{1}{52}$. Проверим выполнение равенства $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$, получим $\frac{1}{52} = \frac{1}{13} \cdot \frac{1}{4}$. Таким образом, A и B независимы в теоретико-вероятностном смысле.

Решим такую же задачу при условии, что в колоду добавлен джокер. Очевидно, что при подсчете вероятностей изменятся лишь знаменатели, так как карт теперь стало 53. Следовательно, проверка равенства (1) дает следующий результат:

$$\frac{1}{53} \neq \frac{13}{53} \cdot \frac{4}{53}.$$

Значит, если в колоде есть джокер, то события A и B становятся зависимыми.

Определение 2. *Случайные события* $A_1, \ldots, A_n$ *называются попарно независимыми, если для любых* $i \neq j$, $i, j = 1, 2, \ldots, n$,

$$\mathrm{P}(A_i \cap A_j) = \mathrm{P}(A_i) \cdot \mathrm{P}(A_j).$$

Определение 3. *Случайные события* $A_1, \ldots, A_n$ *называются независимыми в совокупности, если для любого подмножества индексов* $\{i_1, \ldots, i_k\} \subseteq \{1, 2, \ldots, n\}$

$$\mathrm{P}(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = \prod_{j=1}^{k} \mathrm{P}(A_{i_j}).$$

Пусть даны события A, B, C. Запишем, что означают попарная независимость случайных событий и независимость случайных событий в совокупности:

1) $\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B)$; 2) $\mathrm{P}(A \cap C) = \mathrm{P}(A) \cdot \mathrm{P}(C)$;

3) $\mathrm{P}(B \cap C) = \mathrm{P}(B) \cdot \mathrm{P}(C)$; 4) $\mathrm{P}(A \cap B \cap C) = \mathrm{P}(A) \cdot \mathrm{P}(B) \cdot \mathrm{P}(C)$.

Попарная независимость – это выполнение первых трех условий. Если к тому же выполняется и условие 4, то это означает, что события независимы в совокупности. Очевидно, что независимость в совокупности сильнее, т. е. из нее вытекает попарная независимость. Обратное, вообще говоря, не верно. Это показывает следующий пример.

Пример 2 (Пример Бернштейна). На плоскость бросают тетраэдр, три грани которого окрашены соответственно в красный, зеленый и синий цвета, а на четвертой есть все цвета. Рассмотрим следующие события:

$$A = \{\text{выпала грань, содержащая красный цвет}\},$$

$$B = \{\text{выпала грань, содержащая зеленый цвет}\},$$

$$C = \{\text{выпала грань, содержащая синий цвет}\}.$$

Покажем, что события A, B, C независимы попарно, но зависимы в совокупности.

Имеем

$$\mathrm{P}(A\cap B)=\mathrm{P}(A\cap C)=\mathrm{P}(B\cap C)=\frac{1}{4},\quad \mathrm{P}(A)=\mathrm{P}(B)=\mathrm{P}(C)=\frac{1}{2},\quad \mathrm{P}(A\cap B\cap C)=\frac{1}{4}.$$

Тогда $\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B)$, $\mathrm{P}(A\cap C)=\mathrm{P}(A)\mathrm{P}(C)$, $\mathrm{P}(B\cap C)=\mathrm{P}(B)\mathrm{P}(C)$, т. е. события A, B и C попарно независимы. Но $\mathrm{P}(A\cap B\cap C)\neq \mathrm{P}(A)\mathrm{P}(B)\mathrm{P}(C)$, а значит, A, B, C не являются независимыми в совокупности.

Пусть имеется два класса событий. Под *независимостью этих классов событий* по определению будем понимать, что любой представитель одного класса независим с любым представителем из другого. По аналогии для n классов случайных событий можем ввести определения попарной независимости и независимости в совокупности. Сделайте это самостоятельно.

Определение 4. *Сигма-алгебры* $\mathscr{A}_1, \ldots, \mathscr{A}_n$ *называются независимыми, если для любых* $A_k\in\mathscr{A}_k$, $k=1, 2, \ldots n$,

$$\mathrm{P}(A_1\cap A_2\cap\ldots\cap A_n)=\prod_{k=1}^{n}\mathrm{P}(A_k). \tag{2}$$

Упражнение 5. Установите связь независимости, заданной соотношениями (2), с попарной независимостью и независимостью в совокупности классов событий.

§ 3. Независимые испытания

Цель этого параграфа – построить математическую модель n независимых испытаний, каждое из которых имеет конечное число исходов. Напомним, что под *испытанием* мы понимаем эксперимент со случайными исходами. Рассмотрим сначала следующий простейший пример.

Пример 1. Пусть на отрезок $[0;1]$ независимо друг от друга случайным образом бросаются две частицы (точки), каждая из которых с одинаковой возможностью может попасть в любую точку отрезка $[0;1]$. Событие $\tilde{A}$ – первая частица попала в борелевское подмножество $A\subseteq[0;1]$, событие $\tilde{B}$ – вторая частица попала в борелевское подмножество $B\subseteq[0;1]$. Необходимо построить вероятностное пространство, в котором можно было бы описать эти испытания. Очевидно, что мы не можем в качестве искомого взять «одномерное» геометрическое пространство $(\Omega, \mathscr{A}, \mathrm{P})$, где $\Omega=[0;1]$, $\mathscr{A}=\mathscr{B}[0;1]$, P – мера Лебега. В этом случае независимость событий $\tilde{A}$ и $\tilde{B}$ будет определяться взаимным расположением подмножеств A и B на отрезке $[0;1]$. Для того, чтобы события $\tilde{A}$ и $\tilde{B}$ были независимы при любых борелевских подмножествах A и B, в качестве вероятностного пространства $(\Omega, \mathscr{A}, \mathrm{P})$ рассмотрим следующее: $\Omega = [0;1]\times[0;1]$, $\mathscr{A} = \mathscr{B}([0;1]\times[0;1])$, P – мера Лебега на R^2.

Несложно видеть (см. рис. 2), что при таком подходе события $\tilde{A}$ и $\tilde{B}$ независимы для любых борелевских множеств A и B (проверьте это самостоятельно, используя аксиоматику геометрического вероятностного пространства).

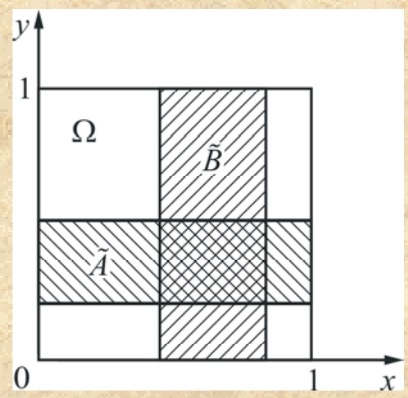

Рис. 2

Эту идею положим в основу для решения задачи данного параграфа. Поскольку математическая модель любого испытания с конечным числом исходов – конечное вероятностное пространство, то в нашем случае n испытаний имеем n вероятностных пространств. Пусть $(\Omega_i, \mathscr{A}_i, \mathrm{P}_i)$, $i=\overline{1,n}$, – конечные вероятностные пространства, где $\Omega_i = \{\omega_i\}$, $\mathscr{A}_i = \mathscr{P}(\Omega_i)$, $\mathrm{P}_i(\omega_i) = p_i$, $i=\overline{1,n}$.

Вложим все эти пространства в более общее, которое конструктивно зададим следующим образом:

$$\Omega = \left\{\omega = (\omega_1, \ldots, \omega_n) : \omega_i \in \Omega_i,\ i=\overline{1,n}\right\},\ \mathscr{A} = \mathscr{P}(\Omega),$$

$$\mathrm{P}(\omega) = \mathrm{P}_1(\omega_1)\cdot\ldots\cdot\mathrm{P}_n(\omega_n),$$

для любого $A \in \mathscr{A}$

$$\mathrm{P}(A) = \sum_{\omega\in A}\mathrm{P}(\omega).$$

Построенная так вероятность P называется *прямым произведением вероятностей* P_i и обозначается $\mathrm{P} = \mathrm{P}_1\times\mathrm{P}_2\times\ldots\times\mathrm{P}_n$. Аналогично в этом случае $\mathscr{A} = \mathscr{A}_1\times\mathscr{A}_2\times\ldots\times\mathscr{A}_n$ есть *прямое произведение* σ*-алгебр*. В построенном вероятностном пространстве выделим класс событий A, называемых *параллелепипедами*, который определим следующим образом: пусть $A_i \in \mathscr{A}_i$, $i=\overline{1,n}$, параллелепипед $A = A_1\times A_2\times\ldots\times A_n$ составлен из тех и только тех элементарных событий $\omega = (\omega_1,\ldots,\omega_n)$, для которых $\omega_i \in A_i$, $i=\overline{1,n}$. Из определения вероятности $\mathrm{P}(\omega)$ вытекает, что вероятность параллелепипеда равна

$$\mathrm{P}(A) = \sum_{\omega\in A}\mathrm{P}(\omega) = \sum_{\omega_1\in A_1}P_1(\omega_1)\cdot\ldots\cdot\sum_{\omega_i\in A_i}P_i(\omega_i)\cdot\ldots\cdot\sum_{\omega_n\in A_n}P_n(\omega_n) = \prod_{k=1}^{n}P_k(A_k).$$

Обозначим через $\mathscr{A}'_m$ подалгебру алгебры $\mathscr{A}$, состоящую из всех тех параллелепипедов, у которых $\mathscr{A}_i = \Omega_i$ для $i \neq m$. Установим изоморфизм

$$A'_i \sim A_i \in \mathscr{A} \qquad \left(A'_i = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_{i-1} \times A_i \times \Omega_{i+1} \times \ldots \times \Omega_n \in \mathscr{A}'_i\right).$$

Поэтому вместо событий A_i из вероятностного пространства $(\Omega_i, \mathscr{A}_i, \mathrm{P}_i)$ будем рассматривать изоморфные события A'_i из подалгебры $\mathscr{A}'_i$ вероятностного пространства $(\Omega, \mathscr{A}, \mathrm{P})$. Очевидно, что $\mathrm{P}(A'_i) = \mathrm{P}_i(A_i)$. Сигма-алгебры $\mathscr{A}'_1, \mathscr{A}'_2, \ldots, \mathscr{A}'_n$ независимы, так как

$$A = A_1 \times A_2 \times \ldots \times A_n = \bigcap_{k=1}^{n} A'_k,$$

и отсюда получаем, что

$$\mathrm{P}\left(\bigcap_{k=1}^{n} A'_k\right) = \prod_{k=1}^{n} \mathrm{P}(A'_k)$$

для любых $A'_k \in \mathscr{A}'_k$.

С х е м а Б е р н у л л и (б и н о м и а л ь н а я с х е м а). Пусть проводятся n независимых испытаний. Известно, что в каждом испытании возможны два исхода: либо происходит событие A (*успех*), либо событие A не происходит (*неуспех*). Данная схема называется *схемой Бернулли*. При этом предполагается, что вероятность p успеха и $q = 1 - p$ неуспеха не изменяются при переходе от испытания к испытанию. Простейшим примером такой схемы является следующий эксперимент.

П р и м е р 2. n раз бросается монета. Очевидно, что в любом испытании возможны два исхода – выпадение орла или решки.

Опишем математическую модель схемы Бернулли, т. е. построим соответствующее вероятностное пространство:

$$\Omega_i = \{0,\ 1\},\ \mathscr{A}_i = \{\varnothing,\ \{0\},\ \{1\},\ \Omega_i\},\ \mathrm{P}_i(1) = p,\ \mathrm{P}_i(0) = q,\ p + q = 1.$$

Тогда прямое произведение вероятностных пространств $(\Omega, \mathscr{A}, \mathrm{P})$ имеет вид:

$$\Omega = \{(\omega_1, \omega_2, \ldots, \omega_n) = \omega;\ \omega_i = 0, 1\},$$

$$\mathscr{A} = \mathscr{P}(\Omega),$$

$$\mathrm{P}(\omega) = \prod_{i=1}^{n} p^{\omega_i} q^{1-\omega_i}.$$

Для любого события A из Ω

$$\mathrm{P}(A)=\sum_{\omega\in A}\mathrm{P}(\omega).$$

Пусть событие B состоит в том, что при n независимых испытаниях в схеме Бернулли произошло ровно k успехов, т. е.

$$B=\left\{\omega\in(\omega_1,\ ...,\ \omega_n)\in\Omega:\ \omega_1+...+\omega_n=k,\ \omega_i=0,\ 1,\ i=\overline{1,n}\right\}.$$

Так как для $\omega\in B$ (см. общую схему)

$$\mathrm{P}(\omega)=p^k\cdot q^{n-k},$$

а число элементарных событий $\omega\in B$ равно C_n^k, то

$$\mathrm{P}(B)=C_n^k p^k q^{n-k}=P_n(k),\ k=\overline{0,n}.$$

Вероятности $P_n(k),\ k=\overline{0,n},$ называются *биномиальными вероятностями.*

П р и м е р 3. Известно, что левши составляют 1 %. Найти вероятность того, что среди 200 человек найдется хотя бы трое левшей.

Для решения этой задачи используется схема Бернулли, в которой проводятся 200 независимых испытаний с вероятностью успеха в отдельном испытании, равной $1/100$. Тогда $P_{200}(0)$ – вероятность того, что среди 200 человек нет левшей, $P_{200}(1)$ – есть один левша, $P_{200}(2)$ – есть двое левшей. Искомая вероятность равна

$$1-P_{200}(0)-P_{200}(1)-P_{200}(2)=$$

$$=1-C_{200}^0\cdot\left(\frac{1}{100}\right)^0\cdot\left(\frac{99}{100}\right)^{200}-C_{200}^1\cdot\left(\frac{1}{100}\right)^1\cdot\left(\frac{99}{100}\right)^{199}-C_{200}^2\cdot\left(\frac{1}{100}\right)^2\cdot\left(\frac{99}{100}\right)^{198}.$$

П о л и н о м и а л ь н а я с х е м а. Более сложная схема n независимых испытаний получается, когда при каждом испытании возможно появление одного из k попарно несовместных исходов $A_1,\ A_2,\ \dots,\ A_k$. Такая схема называется *полиномиальной схемой*. При $k=2$ она превращается в схему Бернулли.

Пусть i-е испытание связано с вероятностным пространством $(\Omega_i,\ \mathcal{A}_i,\ \mathrm{P}_i),\ i=\overline{1,n},$ где $\Omega_i=\{1,\ 2,\ ...,\ k\}$ состоит из номеров исходов $A_1,\ ...,\ A_k,\ \mathcal{A}_i=\mathcal{P}(\Omega),\ \mathrm{P}_i(j)=p_j,\ p_j\geq 0,\ j=\overline{1,k},\ p_1+...+p_k=1.$

Тогда согласно общей схеме прямое произведение $(\Omega,\ \mathcal{A},\ \mathrm{P})$ может быть представлено следующим образом:

$$\Omega=\left\{\omega=(\omega_1,\ ...,\ \omega_n):\ \omega_i=\overline{1,k}\right\},\ \mathcal{A}=\mathcal{P}(\Omega),$$

$$\mathrm{P}(\omega)=\mathrm{P}_1(\omega_1)\cdot\mathrm{P}_2(\omega_2)\cdot...\cdot\mathrm{P}_n(\omega_n).$$

Пусть событие B состоит в том, что в n испытаниях полиномиальной схемы событие A_1 произойдет ровно m_1 раз, A_2 – ровно m_2 раз, …, A_k – ровно m_k раз, $m_1 + \ldots + m_k = n$. Вероятность события B обозначается $P_n(m_1, \ldots, m_k)$ и называется *полиномиальной вероятностью*.

Так как для любого $\omega \in \Omega$ (см. общую схему)

$$\mathrm{P}(\omega) = p_1^{m_1} \cdot p_2^{m_2} \cdot \ldots \cdot p_k^{m_k},$$

а количество элементарных событий $\omega \in B$ равно (см. дополнения к гл. 1, § 1)

$$C_n^{m_1} \cdot C_{n-m_1}^{m_2} \cdot \ldots \cdot C_{n-m_1-\ldots-m_{k-1}}^{m_k} = \frac{n!}{m_1! \ldots m_k!},$$

то для любых $m_i \geq 0$, $m_1 + \ldots + m_k = n$, получаем

$$P_n(m_1, \ldots, m_k) = \frac{n!}{m_1! \ldots m_k!} \cdot p_1^{m_1} \cdot \ldots \cdot p_k^{m_k}.$$

П р и м е р 4. В некотором обществе блондины составляют 40 %, брюнеты – 30 %, шатены – 20 %, рыжие – 10 %. Случайно выбираются 6 человек. Найти вероятность того, что число блондинов и шатенов среди них одинаково.

Очевидно, что возможны только четыре выборки, при которых число блондинов и шатенов одинаково: среди шести выбранных нет блондинов и шатенов, один блондин и один шатен, два шатена и два блондина, три шатена и три блондина. Введем следующие события:

$$A_1 = \{\text{случайно выбранный человек является блондином}\},$$

$$A_2 = \{\text{случайно выбранный человек является шатеном}\},$$

$$A_3 = \overline{A}_1 \cap \overline{A}_2.$$

Вероятности этих событий будут равны соответственно $\mathrm{P}(A_1) = \dfrac{4}{10}$, $\mathrm{P}(A_2) = \dfrac{2}{10}$, $\mathrm{P}(A_3) = \dfrac{4}{10}$. Тогда искомая вероятность запишется в виде

$$P_6(0,0,6) + P_6(1,1,4) + P_6(2,2,2) + P_6(3,3,0),$$

где $P_6(m_1, m_2, m_6) = \dfrac{6!}{m_1! m_2! m_3!} \mathrm{P}(A_1)^{m_1} \mathrm{P}(A_2)^{m_2} \mathrm{P}(A_3)^{m_3}$ – вероятность того, что из шести случайно выбранных человек m_1 являются блондинами, m_2 – шатенами, m_3 – рыжими или брюнетами, причем $m_1 + m_2 + m_3 = 6$.

ДОПОЛНЕНИЯ К ГЛАВЕ 2

§ 1. Предельные теоремы в схеме Бернулли

Схема независимых испытаний служит вероятностной моделью многих реальных явлений, поэтому представляет значительный интерес задача подсчета вероятностей $P_n(m)$ (т. е. того, что в n независимых испытаниях схемы Бернулли произошло ровно m успехов). При больших значениях n и m есть трудности в получении численного значения этих вероятностей. Естественным образом возникает задача нахождения асимптотических формул, позволяющих приближенно вычислять вероятности $P_n(m)$ и их суммы $\sum_{m=m_1}^{m_2} P_n(m)$ для достаточно больших значений *n*. Сначала исследуем случай больших *n* и малых *p*.

Рассмотрим последовательность серий

E_{11},

E_{21}, E_{22},

E_{31}, E_{32}, E_{33},

…………………

$E_{n1}, E_{n2}, E_{n3}, \ldots, E_{nn}$,

………………………,

для которой все события одной серии независимы и вероятность успеха p_n в каждом испытании некоторой серии зависит только от номера серии. Обозначим μ_n – число наступивших успехов в *n*-й серии.

Теорема 1 (**Локальная предельная теорема Пуассона**). *Если* $p_n \to 0$, $n \to \infty$, *так, что* $p_n \cdot n \to a$, *то*

$$\mathrm{P}(\mu_n = m) = C_n^m p_n^m (1 - p_n)^{n-m} \to \frac{a^m}{m!} \cdot e^{-a}.$$

Д о к а з а т е л ь с т в о. Пусть зафиксировано число *m*. Тогда справедливы следующие соотношения:

$$P_n(m) = \mathrm{P}(\mu_n = m) = C_n^m \cdot p_n^m \cdot (1 - p_n)^{n-m} = \frac{n!}{m! \cdot (n-m)!} \cdot p_n^m \cdot (1 - p_n)^{n-m} =$$

$$= \frac{(np_n)^m}{m!} \cdot \frac{n(n-1)\ldots(n-(m-1))}{n^m} \cdot (1 - p_n)^{-m} \cdot (1 - p_n)^n =$$

$$= \frac{(np_n)^m}{m!}\left(1 - \frac{1}{n}\right)\ldots\left(1 - \frac{m-1}{n}\right)(1 - p_n)^{-m} (1 - p_n)^{-\frac{1}{p_n}(-np_n)} \to \frac{a^m}{m!} \cdot e^{-a}.$$

П р и м е р 1. Рассмотрим теперь продолжение прим. 3 (см. гл. 2, § 3). Воспользовавшись локальной предельной теоремой Пуассона, найдем приближенное численное значение вероятности того, что среди 200 человек хотя бы трое левшей.

На практике обычно пользуются пуассоновским приближением уже тогда, когда $n \geq 100$, $p \leq 1/100$.

Итак, $\mathrm{P}(A) = 1 - P_{200}(0) - P_{200}(1) - P_{200}(2)$. Имеем, что $n = 200$, $p = \dfrac{1}{100}$ и $a = np = 2$, поэтому искомая вероятность будет приближенно равна

$$1 - \frac{2^0}{0!}\cdot e^{-2} - \frac{2^1}{1!}\cdot e^{-2} - \frac{2^2}{2!}\cdot e^{-2} = 1 - 5\cdot e^{-2} \approx 0{,}323.$$

Приведем без доказательства теорему, позволяющую определять точность «пуассоновского» приближения.

Теорема 2 (Интегральная предельная теорема Пуассона) [34]. *В схеме Бернулли для любого $n \in N$, любого $p \in (0;1)$ и для любого числового множества B справедливо неравенство*

$$\left|\mathrm{P}(\mu_n \in B) - \sum_{m\in B}\frac{a^m}{m!}\cdot e^{-a}\right| \leq \frac{a^2}{n}, \quad \text{где } a = np.$$

Теорема, которая будет доказана ниже, дает асимптотическую формулу для биномиальной вероятности при p не близких к 0 или 1.

Теорема 3 (Локальная предельная теорема Муавра – Лапласа). *Если в схеме Бернулли $\sigma = \sqrt{npq} \xrightarrow[n\to\infty]{} \infty$, то для любого $C > 0$ равномерно по всем x,*

$$x \in \left\{ x \in R:\ |x| \leq C,\ x = \frac{m - np}{\sigma},\ m \in N \cup \{0\} \right\}$$

справедливо соотношение

$$\mathrm{P}\left\{\frac{\mu_n - np}{\sqrt{npq}} = x\right\} = \frac{1}{\sigma\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}\cdot\left(1 + o(1)\right), \tag{1}$$

где $o(1)$ – бесконечно малая величина при $\sigma \to \infty$

Иллюстрация в Mathematica

Схема доказательства

Д о к а з а т е л ь с т в о. Прологарифмировав равенство

$$P_n(m) = \frac{n!}{m!(n-m)!}\cdot p^m \cdot q^{n-m},$$

получим

$$\ln P_n(m) = \ln\left(\frac{n!}{m!(n-m)!}\cdot p^m\cdot q^{n-m}\right) =$$

$$= \ln n! - \ln m! - \ln(n-m)! + m\ln p + (n-m)\ln q. \tag{2}$$

Известна асимптотическая формула Стирлинга

$$\ln n! = n\ln n + \ln\sqrt{2\pi n} - n + \theta_n, \tag{3}$$

где

$$|\theta_n| \leq \frac{1}{12n}.$$

Воспользовавшись соотношением (3), получим

$$\ln P_n(m) = n\ln n + \ln\sqrt{2\pi n} - n + \theta_n - m\ln m - \ln\sqrt{2\pi m} + m - \theta_m -$$
$$-k\ln k - \ln\sqrt{2\pi k} + k - \theta_k + m\ln p + k\ln q, \tag{4}$$

где $k = n - m$.

Для законности применения формулы Стирлинга убедимся, что $m \to \infty$, $k \to \infty$. Действительно

$$m = np + x\sigma = np\left(1 + \frac{xq\sigma}{npq}\right) = np\left(1 + \frac{xq}{\sigma}\right) \to \infty, \tag{5}$$

$$k = n - m = n - np - x\sigma = nq - x\sigma = nq\left(1 - \frac{xp}{\sigma}\right) \to \infty. \tag{6}$$

Из математического анализа известно, что

$$\ln(1+\varepsilon) = O(\varepsilon) \text{ при } \varepsilon \to 0,$$

$$\ln(1+\varepsilon) = \varepsilon - \frac{\varepsilon^2}{2} + O(\varepsilon^3) \text{ при } \varepsilon \to 0,$$

поэтому

$$\ln\sqrt{2\pi n} - \ln\sqrt{2\pi m} - \sqrt{2\pi k} = \frac{1}{2}\ln\frac{2\pi n}{2\pi m 2\pi k} = \frac{1}{2}\ln\frac{1}{2\pi npq\left(1 + \frac{xq}{\sigma}\right)\left(1 - \frac{xp}{\sigma}\right)} =$$

$$= \ln\frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2}\ln\left(1 + \frac{xq}{\sigma}\right) - \frac{1}{2}\ln\left(1 - \frac{xp}{\sigma}\right) = \ln\frac{1}{\sigma\sqrt{2\pi}} + O\left(\frac{1}{\sigma}\right). \tag{7}$$

Очевидно, что

$$\theta_n = O\left(\frac{1}{\sigma^2}\right), \quad \theta_m = O\left(\frac{1}{\sigma^2}\right), \quad \theta_k = O\left(\frac{1}{\sigma^2}\right),$$

$$\theta_n + \theta_m + \theta_k = O\left(\frac{1}{\sigma}\right). \tag{8}$$

Вернемся к равенству (4). Рассмотрим оставшиеся слагаемые правой части этого равенства:

$$n\ln n - m\ln m - k\ln k + m\ln p + k\ln q = m\ln n + k\ln n - m\ln m - k\ln k + m\ln p + k\ln q =$$

$$= -m\ln\frac{m}{np} - k\ln\frac{k}{nq} = -(np + x\sigma)\ln\frac{np\left(1 + \frac{xq}{\sigma}\right)}{np} - (nq - x\sigma)\ln\frac{nq\left(1 - \frac{xp}{\sigma}\right)}{nq} =$$

$$= -(np + x\sigma)\left(\frac{xq}{\sigma} - \frac{x^2p^2}{2\sigma^2} + O\left(\frac{1}{\sigma^3}\right)\right) - (nq - x\sigma)\left(-\frac{xp}{\sigma} - \frac{x^2p^2}{2\sigma^2} + O\left(\frac{1}{\sigma^3}\right)\right) =$$

$$= -\left(\frac{npxq}{\sigma} + x^2q - \frac{npx^2q^2}{2\sigma^2} - \frac{x^3q^2}{2\sigma} + O\left(\frac{1}{\sigma}\right)\right) - \left(-\frac{nq\,xp}{\sigma} + x^2p - \frac{nqx^2p^2}{2\sigma^2} + \frac{x^3p^2}{2\sigma} + O\left(\frac{1}{\sigma}\right)\right) =$$

$$= -\frac{x^2}{2} + O\left(\frac{1}{\sigma}\right). \tag{9}$$

Подставив в равенство (4) соотношения (7) – (9), получим

$$\ln P_n(m) = \ln\frac{1}{\sigma\sqrt{2\pi}} - \frac{x^2}{2} + O\left(\frac{1}{\sigma}\right),$$

$$P_n(m) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \cdot e^{O\left(\frac{1}{\sigma}\right)} = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}\left(1+O\left(\frac{1}{\sigma}\right)\right) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}\left(o(1)+1\right).$$

Теорема доказана.

Теорема 4 (Интегральная предельная теорема Муавра – Лапласа). *При выполнении условий теоремы* 3 *равномерно по* $-\infty \le a < b \le +\infty$ *выполнено предельное соотношение*

$$\mathrm{P}\left(a \le \frac{\mu_n - np}{\sigma} \le b\right) \to \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}\,dx.$$

Более общий вариант этой теоремы будет доказан в гл. 6, поэтому здесь ее доказательство мы опускаем.

Для просмотра иллюстрации нажмите кнопку **Анимация.**

З а м е ч а н и е 1. Ю. В. Прохоров показал, что

$$\sup_k\left|P_n(k) - \frac{a^k e^{-a}}{k!}\right| \le \frac{2a}{n}\min(2, a).$$

З а м е ч а н и е 2. Теорема Берри – Эсеена утверждает, что в интегральной предельной теореме Муавра – Лапласа

$$\sup_{x\in R}\left|\mathrm{P}\left(\frac{\mu_n - np}{\sqrt{npq}} \le x\right) - \Phi(x)\right| \le \frac{p^2+q^2}{\sqrt{npq}}.$$

§ 2. Цепи Маркова

Непосредственным обобщением схемы независимых испытаний является схема так называемых цепей Маркова. Мы ограничимся изложением элементов этой теории.

Представим себе, что производится последовательность испытаний, в каждом из которых может осуществиться одно и только одно из r несовместных событий $A_1(t)$, $A_2(t)$, $\ldots$, $A_r(t)$, $t = 0,\ 1,\ \ldots,\ T$, – номер испытания. Будем говорить, что последовательность испытаний образует *цепь Маркова*, если условная вероятность того, что в $t+1$-м испытании осуществится событие $A_1(t+1)$, $i = \overline{1,r}$, зависит только от того, какое событие произошло при t-м испытании и не изменяется от добавочных сведений о том, какие события происходили в более ранних испытаниях.

Изложим теперь понятие цепей Маркова, учитывая терминологию и модели, введенные в теме «Независимые испытания» (см. гл. 2, § 3). Пусть параметр t (время) дискретен, $t = 0,\ 1,\ \ldots,\ T$, и алгебры

$$\mathcal{A}_0,\ \mathcal{A}_1,\ \ldots,\ \mathcal{A}_T \qquad (1)$$

описывают последовательность случайных испытаний. Цепь Маркова определим следующим образом. В последовательности испытаний (1) зафиксируем какой-нибудь момент времени t.

Алгебру событий $\mathscr{A}_t$ назовем *настоящим*, алгебру $\mathscr{A}_0^{t-1}$, порожденную алгебрами $\mathscr{A}_0,\ \mathscr{A}_1,\ \ldots,\ \mathscr{A}_{t-1}$ (т. е. $\mathscr{A}_0^{t-1}=\sigma\{\mathscr{A}_0,\ \mathscr{A}_1,\ \ldots,\ \mathscr{A}_{t-1}\}$), назовем *прошлым*, а алгебру событий $\mathscr{A}_{t+1}^T=\sigma\{\mathscr{A}_{t+1},\ \mathscr{A}_1,\ \ldots,\ \mathscr{A}_T\}$ – *будущим*. Любое событие из $\mathscr{A}_0^{t-1}$ также назовем *прошлым*, из $\mathscr{A}_{t+1}^T$ – *будущим*, из $\mathscr{A}_t$ – *настоящим*.

О п р е д е л е н и е 1. *Последовательность испытаний* (1) *будем называть цепью Маркова, если при любом фиксированном настоящем* $A_i(t),\ t=1,\ \ldots,\ T-1,$ *прошлое* $\mathscr{A}_0^{t-1}$ *и будущее* $\mathscr{A}_{t+1}^T$ *независимы, т. е. для любых* $A\in\mathscr{A}_0^{t-1}$ *и* $B\in\mathscr{A}_{t+1}^T$

$$\mathrm{P}\big(AB\mid A_i(t)\big)=\mathrm{P}\big(A\mid A_i(t)\big)P\big(B\mid A_i(t)\big),\ i=\overline{1,r}.. \tag{2}$$

З а м е ч а н и е 1. Так как

$$\frac{\mathrm{P}(AB\mid C)}{\mathrm{P}(A\mid C)}=\mathrm{P}(B\mid AC),$$

если $\mathrm{P}(AC)>0$, то условие (2) равносильно условию

$$\mathrm{P}\big(B\mid A_i(t)\cdot A\big)=\mathrm{P}\big(B\mid A_i(t)\big),\ i=1,\ \ldots,\ r,\ t=1,\ \ldots,\ T-1. \tag{3}$$

Пусть состояние некоторой системы описывается точкой фазового пространства $E=\{1,\ 2,\ \ldots,\ r\}$. Эволюция изучаемой системы описывается траекторией $\omega=\{\omega_0,\ \omega_1,\ \ldots,\ \omega_T\}$, где $\omega_t=i$, если в момент времени t система находилась в состоянии *i*. Обозначим $A_i(t)=\{\omega:\ \omega_t=i\}$, $i=\overline{1,r},\ t=\overline{0,T}$. Вычислим вероятность того, что траектория $\omega=\{\omega_0,\ \omega_1,\ \ldots,\ \omega_T\}$ равна $\{i_0,\ i_1,\ \ldots,\ i_T\}$, $i_k=\overline{1,r}$, $k=\overline{0,T}$. По теореме умножения с учетом условия (3) имеем

$$\mathrm{P}\big(\omega=\{i_0,\ i_1,\ \ldots,\ i_T\}\big)=\mathrm{P}\big(A_{i_0}(0)\big)\prod_{t=1}^{T}\mathrm{P}\big(A_{i_t}(t)\mid A_{i_0}(0)A_{i_1}(1)\ldots A_{i_{t-1}}(t-1)\big)=$$

$$=\mathrm{P}\big(A_{i_0}(0)\big)\prod_{t=1}^{T}\mathrm{P}\big(A_{i_t}(t)\mid A_{i_{t-1}}(t-1)\big). \tag{4}$$

В дальнейшем будем рассматривать *однородные цепи Маркова*, в которых условные вероятности

$$\mathrm{P}\big(A_j(t)\mid A_i(t-1)\big)=p_{ij},$$

называемые *переходными вероятностями*, не зависят от t.

Таким образом, чтобы вычислить вероятность любой траектории ω в однородной цепи Маркова, достаточно задать начальное распределение $p_i(0)=\mathrm{P}\big(A_i(0)\big)$, $i=\overline{1,r}$, и *матрицу переходных вероятностей*

$$P=\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix}. \qquad (5)$$

Формула (4) запишется тогда так:

$$\mathrm{P}\left(\omega=\left(i_0,\ i_1,\ \ldots,\ i_T\right)\right)=p_{i_0}(0)\prod_{t=1}^{T} p_{i_{t-1}\, i_t}. \qquad (6)$$

Отметим, что элементы матрицы (5) обладают следующими свойствами:

$$\sum_{j=1}^{r} p_{ij}=1,\ p_{ij}\ge 0. \qquad (7)$$

Любая квадратная матрица (5) со свойствами (7) называется *стохастической*. Введем переходные вероятности за t шагов

$$p_{ij}(t)=\mathrm{P}\left(A_j(t+s)\middle|A_i(s)\right).$$

Матрица

$$P(t)=\left[p_{ij}(t)\right]_{i,j=1}^{r}$$

так же будет стохастической.

Из формулы полной вероятности, используя (3) и условие однородности, получаем

$$p_{ij}(t+s)=\sum_{k=1}^{r} p_{ik}(t)\,p_{kj}(s)$$

или в матричной форме

$$P(t+s)=P(t)P(s).$$

Через начальные вероятности $p_i(0)$ и переходные вероятности $p_{ij}(t)$ можем выразить распределение вероятностей $p_i(t)=\mathrm{P}\left(A_i(t)\right)$ при любом *t*:

$$p_i(t)=\sum_{k=0}^{r} p_k(0)\,p_{ki}(t).$$

П р и м е р 1. *Блуждание с поглощением.* Пусть по точкам $0,\ 1,\ 2,\ \ldots,\ N$ прямой блуждает частица. Время t дискретно. Если в момент t частица была в точке i, то в момент $t+1$ она, независимо от ее положения в более ранние моменты времени, с вероятностью p_{ij} попадет в точку *j*. Если матрица переходных вероятностей задается равенствами $p_{00}=p_{NN}=1$, $p_{i,i+1}=p$, $p_{i,i-1}=1-p$ для $i=1,\ 2,\ \ldots,\ N-1$ и $p_{ij}=0$ при $|i-j|>1$, то получаем цепь Маркова, которая описывает блуждание частицы по целым точкам отрезка $[0;N]$ с поглощением на концах.

П р и м е р 2. *Блуждание с отражением.* Пусть переходные вероятности $p_{i,i+1},\ p_{i,i-1}$ для $i=1,\ 2,\ \ldots,\ N-1$ и p_{ij} для $|i-j|>1$ остаются такими как и в прим. 1. Если $p_{00}=1-p$, $p_{01}=p$, $p_{NN}=p$, $p_{N,N-1}=1-p$, то полученная цепь

Маркова моделирует блуждание по целым точкам отрезка $[0;N]$ с отражением на концах.

Теорема 1 (**О предельных вероятностях**). *Если при некотором* t_0 *все элементы матрицы* $P(t_0)$ *положительны, то существуют пределы*

$$\lim_{t\to\infty} p_{ij}(t) = p_j,\ j=\overline{1,r}. \tag{8}$$

Предельные вероятности p_j *не зависят от начального состояния* i *и являются единственным решением системы*

$$\sum_{k=1}^{r} x_k p_{kj} = x_j,\ j=\overline{1,r},\ \sum_{j=1}^{r} x_j = 1. \tag{9}$$

Доказательство. Обозначим $M_j(t)=\max_i p_{ij}(t)$, $m_j(t)=\min_i p_{ij}(t)$. Так как $m_j(t)\le p_{kj}(t)\le M_j(t)$ при любом $k=\overline{1,r}$, то из равенства

$$p_{ij}(t+1)=\sum_k p_{ik}p_{kj}(t)$$

следует, что при всех i

$$m_j(t)\le p_{ij}(t+1)\le M_j(t).$$

Отсюда

$$m_j(t)\le m_j(t+1)\le M_j(t+1)\le M_j(t).$$

Таким образом, при $t\to\infty$ имеются пределы у последовательностей $m_j(t)$ и $M_j(t)$. Докажем, что эти пределы совпадают.

Пусть i и j таковы, что $p_{ik}(t+t_0)=M_k(t+t_0)$, $p_{jk}(t+t_0)=m_k(t+t_0)$, тогда

$$M_k(t+t_0)-m_k(t+t_0)=\sum_{l=1}^{r}\left(p_{il}(t_0)-p_{jl}(t_0)\right)p_{lk}(t).$$

Разобьем сумму в правой части этого равенства на сумму Σ^+ положительных слагаемых и сумму Σ^- отрицательных слагаемых

$$M_k(t+t_0)-m_k(t+t_0)\le M_k(t)\sum_l{}^{+}\left(p_{il}(t_0)-p_{jl}(t_0)\right)+$$

$$+m_k(t)\sum_l{}^{-}\left(p_{il}(t_0)-p_{jl}(t_0)\right)=M_k(t)\Sigma^+ + m_k(t)\Sigma^-. \tag{10}$$

Так как

$$0=\sum_l\left(p_{il}(t_0)-p_{jl}(t_0)\right)=\Sigma^+ + \Sigma^-,$$

то

$$\Sigma^+ = -\Sigma^-.$$

Обозначим

$$\sum_l{}^{+}\left(p_{il}(t_0)-p_{jl}(t_0)\right)=d_{ij}.$$

Из условий теоремы следует, что все $d_{ij}<1$, поэтому $d=\max_{i,j} d_{ij}(t)<1$. Из (10) имеем

$$M_k(t+t_0)-m_k(t+t_0)\le d\left(M_k(t)-m_k(t)\right)$$

и

$$0\le M_k(t)-m_k(t)\le d^{\left[\frac{t}{t_0}\right]}\to 0$$

при $t\to\infty$. Так как $m_k(t)\le p_{ik}(t)\le M_k(t)$, то отсюда следует (8). Перейдем в уравнениях $p_{ij}(t+1)=\sum_{k=1}^{r}p_{ik}(t)\,p_{kj}$ к пределу при $t\to\infty$. Получаем $p_j=\sum_{k=1}^{r}p_k p_{kj}$. Кроме того, $\sum_{j=1}^{r}p_j=1$, т. е. предельные вероятности удовлетворяют системе (9). Предположим, что какие-либо $x_1,\ \dots,\ x_r$ удовлетворяют (9). Тогда они при любом t удовлетворяют системе

$$x_j=\sum_{k=1}^{r}x_k p_{kj}(t). \tag{11}$$

Это доказывается по индукции:

$$\sum_{k=1}^{r}x_k p_{kj}(t+1)=\sum_{k=1}^{r}x_k\sum_{i=1}^{r}p_{ki}p_{ij}(t)=\sum_{i=1}^{r}\left(\sum_{k=1}^{r}x_k p_{ki}\right)p_{ij}(t)=\sum_{i=1}^{r}x_i p_{ij}(t)=x_j.$$

Переходя в равенстве (11) к пределу при $t\to\infty$, получаем $x_j=\sum_{k=1}^{r}x_k p_j=p_j$. Теорема доказана.

З а м е ч а н и е 2. Можно проверить, что цепь Маркова, описывающая блуждание с отражением в прим. 2, удовлетворяет условиям теоремы 1. Предельные вероятности в этом случае можно найти с помощью системы уравнений (9).

Задачи

1. Приведите примеры, показывающие, что, вообще говоря, равенства

$$\mathrm{P}(A\,|\,B)+\mathrm{P}(A\,|\,\overline{B})=1,\ \mathrm{P}(A\,|\,B)+\mathrm{P}(\overline{A}\,|\,\overline{B})=1$$

неверны.

2. В урне *M* шаров, из которых M_1 белого цвета. Рассматривается выборка объема *n*. Пусть B_j – событие, состоящее в том, что извлеченный на *j*-м шаге шар белого цвета, а A_k – событие состоящее в том, что в выборке объема *n* имеется ровно *k* белых шаров. Покажите, что как для выбора с возвращением, так и для выбора без возвращения

$$\mathrm{P}(B_j\,|\,A_k)=\frac{k}{n}.$$

3. Пусть $A_1,\ \dots,\ A_n$ – независимые события. Докажите, что

$$\mathrm{P}\left(\bigcup_{k=1}^{n}A_k\right)=1-\prod_{k=1}^{n}\mathrm{P}\left(\overline{A}_k\right).$$

4. Пусть $A_1,\ \dots,\ A_n$ – независимые события. Докажите, что вероятность того, что ни одно из этих событий не произойдет равна

$$\prod_{k=1}^{n}\left(1-\mathrm{P}\left(A_k\right)\right).$$

5. Пусть A и B независимые события. В терминах вероятностей этих событий выразите вероятности событий, состоящих в том, что произойдет в точности k, по меньшей мере k и самое большее k из событий A и B $\left(k=0,\ 1,\ 2\right)$.

6. Пусть событие A таково, что оно не зависит от самого себя. Покажите, что тогда $\mathrm{P}\left(A\right)$ равна 0 или 1.

7. Пусть A и B – независимые события и $\mathrm{P}\left(A\bigcup B\right)=1$. Докажите, что либо A, либо B имеет вероятность, равную 1.

8. Предположим, что события A и B_1 независимы и события A и B_2 также независимы. Покажите, что события A и $B_1\bigcup B_2$ независимы тогда и только тогда, когда события A и $B_1\bigcap B_2$ независимы.

9. Монета бросается n раз. Событие A состоит в том, что герб выпадет не более одного раза. Событие B состоит в том, что герб и решка выпадает не менее одного раза каждый. Покажите, что:

1) при $n=2$ события A и B зависимы,
2) при $n=3$ события A и B независимы,
3) при $n=4$ события A и B зависимы.

10. Докажите, что в теореме Пуассона имеет место следующая скорость сходимости:

$$\sup_k\left|P_n\left(k\right)-\frac{a^k e^{-a}}{k!}\right|\le\frac{2a^2}{n}.$$

11. Два игрока независимым образом подбрасывают (каждый свою) симметричные монеты. Покажите, что вероятность того, что у каждого после n подбрасываний будет одно и тоже число гербов, равна

$$2^{-2n}\sum_{k=0}^{n}\left(C_n^k\right)^2.$$

Выведите равенство

$$C_{2n}^n=\sum_{k=0}^{n}\left(C_n^k\right)^2.$$

12. Игральная кость бросается n раз. Найдите вероятность того, что последовательность выпавших очков является монотонно убывающей.

Глава 3. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

В предыдущих главах в качестве математических моделей исследуемых случайных явлений рассматривались вероятностные пространства $(\Omega, \mathscr{A}, \mathrm{P})$. Однако теория вероятностей не достигла бы такого расцвета и не была бы столь широко используема на практике, если бы изучала только случайные события. Возвращаясь к истокам возникновения теории вероятностей, вспомним, что уже в азартных играх интерес играющих вызывает не наступление случайного исхода, а связанный с ним выигрыш или проигрыш, т. е. определенная числовая величина, поставленная в соответствие этому исходу. Поэтому вполне естественной является задача определения и развития теории функций на вероятностных пространствах.

§ 1. Случайные величины и их распределения

Одним из основных понятий теории вероятностей является понятие случайной величины. Случайная величина – это величина, принимающая те или иные значения, в зависимости от случая. Примерами случайных величин могут быть число очков, выпадающих при однократном подбрасывании игральной кости, число бракованных среди взятых наугад n изделий, число попаданий в цель при n выстрелах и т. д. Случайная величина ξ есть число, которое ставится в соответствие каждому возможному исходу эксперимента, т. е. ее можно рассматривать как функцию $\xi = \xi(\omega)$ на пространстве элементарных событий Ω. Следует отметить, что при изучении случайных величин нам часто придется отвечать на вопрос: какова вероятность того, что значения случайной величины $\xi(\omega)$ принадлежат тому или иному множеству? Следовательно, для достаточно широкого класса множеств $\{B\}$ на числовой прямой мы должны быть уверены, что множество $\{\omega : \xi(\omega) \in B\}$ принадлежит σ-алгебре $\mathscr{A}$, и поэтому можно рассматривать вероятность $\mathrm{P}\{\omega \in \Omega : \xi(\omega) \in B\}$. Оказывается (в этом убедимся далее), достаточно предположить, что для любого $c \in R$ было выполнено $\{\omega \in \Omega : \xi(\omega) \le c\} \in \mathscr{A}$, чтобы для любого борелевского множества B выполнялось $\{\omega \in \Omega : \xi(\omega) \in B\} \in \mathscr{A}$. В силу вышесказанного, естественным является следующее определение.

Определение 1. *Пусть* $(\Omega, \mathscr{A}, \mathrm{P})$ *– произвольное вероятностное пространство. Случайной величиной называется функция* $\xi : \Omega \to R$, *такая что для любого* $c \in R$

$$\{\omega \in \Omega :\ \xi(\omega) \le c\} \in \mathscr{A}. \qquad (1)$$

Таким образом, случайная величина – это числовая измеримая функция. Напомним, что условие (1) – условие измеримости функции.

В дальнейшем будем использовать также другое определение случайной величины.

Определение 2. *Случайной величиной называется функция* $\xi : \Omega \to R$, *такая, что для любого борелевского множества* $B \in \mathscr{B}(R)$

$$\xi^{-1}(B) \in \mathscr{A}. \qquad (2)$$

Здесь по определению

$$\xi^{-1}(B) = \{\omega \in \Omega :\ \xi(\omega) \in B\}.$$

Докажем эквивалентность этих определений. Для этого понадобятся факты, которые сформулируем в виде упражнения.

Упражнение 1. Докажите, что справедливы равенства:

1) $\xi^{-1}(\varnothing) = \varnothing$, 2) $\xi^{-1}(R) = \Omega$, 3) $\xi^{-1}\left(\bigcap\limits_{k=1}^{\infty} A_k\right) = \bigcap\limits_{k=1}^{\infty} \xi^{-1}(A_k)$,

4) $\xi^{-1}\left(\bigcup\limits_{k=1}^{\infty} A_k\right) = \bigcup\limits_{k=1}^{\infty} \xi^{-1}(A_k)$, 5) $\xi^{-1}(A \setminus B) = \xi^{-1}(A) \setminus \xi^{-1}(B)$.

Утверждение 1. *Определения* 1 *и* 2 *эквивалентны.*

Доказательство. Покажем, что условие (1) эквивалентно условию (2).

Необходимость. Пусть выполняется условие (1). Рассмотрим класс K тех множеств $B \in R$, для которых $\xi^{-1}(B) \in \mathscr{A}$, т. е.

$$K = \{B :\ B \subseteq R,\ \ \xi^{-1}(B) \in \mathscr{A}\}.$$

Из упражнения 1 вытекает, что K – σ-алгебра и любые полуинтервалы $(a;b] \in K$. $\mathscr{B}(R)$ по определению наименьшая σ-алгебра, содержащая все полуинтервалы вида $(a;b]$. Таким образом, $\mathscr{B}(R) \subseteq K$ и, следовательно, для любого $B \in \mathscr{B}(R)$ $\xi^{-1}(B) \in \mathscr{A}$, а значит, справедливо (2).

Достаточность. Пусть имеет место (2). Тогда, взяв в качестве борелевского множества B полуинтервал $(-\infty;c]$, где c – произвольное действительное число, приходим к справедливости условия (1). Таким образом, утверждение 1 доказано полностью.

П р и м е р 1. Бросается симметричная монета. Вероятностное пространство $(\Omega, \mathscr{A}, \mathrm{P})$ зададим таким образом: $\Omega = \{г, р\}$, где $\{г\}$ – событие, состоящее в том, что выпал герб, $\{р\}$ – событие, состоящее в том, что выпала решка. В таком случае $\mathscr{A} = \{\varnothing, \Omega, \{г\}, \{р\}\}$ и $\mathrm{P}(\{г\}) = \mathrm{P}(\{р\}) = 1/2$. На этом вероятностном пространстве определим функцию $\xi: \Omega \to R$, $\xi(\{г\}) = 0, \xi(\{р\}) = 1$. Убедимся в том, что ξ – случайная величина. Несложно видеть, что для любых $c < 0$

$$\{\omega \in \Omega: \xi(\omega) \le c\} = \varnothing \in \mathscr{A},$$

если $0 \le c < 1$, то

$$\{\omega \in \Omega: \xi(\omega) \le c\} = \{г\} \in \mathscr{A},$$

и, наконец, для любых $c \ge 1$

$$\{\omega \in \Omega: \xi(\omega) \le c\} = \Omega \in \mathscr{A},$$

т. е. выполняется условие измеримости (1) и, следовательно, ξ – случайная величина.

В качестве модельных вероятностных пространств рассматриваемого примера могут быть взяты различные вероятностные пространства. Графики исходной случайной величины ξ будут зависеть от их вида. Рассмотрим два примера

1) $\Omega = [0;1]$, $\{г\} = [0;1/2]$, $\{р\} = (1/2;1]$, P – мера Лебега на $[0;1]$;

2) $\Omega = [0;1]$, $\{г\} = [1/4;3/4]$, $\{р\} = [0;1/4) \cup [3/4;1]$, P – мера Лебега на $[0;1]$.

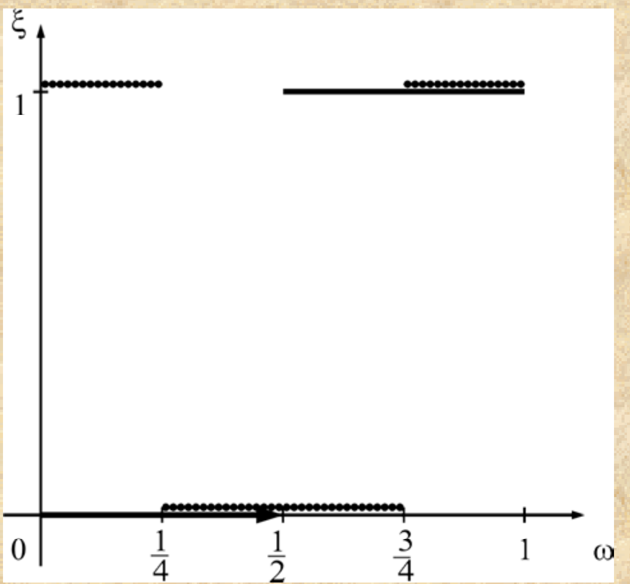

Рис. 1

Тогда график случайной величины ξ можно изобразить двумя способами (см. рис. 1).

Более того, если учесть, что в качестве множества $\{г\}$ можно взять любое подмножество отрезка $[0;1]$, мера Лебега которого равна $1/2$, а в качестве $\{р\}$ – его дополнение: $\{р\} = [0;1] \setminus \{г\}$, то придем к тому, что одна и та же случайная величина ξ имеет бесконечное число графиков. Заметим также, что множество $\{г\}$ на отрезке $[0;1]$ выбирается случайным образом (в силу специфики исходной задачи), поэтому элемент случайности присущ и самим графикам.

Подытоживая все вышесказанное, приходим к выводу: функция ξ – числовая измеримая функция, значения которой в силу случайного расположения элементов из области определения Ω являются случайными числами.

Упражнение 2. Докажите, что функция ξ из примера 1 удовлетворяет соотношению (2).

Пример 2. На отрезок $[0;1]$ числовой прямой наугад бросают частицу, причем считают, что все положения частицы одинаково возможны.

Вероятностное пространство $(\Omega, \mathcal{A}, \mathrm{P})$ зададим стандартным образом: $\Omega = [0;1]$, $\mathcal{A} = \mathcal{B}([0;1])$ – борелевская σ-алгебра на [0;1], P – мера Лебега.

Определим функцию $\xi: \Omega \to R$ следующим образом: $\xi(\omega) = \omega$ для любых $\omega \in \Omega$. Заметим, что

$$\{\omega \in \Omega: \xi(\omega) \le c\} = \begin{cases} \varnothing \in \mathcal{A}, & c < 0, \\ [0;c] \in \mathcal{A}, & 0 \le c < 1, \\ \Omega \in \mathcal{A}, & c \ge 1 \end{cases}$$

т. е. ξ – случайная величина.

Пример 3. Пусть $(\Omega, \mathcal{A}, \mathrm{P})$ – произвольное вероятностное пространство, A – некоторое случайное событие ($A \in \mathcal{A}$) и

$$I_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

Так как

$$\{\omega \in \Omega: \xi(\omega) \le c\} = \begin{cases} \varnothing, & c < 0, \\ \overline{A}, & 0 \le c < 1, \\ \Omega, & c \ge 1, \end{cases}$$

то $\{\omega \in \Omega: \xi(\omega) \le c\} \in \mathcal{A}$ для любых $c \in R$,, т. е. I_A – случайная величина.

Случайную величину I_A называют *индикатором случайного события A.*

Упражнение 3. Покажите, что, если в прим. 3 $A \notin \mathcal{A}$, то функция I_A не является случайной величиной.

Из прим. 1 было видно, что, хотя случайная величина и является числовой измеримой функцией, работать с ней как с обычной (неслучайной) функцией нельзя. Сейчас определим функцию распределения случайной величины, которая, с одной стороны, является неслучайной функцией, а с другой – несет всю информацию, заложенную в случайной величине.

Определение 3. *Функцией распределения случайной величины ξ называется функция $F_\xi: R \to [0;1]$, такая, что для любых $x \in R$*

$$F_\xi(x) = \mathrm{P}\{\omega \in \Omega: \xi(\omega) \le x\}. \tag{3}$$

Замечание 1. Корректность определения 3 непосредственно вытекает из (1).

П р и м е р 4. Найдем функцию распределения случайной величины ξ из прим. 1 и построим ее график. Несложно видеть, что для $x < 0$

$$\{\omega \in \Omega : \ \xi(\omega) \le x\} = \varnothing,$$

если же $0 \le x < 1,$ то

$$\{\omega \in \Omega : \ \xi(\omega) \le x\} = \{\text{г}\},$$

и, наконец, для $x \ge 1$

$$\{\omega \in \Omega : \ \xi(\omega) \le x\} = \Omega,$$

поэтому по определению функции распределения имеем

$$F_\xi(x) = \begin{cases} 0, & x \in (-\infty; 0), \\ 1/2, & x \in [0; 1), \\ 1, & x \in [1; +\infty). \end{cases}$$

График функции $F_\xi(x)$ имеет вид

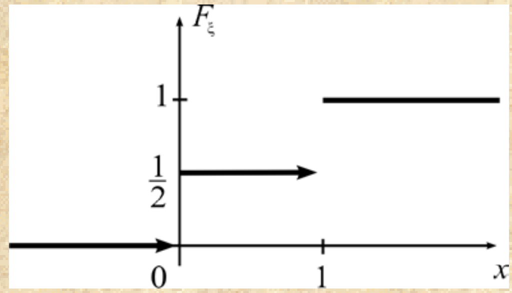

Рис. 2

У п р а ж н е н и е 4. Найдите функции распределения случайных величин из прим. 2 и 3 и постройте их графики.

На практике часто бывают полезными соотношения из следующего упражнения.

У п р а ж н е н и е 5. Докажите, что для любых $a, b \in R$, $a < b$, справедливы следующие равенства:

1) $\mathrm{P}(a < \xi \le b) = F_\xi(b) - F_\xi(a)$;

2) $\mathrm{P}(\xi < b) = F_\xi(b - 0)$;

3) $\mathrm{P}(a \le \xi \le b) = F_\xi(b) - F_\xi(a - 0)$;

4) $\mathrm{P}(\xi = a) = F_\xi(a) - F_\xi(a - 0)$;

5) $\mathrm{P}(a \le \xi < b) = F_\xi(b - 0) - F_\xi(a - 0)$;

6) $\mathrm{P}(a < \xi < b) = F_\xi(b - 0) - F_\xi(a)$.

Р е ш е н и е. 1) Используя свойство 2 (гл. 1, § 2) вероятностей и пункт 5 упражнения 1, имеем

$$\mathrm{P}(a < \xi \le b) = \mathrm{P}(\omega \in \Omega : a < \xi(\omega) \le b) = \mathrm{P}\left(\omega \in \Omega : \xi^{-1}\left((a; b]\right)\right) =$$

$$= \mathrm{P}\left(\xi^{-1}\left(]-\infty;b] \setminus]-\infty;a]\right)\right) = \mathrm{P}\left(\xi^{-1}\left(]-\infty;b]\right)\right) - \mathrm{P}\left(\xi^{-1}\left(]-\infty;a]\right)\right) =$$

$$= \mathrm{P}(\xi \le b) - \mathrm{P}(\xi \le a) = F_\xi(b) - F_\xi(a).$$

2) Несложно видеть, что

$$\{\omega \in \Omega : \xi(\omega) < b\} = \sum_{n=1}^{\infty} \left\{\omega \in \Omega : b - \frac{1}{n-1} < \xi(\omega) \le b - \frac{1}{n}\right\},$$

поэтому по аксиоме счетной аддитивности вероятностей и доказанному выше п. 1) получаем

$$\mathrm{P}(\xi < b) = \mathrm{P}(\omega \in \Omega : \xi(\omega) < b) = \mathrm{P}(\omega \in \Omega : \xi(\omega) \le b - 1) +$$

$$+ \sum_{n=2}^{\infty} \mathrm{P}\left(\omega \in \Omega : b - \frac{1}{n-1} < \xi(\omega) \le b - \frac{1}{n}\right) = F_\xi(b-1) +$$

$$+ \lim_{N\to\infty} \sum_{n=2}^{N} \left(F_\xi\left(b - \frac{1}{n}\right) - F_\xi\left(b - \frac{1}{n-1}\right)\right) = \lim_{N\to\infty} F_\xi\left(b - \frac{1}{N}\right) = F_\xi(b-0).$$

Доказательства остальных пунктов этого упражнения непосредственно вытекают из рассмотренных выше первых двух.

В следующей теореме мы рассмотрим свойства функции распределения.

Теорема 1. *Функция распределения* $F_\xi(x)$, $x \in R$, *обладает следующими свойствами:*

1) F_ξ *монотонно неубывающая функция*;

2) $F_\xi(x)$ *непрерывна справа для любых* $x \in R$;

3) $F_\xi(-\infty) = 0$;

4) $F_\xi(+\infty) = 1$.

З а м е ч а н и е 2. В п. 3 и 4 по определению полагаем

$$F_\xi(-\infty) = \lim_{x\to-\infty} F_\xi(x), \quad F_\xi(+\infty) = \lim_{x\to+\infty} F_\xi(x).$$

Д о к а з а т е л ь с т в о. 1) Убедимся в том, что для любых $x_1 < x_2$ $F_\xi(x_1) \le F_\xi(x_2)$.

Используя п. 4 упражнения 1 и аксиому аддитивности вероятностей, имеем

$$F_\xi(x_2) = \mathrm{P}(\xi \le x_2) = \mathrm{P}\left(\xi^{-1}\left(]-\infty;x_2]\right)\right) = \mathrm{P}\left(\xi^{-1}\left(]-\infty;x_1] +]x_1;x_2]\right)\right) =$$

$$= \mathrm{P}\left(\xi^{-1}\left(]-\infty;x_1]\right) + \xi^{-1}\left(]x_1;x_2]\right)\right) = \mathrm{P}(\xi \le x_1) + \mathrm{P}(x_1 < \xi \le x_2) =$$

$$= F_\xi(x_1) + \mathrm{P}(x_1 < \xi \le x_2),$$

поэтому по аксиоме неотрицательности вероятностей $F_\xi(x_2) \ge F_\xi(x_1)$.

2) Покажем, что для любого $x \in R$

$$\lim_{n\to\infty} F_\xi\left(x+\frac{1}{n}\right)=F_\xi(x).$$

Рассмотрим события $B_n=\left\{\omega\in\Omega:\ x<\xi(\omega)\le x+\frac{1}{n}\right\}$. Очевидно, что

$$B_1 \supseteq B_2 \supseteq \ldots;\quad \bigcap_{n=1}^{\infty} B_n=\varnothing.$$

По аксиоме непрерывности вероятностей

$$\mathrm{P}(B_n)\xrightarrow[n\to\infty]{}0,$$

но

$$\mathrm{P}(B_n)=\mathrm{P}\left(x<\xi(\omega)\le x+\frac{1}{n}\right)=F_\xi\left(x+\frac{1}{n}\right)-F_\xi(x).$$

Следовательно,

$$\lim_{n\to\infty} F_\xi\left(x+\frac{1}{n}\right)=F_\xi(x).$$

3), 4) Очевидно равенство

$$\Omega=\sum_{n=-\infty}^{n=+\infty}\left\{\omega\in\Omega:\ n-1<\xi(\omega)\le n\right\},$$

поэтому, учитывая аксиому σ-аддитивности вероятностей и п. 1 упражнения 4, получаем

$$1=\mathrm{P}(\Omega)=\sum_{n=-\infty}^{\infty}\mathrm{P}\left(\left\{n-1<\xi(\omega)\le n\right\}\right)=\lim_{N\to\infty}\sum_{n=-N+1}^{N}\mathrm{P}\left(n-1<\xi\le n\right)=$$

$$=\lim_{N\to\infty}\sum_{n=-N+1}^{N}\left(F_\xi(n)-F_\xi(n-1)\right)=\lim_{N\to\infty}\left(F_\xi(N)-F_\xi(-N)\right)=$$

$$=F_\xi(+\infty)-F_\xi(-\infty).$$

Свойства 3) и 4) доказаны.

Теорема 2 [29]. *Если функция F удовлетворяет свойствам* 1)–4) *функции распределения, то она является функцией распределения некоторой случайной величины.*

З а м е ч а н и е 3. Очевидно, что множество точек разрыва у функции распределения не более, чем счетно.

Действительно, число точек разрыва, величина скачков в которых больше $1/n$, не превышает n, $n=1,\ 2,\ \ldots$. Пронумеруем сначала точки разрыва, величина скачков которых лежит в промежутке $(1/2;1]$, затем в промежутке $(1/3;1/2]$ и т. д.

В некоторых случаях бывает удобно пользоваться следующей характеристикой случайной величины.

О п р е д е л е н и е 4. *Распределением вероятностей случайной величины* ξ *называется функция*

$$P_\xi : \mathscr{B}(R) \to [0;1]$$

такая, что для любого $B \in \mathscr{B}(R)$

$$P_\xi(B) = \mathrm{P}(\xi \in B).$$

П р и м е р 5. Найдем распределение вероятностей случайной величины из прим. 1:

1) пусть $\{0\} \in B,\ \{1\} \in B$, тогда $P_\xi(B) = 1$;

2) если $\{0\} \in B,\ \{1\} \notin B$, то $P_\xi(B) = 1/2$;

3) в случае, когда $\{0\} \notin B,\ \{1\} \in B\ \ P_\xi(B) = 1/2$;

4) если $\{0\} \notin B,\ \{1\} \notin B$, то $P_\xi(B) = 0$.

У п р а ж н е н и е 6. Найдите распределения вероятностей случайных величин из прим. 2 и 3.

Теорема 3 [29]. *Распределение вероятностей* P_ξ *однозначно определяется по функции распределения* F_ξ.

З а м е ч а н и е 4. Из данной теоремы следует, что функция распределения F_ξ и распределение вероятностей P_ξ взаимно однозначно определяют друг друга (то, что F_ξ однозначно определяется по P_ξ очевидно, для этого достаточно в качестве борелевских множеств B брать полуинтервалы $(-\infty; x]$, $x \in R$).

Всюду в дальнейшем по определению будем считать, что *задать случайную величину* – это значит задать ее функцию распределения или распределение вероятностей. Таким образом, каждая случайная величина дает такое отображение $\xi = \xi(\omega)$ множества Ω в числовую прямую R, которое порождает новое вероятностное пространство $(R, \mathscr{B}(R), P_\xi)$, поэтому в дальнейшем природа вероятностного пространства нас интересовать не будет.

§ 2. Классификация случайных величин

О п р е д е л е н и е 1. *Говорят, что случайная величина* ξ *имеет распределение* Q, *если* $Q(B) = P_\xi(B)$ *при любом* $B \in \mathscr{B}(R)$.

В этом параграфе разобьем множество всех случайных величин на три класса в зависимости от типа их распределения.

1. Дискретные случайные величины (распределения)

О п р е д е л е н и е 2. *Случайная величина* ξ *называется дискретной, если она принимает не более чем счетное число значений.*

Задание такой случайной величины по определению равносильно заданию *закона распределения*:

$$\begin{array}{lllll} \xi: & x_1, & x_2, & \ldots, & x_n, & \ldots \\ \mathrm{P}: & p_1, & p_2, & \ldots, & p_n, & \ldots, \end{array} \qquad (1)$$

где $p_n = \mathrm{P}(\xi = x_n)$, $\sum\limits_n p_n = 1$.

Следующее утверждение отражает связь между функцией распределения и законом распределения.

У т в е р ж д е н и е 1. *Закон распределения и функция распределения взаимно однозначно определяют друг друга.*

Д о к а з а т е л ь с т в о. Пусть ξ имеет закон распределения (1). Найдем функцию распределения. При этом будем полагать, что

$$x_1 < x_2 < \ldots < x_n < \ldots .$$

Тогда по определению функции распределения имеем

$$F_\xi(x) = \begin{cases} 0, & \text{если } x < x_1, \\ p_1, & \text{если } x_1 \le x < x_2, \\ p_1 + p_2, & \text{если } x_2 \le x < x_3, \\ p_1 + p_2 + p_3, & \text{если } x_3 \le x < x_4, \\ \ldots . \end{cases}$$

Следующий график (рис. 3) дает представление о найденной функции $F_\xi(x)$.

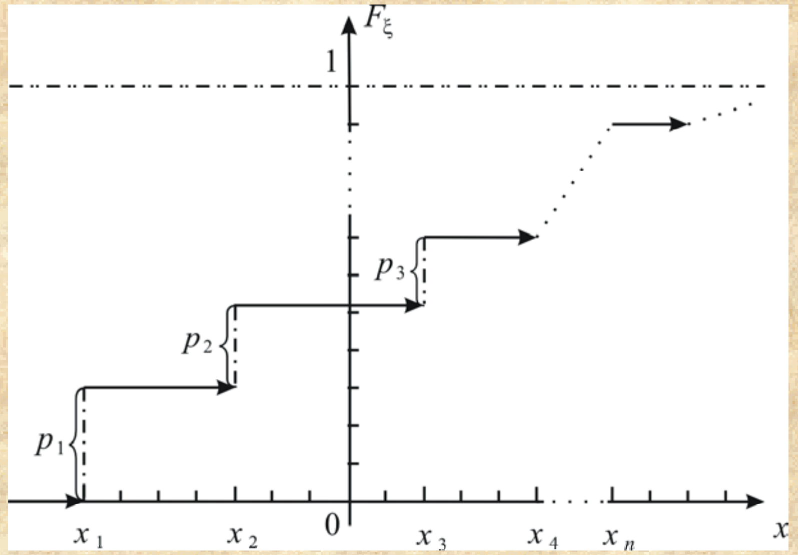

Рис. 3

Пусть теперь известна функция распределения дискретной случайной величины ξ. Тогда, учитывая, что $P(\xi = x) = F_\xi(x) - F_\xi(x-0)$ (см. § 1, п. 4 упражнения 5), легко восстанавливаем закон распределения

$$\begin{array}{lllll} \xi: & x_1 & x_2 & \dots & x_n \quad \dots \\ \mathrm{P}: & F_\xi(x_1) - F_\xi(x_1 - 0) & F_\xi(x_2) - F_\xi(x_2 - 0) & \dots & F_\xi(x_n) - F_\xi(x_n - 0) \dots . \end{array}$$

П р и м е р 1. Найдем закон распределения случайной величины, если известна функция распределения этой величины

$$F_\xi(x) = \begin{cases} 0, & x < 0, \\ 1/2, & 0 \le x < 1, \\ 1, & x \ge 1. \end{cases}$$

Опираясь на выше доказанное утверждение, записываем закон распределения

$$\begin{array}{lll} \xi: & 0 & 1 \\ \mathrm{P}: & 1/2 = F_\xi(0) - F_\xi(0-0) & 1/2 = F_\xi(1) - F_\xi(1-0). \end{array}$$

Примеры дискретных случайных величин (распределений)

1. С л у ч а й н а я в е л и ч и н а Б е р н у л л и (р а с п р е д е л е н и е Б е р н у л л и). Закон распределения имеет вид:

$$\begin{array}{lll} \xi: & 0 & 1, \\ \mathrm{P}: & 1-p & p, \qquad (0 < p < 1). \end{array}$$

Такому распределению соответствует бросание монеты, на одной стороне которой – 0, а на второй – 1.

Иллюстрация в Mathematica

2. Биномиальная случайная величина (биномиальное распределение). Закон распределения запишется следующим образом

$$\begin{array}{llllllll} \xi: & 0 & 1 & \dots & k & \dots & n \\ \mathrm{P}: & P_n(0) & P_n(1) & \dots & P_n(k) & \dots & P_n(n), \end{array}$$

где $P_n(k) = C_n^k \cdot p^k \cdot q^{n-k}$, $0 < p < 1$, $q = 1 - p$.

Число успехов в n испытаниях схемы Бернулли имеет биномиальное распределение.

Иллюстрация в Mathematica

3. Случайная величина Пуассона (распределение Пуассона с параметром λ). Закон распределения задается следующим образом:

$$\mathrm{P}(\xi = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}, \quad k = 0,\ 1,\ 2,\ 3,\ \dots,$$

где $\lambda > 0$ – параметр.

Распределение Пуассона носит название *закона редких событий*, поскольку оно всегда появляется там, где производится большое число испытаний, в каждом из которых с малой вероятностью происходит «редкое» событие. По закону Пуассона распределены, например, число вызовов, поступивших на телефонную станцию, число распавшихся нестабильных частиц и т. д.

Иллюстрация в Mathematica

4. Геометрическая случайная величина (геометрическое распределение). Закон распределения имеет вид:

$$\mathrm{P}(\xi = k) = p(1 - p)^{k-1}, \ 0 < p < 1, \quad k = 1,\ 2,\ \dots .$$

Пусть производятся независимые испытания, причем в каждом испытании возможны два исхода – «успех» с вероятностью p или «неуспех» с вероятностью $1 - p$, $0 < p < 1$. Обозначим через ξ число испытаний до первого появления «успеха», тогда ξ будет геометрической случайной величиной.

Иллюстрация в Mathematica

5. Гипергеометрическая случайная величина (гипергеометрическое распределение). Закон распределения имеет вид:

$$\mathrm{P}(\xi = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, \ \ k = \max(0, M + n - N), \ \dots \ \min(M, n).$$

Пусть имеется N деталей, среди них находится M бракованных. Выбираем наудачу n деталей. Обозначим через ξ число бракованных деталей среди выбранных, тогда ξ будет гипергеометрической случайной величиной.

Иллюстрация в Mathematica

6. Отрицательная биномиальная случайная величина (отрицательное биномиальное распределение). Закон распределения запишется следующим образом

$$\mathrm{P}(\xi = k) = C_{n+k-1}^{n-1} p^n (1-p)^k, \quad 0 < p < 1, \quad k = 0,\ 1,\ \ldots .$$

Данное распределение имеет число «неуспехов» в испытаниях схемы Бернулли до появления n-ого «успеха».

Иллюстрация в Mathematica

7. Равномерная дискретная случайная величина (дискретное равномерное распределение). Закон распределения запишется следующим образом

$$\begin{array}{llllllll} \xi: & 1 & 2 & \ldots & k & \ldots & n \\ \mathrm{P}: & \dfrac{1}{n} & \dfrac{1}{n} & \ldots & \dfrac{1}{n} & \ldots & \dfrac{1}{n}. \end{array}$$

Представляет собой распределение n равновероятных событий.

Иллюстрация в Mathematica

8. Логарифмического ряда случайная величина (логарифмического ряда распределение). Закон распределения имеет вид:

$$\mathrm{P}(\xi = k) = -\frac{\theta^k}{k \ln(1-\theta)}, \quad 0 < \theta < 1, \quad k = 1,\ 2,\ \ldots .$$

Данное распределение возникает в некоторых задачах экономики.

Иллюстрация в Mathematica

2. Абсолютно непрерывные случайные величины (распределения)

Определение 3. *Распределение случайной величины ξ называется абсолютно непрерывным, а сама случайная величина – абсолютно непрерывной случайной величиной, если для любого $B \in \mathcal{B}(R)$*

$$P_\xi(B) = \int_B p_\xi(x)\,dx, \tag{2}$$

где $p_\xi(x)$, $x \in R$, – интегрируемая по Лебегу функция. Функция $p_\xi(x)$ называется плотностью распределения случайной величины ξ.

Теорема 1 [29]. *Для того чтобы случайная величина* ξ *была абсолютно непрерывной, необходимо и достаточно, чтобы для любого* $x \in R$

$$F_{\xi}(x) = \int_{-\infty}^{x} p_{\xi}(t)\,dt. \tag{3}$$

Замечание 1. Из представления (3) видно, что функция распределения абсолютно непрерывной случайной величины является абсолютно непрерывной функцией.

Свойства плотности распределения

1) $\int_{-\infty}^{+\infty} p_{\xi}(t)\,dt = 1$ (вытекает из того, что $F_{\xi}(+\infty) = 1$).

2) $p_{\xi}(t) \geq 0$ почти всюду (следует из того, что F_{ξ} – монотонно неубывающая функция).

3) $F'_{\xi}(x) = p_{\xi}(x)$ для любых x, являющихся точками непрерывности плотности.

Теорема 2. *Для того чтобы функция* $p = p(x)$ *была плотностью распределения некоторой случайной величины* ξ, *необходимо и достаточно, чтобы она удовлетворяла свойствам* 1) *и* 2) *плотности.*

Доказательство. Если функция $p(x)$ удовлетворяет свойствам 1) и 2) плотности распределения, то функция

$$F(x) = \int_{-\infty}^{x} p(t)\,dt$$

удовлетворяет свойствам 1)–4) функции распределения (теорема 1, § 1, гл. 3). Таким образом, по теореме 2 (гл. 3, § 1) существует случайная величина ξ, для которой функция F будет функцией распределения. Тогда p будет плотностью распределения ξ.

Примеры абсолютно непрерывных случайных величин (распределений)

1. Нормальная случайная величина, или случайная величина Гаусса (нормальное распределение). Случайная величина ξ имеет *нормальное распределение (распределение Гаусса)*, если ее плотность распределения имеет вид:

$$p_{\xi}(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad \sigma > 0, \quad a \in R.$$

Если $a=0$, $\sigma=1$, то данное распределение называется *стандартным нормальным распределением*.

С плотностью и функцией распределения стандартного нормального распределения мы уже встречались в локальной и интегральной теоремах Муавра – Лапласа. Важная роль этого распределения объясняется тем, что оно обычно возникает в явлениях, подверженных действию большого числа малых случайных явлений. Так математическая теория выборочного метода в статистике для расчета некоторых показателей (ошибка выборки, доверительный интервал, связь между признаками) широко использует нормальное распределение.

Иллюстрация в Mathematica

2. Экспоненциальная (показательная) случайная величина (экспоненциальное распределение). Случайная величина ξ имеет *экспоненциальное (показательное) распределение с параметром* λ *(*$\lambda>0$*)*, если ее плотность имеет вид:

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

Экспоненциальному распределению подчиняется время распада атомов различных элементов. Оно обладает важным свойством – *отсутствием последействия*. Несложно убедиться в том, что вероятность распада атома за время x_2, при условии, что перед этим он уже прожил время x_1, совпадает с безусловной вероятностью распада того же самого атома за время x_2. Именно это свойство и представляет собой отсутствие последействия.

Иллюстрация в Mathematica

3. Равномерная на $[a;b]$ случайная величина (равномерное на отрезке $[a;b]$ распределение). *Равномерно распределенная на отрезке* $[a;b]$ *случайная величина* ξ имеет плотность распределения

$$p_\xi(x)=\begin{cases}0, & x\notin[a;b],\\ \dfrac{1}{b-a}, & x\in[a;b].\end{cases}$$

Равномерное распределение реализует принцип геометрической вероятности при бросании точки на отрезок $[a;b]$.

Иллюстрация в Mathematica

4. Случайная величина Коши (распределение Коши). Случайная величина ξ имеет *распределение Коши*, если ее плотность представима в виде:

$$p_\xi(x) = \frac{1}{\pi} \cdot \frac{c}{x^2 + c^2}, \quad c > 0, \quad x \in R.$$

Распределение Коши совпадает с распределением вероятностей отношения ξ_1/ξ_2, где ξ_1, ξ_2 – независимые стандартные нормальные случайные величины. Такое же распределение имеет $\mathrm{tg}\,\alpha$ – тангенс случайной величины α, равномерно распределенной на отрезке $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$.

Иллюстрация в Mathematica

5. Гамма-распределение. Случайная величина ξ имеет *гамма-распределение с параметрами* α *и* λ *(*$\alpha, \lambda > 0$*)*, если ее плотность имеет вид:

$$p_\xi(x) = \begin{cases} 0, & x \le 0, \\ \dfrac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & x > 0. \end{cases}$$

Здесь $\Gamma(\alpha) = \int\limits_0^\infty y^{\alpha-1} e^{-y} dy$ – гамма-функция.

При $\alpha = 1$ получаем экспоненциальное распределение, а при $\alpha = n/2$, $\beta = 1/2$, где n целое положительное получаем хи-квадрат распределение с n степенями свободы (см. Дополнение к главе 5, §4).

Иллюстрация в Mathematica

6. Бета-распределение. Случайная величина ξ имеет *бета-распределение с параметрами* α *и* β ($\alpha, \beta > 0$), если ее плотность имеет вид:

$$p_\xi(x) = \begin{cases} 0, & x \notin (0;1), \\ \dfrac{1}{B(\alpha,\beta)} x^{\alpha-1} (1-x)^{\beta-1}, & x \in (0;1). \end{cases}$$

Здесь $B(\alpha,\beta) = \int\limits_0^1 y^{\alpha-1}(1-y)^{\beta-1} dy = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ – бета-функция.

Если η и ζ независимые случайные величины, имеющие гамма-распределение с одинаковыми параметрами, тогда случайная величина $\frac{\eta}{\eta+\zeta}$ имеет бета распределение с теми же параметрами.

Иллюстрация в Mathematica

7. Логарифмически нормальная (логнормальная) случайная величина (логнормальное распределение). Случайная величина ξ имеет *логарифмически нормальное распределение (логнормальное распределение)* с параметрами σ и μ $(\sigma > 0, \mu \in R)$, если ее плотность распределения имеет вид:

$$p_\xi(x) = \begin{cases} 0, & x < 0, \\ \dfrac{1}{\sigma x\sqrt{2\pi}} \cdot e^{-\frac{(\ln(x/\mu))^2}{2\sigma^2}}, & x \ge 0. \end{cases}$$

Такое распределение имеет неотрицательная случайная величина, натуральный логарифм которой, то есть $\ln\xi$, имеет нормальный закон распределения с дисперсией σ^2 и средним $\ln\mu$.

Иллюстрация в Mathematica

8. Отраженная нормальная случайная величина (отраженное нормальное распределение). Случайная величина ξ имеет *отраженное нормальное распределение* с параметром σ $(\sigma > 0)$, если ее плотность распределения имеет вид:

$$p_\xi(x) = \begin{cases} 0, & x < 0, \\ \dfrac{1}{\sigma}\sqrt{\dfrac{2}{\pi}} \cdot e^{-\frac{x^2}{2\sigma^2}}, & x \ge 0. \end{cases}$$

Такое распределение имеет модуль $|\eta|$ случайной величины η, подчиняющейся симметричному нормальному закону распределения с дисперсией σ^2 и средним $a = 0$.

Иллюстрация в Mathematica

9. Распределение экстремальных значений. Случайная величина ξ имеет *распределение экстремальных значений* с параметрами α *и* β $(\alpha \in R, \beta > 0)$, если ее плотность распределения имеет вид:

$$p_\xi(x) = \frac{e^{\frac{\alpha - x}{\beta} - e^{\frac{\alpha - x}{\beta}}}}{\beta}$$

Это предельное распределение для наименьшего или наибольшего значения в больших выборках.

Иллюстрация в Mathematica

10.Случайная величина Лапласа (распределение Лапласа). Случайная величина ξ имеет *распределение Лапласа*, если ее плотность распределения имеет вид:

$$p_\xi(x) = \frac{1}{2b} \cdot e^{-\frac{|x-a|}{b}}, \quad b > 0, \quad a \in R.$$

Такое распределение имеет разность двух независимых одинаково распределенных показательных случайных величин.

Иллюстрация в Mathematica

11.Логистическая случайная величина (логистическое распределение). Случайная величина ξ имеет *логистическое распределение*, если ее плотность имеет вид:

$$p_\xi(x) = \frac{e^{-\frac{x-a}{b}}}{b\left(1 + e^{-\frac{x-a}{b}}\right)^2}, \quad b > 0, \quad a \in R.$$

Данное распределение часто используется вместо нормального распределения при исследовании, например, медико-биологических объектов.

Иллюстрация в Mathematica

12.Случайная величина Парето (распределение Парето). Случайная величина ξ имеет *распределение Парето* с параметрами a и b ($a, b > 0$), если ее плотность распределения имеет вид:

$$p_\xi(x) = \begin{cases} 0, & x < a, \\ a^b b x^{-b-1}, & x \ge a. \end{cases}$$

Это распределение используют в экономике для описания величины дохода, причем a – минимально возможный доход.

Иллюстрация в Mathematica

13.Случайная величина Вейбулла-Гнеденко (распределение Вейбулла-Гнеденко). Случайная величина ξ имеет *распределение Вейбулла-Гнеденко* с параметрами α и λ ($\alpha, \lambda > 0$), если ее плотность распределения имеет вид:

$$p_\xi(x) = \begin{cases} 0, & x \le 0, \\ \alpha\lambda x^{\alpha-1} e^{-\lambda x^\alpha}, & x > 0. \end{cases}$$

Данное распределение используют в статистической теории надежности для описания распределения времени безотказной работы.

Иллюстрация в Mathematica

14. Случайная величина Рэлея (распределение Релея). Случайная величина ξ имеет *распределение Рэлея* с параметром σ^2, если ее плотность распределения имеет вид:

$$p_{\xi}(x)=\begin{cases}0, & x\le 0,\\ \dfrac{1}{\sigma^2}xe^{-\frac{x^2}{2\sigma^2}}, & x>0.\end{cases}$$

Если η и ζ независимые случайные величины, имеющие нормальное распределение с параметрами $a=0$ и σ^2, тогда случайная величина $\sqrt{\eta^2+\zeta^2}$ имеет распределение Рэлея с параметром σ^2. Это распределение находит широкое применение в теории стрельбы и статистической теории связи.

Иллюстрация в Mathematica

15. Сравнение нормальной и равномерной случайных величин.

Иллюстрация в Mathematica

3. Сингулярные случайные величины (распределения)

Точку $x\in R$ называют *точкой роста* функции распределения $F_{\xi}(x)$, если для каждого $\varepsilon>0$ выполняется неравенство

$$F_{\xi}(x+\varepsilon)-F_{\xi}(x-\varepsilon)>0.$$

Определение 4. *Случайная величина ξ имеет сингулярное распределение, если ее функция распределения непрерывная функция, множество точек роста которой имеет нулевую меру Лебега.*

Примером такой случайной величины является случайная величина, функция распределения которой – *функция Кантора*.

Таким образом, множество случайных величин (распределений) можно разбить на три класса: дискретные, абсолютно непрерывные и сингулярные величины (распределения). И эта классификация полная, что подтверждается следующей теоремой, которую мы приводим без доказательства.

Теорема 3 (**Лебег**). *Любая функция распределения* $F(x)$, $x\in R$, *единственным образом представима в виде*

$$F(x)=a_1F_1(x)+a_2F_2(x)+a_3F_3(x),$$

где $F_1(x)$ *– дискретная функция распределения,* $F_2(x)$ *– абсолютно непрерывная функция распределения,* $F_3(x)$ *– сингулярная функция распределения,* $a_i \ge 0$ *и* $a_1 + a_2 + a_3 = 1$.

В дальнейшем нам придется производить алгебраические операции над случайными величинами, в силу чего полезными могут оказаться понятия и факты, которые изложены далее.

Прежде всего, напомним, что функция $g: R^n \to R^m$ называется *борелевской функцией*, если из того, что $B \in \mathscr{B}(R^m)$ следует, что $g^{-1}(B) \in \mathscr{B}(R^n)$.

Упражнение 1. Докажите, что любая непрерывная функция является борелевской.

Утверждение 2. *Пусть* $g: R \to R$ *– борелевская функция,* ξ *– случайная величина, тогда* $g(\xi)$ *также является случайной величиной.*

Доказательство. Покажем, что $g(\xi): \Omega \to R$ – измеримая числовая функция. То, что $g(\xi)$ – числовая функция, очевидно. Проверим по определению 2 выполнимость условия измеримости, т. е., что

$$\left[g(\xi(\omega))\right]^{-1}(B) \in \mathscr{A}$$

для любых $B \in \mathscr{B}(R)$. Имеем

$$\left[g(\xi(\omega))\right]^{-1}(B) = \left\{\omega \in \Omega : g(\xi(\omega)) \in B\right\} =$$
$$= \left\{\omega \in \Omega : \xi(\omega) \in g^{-1}(B)\right\} \in \mathscr{A},$$

так как $g^{-1}(B) \in \mathscr{B}(R)$, то доказательство утверждения завершено.

Рассмотрим несколько примеров, в которых требуется найти функцию распределения функции от случайной величины.

Пример 2. Пусть функция распределения случайной величины ξ равна F_ξ. Найдем функцию распределения случайной величины $\eta = \xi^2$.

Решение. Имеем

$$F_\eta(x) = \mathrm{P}(\eta \le x) = \mathrm{P}(\xi^2 \le x) = \begin{cases} 0, & x < 0, \\ \mathrm{P}\left(-\sqrt{x} \le \xi \le \sqrt{x}\right), & x \ge 0. \end{cases}$$

Следовательно,

$$F_\eta(x)=\begin{cases}0, & x<0,\\ F_\xi\left(\sqrt{x}\right)-F_\xi\left(-\sqrt{x}-0\right), & x\ge 0.\end{cases}$$

Если случайная величина ξ имеет плотность распределения $p_\xi(x)$, то плотность распределения η равна

$$p_\eta(x)=\begin{cases}0, & x<0,\\ \dfrac{1}{2\sqrt{x}}\left[p_\xi\left(\sqrt{x}\right)-p_\xi\left(-\sqrt{x}\right)\right], & x>0.\end{cases}$$

П р и м е р 3. Пусть ξ случайная величина с непрерывной и строго монотонной функцией распределения F_ξ. Найдем функцию распределения случайной величины $\eta=F_\xi(\xi)$.

Р е ш е н и е. Так как при всех $x\in R$ $0\le F(x)\le 1$, то $F_\eta(x)=0$ при $x<0$ и $F_\eta(x)=1$ при $x\ge 1$. Пусть $0\le x\le 1$ и обозначим через z точку $z=F_\xi^{-1}(x)$ такую, что $F_\xi(z)=x$. Событие $\{\eta=F_\xi(\xi)\le x\}$ произойдет тогда и только тогда, когда произойдет событие $\{\xi\le z\}$. Следовательно,

$$F_\eta(x)=\begin{cases}0, & \text{если } x<0,\\ x, & \text{если } 0\le x<1,\\ 1, & \text{если } x\ge 1.\end{cases}$$

Таким образом, случайная величина η имеет равномерное на отрезке $[0;1]$ распределение.

П р и м е р 4. Пусть ξ равномерно распределена на отрезке $[0;1]$. Найдем функцию распределения случайной величины $\eta=\ln\dfrac{1}{\xi}$.

Р е ш е н и е. Так как $\ln\dfrac{1}{\xi}>0$, если $\xi\in(0;1)$, то $\mathrm{P}(\xi\le x)=0$ при $x<0$. Пусть $x\ge 0$, тогда

$$\mathrm{P}\left(\ln\frac{1}{\xi}\le x\right)=\mathrm{P}\left(\frac{1}{\xi}\le e^x\right)=1-e^{-x}.$$

Следовательно, η имеет экспоненциальное распределение с параметром $\lambda=1$.

§ 3. Многомерные случайные величины

Ясно, что на каждом вероятностном пространстве можно определить не одну случайную величину. На практике необходимость учета непредсказуемых воздействий также весьма редко приводит к рассмотрению только одной случайной величины.

Итак, в этом параграфе мы обобщим результаты § 2 на случай нескольких случайных величин.

Пусть $(\Omega,\ \mathscr{A},\ \mathrm{P})$ – произвольное вероятностное пространство, $\xi_1,\ \ldots,\ \xi_n$ – случайные величины, заданные на этом вероятностном пространстве.

О п р е д е л е н и е 1. *Совокупность*

$$\overline{\xi}(\omega)=\left(\xi_1(\omega),\ \xi_2(\omega),\ \ldots,\ \xi_n(\omega)\right)$$

называется n-мерной случайной величиной (случайным вектором).

О п р е д е л е н и е 2. *Функция* $\overline{\xi}:\ \Omega\to R^n$, *такая что для любого* $B\in\mathscr{B}\left(R^n\right)$

$$\overline{\xi}^{-1}(B)\in\mathscr{A},$$

называется n-мерной случайной величиной (случайным вектором).

У п р а ж н е н и е 1. Докажите эквивалентность определений 1 и 2.

П р и м е р 1. Пусть бросаются две игральные кости. Тогда двумерной случайной величиной является вектор $\overline{\xi}=(\xi_1,\ \xi_2)$, где ξ_i – число очков, выпавших на i-ой кости, $i=1,\ 2$.

П р и м е р 2. На квадрат $[0;1]\times[0;1]$ бросается частица. Декартовы координаты $(\xi_1,\ \xi_2)$ точки образуют двумерную случайную величину.

О п р е д е л е н и е 3. *Функция*

$$F_{\overline{\xi}}:\ R^n\to[0;1]$$

называется функцией распределения n-мерной случайной величины (случайного вектора) $\overline{\xi}=(\xi_1,\ \ldots,\ \xi_n)$, *если для любого вектора* $\overline{x}=(x_1,\ \ldots,\ x_n)\in R^n$

$$F_{\overline{\xi}}(\overline{x})=\mathrm{P}\left(\omega\in\Omega:\ \left(\xi_1(\omega)\le x_1\right)\bigcap\left(\xi_2(\omega)\le x_2\right)\bigcap\ldots\bigcap\left(\xi_n(\omega)\le x_n\right)\right).$$

З а м е ч а н и е 1. Функцию распределения n-мерной случайной величины также будем называть *многомерной функцией распределения.*

Свойства многомерной функции распределения

1°. $F_{\overline{\xi}}$ *монотонно не убывает по каждой переменной.*

2°. $F_{\overline{\xi}}$ *непрерывна справа по каждой переменной.*

3°. $F_{\overline{\xi}}(x_1,\ldots,x_{k-1},-\infty,x_{k+1},\ldots,x_n)=0$ *для любого* $k=\overline{1,n}.$

4°. (Свойство согласованности функции распределения)

$$F_{(\xi_1,\xi_2,\ldots,\xi_k,\ldots,\xi_n)}(x_1,\ldots,x_{k-1},+\infty,x_{k+1},\ldots,x_n)=$$
$$=F_{(\xi_1,\xi_2,\ldots,\xi_{k-1},\xi_{k+1},\ldots,\xi_n)}(x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)$$

для любого $k=\overline{1,n}$.

З а м е ч а н и е 2. Доказательство свойств 1°–4° проводится аналогично одномерному случаю.

Для одномерных случайных величин этих свойств было достаточно, чтобы функция, обладающая ими, была функцией распределения. Для n-мерных случайных величин, когда $n>1$, этих свойств недостаточно. Необходимо еще одно свойство, которое мы сейчас рассмотрим.

Обозначим $I_k=(a_k;b_k]$, $\Delta_{I_k}^k$ – преобразование, применение которого к функции $g(x_1,\ \ldots,\ x_n)$ дает:

$$\Delta_{I_k}^k g(x_1,\ ...,\ x_n)=g(x_1,\ \ldots,\ x_{k-1},\ b_k,\ x_{k+1},\ \ldots,\ x_n)-$$
$$-g(x_1,\ \ldots,\ x_{k-1},\ a_k,\ x_{k+1},\ \ldots,\ x_n).$$

5°.*Для любых интервалов* I_k, $k=\overline{1,n}$,

$$\Delta_{I_n}^n\ldots\Delta_{I_1}^1 F_{\bar{\xi}}(x_1,\ \ldots,\ x_n)=\Delta_{I_n}^n\left(\Delta_{I_{n-1}}^{n-1}\ldots\left(\Delta_{I_1}^1 F_{\bar{\xi}}(x_1,\ \ldots,\ x_n)\right)\right)\geq 0.$$

Д о к а з а т е л ь с т в о. Из п. 1 упражнения 1 § 2 следует, что

$$\Delta_{I_n}^n\ldots\Delta_{I_1}^1 F_{\bar{\xi}}(\bar{x})=\mathrm{P}(\xi_1\in I_1\cap\ldots\cap\xi_n\in I_n),$$

а поэтому из аксиомы неотрицательности вероятностей приходим к доказательству свойства 5°.

З а м е ч а н и е 3. Свойство 5° не вытекает из свойств 1°–4°. Покажем это на примере. Пусть

$$F(x_1,\ x_2)=\begin{cases}1, & \text{если } x_1>0,\quad x_2>0,\quad x_1+x_2>1,\\ 0, & \text{для остальных}\quad x_1,x_2.\end{cases}$$

Функция $F(x_1,\ x_2)$, $x_1,\ x_2\in R$, удовлетворяет свойствам 1°–4°. Покажем, что она не удовлетворяет свойству 5°. Пусть $I_1=I_2=(0;1]$. Тогда

$$\Delta_{I_2}^2\Delta_{I_1}^1 F(x_1,\ x_2)=\Delta_{I_2}^2\left(F(1,\ x_2)-F(0,\ x_2)\right)=$$
$$=F(1,\ 1)-F(0,\ 1)-F(1,\ 0)+F(0,\ 0)=-1.$$

Теорема 1 [41]. *Если функция* $F: R^n\to R$ *удовлетворяет свойствам* 1°–5°*, то она является функцией распределения некоторой n-мерной случайной величины.*

У п р а ж н е н и е 2. Приведите пример двумерной функции распределения.

По аналогии с одномерным случаем введем понятие распределения вероятностей n-мерной случайной величины (случайного вектора).

О п р е д е л е н и е 4. *Функция*

$$P_{\overline{\xi}}:\ \mathcal{B}\left(R^n\right)\to[0;1]$$

называется распределением вероятностей n-мерной случайной величины $\overline{\xi}=\left(\xi_1,\ ...,\ \xi_n\right)$, *если для любого* $B\in\mathcal{B}\left(R^n\right)$

$$P_{\overline{\xi}}\left(B\right)=\mathrm{P}\left(\omega\in\Omega:\ \overline{\xi}\left(\omega\right)\in B\right).$$

У п р а ж н е н и е 3. Докажите, что $P_{\overline{\xi}}$ является вероятностной мерой на измеримом пространстве $\left(R^n,\ \mathcal{B}\left(R^n\right)\right)$.

Теорема 2 [29]. *Распределение вероятностей однозначно определяется функцией распределения* $F_{\overline{\xi}}$.

У п р а ж н е н и е 4. Покажите, что для любого $B\in\mathcal{B}\left(R^n\right)$

$$P_{\overline{\xi}}\left(B\right)=\int\limits_B dF_{\overline{\xi}}\left(\overline{x}\right),\quad \overline{x}=\left(x_1,\ x_2,\ \dots,\ x_n\right).$$

У п р а ж н е н и е 5. Можно ли утверждать, что многомерная функция распределения имеет не более чем счетное число точек разрыва?

Классификация n-мерных случайных величин (случайных векторов)

1. Дискретные случайные векторы

Случайный вектор называется *дискретным*, если он принимает не более чем счетное число значений.

Задать дискретный случайный вектор, значит, задать его *закон распределения*, т. е. задать вероятности $\mathrm{P}\left(\overline{\xi}=\overline{x}\right)$ для всех возможных значений $\overline{x}$ случайного вектора $\overline{\xi}$.

П р и м е р 3. П о л и н о м и а л ь н ы й з а к о н р а с п р е д е л е н и я с п а р а м е т р а м и $\left(N,p\right)$ (ср. с полиномиальной схемой). Распределение случайного вектора $\xi=\left(\xi_1,\ \dots,\ \xi_n\right)$ называется *полиномиальным с параметрами* N, $p=\left(p_1,\ \dots,\ p_n\right)$, если

$$\xi_k=\overline{0,N},\ \ k=\overline{1,n};\ \ \sum_{k=1}^{n}\xi_k=N;\ \ \mathrm{P}\left(\xi_1=k_1\cap\ldots\cap\xi_n=k_n\right)=\frac{N!}{k_1!\ \ldots k_n!}p_1^{k_1}\ldots p_n^{k_n}.$$

Рассмотрим понятие *закона распределения* в простейшем случае, а именно: пусть $n=2$,

$$\overline{\xi}=\left(\xi_1,\ \xi_2\right),\quad p_{lm}=\mathrm{P}\left(\left(\xi_1=x_l\right)\cap\left(\xi_2=x_m\right)\right),\quad \sum_{l,m=1}^{\infty}p_{lm}=1,\ \ p_{lm}\ge 0,$$

тогда закон распределения случайного вектора $\overline{\xi}$ (его также называют *совместным законом распределения случайных величин* ξ_1 и ξ_2) можно представить в виде табл. 1:

Таблица 1

$\xi_2 \backslash \xi_1$	x_1	x_2	$\dots$	x_l	$\dots$
y_1	p_{11}	p_{21}	$\dots$	p_{l1}	$\dots$
y_2	p_{12}	p_{22}	$\dots$	p_{l2}	$\dots$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
y_m	p_{1m}	p_{2m}	$\dots$	p_{lm}	$\dots$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$

Пример 4. Пусть случайный вектор $\overline{\xi}=(\xi_1,\ \xi_2)$ имеет приведенный в табл. 1 закон распределения. Найдем законы распределения координат ξ_1, ξ_2. Сначала рассмотрим случай, когда закон распределения имеет следующий вид:

$\xi_2 \backslash \xi_1$	0	1
0	1/4	1/4
1	1/4	1/4

Несложно видеть, что

$$\mathrm{P}(\xi_1=0)=\mathrm{P}\Big((\xi_1=0)\cap\big((\xi_2=0)+(\xi_2=1)\big)\Big)=$$

$$=\mathrm{P}\Big(\big((\xi_1=0)\cap(\xi_2=0)\big)+\big((\xi_1=0)\cap(\xi_2=1)\big)\Big)=$$

$$=\mathrm{P}\big((\xi_1=0)\cap(\xi_2=0)\big)+\mathrm{P}\big((\xi_1=0)\cap(\xi_2=1)\big)=\frac{1}{4}+\frac{1}{4}=\frac{1}{2},$$

поэтому законы распределения ξ_1 и ξ_2 имеют вид:

$$\begin{array}{lcc} \xi_1: & 0 & 1 \\ \mathrm{P}: & 1/2 & 1/2 \end{array} \qquad\qquad \begin{array}{lcc} \xi_2: & 0 & 1 \\ \mathrm{P}: & 1/2 & 1/2 \end{array}.$$

В общем случае, используя данные табл. 1, получаем

$$\begin{array}{lccccc} \xi_1: & x_1 & x_2 & \dots & x_l & \dots \\ \mathrm{P}: & \sum\limits_{i=1}^{m} p_{1i} & \sum\limits_{i=1}^{m} p_{2i} & \dots & \sum\limits_{i=1}^{m} p_{li} & \dots \end{array} \qquad \begin{array}{lccccc} \xi_2: & y_1 & y_2 & \dots & y_m & \dots \\ \mathrm{P}: & \sum\limits_{i=1}^{l} p_{i1} & \sum\limits_{i=1}^{l} p_{i2} & \dots & \sum\limits_{i=1}^{l} p_{im} & \dots \end{array}.$$

Упражнение 6. Пусть случайные величины ξ_1, ξ_2 имеют законы распределения из прим. 4 Можно ли найти закон распределения вектора $\overline{\xi}=(\xi_1,\xi_2)$ (совместный закон распределения)?

Указание. В общем случае – нет. Для ответа нужно знать связь между ξ_1 и ξ_2 (см., например, § 4).

2. Абсолютно непрерывные случайные векторы

Определение 5. *Распределение вероятностей n-мерной случайной величины (случайного вектора)* $\overline{\xi}=(\xi_1, \ldots, \xi_n)$ *называется абсолютно непрерывным распределением, а сама случайная величина* $\overline{\xi}$ *– абсолютно непрерывной n-мерной случайной величиной, если для любого множества* $B\in\mathscr{B}(R^n)$

$$P_{\overline{\xi}}(B)=\int\ldots\int_B p_{\overline{\xi}}(x_1, \ldots, x_n)\,dx_1\ldots dx_n, \tag{1}$$

где $p_{\overline{\xi}}: R^n\to R$ *интегрируемая по Лебегу функция.*

Функция $p_{\overline{\xi}}$ называется *плотностью распределения n-мерной случайной величины* или *многомерной плотностью распределения.*

Несложно видеть, что функция распределения абсолютно непрерывной n-мерной случайной величины для любых $\overline{x}\in R^n$ допускает представление

$$F_{\overline{\xi}}(\overline{x})=\int_{-\infty}^{x_1}\ldots\int_{-\infty}^{x_n} p_{\overline{\xi}}(t_1, \ldots, t_n)\,dt_1\ldots dt_n. \tag{2}$$

Для этого в равенстве (1) достаточно в качестве борелевского множества взять $B=(-\infty;x_1]\times\ldots\times(-\infty;x_n]$.

Таким образом, распределение вероятностей однозначно определяет функцию распределения и в многомерном случае. Справедливо и обратное утверждение (см., напр., [29]).

Свойства многомерной плотности распределения

1°. $p_{\overline{\xi}}(\overline{x})\geq 0$ *для почти всех* $\overline{x}\in R^n$.

Доказательство вытекает из равенства (1) и того, что $P_{\overline{\xi}}(B)\geq 0$ для любых $B\in\mathscr{B}(R^n)$.

2°. $\int_{R^n} p_{\overline{\xi}}(\overline{x})\,d\overline{x}=1.$

Доказательство следует из равенства (2) и того, что $F_{\overline{\xi}}(+\infty, \ldots, +\infty)=1$.

3°. Если существует

$$\frac{\partial^n F_{\overline{\xi}}(x_1, \ldots, x_n)}{\partial x_1\ldots\partial x_n},$$

то

$$p(x_1, \ldots, x_n) = \frac{\partial^n F_{\bar{\xi}}(x_1, \ldots, x_n)}{\partial x_1 \ldots \partial x_n}.$$

Для доказательства смотрите равенство (2).

4°. *Для любого* $k = \overline{1, n}$

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

Доказательство вытекает из равенства (2) и свойства 3° многомерной функции распределения.

Утверждение 1. *Для того чтобы функция* $p(\bar{x})$, $\bar{x} \in R^n$, *была плотностью распределения некоторой n-мерной случайной величины* ξ, *необходимо и достаточно, чтобы она удовлетворяла свойствам* 1° *и* 2°.

Доказательство. Если функция $p(x_1, \ldots, x_2)$ удовлетворяет свойствам 1° и 2°, то функция $F(x_1, \ldots, x_2)$, определенная по формуле (2) будет удовлетворять свойствам 1°–5°, многомерной функции распределения. Тогда по теореме 1 существует n-мерная случайная величина $\bar{\xi}$, для которой $F(x_1, \ldots, x_2)$ будет функцией распределения, а $p(x_1, \ldots, x_2)$ соответственно плотностью распределения. Утверждение доказано.

Пример 5. Равномерное распределение на борелевском множестве G из R^n. Распределение вероятностей n-мерной случайной величины $\bar{\xi}$ называется *равномерным на множестве* G, а величина $\bar{\xi}$ – *равномерно распределенной на* G, если для любых $B \in \mathscr{B}(R^n)$

$$P_{\xi}(B) = \frac{|B \cap G|}{|G|},$$

где $|A|$ – мера Лебега борелевского множества A и $|G| < \infty$.

В этом случае плотность распределения $\bar{\xi}$ имеет вид:

$$p(x_1, \ldots, x_n) = \begin{cases} \dfrac{1}{|G|}, & \text{если} \quad (x_1, \ldots, x_n) \in G, \\ 0, & \text{если} \quad (x_1, \ldots, x_n) \notin G. \end{cases}$$

Пример 6. Многомерное нормальное распределение с параметрами $(\bar{a}, \sigma^2)$. Распределение вероятностей n-мерной случайной величины называется *многомерным нормальным распределением с параметрами*

$$\bar{a} = (a_1, \ldots, a_n) \text{ и } \sigma^2 = \left[\sigma_{ij}\right]_{i,j=1}^{n}$$

(симметричная неотрицательно определенная матрица), если для любых $B \in \mathscr{B}\left(R^n\right)$

$$P_{\bar{\xi}}(B) = \frac{\sqrt{|A|}}{(2\pi)^{\frac{n}{2}}} \int\limits_B \exp\left\{\frac{1}{2}(x-a)A(x-a)^T\right\} dx,$$

где $A = \left[a_{ij}\right]_{i,j=1}^n$, $\sigma^2 = A^{-1}$, $xAx^T = \sum\limits_{i,j=1}^n a_{ij}x_ix_j$, $|A|$ – определитель матрицы A.

3 . Сингулярные случайные векторы

О п р е д е л е н и е 6. *Распределение вероятностей n-мерной случайной величины (случайного вектора)* $\bar{\xi} = (\xi_1, \ldots, \xi_n)$ *называется сингулярным, а сама случайная величина* $\bar{\xi}$ *– сингулярной n-мерной случайной величиной (сингулярным вектором), если*

1) $F_{\bar{\xi}}$ *– непрерывная функция,*

2) *существует множество* $S \in \mathscr{B}\left(R^n\right)$ *нулевой меры Лебега, такое, что* $\mathrm{P}\left(\bar{\xi} \in S\right) = 1$.

У п р а ж н е н и е 7. Сравните определение 6 при $n = 1$ с определением сингулярной случайной величины (см. § 2).

П р и м е р 7. Пусть случайная величина ξ_1 имеет непрерывную функцию распределения $F_{\xi_1}(x)$, $g: R \to R$ – непрерывная, строго монотонно возрастающая функция, g^{-1} – обратная функция, кроме того $\xi_2 = g(\xi_1)$. Тогда случайный вектор $\bar{\xi} = (\xi_1, \xi_2)$ будет сингулярным, так как

$$F_{(\xi_1,\xi_2)}(x_1, x_2) = \mathrm{P}(\xi_1 \le x_1 \cap \xi_2 \le x_2) = \mathrm{P}(\xi_1 \le x_1 \cap g(\xi_1) \le x_2) =$$
$$= \mathrm{P}\left(\xi_1 \le x_1 \cap \xi_1 \le g^{-1}(x_2)\right) = \mathrm{P}\left(\xi_1 \le \min\left(x_1;\, g^{-1}(x_2)\right)\right) =$$
$$= F_{\xi_1}\left(\min\left(x_1;\, g^{-1}(x_2)\right)\right) \text{ – непрерывная функция.}$$

Рассмотрим множество

$$S = \left\{(x_1, x_2):\, x_i \in R,\;\; i = 1,\, 2,\;\; x_2 = g(x_1)\right\} \in \mathscr{B}\left(R^n\right).$$

Множество S – кривая в R^2, мера Лебега которой равна 0 и

$$\mathrm{P}\left((\xi_1, \xi_2) \in S\right) = \mathrm{P}\left((\xi_1, g(\xi_1)) \in S\right) = 1.$$

§ 4. Независимость случайных величин

Понятие независимости случайных величин представляет собой перенос понятия независимости случайных событий на случайные величины и опять-таки отражает отсутствие связи между случайными величинами. Иными словами, независимость случайных величин ξ и η можно охарактеризовать следующим образом: знание значений, которые приняла случайная величина ξ, не дает никакой новой информации о распределении случайной величины η.

О п р е д е л е н и е 1. *Сигма-алгебра (σ-алгебра)*

$$\mathscr{A}_\xi = \xi^{-1}\left(\mathscr{B}(R)\right)$$

называется сигма-алгеброй, порожденной случайной величиной ξ.

Пусть случайные величины $\xi_1, \ldots, \xi_n$ определены на одном и том же вероятностном пространстве $(\Omega, \mathscr{A}, \mathrm{P})$.

О п р е д е л е н и е 2. *Случайные величины $\xi_1, \ldots, \xi_n$ называются независимыми, если независимы σ-алгебры $\mathscr{A}_1, \ldots, \mathscr{A}_n$, порожденные этими случайными величинами, т. е. для любых $B_1, \ldots, B_n \in \mathscr{B}(R)$ имеет место равенство*

$$\mathrm{P}\left(\xi_1 \in B_1 \bigcap \xi_2 \in B_2 \bigcap \ldots \bigcap \xi_n \in B_n\right) = \prod_{k=1}^{n} \mathrm{P}\left(\xi_k \in B_k\right). \qquad (1)$$

Теорема 1 (**Критерий независимости**) [5]. *Случайные величины $\xi_1, \ldots, \xi_n$ независимы тогда и только тогда, когда для любых $x_1, \ldots, x_n \in R$*

$$F_{(\xi_1,\ldots,\xi_n)}\left(x_1, \ldots, x_n\right) = \prod_{k=1}^{n} F_{\xi_k}\left(x_k\right). \qquad (2)$$

Теорема 2. *Пусть случайные величины $\xi_1, \ldots, \xi_n$ абсолютно непрерывны и существует плотность распределения $p_{(\xi_1, \ldots, \xi_n)}\left(x_1, \ldots, x_n\right)$ случайного вектора $\overline{\xi} = \left(\xi_1, \ldots, \xi_n\right)$.*

Они независимы тогда и только тогда, когда для любых $x_1, \ldots, x_n \in R$

$$p_{(\xi_1,\ldots,\xi_n)}\left(x_1, \ldots, x_n\right) = \prod_{k=1}^{n} p_{\xi_k}\left(x_k\right). \qquad (3)$$

Д о к а з а т е л ь с т в о. Н е о б х о д и м о с т ь. Пусть случайные величины $\xi_1, \ldots, \xi_n$ абсолютно непрерывны, независимы и существует плотность $p_{(\xi_1, \ldots, \xi_n)}$. Тогда, во-первых, в силу справедливости равенства (2) и абсолютной непрерывности случайных величин $\xi_1, \ldots, \xi_n$ для любых $x_1, \ldots, x_n \in R$

$$F_{(\xi_1, \ldots, \xi_n)}\left(x_1, \ldots, x_n\right) = \prod_{k=1}^{n} F_{\xi_k}\left(x_k\right) = \prod_{k=1}^{n} \int_{-\infty}^{x_k} p_{\xi_k}\left(t_k\right) dt_k =$$

$$= \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_n} p_{\xi_1}(t_1) \ldots p_{\xi_n}(t_n)\, dt_1 \ldots dt_n, \tag{4}$$

во-вторых, учитывая существование плотности $p_{(\xi_1, \ldots, \xi_n)}$ для любых $x_1, \ldots, x_n \in R$

$$F_{(\xi_1, \ldots, \xi_n)}(x_1, \ldots, x_n) = \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_n} p_{(\xi_1, \ldots, \xi_n)}(t_1, \ldots, t_n)\, dt_1 \ldots dt_n. \tag{5}$$

Из равенства для любых $x_1, \ldots, x_n \in R$ правых частей записанных выше соотношений приходим к справедливости равенства (3). Необходимость доказана.

Д о с т а т о ч н о с т ь. Пусть справедливо равенство (3). Тогда, воспользовавшись соотношениями (4) и (5), приходим к справедливости равенства (2), а это означает, что случайные величины $\xi_1, \ldots, \xi_n$ независимы. Теорема доказана.

У п р а ж н е н и е 1. Можно ли заменить равенство (3), выполняющееся для любых $x_1, \ldots, x_n \in R$, на аналогичное, но справедливое для почти всех по мере Лебега $x_1, \ldots, x_n \in R^n$?

У п р а ж н е н и е 2. Требование существования плотности случайного вектора $(\xi_1, \ldots, \xi_n)$ достаточно для доказательства в одну сторону. Зададим вопрос: в какую?

П р и м е р 1. Пусть случайная величина ξ – стандартная нормальная случайная величина, т. е.

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in R.$$

Приведите пример абсолютно непрерывной случайной величины η, независимой со случайной величиной ξ.

Р е ш е н и е. В качестве искомой случайной величины ξ можно взять произвольную случайную величину с плотностью $p_\eta(y)$ и совместную плотность распределения $p_{(\xi, \eta)}(x, y)$ задать следующим образом:

$$p_{(\xi, \eta)}(x, y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} p_\eta(y), \quad x, y \in R.$$

Теорема 3. *Дискретные случайные величины $\xi_1, \ldots, \xi_n$ независимы тогда и только тогда, когда для любых $x_1, \ldots, x_n$ возможных значений соответственно $\xi_1, \ldots, \xi_n$ выполнено равенство*

$$\mathrm{P}(\xi_1 = x_1 \cap \xi_2 = x_2 \cap \ldots \cap \xi_n = x_n) = \prod_{k=1}^{n} \mathrm{P}(\xi_k = x_k). \tag{6}$$

Д о к а з а т е л ь с т в о проведем для случая $n = 2$. Общий случай доказывается методом математической индукции. Итак, покажем, что для любых $x_1,\ x_2 \in R$

$$\mathrm{P}\left(\xi_1 = x_1 \cap \xi_2 = x_2\right) = \mathrm{P}\left(\xi_1 = x_1\right)\mathrm{P}\left(\xi_2 = x_2\right). \tag{7}$$

Н е о б х о д и м о с т ь. Пусть случайные величины ξ_1 и ξ_2 независимы, т. е.

$$\mathrm{P}\left(\xi_1 \in B_1 \cap \xi_2 \in B_2\right) = \mathrm{P}\left(\xi_1 \in B_1\right)\mathrm{P}\left(\xi_2 \in B_2\right)$$

для любых борелевских множеств B_1 и B_2. Если взять $B_1 = \{x_1\}$, $B_2 = \{x_2\}$, то приходим к справедливости (7).

Д о с т а т о ч н о с т ь. Пусть выполняется соотношение (7), тогда

$$\mathrm{P}\left(\xi_1 \in B_1 \cap \xi_2 \in B_2\right) = \sum_{x_1 \in B_1}\sum_{x_2 \in B_2} \mathrm{P}\left(\xi_1 = x_1 \cap \xi_2 = x_2\right) =$$

$$= \sum_{x_1 \in B_1}\sum_{x_2 \in B_2} \mathrm{P}\left(\xi_1 = x_1\right)\mathrm{P}\left(\xi_2 = x_2\right) = \sum_{x_1 \in B_1} \mathrm{P}\left(\xi_1 = x_1\right) \sum_{x_2 \in B_2} \mathrm{P}\left(\xi_2 = x_2\right) =$$

$$= \mathrm{P}\left(\xi_1 \in B_1\right)\mathrm{P}\left(\xi_2 \in B_2\right)$$

для любых борелевских множеств B_1 и B_2. Теорема доказана.

У т в е р ж д е н и е 1. *Если $\xi_1,\ \ldots,\ \xi_n$ – независимые случайные величины, $g_i : R \to R,\ i = \overline{1,n},$ – борелевские функции, то случайные величины $g_1(\xi_1),\ \ldots,\ g_n(\xi_n)$ также независимы.*

Д о к а з а т е л ь с т в о. Пусть $B_k = \left\{x :\ g_k(x) \le y_k\right\},\ \ k = \overline{1,n},$ – борелевские множества. Тогда

$$\mathrm{P}\left(\bigcap_{k=1}^{n}\left(g_k(\xi_k) \le y_k\right)\right) = \mathrm{P}\left(\bigcap_{k=1}^{n}\left(\xi_k \in B_k\right)\right) = \prod_{k=1}^{n}\mathrm{P}\left(\xi_k \in B_k\right) =$$

$$= \prod_{k=1}^{n}\mathrm{P}\left(g_k(\xi_k) \le y_k\right)$$

для любых $y_k \in R,\ \ k = \overline{1,n}.$ В силу критерия независимости (2) $g_1(\xi_1),\ \ldots,\ g_n(\xi_n)$ являются независимыми случайными величинами. Утверждение доказано.

П р и м е р 2. Пусть вероятностное пространство $(\Omega,\ \mathscr{A},\ \mathrm{P})$ задано следующим образом: $\Omega = [0;1]$, $\mathscr{A} = \mathscr{B}([0;1])$, P – мера Лебега. Рассмотрим на этом вероятностном пространстве случайную величину $\xi(\omega) = \omega.$ Найдем все случайные величины η независимые с ξ на данном вероятностном пространстве.

Используя утверждение 1, несложно показать, что только тривиальные (вырожденные) случайные величины η будут независимы с ξ.

Напомним, что случайная величина η называется *вырожденной*, если существует такая константа $c \in R$, что $\mathrm{P}(\xi = c) = 1$.

Данный пример показывает, что необходимо с должной осторожностью относиться к фразе «пусть на произвольном вероятностном пространстве $(\Omega, \mathscr{A}, \mathrm{P})$ заданы независимые случайные величины ξ и η (произвольные!)». На самом деле ситуация несколько иная: случайные величины ξ и η действительно могут рассматриваться на произвольных вероятностных пространствах $(\Omega_1, \mathscr{A}_1, \mathrm{P}_1)$ и $(\Omega_2, \mathscr{A}_2, \mathrm{P}_2)$ соответственно. А вот пространство $(\Omega, \mathscr{A}, \mathrm{P})$ строится по ним так, чтобы случайные величины ξ и η были независимы.

ДОПОЛНЕНИЯ К ГЛАВЕ 3

§ 1. Функциональные преобразования случайных величин

Пусть на произвольном вероятностном пространстве $(\Omega,\ \mathcal{A},\ \mathrm{P})$ задана дискретная случайная величина ξ с законом распределения

$$\begin{aligned}&\xi:\ \ x_1,\ \ldots,\ x_n,\ \ldots\\ &p:\ \ p_1,\ \ldots,\ p_n,\ \ldots,\quad p_i>0,\quad \sum_i p_i=1.\end{aligned}$$

Функция $g: R\to R$ – борелевская функция. Тогда очевидно, что закон распределения случайной величины $g(\xi)$ имеет вид

$$\begin{aligned}&g(\xi):\ \ g(x_1),\ \ g(x_2),\ \ \ldots,\ \ g(x_n),\ \ \ldots\\ &\mathrm{P}:\qquad p_1,\qquad p_1,\qquad \ldots,\qquad p_n,\qquad \ldots,\quad p_i>0,\quad \sum_i p_i=1.\end{aligned}$$

Многомерный случай описывается аналогично.

Пусть $\overline{\xi}=(\xi_1,\ \ldots,\ \xi_n)$ – n-мерная абсолютно непрерывная случайная величина с плотностью распределения

$$p_{\overline{\xi}}(\overline{x}),\ \ \overline{x}=(x_1,\ \ldots,\ x_n)\in R^n.$$

Рассмотрим случайный вектор $\overline{\eta}=(\eta_1,\ \ldots,\ \eta_n)$, где $\eta_i=g_i(\xi_i)$, $i=\overline{1,n}$. Функции $g_i: R\to R$ – непрерывные, для которых существуют обратные преобразования:

$$\xi_i=g_i^{-1}(\eta_i),\ \ i=\overline{1,n}.$$

Обозначим

$$g(\overline{x})=(g_1(\overline{x}),\ \ldots,\ g_n(\overline{x})),\qquad g^{-1}(\overline{x})=\left(g_1^{-1}(\overline{x}),\ \ldots,\ g_n^{-1}(\overline{x})\right).$$

Теорема 1. *Если случайный вектор $\overline{\xi}$ имеет плотность распределения $p_{\overline{\xi}}$, $\overline{\eta}=g\left(\overline{\xi}\right)$, где функция $g=(g_1,\ \ldots,\ g_n)$ непрерывно дифференцируема и имеет обратную функцию $g^{-1}=\left(g_1^{-1},\ \ldots,\ g_n^{-1}\right)$,, то случайный вектор η имеет плотность распределения $p_{\overline{\eta}}$, которая определяется по формуле*

$$p_{\overline{\eta}}(\overline{x})=p_{\xi}\left(g^{-1}(\overline{x})\right)\left|Ig^{-1}(\overline{x})\right|,$$

где

$$Ig^{-1}(\overline{x})=\left[\frac{\partial g_i^{-1}(\overline{x})}{\partial x_j}\right]_{i,j=1}^{n},$$

а $|A|$ – определитель матрицы A.

Доказательство. Пусть $B\in\mathcal{B}\left(R^n\right)$. Тогда из определения плотности распределения и формулы замены переменных в многомерном интеграле имеем

$$\mathrm{P}\left(\overline{\eta}\in B\right)=\mathrm{P}\left(g\left(\overline{\xi}\right)\in B\right)=\int\limits_{g^{-1}(B)} p_{\overline{\xi}}\left(\overline{x}\right)d\,\overline{x}=\int\limits_{B} p_{\xi}\left(g^{-1}\left(\overline{y}\right)\right)\left|Ig^{-1}\left(\overline{y}\right)\right|dy.$$

Из этого равенства и определения абсолютно непрерывного случайного вектора приходим к доказательству теоремы 1.

Формулы композиции (свертки)

Пусть на вероятностном пространстве $(\Omega,\ \mathscr{A},\ \mathrm{P})$ даны две случайные величины ξ и η и известна связь между ними, т. е. известна совместная функция распределения $F_{\xi,\eta}(x,\ y)$. Найдем функцию распределения суммы $\xi+\eta$. Для этого воспользуемся формулой (см. § 3 упражнение 4)

$$\mathrm{P}\left(\overline{\xi}\in B\right)=\int\limits_{B} dF_{\overline{\xi}}\left(\overline{x}\right)$$

для любых $B\in\mathscr{B}\left(R^{n}\right)$. Пусть $\overline{\xi}=(\xi,\eta)$, тогда из этой формулы получаем

$$F_{\xi+\eta}(z)=\mathrm{P}(\xi+\eta\le z)=\mathrm{P}\left((\xi,\eta)\in B:\ B=\left\{(x,y):\ \ x+y\le z\right\}\right)=$$
$$=\iint\limits_{x+y\le z} dF_{(\xi,\ \eta)}(x,\ y),$$

т. е.

$$F_{\xi+\eta}(z)=\iint\limits_{x+y\le z} dF_{(\xi,\ \eta)}(x,\ y). \tag{1}$$

Пусть ξ и η независимы, тогда из (1) имеем

$$F_{\xi+\eta}(z)=\iint\limits_{x+y\le z} dF_{\xi}(x)dF_{\eta}(y)=\int\limits_{-\infty}^{+\infty}\left(\int\limits_{-\infty}^{z-y} dF_{\xi}(x)\right)dF_{\eta}(y)=$$
$$=\int\limits_{-\infty}^{+\infty} F_{\xi}(z-y)-F_{\xi}(-\infty)dF_{\eta}(y)=\int\limits_{-\infty}^{+\infty} F_{\xi}(z-y)dF_{\eta}(y).$$

Учитывая, что $F_{\xi+\eta}(z)$ симметрично относительно ξ и η, получаем

$$F_{\xi+\eta}(z)=\int\limits_{-\infty}^{+\infty} F_{\xi}(z-y)dF_{\eta}(y)=\int\limits_{-\infty}^{+\infty} F_{\eta}(z-y)dF_{\xi}(y). \tag{2}$$

Пусть случайная величина ξ произвольна, а случайная величина η абсолютно непрерывна и они независимы, тогда из (2) следует

$$F_{\xi+\eta}(z)=\int\limits_{-\infty}^{+\infty} F_{\xi}(z-y)dF_{\eta}(y)=\int\limits_{-\infty}^{+\infty} F_{\xi}(z-y)p_{\eta}(y)dy. \tag{3}$$

С другой стороны

$$F_{\xi+\eta}(z)=\int\limits_{-\infty}^{+\infty} F_{\eta}(z-y)dF_{\xi}(y)=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{z-y} p_{\eta}(t)dt\,dF_{\xi}(y)=$$
$$=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{z} p_{\eta}(t-y)dt\,dF_{\xi}(y)=\int\limits_{-\infty}^{z}\left(\int\limits_{-\infty}^{+\infty} p_{\eta}(t-y)dF_{\xi}(y)\right)dt,\quad z\in R. \tag{4}$$

Следовательно, можно сделать вывод, что сумма независимых случайных величин, одна из которых абсолютно непрерывна, а вторая произвольна, является абсолютно непрерывной случайной величиной. В нашем случае из соотношения (4) следует, что существует плотность $p_{\xi+\eta}$ и она равна

$$p_{\xi+\eta}(t)=\int\limits_{-\infty}^{+\infty} p_\eta(t-y)\,dF_\xi(y), \quad t\in R. \tag{5}$$

Пусть случайные величины ξ и η абсолютно непрерывны и независимы, тогда из равенства (5), учитывая то, что его левая часть симметрична относительно ξ и η, получаем

$$p_{\xi+\eta}(t)=\int\limits_{-\infty}^{+\infty} p_\eta(t-y)\,p_\xi(y)\,dy=\int\limits_{-\infty}^{+\infty} p_\xi(t-y)\,p_\eta(y)\,dy. \tag{6}$$

Формулы (1) – (6) называют *формулами композиции,* или *свертками.*

Задачи

1. Проверьте следующие свойства индикаторов $I_A = I_A(\omega)$:

$$I_\varnothing\equiv 0,\ \ I_\Omega\equiv 1,\ \ I_A+I_{\bar A}\equiv 1,\ \ I_A I_B = I_{AB},\ \ I_{A\cup B}=I_A+I_B-I_{AB},$$

$$I_{\bigcup\limits_{k=1}^{n}A_k}=1-\prod_{k=1}^{n}\left(1-I_{A_k}\right),\ \ I_{\overline{\bigcup\limits_{k=1}^{n}A_k}}=\prod_{k=1}^{n}\left(1-I_{A_k}\right),$$

$$I_{\sum\limits_{k=1}^{n}A_k}=\sum_{k=1}^{n}I_{A_k},\quad I_{A\Delta B}=\left(I_A-I_B\right)^2.$$

2. Пусть $\xi_1,\ \ldots,\ \xi_n$ – независимые случайные величины и

$$\xi_{\min}=\min\left(\xi_1,\ \ldots,\ \xi_n\right),\ \ \xi_{\max}=\max\left(\xi_1,\ \ldots,\ \xi_n\right)..$$

Покажите, что

$$\mathrm{P}\left(\xi_{\min}>x\right)=\prod_{k=1}^{n}\mathrm{P}\left(\xi_k>x\right),\ \ \mathrm{P}\left(\xi_{\max}\le x\right)=\prod_{k=1}^{n}\mathrm{P}\left(\xi_k\le x\right),\ \ x\in R.$$

3. Пусть $\xi^+=\max(\xi,0)$, $\xi^-=\min(\xi,0)$. Выразите функции распределения ξ^+ и ξ^- через функцию распределения ξ.

4. Покажите, что случайная величина ξ не зависит от самой себя в том и только в том случае, когда $\mathrm{P}(\xi=\mathrm{const})=1$.

5. При каких условиях на ξ случайные величины ξ и $\sin\xi$ независимы?

6. Пусть ξ и η независимые случайные величины с известными функциями распределения. Найдите функции распределения случайных величин $\xi\eta$ и ξ/η.

7. Покажите, что, если для любого $x\in R$ $\mathrm{P}(\xi=x)=0$, то случайная величина ξ – абсолютно непрерывная или сингулярная. Верно ли обратное утверждение?

8. Пусть $|\xi|$ – случайная величина. Будет ли ξ случайной величиной?

9. Докажите, что функции $x^+=\max(x,0)$, $x^-=\min(x,0)$, $|x|=x^++x^-$, $x\in R$, являются борелевскими.

10. На отрезке $[0;a]$ независимо друг от друга берут две точки с равномерным распределением. Найдите функцию и плотность распределения расстояния между ними.

11. Пусть случайные величины $\xi_1,\ \ldots,\ \xi_n,\ n \geq 0,$ независимы и одинаково распределены с функцией распределения $F(x)$ и плотностью $p(x),\ x \in R,$ (если она существует); $\xi^- = \min(\xi_1,\ \ldots,\ \xi_n),\ \xi^+ = \max(\xi_1,\ \ldots,\ \xi_n),\ \overline{\xi} = \xi^+ - \xi^-$. Докажите, что

$$F_{(\xi^+,\ \xi^-)}(x,\ y) = \begin{cases} (F(y))^n - (F(y) - F(x))^n,\ y > x, \\ (F(y))^n,\ y \leq x, \end{cases}$$

$$p_{(\xi^+,\ \xi^-)}(x,\ y) = \begin{cases} n(n-1)\left[F(y) - F(x)\right]^{n-2} p(x)p(y),\ y > x, \\ 0,\ y \leq x, \end{cases}$$

$$F_{\overline{\xi}}(x) = \begin{cases} n\int\limits_{-\infty}^{\infty}\left[F(y) - F(y-x)\right]^{n-1} p(y)dy,\ x > 0, \\ 0,\ x \leq 0, \end{cases}$$

$$p_{\overline{\xi}}(x) = \begin{cases} n(n-1)\int\limits_{-\infty}^{\infty}\left[F(y) - F(y-x)\right]^{n-2} p(y-x)p(y)dy,\ x > 0, \\ 0,\ x \leq 0. \end{cases}$$

12. Пусть случайные величины $\xi_1,\ \ldots,\ \xi_n,\ n \geq 2,$ независимы и одинаково распределены с функцией распределения $F(x)$ и плотностью $p(x)$. Упорядочив их по возрастанию, образуем «вариационный» ряд $\xi_{(1)} \leq \xi_{(2)} \leq \ldots \leq \xi_{(n)}$. Найдите плотность распределения $\xi_{(k)},\ k = \overline{1,n},$ и двумерную плотность распределения $(\xi_{(k)},\ \xi_{(l)}),\ 1 \leq k < l \leq n.$

13. На прямоугольнике $0 \leq x \leq a,\ 0 \leq y \leq b$ случайно с равномерным распределением берется точка. Докажите, что ее координаты $(\xi,\ \eta)$ независимы.

14. На круге $x^2 + y^2 \leq R^2$ случайно с равномерным распределением берется точка. Покажите, что ее координаты $(\xi,\ \eta)$ зависимы.

15. Пусть $F(x),\ x \in R,$ – функция распределения случайной величины ξ. Найдите функцию распределения случайной величины $F(\xi)$.

16. Пусть $\xi_1,\ \ldots,\ \xi_n$ – независимые одинаково распределенные экспоненциальные случайные величины с параметром $\lambda,\ \lambda > 0$. Докажите, что

$$p_{\xi_1 + \ldots + \xi_n}(t) = \begin{cases} \dfrac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!},\ t \geq 0, \\ 0,\quad t < 0. \end{cases}$$

17. Случайные величины $\xi_1,\ \xi_2$ независимы и экспоненциально распределены с параметром $\lambda = 1$. Найдите плотность распределения случайной величины η, если

$$\eta = \frac{\xi_1}{\xi_1 + \xi_2}.$$

Глава 4. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

В предыдущей главе было установлено, что каждая случайная величина характеризуется своей функцией распределения. Так, две случайные величины, имеющие одинаковые функции распределения, неразличимы, несмотря на то, что они могут быть заданы на различных вероятностных пространствах и описывать разные явления. Однако в том случае, когда случайные величины имеют различные функции распределения и их необходимо сравнить, возникают определенные трудности. В общем случае непонятно, как сравнивать функции распределения, и поэтому хотелось бы характеризовать каждую случайную величину некоторым (неслучайным) числом, возможно, несколькими числами, которые и позволили бы произвести упорядочение случайных величин в некотором смысле.

Такие числовые характеристики будут рассмотрены в этой главе.

§ 1. Математическое ожидание и его свойства

Пусть $(\Omega, \mathscr{A}, \mathrm{P})$ – произвольное вероятностное пространство, $\xi = \xi(\omega)$, $\omega \in \Omega$, – случайная величина на этом пространстве. Цель этого параграфа – определить число, которое играло бы роль вероятностного среднего значения случайной величины.

О п р е д е л е н и е 1. *Математическим ожиданием случайной величины* ξ *называется число*

$$\mathrm{M}\xi = \int_{\Omega} \xi(\omega)\, d\mathrm{P}(\omega) = \int_{\Omega} \xi(\omega)\, \mathrm{P}(d\omega).$$

З а м е ч а н и е 1. Математическое ожидание случайной величины ξ – это интеграл Лебега от этой величины по вероятностной мере (см. дополнения к гл. 4, § 1).

Свойства математического ожидания

Поскольку свойства математического ожидания – это свойства интеграла Лебега, то большинство из них приведем без доказательства.

1. *Если* $\mathrm{P}(\xi = c) = 1$, *то* $\mathrm{M}\xi = c$.

2. *Если* ξ *– интегрируемая неотрицательная случайная величина, то* $\mathrm{M}\xi \geq 0$.

3. *Если* $|\xi| \leq c$, *то* $\mathrm{M}|\xi| \leq c$.

4. *Если* $|\xi| \le \eta$ *и* η *– интегрируема, то* ξ *также интегрируема и*

$$\mathrm{M}|\xi| \le \mathrm{M}\eta.$$

5. *Если* ξ *и* η *– интегрируемые случайные величины, то для любых* a, $b \in R$ *случайная величина* $a\xi + b\eta$ *также интегрируема и*

$$\mathrm{M}(a\xi + b\eta) = a \cdot \mathrm{M}\xi + b \cdot \mathrm{M}\eta.$$

6. *Случайные величины* ξ *и* $|\xi|$ *одновременно интегрируемы и*

$$|\mathrm{M}\xi| \le \mathrm{M}|\xi|.$$

7. **Теорема о монотонной сходимости**. *Пусть* $\{\xi_n\}$, $n \ge 1$, *– последовательность неотрицательных интегрируемых случайных величин таких, что* $\xi_n \uparrow \xi$ *поточечно при* $n \to \infty$. *Если* $\sup\limits_n \mathrm{M}\xi_n < \infty$,*, то* ξ *также интегрируема и*

$$\mathrm{M}\xi = \lim_{n\to\infty} \mathrm{M}\xi_n.$$

8. **Теорема о мажорируемой сходимости**. *Если для всех* $\omega \in \Omega$

$$\xi_n \xrightarrow[n\to\infty]{} \xi,$$

$|\xi_n| \le \eta$, $n = 1,\ 2,\ \ldots$ *и* $\mathrm{M}\eta < \infty$, *то*

$$\mathrm{M}\xi = \lim_{n\to\infty} \mathrm{M}\xi_n.$$

9. **Свойство мультипликативности математических ожиданий**. *Пусть* $\xi_1,\ \ldots,\ \xi_n$ *– независимые случайные величины,* $\mathrm{M}\xi_i < \infty$, $i = \overline{1,n}$. *Тогда существует* $\mathrm{M}(\xi_1\xi_2\ldots\xi_n)$ *и*

$$\mathrm{M}(\xi_1\xi_2\ldots\xi_n) = \mathrm{M}\xi_1 \cdot \mathrm{M}\xi_2 \cdot \ldots \cdot \mathrm{M}\xi_n.$$

З а м е ч а н и е 1. Обратное утверждение, вообще говоря, неверно (приведите пример, показывающий это, см. также гл. 4, § 2 прим. 2).

Схема доказательства

Д о к а з а т е л ь с т в о проведем для случая $n = 2$. Общий случай использует математическую индукцию. Отметим, что доказательство проводится по схеме построения интеграла Лебега (см. дополнения к гл. 4).

1) Пусть ξ_1, ξ_2 – простые случайные величины, т. е.

$$\xi_1(\omega) = \sum_{k=1}^{m} c_k I_{A_k}(\omega), \quad \xi_2(\omega) = \sum_{j=1}^{l} b_j I_{B_j}(\omega),$$

где

$$\sum_{k=1}^{m} A_k = \Omega,\ \sum_{j=1}^{l} B_j = \Omega.$$

Из того, что случайные величины ξ_1, ξ_2 независимы, следует, что A_k и B_j независимы для любых $k=\overline{1,m}$, $j=\overline{1,l}$. Покажем это.

По критерию независимости дискретных случайных величин имеем

$$\mathrm{P}\left(\left(\xi_1=c_k\right)\cap\left(\xi_2=b_j\right)\right)=\mathrm{P}\left(\xi_1=c_k\right)\mathrm{P}\left(\xi_2=b_j\right)$$

для любых $k=\overline{1,m}$, $j=\overline{1,l}$. Но

$$\mathrm{P}(\xi_1=c_k\cap\xi_2=b_j)=\mathrm{P}(A_k\cap B_j),\ \mathrm{P}\left(\xi_1=c_k\right)=\mathrm{P}\left(A_k\right),$$

$$\mathrm{P}\left(\xi_2=b_j\right)=\mathrm{P}\left(B_j\right),$$

поэтому A_k и B_j независимы для любых $k=\overline{1,m}$, $j=\overline{1,l}$. Очевидно, что

$$\xi_1(\omega)\xi_2(\omega)=\sum_{k=1}^{m}\sum_{j=1}^{l}c_kb_jI_{A_k}(\omega)I_{B_j}(\omega)=\sum_{k=1}^{m}\sum_{j=1}^{l}c_kb_jI_{A_k\cap B_j}(\omega).$$

Теперь, учитывая все вышесказанное и определение математического ожидания для простых случайных величин, получаем

$$\mathrm{M}\xi_1\xi_2=\sum_{k=1}^{m}\sum_{j=1}^{l}c_kb_j\mathrm{P}\left(A_k\cap B_j\right)=\sum_{k=1}^{m}\sum_{j=1}^{l}c_kb_j\mathrm{P}\left(A_k\right)\mathrm{P}\left(B_j\right)=$$

$$=\sum_{k=1}^{m}c_k\mathrm{P}\left(A_k\right)\ \sum_{j=1}^{l}b_j\mathrm{P}\left(B_j\right)=\mathrm{M}\xi_1\,\mathrm{M}\xi_2.$$

2) Пусть ξ_1, $\xi_2\geq 0$ и независимы. Тогда найдутся последовательности простых случайных величин ξ_{1n} и ξ_{2n}, такие что

$$\xi_{1n}\uparrow\xi_1,\ \xi_{2n}\uparrow\xi_2$$

при $n\to\infty$ поточечно. В частности, в качестве искомых случайных величин можно взять следующие:

$$\xi_{in}(\omega)=\sum_{k=1}^{n2^n}\frac{k-1}{2^n}I_{\left(\frac{k-1}{2^n}<\xi_i\leq\frac{k}{2^n}\right)}(\omega),\qquad i=\overline{1,2}.$$

Несложно видеть, что случайные величины ξ_{1n}, ξ_{2n} будут независимыми, $n=1,\ 2,\ \ldots$. Поэтому в силу п. 1) этого доказательства

$$\mathrm{M}\xi_{1n}\xi_{2n}=\mathrm{M}\xi_{1n}\,\mathrm{M}\xi_{2n},\quad n=1,\ 2,\ldots\ .$$

Кроме того,

$$\xi_{1n}\cdot\xi_{2n}\underset{n\to\infty}{\uparrow}\xi_1\cdot\xi_2.$$

Теперь, используя теорему о монотонной сходимости (свойство 7 математических ожиданий), окончательно получим

$$\mathrm{M}\,\xi_1\xi_2 = \lim_{n\to\infty}\mathrm{M}\,\xi_{1n}\xi_{2n} = \lim_{n\to\infty}\mathrm{M}\,\xi_{1n}\lim_{n\to\infty}\mathrm{M}\,\xi_{2n} = \mathrm{M}\,\xi_1\,\mathrm{M}\,\xi_2.$$

3) Докажем свойство мультипликативности для произвольных независимых случайных величин ξ_1 и ξ_2. Для этого рассмотрим случайные величины ξ_1^+, ξ_1^-, ξ_2^+, ξ_2^-, положив по определению

$$\xi_i^+(\omega) = \max\left(0,\ \xi_i(\omega)\right),$$

$$\xi_i^-(\omega) = -\min\left(0,\ \xi_i(\omega)\right), \quad i=\overline{1,2}.$$

Очевидно, что $\xi_i^+ \geq 0$, $\xi_i^- \geq 0$, $\xi_i = \xi_i^+ - \xi_i^-$, $i=\overline{1,2}$. Кроме того, так как функции $g_1(x) = \max(0,\ x)$ и $g_2(x) = -\min(0,\ x)$ – борелевские (хотя бы потому, что они непрерывны), то из независимости случайных величин ξ_1 и ξ_2 вытекает независимость случайных величин ξ_1^+, ξ_1^- и ξ_2^+, ξ_2^-, а поэтому в силу п. 2) доказательства

$$\mathrm{M}\xi_1^+\xi_2^+ = \mathrm{M}\xi_1^+\mathrm{M}\xi_2^+,\quad \mathrm{M}\xi_1^+\xi_2^- = \mathrm{M}\xi_1^+\mathrm{M}\xi_2^-,\quad \mathrm{M}\xi_1^-\xi_2^+ = \mathrm{M}\xi_1^-\mathrm{M}\xi_2^+$$

и

$$\mathrm{M}\,\xi_1^-\xi_2^- = \mathrm{M}\,\xi_1^-\,\mathrm{M}\,\xi_2^-.$$

Учитывая все вышесказанное, получаем

$$\mathrm{M}\,\xi_1\xi_2 = \mathrm{M}\left(\xi_1^+ - \xi_1^-\right)\left(\xi_2^+ - \xi_2^-\right) = \mathrm{M}\,\xi_1^+\xi_2^+ - \mathrm{M}\,\xi_1^-\xi_2^+ - \mathrm{M}\,\xi_1^+\xi_2^- + \mathrm{M}\,\xi_1^-\xi_2^- =$$

$$= \mathrm{M}\,\xi_1^+\,\mathrm{M}\,\xi_2^+ - \mathrm{M}\,\xi_1^-\,\mathrm{M}\,\xi_2^+ - \mathrm{M}\,\xi_1^+\,\mathrm{M}\,\xi_2^- + \mathrm{M}\,\xi_1^-\,\mathrm{M}\,\xi_2^- =$$

$$= \left(\mathrm{M}\,\xi_1^+ - \mathrm{M}\,\xi_1^-\right)\left(\mathrm{M}\,\xi_2^+ - \mathrm{M}\,\xi_2^-\right) = \mathrm{M}\left(\xi_1^+ - \xi_1^-\right)\cdot\mathrm{M}\left(\xi_2^+ - \xi_2^-\right) = \mathrm{M}\,\xi_1\,\mathrm{M}\,\xi_2.$$

Свойство мультипликативности математических ожиданий доказано.

10. **Формула замены переменных в интеграле Лебега**. *Пусть функция* $g: R \to R$ *непрерывна и* $\mathrm{M}\,g(\xi) < \infty$, *тогда*

$$\mathrm{M}\,g(\xi) = \int_\Omega g\left(\xi(\omega)\right)d\,\mathrm{P}(\omega) = \int_{-\infty}^{\infty} g(x)\,dF_\xi(x). \qquad (1)$$

З а м е ч а н и е 2. Данное утверждение справедливо и в случае, когда g – борелевская функция.

Схема доказательства

Д о к а з а т е л ь с т в о проведем в три этапа.

1) Пусть $g(x) = 0$ вне отрезка $[a;b]$. Рассмотрим разбиение $[a;b]$, положив по определению

$$x_{nk} = a + \frac{b-a}{n}\times k, \quad k=\overline{0,n},$$

и введем функцию

$$g_n(x)=\begin{cases}0, & x\notin(a;b],\\ g(x_{nk}), & x\in\left(x_{n,k-1};x_{nk}\right].\end{cases}$$

Тогда для любых $x\in R$ и $\varepsilon>0$ существует такое натуральное число n_0, что для любых $n>n_0$

$$\left|g_n(x)-g(x)\right|<\varepsilon,\quad \forall x\in[a;b]. \tag{2}$$

Неравенство (2) следует из того, что функция g равномерно непрерывна на отрезке $[a;b]$.

Из теоремы Лебега о предельном переходе под знаком интеграла в силу справедливости (2) получим

$$\mathrm{M}\,g(\xi)=\lim_{n\to\infty}\mathrm{M}\,g_n(\xi). \tag{3}$$

Так как $g_n(\xi)$ – простая случайная величина, то

$$\mathrm{M}\,g_n(\xi)=\sum_{k=1}^{n}g(x_{nk})\cdot\mathrm{P}\left(x_{n,k-1}<\xi\le x_{nk}\right)=\sum_{k=1}^{n}g(x_{nk})\times$$

$$\times\left[F_\xi(x_{nk})-F_\xi\left(x_{n,k-1}\right)\right]=\sum_{k=1}^{n}g(x_{nk})\int\limits_{x_{n,k-1}}^{x_{nk}}dF_\xi(x)=$$

$$=\sum_{k=1}^{n}\int\limits_{x_{n,k-1}}^{x_{nk}}g(x_{nk})\,dF_\xi(x)=\int\limits_a^b g_n(x)\,dF_\xi(x). \tag{4}$$

Из неравенства (2) следует, что

$$\left|\int\limits_a^b g(x)\,dF_\xi(x)-\int\limits_a^b g_n(x)\,dF_\xi(x)\right|<\varepsilon, \tag{5}$$

а из соотношений (3) – (5) следует доказательство п. 1).

2) Пусть $g(x)\ge 0,\ x\in R$ – непрерывная функция. Положим по определению

$$g_N(x)=\begin{cases}0, & |x|>N,\\ g(x), & |x|\le N.\end{cases}$$

Тогда очевидно, что $g_N(x)\underset{n\to\infty}{\uparrow}g(x),\ x\in R,$ и поэтому по теореме о монотонной сходимости

$$\mathrm{M}\,g(\xi)=\lim_{N\to\infty}\mathrm{M}\,g_N(\xi). \tag{6}$$

В силу уже доказанного п. 1)

$$\mathrm{M}g_N(\xi) = \int_{-N}^{N} g(x)\,dF_\xi(x). \tag{7}$$

Так как

$$\int_{-N}^{N} g(x)\,dF_\xi(x) \xrightarrow[N\to\infty]{} \int_{-\infty}^{\infty} g(x)F_\xi(x), \tag{8}$$

то из соотношений (6) – (8) следует доказательство п. 2).

3) Пусть $g: R \to R$ – непрерывная функция. Положим по определению

$$g^+(x) = \max\big(0,\ g(x)\big),\ \ g^-(x) = -\min\big(0,\ g(x)\big),\ \ x \in R.$$

Тогда

$$g^+(x) \ge 0,\ \ g^-(x) \ge 0$$

и

$$g(x) = g^+(x) - g^-(x),\ \ x \in R.$$

Согласно доказанному п. 2)

$$\mathrm{M}\,g^+(\xi) = \int_{-\infty}^{\infty} g^+(x)\,dF_\xi(x), \qquad \mathrm{M}\,g^-(\xi) = \int_{-\infty}^{\infty} g^-(x)\,dF_\xi(x),$$

а поэтому в силу линейности интеграла Римана – Стилтьеса окончательно получаем

$$\mathrm{M}\,g(\xi) = \mathrm{M}\big(g^+(\xi) - g^-(\xi)\big) = \int_{-\infty}^{\infty} g^+(x)\,dF_\xi(x) - \int_{-\infty}^{\infty} g^-(x)\,dF_\xi(x) =$$

$$= \int_{-\infty}^{\infty} \big(g^+(x) - g^-(x)\big)\,dF_\xi(x) = \int_{-\infty}^{\infty} g(x)\,dF_\xi(x).$$

Свойство 10 математического ожидания доказано.

З а м е ч а н и е 3. Свойство 10 справедливо и в том случае, когда $\overline{\xi} = (\xi_1,\ \ldots,\ \xi_n)$ – случайный вектор, $g: R^n \to R$ – борелевская функция и $\mathrm{M}\,g\big(\overline{\xi}\big) < \infty$.

Из свойства 10 и соответствующих свойств интеграла Лебега – Стилтьеса вытекают следующие утверждения.

С л е д с т в и е 1. *Пусть* ξ *– абсолютно непрерывная случайная величина с плотностью распределения* $p_\xi(x)$, $g: R \to R$ *– борелевская функция,* $\mathrm{M}\,g(\xi) < \infty$. *Тогда*

$$\mathrm{M}g(\xi)=\int_{-\infty}^{\infty}g(x)\,p_\xi(x)\,dx.$$

С л е д с т в и е 2. *Пусть* ξ *– дискретная случайная величина с законом распределения*

$$\begin{array}{llllll} \xi: & x_1 & x_2 & \dots & x_n & \dots, \\ \mathrm{P}: & p_1 & p_2 & \dots & p_n & \dots, \end{array}$$

$g: R \to R$ *– борелевская функция,* $\mathrm{M}\,g(\xi)<\infty$. *Тогда*

$$\mathrm{M}g(\xi)=\sum_{n=1}^{\infty}g(x_n)\,p_n.$$

П р и м е р 1. Пусть случайная величина ξ принимает конечное число различных значений $a_1,\ a_2,\ \dots,\ a_m;\ \ p_k=\mathrm{P}(\xi=a_k),\ \ k=\overline{1,m}.$ Будем повторять n раз эксперимент, в котором наблюдается величина ξ. Через $x_1,\ \dots,\ x_n$ обозначим наблюдаемые значения этой случайной величины. Тогда среднее значение наблюдаемых чисел равно

$$\frac{1}{n}\sum_{k=1}^{n}x_k=\sum_{i=1}^{m}a_i\frac{k_i}{n},$$

где k_i – число появлений события $(\xi=a_i)$ в этой серии экспериментов. Т. е.

$$\frac{1}{n}\sum_{k=1}^{n}x_k=\sum_{i=1}^{m}a_i\nu_n(\xi=a_i), \tag{9}$$

где $\nu_n(A)$ – частота появления события A в n испытаниях. Заметим, что при $n\to\infty$

$$\nu_n(A)\to\mathrm{P}(A),$$

(см. дополнения к гл.1, § 3), т. е.

$$\frac{1}{n}\sum_{k=1}^{n}x_k\to\mathrm{M}\xi.$$

Таким образом, математическое ожидание случайной величины ξ можно трактовать как *вероятностное среднее* этой величины.

П р и м е р 2. Математическое ожидание дискретной случайной величины имеет аналог в теоретической механике. Пусть на прямой расположена система материальных точек с массами p_n и пусть x_n – координата n-ой точки. Тогда центр тяжести системы будет иметь координату

$$\overline{x}=\frac{\sum\limits_n x_n p_n}{\sum\limits_n p_n}=\frac{\sum\limits_n x_n p_n}{1}=\sum_n x_n p_n,$$

совпадающую с математическим ожиданием случайной величины ξ, имеющей закон распределения

$$\xi: \quad x_1 \quad x_2 \quad \dots \quad x_n \quad \dots,$$

$$\mathrm{P}: \quad p_1 \quad p_2 \quad \dots \quad p_n \quad \dots .$$

П р и м е р 3. Пусть случайная величина ξ имеет распределение Бернулли

$$\xi: \quad 0 \quad 1$$

$$\mathrm{P}: \quad 1-p \quad p,$$

где $0<p<1$. Тогда

$$\mathrm{M}\xi = 0\cdot(1-p)+1\cdot p = p.$$

П р и м е р 4. М а т е м а т и ч е с к о е о ж и д а н и е б и н о м и а л ь н о й с л у ч а й н о й в е л и ч и н ы. Пусть закон распределения случайной величины ξ имеет следующий вид

$$\mathrm{P}(\xi=k)=C_n^k p^k (1-p)^{n-k}, \quad k=\overline{0,n}, \quad 0<p<1.$$

Тогда согласно следствию 2

$$\mathrm{M}\xi=\sum_{k=0}^{n} kC_n^k p^k (1-p)^{n-k} = np\sum_{k=1}^{n}\frac{k}{n}C_n^k p^k (1-p)^{n-k} =$$

$$= np\sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} = np\sum_{k=0}^{n-1} C_{n-1}^{k} p^{k} (1-p)^{n-k-1} = np.$$

П р и м е р 5. М а т е м а т и ч е с к о е о ж и д а н и е г е о м е т р и ч е с к о й с л у ч а й н о й в е л и ч и н ы. Пусть

$$\mathrm{P}(\xi=k)=(1-p)^k p, \quad k=1,\ 2,\ \dots,\ 0<p<1.$$

Тогда

$$\mathrm{M}\xi=\sum_{k=0}^{\infty} k(1-p)^k p = (1-p)p\left(1+2(1-p)+3(1-p)^2+\dots\right)=$$

$$=\frac{(1-p)p}{p^2}=\frac{1-p}{p}.$$

П р и м е р 6. М а т е м а т и ч е с к о е о ж и д а н и е с л у ч а й н о й в е л и ч и н ы П у а с с о н а с п а р а м е т р о м λ. Пусть

$$\mathrm{P}(\xi=k)=\frac{\lambda^k}{k!}e^{-\lambda}, \quad k=0,\ 1,\ 2,\dots,\ \lambda>0.$$

Тогда

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

П р и м е р 7. М а т е м а т и ч е с к о е о ж и д а н и е р а в н о м е р н о й н а о т р е з к е $[a;b]$ с л у ч а й н о й в е л и ч и н ы. Пусть плотность распределения ξ равна

$$p_\xi(x)=\begin{cases}\dfrac{1}{b-a}, & x\in[a;b],\\ 0, & x\notin[a;b].\end{cases}$$

Тогда по следствию 1

$$\mathrm{M}\,\xi=\int_{-\infty}^{\infty} x p_{\xi}(x)\,dx=\int_{a}^{b}\frac{x}{b-a}\,dx=\frac{a+b}{2}.$$

Пример 8. Математическое ожидание показательной с параметром $\lambda\,(\lambda>0)$ случайной величины. Пусть

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

Тогда

$$\mathrm{M}\,\xi=\int_{-\infty}^{\infty} x p_{\xi}(x)\,dx=\int_{0}^{\infty} x\lambda e^{-\lambda x}dx=-\int_{0}^{\infty} x d\left(e^{-\lambda x}\right)=-xe^{-\lambda x}\Big|_{0}^{\infty}+\int_{0}^{\infty} e^{-\lambda x}dx=\frac{1}{\lambda}.$$

Пример 9. Математическое ожидание нормальной случайной величины с параметрами $\left(a,\,\sigma^2\right)$. Пусть

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

Тогда

$$\mathrm{M}\,\xi=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} xe^{-\frac{(x-a)^2}{2\sigma^2}}dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-a)e^{-\frac{(x-a)^2}{2\sigma^2}}dx+$$

$$+\,a\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2\sigma^2}}dx=a,$$

так как

$$\int_{-\infty}^{\infty}(x-a)e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\int_{-\infty}^{\infty} ze^{-\frac{z^2}{2\sigma^2}}dz=0,\qquad \int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\sigma\sqrt{2\pi}.$$

Пример 10. Математическое ожидание стандартной случайной величины Коши. Пусть

$$p_{\xi}(x)=\frac{1}{\pi}\cdot\frac{1}{1+x^2}.$$

В этом случае

$$\int_{-\infty}^{\infty}|x|\,p_{\xi}(x)\,dx=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{|x|}{1+x^2}dx=+\infty,$$

и, следовательно, математическое ожидание случайной величины ξ не существует.

§ 2. Моменты случайных величин

Математическое ожидание не всегда является удовлетворительной характеристикой случайной величины. Наряду со средним значением хотелось бы, по крайней мере, иметь и число (числа), характеризующее «разброс» случайной величины относительно своего среднего значения, что и будет рассмотрено в этом параграфе.

Пусть ξ – случайная величина, определенная на вероятностном пространстве $(\Omega, \mathscr{A}, \mathrm{P})$. Для натуральных m величина $\mathrm{M}\xi^m$, если она определена, называется *моментом m-го порядка* случайной величины ξ. В частности, момент первого порядка ξ – это ее математическое ожидание. Величина $\mathrm{M}|\xi|^m$ называется *абсолютным моментом m-го порядка* случайной величины ξ. Моменты (абсолютные моменты) случайной величины $(\xi - \mathrm{M}\xi)$ называются *центральными моментами (абсолютными центральными моментами)* случайной величины ξ.

Несложно видеть, что центральные моменты четного порядка и абсолютные центральные моменты случайной величины ξ характеризуют степень «разброса» значений этой величины относительного ее среднего значения.

Наиболее употребительной числовой характеристикой случайной величины является ее второй центральный момент.

О п р е д е л е н и е 1. *Дисперсией случайной величины ξ называется число*

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

число $\sigma_\xi = \sqrt{\mathrm{D}\xi}$ называется среднеквадратическим отклонением случайной величины ξ.

Свойства дисперсии

1. $\mathrm{D}\xi = \mathrm{M}\xi^2 - (\mathrm{M}\xi)^2$.

Д о к а з а т е л ь с т в о. Учитывая определение дисперсии и линейность математического ожидания, имеем

$$\mathrm{D}\xi = \mathrm{M}(\xi - \mathrm{M}\xi)^2 = \mathrm{M}\left(\xi^2 - 2\xi\mathrm{M}\xi + (\mathrm{M}\xi)^2\right) =$$

$$= \mathrm{M}\xi^2 - 2(\mathrm{M}\xi)^2 + (\mathrm{M}\xi)^2 = \mathrm{M}\xi^2 - (\mathrm{M}\xi)^2.$$

2. $\mathrm{D}\xi \geq 0$.

Д о к а з а т е л ь с т в о. Так как подынтегральная функция $(\xi - \mathrm{M}\xi)^2 \geq 0$, то по свойству 2 интеграла Лебега $\mathrm{D}\xi \geq 0$.

3. $\mathrm{D}(\xi + c) = \mathrm{D}\xi$, *где c – константа.*

Д о к а з а т е л ь с т в о. Из определения дисперсии и линейности математического ожидания следует

$$\mathrm{D}(\xi+c)=\mathrm{M}\left((\xi+c)-\mathrm{M}(\xi+c)\right)^2=\mathrm{M}(\xi+c-\mathrm{M}\xi-c)^2=$$
$$=\mathrm{M}(\xi-\mathrm{M}\xi)^2=\mathrm{D}\xi.$$

4. $\mathrm{D}(c\xi)=c^2\mathrm{D}\xi$, *где c – константа.*

Доказательство аналогично доказательству свойства 3:

$$\mathrm{D}(c\xi)=\mathrm{M}\left(c\xi-\mathrm{M}(c\xi)\right)^2=\mathrm{M}(c\xi-c\mathrm{M}\xi)^2=c^2\mathrm{M}(\xi-\mathrm{M}\xi)^2=c^2\mathrm{D}\xi.$$

5. $\mathrm{D}\xi=0$ *тогда и только тогда, когда существует константа c такая, что* $\mathrm{P}(\xi=c)=1$.

Доказательство. Необходимость. Пусть $\mathrm{D}\xi=0$, т. е. $\mathrm{M}(\xi-\mathrm{M}\xi)^2=0$. По свойствам интеграла Лебега $\mathrm{P}(\xi-\mathrm{M}\xi=0)=1$, значит, $\mathrm{P}(\xi=\mathrm{M}\xi)=1$. Таким образом, существует константа $c=\mathrm{M}\xi$ такая, что $\mathrm{P}(\xi=c)=1$. Необходимость доказана.

Достаточность. Пусть существует константа c такая, что $\mathrm{P}(\xi=c)=1$. Тогда $\mathrm{D}\xi=\mathrm{D}c=\mathrm{M}(c-\mathrm{M}c)^2=\mathrm{M}0=0$. Свойство 5 доказано.

Для формулировки следующего свойства понадобится числовая характеристика случайных величин ξ и η, устанавливающая вероятностную связь между ними.

Определение 2. *Ковариацией случайных величин ξ и η называется число*

$$\mathrm{cov}(\xi,\ \eta)=\mathrm{M}(\xi-\mathrm{M}\xi)(\eta-\mathrm{M}\eta).$$

6. $\mathrm{D}(\xi+\eta)=\mathrm{D}\xi+\mathrm{D}\eta+2\mathrm{cov}(\xi,\ \eta)$.

Доказательство. Воспользовавшись определениями дисперсии, ковариации и линейностью математического ожидания, получаем

$$\mathrm{D}(\xi+\eta)=\mathrm{M}\left((\xi+\eta)-\mathrm{M}(\xi+\eta)\right)^2=\mathrm{M}\left((\xi-\mathrm{M}\xi)+(\eta-\mathrm{M}\eta)\right)^2=$$
$$=\mathrm{M}(\xi-\mathrm{M}\xi)^2+2\mathrm{M}(\xi-\mathrm{M}\xi)(\eta-\mathrm{M}\eta)+\mathrm{M}(\eta-\mathrm{M}\eta)^2=$$
$$\mathrm{D}\xi+2\mathrm{cov}(\xi,\eta)+\mathrm{D}\eta.$$

7. *Если случайные величины ξ и η независимы, то*

$$\mathrm{D}(\xi+\eta)=\mathrm{D}\xi+\mathrm{D}\eta.$$

Доказательство. Учитывая свойство 6, для доказательства свойства 7 достаточно показать, что $\mathrm{cov}(\xi,\eta)=0$, если ξ и η независимы.

Отметим, что из независимости случайных величин ξ и η вытекает независимость $\xi - \mathrm{M}\xi$ и $\eta - \mathrm{M}\eta$, так как $\xi - \mathrm{M}\xi = g_1(\xi)$, $\eta - \mathrm{M}\eta = g_2(\eta)$, где $g_1(x) = x - \mathrm{M}\xi$, $g_2(x) = x - \mathrm{M}\eta$ – борелевские функции. Поэтому по свойству мультипликативности математических ожиданий получаем

$$\operatorname{cov}(\xi,\eta) = \mathrm{M}(\xi - \mathrm{M}\xi)\mathrm{M}(\eta - \mathrm{M}\eta) =$$
$$= \mathrm{M}(\xi - \mathrm{M}\xi)(\eta - \mathrm{M}\eta) = (\mathrm{M}\xi - \mathrm{M}\xi)(\mathrm{M}\eta - \mathrm{M}\eta) = 0.$$

Упражнение 1. Воспользовавшись свойством 10 математического ожидания и следствиями 1 и 2 § 1, докажите, что

1) $$\mathrm{M}\xi^k = \int_{-\infty}^{\infty} x^k dF_\xi(x),$$

$$\mathrm{M}|\xi|^k = \int_{-\infty}^{\infty} |x|^k dF_\xi(x),$$

$$\mathrm{M}(\xi - \mathrm{M}\xi)^k = \int_{-\infty}^{\infty} (x - \mathrm{M}\xi)^k dF_\xi(x),$$

$$\mathrm{M}|\xi - \mathrm{M}\xi|^k = \int_{-\infty}^{\infty} |x - \mathrm{M}\xi|^k dF_\xi(x), \qquad k = 1,\ 2,\ \ldots;$$

2) если случайная величина ξ – абсолютно непрерывна с плотностью распределения p_ξ, то

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

$$\mathrm{M}|\xi|^k = \int_{-\infty}^{\infty} |x|^k p_\xi(x)\,dx,$$

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

$$\mathrm{M}|\xi - \mathrm{M}\xi|^k = \int_{-\infty}^{\infty} |x - \mathrm{M}\xi|^k p_\xi(x)\,dx, \qquad k = 1,\ 2,\ \ldots;$$

3) если случайная величина ξ дискретна с законом распределения

$$\begin{array}{llll} \xi: & x_1 \ \ldots & x_n & \ldots \\ \mathrm{P}: & p_1 \ \ldots & p_n & \ldots, \end{array}$$

то

$$\mathrm{M}\xi^k = \sum_{n=1}^{\infty} x_n^k p_n,$$

$$\mathrm{M}|\xi|^k = \sum_{n=1}^{\infty} |x_n|^k \, p_n,$$

$$\mathrm{M}(\xi - \mathrm{M}\xi)^k = \sum_{n=1}^{\infty} (x_n - \mathrm{M}\xi)^k \, p_n,$$

$$\mathrm{M}|\xi - \mathrm{M}\xi|^k = \sum_{n=1}^{\infty} |x_n - \mathrm{M}\xi|^k \, p_n, \qquad k = 1,\ 2,\ \ldots\ .$$

П р и м е р 1. Пусть ξ – случайная величина, имеющая дисперсию $\mathrm{D}\xi$. Введем новую случайную величину $\eta = \xi - a$. Найдем число a, доставляющее минимум $\mathrm{M}\eta^2$. Воспользовавшись свойствами дисперсии, имеем

$$\mathrm{M}\eta^2 = \mathrm{D}\eta + (\mathrm{M}\eta)^2 = \mathrm{D}\xi + (\mathrm{M}\xi - a)^2.$$

Первое слагаемое от a не зависит, а второе принимает минимальное значение при $a = \mathrm{M}\xi$. Таким образом, в качестве a следует взять $\mathrm{M}\xi$, а минимальное значение $\mathrm{M}\eta^2$ совпадает с дисперсией $\mathrm{D}\xi$.

П р и м е р 2. Бросаются две игральные кости. Пусть ξ_1 – число очков, выпавших на одной кости, а ξ_2 – на другой. Рассмотрим случайные величины $\eta_1 = \xi_1 + \xi_2$ (сумма очков на обеих костях), $\eta_2 = \xi_1 - \xi_2$ (разность очков). Тогда

$$\operatorname{cov}(\eta_1, \eta_2) = \mathrm{M}\big((\xi_1 + \xi_2) - \mathrm{M}(\xi_1 + \xi_2)\big)\big((\xi_1 - \xi_2) - \mathrm{M}(\xi_1 - \xi_2)\big) =$$
$$= \mathrm{M}\big((\xi_1 - \mathrm{M}\xi_1)^2 - (\xi_2 - \mathrm{M}\xi_2)^2\big) = \mathrm{D}\xi_1 - \mathrm{D}\xi_2.$$

Случайные величины ξ_1 и ξ_2 одинаково распределены, значит, $\mathrm{D}\xi_1 = \mathrm{D}\xi_2$ и $\operatorname{cov}(\eta_1, \eta_2) = 0$. Однако, случайные величины η_1 и η_2 зависимы, поскольку, например, из равенства $\eta_1 = 2$ обязательно следует равенство $\eta_2 = 0$. Таким образом, в свойстве 7 дисперсии обратное утверждение, вообще говоря, неверно.

П р и м е р 3. Д и с п е р с и я с л у ч а й н о й в е л и ч и н ы Б е р н у л л и. Поскольку $\mathrm{M}\xi = p$ (см. гл. 4, § 1), то по определению дисперсии, учитывая упражнение 1, имеем

$$\mathrm{D}\xi = \mathrm{M}(\xi - p)^2 = (0 - p)^2 (1 - p) + (1 - p)^2 p =$$
$$= p^2 - p^3 + p - 2p^2 + p^3 = p(1 - p).$$

П р и м е р 4. Д и с п е р с и я б и н о м и а л ь н о й с л у ч а й н о й в е л и ч и н ы . Учитывая то, что биномиальная случайная величина ξ имеет то же распределение, что и сумма $\xi_1 + \ldots + \xi_n$ независимых случайных величин Бернулли, по свойствам дисперсии имеем

$$\mathrm{D}\xi = \mathrm{D}(\xi_1 + \ldots + \xi_n) = \sum_{k=1}^{n} \mathrm{D}\xi_k = np(1 - p).$$

П р и м е р 5. Д и с п е р с и я   с л у ч а й н о й   в е л и ч и н ы   П у а с с о н а   с   п а р а м е т р о м λ. Математическое ожидание этой случайной величины было найдено в § 1 прим. 6, $\mathrm{M}\,\xi=\lambda.$ Вычислим второй момент

$$\mathrm{M}\,\xi^2=\sum_{k=0}^{\infty}k^2\frac{\lambda^k}{k!}e^{-\lambda}=\lambda\sum_{k=1}^{\infty}k\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}=\lambda\sum_{j=0}^{\infty}(j+1)\frac{\lambda^j}{j!}e^{-\lambda}=$$

$$=\lambda\left(\sum_{j=0}^{\infty}j\frac{\lambda^j}{j!}e^{-\lambda}+\sum_{j=0}^{\infty}\frac{\lambda^j}{j!}e^{-\lambda}\right)=\lambda(\mathrm{M}\,\xi+1)=\lambda^2+\lambda.$$

Следовательно,

$$\mathrm{D}\,\xi=\mathrm{M}\,\xi^2-(\mathrm{M}\,\xi)^2=\lambda^2+\lambda-\lambda^2=\lambda.$$

П р и м е р 6. Д и с п е р с и я   г е о м е т р и ч е с к о й   с л у ч а й н о й   в е л и ч и н ы. Известно, что (см. гл. 4, § 1, прим. 5)

$$\mathrm{M}\,\xi=\frac{1-p}{p}.$$

Найдем момент второго порядка, учитывая следующие равенства:

$$\sum_{k=0}^{\infty}(1-p)^k=\frac{1}{p},\qquad \sum_{k=0}^{\infty}k(1-p)^{k-1}=\frac{1}{p^2},\qquad \sum_{k=1}^{\infty}k(k-1)(1-p)^{k-2}=\frac{2}{p^3}.$$

Заметим, что второе равенство получается в результате дифференцирования по переменной p обеих частей первого равенства, а третье аналогичным образом получается из второго. Таким образом

$$\mathrm{M}\,\xi^2=\sum_{k=0}^{\infty}k^2(1-p)^k\,p=p\sum_{k=0}^{\infty}k(k-1)(1-p)^k+\sum_{k=0}^{\infty}k(1-p)^k\,p=$$

$$p(1-p)^2\frac{2}{p^3}+\frac{1-p}{p}=\frac{2(1-p)^2}{p^2}+\frac{1-p}{p},$$

а

$$\mathrm{D}\,\xi=\mathrm{M}\,\xi^2-(\mathrm{M}\,\xi)^2=\frac{2(1-p)^2}{p^2}+\frac{1-p}{p}-\frac{(1-p)^2}{p^2}=\frac{(1-p)^2}{p^2}+\frac{1-p}{p}=\frac{1-p}{p^2}.$$

П р и м е р 7. Д и с п е р с и я   р а в н о м е р н о   р а с п р е д е л е н н о й   н а   о т р е з к е $[a;b]$ с л у ч а й н о й   в е л и ч и н ы. Известно, что (см. гл. 4, § 1, прим. 7)

$$\mathrm{M}\,\xi=\frac{a+b}{2}.$$

Найдем $\mathrm{M}\,\xi^2$. Имеем

$$\mathrm{M}\,\xi^2=\frac{1}{b-a}\int_a^b x^2dx=\frac{1}{3}\frac{b^3-a^3}{b-a}=\frac{1}{3}\left(b^2+ab+a^2\right),$$

поэтому

$$\mathrm{D}\,\xi=\mathrm{M}\,\xi^2-(\mathrm{M}\,\xi)^2=\frac{1}{3}\left(b^2+ab+a^2\right)-\frac{1}{4}\left(b^2+2ab+a^2\right)=\frac{(a-b)^2}{12}.$$

Пример 8. Центральные моменты нормальной случайной величины с параметрами (a, σ^2). Известно (см. гл. 4, § 1, прим. 9), что $\mathrm{M}\xi = a$. Используя упражнение 1, имеем

$$\mathrm{M}(\xi - \mathrm{M}\xi)^n = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-a)^n e^{-\frac{1}{2}\frac{(x-a)^2}{\sigma^2}} dx.$$

В результате подстановки $\dfrac{x-a}{\sigma} = z$ получим

$$\mathrm{M}(\xi - \mathrm{M}\xi)^n = \frac{\sigma^n}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^n e^{-\frac{z^2}{2}} dz.$$

Если n – нечетное, то $\mathrm{M}(\xi - \mathrm{M}\xi)^n = 0$, если же n – четное, то

$$\mathrm{M}(\xi - \mathrm{M}\xi)^{2k} = \frac{2\sigma^{2k}}{\sqrt{2\pi}} \int_{0}^{\infty} z^{2k} e^{-\frac{z^2}{2}} dz,$$

и, выполнив замену переменной $\dfrac{1}{2}z^2 = t$, получим

$$\mathrm{M}(\xi - \mathrm{M}\xi)^{2k} = \frac{2^k\sigma^{2k}}{\sqrt{\pi}} \int_{0}^{\infty} t^{k-\frac{1}{2}} e^{-t} dt = \frac{2^k\sigma^{2k}}{\sqrt{\pi}} \Gamma\left(\mathrm{k} + \frac{1}{2}\right) = (2k-1)!!\sigma^{2k}.$$

В частности, $\mathrm{D}\xi = \sigma^2$.

Пример 9. Моменты показательной с параметром λ случайной величины. По формулам для абсолютно непрерывного случая (см. упражнение 1) имеем

$$\mathrm{M}\xi^k = \int_0^{\infty} x^k \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} x^k e^{-\lambda x} dx = \frac{1}{\lambda^k} \int_0^{\infty} t^k e^{-t} dt = \frac{\Gamma(k+1)}{\lambda^k} = \frac{k!}{\lambda^k}.$$

В частности,

$$\mathrm{M}\xi = \frac{1}{\lambda}, \ \mathrm{M}\xi^2 = \frac{2}{\lambda^2}.$$

§ 3. Неравенства. Коэффициент корреляции

В этом параграфе рассмотрены неравенства для математических ожиданий, которые часто используются в дальнейшем изложении, а также коэффициент корреляции – число, характеризующее степень зависимости двух случайных величин.

Неравенство Чебышева. *Если* $\mathrm{g} = \mathrm{g}(x)$ *– положительная неубывающая на* $[0;\infty)$ *функция, то для любого* $\varepsilon > 0$

$$\mathrm{P}(|\xi| \ge \varepsilon) \le \frac{\mathrm{M}g(\xi)}{g(\varepsilon)}. \qquad (1)$$

Д о к а з а т е л ь с т в о. Из свойств математического ожидания имеем

$$\mathrm{M}g(\xi)\geq \mathrm{M}g(\xi)I_{\{|\xi|\geq\varepsilon\}}\geq \mathrm{M}g(\varepsilon)I_{\{|\xi|\geq\varepsilon\}}=$$

$$=g(\varepsilon)\mathrm{M}I_{\{|\xi|\geq\varepsilon\}}=g(\varepsilon)\mathrm{P}(|\xi|\geq\varepsilon).$$

Неравенство (1) доказано.

С л е д с т в и е 1. *Для любого* $\varepsilon>0$

$$\mathrm{P}(|\xi|\geq\varepsilon)\leq\frac{\mathrm{M}|\xi|}{\varepsilon},\quad \mathrm{P}(|\xi-\mathrm{M}\xi|\geq\varepsilon)\leq\frac{\mathrm{D}\xi}{\varepsilon^2}.$$

Д о к а з а т е л ь с т в о. Первое неравенство непосредственно вытекает из неравенства Чебышева, если $g(x)=x$, а второе – доказывается также, как и неравенство (1).

Н е р а в е н с т в о И е н с е н а. *Пусть* $g: R\to R$ *– борелевская выпуклая вниз (вверх) функция и* $\mathrm{M}\xi<\infty$, *тогда*

$$\mathrm{M}g(\xi)\geq g(\mathrm{M}\xi)\quad (g(\mathrm{M}\xi)\geq \mathrm{M}g(\xi)).$$

Д о к а з а т е л ь с т в о. Напомним, что функция $g:\ (c;d)\to R$, $-\infty\leq c<d\leq+\infty$ называется *выпуклой вниз функцией*, если для любых $x_1,\ x_2\in(c;d)$ и любого $\theta\in[0;1]$

$$g(\theta x_1+(1-\theta)x_2)\leq\theta g(x_1)+(1-\theta)g(x_2).\tag{2}$$

Выпуклая вниз функция обладает следующим свойством: для любых $x,\ a\in(c;d)$ существует константа m такая, что

$$g(x)-g(a)\geq m(x-a).\tag{3}$$

Докажем неравенство (3). Пусть точки $x_1,\ x_2\in(c;d)$ таковы, что $x_1<a<x_2$. Возьмем число $\theta=\dfrac{x_2-a}{x_2-x_1}$.. Подставим так определенное θ в неравенство (2). Тогда

$$g(a)\leq\frac{x_2-a}{x_2-x_1}\cdot g(x_1)+\frac{a-x_1}{x_2-x_1}\cdot g(x_2)$$

или после очевидных преобразований

$$(x_2-a)(g(a)-g(x_1))\leq(a-x_1)(g(x_2)-g(a)).$$

Из последнего неравенства следует, что существует такая константа m, что

$$\frac{g(x_1)-g(a)}{x_1-a}\leq m\leq\frac{g(x_2)-g(a)}{x_2-a}.$$

Неравенство (3) для $x<a$ вытекает из первого неравенства, а для $x>a$ – из второго.

Подставив в неравенство (3) $x=\xi$ и $a=\mathrm{M}\xi$, получим

$$g(\xi)\ge g(\mathrm{M}\xi)+m(\xi-\mathrm{M}\xi).$$

Взяв математические ожидания от обеих частей последнего неравенства, приходим к доказательству неравенства Иенсена для выпуклой вниз функции. Если же функция g выпуклая вверх, то все неравенства в доказательстве поменяются на противоположные.

У п р а ж н е н и е 1. Можно ли в неравенстве Иенсена снять условие борелевости на функцию g?

Н е р а в е н с т в о Л я п у н о в а. *Для любых положительных* $\alpha<\beta$

$$\left(\mathrm{M}|\xi|^{\alpha}\right)^{1/\alpha}\le\left(\mathrm{M}|\xi|^{\beta}\right)^{1/\beta}.$$

Д о к а з а т е л ь с т в о получается из неравенства Иенсена $\mathrm{M}g(\eta)\ge g(\mathrm{M}\eta)$, если в качестве g взять функцию $g(x)=x^{\beta/\alpha}$, $x\in[0;\infty)$, а в качестве случайной величины η – случайную величину $|\xi|^{\alpha}$.

З а м е ч а н и е 1. Из неравенства Ляпунова следует, что если конечен момент порядка k, то конечны и все моменты порядка меньше k, $k\in N$.

Н е р а в е н с т в о Г ё л ь д е р а. *Если*

$$p>1,\ q>1,\ \frac{1}{p}+\frac{1}{q}=1,\ \mathrm{M}|\xi|^{p}<\infty,\ \mathrm{M}|\eta|^{q}<\infty,$$

то

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

Д о к а з а т е л ь с т в о. Так как функция $y=\ln x$ является выпуклой вверх на $(0;\infty)$, то для любых $x_1>0$, $x_2>0$ и $\theta\in[0;1]$

$$\ln\left(\theta x_1+(1-\theta)x_2\right)\ge\theta\ln x_1+(1-\theta)\ln x_2.$$

Отсюда

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

Взяв

$$x_1=\frac{|\xi|^{p}}{\mathrm{M}|\xi|^{p}},\ x_2=\frac{|\eta|^{q}}{\mathrm{M}|\eta|^{q}},\ \theta=\frac{1}{p},\ 1-\theta=\frac{1}{q},$$

получим

$$\frac{1}{p}\frac{|\xi|^p}{\mathrm{M}|\xi|^p}+\frac{1}{q}\frac{|\eta|^q}{\mathrm{M}|\eta|^q}\geq\frac{|\xi|}{\left(\mathrm{M}|\xi|^p\right)^{1/p}}\frac{|\eta|}{\left(\mathrm{M}|\eta|^q\right)^{1/q}}.$$

Если взять математические ожидания от обеих частей последнего неравенства, то приходим к доказательству неравенства Гёльдера.

З а м е ч а н и е 2. Неравенство Гёльдера для $p=q=2$ называется н е р а в е н с т в о м К о ш и – Б у н я к о в с к о г о, или Ш в а р ц а и имеет вид

$$\mathrm{M}|\xi\eta|\leq\sqrt{\mathrm{M}\xi^2\,\mathrm{M}\eta^2}.$$

Н е р а в е н с т в о М и н к о в с к о г о. *Если*

$$\mathrm{M}|\xi|^p<\infty,\ \ \mathrm{M}|\eta|^p<\infty,\ \ p>1,$$

то

$$\mathrm{M}|\xi+\eta|^p<\infty$$

и

$$\left(\mathrm{M}|\xi+\eta|^p\right)^{1/p}\leq\left(\mathrm{M}|\xi|^p\right)^{1/p}+\left(\mathrm{M}|\eta|^p\right)^{1/p}.$$

Д о к а з а т е л ь с т в о. Так как для $p>1$

$$|\xi+\eta|^p\leq 2^{p-1}\left(|\xi|^p+|\eta|^p\right),$$

то получаем, что $\mathrm{M}|\xi+\eta|^p<\infty$. Используя неравенство Гёльдера, имеем

$$\mathrm{M}|\xi+\eta|^p=\mathrm{M}|\xi+\eta||\xi+\eta|^{p-1}\leq\mathrm{M}|\xi||\xi+\eta|^{p-1}+\mathrm{M}|\eta||\xi+\eta|^{p-1}\leq$$

$$\leq\left(\mathrm{M}|\xi|^p\right)^{1/p}\left(\mathrm{M}|\xi+\eta|^{(p-1)q}\right)^{1/q}+\left(\mathrm{M}|\eta|^p\right)^{1/p}\left(\mathrm{M}|\xi+\eta|^{(p-1)q}\right)^{1/q}=$$

$$=\left(\mathrm{M}|\xi+\eta|^p\right)^{(p-1)/p}\left(\left(\mathrm{M}|\xi|^p\right)^{1/p}+\left(\mathrm{M}|\eta|^p\right)^{1/p}\right).$$

Разделив обе части неравенства на $\left(\mathrm{M}|\xi+\eta|^p\right)^{(p-1)/p}$, получим доказательство.

Пусть на вероятностном пространстве $(\Omega,\ \mathcal{A},\ \mathrm{P})$ заданы случайные величины ξ и η с конечными дисперсиями, причем

$$\mathrm{D}\xi>0,\qquad \mathrm{D}\eta>0.$$

О п р е д е л е н и е 1. *Коэффициентом корреляции случайных величин* ξ *и* η *называется число*

$$\rho(\xi,\eta)=\frac{\operatorname{cov}(\xi,\ \eta)}{\sqrt{D\xi D\eta}}=\frac{M(\xi-M\xi)(\eta-M\eta)}{\sqrt{D\xi D\eta}}.$$

Свойства коэффициента корреляции

1. $|\rho(\xi,\ \eta)|\le 1.$

Доказательство вытекает из неравенства Коши – Буняковского

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

если положить по определению $\bar{\xi}=\xi-M\xi$ и $\bar{\eta}=\eta-M\eta$.

2. *Если случайные величины* ξ *и* η *независимы, то* $\rho(\xi,\ \eta)=0.$

Замечание 3. Обратное утверждение, вообще говоря, неверно. Приведите пример, показывающий это.

В случае, когда $\rho(\xi,\eta)=0$, случайные величины ξ и η называются *некоррелированными*.

Доказательство свойства 2 следует из того, что

$$\rho(\xi,\ \eta)=\frac{M\xi\eta-M\xi M\eta}{\sqrt{D\xi D\eta}}=\frac{M\xi M\eta-M\xi M\eta}{\sqrt{D\xi D\eta}}=0.$$

3. $|\rho(\xi,\ \eta)|=1$ *тогда и только тогда, когда существуют такие константы* $a\ne 0$ *и* b, *что* $P(\eta=a\xi+b)=1.$

Доказательство. Необходимость. а) Пусть

$$\rho(\xi,\eta)=-1.$$

Введем в рассмотрение нормированные случайные величины ξ_1 и η_1, положив по определению

$$\xi_1=\frac{\xi-M\xi}{\sqrt{D\xi}},\ \ \eta_1=\frac{\eta-M\eta}{\sqrt{D\eta}}.$$

Используя свойства математического ожидания и дисперсии, несложно видеть, что

$$M\xi_1=M\eta_1=0,\ \ D\xi_1=D\eta_1=1,\ \ \rho(\xi,\ \eta)=M\xi_1\eta_1.$$

Поэтому

$$D(\xi_1+\eta_1)=D\xi_1+D\eta_1+2\operatorname{cov}(\xi_1,\ \eta_1)=1+1+2\rho(\xi,\ \eta)=0.$$

Из последнего соотношения по свойству 5 дисперсии делаем вывод, что существует константа c такая, что

$$P(\xi_1+\eta_1=c)=1,$$

или

$$\mathrm{P}\left(\frac{\xi-\mathrm{M}\xi}{\sqrt{\mathrm{D}\xi}}+\frac{\eta-\mathrm{M}\eta}{\sqrt{\mathrm{D}\eta}}=c\right)=1.$$

Таким образом, существуют константы

$$a=-\frac{\sqrt{\mathrm{D}\eta}}{\sqrt{\mathrm{D}\xi}}\neq 0,\ \ b=\mathrm{M}\eta+\frac{\mathrm{M}\xi\sqrt{\mathrm{D}\eta}}{\sqrt{\mathrm{D}\xi}}+c\sqrt{\mathrm{D}\eta},$$

такие что $\mathrm{P}(\eta=a\xi+b)=1$.

б) Случай, когда $\rho(\xi,\ \eta)=1$, доказывается аналогично, если исследовать $\mathrm{D}(\xi_1-\eta_1)$. Необходимость доказана.

Достаточность. Пусть существуют $a\neq 0$ и b, такие что

$$\mathrm{P}(\eta=a\xi+b)=1.$$

Учитывая свойства математического ожидания и дисперсии, получаем

$$\rho(\xi,\ \eta)=\frac{\mathrm{M}(\xi-\mathrm{M}\xi)(\eta-\mathrm{M}\eta)}{\sqrt{\mathrm{D}\xi}\sqrt{\mathrm{D}\eta}}=\frac{\mathrm{M}(\xi-\mathrm{M}\xi)(a\xi+b-\mathrm{M}(a\xi+b))}{\sqrt{\mathrm{D}\xi}\sqrt{\mathrm{D}(a\xi+b)}}=$$

$$=\frac{\mathrm{M}(\xi-\mathrm{M}\xi)(a\xi-a\mathrm{M}\xi)}{\sqrt{\mathrm{D}\xi}\sqrt{\mathrm{D}(a\xi)}}=\frac{a\mathrm{D}\xi}{|a|\mathrm{D}\xi}=\frac{a}{|a|}. \tag{4}$$

Взяв модули от левой и правой частей соотношения (4), приходим к доказательству достаточности. Свойство доказано.

Замечание 4. Из свойства 3 следует, что коэффициент корреляции является мерой линейной зависимости случайных величин. Чем ближе $|\rho(\xi,\eta)|$ к единице, тем ближе к прямой будут размещаться в среднем значения случайного вектора (ξ,η).

Пример 1. Некоторые применения неравенства Чебышева (правило трех сигм).

Неравенство Чебышева дает оценку вероятностей «значительных» отклонений случайной величины от ее математического ожидания. Запишем это неравенство в следующей форме:

$$\mathrm{P}\left(|\xi-\mathrm{M}\xi|>k\sqrt{\mathrm{D}\xi}\right)\leq\frac{\mathrm{D}\xi}{k^2\,\mathrm{D}\xi}=\frac{1}{k^2},$$

где k – произвольное положительное число. Тогда

$$\mathrm{P}\left(|\xi-\mathrm{M}\xi|\leq k\sqrt{\mathrm{D}\xi}\right)=1-\mathrm{P}\left(|\xi-\mathrm{M}\xi|>k\sqrt{\mathrm{D}\xi}\right)\geq 1-\frac{1}{k^2}.$$

Например, если взять $k = 3,$ то с вероятностью большей, чем $1 - \frac{1}{9} = \frac{8}{9},$ можно утверждать, что значение случайной величины ξ принадлежит интервалу

$$\left(\mathrm{M}\,\xi - 3\sqrt{\mathrm{D}\,\xi};\ \mathrm{M}\,\xi + 3\sqrt{\mathrm{D}\,\xi}\right).$$

Предположим, что случайная величина ξ нормально распределена с параметрами a и σ^2. Тогда

$$\mathrm{P}\left(\left|\xi - a\right| \le 3\sigma\right) = \mathrm{P}\left(\left|\frac{\xi - a}{\sigma}\right| \le 3\right) = 2\Phi_0\left(3\right) \approx 0{,}997,$$

где $\Phi_0\left(x\right) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-t^2/2} dt,$ и ее значение находим по табл. 2 (см. Приложения).

ДОПОЛНЕНИЯ К ГЛАВЕ 4

§ 1. Интеграл Лебега

Случайная величина ξ называется *простой*, если она может принимать только конечное число значений.

О п р е д е л е н и е 1. *Интегралом Лебега от простой случайной величины* ξ, *принимающей значения* $x_1,\ \ldots,\ x_n$ *с вероятностями* $p_i = \mathrm{P}(\xi = x_i)$ *называется число*

$$\int\limits_{\Omega} \xi(\omega)\, d\,\mathrm{P}(\omega) = \sum_{i=1}^{n} x_i p_i .$$

О п р е д е л е н и е 2. *Индикатором* I_A *случайного события* A *называется случайная величина*

$$I_A = I_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

Индикаторы обладают следующими очевидными с в о й с т в а м и:

1. $I_{\Omega}(\omega) \equiv 1$;
2. *Если события* $A_1,\ \ldots,\ A_n$ *попарно несовместны, то*

$$I_{\sum\limits_{k=1}^{n} A_k}(\omega) = \sum_{k=1}^{n} I_{A_k}(\omega).$$

У п р а ж н е н и е 1. Докажите, что для того, чтобы случайная величина ξ была простой, необходимо и достаточно, чтобы она была представима в виде

$$\xi(\omega) = \sum_{k=1}^{n} x_k I_{A_k}(\omega), \tag{1}$$

где случайные события $A_1,\ \ldots,\ A_n$ попарно несовместны, а среди чисел $x_1,\ \ldots,\ x_n$ могут быть равные.

У п р а ж н е н и е 2. Докажите, что если случайную величину ξ можно записать в виде (1), то

$$\int\limits_{\Omega} \xi(\omega)\, d\mathrm{P}(\omega) = \sum_{i=1}^{n} x_k \mathrm{P}(A_k).$$

Пусть ξ – неотрицательная случайная величина. Обозначим

$$A_j^n = \left(\omega \in \Omega :\ \frac{j-1}{2^n} < \xi(\omega) \le \frac{j}{2^n}\right),\quad B_j^n = \left(\omega \in \Omega :\ \xi(\omega) > n\right),\qquad j = 1,\ 2,\ \ldots,$$

$$\xi_n(\omega) = \sum_{j=1}^{n\cdot 2^n} \frac{j-1}{2^n} I_{A_n^j}(\omega) + n I_{B_n^j}(\omega). \tag{2}$$

Заметим, что

1. *При любом* n *случайная величина* ξ_n *является простой:*

$$\xi_n:\ 0,\ \frac{1}{2^n},\ \ldots,\ \frac{n\cdot 2^n-1}{2^n},\ n.$$

2. $\xi_n(\omega)\le\xi_{n+1}(\omega)$ *при любом* $\omega\in\Omega$.

3. $\xi_n(\omega)\le\xi(\omega)$ *для всех* $n\in N$ *и* $\omega\in\Omega$.

4. *Для любого* $\omega\in\Omega$ *такого, что* $\xi(\omega)\le n$

$$\left|\xi(\omega)-\xi_n(\omega)\right|\le\frac{1}{2^n}.$$

Таким образом, для всех $\omega\in\Omega$ $\xi_n(\omega)\uparrow\xi(\omega)$ *при* $n\to\infty$.

Определение 3. *Интегралом Лебега от неотрицательной случайной величины* ξ *называется число*

$$\int\limits_{\Omega}\xi(\omega)\,d\mathrm{P}(\omega)=\lim_{n\to\infty}\xi_n(\omega)\,d\mathrm{P}(\omega),$$

где элементы последовательности $\{\xi_n\}$ *определяются соотношениями* (2).

Пусть ξ – произвольная случайная величина. Обозначим, как и раньше,

$$\xi^+(\omega)=\max\bigl(0,\ \xi(\omega)\bigr),\ \ \xi^-(\omega)=\max\bigl(0,\ -\xi(\omega)\bigr).$$

Тогда

$$\xi^+\ge 0,\ \ \xi^-\ge 0,,\ \xi=\xi^+-\xi^-.$$

Определение 4. *Интегралом Лебега от произвольной случайной величины* ξ *называется число*

$$\int\limits_{\Omega}\xi(\omega)\,d\mathrm{P}(\omega)=\int\limits_{\Omega}\xi^+(\omega)\,d\mathrm{P}(\omega)-\int\limits_{\Omega}\xi^-(\omega)\,d\mathrm{P}(\omega),$$

если оба выражения в правой части конечны.

Доказательства корректности этих определений и свойств интеграла Лебега можно найти, например, в [41].

§ 2. Интегралы Римана – Стилтьеса и Лебега – Стилтьеса

Пусть $\xi=\xi(\omega)$ – случайная величина, определенная на вероятностном пространстве $(\Omega,\ \mathcal{A},\ \mathrm{P})$. Рассмотрим другое вероятностное пространство – $\bigl(R,\ \mathcal{B}(R),\ \mathrm{P}_\xi\bigr)$, порожденное случайной величиной ξ. Случайные величины, определенные на этом пространстве, являются борелевскими функциями $g=g(x)$ действительной переменной $x\in R$. Интеграл Лебега от функции g по мере P_ξ обозначим

$$\int\limits_R g(x)\,\mathrm{P}_\xi(dx)=\int g\,d\mathrm{P}_\xi. \qquad (1)$$

Так как мера P_ξ однозначно определяется по функции распределения F_ξ, то интеграл (1) записывают также так:

$$\int\limits_R g(x)\,dF_\xi(x). \qquad (2)$$

В такой записи данный интеграл называют *интегралом Лебега – Стилтьеса от функции* g *по функции* F_ξ.

Отметим, что если мера P_ξ является мерой Лебега, то искомый интеграл записывается в виде

$$\int\limits_R g(x)\,dx \qquad (3)$$

и называется *интегралом Лебега.*

Пусть функция $F = F(x)$ – функция распределения.

О п р е д е л е н и е 1. *Интегралом Римана – Стилтьеса от функции* g *по функции* F*, если* $x \in [a;b]$, *называется число*

$$\int\limits_a^b g(x)\,dF(x) = \lim_{\Delta x \to 0} \sum_{i=0}^{N-1} g(\tilde{x}_i)\big(F(x_{i+1}) - F(x_i)\big),$$

если предел существует и не зависит от выбора точек x_i, $\tilde{x}_i$, *где* $a = x_0 < x_1 < \ldots$ $< x_N = b,\ \tilde{x}_i \in [x_i, x_{i+1}],\ \Delta x = \max(x_{i+1} - x_i),\ i = \overline{0,\ N-1}.$

О п р е д е л е н и е 2. *Под интегралом*

$$\int\limits_{-\infty}^{+\infty} g(x)\,dF(x)$$

понимается число, равное пределу

$$\lim_{\substack{a \to -\infty \\ b \to +\infty}} \int\limits_a^b g(x)\,dF(x).$$

З а м е ч а н и е 1. Интеграл

$$\int\limits_{-\infty}^{+\infty} g(x)\,dF(x)$$

существует, если функция g непрерывная или кусочно-непрерывная.

З а м е ч а н и е 2. Интеграл Римана – Стилтьеса обладает всеми свойствами интеграла Римана.

У т в е р ж д е н и е 1. *Пусть*

$$F(x) = a_1 F_1(x) + a_2 F_2(x).$$

Тогда

$$\int\limits_a^b g(x)\,dF(x) = a_1 \int\limits_a^b g(x)\,dF_1(x) + a_2 \int\limits_a^b g(x)\,dF_2(x).$$

Д о к а з а т е л ь с т в о вытекает из равенства

$$F(x_{i+1}) - F(x_i) = a_1\big(F_1(x_{i+1}) - F_1(x_i)\big) + a_2\big(F_2(x_{i+1}) - F_2(x_i)\big)$$

и определения интеграла.

У т в е р ж д е н и е 2. *Если функция* $F(x)$ *– непрерывно дифференцируемая функция и* $p(x) = F'(x)$, *то*

$$\int_a^b g(x)\,dF(x)=\int_a^b g(x)\,p(x)\,dx.$$

В правой части этого равенства – интеграл Римана.

Д о к а з а т е л ь с т в о. Так как по теореме о среднем

$$F(x_{i+1})-F(x_i)=\int_{x_i}^{x_{i+1}} p(x)\,dx=p\left(\tilde{\tilde{x}}_i\right)(x_{i+1}-x_i),$$

где $\tilde{\tilde{x}}_i\in(x_i;x_{i+1}]$, то из определения интеграла Римана – Стилтьеса имеем

$$\int_a^b g(x)\,dF(x)=\lim_{\Delta x\to 0}\sum_{i=0}^{N-1} g(\tilde{x}_i)\big(F(x_{i+1})-F(x_i)\big)=\lim_{\Delta x\to 0}\sum_{i=0}^{N-1} g(\tilde{x}_i)\,p\left(\tilde{\tilde{x}}_i\right)(x_{i+1}-x_i).$$

Взяв $\tilde{x}_i=\tilde{\tilde{x}}_i$, получим доказательство этого утверждения.

У т в е р ж д е н и е 3. *Если функция* $F(x)$ *является ступенчатой со скачками в точках* x_k, *причем величина скачка в точке* x_k *равна* p_k, *и*

$$\sum_{k=1}^{\infty}\left|g(x_k)\right|p_k<\infty,$$

то

$$\int_{-\infty}^{+\infty} g(x)\,dF(x)=\sum_{k=1}^{\infty} g(x_k)p_k.$$

Д о к а з а т е л ь с т в о. Пусть точки разбиения $a=t_0<t_1<\ldots<t_N=b$ и диаметр разбиения $\Delta t=\max(t_{i+1}-t_i)$ выбраны так, что при каждом $i=\overline{0,N-1}$ промежуток $(t_i;t_{i+1}]$ содержит не больше одной точки из множества $\{x_1,\,x_2,\,\ldots\}$. Тогда $F(t_{i+1})-F(t_i)=p_k$, если промежуток $(t_i;t_{i+1}]$ содержит точку x_k, и $F(t_{i+1})-$ $-F(t_i)=0$, если ни одна из точек $x_1,\,x_2,\,\ldots$ не попадает в промежуток $(t_i;t_{i+1}]$, $i=\overline{0,N-1}$. Взяв $\tilde{t}_i=x_k$, если $x_k\in(t_i;t_{i+1}]$ по определению интеграла Римана – Стилтьеса получаем

$$\int_a^b g(x)\,dF(x)=\lim_{\Delta t\to 0}\sum_{i=0}^{N-1} g(x_k)\big(F(t_{i+1})-F(t_i)\big)=\sum_{k:x_k\in[a;b]} g(x_k)\,p_k.$$

Отсюда

$$\int_{-\infty}^{+\infty} g(x)\,dF(x)=\lim_{\substack{a\to-\infty\\ b\to+\infty}}\int_a^b g(x)\,dF(x)=\sum_{k=1}^{\infty} g(x_k)p_k,$$

если $\sum\limits_{k=1}^{\infty}|g(x_k)|p_k<\infty.$

З а м е ч а н и е 3. Многомерные интегралы Римана – Стилтьеса и Лебега – Стилтьеса определяются аналогичным образом (см., напр., [13]).

§ 3. Условные математические ожидания

1. **Условное математическое ожидание относительно событий и дискретных случайных величин.** Ранее было установлено, что условная вероятность

$$\mathrm{P}_B = \mathrm{P}_B(A) = \mathrm{P}(A \mid B), \quad \mathrm{P}(B) > 0, \qquad A, B \in \mathcal{A}$$

обладает всеми свойствами вероятности, поэтому корректно следующее определение.

Определение 1. *Условным математическим ожиданием* $\mathrm{M}(\xi \mid B)$ *случайной величины* ξ *относительно события* B *называется число*

$$\mathrm{M}(\xi \mid B) = \int_\Omega \xi(\omega)\mathrm{P}_B(d\omega) = \int_\Omega \xi(\omega)\, d\,\mathrm{P}_B(\omega), \quad \mathrm{P}(B) > 0.$$

Замечание 1. Из определения 1 вытекает, что это условное математическое ожидание обладает всеми свойствами математического ожидания.

Утверждение 1. *Если* $\mathrm{M}\xi < \infty$, *то*

$$\mathrm{M}(\xi \mid B) = \frac{1}{\mathrm{P}(B)}\mathrm{M}\, I_B \xi..$$

Доказательство. Пусть $\xi = I_A$, тогда

$$\mathrm{M}(\xi \mid B) = \int_\Omega I_A(\omega)\mathrm{P}_B(d\omega) = \mathrm{P}_B(A) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{1}{\mathrm{P}(B)}\int_\Omega I_{A\cap B}(\omega)\mathrm{P}(d\omega) =$$

$$= \frac{1}{\mathrm{P}(B)}\int_B I_A(\omega)\mathrm{P}(d\omega) = \frac{1}{\mathrm{P}(B)}\mathrm{M}\,\xi I_B.$$

Для произвольных случайных величин ξ доказательство совпадает с доказательством формулы замены переменной в интеграле Лебега.

Утверждение 2. (Формула полного математического ожидания). *Если события* B_1, B_2, … *образуют полную группу событий и* $\mathrm{M}\xi < \infty$, *то*

$$\mathrm{M}\xi = \sum_{k=1}^{\infty} \mathrm{P}(B_k)\mathrm{M}(\xi \mid B_k).$$

Доказательство. Очевидно, что

$$1 \equiv I_\Omega(\omega) = I_{\sum_{k=1}^{\infty} B_k}(\omega) = \sum_{k=1}^{\infty} I_{B_k}(\omega), \qquad \mathrm{M}\sum_{k=1}^{\infty} \xi I_{B_k} = \mathrm{M}\xi < \infty.$$

Поэтому по теореме о мажорируемой сходимости получаем

$$\mathrm{M}\xi = \mathrm{M}\left(\sum_{k=1}^{\infty} \xi I_{B_k}\right) = \sum_{k=1}^{\infty} \mathrm{M}\,\xi I_{B_k} = \sum_{k=1}^{\infty} \mathrm{P}(B_k)\mathrm{M}(\xi \mid B_k).$$

Утверждение 2 доказано.

Утверждение 3. *Если событие* B *не зависит от событий, порожденных случайной величиной* ξ, *то*

$$\mathrm{M}(\xi \mid B) = \mathrm{M}\xi.$$

Д о к а з а т е л ь с т в о. Из условия утверждения 3 вытекает, что случайные величины ξ и I_B независимы, поэтому

$$\mathrm{M}\,\xi I_B = \mathrm{M}\,\xi\ \mathrm{M}\, I_B = \mathrm{P}(B)\ \mathrm{M}\,\xi.$$

Отсюда, учитывая утверждение 1, имеем

$$\mathrm{M}(\xi \mid B) = \frac{1}{\mathrm{P}(B)} \mathrm{M}\,\xi \cdot I_B = \mathrm{M}\,\xi.$$

Утверждение 3 доказано.

О п р е д е л е н и е 2. *Пусть* η *– дискретная случайная величина, которая принимает значения* $y_1,\ y_2,\ y_3,\ \ldots$ *с вероятностями* $\mathrm{P}(\eta = y_i) \neq 0,\ i = 1,\ 2,\ 3\ldots$.

Условным математическим ожиданием случайной величины ξ *относительно случайной величины* η *называется такая случайная величина* $\mathrm{M}(\xi \mid \eta)$, *которая для* $\omega \in (\eta = y_i)$ *равна*

$$\mathrm{M}(\xi \mid \eta)(\omega) = \mathrm{M}(\xi \mid \eta = y_i),\ \ i = 1,\ 2,\ 3,\ldots .$$

З а м е ч а н и е 2. Из данного определения вытекает, что $\mathrm{M}(\xi \mid \eta) = g(\eta)$, где g – некоторая борелевская функция, которая однозначно определяется в точках $y_1,\ y_2,\ y_3,\ldots$.

У т в е р ж д е н и е 4. *Если* ξ *и* η *независимые случайные величины, то*

$$\mathrm{M}(\xi \mid \eta) = \mathrm{M}\,\xi.$$

Д о к а з а т е л ь с т в о аналогично доказательству утверждения 3.

У т в е р ж д е н и е 5. *Если* $\mathrm{P}(\eta = y) \neq 0$ *и* $\mathrm{M}\,g(\xi,\ \eta) < \infty$, *то*

$$\mathrm{M}\big(g(\xi,\ \eta) \big| \eta = y\big) = \mathrm{M}\big(g(\xi,\ y) \big| \eta = y\big).$$

Д о к а з а т е л ь с т в о непосредственно следует из утверждения 1.

У т в е р ж д е н и е 6. *Пусть случайная величина* η *является дискретной и* $\mathrm{M}|\xi| < \infty$, *тогда для любой борелевской функции g*

$$\mathrm{M}\big(\xi g(\eta) \mid \eta\big) = g(\eta)\mathrm{M}(\xi \mid \eta).$$

Д о к а з а т е л ь с т в о. Пусть случайная величина η принимает значения $y_1,\ y_2,\ y_3,\ \ldots$ и $\mathrm{P}(\eta = y_i) \neq 0$, тогда для $\omega \in (\eta = y_i)$ по утверждению 5 имеем

$$\mathrm{M}\big(\xi g(\eta) \mid \eta\big)(\omega) = \mathrm{M}\big(\xi g(\eta) \mid \eta = y_i\big) = \mathrm{M}\big(\xi g(y_i) \mid \eta = y_i\big) =$$
$$= g(y_i)\mathrm{M}(\xi \mid \eta = y_i) = g\big(\eta(\omega)\big)\mathrm{M}(\xi \mid \eta)(\omega), \quad i = 1,\ 2,\ 3,\ \ldots .$$

Отсюда следует равенство для всех $\omega \in \Omega$.

У т в е р ж д е н и е 6. (Т о ж д е с т в о В а л ь д а). *Если случайные величины* $\xi_1,\ \xi_2,\ \ldots,\ \xi_n$ *– независимы, одинаково распределены и* $\mathrm{M}|\xi_i| < \infty,\ i = \overline{1, n}$, *а случайная величина* τ *независима с ними и принимающая значения* 1, 2, … *и* $\mathrm{M}\,\tau < \infty$, *то*

$$\mathrm{M}\,S_\tau = \mathrm{M}\,\xi_1 \tau,$$

где $S_\tau = \xi_1 + \ldots + \xi_\tau$.

Д о к а з а т е л ь с т в о. Из доказанных утверждений имеем

$$\mathrm{M}\, S_\tau = \sum_{k=1}^{\infty} \mathrm{P}(\tau = k)\mathrm{M}(S_\tau \mid \tau = k) = \sum_{k=1}^{\infty} \mathrm{P}(\tau = k)\mathrm{M}(S_k \mid \tau = k) = \sum_{k=1}^{\infty} \mathrm{P}(\tau = k)\mathrm{M}\, S_k =$$

$$= \mathrm{M}\,\xi_1 \sum_{k=1}^{\infty} k\mathrm{P}(\tau = k) = \mathrm{M}\,\xi_1\, \mathrm{M}\,\tau.$$

Утверждение 6 доказано.

У т в е р ж д е н и е 7. *Пусть* $\mathrm{M}|\xi| < \infty,$ η *– дискретная случайная величина. Тогда для любой ограниченной борелевской функции* h

$$\mathrm{M}\big(h(\eta)\mathrm{M}(\xi \mid \eta)\big) = \mathrm{M}\, h(\eta)\xi.$$

Д о к а з а т е л ь с т в о следует из формулы полного математического ожидания.

2. **Условное математическое ожидание относительно непрерывных случайных величин.** Пусть случайная величина η имеет непрерывную плотность распределения p_η, причем $p_\eta(y) \neq 0$ для любых $y \in R.$

О п р е д е л е н и е 3. *Пусть* $\mathrm{M}|\xi| < \infty.$ *Под условным математическим ожиданием* $\mathrm{M}(\xi \mid \eta = y)$ *будем понимать число*

$$\mathrm{M}(\xi \mid \eta = y) = \lim_{\Delta y \to +0} \mathrm{M}(\xi \mid y \le \eta \le y + \Delta y).$$

Теорема 1. *Если случайные величины* ξ *и* η *имеют непрерывную совместную плотность распределения* $p_{(\xi,\eta)}(x, y),$ $\mathrm{M}\,\xi < \infty,$ $p_\eta(y) \neq 0,$ $y \in R,$ *то*

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

Д о к а з а т е л ь с т в о. Из условия теоремы и утверждения 1 имеем

$$\lim_{\Delta y \to +0} \mathrm{M}(\xi \mid y \le \eta \le y + \Delta y) = \lim_{\Delta y \to +0} \frac{1}{\mathrm{P}(\xi \mid y \le \eta \le y + \Delta y)} \mathrm{M}\,\xi\, I_{(y \le \eta \le y + \Delta y)} =$$

$$= \lim_{\Delta y \to +0} \frac{1}{\int_y^{y+\Delta y} p_\eta(z)\,dz} \int_{-\infty}^{\infty} x \int_y^{y+\Delta y} p_{(\xi,\eta)}(x, z)\,dz dx = \int_{-\infty}^{\infty} x \frac{p_{(\xi,\eta)}(x, y)}{p_\eta(y)} dx.$$

Теорема 1 доказана.

О п р е д е л е н и е 4. *Условной плотностью распределения случайной величины* ξ *при условии, что случайная величина* η *приняла значение* y, *называется функция*

$$p_{\xi|\eta}(x \mid y) = \frac{p_{(\xi,\eta)}(x, y)}{p_\eta(y)}, \qquad x, y \in R.$$

З а м е ч а н и е 3. Если $p_\eta(y) = 0,$ $y \in R,$ то по определению полагаем, что $p_{\xi|\eta}(x \mid y) = 0.$

З а м е ч а н и е 4. Воспользовавшись доказательством теоремы 1, для любого $B \in \mathscr{B}(R)$ получаем:

$$\mathrm{P}(\xi \in B \mid \eta = y) = \lim_{\Delta y \to 0} \mathrm{P}(\xi \in B \mid y \leq \eta \leq y + \Delta y) = \int_B p_{\xi|\eta}(x \mid y)\, dx.$$

О п р е д е л е н и е 5. *Условным математическим ожиданием случайной величины* ξ *относительно случайной величины* η, *имеющих непрерывную совместную плотность распределения, называется случайная величина* $\mathrm{M}(\xi \mid \eta)(\omega)$, *которая для* $\omega \in (\eta = y)$ *принимает значения* $\mathrm{M}(\xi \mid \eta = y)$.

З а м е ч а н и е 5. Из определения 5 следует, что $\mathrm{M}(\xi \mid \eta)$ является функцией от η, т. е. $\mathrm{M}(\xi \mid \eta)(\omega) = g(\eta(\omega))$, $\omega \in \Omega$. Из теоремы 1 получаем, что

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

Теорема 2. *Если случайные величины* ξ *и* η *имеют непрерывную совместную плотность распределения и* $\mathrm{M}\xi < \infty$, *то для любой непрерывной ограниченной функции* $h: R \to R$

$$\mathrm{M}(\mathrm{M}(\xi \mid \eta) h(\eta)) = \mathrm{M} h(\eta)\xi. \tag{1}$$

Д о к а з а т е л ь с т в о. Из формулы замены переменных в интеграле Лебега имеем

$$\mathrm{M} h(\eta)\xi = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(y)\, x\, p_{(\xi,\eta)}(x, y)\, dx\, dy = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} x\, p_{\xi|\eta}(x|y)\, dx\right) h(y)\, p_\eta(y)\, dy =$$

$$= \int_{-\infty}^{\infty} \mathrm{M}(\xi \mid \eta = y) h(y)\, p_\eta(y)\, dy = \mathrm{M}(\mathrm{M}(\xi \mid \eta) h(\eta)).$$

Теорема 2 доказана.

3. **Условное математическое ожидание относительно произвольных случайных величин.** Пусть $(\Omega, \mathscr{A}, \mathrm{P})$ – произвольное вероятностное пространство, $\mathscr{F}$ – σ-подалгебра $\mathscr{A}$, ξ – случайная величина, заданная на $(\Omega, \mathscr{A}, \mathrm{P})$ и имеющая конечное математическое ожидание.

О п р е д е л е н и е 6. *Условным математическим ожиданием случайной величины* ξ *относительно σ-алгебры* $\mathscr{F}$, *называется* $\mathscr{F}$*-измеримая случайная величина* $\mathrm{M}(\xi \mid \mathscr{F})$ *такая, что для любых* $A \in \mathscr{F}$

$$\int_A \xi(\omega) \mathrm{P}(d\omega) = \int_A \mathrm{M}(\xi \mid \mathscr{F})(\omega) \mathrm{P}(d\omega).$$

З а м е ч а н и е 6. Корректность данного определения вытекает непосредственно из теоремы Радона – Никодима.

Зададим на вероятностном пространстве еще одну случайную величину – $\eta(\omega)$, $\omega \in \Omega$. Через $\mathscr{A}_\eta$ обозначим σ-алгебру, порожденную случайной величиной η. Напомним, что по определению $\mathscr{A}_\eta = \eta^{-1}(\mathscr{B}(R))$.

О п р е д е л е н и е 6. *Условное математическое ожидание* $\mathrm{M}\left(\xi \mid \mathcal{A}_\eta\right)$ *называется условным математическим ожиданием случайной величины* ξ *относительно случайной величины* η *и обозначается* $\mathrm{M}(\xi \mid \eta)$.

Приведем без доказательств основные свойства условного математического ожидания. Доказательства можно найти, например, в [5].

Свойства условного математического ожидания

1. *Если* $\mathrm{P}(\xi = C) = 1$, *то* $\mathrm{M}(\xi \mid \eta) = C$, $C \in R$.
2. *Если* $\mathrm{M}\xi_i < \infty$; $i = 1,\ 2$; $a,\ b \in R$, *то*
$$\mathrm{M}(a\,\xi_1 + b\,\xi_2 \mid \eta) = a\,\mathrm{M}(\xi_1 \mid \eta) + b\,\mathrm{M}(\xi_2 \mid \eta).$$
3. *Если* $\xi_1 \le \xi_2$ *и* $\mathrm{M}\xi_2 < \infty$, *то*
$$\mathrm{M}(\xi_1 \mid \eta) \le \mathrm{M}(\xi_2 \mid \eta).$$
4. *Если* $\xi_n \ge 0$, $n = 1,\ 2,\ \dots$, *и* $\xi_n \uparrow \xi$ *при* $n \to \infty$, *причем* $\mathrm{M}\xi < \infty$, *то*
$$\mathrm{M}(\xi_n \mid \eta) \underset{n\to\infty}{\uparrow} \mathrm{M}(\xi \mid \eta).$$
5. *Если случайные величины* ξ *и* η *независимы и* $\mathrm{M}\xi < \infty$, *то*
$$\mathrm{M}(\xi \mid \eta) = \mathrm{M}\xi.$$
6. *Если* $\mathrm{M}\xi < \infty$, *то*
$$\mathrm{M}\big(\mathrm{M}(\xi \mid \eta)\big) = \mathrm{M}\xi.$$
7. *Для любой ограниченной борелевской функции* $g:\ R \to R$
$$\mathrm{M}\big(\xi\, g(\eta) \mid \eta\big) = g(\eta)\,\mathrm{M}(\xi \mid \eta),$$
если $\mathrm{M}\xi < \infty$.

Задачи

1. Пусть ξ – неотрицательная случайная величина с функцией распределения $F(x)$. Покажите, что
$$\mathrm{M}\xi = \int_0^\infty \big(1 - F(x)\big)\,dx$$
и для любой неотрицательной константы c
$$\mathrm{M}\min(\xi,\ c) = \int_0^c \big(1 - F(x)\big)\,dx.$$

2. Найдите $\mathrm{M}\xi_i$, $\mathrm{D}\xi_i$ и $\operatorname{cov}\left(\xi_i,\ \xi_j\right)$ в полиномиальном распределении
$$\mathrm{P}(\xi_1 = m_1,\ \dots,\ \xi_k = m_k) = \frac{n!}{m_1!\ m_2!\ \dots\ m_k!}\, p_1^{m_1} \dots p_k^{m_k},$$
где m_i, $i = \overline{1,k}$, – целые неотрицательные числа, $m_1 + \ldots + m_k = n$, $p_1 + \ldots + p_k = 1$.

3. Случайные величины ξ, η – координаты точки, равномерно распределенной в круге $x^2 + y^2 \le R^2$. Найдите коэффициент корреляции ξ и η.

4. Пусть случайная величина ξ принимает конечное число неотрицательных значений $x_1, \ldots, x_k$. Докажите, что

$$\lim_{n\to\infty} \frac{\mathrm{M}\xi^{n+1}}{\mathrm{M}\xi^n} = \max\left(x_1, \ldots, x_k\right), \ \lim_{n\to\infty} \sqrt[n]{\mathrm{M}\xi^n} = \max\left(x_1, \ldots, x_k\right).$$

5. Пусть случайная величина ξ принимает целые неотрицательные значения и $\mathrm{M}\xi < \infty$. Покажите, что

$$\mathrm{M}\xi = \sum_{k=1}^{\infty} \mathrm{P}\left(\xi \ge k\right).$$

6. Покажите, что простые случайные величины ξ и η некоррелированы тогда и только тогда, когда они независимы.

7. Имеется n писем, адреса на конвертах которых написаны в случайном порядке. Пусть ξ – число писем, которые будут получены теми, кому они предназначены. Покажите, что $\mathrm{M}\xi = 1$.

8. Пусть ξ и η – независимые случайные величины, принимающие целые неотрицательные значения и $\mathrm{M}\xi < \infty$. Докажите, что

$$\mathrm{M}\min\left(\xi, \eta\right) = \sum_{k=1}^{\infty} \mathrm{P}\left(\xi \ge k\right)\mathrm{P}\left(\eta \ge k\right).$$

9. Пусть $\mathrm{M}\xi^2 < \infty$. Покажите, что для любой $c \in R$

$$\mathrm{M}\left(\xi - c\right)^2 \ge \mathrm{D}\xi.$$

10. Пусть ξ и η – случайные величины с конечными моментами второго порядка. Докажите, что для любых a, $b \in R$

$$\mathrm{M}\left(\eta - a\xi - b\right)^2 \ge \mathrm{M}\left(\eta - a_0\xi - b_0\right)^2 = \left(1 - \rho^2\left(\xi, \eta\right)\right)\mathrm{D}\eta,$$

где

$$a_0 = \frac{\operatorname{cov}\left(\xi, \eta\right)}{\mathrm{D}\xi}, \ b_0 = \mathrm{M}\eta - a_0\mathrm{M}\xi;$$

если $\mathrm{D}\xi = 0$, то $a_0 = 0$.

11. Пусть случайные величины ξ_1 и ξ_2 независимы, одинаково распределены и имеют конечные вторые моменты. Покажите, что случайные величины $\eta_1 = \xi_1 + \xi_2$ и $\eta_2 = \xi_1 - \xi_2$ некоррелированы.

12. Случайная величина ξ имеет биномиальное распределение с $p = \frac{1}{2}$. Найдите $\mathrm{M}\eta$, где $\eta = \sin\frac{\pi\xi}{2}$.

13. В N телефонах-автоматах ведутся разговоры. Длительность разговора, измеряемая в секундах, имеет геометрическое распределение с математическим ожиданием μ. Найдите среднее время ожидания до первого освобождения телефона-автомата.

14. Случайные величины $\xi_1, \ldots, \xi_{n+1}$ – независимые случайные величины Бернулли. Пусть $\eta_i \equiv \xi_i + \xi_{i+1} \pmod 2$, $\eta = \sum_{i=1}^{n} \eta_i$.
Найдите $\mathrm{M}\eta$ и $\mathrm{D}\eta$.

15. Пусть ξ – невырожденная нетривиальная случайная величина, $\mathrm{D}\xi < \infty$. Покажите, что

1) $\mathrm{P}\left(-3,2 < \dfrac{\xi - \mathrm{M}\xi}{\sqrt{\mathrm{D}\xi}} < 3,2 \right) > 0,9;$

2) $\mathrm{P}\left(-2 < \dfrac{\xi - a}{b} < 2 \right) > 0,999,$

где $a = \mathrm{M}\xi, b = \left(\mathrm{M}\left(\xi - a\right)^{10} \right)^{\frac{1}{10}} < \infty.$

16. Пусть случайные величины ξ_1, ξ_2 имеют совместное нормальное распределение. Покажите, что они независимы тогда и только тогда, когда некоррелированы.

17. Случайные величины ξ и η независимы и нормально распределены с одними и теми же параметрами a и σ^2. Докажите, что

$$\mathrm{M}\max\left(\xi, \eta\right) = a + \frac{\sigma}{\sqrt{\pi}}.$$

18. Пусть случайные величины $\xi_1, \ldots, \xi_n$ независимы и одинаково распределены, причем $\mathrm{M}\left(\xi_i - \mathrm{M}\xi_i\right)^3 = 0$. Докажите, что случайные величины $\overline{\xi} = \frac{1}{n}\sum_{i=1}^{n} \xi_i$ и $\frac{1}{n}\sum_{i=1}^{n} \left(\xi_i - \overline{\xi}\right)^2$ независимы.

19. Случайные величины $\xi_1, \ldots, \xi_n, \ldots$ независимы и равномерно распределены в $(0;1)$. Пусть ν – случайная величина, равная такому k, при котором впервые сумма $\xi_1 + \ldots + \xi_k$ превосходит 1. Покажите, что $\mathrm{M}\nu = e$.

20. Пусть ξ – стандартная случайная величина Коши. Найдите $\mathrm{M}\min\left(|\xi|, 1\right)$.

Глава 5. ХАРАКТЕРИСТИЧЕСКИЕ ФУНКЦИИ

В гл. 3 было установлено, что случайная величина полностью определяется своей функцией распределения (либо плотностью распределения, если она абсолютно непрерывна, либо законом распределения – в дискретном случае). В этой главе введено и изучено еще одно понятие – характеристическая функция случайной величины. В частности, показано, что множества функций распределения и характеристических функций гомеоморфны, т. е. между ними существует взаимно однозначное и взаимно непрерывное соответствие. Отметим так же, что характеристическая функция есть преобразование Фурье – Стилтьеса функции распределения (преобразование Фурье плотности распределения в абсолютно непрерывном случае). Роль этих преобразований в неслучайном анализе нельзя переоценить. Особая эффективность аппарата характеристических функций в теории вероятностей будет показана в гл. 6 при исследовании предельного поведения сумм независимых случайных величин.

§ 1. Определение и простейшие свойства

Для определения характеристических функций необходимо распространить понятие математического ожидания на комплексные случайные величины. Пусть ξ и η случайные величины, определенные на вероятностном пространстве $(\Omega, \mathscr{A}, \mathrm{P})$, с конечными математическими ожиданиями, $\varsigma = \xi + i\eta$, где i – мнимая единица. Тогда по определению математическое ожидание комплексной случайной величины ς есть

$$\mathrm{M}\varsigma = \mathrm{M}\xi + i\mathrm{M}\eta.$$

Несложно видеть, что определяемое таким образом математическое ожидание обладает всеми свойствами математического ожидания действительных случайных величин, которые не связаны с упорядоченностью.

О п р е д е л е н и е 1. *Характеристической функцией действительнозначной случайной величины* ξ *называется функция* $f_\xi(t)$ *действительного переменного* t

$$f_\xi(t) = \mathrm{M}e^{it\xi}.$$

З а м е ч а н и е 1. Характеристическая функция существует для любой действительнозначной случайной величины ξ, так как для любых $t \in R$

$$\left|e^{it\xi}\right| = \left|\cos t\xi + i\sin t\xi\right| = 1$$

и

$$\left|f_\xi(t)\right| = \left|\mathrm{M}\, e^{it\xi}\right| \le \mathrm{M}\left|e^{it\xi}\right| = \mathrm{M}1 = 1.$$

З а м е ч а н и е 2. Из формулы замены переменных в интеграле Лебега имеем

$$f_\xi(t) = \int_{-\infty}^{+\infty} e^{itx} dF_\xi(x), \tag{1}$$

где F_ξ – функция распределения случайной величины ξ. Отсюда видно, что характеристическая функция есть преобразование Фурье – Стилтьеса функции распределения.

Пусть ξ – абсолютно непрерывная случайная величина с плотностью $p_\xi(x)$. В этом случае

$$f_\xi(t) = \int_{-\infty}^{+\infty} e^{itx} p_\xi(x)\, dx, \tag{2}$$

т. е. характеристическая функция – это преобразование Фурье плотности.

Наконец, если ξ дискретная случайная величина с законом распределения

$$\begin{array}{ll} \xi: & x_1,\ \ldots,\ x_n,\ \ldots \\ \mathrm{P}: & p_1,\ \ldots,\ p_n,\ \ldots, \end{array}$$

то

$$f_\xi(t) = \sum_{n=1}^{\infty} e^{itx_n} p_n. \tag{3}$$

Свойства характеристических функций

1°. $f_\xi(0) = 1$, $\left|f_\xi(t)\right| \le 1$ *для любого действительного* t.

Д о к а з а т е л ь с т в о. Из определения характеристической функции и свойств математического ожидания следует:

$$f_\xi(0) = \mathrm{M}e^0 = \mathrm{M}1 = 1, \qquad \left|f_\xi(t)\right| = \left|\mathrm{M}\, e^{it\xi}\right| \le \mathrm{M}\left|e^{it\xi}\right| = \mathrm{M}1 = 1.$$

2°. *Характеристическая функция* $f_\xi(t)$ *равномерно непрерывна на* R.

Для доказательства свойства 2 воспользуемся леммой.

Л е м м а 1. *Для любых* $\varphi \in R$ *и любых* $n \in N$ *справедливо неравенство*

$$\left| e^{i\varphi} - \sum_{k=0}^{n-1} \frac{(i\varphi)^k}{k!} \right| \le \frac{|\varphi|^n}{n!}. \qquad (4)$$

Доказательство леммы. Используем метод математической индукции.

Пусть $n=1$, докажем, что $\left| e^{i\varphi} - 1 \right| \le |\varphi|$. Несложно видеть, что

$$\int_0^{\varphi} e^{iu} du = \frac{1}{i}\left(e^{i\varphi} - 1\right),$$

а поэтому, взяв модули от обеих частей, получим

$$\left| e^{i\varphi} - 1 \right| = \left| \int_0^{\varphi} e^{iu} du \right| \le \int_0^{|\varphi|} \left| e^{iu} \right| du = |\varphi|.$$

Предположим, что неравенство (4) справедливо и докажем, что

$$\left| e^{i\varphi} - \sum_{k=0}^{n} \frac{(i\varphi)^k}{k!} \right| \le \frac{|\varphi|^{n+1}}{(n+1)!}. \qquad (5)$$

Очевидно, что

$$\int_0^{\varphi} \left(e^{iu} - \sum_{k=0}^{n-1} \frac{(iu)^k}{k!} \right) du = \frac{1}{i} \left(e^{i\varphi} - \sum_{k=0}^{n} \frac{(i\varphi)^k}{k!} \right).$$

Поэтому, учитывая справедливость неравенства (4), имеем

$$\left| e^{i\varphi} - \sum_{k=0}^{n} \frac{(i\varphi)^k}{k!} \right| \le \int_0^{|\varphi|} \left| e^{iu} - \sum_{k=0}^{n-1} \frac{(iu)^k}{k!} \right| du \le$$

$$\le \int_0^{|\varphi|} \frac{u^n}{n!} du = \frac{|\varphi|^{n+1}}{(n+1)!},$$

таким образом, лемма доказана.

Доказательство свойства 2°. Покажем, что для любого $\varepsilon > 0$ существует $\delta > 0$, , что для всех $|h| < \delta$ и $t \in R$

$$\left| f_\xi(t+h) - f_\xi(t) \right| < \varepsilon.$$

Пусть $A = \{ |\xi| \le c \}$ и

$$I_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

Тогда

$$\left| f_\xi(t+h) - f_\xi(t) \right| = \left| \mathrm{M} e^{i(t+h)\xi} - \mathrm{M} e^{it\xi} \right| = \left| \mathrm{M} e^{it\xi} \left(e^{ih\xi} - 1 \right) \right| \le \mathrm{M} \left| e^{ih\xi} - 1 \right| =$$

$$= \mathrm{M}\left\{\left|e^{ih\xi}-1\right|\cdot I_A\right\}+\mathrm{M}\left\{\left|e^{ih\xi}-1\right|\cdot I_{\overline{A}}\right\}=I_1+I_2.$$

Так как согласно лемме 1 $|e^{ih\xi}-1|\le|h||\xi|$, то $I_1\le c|h|$. В свою очередь

$$I_2=\mathrm{M}\left|e^{ih\xi}-1\right|I_{\overline{A}}\le 2\,\mathrm{M}\,I_{\overline{A}}=2\mathrm{P}\left(\overline{A}\right),$$

таким образом

$$\left|f_\xi(t+h)-f_\xi(t)\right|\le c|h|+2\mathrm{P}\left(|\xi|>c\right).$$

Убедимся, что число $\mathrm{P}\left(|\xi|>c\right)$ за счет выбора c может быть сколь угодно малым. Имеем

$$\mathrm{P}\left(|\xi|>c\right)=\mathrm{P}(\xi>c)+\mathrm{P}(\varepsilon<-c)=1-F_\xi(c)+F_\xi(-c-0).$$

Из этого равенства следует, что за счет выбора c можно сделать

$$\mathrm{P}\left(|\xi|>c\right)<\frac{\varepsilon}{2}.$$

Теперь в качестве δ возьмем $\delta=\varepsilon/2c$. Свойство 2° доказано.

3°. *Если случайные величины* ξ *и* η *связаны следующим образом*

$$\mathrm{P}(\eta=a\xi+b)=1,\qquad a,\ b\in R,$$

то

$$f_\eta(t)=e^{ibt}f_\xi(at).$$

Д о к а з а т е л ь с т в о. По определению характеристической функции и свойствам математического ожидания имеем

$$f_\eta(t)=\mathrm{M}\,e^{it\eta}=\mathrm{M}\,e^{it(a\xi+b)}=\mathrm{M}\,e^{ita\xi}e^{ibt}=e^{ibt}f_\xi(at).$$

Свойство 3° доказано.

4°. *Если случайные величины* $\xi_1,\ \ldots,\ \xi_n$ *независимы, то для любых* $t\in R$

$$f_{\xi_1+\ldots+\xi_n}(t)=\prod_{k=1}^{n}f_{\xi_k}(t).$$

Д о к а з а т е л ь с т в о. Известно, что если случайные величины $\xi_1,\ \xi_2,\ \ldots,\ \xi_n$ – независимы, а $g_1,\ g_2,\ \ldots,\ g_n$ – борелевские функции, то $g_1(\xi_1),\ g_2(\xi_2),\ \ldots,\ g_n(\xi_n)$ также независимы. В нашем случае $g_1(x)=g_2(x)=\ldots=g_n(x)=e^{itx},\ \ x\in R$. Поэтому по свойству мультипликативности математического ожидания имеем

$$f_{\xi_1+\ldots+\xi_n}(t)=\mathrm{M}\,e^{it(\xi_1+\ldots+\xi_n)}=\mathrm{M}\prod_{k=1}^{n}e^{it\xi_k}=\prod_{k=1}^{n}\mathrm{M}\,e^{it\xi_k}=\prod_{k=1}^{n}f_{\xi_k}(t).$$

Свойство 4° доказано.

5°. $\overline{f_\xi(t)} = f_\xi(-t)$, *где* $\overline{f_\xi(t)}$ *комплексное сопряжение к* $f_\xi(t)$.

Д о к а з а т е л ь с т в о. По свойствам операции комплексного сопряжения получаем

$$\overline{f_\xi(t)} = \overline{\mathrm{M}\, e^{it\xi}} = \mathrm{M}\, \overline{e^{it\xi}} = \mathrm{M}\, e^{-it\xi} = f_\xi(-t).$$

Свойство 5° доказано.

6°. *Пусть момент n-го порядка* $m_n = \mathrm{M}\xi^n$ *конечен, тогда для всех* $k \le n$ *существует производная k-го порядка характеристической функции* $f_\xi^{(k)}(t)$ *и*

$$f_\xi^{(k)}(0) = i^k m_k. \tag{6}$$

Кроме того, имеет место следующее разложение по формуле Тейлора характеристической функции с остаточным членом в форме Пеано

$$f_\xi(t) = \sum_{k=0}^{n} \frac{(it)^k}{k!} m_k + R_n(t), \tag{7}$$

где $R_n(t) = o(t^n)$, $t \to 0$.

З а д а ч а. Является ли $f(t) = e^{-t^4}$ характеристической функцией какой-либо случайной величины?

Р е ш е н и е. Несложно убедиться в том, что

$$f^{(1)}(0) = 0, \quad f^{(2)}(0) = 0.$$

Предположим, что существует такая характеристическая функция f_ξ, что $f_\xi(t) = f(t)$. Из того, что $f^{(1)}(0) = 0$, $f^{(2)}(0) = 0$ и соотношения (6) следует

$$\mathrm{M}\xi = 0, \quad \mathrm{M}\xi^2 = 0,$$

а поэтому $\mathrm{D}\xi = 0$. По свойству дисперсии в итоге имеем $\mathrm{P}(\xi=0) = 1$ и $f_\xi(t) \equiv 1$. Получили противоречие.

Данное решение – это частичное решение для случайных величин, имеющих конечный момент второго порядка, но оно будет полным, если учесть, что справедлив следующий факт [5]: *если существует производная четного порядка характеристической функции*

$$f_\xi^{(2n)}(t), \quad n \in N,$$

то конечен момент m_{2n}.

Д о к а з а т е л ь с т в о с в о й с т в а 6°. Имеем

$$f_\xi(t) = \mathrm{M}\, e^{it\xi}.$$

Формально продифференцируем обе части этого равенства k раз по t. Тогда

$$f_\xi^{(k)}(t) = i^k \,\mathrm{M}\,\xi^k e^{it\xi}, \qquad (8)$$

и если подставить $t = 0$ в соотношение (8), то получим равенство (6).

Докажем правомочность дифференцирования, используя метод математической индукции.

Предположим, что существует $f_\xi^{(k)}$ и выполняется равенство (8) для $k < n$. Нужно доказать, что существует

$$f_\xi^{(k+1)}(t) \quad \text{и} \quad f_\xi^{(k+1)}(t) = i^{k+1}\,\mathrm{M}\,\xi^{k+1} e^{it\xi}, \qquad t \in R.$$

По определению производной $(k+1)$-го порядка имеем

$$f_\xi^{(k+1)}(t) = \lim_{h\to 0} \frac{f_\xi^{(k)}(t+h) - f_\xi^{(k)}(t)}{h} \overset{(5)}{=} \lim_{h\to 0} i^k\,\mathrm{M}\,\xi^k e^{it\xi}\frac{e^{ih\xi-1}}{h} = \lim_{h\to 0} \mathrm{M}\,\eta_h =$$

$$= i^{k+1}\,\mathrm{M}\,\xi^{k+1} e^{it\xi}.$$

Последнее равенство получено по теореме о мажорируемой сходимости, так как

$$\eta_h \xrightarrow[h\to 0]{} i^{k+1}\xi^{k+1}e^{it\xi}, \qquad |\eta_h| \le \left|\xi^{k+1}\right|,$$

а случайная величина $|\xi|^{k+1}$ интегрируема по условию для $k < n$.

Для доказательства равенства (7) введем, как и раньше, событие $A = \{|\,\xi\,| \le c\}$ и воспользуемся леммой 1. В результате получим

$$\left| f_\xi(t) - \sum_{k=0}^{n} \frac{(it)^k}{k!} m_k \right| \le \mathrm{M}\left| e^{it\xi} - \sum_{k=0}^{n} \frac{(it\xi)^k}{k!} \right| = \mathrm{M}\left| e^{it\xi} - \sum_{k=0}^{n} \frac{(it\xi)^k}{k!} \right| I_A +$$

$$+\mathrm{M}\left| \left(e^{it\xi} - \sum_{k=0}^{n-1} \frac{(it\xi)^k}{k!} \right) - \frac{(it\xi)^n}{n!} \right| I_{\bar A} \le \mathrm{M}\frac{|t\xi|^{n+1}}{(n+1)!} I_A +$$

$$+\mathrm{M}\left| e^{it\xi} - \sum_{k=0}^{n-1} \frac{(it\xi)^k}{k!} \right| I_{\bar A} + \mathrm{M}\frac{|t\xi|^n}{n!} I_{\bar A} \le c\frac{|t|^{n+1}}{(n+1)!}\mathrm{M}|\xi|^n I_A + 2\frac{|t|^n}{n!}\mathrm{M}|\xi|^n I_{\bar A}.$$

В силу конечности $m_n = \mathrm{M}\,\xi^n$ первое слагаемое в последнем выражении есть $o\left(t^n\right)$ при $t \to 0$, а так как

$$\mathrm{M}|\xi|^n I_{\bar A} = \mathrm{M}|\xi|^n - \mathrm{M}|\xi|^n I_A \to 0$$

при $c \to 0$, то за счет выбора c второе слагаемое есть $o(t^n)$ при $t \to 0$. Таким образом, свойство 6° доказано.

7°. *Характеристическая функция* $f_\xi(t)$ *является неотрицательно определенной функцией, т. е. для любого* $n \in N$, *любых действительных чисел* $t_1, t_2, \ldots, t_n$ *и любых комплексных чисел* $z_1, z_2, \ldots, z_n$

$$\sum_{k=1}^{n}\sum_{j=1}^{n} f_\xi\left(t_k - t_j\right) z_k \overline{z}_j \ge 0.$$

Доказательство. Из определения характеристической функции и свойств операции комплексного сопряжения следует:

$$\sum_{k=1}^{n}\sum_{j=1}^{n} f_\xi\left(t_k - t_j\right) z_k \overline{z}_j = \sum_{k=1}^{n}\sum_{j=1}^{n} \mathrm{M} e^{i\left(t_k - t_j\right)\xi} z_k \overline{z}_j = \mathrm{M}\sum_{k=1}^{n} e^{it_k\xi} z_k \sum_{j=1}^{n} \overline{e^{it_j\xi} z_j} =$$

$$= \mathrm{M}\left|\sum_{k=1}^{n} e^{it_k\xi} z_k\right|^2 \ge 0,$$

что и требовалось доказать.

Теорема Бохнера [11]. *Если функция* $f(t)$ *– непрерывная, неотрицательно определенная и* $f(0) = 1$, *то она является характеристической функцией некоторой случайной величины.*

Примеры характеристических функций

Пример 1. Характеристическая функция случайной величины Бернулли. Так как закон распределения ξ имеет вид

$$\begin{aligned} &\xi: \quad 0 \qquad 1, \\ &\mathrm{P}: \quad 1-p \quad p, \qquad 0 < p < 1, \end{aligned}$$

то, используя равенство (3), имеем

$$f_\xi(t) = e^{it0}(1-p) + e^{it1}p = (1-p) + pe^{it}.$$

Пример 2. Характеристическая функция биномиальной случайной величины. Имеем

$$\mathrm{P}(\xi = k) = C_n^k p^k (1-p)^{n-k}, \quad k = \overline{0, n},$$

поэтому по определению характеристической функции дискретной случайной величины

$$f_\xi(t) = \sum_{k=0}^{n} e^{itk}\mathrm{P}(\xi = \mathrm{k}) = \sum_{k=0}^{n} C_n^k (e^{it}p)^k (1-p)^{n-k} = (e^{it}p + 1 - p)^n.$$

В последнем равенстве использована формула бинома Ньютона.

Пример 3. Характеристическая функция случайной величины Пуассона. Так как

$$\mathrm{P}(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda}, \qquad \lambda > 0, \quad k = 0,\ 1,\ \ldots,$$

то

$$f_\xi(t) = \sum_{k=0}^{\infty} e^{itk}\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=0}^{\infty}\left(e^{it}\lambda\right)^k\frac{e^{-\lambda}}{k!} = \exp\left\{e^{it}\lambda - \lambda\right\} = \exp\left\{\lambda\left(e^{it} - 1\right)\right\}.$$

П р и м е р 4. Х а р а к т е р и с т и ч е с к а я ф у н к ц и я г е о м е т р и ч е с к о й с л у ч а й н о й в е л и ч и н ы. Известно, что

$$\mathrm{P}(\xi = \mathrm{k}) = p^k (1-p)^{n-k}, \quad k = 0,\ 1,\ \ldots,$$

поэтому

$$f_\xi(t) = \sum_{k=0}^{\infty} e^{itk} (1-p)^k\ p = p\sum_{k=0}^{\infty} \left((1-p)e^{it}\right)^k = p\frac{1}{1-(1-p)e^{it}}.$$

П р и м е р 5. Х а р а к т е р и с т и ч е с к а я ф у н к ц и я р а в н о м е р н о й н а о т р е з к е $[a;b]$ с л у ч а й н о й в е л и ч и н ы. Так как плотность распределения случайной величины ξ имеет вид

$$p_\xi(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a;b], \\ 0, & x \notin [a;b], \end{cases}$$

то по определению характеристической функции для абсолютно непрерывных случайных величии имеем (см. (2))

$$f_\xi(t) = \int_{-\infty}^{+\infty} e^{itx} p_\xi(x)\,dx = \int_a^b e^{itx} \frac{1}{b-a}\,dx = \frac{e^{ibt} - e^{iat}}{it(b-a)}.$$

Пусть случайная величина ξ равномерно распределена на отрезке $[-1;1]$, тогда

$$f_\xi(t) = \frac{\sin t}{t}, \qquad t \in R.$$

П р и м е р 6. Х а р а к т е р и с т и ч е с к а я ф у н к ц и я н о р м а л ь н о й с л у ч а й н о й в е л и ч и н ы с п а р а м е т р а м и $(a,\ \sigma^2)$. Рассмотрим сначала случай, когда $a = 0,\ \ \sigma = 1,$ тогда

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, \quad x \in R,$$

и

$$f_\xi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx} e^{\frac{-x^2}{2}}\,dx.$$

Продифференцируем обе части последнего равенства по t

$$f'_\xi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} ixe^{itx-\frac{x^2}{2}}\,dx \tag{9}$$

(дифференцирование под знаком интеграла по t законно, так как интеграл в равенстве (9) сходится равномерно относительно $x \in R$). Интегрируя (9) по частям, получим

$$f'_\xi(t) = -\frac{i}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx} d\left(e^{-\frac{x^2}{2}}\right) = -\frac{i}{\sqrt{2\pi}} e^{itx} e^{-\frac{x^2}{2}}\Bigg|_{-\infty}^{+\infty} - \frac{t}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx} e^{-\frac{x^2}{2}}\,dx = -t\,f_\xi(t).$$

Следовательно, $f_\xi(t)$ удовлетворяет следующей задаче Коши:

$$\begin{cases} f'_\xi(t) = -t\, f_\xi(t), \\ f_\xi(0) = 1. \end{cases}$$

Решая ее, находим $f_\xi(t) = e^{\frac{-t^2}{2}}$, $t \in R$.

Пусть случайная величина ξ имеет нормальное распределение с параметрами $(a,\ \sigma^2)$. Рассмотрим случайную величину

$$\eta = \frac{\xi - a}{\sigma}.$$

Она является нормальной случайной величиной с параметрами $(0,\ 1)$ (докажите это самостоятельно). Следовательно,

$$f_\eta(t) = e^{\frac{-t^2}{2}},$$

но, так как $\xi = \sigma\eta + a$, то по свойству 3° характеристических функций имеем

$$f_\xi(t) = e^{iat - \frac{\sigma^2 t^2}{2}}, \quad t \in R.$$

П р и м е р 7. Х а р а к т е р и с т и ч е с к а я ф у н к ц и я п о к а з а - т е л ь н о й с п а р а м е т р о м λ $(\lambda > 0)$ с л у ч а й н о й в е л и ч и н ы. Так как

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

то

$$f_\xi(t) = \lambda \int_0^{+\infty} e^{itx - \lambda x} dx = \lambda \left. \frac{e^{x(it - \lambda)}}{it - \lambda} \right|_0^{+\infty} = \frac{\lambda}{\lambda - it}, \quad t \in R.$$

П р и м е р 8. Х а р а к т е р и с т и ч е с к а я ф у н к ц и я с л у ч а й н о й в е л и ч и н ы К о ш и. Имеем

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

следовательно,

$$f_\xi(t) = \frac{a}{\pi} \int_{-\infty}^{+\infty} e^{itx} \frac{dx}{x^2 + a^2} = \frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{\cos tx}{x^2 + a^2} dx.$$

Заметим, что $f_\xi(t)$ – четная функция, поэтому достаточно знать ее значения при $t > 0$. Дифференцируя по t обе части последнего равенства, получим

$$f'_\xi(t) = \frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{-x \sin tx}{x^2 + a^2} dx. \tag{10}$$

Из курса математического анализа известно, что

$$\int_{-\infty}^{+\infty} \frac{\sin tx}{x} dx = \pi \quad (t > 0), \qquad \text{поэтому} \qquad a = \frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{\sin tx}{x} dx. \tag{11}$$

Складывая почленно равенства (10) и (11), получим

$$f'_\xi(t)+a=\frac{a}{\pi}\int_{-\infty}^{+\infty}\frac{a^2}{x^2+a^2}\frac{\sin tx}{x}dx.$$

Продифференцируем обе части последнего равенства по t. В результате получим $f''_\xi(t)=a^2 f_\xi(t)$. Следовательно $f_\xi(t)=c_1e^{at}+c_2e^{-at}$, $t>0$. Так как $f_\xi(t)$ ограниченная функция на R, то $c_1=0$, а из условия $f_\xi(0)=1$ получаем $c_2=1$. Итак, при $t>0$ $f_\xi(t)=e^{-at}$, но учитывая четность этой функции, окончательно имеем $f_\xi(t)=e^{-a|t|}$, $t\in R$.

У п р а ж н е н и е 1. Пусть плотность распределения случайной величины ξ имеет вид

$$p_\xi(x)=\frac{a}{\pi\left(a^2+(x-\theta)^2\right)},\ a>0,\ \theta\in R.$$

Убедитесь в том, что $f_\xi(t)=\exp\left(it\theta-a|t|\right)$ – ее характеристическая функция.

§ 2. Формулы обращения для характеристических функций

Задача этого параграфа – установить биекцию между множеством характеристических функций и множеством функций распределения.

Пусть случайная величина ξ абсолютно непрерывна, т. е. существует плотность распределения $p_\xi(x)$. По определению характеристическая функция

$$f_\xi(t)=\int_{-\infty}^{+\infty}e^{itx}p_\xi(x)dx$$

есть преобразование Фурье плотности. Из курса математического анализа известно (см. также следствие 2), что, если $f_\xi(t)\in L^1(R)$, то существует обратное преобразование Фурье, т. е.

$$p_\xi(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}f_\xi(t)dt.$$

Из приведенных формул видно, что в данном случае существует биекция между множеством характеристических функций и множеством плотностей распределения.

В общем случае

$$f_\xi(t)=\int_{-\infty}^{+\infty}e^{itx}dF_\xi(x).$$

Из этой формулы следует, что каждой функции распределения соответствует единственная характеристическая функция.

Покажем, что справедливо и обратное утверждение, т. е. что характеристическая функция однозначно определяет функцию распределения. Для этого сначала дадим явное представление функции распределения через характеристическую функцию. Формулы, описывающие это представление, называют *формулами обращения для характеристических функций*. В этом параграфе приведем две такие формулы. Вывод первой будем проводить по схеме, предложенной в [34]. Сначала приведем вспомогательные утверждения.

Лемма 1. *Пусть случайные величины ξ и η являются независимыми, $F(x)$ – функция распределения случайной величины ξ, а η – равномерно распределена на отрезке $[a;b]$. Тогда существует плотность распределения $p_{\xi+\eta}$, которая выражается формулой:*

$$p_{\xi+\eta}(x)=\frac{F(x-a)-F(x-b)}{b-a}.$$

Доказательство. По формуле свертки, учитывая вид плотности случайной величины η, имеем

$$F_{\xi+\eta}(x)=\int_{-\infty}^{+\infty}F_{\xi}(x-t)\,p_{\eta}(t)\,dt=\int_{-\infty}^{+\infty}F(x-t)\,p_{\eta}(t)\,dt=\frac{1}{b-a}\int_{a}^{b}F(x-t)\,dt.$$

После дифференцирования обеих частей этого равенства по x получаем доказательство леммы 1.

Лемма 2 [34]. *Пусть случайные величины ξ и η независимы и абсолютно непрерывны, причем плотность распределения $p_{\xi}(x)=p(x)$ – ограниченная функция. Через $p_{\theta}(x)$ обозначим плотность распределения случайной величины $\xi+\theta\eta$, $\theta\in R$. Тогда в точках непрерывности $p(x)$ справедливо равенство*

$$p(x)=\lim_{\theta\to 0}p_{\theta}(x).$$

Доказательство формулы обращения. Пусть случайные величины ξ, η, ς независимы, причем ξ – произвольная случайная величина, $F(x)$ – ее функция распределения, а $f(t)$ – характеристическая функция; случайная величина η равномерно распределена на отрезке $[-1;1]$, а ς – нормальная случайная величина с параметрами $(0,\ 1)$.

Рассмотрим линейную комбинацию $\xi + l\eta + \sigma\varsigma$, где $l, \sigma \in R$. Несложно видеть, что данная случайная величина абсолютно непрерывна, ибо 2-я и 3-я компоненты суммы – абсолютно непрерывные случайные величины. Известно (см. дополнения к гл. 3, § 1), что, если в сумме независимых случайных величин хотя бы одна компонента абсолютно непрерывна, то сумма – абсолютно непрерывная случайная величина.

Найдем представление характеристической функции данной случайной величины.

В гл. 5, § 1 было показано, что характеристическая функция η

$$f_\eta(t) = \frac{\sin t}{t},$$

поэтому

$$f_{l\eta}(t) = f_\eta(lt) = \frac{\sin tl}{tl}.$$

Аналогично, так как

$$f_\varsigma(t) = e^{-\frac{t^2}{2}}, \quad \text{то} \quad f_{\sigma\varsigma}(t) = e^{-\frac{\sigma^2 t^2}{2}}.$$

Таким образом,

$$f_{\xi+l\eta+\sigma\varsigma} = f(t)\,\frac{\sin tl}{tl}\,e^{-\frac{\sigma^2 t^2}{2}} \in L^1(R),$$

следовательно, существует обратное преобразование Фурье

$$p_{\xi+l\eta+\sigma\varsigma}(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx} f(t)\frac{\sin tl}{tl} e^{-\frac{\sigma^2 t^2}{2}}\,dt. \tag{1}$$

Из леммы 1 следует:

$$p_{\xi+l\eta}(x) = \frac{F(x+l)-F(x-l)}{2l}. \tag{2}$$

Из леммы 2 получаем, что для любой точки x, являющейся точкой непрерывности плотности $p_{\xi+l\eta}$

$$p_{\xi+l\eta}(x) = \lim_{\sigma\to 0} p_{\xi+l\eta+\sigma\varsigma}(x). \tag{3}$$

Из формул (1) – (3) следует

$$F(x+l)-F(x-l) = \lim_{\sigma\to 0}\frac{1}{\pi}\int_{-\infty}^{+\infty} e^{-itx} f(t)\frac{\sin tl}{t} e^{-\frac{\sigma^2 t^2}{2}}\,dt, \tag{4}$$

где $x+l$, $x-l$ – произвольные точки непрерывности функции распределения $F(x)$.

Формула (4) – формула обращения. Таким образом, доказана

Теорема 1 (**Формула обращения**). *Пусть функции F и f – соответственно функция распределения и характеристическая функция случайной величины ξ. Тогда для любых точек непрерывности $x+l$ и $x-l$ (для $l>0$) функции F справедливо равенство* (4).

Приведем еще одно представление формулы обращения для характеристических функций.

Теорема 2 (**Формула обращения**). *Пусть функции F и f – соответственно функция распределения и характеристическая функция случайной величины ξ. Тогда для любых точек непрерывности x_1 и x_2 $(x_1<x_2)$ функции F*

$$F(x_2)-F(x_1)=\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-itx_1}-e^{-itx_2}}{it}f(t)\,dt. \tag{5}$$

Д о к а з а т е л ь с т в о. Из определения характеристической функции $f(t)$ и теоремы Фубини имеем

$$Q(T)=\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-itx_1}-e^{-itx_2}}{it}f(t)\,dt=\frac{1}{2\pi}\int_{-T}^{T}\left(\frac{e^{-itx_1}-e^{-itx_2}}{it}\int_{-\infty}^{+\infty}e^{itx}dF(x)\right)dt=$$

$$=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left(\int_{-T}^{T}\frac{e^{it(x-x_1)}-e^{it(x-x_2)}}{it}dt\right)dF(x)=$$

$$=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-T}^{T}\frac{\sin t(x-x_1)-\sin t(x-x_2)}{t}dt\,dF(x)=$$

$$=\frac{1}{\pi}\int_{-\infty}^{+\infty}\int_{0}^{T}\frac{\sin t(x-x_1)-\sin t(x-x_2)}{t}dt\,dF(x)=\frac{1}{\pi}\int_{-\infty}^{+\infty}\int_{T(x-x_2)}^{T(x-x_1)}\frac{\sin u}{u}du\,dF(x).$$

Так как

$$\lim_{T\to+\infty}\int_{-T}^{T}\frac{\sin u}{u}du=\pi,$$

то

$$\psi(x)=\lim_{T\to\infty}\frac{1}{\pi}\int_{T(x-x_2)}^{T(x-x_1)}\frac{\sin u}{u}du=\begin{cases}0, & \text{если } x<x_1 \text{ или } x>x_2;\\ \frac{1}{2}, & \text{если } x=x_1 \text{ или } x=x_2;\\ 1, & \text{если } x_1<x<x_2.\end{cases}$$

Поэтому

$$\lim_{T\to+\infty}Q(T)=\int_{-\infty}^{+\infty}\psi(x)dF(x)=\mathrm{M}\psi(\xi)=$$

$$=\frac{1}{2}\mathrm{P}(\xi=x_1)+\frac{1}{2}\mathrm{P}(\xi=x_2)+\mathrm{P}(x_1<\xi<x_2),$$

но так как x_1 и x_2 – точки непрерывности функции распределения F, то

$$\mathrm{P}(\xi=x_1)=\mathrm{P}(\xi=x_2)=0, \quad \text{а} \quad \mathrm{P}(x_1<\xi<x_2)=F(x_2)-F(x_1).$$

Теорема доказана.

Следствие 1 (**Теорема единственности**). *Каждой характеристической функции соответствует единственная функция распределения.*

Доказательство. Из формулы обращения (5) для любых $x\in R$ имеем

$$F(x)=\lim_{y\to-\infty}\lim_{T\to+\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ity}-e^{-itx}}{it}f_\xi(t)\,dt,$$

где y пробегает точки непрерывности функции F. Теорема единственности доказана.

Замечание 1. Из теоремы единственности и определения характеристической функции вытекает существование биекции между множествами функций распределения и характеристических функций.

Следствие 2. *Пусть характеристическая функция f_ξ случайной величины ξ интегрируема на R, тогда существует плотность распределения p_ξ и*

$$p_\xi(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}f_\xi(t)\,dt, \quad x\in R.$$

Доказательство. Пусть точки a и b являются точками непрерывности функции распределения F_ξ, тогда по формуле обращения (5) имеем

$$F_\xi(b)-F_\xi(a)=\lim_{T\to+\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}f_\xi(t)\,dt=$$

$$=\lim_{T\to+\infty}\frac{1}{2\pi}\int_{-T}^{T}f_\xi(t)\int_a^b e^{-itx}\,dx\,dt=\int_a^b\frac{1}{2\pi}\left(\int_{-\infty}^{+\infty}e^{-itx}f_\xi(t)\,dt\right)dx=\int_a^b p_\xi(x)\,dx.$$

Таким образом,

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

т. е. p_ξ – плотность распределения случайной величины ξ. Следствие 2 доказано.

С л е д с т в и е 3. *Пусть* ξ *– целочисленная случайная величина,* $p_k = \mathrm{P}(\xi = k)$, *тогда*

$$p_k = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} e^{-ikt} f_\xi(t)\,dt, \quad k = 0,\ \pm 1,\ \pm 2,\ \ldots,$$

где f_ξ *– характеристическая функция* ξ.

Д о к а з а т е л ь с т в о. По определению характеристической функции дискретной случайной величины имеем (см. гл. 5, § 1, равенство (3))

$$f_\xi(t) = \sum_{j=-\infty}^{+\infty} e^{itj} p_j.$$

Умножая это равенство на e^{-ikt} и интегрируя ряд почленно (он сходится равномерно, так как $\left| p_j e^{itj} \right| = p_j$, $\sum\limits_{j=-\infty}^{+\infty} p_j = 1$) на отрезке $[-\pi; \pi]$, приходим к доказательству следствия 3.

§ 3. Непрерывность соответствия между множествами функций распределения и характеристических функций

В § 2 показано, что между множествами функций распределения и характеристических функций существует взаимно однозначное соответствие (биекция). Цель этого параграфа – установить, что это соответствие и взаимно непрерывно, т. е. является гомеоморфизмом. Для этого, прежде всего, определимся с топологиями на данных множествах. На множестве характеристических функций будем рассматривать топологию поточечной или равномерной на компактах сходимости. На множестве функций распределения будет задействована топология, так называемой, слабой сходимости.

О п р е д е л е н и е 1. *Говорят, что последовательность функций распределения* $\{F_n\}$ *слабо сходится к функции* F *и обозначают*

$$F_n \underset{n\to\infty}{\Rightarrow} F,$$

если

$$F_n(x) \xrightarrow[n\to\infty]{} F(x)$$

для любых x, *являющихся точками непрерывности функции* F.

З а м е ч а н и е 1. Данным определением мы будем пользоваться и в том случае, когда F не является функцией распределения.

То, что слабый предел последовательности функций распределения не обязательно является функцией распределения, показывает следующий пример.

П р и м е р 1. Рассмотрим последовательность функций распределения $\{F_n\}$, где

$$F_n(x)=\begin{cases}0, & x<n,\\ 1, & x\ge n,\end{cases}$$

тогда $F_n \underset{n\to\infty}{\Rightarrow} F$, где $F(x)\equiv 0,\ x\in R$.

З а м е ч а н и е 2. Очевидно, что любая последовательность функций распределения слабо сходится к функции Дирихле.

Л е м м а 1. *Пусть последовательность функций распределения $\{F_n\}$ для любых $x\in D$, где множество D всюду плотно на R, сходится к функции распределения F при $n\to\infty$, тогда $F_n \underset{n\to\infty}{\Rightarrow} F$.*

Д о к а з а т е л ь с т в о. Пусть точка x – произвольная, но фиксированная точка непрерывности функции F. Покажем, что

$$F_n(x)\xrightarrow[n\to\infty]{}F(x).$$

Так как $\bar{D}=R$, то существуют точки $x',x''\in D$, такие что $x'<x<x''$.

Из условия леммы непосредственно вытекает, что

$$F(x')=\lim_{n\to\infty}F_n(x')\le\underline{\lim}_{n\to\infty}F_n(x)\le\overline{\lim}_{n\to\infty}F_n(x)\le\lim_{n\to\infty}F_n(x'')=F(x''),$$

но x – точка непрерывности F, следовательно, числа $F(x')$, $F(x'')$ могут быть сделаны сколь угодно близкими. Лемма 1 доказана.

Л е м м а 2 (**Первая теорема Хелли**). *Из любой последовательности функций распределения можно извлечь слабо сходящуюся подпоследовательность, причем предельная функция F обладает следующими свойствами:*

1. *$F(x)$ монотонно не убывает,*
2. *$F(x)$ непрерывна справа,*
3. *$0\le F(x)\le 1$ для любых $x\in R$.*

Д о к а з а т е л ь с т в о проведем методом диагонализации Кантора.

Пусть $D=\{x_k\}$ – счетно и всюду плотно на R. Рассмотрим числовую последовательность $\{F_n(x_1)\}$. Учитывая то, что F_n – функция распределения, имеем

$$0 \le \{F_n(x_1)\} \le 1, \ n \ge 1.$$

По теореме Вейерштрасса из этой числовой последовательности можно извлечь сходящуюся подпоследовательность $\{F_{n_1}(x_1)\}$, предел которой обозначим через $F(x_1)$.

Теперь рассмотрим числовую последовательность $\{F_{n_1}(x_2)\}$, $0 \le \{F_{n_1}(x_2)\} \le 1$. Из этой ограниченной числовой последовательности выделим сходящуюся подпоследовательность $\{F_{n_2}(x_2)\}$, а ее предел обозначим через $F(x_2)$. И так далее.

Наконец, рассмотрим диагональную последовательность функций распределения $\{F_{n_n}\}$, $\{F_{n_n}\} \subset \{F_n\}$, и по построению

$$F_{n_n}(x_k) \xrightarrow[n\to\infty]{} F(x_k)$$

для любых $x_k \in D$. Следовательно, по лемме 1, $F_{n_n} \underset{n\to\infty}{\Rightarrow} F$.

Доказательство того, что функция F обладает свойствами из условия леммы 2, следует из свойств пределов.

Лемма 3 (**Вторая теорема Хелли**). *Пусть последовательность функций распределения* $\{F_n\}$ *при* $n \to \infty$ *слабо сходится к функции* F *такой, что* $F(+\infty) - F(-\infty) = 1$. *Тогда для любой непрерывной и ограниченной функции* $g: R \to R$

$$\int_{-\infty}^{+\infty} g(x)\,dF_n(x) \xrightarrow[n\to\infty]{} \int_{-\infty}^{+\infty} g(x)\,dF(x). \tag{1}$$

Замечание 3. Несложно видеть, что лемму 3 можно сформулировать следующим образом: *пусть* $F_n \underset{n\to\infty}{\Rightarrow} F$, *где* F_n *и* F *– функции распределения. Тогда для любой непрерывной ограниченной функции* $g: R \to R$ *выполняется соотношение* (1).

Схема доказательства

Доказательство проведем в два этапа.

1) Покажем, что для любых $a < b$ – точек непрерывности функции F выполняется соотношение

$$\int_a^b g(x)\,dF_n(x) \xrightarrow[n\to\infty]{} \int_a^b g(x)\,dF(x). \tag{2}$$

Так как $g(x)$ – непрерывна на отрезке $[a;b]$, следовательно, она и равномерно непрерывна на $[a;b]$. Поэтому для любого $\varepsilon > 0$ существует такое разбиение отрезка $[a;b]$: $a = x_0 < x_2 < \ldots < x_L = b$, что $\left|g(x) - g_\varepsilon(x)\right| < \varepsilon$, для любых $x \in [a;b]$, где по определению

$$g_\varepsilon(x) = g(x_k), \qquad \forall x \in (x_{k-1}; x_k] \qquad k = \overline{1, L}.$$

В силу того, что множество точек непрерывности функции F всюду плотно, будем считать, что все x_k – точки непрерывности функции F. Теперь, учитывая ограниченность функции g и свойства интеграла Римана – Стилтьеса, имеем

$$\left|\int_a^b g(x)dF_n(x) - \int_a^b g(x)dF(x)\right| =$$

$$= \left|\int_a^b g(x)dF_n(x) - \int_a^b g_\varepsilon(x)dF_n(x) + \int_a^b g_\varepsilon(x)dF_n(x) - \int_a^b g_\varepsilon(x)dF(x) + \right.$$

$$\left. + \int_a^b g_\varepsilon(x)dF(x) - \int_a^b g(x)dF(x)\right| \le \int_a^b \left|g(x) - g_\varepsilon(x)\right| dF_n(x) +$$

$$+ \left|\sum_{k=1}^n \int_{x_{k-1}}^{x_k} g(x_k)dF_n(x) - \sum_{k=1}^n \int_{x_{k-1}}^{x_k} g(x_k)dF(x)\right| + \int_a^b \left|g(x) - g_\varepsilon(x)\right| dF(x) \le$$

$$\le \varepsilon\big(F_n(b) - F_n(a)\big) + \varepsilon\big(F(b) - F(a)\big) +$$

$$+ \left|\sum_{k=1}^L g(x_k)\big[F_n(x_k) - F_n(x_{k-1}) - F(x_k) + F(x_{k-1})\big]\right| \le$$

$$\le 2\varepsilon + M\sum_{k=1}^L \left|\big(F_n(x_k) - F(x_k)\big) - \big(F_n(x_{k-1}) - F(x_{k-1})\big)\right|.$$

Для доказательства соотношения (2) сначала выбираем L достаточно большим, а затем устремляем $n \to \infty$ и учитываем сходимость последовательности функции распределения F_n к функции F в точках x_k, $k = \overline{1, L}$, при $n \to \infty$.

2) Покажем, что справедливо соотношение (1). Так как по замечанию 3 F – функция распределения, то для любого $\varepsilon > 0$ существуют точки $-X$, X – точки непрерывности функции F такие что

$$F(-X) < \frac{\varepsilon}{2}, \qquad 1 - F(X) < \frac{\varepsilon}{2}.$$

Из того, что $F_n \Rightarrow F$ при $n \to \infty$ найдется такое $n_0 \in N$, что для всех $n > n_0$

$$F_n(-X) < \varepsilon, \qquad 1 - F_n(X) < \varepsilon.$$

Учитывая все вышесказанное, а также ограниченность функции g, получаем

$$\left| \int_{-\infty}^{+\infty} g(x)\,dF_n(x) - \int_{-\infty}^{+\infty} g(x)\,dF(x) \right| \le$$

$$\le \int_{|x|>X} |g(x)|\,dF_n(x) + \left| \int_{|x|\le X} g(x)\,dF_n(x) - \int_{|x|\le X} g(x)\,dF(x) \right| +$$

$$+ \int_{|x|>X} |g(x)|\,dF(x) \le M\left(F_n(-X) + 1 - F_n(X)\right) + \varepsilon +$$

$$+ M\left(F(-X) + 1 - F(X)\right) \le 3M\varepsilon + \varepsilon.$$

Лемма 3 доказана.

Заметим, что справедливо утверждение, обратное второй теореме Хелли:

Утверждение 1 [13]. *Если для любой непрерывной и ограниченной функции* $g:\ R \to R$

$$\int_{-\infty}^{+\infty} g(x)\,dF_n(x) \xrightarrow[n\to+\infty]{} \int_{-\infty}^{+\infty} g(x)\,dF(x),$$

где $\{F_n\}$, $n = 1,\ 2,\ \ldots,\ F$ *– функции распределения, то* $F_n \underset{n\to\infty}{\Rightarrow} F$.

Теорема 1 (**Прямая предельная теорема**). *Если последовательность функций распределения* $\{F_n\}$ *слабо сходится к функции распределения* F *при* $n \to \infty$, *то последовательность соответствующих характеристических функций* $\{f_n\}$ *сходится поточечно к характеристической функции* f.

Замечание 4. Поточечную сходимость последовательности характеристических функций в этой теореме можно заменить равномерной сходимостью на любом компакте из R.

Доказательство этой теоремы непосредственно вытекает из 2-й теоремы Хелли и определения характеристической функции: в качестве функции g возьмем $g(x) = e^{itx}$, $x \in R$, а на i, t смотрим как на параметры.

Теорема 2 (**Обратная предельная теорема**). *Пусть последовательность характеристических функции* $\{f_n\}$ *сходится поточечно к функции* f, *непрерывной в точке* 0. *Тогда последовательность соот-*

ветствующих функций распределения $\{F_n\}$ слабо сходится к функции распределения F и f является характеристической функцией, соответствующей функции распределения F.

Схема доказательства

Доказательство. Пусть $\{F_n\}$ – последовательность функций распределения соответствующих последовательности характеристических функций $\{f_n\}$. Из 1-ой теоремы Хелли следует, что существует слабо сходящаяся подпоследовательность

$$\{F_{n_k}\} \subset \{F_n\},$$

такая что $F_{n_k} \Rightarrow F$. Докажем теперь, что F является функцией распределения. Для этого достаточно показать что $F(+\infty) - F(-\infty) = 1$.

В дальнейшем для доказательства понадобится следующее неравенство: *пусть ξ – произвольная случайная величина, f – ее характеристическая функция, тогда для любых $\tau > 0$ и $x > 0$*

$$\mathrm{P}(|\xi| \le x) \ge \frac{\left|\frac{1}{2\tau}\int_{-\tau}^{\tau} f(t)\,dt\right| - \frac{1}{\tau x}}{1 - \frac{1}{\tau x}}. \quad (3)$$

Положим $\tau x = 2$, тогда неравенство (3) примет вид

$$\mathrm{P}(|\xi| \le x) \ge \frac{\left|\frac{1}{2\tau}\int_{-\tau}^{\tau} f(t)\,dt\right| - \frac{1}{2}}{1 - \frac{1}{2}} = 2\left|\frac{1}{2\tau}\int_{-\tau}^{\tau} f(t)\,dt\right| - 1. \quad (4)$$

Докажем неравенство (3). Из определения характеристической функции и теоремы Фубини следует

$$\left|\frac{1}{2\tau}\int_{-\tau}^{\tau} f(t)\,dt\right| = \left|\frac{1}{2\tau}\int_{-\tau}^{\tau} \mathrm{M} e^{it\xi}\,dt\right| = \left|\frac{1}{2\tau}\mathrm{M}\int_{-\tau}^{\tau} e^{it\xi}\,dt\right| = \left|\frac{1}{2\tau}\mathrm{M}\frac{\sin\tau\xi}{\xi}\right| =$$

$$= \left|\frac{1}{\tau}\mathrm{M}\frac{\sin\tau\xi}{\xi}\right| I_{\{|\xi|\le x\}} + \left|\frac{1}{\tau}\mathrm{M}\frac{\sin\tau\xi}{\xi}\right| I_{\{|\xi|> x\}} \le$$

$$\le \mathrm{M}\left|\frac{\sin\tau\xi}{\tau\xi}\right| I_{\{|\xi|\le x\}} + \mathrm{M}\left|\frac{\sin\tau\xi}{\tau\xi}\right| I_{\{|\xi|> x\}} \le \mathrm{P}(|\xi| \le x) + \frac{1}{\tau x}\big(1 - \mathrm{P}(|\xi| \le x)\big).$$

Несложно видеть, что последнее неравенство совпадает с неравенством (3).

Так как функция f непрерывна в точке 0 и является поточечным пределом характеристических функций $\{f_n\}$, то $f(0)=1$ и для любого $\varepsilon>0$ существует такое $\tau_0>0$, что для всех τ, удовлетворяющих неравенству $0<\tau\le\tau_0$, выполнено

$$\left|\frac{1}{2\tau}\int_{-\tau}^{\tau} f(t)\,dt\right|>1-\frac{\varepsilon}{4}. \qquad (5)$$

Из того, что $f_n\to f$ при $n\to\infty$ вытекает для всех $n>n_0$ и $\tau\in(0;\tau_0]$,

$$\left|\frac{1}{2\tau}\int_{-\tau}^{\tau} f_n(t)\,dt-\frac{1}{2\tau}\int_{-\tau}^{\tau} f(t)\,dt\right|<\frac{\varepsilon}{4}. \qquad (6)$$

Из неравенств (5) и (6) следует, что для любых $n>n_0$ и τ, таких что $0<\tau\le\tau_0$

$$\left|\frac{1}{2\tau}\int_{-\tau}^{\tau} f_n(t)\,dt\right|>1-\frac{\varepsilon}{2}. \qquad (7)$$

Из неравенств (7) и (5) имеем

$$F_{n_k}\left(\frac{2}{\tau}\right)-F_{n_k}\left(-\frac{2}{\tau}-0\right)\ge \mathrm{P}\left(\left|\xi_{n_k}\right|\le\frac{2}{\tau}\right)\ge 2\left(1-\frac{\varepsilon}{2}\right)-1=1-\varepsilon,$$

для всех $n>n_0$ и $0<\tau\le\tau_0$. Из последнего неравенства в силу произвольности τ и ε получаем

$$F(+\infty)-F(-\infty)=1,$$

т. е. F – функция распределения. По прямой предельной теореме из доказанного следует

$$f_{n_k}(t)\xrightarrow[n_k\to\infty]{}\int_{-\infty}^{\infty} e^{itx}\,dF(x),\quad t\in R.$$

Но по условию теоремы

$$f_n(t)\xrightarrow[n\to\infty]{} f(t),\quad t\in R,$$

следовательно,

$$f(t)=\int_{-\infty}^{\infty} e^{itx}\,dF(x)$$

– характеристическая функция, соответствующая функции распределения F.

Докажем теперь, что

$$F_n \underset{n\to\infty}{\Rightarrow} F.$$

Предположим противное: пусть

$$F_n \not\Rightarrow F$$

при $n \to \infty$. Тогда существует подпоследовательность

$$\left\{F_{m_k}\right\} \subset \left\{F_n\right\},\ F_{m_k} \Rightarrow F^*,\ F^* \neq F,$$

причем F и F^* – функции распределения. По прямой предельной теореме имеем

$$f_{n_k}(t) \to f(t),$$

$$f_{m_k}(t) \to f^*(t),\ \ k \to \infty,$$

и по теореме единственности $f(t) \neq f^*(t)$, но этого не может быть, так как

$$f_n(t) \xrightarrow[n\to\infty]{} f(t),$$

следовательно,

$$f(t) = f^*(t).$$

Теорема доказана.

ДОПОЛНЕНИЯ К ГЛАВЕ 5

§ 1. Производящие функции

Для неотрицательных целочисленных случайных величин удобно пользоваться производящими функциями.

О п р е д е л е н и е 1. *Пусть* ξ *– целочисленная неотрицательная случайная величина. Производящей функцией* φ_ξ *случайной величины* ξ *называется функция комплексного переменного* z, $|z| \le 1$,

$$\varphi_\xi(z) = \mathrm{M}z^\xi = \sum_{k=0}^{\infty} z^k \mathrm{P}(\xi = k). \tag{1}$$

З а м е ч а н и е 1. Так как $|e^{it}| = 1$, $t \in R$, то для $z = e^{it}$ очевидно равенство $\varphi_\xi(z) = f_\xi(t)$, где f_ξ – характеристическая функция случайной величины ξ.

Свойства производящих функций

1°. *Если* $\xi_1, \ \ldots, \ \xi_n$ *– независимые случайные величины, то*

$$\varphi_{\xi_1+\ldots+\xi_n}(z) = \varphi_{\xi_1}(z)\cdot\ldots\cdot\varphi_{\xi_n}(z), \qquad |z| \le 1.$$

Д о к а з а т е л ь с т в о. Из определения производящей функции и независимости случайных величин $z^{\xi_1}, \ \ldots, \ z^{\xi_n}$ имеем

$$\varphi_{\xi_1+\ldots+\xi_n}(z) = \mathrm{M}\, z^{\xi_1+\ldots+\xi_n} = \mathrm{M}\, z^{\xi_1}\cdot\ldots\cdot z^{\xi_n} = \mathrm{M}\, z^{\xi_1}\cdot\ldots\cdot\mathrm{M}\, z^{\xi_n} = \varphi_{\xi_1}(z)\cdot\ldots\cdot\varphi_{\xi_n}(z).$$

2°. *Если* $\mathrm{P}(\eta = a\xi + b) = 1$, *то* $\varphi_\eta(z) = z^b \varphi_{a\xi}(z)$.

Д о к а з а т е л ь с т в о. По определению производящей функции и свойствам математического ожидания получаем

$$\varphi_\eta(z) = \mathrm{M}\, z^\eta = \mathrm{M}\, z^{a\xi+b} = z^b\, \mathrm{M}\, z^{a\xi} = z^b \varphi_{a\xi}(z).$$

3°. *Если* φ_ξ *– производящая функция случайной величины* ξ, *то*

$$\mathrm{P}(\xi = k) = \varphi_\xi^{(k)}(0)\frac{1}{k!}, \qquad k = 0, \ 1, \ 2, \ \ldots \ .$$

Д о к а з а т е л ь с т в о непосредственно вытекает из определения производящей функции.

4°. *Если* $\mathrm{M}\xi^n < \infty$, *то* $\varphi_\xi^{(n)}(1) = \mathrm{M}\,\xi(\xi-1)\ldots(\xi-n+1)$.

Д о к а з а т е л ь с т в о. Так как

$$\mathrm{M}\xi^n = \sum_{k=0}^{\infty} k^n \mathrm{P}(\xi = k) < \infty,$$

то при всех $|z| \le 1$ будет сходиться ряд

$$\sum_{k=0}^{\infty} k(k-1)\ldots(k-n+1) z^{k-n} \mathrm{P}(\xi = k) = \varphi_\xi^{(n)}(z).$$

Отсюда

$$\varphi_\xi^{(n)}(1)=\sum_{k=0}^{\infty}k(k-1)\ldots(k-n+1)\mathrm{P}(\xi=k)=\mathrm{M}\xi(\xi-1)\ldots(\xi-n+1).$$

Свойство 4° доказано.

5°. *Пусть* $\xi_1,\ \xi_2,\ \ldots$ *– последовательность независимых неотрицательных целочисленных одинаково распределенных случайных величин, а* τ *– независимая с ними целочисленная положительная случайная величина и* $\mathrm{M}\tau<\infty$. *Тогда*

$$\varphi_{S_\tau}(z)=\varphi_\tau\left(\varphi_{\xi_1}(z)\right),$$

где $S_\tau=\xi_1+\ldots+\xi_\tau$ *(это значит* $S_\tau(\omega)=\xi_1(\omega)+\ldots+\xi_n(\omega)$,, *если* $\omega\in(\tau=n)$, $n\in N$*).*

Доказательство. Из свойств условного математического ожидания имеем

$$\varphi_{S_\tau}(z)=\mathrm{M}z^{S_\tau}=\sum_{n=1}^{\infty}\mathrm{M}\left(z^{S_n}\mid\tau=n\right)\mathrm{P}(\tau=n)=\sum_{n=1}^{\infty}\mathrm{M}z^{S_n}\mathrm{P}(\tau=n)=$$

$$=\sum_{n=1}^{\infty}\varphi_{\xi_1}^{n}(z)\mathrm{P}(\tau=n)=\varphi_\tau\left(\varphi_{\xi_1}(z)\right).$$

Свойство 5° доказано.

Пример 1. Найдем производящую функцию биномиальной случайной величины. По определению производящей функции имеем

$$\varphi_\xi(z)=\mathrm{M}z^\xi=\sum_{k=0}^{n}z^k\mathrm{P}(\xi=k)=\sum_{k=0}^{n}z^kC_n^kp^k(1-p)^{n-k}=\sum_{k=0}^{n}C_n^k(pz)^k(1-p)^{n-k}=$$

$$=(1-p+pz)^k,\qquad |z|\le 1.$$

Пример 2. Производящая функция геометрической случайной величины. Так как

$$\mathrm{P}(\xi=k)=(1-p)^kp,\qquad k=0,\ 1,\ 2,\ \ldots,$$

то

$$\varphi_\xi(z)=\mathrm{M}\,z^\xi=\sum_{k=0}^{\infty}z^k(1-p)^kp=\frac{p}{1-(1-p)z},\qquad |z|\le 1.$$

Пример 3. Однородное случайное блуждание на прямой. Рассмотрим случайное блуждание частицы по целочисленным точкам числовой прямой, описываемое следующей схемой: в начальный момент времени $t=0$ частица находится в точке $x=0$, а затем в моменты времени $t=1,\ 2,\ \ldots$ совершает скачки на единицу вправо с вероятностью p или на единицу влево с вероятностью $q=1-p$.

С помощью аппарата производящих функций можно показать следующее (см., напр., [10]):

1) вероятность того, что частица когда-либо вернется в исходную точку равна $1-|p-q|$;

2) математическое ожидание числа шагов до первого возвращения частицы в исходное положение равно $\frac{4pq}{|p-q|}$, в частности, при $p=q=\frac{1}{2}$, среднее число шагов до первого возвращения бесконечно.

§ 2. Решетчатые распределения

Будем говорить, что дискретная случайная величина ξ имеет *решетчатое распределение*, если существуют числа a и h ($h>0$), такие что все возможные значения ξ могут быть представлены в виде $a+kh$, $k\in Z$. Число h называют *шагом распределения*.

Условие решетчатости можно выразить в терминах характеристических функций.

Теорема 1. *Для того чтобы случайная величина ξ имела решетчатое распределение, необходимо и достаточно, чтобы при некотором $t\neq 0$ модуль ее характеристической функции был равен единице.*

Доказательство. Необходимость. Пусть случайная величина ξ имеет решетчатое распределение и

$$p_k = \mathrm{P}(\xi = a+kh), \quad k\in Z.$$

Тогда характеристическая функция ξ равна

$$f_\xi(t) = \sum_{k=-\infty}^{\infty} p_k e^{it(a+kh)} = e^{ita}\sum_{k=-\infty}^{\infty} p_k e^{ikht}.$$

Поэтому

$$f_\xi\left(\frac{2\pi}{h}\right) = e^{ia\frac{2\pi}{h}}\sum_{k=-\infty}^{\infty} p_k e^{i2\pi k} = e^{ia\frac{2\pi}{h}}$$

и

$$\left|f_\xi\left(\frac{2\pi}{h}\right)\right| = 1.$$

Достаточность. Пусть при некотором $t_1\neq 0$ $|f_\xi(t_1)|=1$. Тогда

$$f_\xi(t_1) = e^{it_1 a}, \quad a\in R,$$

и

$$\int_{-\infty}^{\infty} e^{it_1 x} dF_\xi(x) = e^{it_1 a}.$$

Отсюда

$$\int_{-\infty}^{\infty} e^{it_1(x-a)} dF_\xi(x) = 1,$$

следовательно,

$$\int_{-\infty}^{\infty} \cos t_1(x-a)\, dF_\xi(x) = 1,$$

или

$$\int_{-\infty}^{\infty}\left(1-\cos t_1\left(x-a\right)\right)dF_{\xi}\left(x\right)=0.$$

Последнее равенство можно записать в виде $\mathrm{M}\left(1-\cos t_1\left(\xi-a\right)\right)=0,$ но так как $1-\cos t_1\left(\xi-a\right)\geq 0,$ то по свойству интеграла Лебега $\mathrm{P}\left(1-\cos t_1\left(\xi-a\right)=0\right)=1,$ поэтому возможные значения ξ имеют вид

$$a+k\frac{2\pi}{t_1},\qquad k\in Z,$$

т. е. случайная величина ξ имеет решетчатое распределение.

Теорема 1 доказана.

Шаг распределения h называется *максимальным*, если ни при каких $h_1>h$ нельзя представить все возможные значения ξ в виде $b+kh_1,\ k\in Z.$

Теорема 2. *Шаг распределения h будет максимальным тогда и только тогда, когда модуль характеристической функции при $0<t<\frac{2\pi}{h}$ строго меньше единицы и*

$$\left|f_{\xi}\left(\frac{2\pi}{h}\right)\right|=1.$$

Доказательство. Если при $0<t_1<\frac{2\pi}{h}$

$$\left|f_{\xi}\left(t_1\right)\right|=1,$$

то из теоремы 1 $\frac{2\pi}{t_1}$ – шаг распределения, а так как $h<\frac{2\pi}{t_1},$ то шаг h не может быть максимальным.

Упражнение 1. Докажите, что, если $\left|f_{\xi}\left(t\right)\right|=\left|f_{\xi}\left(\alpha t\right)\right|=1$ для двух различных точек t и $\alpha t,$ где α – иррациональное число, то случайная величина ξ является вырожденной, т. е. существует $a\in R$ такое, что

$$\mathrm{P}\left(\xi=a\right)=1.$$

Упражнение 2. Докажите, что если $\left|f_{\xi}\left(t\right)\right|\equiv 1,$ то случайная величина ξ – вырожденная.

§ 3. Многомерные характеристические функции

Пусть $\overline{\xi}=\left(\xi_1,\ \ldots,\ \xi_m\right)$ – случайный вектор, определенный на вероятностном пространстве $\left(\Omega,\ \mathcal{A},\ \mathrm{P}\right).$

Определение 1. *Многомерной характеристической функцией случайного вектора $\overline{\xi}$ назовем функцию*

$$f_{\overline{\xi}}\left(\overline{t}\right)=\mathrm{M}\,e^{i\left(\overline{t},\,\overline{\xi}\right)},$$

где $\overline{t}=(t_1,\ \dots,\ t_n)\in R^n$, $(\overline{t},\ \overline{\xi})=\sum_{k=1}^{n}t_k\xi_k$.

По аналогии с одномерным случаем отметим также, что:

1) в общем случае многомерная характеристическая функция допускает и такое представление

$$f_{\overline{\xi}}(\overline{t})=\int_{R^n}e^{i(\overline{t},\overline{x})}dF_{\overline{\xi}}(\overline{x}),\qquad \overline{t}\in R^n,$$

где $F_{\overline{\xi}}$ – функция распределения случайного вектора $\overline{\xi}$;

2) если случайный вектор $\overline{\xi}$ абсолютно непрерывен, то

$$f_{\overline{\xi}}(\overline{t})=\int_{R^n}e^{i(\overline{t},\overline{x})}p_{\overline{\xi}}(\overline{x})d\overline{x},\qquad \overline{t}\in R^n,$$

где $p_{\overline{\xi}}$ – плотность распределения случайного вектора $\overline{\xi}$;

3) если $\overline{\xi}$ – дискретен с возможными значениями $\overline{x}_1,\ \overline{x}_2,\ \dots,\ \overline{x}_k,\ \dots$ из R^n, то

$$f_{\overline{\xi}}(\overline{t})=\sum_{k=1}^{\infty}e^{i(\overline{t},\ \overline{x}_k)}\mathrm{P}(\overline{\xi}=\overline{x}_k).$$

Свойства многомерных характеристических функций

1°. *При всех* $\overline{t}\in R^n$ $\left|f_{\overline{\xi}}(\overline{t})\right|\le 1$, $f_{\overline{\xi}}(\overline{0})=1$.

2°. *Многомерная характеристическая функция равномерно непрерывна в* R^n *по всем переменным.*

3°. *Если* $\overline{\xi}_1,\ \overline{\xi}_2,\ \dots,\ \overline{\xi}_n$ *– независимые случайные векторы, то*

$$f_{\overline{\xi}_1+\dots+\overline{\xi}_n}(\overline{t})=\prod_{k=1}^{n}f_{\overline{\xi}_k}(\overline{t}).$$

4°. *Характеристическая функция случайного вектора* $(\xi_1,\ \dots,\ \xi_m)$, $m<n$, *равна*

$$f_{(\xi_1,\ \dots,\ \xi_m)}(t_1,\ \dots,\ t_m)=f_{(\xi_1,\ \dots,\ \xi_n)}(t_1,\ \dots,\ t_m,\ 0,\ 0,\ \dots,\ 0).$$

5°. $f_{\xi_1+\dots+\xi_m}(t)=f_{(\xi_1,\ \dots,\ \xi_m)}(t,\ t,\ \dots,\ t)$, $t\in R$.

6°. *Случайные величины* $\xi_1,\ \dots,\ \xi_n$ *независимы тогда и только тогда, когда*

$$f_{(\xi_1,\ \dots,\ \xi_n)}(t_1,\ \dots,\ t_n)=\prod_{k=1}^{n}f_{\xi_k}(t_k)$$

для всех $\overline{t}=(t_1,\ \dots,\ t_n)\in R^n$.

7°. *Если* $\overline{\eta}=C\overline{\xi}$ *– линейное преобразование*

$$\eta_l=\sum_{k=1}^{n}C_{lk}\xi_k,\qquad l=\overline{1,m},$$

с матрицей $C=[C_{lk}]$, $l=\overline{1,m}$, $k=\overline{1,n}$, *то*

$$f_{\overline{\eta}}\left(\overline{t}\right)=f_{\overline{\xi}}\left(C^{*}\overline{t}\right),$$

где C^{*} *– сопряженная к* C *матрица, преобразующая вектор* $\overline{t}=\left(t_{1},\ \ldots,\ t_{m}\right)$ *по формуле*

$$\sum_{k=1}^{m}C_{lk}t_{l},\qquad l=\overline{1,n}.$$

З а м е ч а н и е 1. Если $m=n$, $\det C\neq 0$ и существует плотность $p_{\overline{\xi}}(\overline{x})$, то существует плотность распределения случайного вектора $\overline{\eta}=C\overline{\xi}$ и

$$p_{\overline{\eta}}(\overline{x})=\frac{1}{|C|}\,p_{\overline{\xi}}(C^{-1}\overline{x}).\tag{1}$$

Доказательство равенства (1) непосредственно вытекает из формулы замены переменных в n-мерном интеграле.

З а м е ч а н и е 2. Из свойств 3° и 7° следует, что если $\overline{\eta}=C\overline{\xi}+b$, то

$$f_{\overline{\eta}}\left(\overline{t}\right)=e^{i(\overline{a},\,\overline{t})}f_{\overline{\xi}}\left(C^{*}\overline{t}\right).$$

8°. $f_{\overline{\xi}}\left(-\overline{t}\right)=\overline{f_{\overline{\xi}}}\left(\overline{t}\right)=f_{-\overline{\xi}}\left(\overline{t}\right)$.

Введем в рассмотрение смешанные моменты, положив по определению

$$m_{\alpha_{1}\,\ldots\,\alpha_{n}}=\mathrm{M}\,\xi_{1}^{\alpha_{1}}\ldots\ \xi_{n}^{\alpha_{n}},$$

где α_{i} *– целые неотрицательные числа,* $i=\overline{1,n}$.

9°. *Если конечны все смешанные моменты* $m_{\alpha_{1}\,\ldots\,\alpha_{n}}$ *с* $\alpha_{1}+\ldots+\alpha_{n}=r$, *то*

$$m_{\alpha_{1}\,\ldots\,\alpha_{n}}=i^{\alpha}\,\frac{\partial^{\alpha}f_{\overline{\xi}}\left(0,\ \ldots,\ 0\right)}{\partial t_{1}^{\alpha_{1}}\ldots\ \partial t_{n}^{\alpha_{n}}},\qquad \alpha=\alpha_{1}+\ldots+\alpha_{n}\leq r,$$

и

$$f_{\overline{\xi}}\left(\overline{t}\right)=\sum_{\alpha=0}^{r}i^{\alpha}\sum_{\alpha=\alpha_{1}+\ldots+\alpha_{n}}\frac{t_{1}^{\alpha_{1}}\ldots\ t_{n}^{\alpha_{n}}}{\alpha_{1}!\ \ldots\ \alpha_{n}!}m_{\alpha_{1}\,\ldots\,\alpha_{n}}+R_{r}\left(\overline{t}\right),$$

где $R_{r}\left(\overline{t}\right)=o\left(\left|\,t\,\right|^{r}\right)$ *при* $\left|\,t\,\right|=\left|t_{1}\right|+\ldots+\left|t_{n}\right|\to 0.$

§ 4. Многомерное нормальное распределение и связанные с ним распределения

Будем говорить, что случайный вектор $\overline{\xi}=\left(\xi_{1},\ \ldots,\ \xi_{n}\right)$ имеет *нормальное или гауссовское распределение*, если характеристическая функция имеет вид

$$f_{\overline{\xi}}\left(\overline{t}\right)=e^{i\left(\overline{a},\,\overline{t}\right)-\frac{1}{2}\left(B\overline{t},\,\overline{t}\right)},\qquad \overline{t}\in R^{n},\tag{1}$$

где $\overline{a}=\left(a_{1},\ \ldots,\ a_{n}\right)\in R^{n}$, а $B=\left[b_{lk}\right]$ – симметричная $n\times n$-матрица, неотрицательно определенная, т. е. $\left(B\overline{t},\ \overline{t}\right)\geq 0$ для любых $\overline{t}\in R^{n}$. Случайный вектор с характеристической функцией вида (1) называется также $\left(\overline{a},\,B\right)$*-нормальным случайным вектором.*

Из равенства (1) и свойства 4° многомерной характеристической функции (§ 3) следует, что каждая компонента ξ_k, $k=\overline{1,n}$, нормального случайного вектора $\overline{\xi}$ имеет характеристическую функцию

$$f_{\xi_k}(t)=e^{ita_k-\frac{b_{kk}}{2}t^2}, \quad t\in R,$$

т. е. нормально распределена с $\mathrm{M}\,\xi_k=a_k$, $\mathrm{D}\,\xi_k=b_{kk}$.

Поскольку конечны все $\mathrm{M}\,\xi_k^2$, $k=\overline{1,n}$, то конечны и смешанные моменты $\mathrm{M}\,\xi_k\ \ldots\ \xi_l$, $k,\ \ldots,\ l=\overline{1,n}$, поэтому по свойству 9° многомерной характеристической функции имеем

$$\operatorname{cov}(\xi_k,\ \xi_l)=\mathrm{M}(\xi_k-a_k)(\xi_l-a_l)=-\frac{\partial^2 f_{\overline{\xi}_0}(\overline{0})}{\partial t_k\partial t_l}=b_{kl},$$

где $\overline{\xi}_0=\overline{\xi}-\overline{a}$.

Таким образом матрица $B=[b_{lk}]$ – это *ковариационная матрица* случайного вектора $\overline{\xi}=(\xi_1,\ \ldots,\ \xi_n)$.

Пусть ковариационная матрица B нормального случайного вектора $\overline{\xi}$ диагональна с одинаковыми диагональными элементами $b_{kk}=\sigma^2>0$ и $\mathrm{M}\,\overline{\xi}=\overline{0}$. Такое нормальное распределение называется *сферическим*. Несложно получить, что плотность $\overline{\xi}$ в этом случае имеет вид

$$p_{\overline{\xi}}(\overline{x})=\frac{1}{\sigma^n(2\pi)^{\frac{n}{2}}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}x_k^2\right\}. \qquad (2)$$

Из сферического нормального распределения получим вид плотностей нескольких стандартных распределений, имеющих большое значение в математической статистике и других приложениях теории вероятностей.

Хи-квадрат распределение (χ^2-распределение).

Рассмотрим случайную величину

$$\chi_n^2=\xi_1^2+\ldots+\xi_n^2,$$

где $\xi_1,\ \ldots,\ \xi_n$ – независимые нормально распределенные с параметрами $(0,\ 1)$ случайные величины. Найдем плотность распределения случайной величины χ_n. Вероятность события $x<\chi_n<x+dx$ можно получить из n-мерной сферической нормальной плотности (2) с $\sigma=1$, интегрируя ее по n-мерному сферическому слою радиуса x и толщины dx, в результате, поскольку $(n-1)$-мерный объем $(n-1)$-мерной сферы радиуса x пропорционален x^{n-1}, получим

$$p_{\chi_n}(x)=C_n\,x^{n-1}e^{-\frac{x^2}{2}}, \quad x\geq 0.$$

Для определения C_n воспользуемся тем, что

$$\int_0^\infty p_{\chi_n}(x) = 1,$$

откуда получим

$$C_n \int_0^\infty x^{n-1} e^{-\frac{x^2}{2}} dx = 2^{\frac{n}{2}-1} \Gamma\left(\frac{n}{2}\right) C_n = 1$$

и

$$p_{\chi_n}(x) = \begin{cases} 0, & x < 0, \\ \dfrac{x^{n-1}}{2^{\frac{n}{2}-1}\Gamma\left(\frac{n}{2}\right)} e^{-\frac{x^2}{2}}, & x \geq 0. \end{cases} \tag{3}$$

Иллюстрация в Mathematica

Используя

$$p_{\chi_n^2}(x) = \frac{1}{2\sqrt{x}} p_{\chi_n}\left(\sqrt{x}\right), \qquad x \geq 0,$$

имеем

$$p_{\chi_n^2}(x) = \begin{cases} 0, & x < 0, \\ \dfrac{x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} e^{-\frac{x}{2}}, & x \geq 0. \end{cases} \tag{4}$$

Иллюстрация в Mathematica

Распределение с плотностью (4) называется *хи-квадрат распределением (χ^2-распределением) с n степенями свободы*. Плотность (3) задает *хи-распределение (χ-распределение) с n степенями свободы*. При $n = 3$ выражение (3) называется *плотностью распределения Максвелла* и в кинетической теории газов дает распределение абсолютной величины скорости частиц.

Распределение Стьюдента. Пусть случайные величины $\xi_0, \xi_1, \ldots, \xi_n$ независимы и нормально распределены с параметрами $(0, 1)$. В статистике часто используют случайную величину

$$\tau_n = \frac{\xi_0}{\sqrt{\frac{1}{n}\sum_{k=1}^{n}\xi_k^2}},$$

называемую *отношением Стьюдента*. Распределение случайной величины τ_n, называемое *распределением Стьюдента с n степенями свободы*, имеет следующую плотность

$$s_n(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \qquad x \in R. \tag{5}$$

Иллюстрация в Mathematica

Формулу (5) можно вывести следующим образом. Пусть

$$\chi_n = \sqrt{\xi_1^2 + \ldots + \xi_n^2},$$

тогда

$$\tau_n = \frac{\xi_0 \sqrt{n}}{\chi_n}$$

есть отношение двух независимых случайных величин, распределение которых известно (числитель имеет нормальное распределение с параметрами $\left(0,\ \sigma = \sqrt{n}\right)$, а знаменатель – распределение (3)).

Функция распределения $S_n(x)$ случайной величины τ_n равна

$$S_n(x) = \mathrm{P}(\tau_n \le x) = \iint\limits_{\frac{u}{v} \le x,\, v>0} p_{\xi_0\sqrt{n}}(u)\, p_{\chi_n}(v)\, du dv = a_n \iint\limits_{\frac{u}{v} \le x,\, v>0} v^{n-1} e^{-\frac{1}{2}\left(\frac{u^2}{n} + v^2\right)} du\, dv,$$

где

$$a_n = \frac{1}{\sqrt{2n\pi}} \frac{1}{2^{\frac{n}{2}-1}\Gamma\left(\frac{n}{2}\right)}.$$

От переменных $(u,\ v)$ перейдем к новым переменным $(y,\ z)$ по формулам

$$u = yz, \qquad v = z. \tag{6}$$

Якобиан преобразования равен

$$\frac{\partial(u,v)}{\partial(y,z)} = z,$$

поэтому

$$\iint\limits_{\frac{u}{v} \le x,\, v>0} v^{n-1} e^{-\frac{1}{2}\left(\frac{u^2}{n} + v^2\right)} du dv = \int\limits_{-\infty}^{x} dy \int\limits_{0}^{\infty} z^n e^{-\frac{z^2}{2}\left(\frac{y^2}{n}+1\right)} dz =$$

$$= \int\limits_{-\infty}^{x} \left(\frac{y^2}{n} + 1\right)^{-\frac{n+1}{2}} dy \int\limits_{0}^{\infty} \omega^n e^{-\frac{\omega^2}{2}} d\omega = 2^{\frac{n-1}{2}} \Gamma\left(\frac{n+1}{2}\right) \int\limits_{-\infty}^{x} \left(\frac{y^2}{n} + 1\right)^{-\frac{n+1}{2}} dy,$$

откуда следует формула (5). Заметим, что при $n \to \infty$ плотность $s_n(x)$ сходится к нормальной плотности

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

F-распределение (распределение Фишера). Пусть $\xi_1,\ \ldots,\ \xi_m,\ \eta_1,\ \ldots,\ \eta_n$ – независимые нормальные случайные величины с параметрами $(0,\ 1)$. Обозначим

$$F_{mn}=\frac{\frac{1}{m}\sum_{k=1}^{m}\xi_k^2}{\frac{1}{n}\sum_{j=1}^{n}\eta_j^2}.$$

Распределение случайной величины F_{mn} имеет плотность

$$p_{F_{mn}}(x)=m^{\frac{m}{2}}n^{\frac{n}{2}}\frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\cdot\frac{x^{\frac{m}{2}-1}}{(n+mx)^{\frac{m+n}{2}}},\qquad x\ge 0, \tag{7}$$

($p_{F_{mn}}(x)=0$, если $x<0$) и называется *F-распределением (распределением Фишера, F-распределением Фишера).*

Иллюстрация в Mathematica

Для вывода (7) воспользуемся тем, что $\frac{m}{n}F_{mn}$ есть отношение двух независимых случайных величин, имеющих распределение χ^2 с m и n степенями свободы соответственно. Поэтому

$$\mathrm{P}\left(\frac{m}{n}F_{mn}\le x\right)=\frac{1}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\iint\limits_{\frac{u}{v}\le x,\ u\ge 0,\ v\ge 0} u^{\frac{m}{2}-1}v^{\frac{n}{2}-1}e^{-\frac{u+v}{2}}\,du\,dv.$$

Делая опять под знаком интеграла замену (6), получаем

$$\iint\limits_{\frac{u}{v}\le x,\ u\ge 0,\ u\ge 0} u^{\frac{m}{2}-1}v^{\frac{n}{2}-1}e^{-\frac{u+v}{2}}\,du\,dv=\int_0^x y^{\frac{m}{2}-1}dy\int_0^\infty z^{\frac{m+n}{2}-1}e^{-\frac{z(1+y)}{2}}dz=$$

$$=2^{\frac{m+n}{2}}\Gamma\left(\frac{m+n}{2}\right)\int_0^x\frac{y^{\frac{m}{2}-1}}{(1+y)^{\frac{m+n}{2}}}dy$$

и

$$\mathrm{P}\left(\frac{m}{n}F_{mn}\le x\right)=\frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\int_0^x\frac{y^{\frac{m}{2}-1}}{(1+y)^{\frac{m+n}{2}}}\,dy,$$

откуда уже несложно вывести (7).

З а м е ч а н и е 1. Функции χ^2-распределения (Приложения, табл. 6), распределения Стьюдента (Приложения, табл. 5) и F-распределения Фишера табулированы.

Задачи

1. Докажите, что функция, определяемая равенствами

$$f(t)=f(-t),\ f(t+2a)=f(t),\ f(t)=\frac{a-t}{a}$$

при $0 \le t \le a$ является характеристической.

2. Докажите, что можно найти такие независимые случайные величины ξ_1, ξ_2, ξ_3, что распределения ξ_2 и ξ_3 различны, а функции распределения сумм $\xi_1 + \xi_2$ и $\xi_1 + \xi_3$ одинаковы.

3. Докажите, что если $f(t)$, $t \in R$, является характеристической функцией, то функции

$$\varphi_1(t) = \begin{cases} f(t), & |t| \le a, \\ f(t+2a), & |t| > a; \end{cases}$$

$$\varphi_2(t) = \frac{1}{t}\int_0^t f(t)\,dt;$$

$$\varphi_3(t) = \exp\{f(t) - 1\}$$

также являются характеристическими.

4. Покажите, что для характеристической функции $f(t)$, $t \in R$, имеет место неравенство

$$1 - |f(2t)|^2 \le 4\left(1 - |f(t)|^2\right), \quad t \in R.$$

5. Докажите, что для любой вещественной характеристической функции $f(t)$, $t \in R$, справедливо неравенство

$$1 + f(2t) \ge 2\left[f(t)\right]^2.$$

6. Докажите, что если *F* – функция распределения и *f* – ее характеристическая функция, то при любом $x \in R$

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f(t)\, e^{-itx} dt = F(x) - F(x-0).$$

7. Докажите, что если *F* – функция распределения, *f* – ее характеристическая функция и x_k – абсциссы скачков *F*, то

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} |f(t)|^2 dt = \sum_k \left[F(x_k) - F(x_k - 0)\right].$$

8. Докажите, что если случайная величина имеет плотность распределения, то ее характеристическая функция при $t \to \infty$ стремится к 0.

9. Пусть f_1, f_2, f_3 – характеристические функции и $f_1 f_2 = f_1 f_3$. Следует ли отсюда, что $f_2 = f_3$?

10. Пусть f_1 и f_2 – характеристические функции, $0 \le \theta \le 1$. Покажите, что функция $(1-\theta) f_1 + \theta f_2$ является характеристической.

11. Докажите, что если $\mathrm{M}\xi = 0$, то

$$\mathrm{M}|\xi| = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1 - \operatorname{Re} f(t)}{t^2}\, dt,$$

где f – характеристическая функция ξ, а $\operatorname{Re} f$ – вещественная часть функции f. Если существует $\mathrm{D}\xi$, то

$$\mathrm{M}|\xi| = -\frac{2}{\pi}\int_{-\infty}^{\infty}\frac{1-\operatorname{Re} f'(t)}{t}\,dt.$$

12. Закон распределения $F(x)$, $x \in R$, называется *устойчивым*, если для любых a_1, b_1, a_2, b_2 найдутся такие a_3, b_3, что

$$F(a_1x+b_1) * F(a_2x+b_2) = F(a_3x+b_3),$$

где $*$ – операция свертки.

Выясните, какие из нижеприведенных законов принадлежат к устойчивому типу:

а) биномиальное распределение;

б) равномерное на отрезке $[a;b]$ распределение;

в) нормальное распределение;

г) экспоненциальное распределение;

д) распределение Пуассона.

13. Докажите, что при сложении независимых случайных величин третьи центральные моменты суммируются, а четвертые – нет.

14. Докажите, что из равенства

$$f_{(\xi,\,\eta)}(t,\,t) = f_\xi(t)\,f_\eta(t),\ t \in R,$$

не следует независимость случайных величин ξ и η.

15. Докажите, что из равенства

$$f_{\xi+\eta}(t) = f_\xi(t)\,f_\eta(t),\ t \in R,$$

не следует независимость случайных величин ξ и η.

16. Покажите, что из дифференцируемости в нуле характеристической функции случайной величины ξ не следует, что существует $\mathrm{M}\xi$.

Глава 6. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

На практике довольно часто встречается ситуация, когда многократно наблюдается «одна и та же» случайная величина, а конечный результат представляет собой сумму наблюдаемых в каждом испытании значений этих величин. Наряду с уже упоминавшимися азартными играми сюда можно отнести: повторные замеры одного и того же параметра с последующим усреднением результатов для повышения точности измерений, многократное воздействие однородных причин на некоторый протекающий во времени физический процесс и т. д.

В настоящей главе будет дано описание этих явлений с единых вероятностных позиций на основе следующей схемы: имеется последовательность независимых случайных величин, возможно одинаково распределенных. Нас будет интересовать поведение суммы первых n членов этой последовательности, в частности поведение среднего арифметического этих членов, если n велико. Оказывается при больших n среднее арифметическое случайных величин теряет свойство случайности и приближается к среднему арифметическому математических ожиданий. Этот факт называется законом больших чисел.

Уточнение закона больших чисел происходило в двух направлениях. Первое связано с динамикой поведения средних арифметических. К основным результатам этого направления следует отнести усиленный закон больших чисел и закон повторного логарифма. Исходным пунктом второго направления, называемого иногда центральной предельной проблемой, являются уже известные нам теоремы Муавра – Лапласа. Решение центральной предельной проблемы позволило описать класс всех распределений, которые могут выступать в качестве предельных для функций распределения сумм независимых случайных величин в том случае, когда вклад каждого слагаемого бесконечно мал по сравнению с вкладом суммы.

§ 1. Центральная предельная теорема

Настоящий параграф посвящен одному из самых замечательных результатов теории вероятностей: при широких условиях суммы большого числа независимых малых случайных слагаемых имеют распределение, близкое к нормальному (гауссовскому). Значение этого результата выходит далеко за рамки теории вероятностей. Он является основой применения нормального распределения при решении многих практических задач.

Пусть $(\Omega, \mathscr{A}, \mathrm{P})$ – произвольное вероятностное пространство, $\xi_1, \ldots, \xi_n, \ldots$ – последовательность случайных величин на $(\Omega, \mathscr{A}, \mathrm{P})$.

Определение 1. *Последовательность случайных величин* $\xi_1, \ldots, \xi_n, \ldots$ *удовлетворяет центральной предельной теореме, если для любого* $x \in R$

$$\mathrm{P}\left(\frac{\xi_1 + \ldots + \xi_n - \mathrm{M}(\xi_1 + \ldots + \xi_n)}{\sqrt{\mathrm{D}(\xi_1 + \ldots + \xi_n)}} \le x\right) \xrightarrow[n\to\infty]{} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Сначала приведем простейший вариант центральной предельной теоремы, относящийся к суммам одинаково распределенных слагаемых.

Теорема 1. *Пусть случайные величины* $\xi_1, \ldots, \xi_n, \ldots$ *независимы, одинаково распределены и имеют конечные математические ожидания* $\mathrm{M}\xi_n = a$ *и дисперсии* $\mathrm{D}\xi_n = \sigma^2 > 0$, $n \ge 1$, *тогда они удовлетворяют центральной предельной теореме, т. е. для любого* $x \in R$

$$\mathrm{P}\left(\frac{\xi_1 + \ldots + \xi_n - na}{\sigma\sqrt{n}} \le x\right) \xrightarrow[n\to\infty]{} \Phi(x).$$

Иллюстрация в Mathematica

Доказательство. Пусть по определению $\overline{\xi}_k = \xi_k - a$. Тогда $\mathrm{M}\overline{\xi}_k = 0$, $\mathrm{M}\overline{\xi}_k^2 = \sigma^2$.

По свойству 6° характеристической функции (гл. 5, § 1) имеем при $t \to 0$

$$f_{\overline{\xi}_k}(t) = 1 - \frac{\sigma^2 t^2}{2} + o(t^2). \tag{1}$$

Несложно видеть, что

$$S_n = \frac{\xi_1 + \ldots + \xi_n - na}{\sigma\sqrt{n}} = \sum_{k=1}^{n} \frac{\overline{\xi}_k}{\sigma\sqrt{n}}.$$

Поэтому характеристическая функция S_n имеет вид:

$$f_{S_n(t)} = \prod_{k=1}^{n} f_{\frac{\overline{\xi}_k}{\sigma\sqrt{n}}}(t) = \prod_{k=1}^{n} f_{\overline{\xi}_k}\left(\frac{t}{\sigma\sqrt{n}}\right) = \left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n,$$

и, следовательно, при $n \to \infty$ поточечно сходится к $e^{-t^2/2}$, но $e^{-t^2/2}$ – характеристическая функция нормальной случайной величины с параметрами $(0, 1)$. Таким образом, из обратной предельной теоремы для характеристических функций вытекает доказательство данной теоремы.

П р и м е р 1. Рассмотрим схему Бернулли с вероятностью успеха p в каждом испытании. Пусть μ_i – число успехов в i-м испытании, тогда

$$\mathrm{M}\mu_i = p, \quad \mathrm{D}\mu_i = p(1-p), \quad i = \overline{1,n}.$$

Обозначим

$$S_n = \mu_1 + \ldots + \mu_n.$$

По теореме 1 имеем для любого $x \in R$

$$\mathrm{P}\left(\frac{S_n - np}{\sqrt{np(1-p)}} \le x\right) \xrightarrow[n\to\infty]{} \Phi(x).$$

Это утверждение представляет собой не что иное, как интегральную теорему Муавра – Лапласа (см. дополнения к гл. 2).

П р и м е р 2. О ш и б к и и з м е р е н и я. При измерении некоторой величины a получаем приближенное значение ξ. Сделанная ошибка $\delta = \xi - a$ может быть представлена в виде суммы двух ошибок

$$\delta = (\xi - \mathrm{M}\xi) + (\mathrm{M}\xi - a),$$

первая из которых $\xi - \mathrm{M}\xi$ называется *случайной ошибкой*, а вторая $\mathrm{M}\xi - a$ – *систематической ошибкой*. Хорошие методы измерения не должны иметь систематической ошибки, поэтому далее будем полагать $\mathrm{M}\xi = a$ и, следовательно, $\mathrm{M}\xi - a = 0$. Пусть $\mathrm{D}\delta = \tau^2$. Для уменьшения случайной ошибки δ производят n независимых измерений $\xi_1, \xi_2, \ldots, \xi_n$ и принимают за значение измеряемой величины a среднее арифметическое

$$S_n = \frac{\xi_1 + \ldots + \xi_n}{n}.$$

Какая при этом допускается погрешность? По теореме 1

$$\mathrm{P}\left(\left|S_n - a\right| \le \varepsilon\right) \approx \frac{1}{\sqrt{2\pi}} \int\limits_{-\varepsilon\sqrt{n}/a}^{\varepsilon\sqrt{n}/a} e^{-\frac{t^2}{2}} dt.$$

Значения функции, стоящей в правой части, протабулированы (см. Приложения, табл. 2)

П р и м е р 3. Л о г а р и ф м и ч е с к и-н о р м а л ь н о е р а с п р е д е л е н и е. В антропологии обычно рост или вес человека определенного возраста и пола считают нормальной случайной величиной. Однако во многих случаях с гораздо большим основанием можно считать, что логарифмы этих параметров имеют нормальное распределение. Если случайная величина η такова, что $\xi = \log\eta$ имеет нормальное распределение, то говорят, что η имеет *логарифмически-нормальное распределение* или, короче *лог-нормальное распределение*. Лог-нормальности роста и веса можно дать теоретическое обоснование. Например, вес получается в результате воздействия многих независимых причин, которые воздействуют на вес мультипликативно, т. е.

$$\eta = \eta_1 \ldots \eta_n,$$

где η_i близкие к единице независимые случайные величины. В этом случае

$$\log \eta = \sum_{i=1}^{n} \log \eta_i$$

и $\log \eta$ в силу теоремы 1 имеет в пределе нормальное распределение.

П р и м е р 4. С помощью центральной предельной теоремы можно доказать и чисто аналитические факты. Докажем, например, что

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

В самом деле, пусть S_n есть случайная величина Пуассона с параметром n. Тогда

$$\mathrm{P}\left(S_n \le n\right) = e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!}.$$

Но $S_n = \xi_1 + \ldots + \xi_n$, где случайные величины $\xi_1, \ \ldots, \ \xi_n$ – независимые пуассоновские случайные величины, $\mathrm{M}\,\xi_k = 1$, $k = \overline{1, n}$. По теореме 1 имеем

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \lim_{n\to\infty} \mathrm{P}\left(S_n \le n\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{t^2}{2}} dt = \frac{1}{2}.$$

Центральная предельная теорема имеет место при некоторых условиях и для неодинаково распределенных независимых слагаемых.

Будем говорить, что для последовательности случайных величин $\{\xi_n\}$, $n \in N$, *выполнено условие Линдеберга*, если для произвольного $\tau > 0$

$$\frac{1}{\mathrm{B}_n^2} \sum_{k=1}^{n} \int\limits_{|x - a_k| > \tau \mathrm{B}_n} \left(x - a_k\right)^2 dF_k\left(x\right) \xrightarrow[n\to\infty]{} 0,$$

где $a_k = \mathrm{M}\,\xi_k$, $\mathrm{B}_n^2 = \sum_{k=1}^{n} \mathrm{D}\,\xi_k$, F_k – функция распределения случайной величины ξ_k.

Теорема 2 (Линдеберг). *Если для последовательности независимых случайных величин* $\{\xi_n\}$, $n \in N$, *выполнено условие Линдеберга, то для любого* $x \in R$

$$\mathrm{P}\left(\frac{\left(\xi_1 + \ldots + \xi_n - \left(a_1 + \ldots + a_n\right)\right)}{\mathrm{B}_n} \le x\right) \xrightarrow[n\to\infty]{} \Phi\left(x\right).$$

Схема доказательства

Д о к а з а т е л ь с т в о. Введем в рассмотрение следующие случайные величины

$$\xi_{kn} = \frac{\xi_k - a_k}{\mathrm{B}_n}, \ k = 1,\ 2,\ \ldots,\ n = 1,\ 2,\ \ldots .$$

Очевидно, что

$$\mathrm{M}\xi_{kn} = 0, \ \sum_{k=1}^{n} \mathrm{D}\,\xi_{kn} = \frac{1}{\mathrm{B}_n^2} \sum_{k=1}^{n} \mathrm{D}\,\xi_k = 1.$$

Рассмотрим характеристическую функцию случайной величины

$$S_n = \frac{\left(\xi_1 + \ldots + \xi_n - \left(a_1 + \ldots + a_n\right)\right)}{\mathrm{B}_n} = \sum_{k=1}^{n} \frac{\xi_k - a_k}{\mathrm{B}_n} = \sum_{k=1}^{n} \xi_{kn}.$$

Так как случайные величины ξ_{kn}, $k = 1,\ 2,\ \ldots,$ независимы, то

$$\varphi_{S_n}(t) = \prod_{k=1}^{n} \varphi_{\xi_{kn}}(t) = \prod_{k=1}^{n} \varphi_{kn}(t). \tag{2}$$

Прологарифмируем обе части (2)

$$\ln \varphi_{S_n}(t) = \sum_{k=1}^{n} \ln \varphi_{kn}(t).$$

Докажем, что

$$\ln \varphi_{S_n}(t) \xrightarrow[n\to\infty]{} -\frac{t^2}{2}.$$

Для этого сначала докажем, что

$$\ln \varphi_{kn}(t) \xrightarrow[n\to\infty]{} 0,$$

или

$$\varphi_{kn}(t) \xrightarrow[n\to\infty]{} 1,$$

равномерно относительно k $(1 \le k \le n)$ и t в любом конечном интервале $|t| \le T$. Несложно видеть, что

$$\left|\varphi_{kn}(t) - 1\right| = \left|\int_{-\infty}^{\infty} \left(e^{itx} - 1\right) dF_{kn}(x)\right| = \left|\int_{-\infty}^{\infty} \left(e^{itx} - 1 - itx\right) dF_{kn}(x)\right| \le$$

$$\le \int_{-\infty}^{\infty} \frac{(tx)^2}{2} dF_{kn}(x) = \frac{t^2}{2}\left(\int_{|x|\le\varepsilon} + \int_{|x|>\varepsilon}\right) x^2 dF_{kn}(x). \tag{3}$$

Для доказательства соотношений (3) использовались неравенство (4) (см. гл. 5, § 1) и то, что

$$\int_{-\infty}^{\infty} itx\, dF_{kn}(x) = \mathrm{M}it\xi_{kn} = 0.$$

Заметим, что

$$F_{kn}\left(\frac{x-a_k}{\mathrm{B}_n}\right)=\mathrm{P}\left(\xi_{kn}\le\frac{x-a_k}{\mathrm{B}_n}\right)=\mathrm{P}\left(\frac{\xi_k-a_k}{\mathrm{B}_n}\le\frac{x-a_k}{\mathrm{B}_n}\right)=\mathrm{P}\left(\xi_k\le x\right)=F_k\left(x\right).$$

Тогда

$$\frac{1}{\mathrm{B}_n^2}\sum_{k=1}^{n}\int\limits_{|x-a_k|>\tau B_n}\left(x-a_k\right)^2 dF_k\left(x\right)=\sum_{k=1}^{n}\int\limits_{\frac{|x-a_k|}{\mathrm{B}_n}>\tau}\left(\frac{x-a_k}{\mathrm{B}_n}\right)^2 dF_{kn}\left(\frac{x-a_k}{\mathrm{B}_n}\right)=$$

$$=\sum_{k=1}^{n}\int\limits_{|z|>\tau} z^2 dF_{kn}\left(z\right)\xrightarrow[n\to\infty]{}0$$

для любого $\tau>0$.

Правую часть неравенства (3) можно оценить сверху следующим образом

$$\frac{t^2}{2}\varepsilon^2+\frac{t^2}{2}\int\limits_{|z|>\varepsilon} z^2 dF_{kn}\left(z\right).$$

Таким образом, для всех достаточно больших n равномерно относительно k $(1\le k\le n)$ и t в любом конечном интервале $|t|\le T$

$$\left|\varphi_{kn}\left(t\right)-1\right|<\varepsilon^2 T^2.$$

Отсюда заключаем, что равномерно относительно $(1\le k\le n)$

$$\lim_{n\to\infty}\varphi_{kn}\left(t\right)=1 \tag{4}$$

и что при всех достаточно больших n для t, лежащих в произвольном конечном интервале $|t|\le T$, выполняется неравенство

$$\left|\varphi_{kn}\left(t\right)-1\right|<\frac{1}{2}. \tag{5}$$

Следовательно, в промежутке $|t|\le T$ можем записать разложение

$$\ln\varphi_{S_n}\left(t\right)=\sum_{k=1}^{n}\ln\varphi_{kn}\left(t\right)=\sum_{k=1}^{n}\ln\left(1+\left(\varphi_{kn}\left(t\right)-1\right)\right)=$$

$$=\sum_{k=1}^{n}\left(\varphi_{kn}\left(\left(t\right)-1\right)\right)+R_n, \tag{6}$$

где

$$R_n=\sum_{k=1}^{n}\sum_{s=2}^{\infty}\frac{\left(-1\right)^{s-1}}{s}\left(\varphi_{kn}\left(t\right)-1\right)^s.$$

В силу (5)

$$|R_n| \le \sum_{k=1}^{n}\sum_{s=2}^{\infty}\frac{|\varphi_{kn}(t)-1|^s}{2} \le \frac{1}{2}\sum_{k=1}^{n}\frac{|\varphi_{kn}(t)-1|^2}{1-|\varphi_{kn}(t)-1|} \le \sum_{k=1}^{n}|\varphi_{kn}(t)-1|^2,$$

но

$$\sum_{k=1}^{n}|\varphi_{kn}(t)-1| = \sum_{k=1}^{n}\left|\int_{-\infty}^{+\infty}\left(e^{itx}-1-itx\right)dF_{kn}(x)\right| \le \frac{t^2}{2}\sum_{k=1}^{n}\int_{-\infty}^{+\infty}x^2 dF_{kn}(x) = \frac{t^2}{2}.$$

Тогда

$$|R_n| \le \frac{t^2}{2}\max_{1\le k\le n}|\varphi_{kn}(t)-1|.$$

Из (5) следует, что равномерно относительно $t \in [-T;T]$ для любого $T>0$ при $n\to\infty$

$$R_n \to 0, \tag{7}$$

но

$$\sum_{k=1}^{n}\left(\varphi_{kn}(t)-1\right) = -\frac{t^2}{2}+\rho_n, \tag{8}$$

где

$$\rho_n = \frac{t^2}{2}+\sum_{k=1}^{n}\int_{-\infty}^{+\infty}\left(e^{itx}-1-itx\right)dF_{kn}(x).$$

Пусть $\varepsilon>0$ произвольно. Тогда с учетом того, что

$$\sum_{k=1}^{n}\mathrm{D}\xi_{kn}=1,$$

получим

$$\rho_n = \sum_{k=1}^{n}\int_{|x|\le\varepsilon}\left(e^{itx}-1-itx-\frac{(itx)^2}{2}\right)dF_{kn}(x)+$$

$$+\sum_{k=1}^{n}\int_{|x|>\varepsilon}\left(\frac{t^2x^2}{2}+e^{itx}-1-itx\right)dF_{kn}(x).$$

Неравенство (4) (см. гл. 5, § 1) позволяет получить следующую оценку:

$$|\rho_n| \le \frac{|t|^3}{6}\sum_{k=1}^{n}\int_{|x|\le\varepsilon}|x|^3 dF_{kn}(x)+t^2\sum_{k=1}^{n}\int_{|x|>\varepsilon}|x|^2 dF_{kn}(x) \le$$

$$\le \frac{|t|^3}{6}\varepsilon\sum_{k=1}^{n}\int_{|x|\le\varepsilon}x^2 dF_{kn}(x)+t^2\sum_{k=1}^{n}\int_{|x|>\varepsilon}x^2 dF_{kn}(x) =$$

$$= \frac{|t|^3}{6}\varepsilon + t^2\left(1 - \frac{|t|}{6}\varepsilon\right)\sum_{k=1}^{n}\int_{|x|>\varepsilon} x^2 dF_{kn}(x).$$

Согласно условию Линдеберга слагаемое

$$t^2\left(1 - \frac{|t|}{6}\varepsilon\right)\sum_{k=1}^{n}\int_{|x|>\varepsilon} x^2 dF_{kn}(x)$$

при любом $\varepsilon > 0$ может быть сделано меньше любого $\delta > 0$ для достаточно больших n. В силу произвольности ε выберем его таким, чтобы для всех $\delta > 0,\ T,\ t \in [-T;T]$ выполнялось

$$|\rho_n| < 2\delta \qquad \left(n \geq n_0(\varepsilon,\ \delta,\ T)\right).$$

Это неравенство показывает, что равномерно в каждом конечном интервале значений t

$$\lim_{n\to\infty}\rho_n = 0. \tag{9}$$

Из соотношений (6) – (9) следует, что равномерно на каждом промежутке $[-T;T]$

$$\lim_{n\to\infty}\ln\varphi_{S_n}(t) = -\frac{t^2}{2}.$$

Теорема доказана.

З а м е ч а н и е 1. Вместо равномерной сходимости на компактах последовательности характеристических функций в доказательстве можно рассматривать поточечную сходимость.

С л е д с т в и е 1. (**Теорема Ляпунова**). *Если для последовательности случайных величин* $\{\xi_n\},\ n \in N,$ *существует такое положительное число* δ, *что при* $n \to \infty$

$$\frac{1}{\mathrm{B}_n^{2+\delta}}\sum_{k=1}^{n}\mathrm{M}|\xi_k - a_k|^{2+\delta} \to 0 \ (\textit{условие Ляпунова}),$$

то для этой последовательности имеет место центральная предельная теорема.

Д о к а з а т е л ь с т в о. Убедимся, что из выполнения условия Ляпунова следует справедливость условия Линдеберга

$$\frac{1}{\mathrm{B}_n^2}\sum_{k=1}^{n}\int_{|x-a_k|>\tau\mathrm{B}_n}(x-a_k)^2 dF_k(x) \leq \frac{1}{\mathrm{B}_n^2(\tau\mathrm{B}_n)^\delta}\sum_{k=1}^{n}\int_{|x-a_k|>\tau\mathrm{B}_n}(x-a_k)^{2+\delta} dF_k(x) \leq$$

$$\leq \frac{1}{\tau^\delta}\frac{1}{\mathrm{B}_n^{2+\delta}}\sum_{k=1}^{n}\mathrm{M}|\xi_k - a_k|^{2+\delta} \xrightarrow[n\to\infty]{} 0.$$

У п р а ж н е н и е 1. Иногда теорему Ляпунова формулируют следующим образом: *пусть случайные величины* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *независимы, имеют конечные*

$$a_k = \mathrm{M}\xi_k,\ b_k^2 = \mathrm{M}\left(\xi_k - a_k\right)^2,\ c_k^3 = \mathrm{M}\left|\xi_k - a_k\right|^3$$

и

$$\frac{C_n}{\mathrm{B}_n} \xrightarrow[n\to\infty]{} 0,$$

где $C_n^3 = \sum_{k=1}^{n} c_k^3$, $\mathrm{B}_n^2 = \sum_{k=1}^{n} b_k^2$. *Тогда они удовлетворяют центральной предельной теореме.* Докажите, что и это утверждение является следствием теоремы 2.

У п р а ж н е н и е 2. Докажите, что теорема 1 есть следствие теоремы Линдеберга.

О п р е д е л е н и е 2. *Будем говорить, что случайные величины*

$$\xi_{kn} = \frac{\xi_k - a_k}{\mathrm{B}_n},\ k = \overline{1,n},\ n \ge 1,$$

пренебрежимо малы (равномерно сходятся к нулю), если для любого $\varepsilon > 0$ *при* $n \to \infty$

$$\max_{k\le n} \mathrm{P}\left(\left|\xi_{kn}\right| > \varepsilon\right) \to 0. \tag{10}$$

Условие Линдеберга не является необходимым для сходимости распределения

$$\sum_{k=1}^{n} \left(\xi_k - a_k\right)/\mathrm{B}_n$$

к нормальному $N(0,\ 1)$ распределению. Это показывает следующий пример.

П р и м е р 5. Пусть $\xi_{1n} = \eta,\ \xi_{2n} = 0,\ \ldots,\ \xi_{nn} = 0,\ \eta$ – нормально распределенная случайная величина с параметрами $(0,\ 1)$. Тогда

$$\mathrm{M}\,\xi_{kn} = 0, \quad \sum_{k=1}^{n} \mathrm{D}\,\xi_{kn} = 1, \quad \mathrm{P}\left(\sum_{k=1}^{n} \xi_{kn} \le x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt,$$

но условие Линдеберга не выполнено.

Однако, если потребовать, чтобы вместе со сходимостью

$$\mathrm{P}\left(\sum_{k=1}^{n} \xi_{kn} \le x\right) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

было выполнено условие пренебрежимой малости, то условие Линдеберга становится необходимым.

Теорема 3. *Если последовательность независимых случайных величин* $\{\xi_{kn}\}$, *удовлетворяет условию* (10) *и для любого* $x \in R$

$$\mathrm{P}\left(\sum_{k=1}^{n} \xi_{kn} \le x\right) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt,$$

то выполнено условие Линдеберга.

Д о к а з а т е л ь с т в о. Очевидны следующие неравенства

$$\left|\varphi_{kn}(t)-1\right| = \left|\mathrm{M}\left(e^{it\xi_{kn}}-1\right)\right| \le \int_{|x|>\varepsilon} \left|e^{itx}-1\right| dF_{kn}(x) + \int_{|x|\le\varepsilon} \left|e^{itx}-1\right| dF_{kn}(x) \le$$

$$\le 2\int_{|x|>\varepsilon} dF_{kn}(x) + \int_{|x|\le\varepsilon} |t\varepsilon| dF_{kn}(x) \le 2\mathrm{P}\left(\left|\xi_{kn}\right|>\varepsilon\right) + |t\varepsilon|.$$

В силу условия (10)

$$\max_{k\le n}\left|\varphi_{kn}(t)-1\right| \xrightarrow[n\to\infty]{} 0$$

для любого $t \in R$. Используя технику доказательства теоремы Линдеберга, имеем

$$\left|R_n\right| = \left|\ln\varphi_n(t) - \sum_{k=1}^{n}\left(\varphi_{kn}(t)-1\right)\right| \le \sum_{k=1}^{n}\left|\varphi_{kn}(t)-1\right|^2 \le \frac{t^2}{2}\max_{k\le n}\left|\varphi_{kn}(t)-1\right|.$$

Очевидно, что

$$\sum_{k=1}^{n}\left(\varphi_{kn}(t)-1\right) = \int_{-\infty}^{+\infty} \frac{e^{itx}-1-itx}{x^2} \sum_{k=1}^{n} x^2 dF_{kn}(x) = \int_{-\infty}^{+\infty} \frac{e^{itx}-1-itx}{x^2} dG_n(x),$$

где $G_n(x) = \sum_{k=1}^{n} \int_{t\le x} t^2 dF_{kn}(t)$.

Значит, из того, что

$$\ln\varphi_n(t) \to -\frac{t^2}{2},\ n\to\infty,$$

следует, что

$$\sum_{k=1}^{n}\left(\varphi_{kn}(t)-1\right) \xrightarrow[n\to\infty]{} -\frac{t^2}{2}$$

поточечно по $t \in R$. Поэтому

$$\int_{-\infty}^{+\infty} \frac{e^{itx}-1-itx}{x^2} dG_n(x) \xrightarrow[n\to\infty]{} -\frac{t^2}{2}$$

и

$$\int_{-\infty}^{+\infty} \left(\frac{e^{i1x}-1-i1x}{x^2}+\frac{1}{2}\right) dG_n(x) \xrightarrow[n\to\infty]{} 0,$$

так как G_n является функцией распределения.

Далее

$$\left|\frac{e^{ix}-1-ix}{x^2}\right| \le \frac{1}{2},$$

и это неравенство является строгим при $x \ne 0$. Поэтому, если $\tau > 0$, то

$$\sup_{|x|>\tau}\left|\frac{e^{ix}-1-ix}{x^2}\right| \le \frac{1}{2}-\delta(\tau),$$

где $\delta(\tau) > 0$. Значит,

$$\operatorname{Re}\left(\frac{e^{ix}-1-ix}{x^2}+\frac{1}{2}\right) \ge \delta(\tau) > 0$$

для всех x, $|x| > \tau$. Тогда

$$\int_{|x|>\tau} \delta(\tau)\, dG_n(x) \xrightarrow[n\to\infty]{} 0.$$

Очевидно, что

$$\int_{|x|>\tau} \delta(\tau)\, dE(x) = 0,$$

где $E(x)$ – функция вырожденного распределения, сосредоточенного в точке 0,

$$E(x) = \begin{cases} 1, & x \ge 0, \\ 0, & x < 0. \end{cases}$$

Тогда $G_n(x) \Rightarrow E(x)$, что равносильно выполнению условия Линдеберга.

Будем говорить, что для последовательности ξ_k, $k \ge 1$, выполняется *условие равномерной малости*, если

$$\lim_{n\to\infty} \max_{1\le k\le n} \frac{\sigma_k}{\mathrm{B}_n} = 0, \tag{11}$$

где $\sigma_k = \mathrm{D}\xi_k$.

У п р а ж н е н и е 3. Докажите, что теорема 3 будет справедлива, если условие (10) заменить на (11).

У п р а ж н е н и е 4. Докажите, что условие (10) в теореме 3 можно заменить на следующие:

$$\frac{\sigma_n}{\mathrm{B}_n} \to 0,\ \mathrm{B}_n \to \infty\ (n \to \infty).$$

§ 2. Сходимость случайных величин

В этом параграфе рассмотрены различные виды сходимости случайных величин и установлена связь между ними.

1. Сходимость почти наверное (с вероятностью единица)

О п р е д е л е н и е 1. *Будем говорить, что последовательность случайных величин* $\{\xi_n\}$, $n \in N$, *сходится почти наверное (с вероятностью единица) к случайной величине* ξ *и обозначать*

$$\xi_n \xrightarrow[n\to\infty]{п.н.} \xi,$$

если

$$\mathrm{P}\left\{\omega \in \Omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right\} = 1.$$

Приведем еще одно определение сходимости почти наверное, эквивалентное определению 1. Несложно видеть, что справедливо следующее равенство

$$\left\{\omega \in \Omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right\} = \bigcap_{r=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \left\{\omega \in \Omega : \left|\xi_m(\omega) - \xi(\omega)\right| \le 1/r\right\}.$$

Предположим, что

$$\mathrm{P}\left\{\omega : \lim_{n\to\infty} \xi_n(\omega) \ne \xi(\omega)\right\} = 0.$$

Это эквивалентно тому, что

$$\mathrm{P}\left(\bigcup_{r=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \left\{\left|\xi_m(\omega) - \xi(\omega)\right| > 1/r\right\}\right) = 0.$$

Последнее равенство эквивалентно тому, что для любых $r > 0$

$$\mathrm{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \left\{\left|\xi_m - \xi\right| > 1/r\right\}\right) = 0,$$

или

$$\lim_{n\to\infty} \mathrm{P}\left(\bigcup_{m=n}^{\infty} \left\{\left|\xi_m - \xi\right| > 1/r\right\}\right) = 0,$$

что равносильно

$$\lim_{n\to\infty} \mathrm{P}\left(\sup_{m \ge n} \left|\xi_m - \xi\right| > 1/r\right) = 0$$

для любого $r > 0$.

Из всего вышесказанного следует

Утверждение 1. *Последовательность случайных величин* $\{\xi_n\}$, $n \in N$, *сходится почти наверное (с вероятностью единица) к случайной величине* ξ *тогда и только тогда, когда для любого* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}\left(\omega \in \Omega : \sup_{m \ge n} \left| \xi_m(\omega) - \xi(\omega) \right| \le \varepsilon \right) = 1.$$

2. Сходимость по вероятности

Определение 2. *Будем говорить, что последовательность случайных величин* $\{\xi_n\}$, $n \in N$, *сходится по вероятности к случайной величине* ξ *и обозначать*

$$\xi_n \xrightarrow[n\to\infty]{\mathrm{P}} \xi,$$

если для любого $\varepsilon > 0$

$$\mathrm{P}\left(\omega \in \Omega : \left| \xi_n(\omega) - \xi(\omega) \right| \le \varepsilon \right) \xrightarrow[n\to\infty]{} 1.$$

Утверждение 2. *Если*

$$\xi_n \xrightarrow[n\to\infty]{\text{п.н.}} \xi,$$

то

$$\xi_n \xrightarrow[n\to\infty]{\mathrm{P}} \xi.$$

Доказательство. Утверждение 2 вытекает из утверждения 1.

Упражнение 1. Докажите, что, если

$$\xi_n \xrightarrow[n\to\infty]{\mathrm{P}} \xi,$$

то существует подпоследовательность $\{\xi_{n_k}\}$, такая что

$$\xi_{n_k} \xrightarrow[k\to\infty]{\text{п.н.}} \xi.$$

Упражнение 2. Докажите, что последовательность случайных величин $\{\xi_n\}$ сходится по вероятности тогда и только тогда, когда для любого $\varepsilon > 0$

$$\mathrm{P}\left(\left| \xi_n - \xi_m \right| > \varepsilon \right) \to 0, \qquad n, m \to \infty.$$

3. Слабая сходимость

Определение 3. *Говорят, что последовательность случайных величин* $\{\xi_n\}$, $n \in N$, *слабо сходится к случайной величине* ξ *и обозначают* $\xi_n \Rightarrow \xi$, *если последовательность функций распределения* F_{ξ_n} *слабо сходится к* F_ξ *при* $n \to \infty$.

Утверждение 3. *Если*

$$\xi_n \xrightarrow[n\to\infty]{\mathrm{P}} \xi,$$

то

$$\xi_n \underset{n\to\infty}{\Rightarrow} \xi.$$

Д о к а з а т е л ь с т в о. Нужно доказать, что для любого x, являющегося точкой непрерывности функции F_ξ, при $n\to\infty$

$$F_{\xi_n}(x) \to F_\xi(x).$$

Введем в рассмотрение множество

$$A_n = \{\omega\in\Omega:\ |\xi_n(\omega)-\xi(\omega)|\le\varepsilon\}.$$

Из того, что

$$\xi_n \xrightarrow[n\to\infty]{\mathrm{P}} \xi$$

следует

$$\mathrm{P}(\overline{A}_n) \xrightarrow[n\to\infty]{} 0.$$

Очевидны следующие включения:

$$\{\xi_n\le x\}\subseteq\{\xi\le x+\varepsilon\}\cup\overline{A}_n,\quad \{\xi\le x-\varepsilon\}\subseteq\{\xi_n\le x\}\cup\overline{A}_n,$$

поэтому

$$\mathrm{P}(\xi_n\le x)\le \mathrm{P}(\xi\le x+\varepsilon)+\mathrm{P}(\overline{A}_n),\quad \mathrm{P}(\xi\le x-\varepsilon)-\mathrm{P}(\overline{A}_n)\le \mathrm{P}(\xi_n\le x).$$

Переходя к пределу при $n\to\infty$, получаем

$$F_\xi(x-\varepsilon)\le \varliminf_{n\to\infty}\mathrm{P}(\xi_n\le x)\le \varlimsup_{n\to\infty}\mathrm{P}(\xi_n\le x)\le F_\xi(x+\varepsilon).$$

В силу того, что x – точка непрерывности функции F_ξ, а ε сколь угодно мало, можно сделать вывод, что утверждение 3 доказано.

У т в е р ж д е н и е 4. *Если*

$$\xi_n \underset{n\to\infty}{\Rightarrow} \xi$$

и $\mathrm{P}(\xi=c)=1$, *то*

$$\xi_n \xrightarrow[n\to\infty]{\mathrm{P}} \xi.$$

Д о к а з а т е л ь с т в о. Пусть x – точка непрерывности F_ξ (т. е. $x\ne c$). Тогда

$$F_{\xi_n}(x) \xrightarrow[n\to\infty]{} F_\xi(x) = \begin{cases} 0, & x<c, \\ 1, & x\ge c. \end{cases}$$

Но для любого $\varepsilon>0$

$$\mathrm{P}(|\xi_n-\xi|\le\varepsilon) = \mathrm{P}(\xi_n\le\xi+\varepsilon)-\mathrm{P}(\xi_n<\xi-\varepsilon)=$$
$$=\mathrm{P}(\xi_n\le c+\varepsilon)-\mathrm{P}(\xi_n<c-\varepsilon)\to F_\xi(c+\varepsilon)-F_\xi(c-\varepsilon)=1.$$

Утверждение 4 доказано.

4. Сходимость в среднем порядка *r*

О п р е д е л е н и е 4. *Говорят, что последовательность случайных величин* $\{\xi_n\}$, $n \in N$, *сходится в среднем порядка* r $(r > 0)$ *к случайной величине* ξ *и обозначают*

$$\xi_n \xrightarrow[n\to\infty]{r} \xi,$$

если

$$\mathrm{M}\,|\,\xi_n - \xi\,|^r \xrightarrow[n\to\infty]{} 0.$$

У т в е р ж д е н и е 5. *Если*

$$\xi_n \xrightarrow[n\to\infty]{r} \xi,$$

то

$$\xi_n \xrightarrow[n\to\infty]{\mathrm{P}} \xi.$$

Д о к а з а т е л ь с т в о вытекает из неравенства Чебышева:

$$\mathrm{P}\big(|\,\xi_n - \xi\,| > \varepsilon\big) = \mathrm{P}\Big(|\,\xi_n - \xi\,|^r > \varepsilon^r\Big) \le \frac{\mathrm{M}\,|\,\xi_n - \xi\,|^r}{\varepsilon^r} \xrightarrow[n\to\infty]{} 0,$$

т. е.

$$\xi_n \xrightarrow{\mathrm{P}} \xi$$

при $n \to \infty$.

У п р а ж н е н и е 3. Докажите, что последовательность случайных величин $\{\xi_n\}$, $n \in N$, для которых $\mathrm{M}\xi_n < \infty$, $n \in N$, сходится в среднем порядка r тогда и только тогда, когда $\mathrm{M}\,|\,\xi_n - \xi_m\,|^r \to 0$, $m,\ n \to \infty$.

Таким образом, мы установили соотношения между разными видами сходимости случайных величин (см. диаграмму).

$$\boxed{\xi_n \xrightarrow{п.н.} \xi,\ \ n \to \infty} \Rightarrow \boxed{\xi_n \xrightarrow{P} \xi,\ \ n \to \infty} \Leftarrow \boxed{\xi_n \xrightarrow{r} \xi,\ \ n \to \infty}$$

$$\Downarrow$$

$$\boxed{\xi_n \Rightarrow \xi,\ \ n \to \infty}$$

Покажем, что в данной диаграмме обратные направления, вообще говоря, не имеют места.

П р и м е р 1. Пусть $\Omega = [0;1]$, $\mathscr{A} = \mathscr{B}\big([0;1]\big)$, P – мера Лебега. Положим

$$A_n^i = \left[\frac{i-1}{n};\frac{i}{n}\right], \qquad \xi_n^i(\omega) = I_{A_n^i}(\omega), \qquad i = \overline{1,n}, \qquad n \in N.$$

Тогда последовательность случайных величин

$$\xi_1^1,\ \xi_2^1,\ \xi_2^2,\ \xi_3^1,\ \xi_3^2,\ \xi_3^3,\ \ldots$$

сходится и по вероятности, и в среднем порядка $r>0$, но не сходится ни в одной точке $\omega\in[0;1]$.

П р и м е р 2. Пусть $(\Omega,\mathscr{A},\mathrm{P})$ такое же, как и в прим. 1, а

$$\xi_n(\omega)=\begin{cases} e^n, & 0\le\omega\le 1/n,\\ 0, & \omega>1/n.\end{cases}$$

Тогда последовательность $\{\xi_n\}$, $n\in N$, сходится почти наверное (и, следовательно, по вероятности и слабо) к нулю, однако для любого $r>0$

$$\mathrm{M}\,|\,\xi_n\,|^r=e^{nr}/n\to\infty,\qquad n\to\infty.$$

П р и м е р 3. Пусть $\{\xi_n\}$, $n\in N$, – последовательность независимых случайных величин и $\mathrm{P}(\xi_n=1)=p_n$, $\mathrm{P}(\xi_n=0)=1-p_n$. Тогда

$$\xi_n\xrightarrow{\mathrm{P}}0\Leftrightarrow p_n\to 0,\qquad n\to\infty;$$

$$\xi_n\xrightarrow{r}0\Leftrightarrow p_n\to 0,\qquad n\to\infty;$$

$$\xi_n\xrightarrow{п.н.}0\Leftrightarrow\sum_{n=1}^{\infty}p_n<\infty.$$

В частности, при $p_n=1/n$, последовательность $\{\xi_n\}$, $n\in N$, сходится в среднем порядка r для любого $r>0$, но не сходится почти наверное.

§ 3. Законы больших чисел

В этом параграфе исследовано предельное поведение при $n\to\infty$ среднего арифметического n случайных величин. В зависимости от вида сходимости, по вероятности или почти наверное, материал разбит на два пункта.

1. Закон больших чисел

Пусть на произвольном вероятностном пространстве $(\Omega,\mathscr{A},\mathrm{P})$ задана последовательность случайных величин $\{\xi_n\}$, $n\in N$.

О п р е д е л е н и е 1. *Говорят, что последовательность случайных величин $\{\xi_n\}$, $n\in N$, удовлетворяет закону больших чисел, если*

$$\frac{\xi_1+\ldots+\xi_n}{n}-\frac{\mathrm{M}\xi_1+\ldots+\mathrm{M}\xi_n}{n}\xrightarrow[n\to\infty]{\mathrm{P}}0,\tag{1}$$

т. е. для любого $\varepsilon>0$

$$\mathrm{P}\left(\left|\frac{\xi_1+\ldots+\xi_n}{n}-\frac{\mathrm{M}\xi_1+\ldots+\mathrm{M}\xi_n}{n}\right|>\varepsilon\right)\xrightarrow[n\to\infty]{}0.$$

Иллюстрация в Mathematica

Теорема 1. *Для того, чтобы последовательность случайных величин удовлетворяла закону больших чисел, необходимо и достаточно, чтобы*

$$\mathrm{M}\frac{\left(\sum_{k=1}^{n}(\xi_k-\mathrm{M}\xi_k)\right)^2}{n^2+\left(\sum_{k=1}^{n}(\xi_k-\mathrm{M}\xi_k)\right)^2}\xrightarrow[n\to\infty]{}0. \tag{2}$$

Доказательство. Необходимость. Обозначим

$$\eta_n=\frac{1}{n}\sum_{k=1}^{n}(\xi_k-\mathrm{M}\xi_k).$$

Пусть выполняется условие (1), т. е.

$$\eta_n\xrightarrow{\mathrm{P}}0$$

при $n\to\infty$. Тогда для любого $\varepsilon>0$

$$\mathrm{M}\frac{\eta_n^2}{1+\eta_n^2}=\mathrm{M}\frac{\eta_n^2}{1+\eta_n^2}I_{(|\eta_n|\le\varepsilon)}+\mathrm{M}\frac{\eta_n^2}{1+\eta_n^2}I_{(|\eta_n|>\varepsilon)}\le\varepsilon^2+\mathrm{M}I_{(|\eta_n|>\varepsilon)}=$$

$$=\varepsilon^2+\mathrm{P}\left(|\eta_n|>\varepsilon\right)\xrightarrow[n\to\infty]{}\varepsilon^2,$$

отсюда получаем справедливость (2).

Достаточность. Пусть выполняется условие (2), т. е.

$$\mathrm{M}\frac{\eta_n^2}{1+\eta_n^2}\xrightarrow[n\to\infty]{}0,$$

тогда

$$\mathrm{P}\left(|\eta_n|>\varepsilon\right)=\mathrm{M}I_{(|\eta_n|>\varepsilon)}=\mathrm{M}\frac{\eta_n^2}{1+\eta_n^2}\frac{1+\eta_n^2}{\eta_n^2}I_{(|\eta_n|>\varepsilon)}\le\frac{1+\varepsilon^2}{\varepsilon^2}\mathrm{M}\frac{\eta_n^2}{1+\eta_n^2}I_{(|\eta_n|>\varepsilon)}\le$$

$$\le\frac{1+\varepsilon^2}{\varepsilon^2}\mathrm{M}\frac{\eta_n^2}{1+\eta_n^2}\xrightarrow[n\to\infty]{}0.$$

Теорема 1 доказана.

Теорема 2 (**Марков**). *Если*

$$\frac{1}{n^2}\mathrm{D}\sum_{k=1}^{n}\xi_k\xrightarrow[n\to\infty]{}0, \tag{3}$$

то последовательность случайных величин $\{\xi_n\}$, $n\in N$, *удовлетворяет закону больших чисел.*

Доказательство. Данная теорема – следствие теоремы 1. Так как

$$\frac{1}{n^2}\mathrm{D}\sum_{k=1}^{n}\xi_k=\frac{1}{n^2}\mathrm{M}\left(\sum_{k=1}^{n}\left(\xi_k-\mathrm{M}\xi_k\right)\right)^2\ge\mathrm{M}\frac{\left(\sum_{k=1}^{n}\left(\xi_k-\mathrm{M}\xi_k\right)\right)^2}{n^2+\left(\sum_{k=1}^{n}\left(\xi_k-\mathrm{M}\xi_k\right)\right)^2},$$

то из справедливости условия (3) вытекает справедливость условия (2). Таким образом, теорема 2 доказана.

Теорема 3 (**Чебышев**). *Пусть случайные величины* $\xi_1,\xi_2,\ldots,\xi_n,\ldots$ *независимы и существует константа* $C>0$ *такая, что* $\mathrm{D}\xi_n\le C$ *для всех* $n=1,\ 2,\ \ldots$. *Тогда они удовлетворяют закону больших чисел.*

Доказательство непосредственно вытекает из неравенства Чебышева:

$$\mathrm{P}\left(\left|\frac{\xi_1+\ldots+\xi_n}{n}-\frac{\mathrm{M}\xi_1+\ldots+\mathrm{M}\xi_n}{n}\right|>\varepsilon\right)\le\frac{\mathrm{D}\left(\xi_1+\ldots+\xi_n\right)}{\varepsilon^2n^2}=$$

$$=\frac{\mathrm{D}\xi_1+\ldots+\mathrm{D}\xi_n}{\varepsilon^2n^2}\le\frac{Cn}{\varepsilon^2n^2}=\frac{C}{\varepsilon^2n}\xrightarrow[n\to\infty]{}0$$

для любого $\varepsilon>0$.

Замечание 1. Для справедливости теоремы достаточно, чтобы

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{k=1}^{n}\mathrm{D}\xi_k=0.$$

Это будет выполнено, если

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{D}\xi_n=0,$$

так как на основании теоремы Штольца[1]

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{k=1}^{n}\mathrm{D}\xi_k=\lim_{n\to\infty}\frac{\mathrm{D}\xi_n}{n^2-(n-1)^2}=\lim_{n\to\infty}\frac{\mathrm{D}\xi_n}{2n-1}=0.$$

Упражнение 1. Докажите, что теорема 3 является следствием теоремы 1.

Теорема 4 (**Хинчин**). *Пусть случайные величины последовательности* $\{\xi_n\}$, $n\in N$, *независимы, одинаково распределены и имеют конечные* $\mathrm{M}\xi_n=a$, $n\in N$. *Тогда они удовлетворяют закону больших чисел, т. е.*

[1] Фихтенгольц Г.М. Курс дифференциального и интегрального исчисления. Т.3. М. 1969. С. 67

$$\frac{\xi_1 + \ldots + \xi_n}{n} \xrightarrow[n\to\infty]{\mathrm{P}} a. \qquad (4)$$

Доказательство проведем методом характеристических функций. Для этого введем случайную величину $\overline{\xi}_n = \xi_n - a$, тогда $\mathrm{M}\,\overline{\xi}_n = 0$, $n \in N$, и по свойству 6° характеристических функций (гл. 5, § 1) $f_{\overline{\xi}_n}(t) = 1 + o(t)$ при $t \to 0$.

Обозначим

$$S_n = \sum_{k=1}^{n} \frac{\overline{\xi}_k}{n} \xrightarrow[n\to\infty]{\mathrm{P}} 0. \qquad (5)$$

Очевидно, что условие (4) равносильно условию (5).

Несложно видеть, что

$$f_{S_n}(t) = \prod_{k=1}^{n} f_{\frac{\overline{\xi}_k}{n}}(t) = \prod_{k=1}^{n} f_{\overline{\xi}_k}\left(\frac{t}{n}\right) = \left(1 + o\left(\frac{t}{n}\right)\right)^n \xrightarrow[n\to\infty]{} 1,$$

но $f(t) \equiv 1$, так как это характеристическая функция вырожденной случайной величины ξ: $\mathrm{P}(\xi = 0) = 1$.

По обратной предельной теореме получим, что $S_n \underset{n\to\infty}{\Rightarrow} \xi$.

В силу утверждения 4 § 2, последнее утверждение эквивалентно (4). Теорема доказана.

Теорема 5 (**Бернулли**). *Пусть* μ_n *– число успехов в* n *испытаниях схемы Бернулли с вероятностью* p *успеха в каждом испытании, тогда*

$$\frac{\mu_n}{n} \xrightarrow[n\to\infty]{\mathrm{P}} p.$$

Доказательство вытекает из теоремы Хинчина, если рассмотреть независимые одинаково распределенные случайные величины ξ_i, $i = 1, 2, \ldots$, где $\mathrm{P}(\xi_i = 0) = 1 - p$, $\mathrm{P}(\xi_i = 1) = p$. Тогда

$$\mu_n = \sum_{i=1}^{n} \xi_i, \quad \mathrm{M}\xi_i = p$$

и теорема доказана.

2. **Усиленный закон больших чисел**

Определение 2. *Говорят, что последовательность случайных величин* $\{\xi_n\}$, $n \in N$, *удовлетворяет усиленному закону больших чисел, если*

$$\frac{\xi_1+\ldots+\xi_n}{n}-\frac{M\xi_1+\ldots+M\xi_n}{n}\xrightarrow[n\to\infty]{п.н.}0.$$

Теорема 6 (Неравенство Гаека – Реньи). *Если случайные величины* $\xi_1,\ \xi_2,\ \ldots,\ \xi_n,\ \ldots$ *являются независимыми,* $\mathrm{M}\xi_k=a_k$, $\mathrm{D}\xi_k=\sigma_k^2$, $k=1,\ 2,\ \ldots,$ *а* $C_1,\ C_2,\ \ldots$ *– невозрастающая последовательность неотрицательных чисел, то для любого* $\varepsilon>0$ *и для всех* $m,n\in N,\ m<n$, *выполнено*

$$\mathrm{P}\left(\max_{m\le k\le n}C_k\left|\sum_{i=1}^{k}(\xi_i-a_i)\right|>\varepsilon\right)\le\frac{1}{\varepsilon^2}\left(C_m^2\sum_{k=1}^{m}\sigma_k^2+\sum_{k=m+1}^{n}C_k^2\sigma_k^2\right).$$

Доказательство. Введем следующие обозначения

$$S_k=\sum_{i=1}^{k}(\xi_i-a_i),\qquad \eta=\sum_{k=m}^{n-1}S_k^2\left(C_k^2-C_{k+1}^2\right)+S_n^2C_n^2.$$

Найдем математическое ожидание η и преобразуем его к удобному виду

$$\mathrm{M}\eta=\sum_{k=m}^{n-1}\left(C_k^2-C_{k+1}^2\right)\mathrm{M}S_k^2+C_n^2\ \mathrm{M}S_n^2=\sum_{k=m}^{n-1}\sum_{i=1}^{k}\sigma_i^2\left(C_k^2-C_{k+1}^2\right)+C_n^2\sum_{i=1}^{n}\sigma_i^2=$$

$$=\sum_{i=1}^{m}\sum_{k=m}^{n-1}\sigma_i^2\left(C_k^2-C_{k+1}^2\right)+\sum_{i=m+1}^{n-1}\sum_{k=i}^{n-1}\sigma_i^2\left(C_k^2-C_{k+1}^2\right)+C_n^2\sum_{i=1}^{n}\sigma_i^2=\sum_{i=1}^{m}\sigma_i^2\left(C_m^2-C_n^2\right)+$$

$$+\sum_{i=m+1}^{n-1}\sigma_i^2\left(C_i^2-C_n^2\right)+C_n^2\sum_{i=1}^{n}\sigma_i^2=C_m^2\sum_{i=1}^{m}\sigma_i^2+\sum_{i=m+1}^{n}\sigma_i^2C_i^2.$$

Рассмотрим следующие случайные события для некоторого $\varepsilon>0$

$$A_i=\left\{\omega\in\Omega:\ C_k\left|S_k(\omega)\right|\le\varepsilon,\ m\le k\le i-1,\ \ C_i\left|S_i(\omega)\right|>\varepsilon\right\},\quad i=\overline{m,n}.$$

События $A_i,\ i=\overline{m,n}$, являются несовместными. Значит,

$$\mathrm{P}\left(\max_{m\le k\le n}C_k\left|\sum_{i=1}^{k}(\xi_i-a_i)\right|>\varepsilon\right)=\mathrm{P}\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=m}^{n}\mathrm{P}(A_i).$$

Теорема будет доказана, если будет установлено неравенство

$$\mathrm{M}\eta\ge\varepsilon^2\sum_{i=m}^{n}\mathrm{P}(A_i).$$

Докажем его:

$$\mathrm{M}\eta\ge\mathrm{M}\eta\sum_{i=m}^{n}I_{A_i}=\sum_{i=m}^{n}\mathrm{M}\eta I_{A_i},$$

$$\mathrm{M}\eta I_{A_i}=\sum_{k=m}^{n-1}\left(C_k^2-C_{k+1}^2\right)\mathrm{M}S_k^2I_{A_i}+C_n^2\,\mathrm{M}S_n^2I_{A_i},$$

$$\mathrm{M} S_k^2 I_{A_i} = \mathrm{M}\left(S_k - S_i + S_i\right)^2 I_{A_i} \ge \mathrm{M} S_i^2 I_{A_i} + 2\mathrm{M}\left(S_k - S_i\right) S_i I_{A_i} =$$

$$= \mathrm{M} S_i^2 I_{A_i} + 2\mathrm{M}\left(S_k - S_i\right)\mathrm{M} S_i I_{A_i} = \mathrm{M} S_i^2 I_{A_i} \ge \mathrm{M}\frac{\varepsilon^2}{C_i^2} I_{A_i} = \frac{\varepsilon^2}{C_i^2}\mathrm{P}\left(A_i\right).$$

С л е д с т в и е 1 (**Неравенство Колмогорова**). *Если случайные величины* $\xi_1, \ldots, \xi_n, \ldots$ *независимы и имеют конечные математические ожидания и дисперсии, то*

$$\mathrm{P}\left(\max_{1\le k\le n}\frac{1}{k}\left|\sum_{i=1}^{k}\left(\xi_i - \mathrm{M}\xi_i\right)\right| > \varepsilon\right) \le \frac{1}{\varepsilon^2}\sum_{k=1}^{n}\frac{\mathrm{D}\xi_k}{k^2}.$$

Д о к а з а т е л ь с т в о вытекает из неравенства Гаека – Реньи, если

$$C_k = 1/k, \quad m = 1.$$

Это неравенство можно записать в виде

$$\mathrm{P}\left(\max_{1\le k\le n}\left|\sum_{i=1}^{k}\left(\xi_i - \mathrm{M}\xi_i\right)\right| > \varepsilon\right) \le \frac{1}{\varepsilon^2}\sum_{k=1}^{n}\mathrm{D}\xi_k.$$

Теорема 7. *Пусть* $\xi_1, \ldots, \xi_n, \ldots$ – *независимые случайные величины,*

$$\mathrm{M}\xi_n = 0,\ \mathrm{D}\xi_n = \sigma_n^2,\ \sum_{n=1}^{\infty}\frac{\sigma_n^2}{n^2} < \infty.$$

Тогда

$$\frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n} \xrightarrow[n\to\infty]{\text{п.н.}} 0,$$

т. е. для этой последовательности справедлив усиленный закон больших чисел.

Д о к а з а т е л ь с т в о. Пусть $S_n = \xi_1 + \xi_2 + \ldots + \xi_n$, тогда сходимость

$$\frac{S_n}{n} \xrightarrow[n\to\infty]{\text{п.н.}} 0$$

равносильна тому, что для любого $\varepsilon > 0$

$$\mathrm{P}\left\{\sup_{k\ge n}\left|\frac{S_k}{k}\right| > \varepsilon\right\} \xrightarrow[n\to\infty]{} 0. \tag{6}$$

Введем обозначение

$$A_n = \left\{\max_{2^{n-1}\le k<2^n}\left|\frac{S_k}{k}\right| > \varepsilon\right\}.$$

Тогда (6) будет равносильно

$$\mathrm{P}\left(\bigcup_{i=n}^{\infty} A_i\right) \xrightarrow[n\to\infty]{} 0.$$

Используем неравенство Колмогорова

$$\mathrm{P}(A_n) = \mathrm{P}\left(\max_{2^{n-1}\le k<2^n}\left|\frac{S_k}{k}\right| > \varepsilon\right) \le \mathrm{P}\left(\max_{2^{n-1}\le k\le 2^n}|S_k| > \varepsilon 2^{n-1}\right) \le$$

$$\le \mathrm{P}\left(\max_{1\le k\le 2^n}|S_k| > \varepsilon 2^{n-1}\right) \le 4\frac{\mathrm{D}S_{2^n}}{\varepsilon^2 2^{2n}},$$

далее

$$\sum_{k=1}^{\infty}\mathrm{P}(A_k) \le 4\varepsilon^{-2}\sum_{k=1}^{\infty}2^{-2k}\sum_{n=1}^{2^k}\sigma_n^2 \le 4\varepsilon^{-2}\sum_{n=1}^{\infty}\sigma_n^2\sum_{\{k:2^k\ge n\}}2^{-2k} =$$

$$= 4\varepsilon^{-2}\sum_{n=1}^{\infty}\sigma_n^2\sum_{\{k:2^k\ge n\}}\frac{1}{\left(2^k\right)^2} \le 8\varepsilon^{-2}\sum_{n=1}^{\infty}\frac{\sigma_n^2}{n^2} < \infty,$$

потому что

$$\sum_{k=k_0}^{\infty} 2^{-2k} \le 2\cdot 2^{-2k_0}.$$

Из этого следует, что ряд

$$\sum_{k=1}^{\infty}\mathrm{P}(A_k)$$

сходится, а, значит,

$$\mathrm{P}\left(\bigcup_{k=n}^{\infty}A_k\right) \le \sum_{k=n}^{\infty}\mathrm{P}(A_k) \xrightarrow[n\to\infty]{} 0,$$

что равносильно (6).

Теорема доказана.

Лемма 1. *Математическое ожидание случайной величины* ξ *конечно тогда и только тогда, когда*

$$\sum_{n=1}^{\infty}\mathrm{P}\{|\xi| > n\} < \infty.$$

Доказательство. Из конечности $\mathrm{M}\xi$ следует конечность $\mathrm{M}|\xi|$, и наоборот. Очевидно следующее неравенство

$$\sum_{n=1}^{\infty}(n-1)I_{\{n-1<|\xi|\le n\}} \le |\xi| I_{\{\xi=0\}} + \sum_{n=1}^{\infty}|\xi| I_{\{n-1<|\xi|\le n\}} \le \sum_{n=1}^{\infty} nI_{\{n-1<|\xi|\le n\}},$$

поэтому

$$\sum_{n=1}^{\infty}(n-1)\mathrm{P}\left(n-1<|\xi|\le n\right)\le \mathrm{M}|\xi|\le \sum_{n=1}^{\infty} n\mathrm{P}\left(n-1<|\xi|\le n\right), \tag{7}$$

но

$$\sum_{n=1}^{\infty} n\mathrm{P}\left(n-1<|\xi|\le n\right)=\sum_{n=0}^{\infty}\mathrm{P}\left(|\xi|>n\right)\le 1+\sum_{n=1}^{\infty}\mathrm{P}\left(|\xi|>n\right), \tag{8}$$

$$\sum_{n=1}^{\infty}(n-1)\mathrm{P}\left(n-1<|\xi|\le n\right)=\sum_{n=1}^{\infty} n\mathrm{P}\left(n-1<|\xi|\le n\right)-$$

$$-\mathrm{P}\left(|\xi|>0\right)=\sum_{n=1}^{\infty}\mathrm{P}\left(|\xi|>n\right). \tag{9}$$

Из (7) – (9) следуют неравенства

$$\sum_{n=1}^{\infty}\mathrm{P}\left(|\xi|>n\right)\le \mathrm{M}|\xi|\le 1+\sum_{n=1}^{\infty}\mathrm{P}\left(|\xi|>n\right),$$

из которых вытекает доказательство леммы 1.

Теорема 8 (**Колмогоров**). *Пусть* $\{\xi_n\}$, $n\in N$, – *последовательность независимых одинаково распределенных случайных величин. Для выполнения усиленного закона больших чисел, т. е. для того, чтобы*

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k \xrightarrow[n\to\infty]{п.н.} a,$$

необходимо и достаточно существование конечного математического ожидания $\mathrm{M}\xi_k=a,\ k\in N$.

Иллюстрация в Mathematica

Схема доказательства

Доказательство. Достаточность. Введем случайные величины

$$\overline{\xi}_n=\xi_n I_{\{|\xi_n|\le n\}},\quad \overline{S}_n=\overline{\xi}_1+\overline{\xi}_2+\ldots+\overline{\xi}_n.$$

Случайные величины $\overline{\xi}_1,\ \overline{\xi}_2,\ \ldots,\ \overline{\xi}_n,\ \ldots$ являются независимыми, так как $x\,I_{\{|x|\le n\}}$ является борелевской функцией для каждого n.

Рассмотрим равенство

$$\frac{S_n-na}{n}=\frac{S_n-\overline{S}_n}{n}+\frac{\overline{S}_n-\mathrm{M}\overline{S}_n}{n}+\left(\frac{\mathrm{M}\overline{S}_n}{n}-a\right)=J_1^n+J_2^n+J_3^n,$$

где $S_n=\xi_1+\xi_2+\ldots+\xi_n$.

Для доказательства достаточности покажем, что все три слагаемых почти наверное сходятся к нулю. Очевидно, что

$$J_3^n=\frac{1}{n}\mathrm{M}\sum_{k=1}^{n}\xi_k I_{\{|\xi_k|\le k\}}-a=\frac{1}{n}\mathrm{M}\sum_{k=1}^{n}\xi_k I_{\{|\xi_k|\le k\}}-$$

$$-\frac{1}{n}\sum_{k=1}^{n}\mathrm{M}\left(\xi_k\left(I_{\{|\xi_k|\le k\}}+I_{\{|\xi_k|>k\}}\right)\right)=-\frac{1}{n}\sum_{k=1}^{n}\mathrm{M}\left(\xi_k I_{\{|\xi_k|>k\}}\right),$$

но

$$\mathrm{M}\left(\xi_k I_{\{|\xi_k|>k\}}\right)\xrightarrow[k\to\infty]{}0.$$

Из теоремы Штольца следует, что $J_3^n \xrightarrow[n\to\infty]{} 0.$

Обозначим $A_n=\left\{\xi_n\ne\overline{\xi}_n\right\}$. Получим, что

$$\sum_{n=1}^{\infty}\mathrm{P}(A_n)=\sum_{n=1}^{\infty}\mathrm{P}\left(|\xi_n|>n\right)=\sum_{n=1}^{\infty}\mathrm{P}\left(|\xi_1|>n\right)<\infty$$

по предыдущей лемме, так как $\mathrm{M}\,\xi_n<\infty$ для каждого n.

Далее

$$0\le\mathrm{P}\left(A^*\right)=\mathrm{P}\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}A_m\right)=\lim_{n\to\infty}\mathrm{P}\left(\bigcup_{m=n}^{\infty}A_m\right)\le$$

$$\le\lim_{n\to\infty}\sum_{m=n}^{\infty}\mathrm{P}(A_m)=0,$$

поэтому $\mathrm{P}\left(A^*\right)=0$, т. е. лишь для конечного числа номеров n

$$\xi_n\ne\overline{\xi}_n.$$

Значит,

$$J_1^n\xrightarrow[n\to\infty]{п.н.}0.$$

Покажем, что

$$J_2^n=\frac{\overline{S_n}-\mathrm{M}\,\overline{S_n}}{n}\xrightarrow[n\to\infty]{п.н.}0.$$

Для этого воспользуемся теоремой 7, дающей достаточное условие для справедливости усиленного закона больших чисел. Будем доказывать, что

$$\sum_{n=1}^{\infty}\frac{\mathrm{D}\,\overline{\xi}_n}{n^2}<\infty.$$

Поскольку

$$\mathrm{D}\,\overline{\xi}_n\le\mathrm{M}\overline{\xi}_n^2\le\sum_{k=1}^{n}k^2\mathrm{P}\left(k-1<|\xi_n|\le k\right),$$

то

$$\sum_{n=1}^{\infty}\frac{\mathrm{D}\,\overline{\xi}_n}{n^2}\le\sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{k^2}{n^2}\mathrm{P}\left(k-1<|\xi_1|\le k\right)=\sum_{k=1}^{\infty}k^2\mathrm{P}\left(k-1<|\xi_1|\le k\right)\sum_{n\ge k}\frac{1}{n^2}.$$

Очевидно, что

$$\sum_{n>k}\frac{1}{n^2}\le\int_k^\infty\frac{dx}{x^2}=\frac{1}{k}\qquad\text{и}\qquad\sum_{n\ge k}\frac{1}{n^2}\le\frac{1}{k}+\frac{1}{k^2}=\frac{k+1}{k^2},$$

поэтому

$$\sum_{n=1}^\infty\frac{\mathrm{D}\overline{\xi}_n}{n^2}\le\sum_{k=1}^\infty\frac{k^2(k+1)}{k^2}\mathrm{P}\left(k-1<|\xi_1|\le k\right)\le\sum_{k=1}^\infty(k+1)\mathrm{P}\left(k-1<|\xi_1|\le k\right)\le$$

$$\le 2+\sum_{k=1}^\infty(k-1)\mathrm{P}\left(k-1<|\xi_1|\le k\right)\le 2+\mathrm{M}|\xi_1|<\infty.$$

Необходимость. Если же

$$\frac{S_n}{n}\xrightarrow[n\to\infty]{п.н.}a,$$

то

$$\frac{\xi_n}{n}=\frac{S_n}{n}-\frac{n-1}{n}\frac{S_{n-1}}{n-1}\xrightarrow[n\to\infty]{п.н.}0,$$

т. е. почти наверное происходит конечное число событий

$$\left\{\omega\in\Omega:\ \left|\frac{\xi_n}{n}\right|>1\right\}.$$

Покажем, что

$$\sum_{n=1}^\infty\mathrm{P}\left\{|\xi_n|>n\right\}=\sum_{n=1}^\infty\mathrm{P}\left\{|\xi_1|>n\right\}<\infty.$$

Предположим, что

$$\sum_{n=1}^\infty\mathrm{P}\left\{|\xi_n|>n\right\}=\infty.$$

Обозначим B^* случайное событие, заключающееся в том, что происходит бесконечное число событий

$$\left\{\omega\in\Omega:\ \left|\frac{\xi_n}{n}\right|>1\right\},$$

и, воспользовавшись независимостью случайных величин, получим

$$\mathrm{P}\left(B^*\right)=\lim_{n\to\infty}\mathrm{P}\left(\bigcup_{m=n}^\infty\left(\frac{|\xi_m|}{m}>1\right)\right)=$$

$$=1-\lim_{n\to\infty}\mathrm{P}\left(\bigcap_{m=n}^\infty\left(\frac{|\xi_m|}{m}\le 1\right)\right)=1-\lim_{n\to\infty}\lim_{k\to\infty}\mathrm{P}\left(\bigcap_{m=n}^k\left(\frac{|\xi_m|}{m}\le 1\right)\right)=$$

$$=1-\lim_{n\to\infty}\lim_{k\to\infty}\prod_{m=n}^{k}\left(1-\mathrm{P}\left(\frac{|\xi_m|}{m}>1\right)\right)=1-\lim_{n\to\infty}\prod_{m=n}^{\infty}\left(1-\mathrm{P}\left(\frac{|\xi_m|}{m}>1\right)\right)=1.$$

Тогда $\mathrm{M}\xi_1<+\infty$.

Теорема доказана.

С л е д с т в и е 2 (**Теорема Бореля**). *В n испытаниях схемы Бернулли с вероятностью p успеха в каждом испытании для числа успехов μ_n имеет место усиленный закон больших чисел, т. е.*

$$\frac{\mu_n}{n}\xrightarrow[n\to\infty]{п.н.}p.$$

П р и м е р 1. П о л и н о м ы Б е р н ш т е й н а. Закон больших чисел используется для доказательства известной из курса математического анализа теоремы Вейерштрасса о равномерном приближении непрерывной функции многочленами.

Проводятся независимые испытания, в каждом из которых «успех» наступает с вероятностью x, а противоположное событие – с вероятностью $1-x$ ($0<x<1$). Пусть μ_n – количество «успехов» в n испытаниях, $f\in C[a;b]$. Известно, что

$$\mathrm{P}(\mu_n=k)=C_n^k x^k(1-x)^{n-k},\quad k=\overline{0,n},$$

поэтому

$$B_n(x)=\mathrm{M}f\left(\frac{\mu_n}{n}\right)=\sum_{k=0}^{n}f\left(\frac{k}{n}\right)C_n^k x^k(1-x)^{n-k}. \qquad (10)$$

Многочлен $B_n(x)$ называется *полиномом Бернштейна* для функции $f(x)$. Теорема Бернулли утверждает, что

$$\frac{\mu_n}{n}\xrightarrow[n\to\infty]{\mathrm{P}}x.$$

Естественно ожидать, что

$$\mathrm{M}f\left(\frac{\mu_n}{n}\right)\xrightarrow[n\to\infty]{}f(x).$$

Теорема (Бернштейн). *Последовательность многочленов*

$$\{B_n(x),\ n\in N\},$$

определенных равенством (10)*, сходится к непрерывной функции f равномерно относительно $x\in[0;1]$.*

Д о к а з а т е л ь с т в о. Так как f – непрерывна на $[0;1]$, то f – равномерно непрерывна на $[0;1]$, т. е. для любого $\varepsilon>0$ найдется $\delta(\varepsilon)$ такое, что

$$|f(x_1)-f(x_2)|<\frac{\varepsilon}{2},$$

если только $|x_1 - x_2| < \delta(\varepsilon)$. Функция f – ограничена на $[0;1]$, поэтому найдется C, что $f(x) \le C$ для всех $x \in [0;1]$. Заметим, что

$$\sum_{k=0}^{n} C_n^k x^k (1-x)^{n-k} = 1,$$

поэтому

$$B_n(x) - f(x) = \sum_{k=0}^{m} \left(f\left(\frac{k}{n}\right) - f(x) \right) C_n^k x^k (1-x)^{n-k}.$$

Далее получим

$$\left|B_n(x) - f(x)\right| = \sum_{k=0}^{m} \left| f\left(\frac{k}{n}\right) - f(x) \right| C_n^k x^k (1-x)^{n-k} \le$$

$$\le \sum_{k:\left|\frac{k}{n}-x\right|<\delta} \left| f\left(\frac{k}{n}\right) - f(x) \right| C_n^k x^k (1-x)^{n-k} +$$

$$+ \sum_{k:\left|\frac{k}{n}-x\right|\ge\delta} \left| f\left(\frac{k}{n}\right) - f(x) \right| C_n^k x^k (1-x)^{n-k} \le$$

$$\le \frac{\varepsilon}{2} + 2c \sum_{k:\left|\frac{k}{n}-x\right|\ge\delta} C_n^k x^k (1-x)^{n-k} = \frac{\varepsilon}{2} + 2c\,\mathrm{P}\left(\left| \frac{\mu_n}{n} - x \right| \ge \delta \right).$$

Из неравенства Чебышева следует, что

$$\mathrm{P}\left(\left| \frac{\mu_n}{n} - x \right| \ge \delta \right) \le \frac{\mathrm{D}\left(\frac{\mu_n}{n}\right)}{\delta^2} = \frac{x(1-x)}{n\delta^2} \le \frac{1}{4n\delta^2}$$

($x(1-x) \le 1/4$ для всех $0 < x < 1$).

Пусть $N(\delta)$ таково, что

$$\frac{1}{4N(\delta)\delta^2} < \frac{\varepsilon}{2},$$

тогда для всех $n \ge N(\delta)$

$$\left|B_n(x) - f(x)\right| < \varepsilon$$

для всех $x \in [0,1]$.

Что и требовалось доказать.

П р и м е р 2. М е т о д М о н т е-К а р л о. Вычислим интеграл

$$\int_0^1 g(x)\,dx$$

для некоторой непрерывной функции g. Пусть $\{\xi_n\}$, $n \in N$, – последовательность независимых равномерно распределенных на $[0;1]$ случайных величин. Заметим, что

$$\mathrm{M}\,g(\xi_n) = \int_{-\infty}^{+\infty} g(x)\,p_{\xi_n}(x)\,dx = \int_0^1 g(x)\,dx.$$

Можно утверждать, что с вероятностью 1

$$\frac{g(\xi_1)+g(\xi_2)+\ldots+g(\xi_n)}{n}\xrightarrow[n\to\infty]{}\mathrm{M}\,g(\xi_n)=\int_0^1 g(x)\,dx.$$

Таким образом, теоремы о законе больших чисел дают теоретическое обоснование алгоритма для приближенного подсчета интегралов

$$\int_0^1 g(x)\,dx.$$

1. моделируем последовательность $\{\xi_n\}$, $n\in N$, независимых равномерно распределенных на отрезке $[0;1]$ случайных величин.

2. полагаем

$$\int_0^1 g(x)\,dx\approx\frac{g(\xi_1)+g(\xi_2)+\ldots+g(\xi_n)}{n}.$$

Описанный метод называется *методом статистических испытаний вычисления интегралов (методом Монте-Карло)*.

Метод Монте-Карло особенно эффективен при вычислении интегралов большой кратности; для вычисления таких интегралов другие методы приближенного анализа непригодны.

П р и м е р 3. П р и м е н е н и е у с и л е н н о г о з а к о н а б о л ь ш и х ч и с е л к т е о р и и ч и с е л. Пусть $\Omega=[0;1)$, $\mathscr{A}=\mathscr{B}([0;1))$, P – мера Лебега. Рассмотрим двоичное разложение $\omega=0,\ \omega_1,\ \omega_2,\ \ldots$ чисел $\omega\in\Omega$ (с бесконечным количеством нулей) и определим случайные величины $\xi_1(\omega),\ \xi_2(\omega),\ \ldots$, полагая $\xi_n(\omega)=\omega_n$. Поскольку для любого $n\in N$ и любых $x_1,\ \ldots,\ x_n$, принимающих значения 0 или 1,

$$(\omega:\ \xi_1(\omega)=x_1,\ \ldots,\ \xi_n(\omega)=x_n)=\left(\omega:\ \frac{x_1}{2}+\frac{x_2}{2^2}+\ldots+\frac{x_n}{2^n}\le\omega\le\frac{x_1}{2}+\frac{x_2}{2^2}+\ldots+\frac{x_n}{2^n}+\frac{1}{2^n}\right),$$

то вероятность этого события равна $1/2^n$. Из этого вытекает, что $\xi_1,\ \xi_2,\ \ldots$ – последовательность независимых одинаково распределенных случайных величин, таких, что

$$\mathrm{P}(\xi_n=0)=\mathrm{P}(\xi_1=0)=1/2,\qquad n\in N.$$

Отсюда и из усиленного закона больших чисел следует следующий результат Бореля: *почти все числа полуинтервала* $[0;1)$ *нормальны в том смысле, что с вероятностью единица доля нулей и единиц в их двоичном разложении стремиться к* $1/2$, *т. е.*

$$\frac{1}{n}\sum_{k=1}^{n}I_{\{\xi_k=1\}}\xrightarrow[n\to\infty]{\text{п.н.}}\frac{1}{2}.$$

ДОПОЛНЕНИЯ К ГЛАВЕ 6

§ 1. Сходимость рядов

Цель настоящего параграфа – дать критерии, позволяющие определять, сходится или расходится ряд независимых случайных величин.

Для дальнейшего изложения материала напомним некоторые сведения из гл. 1.

Пусть $\{A_n\}$, $n \in N$, – последовательность случайных событий на вероятностном пространстве $(\Omega, \mathscr{A}, \mathrm{P})$. Обозначим через A^* множество всех тех и только тех элементарных событий, которые принадлежат бесконечному числу множеств A_n. Тогда

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

Множество A^* можно интерпретировать как событие, состоящее в том, что произойдет бесконечно много событий из последовательности $\{A_n\}_{n=1}^{\infty}$.

Л е м м а 1 (**Борель – Кантелли**)

1) *Если*

$$\sum_{n=1}^{\infty} \mathrm{P}(A_n) < \infty,$$

то

$$\mathrm{P}(A^*) = 0;$$

2) *Если события* $A_1, A_2, \ldots, A_n, \ldots$ *независимы и*

$$\sum_{n=1}^{\infty} \mathrm{P}(A_n) = \infty,$$

то

$$\mathrm{P}(A^*) = 1.$$

Д о к а з а т е л ь с т в о. 1) Обозначим

$$C_n = \bigcup_{m=n}^{\infty} A_m.$$

Тогда $C_1 \supseteq C_2 \supseteq \ldots$ и по упражнению 5 гл.1, §1

$$C_n \downarrow \bigcap_{k=1}^{\infty} C_k, \ n \to \infty.$$

Поэтому

$$\lim_{n\to\infty} \mathrm{P}\left(\bigcup_{m=n}^{\infty} A_m\right) \le \lim_{n\to\infty} \sum_{m=n}^{\infty} \mathrm{P}(A_m) = 0$$

в силу сходимости ряда, но по свойству непрерывности вероятности

$$\mathrm{P}(A^*)=\mathrm{P}\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}A_m\right)=\mathrm{P}\left(\bigcap_{n=1}^{\infty}C_n\right)=\lim_{n\to\infty}\mathrm{P}(C_n)=0.$$

2) Воспользуемся независимостью $A_1, \ldots, A_m, \ldots$, а, значит, и $\overline{A}_1, \ldots, \overline{A}_m, \ldots$. Получим

$$\mathrm{P}(A^*)=\mathrm{P}\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}A_m\right)=1-\mathrm{P}\left(\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty}\overline{A}_m\right)=$$

$$=1-\lim_{n\to\infty}\prod_{m=n}^{\infty}\mathrm{P}(\overline{A}_m)=1-\lim_{n\to\infty}\prod_{m=n}^{\infty}\left(1-\mathrm{P}(A_m)\right)..$$

Очевидно, что $1-\mathrm{P}(A_m)\le e^{-\mathrm{P}(A_m)}$, значит,

$$\mathrm{P}(A^*)\ge 1-\lim_{n\to\infty}e^{-\sum\limits_{m=n}^{\infty}\mathrm{P}(A_m)},$$

но

$$-\sum_{m=n}^{\infty}\mathrm{P}(A_m)=-\infty,$$

тогда $\mathrm{P}(A^*)\ge 1$, т. е. $\mathrm{P}(A^*)=1$.

С л е д с т в и е 1. *Если случайные события* $A_1, A_2, \ldots, A_n, \ldots$ *независимы, то* $\mathrm{P}(A^*)$ *равно* 0 *или* 1 *в зависимости от того, сходится или расходится ряд*

$$\sum_{n=1}^{\infty}\mathrm{P}(A_n).$$

Теперь рассмотрим более общий закон «нуля или единицы» Колмогорова. Пусть на вероятностном пространстве $(\Omega,\mathscr{A},\mathrm{P})$ определена последовательность $\xi_1, \xi_2, \ldots$ независимых случайных величин. Определим σ-алгебру $\mathscr{A}_{\xi_1\xi_2\ldots\xi_n}$, порожденную случайными величинами $\xi_1, \xi_2, \ldots, \xi_n$, как σ-алгебру всех событий A, представимых в виде

$$A=\left\{\omega\in\Omega:\ \left(\xi_1(\omega),\ \xi_2(\omega),\ \ldots,\ \xi_n(\omega)\right)\in B\right\}$$

для некоторого борелевского множества $B\in\mathscr{B}(R^n)$. Аналогично определяются $\mathscr{A}_{\xi_n}\subset\mathscr{A}_{\xi_n\xi_{n+1}}\subset\ldots$. Минимальную σ-алгебру, содержащую $\mathscr{A}_{\xi_n}$, $\mathscr{A}_{\xi_n\xi_{n+1}}$, ... (т. е. σ-алгебру, порожденную $\mathscr{A}_{\xi_n}$, $\mathscr{A}_{\xi_n\xi_{n+1}}$, …), обозначим $\mathscr{A}_{\xi_n\xi_{n+1}}\ldots$.

Рассмотрим последовательность $\mathscr{A}_{\xi_n\xi_{n+1}\ldots}$, $\mathscr{A}_{\xi_{n+1}\xi_{n+2}\ldots}$, …. Она является последовательностью невозрастающих σ-алгебр. Рассмотрим

$$\mathscr{F}=\bigcap_{n=1}^{\infty}\mathscr{A}_{\xi_n\xi_{n+1}\ldots}$$

Эта σ-алгебра называется *остаточной σ-алгеброй последовательности* $\{\xi_n(\omega)\}$, $n\in N$. События $A\in\mathscr{F}$ также называются остаточными. Это название отражает тот факт, что $A\in\mathscr{F}$ независимо с любым числом конечных случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ и определяется лишь «бесконечно далекими» значениями последовательности $\xi_1, \xi_2, \ldots$. Примерами остаточных событий являются следующие события:

$$\left\{\omega \in \Omega : \text{последовательность } \left\{\xi_n(\omega)\right\}_{n=1}^{\infty} \text{ сходится}\right\},$$

$$\left\{\omega \in \Omega : \sum_{n=1}^{\infty} \xi_n(\omega) < +\infty\right\}.$$

Теорема 1 (**Закон «нуля или единицы» Колмогорова**). *Если* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *– независимые случайные величины, то всякое остаточное событие* $A \in \mathscr{F}$ *имеет вероятность* $\mathrm{P}(A)$, *равную* 0 *или* 1.

Доказательство. Пусть $A \in \mathscr{F}$. Значит, для всех $n \in N$ $A \in \mathscr{A}_{\xi_n \xi_{n+1} \ldots}$. Из независимости случайных величин $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ следует, что $\mathscr{A}_{\xi_1 \ldots \xi_{n-1}}$ и $\mathscr{A}_{\xi_n \xi_{n+1} \ldots}$ являются независимыми σ-алгебрами, т. е. для любого $B \in \mathscr{A}_{\xi_1 \ldots \xi_n}$

$$\mathrm{P}(AB) = \mathrm{P}(A)\mathrm{P}(B).$$

И это выполняется для всех $n \in N$. Тогда A не зависит от любого $B \in \mathscr{A}_{\xi_1 \xi_2 \ldots}$, но $\mathscr{F} \subset \mathscr{A}_{\xi_1 \xi_2 \ldots}$, значит, A не зависит от самого себя, т. е. $\mathrm{P}(A) = 0$ или $\mathrm{P}(A) = 1$. Что и требовалось доказать.

Таким образом, из этой теоремы следует, что для независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ ряд

$$\sum_{k=1}^{\infty} \xi_k(\omega)$$

либо с вероятностью 1 сходится, либо с вероятностью 1 расходится. То же самое можно сказать и о самих последовательностях.

Пусть $\xi_1, \ldots, \xi_n, \ldots$ – последовательность независимых случайных величин, определенных на вероятностном пространстве $(\Omega, \mathscr{A}, \mathrm{P})$,

$$S_n = \xi_1 + \ldots + \xi_n$$

и A – множество тех элементарных исходов ω, где ряд $\sum \xi_n(\omega)$ сходится к конечному пределу. Из закона «0 или 1» Колмогорова следует, что $\mathrm{P}(A) = 0$ или $\mathrm{P}(A) = 1$, т. е. с вероятностью единица ряд $\sum \xi_n(\omega)$ сходится или расходится.

Докажем критерии, позволяющие определять, сходится или расходится ряд из независимых случайных величин.

Начнем изложение с одного вспомогательного утверждения, которое представляет и самостоятельный интерес.

Лемма 2 (**Неравенство Колмогорова**). *Пусть* $\xi_1, \ldots, \xi_n$ *– независимые случайные величины,* $\mathrm{M}\,\xi_i = 0$, $\mathrm{P}\left(|\xi_i| \le c\right) = 1$, $i = \overline{1, n}$, *тогда для любых* $\varepsilon > 0$

$$\mathrm{P}\left(\max_{1 \le k \le n} |S_k| \ge \varepsilon\right) \ge 1 - \frac{(c+\varepsilon)^2}{\mathrm{M}\, S_n^2}. \qquad (1)$$

Д о к а з а т е л ь с т в о. Обозначим

$$A=\left(\max_{1\le k\le n}|S_k|\ge\varepsilon\right),\qquad A_k=\left(|S_i|<\varepsilon,\,|S_k|\ge\varepsilon,\,i=\overline{1,k-1}\right),\qquad k=\overline{1,n},$$

тогда $A=\sum_{k=1}^{n}A_k$. Несложно видеть, что

$$\mathrm{M}\,S_n^2 I_A=\mathrm{M}\,S_n^2-\mathrm{M}\,S_n^2 I_{\overline{A}}\ge\mathrm{M}\,S_n^2-\varepsilon^2\mathrm{P}\left(\overline{A}\right)=\mathrm{M}\,S_n^2-\varepsilon^2+\varepsilon^2\mathrm{P}(A),\tag{2}$$

с другой стороны на множествах A_k, $k=\overline{1,n}$,

$$|S_{k-1}|\le\varepsilon,\qquad |S_k|\le|S_{k-1}|+|\xi_k|\le\varepsilon+c,$$

поэтому

$$\mathrm{M}\,S_n^2 I_A=\sum_{k=1}^{n}\mathrm{M}\,S_k^2 I_{A_k}+\sum_{k=1}^{n}\mathrm{M}\left(I_{A_k}\left(S_n-S_k\right)^2\right)\le$$

$$\le(c+\varepsilon)^2\sum_{k=1}^{n}\mathrm{P}(A_k)+\sum_{k=1}^{n}\mathrm{P}(A_k)\sum_{j=k+1}^{n}\mathrm{M}\,\xi_j^2\le$$

$$\le\mathrm{P}(A)\left[(c+\varepsilon)^2+\sum_{j=1}^{n}\mathrm{M}\,\xi_j^2\right]=\mathrm{P}(A)\left[(c+\varepsilon)^2+\mathrm{M}\,S_n^2\right].\tag{3}$$

Из (2) и (3) находим, что

$$\mathrm{P}(A)=\frac{\mathrm{M}\,S_n^2-\varepsilon^2}{(c+\varepsilon)^2+\mathrm{M}\,S_n^2-\varepsilon^2}=1-\frac{(c+\varepsilon)^2}{(c+\varepsilon)^2+\mathrm{M}\,S_n^2-\varepsilon^2}\ge1-\frac{(c+\varepsilon)^2}{\mathrm{M}\,S_n^2},$$

Неравенство (1) доказано.

Теорема 2 (**Колмогоров, Хинчин**). *Пусть* $\mathrm{M}\,\xi_n=0,\ n\in N$.

1) *Тогда, если*

$$\sum_{n=1}^{\infty}\mathrm{M}\,\xi_n^2<\infty,\tag{4}$$

то ряд

$$\sum_{n=1}^{\infty}\xi_n\tag{5}$$

сходится с вероятностью единица.

2) *Если существует* $c<\infty$ *такое, что* $\mathrm{P}\left(|\xi_k|\le c\right)=1,\ k=1,\,2,\,\ldots,$ *(равномерная ограниченность последовательности случайных величин), то верно и обратное:из сходимости с вероятностью единица ряда* (5) *следует условие* (4).

Д о к а з а т е л ь с т в о. 1) Последовательность $\{S_n\}$, $n\in N$, сходится с вероятностью единица тогда и только тогда, когда она фундаментальна с вероятностью единица. Известно [41], что $\{S_n\}$, $n\in N$, фундаментальна с вероятностью единица в том и только том случае, когда для любого $\varepsilon>0$

$$\mathrm{P}\left(\sup_{k\ge1}|S_{n+k}-S_n|\ge\varepsilon\right)\xrightarrow[n\to\infty]{}0.\tag{6}$$

В силу неравенства Колмогорова – следствия 1 теоремы 6 (гл. 6, § 3)

$$P\left(\sup_{k\ge 1}|S_{n+k}-S_n|\ge\varepsilon\right)=\lim_{N\to\infty}P\left(\max_{1\le k\le N}|S_{n+k}-S_n|\ge\varepsilon\right)\le$$

$$\le\lim_{N\to\infty}\frac{\sum_{k=n}^{N}M\,\xi_k^2}{\varepsilon^2}=\frac{\sum_{k=n}^{\infty}M\,\xi_k^2}{\varepsilon^2},$$

поэтому, если выполнено условие (4), то справедливо (6) и, следовательно, ряд (5) сходится с вероятностью единица.

Пусть ряд (5) сходится почти наверное. Тогда в силу (6) для достаточно больших n

$$P\left(\sup_{k\ge 1}|S_{n+k}-S_n|\ge\varepsilon\right)<\frac{1}{2}. \tag{7}$$

Из неравенства (1) следует, что

$$P\left(\sup_{k\ge 1}|S_{n+k}-S_n|\ge\varepsilon\right)\ge 1-\frac{(c+\varepsilon)^2}{\sum_{k=n}^{\infty}M\,\xi_k^2}.$$

Поэтому, если допустить, что

$$\sum_{k=1}^{\infty}M\,\xi_k^2=\infty,$$

то получим

$$P\left(\sup_{k\ge 1}|S_{n+k}-S_n|\ge\varepsilon\right)=1,$$

что противоречит неравенству (7).

Теорема 1 доказана.

П р и м е р 1. Пусть $\xi_1,\ \dots,\ \xi_n,\ \dots$ – последовательность независимых случайных величин Бернулли, таких, что

$$P(\xi_n=1)=P(\xi_n=-1)=1/2.$$

Тогда ряд

$$\sum_{n=1}^{\infty}a_n\xi_n,$$

где $|a_n|\le c$ для $n=1,\ 2,\ \dots$, сходится с вероятностью единица тогда и только тогда, когда

$$\sum_{n=1}^{\infty}a_n^2<\infty.$$

Теорема 3 (**О «двух рядах»**). *Для сходимости с вероятностью единица ряда* $\sum_{n=1}^{\infty}\xi_n$ *из независимых случайных величин достаточно, чтобы одновременно сходились два ряда*

$$\sum_{n=1}^{\infty}M\,\xi_n\ \text{ и }\ \sum_{n=1}^{\infty}D\,\xi_n.$$

Если к тому же существует $c<\infty$ *такое, что* $P\left(|\xi_k|\le c\right)=1$ *для всех* $k\in N$, *то это условие является и необходимым.*

Доказательство. Достаточность. Если

$$\sum_{n=1}^{\infty} \mathrm{D}\,\xi_n < \infty,$$

то по теореме 2 ряд

$$\sum_{n=1}^{\infty} (\xi_n - \mathrm{M}\,\xi_n)$$

сходится с вероятностью единица. Но по предположению теоремы ряд

$$\sum_{n=1}^{\infty} \mathrm{M}\,\xi_n$$

сходится, следовательно, искомый ряд из случайных величин ξ_n сходится с вероятностью единица.

Необходимость. Для доказательства воспользуемся приемом «симметризации». Наряду с последовательностью $\xi_1, \ldots, \xi_n, \ldots$ рассмотрим не зависящую от нее последовательность независимых случайных величин $\overline{\xi}_1, \ldots, \overline{\xi}_n, \ldots$ таких, что $\overline{\xi}_n$ имеет то же распределение, что и ξ_n, $n \in N$.

Тогда, если сходится с вероятностью единица ряд $\sum \xi_n$, то сходится и ряд $\sum \overline{\xi}_n$, а значит, и ряд $\sum \left(\xi_n - \overline{\xi}_n\right)$. Но

$$\mathrm{M}\left(\xi_n - \overline{\xi}_n\right) = 0$$

и

$$\mathrm{P}\left(\left|\xi_n - \overline{\xi}_n\right| \le 2c\right) = 1, \ n \in N.$$

Поэтому по теореме 2

$$\sum_{n=1}^{\infty} \mathrm{D}\left(\xi_n - \overline{\xi}_n\right) < \infty.$$

Далее

$$\sum_{n=1}^{\infty} \mathrm{D}\,\xi_n \le \frac{1}{2}\sum_{n=1}^{\infty} \mathrm{D}\left(\xi_n - \overline{\xi}_n\right) < \infty,$$

поэтому по теореме 2 с вероятностью единица сходится ряд $\sum(\xi_n - \mathrm{M}\,\xi_n)$, а значит, сходится и ряд $\sum \mathrm{M}\,\xi_n$.

Теорема 3 доказана.

Следующая теорема дает необходимое и достаточное условие сходимости ряда без предположения об ограниченности случайных величин.

Пусть c – некоторая константа и

$$\xi^c = \begin{cases} \xi, & |\xi| \le c, \\ 0, & |\xi| > c. \end{cases}$$

Теорема 4 (Колмогорова о «трех рядах»). *Пусть* $\xi_1, \xi_2, \ldots$ *– последовательность независимых случайных величин. Для сходимости почти наверное ряда*

$\sum_{n=1}^{\infty}\xi_n$ *необходимо, чтобы для любого* $c>0$ *сходились ряды*

$$\sum_{n=1}^{\infty}\mathrm{M}\,\xi_n^c,\ \sum_{n=1}^{\infty}\mathrm{D}\,\xi_n^c,\ \sum_{n=1}^{\infty}\mathrm{P}\left(\left|\xi_n\right|\ge c\right)$$

и достаточно, чтобы эти ряды сходились при некотором $c>0$.

Доказательство. Достаточность. По теореме о «двух рядах» ряд $\sum\xi_n^c$ сходится с вероятностью единица, но если

$$\sum_{n=1}^{\infty}\mathrm{P}\left(\left|\xi_n\right|\ge c\right)<\infty,$$

то по лемме Бореля – Кантелли с вероятностью единица

$$\sum_{n=1}^{\infty}I_{\{|\xi_n|\ge c\}}<\infty,$$

а значит $\xi_n=\xi_n^c$ для всех n, за исключением, быть может, конечного числа. Поэтому ряд $\sum\xi_n$ также сходится с вероятностью единица.

Необходимость. Если ряд $\sum\xi_n$ сходится с вероятностью единица, то с вероятностью единица $\xi_n\to 0$ при $n\to\infty$ и для любого $c>0$ может произойти не более конечного числа событий $\left(\left|\xi_n\right|\ge c\right)$. Поэтому

$$\mathrm{P}\left(\sum_{n=1}^{\infty}I_{(|\xi_n|\ge c)}<\infty\right)=1$$

и по лемме Бореля – Кантелли

$$\sum_{n=1}^{\infty}\mathrm{P}\left(\left|\xi_n\right|\ge c\right)<\infty.$$

Далее, из сходимости ряда $\sum\xi_n$ следует и сходимость ряда $\sum\xi_n^c$, поэтому по теореме о «двух рядах» каждый из рядов $\sum\mathrm{M}\,\xi_n^c$ и $\sum\mathrm{D}\,\xi_n^c$ сходится с вероятностью единица.

Теорема 4 доказана.

Упражнение 1. Пусть $\xi_1,\ \ldots,\ \xi_n,\ \ldots$ – независимые случайные величины с $\mathrm{M}\,\xi_n=0,\ n\in N$. Докажите, что если

$$\sum_{n=1}^{\infty}\mathrm{M}\frac{\xi_n^c}{1+\left|\xi_n\right|}<\infty$$

при $c=1$, то ряд $\sum\xi_n$ сходится с вероятностью единица.

Задачи

1. Пусть последовательность функций распределения $\{F_n\}$, $n\in N$, слабо сходится к непрерывной функции распределения F. Покажите, что эта сходимость равномерна.

2. Пусть $\xi_1, \xi_2, \ldots$ – последовательность независимых нормально распределенных случайных величин,

$$\mathrm{M}\,\xi_k = 0,\ k \in N,\ \mathrm{D}\,\xi_1 = 1,\ \mathrm{D}\,\xi_k = 2^{k-2},\ k \geq 2.$$

Покажите, что в этом случае условие Линдеберга не выполнено, но в то же время справедлива центральная предельная теорема.

3. Пусть $\xi_1, \xi_2, \ldots$ – последовательность независимых одинаково распределенных случайных величин, $\mathrm{M}\,\xi_1 = 0,\ \mathrm{M}\,\xi_1^2 = 1$. Докажите, что

$$\max\left(\frac{|\xi_1|}{\sqrt{n}}, \ldots, \frac{|\xi_n|}{\sqrt{n}}\right) \Rightarrow 0,\ n \to \infty.$$

4. Пусть $\xi_1, \xi_2, \ldots$ – последовательность независимых случайных величин. Докажите, что случайные величины $\overline{\lim\limits_{n\to\infty}}\,\xi_n$ и $\underline{\lim\limits_{n\to\infty}}\,\xi_n$ являются вырожденными.

5. Случайные величины $\xi_1, \xi_2, \ldots$ независимы и одинаково распределены. Докажите, что с вероятностью единица произойдет конечное число событий $A_n = \left(|\xi_n| \geq \sqrt{n}\right)$ тогда и только тогда, когда $\mathrm{D}\,\xi_1$ конечна.

6. Пусть $\xi_1, \xi_2, \ldots$ – последовательность независимых случайных величин,

$$\mathrm{P}\left(\xi_n = 0\right) = 1 - 1/n,\ \mathrm{P}\left(\xi_n = 1\right) = 1/n,\ n \in N.$$

Докажите, что эта последовательность сходится в среднем порядка $r > 0$, но не сходится почти наверное.

7. Пусть $\xi_1, \xi_2, \ldots$ – последовательность независимых случайных величин,

$$\mathrm{P}\left(\xi_n = 0\right) = 1 - 1/n,\quad \mathrm{P}\left(\xi_n = n^{2/r}\right) = 1/n,\ n \in N,\ r > 0.$$

Покажите, что эта последовательность сходится по вероятности, но не сходится в среднем порядка r.

8. Пусть $\xi_n = a_n\eta$, где числовая последовательность $\{a_n\}$, $n \in N$, сходится, а $\mathrm{M}|\eta|^r = \infty$, $r > 0$. Докажите, что последовательность случайных величин $\{\xi_n\}$, $n \in N$, сходится почти наверное, но не сходится в среднем порядка r.

9. Пусть

$$\xi_n = n^2\xi\exp(-n\xi),\ n \in N,$$

где ξ – экспоненциальная случайная величина. Покажите, что $\xi_n \to 0$ при $n \to \infty$ поточечно, но $\mathrm{M}\,\xi_n$ не сходится к нулю при $n \to \infty$.

10. Пусть

$$S_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\xi_j,$$

где ξ_j – независимые одинаково распределенные случайные величины, $\mathrm{M}\,\xi_j = 0$, $\mathrm{D}\,\xi_j = 1$, $j = \overline{1,n}$. Докажите, что последовательность $\{S_n\}$, $n \in N$, сходится слабо, но не сходится в среднем порядка 2.

11. Пусть ξ – случайная величина Пуассона с параметром $\lambda > 0$. Найдите слабый предел при $\lambda \to \infty$ последовательности

$$\left(\frac{\xi - \lambda}{\sqrt{\lambda}} \right).$$

12. Пусть $\xi_1, \xi_2, \ldots$ – последовательность случайных величин, для которой выполняется закон больших чисел. Обязан ли выполняться закон больших чисел для последовательности $|\xi_1|, |\xi_2|, \ldots$?

13. Пусть $\xi_1, \xi_2, \ldots$ – независимые одинаково распределенные случайные величины,

$$S_n = \sum_{k=1}^{n} \xi_k, \qquad \eta_n = \frac{S_n}{n}, \qquad \chi_n = \frac{S_n}{\sqrt{n}}.$$

Найдите слабые пределы при $n \to \infty$ последовательностей случайных величин $\{S_n\}$, $\{\eta_n\}$, $\{\chi_n\}$, $n \in N$, в том случае, если случайная величина ξ_1 имеет:

а) биномиальное распределение;

б) распределение Пуассона;

в) равномерное на отрезке $[a;b]$ распределение;

г) нормальное распределение;

д) распределение Коши.

14. Пусть $\xi_1, \xi_2, \ldots$ – последовательность независимых одинаково распределенных случайных величин с конечными дисперсиями. Найдите

$$\lim_{n \to \infty} \mathrm{P}\left(\xi_1 + \ldots + \xi_n \le x \right), \quad x \in R.$$

15. Пусть $\{f(m)\}$, $m \in N$, – произвольная последовательность действительных чисел, $v_n(\ldots)$ – частота всех натуральных чисел $m \le n$, подчиненных условиям, которые написаны в скобках,

$$M_n = \frac{1}{n} \sum_{m=1}^{n} f(m), \;\; D_n = \frac{1}{n} \sum_{m=1}^{n} \left(f(m) - M_n \right)^2,$$

$\psi(n)$ – произвольная неограниченно возрастающая при $n \to \infty$ функция. Докажите аналог закона больших чисел

$$\nu_n \left(\left| f(m) - M_n \right| \le \psi(n) \sqrt{D_n} \right) \to 1, \quad (n \to \infty).$$

16. На отрезке $[0;1]$ наудачу выбирается число ξ и разлагается в десятичную дробь

$$\xi = \sum_{n=1}^{\infty} \frac{l_n(\xi)}{10^n}.$$

Докажите, что при надлежащей нормировке $l_1(\xi) + \ldots + l_n(\xi)$ при $n \to \infty$ слабо сходится к нормальной случайной величине.

Глава 7. ОСНОВЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

Теория случайных процессов – сравнительно молодая ветвь теории вероятностей, имеющая важные приложения в целом ряде разделов физики, техники и других областей естественных наук. Именно запросы со стороны этих дисциплин и привели к бурному развитию рассматриваемой теории в течение последних десятилетий. К настоящему времени в распоряжении читателя имеется большое количество книг по этой теории. Следует отметить, что теория случайных процессов настолько многообразна. Что многие из ее разделов еще не нашли отражения в учебной и монографической литературе.

В настоящей главе излагаются основы теории случайных процессов. В четырех параграфах основного текста дается определение случайного процесса, рассматриваются случайные процессы с независимыми приращениями, стационарные диффузионные, марковские случайные процессы, приводятся основные положения корреляционной теории. В дополнениях к главе представлены элементы теории обобщенных случайных процессов. Более полное изложение теории случайных процессов можно найти в книгах [7], [9], [12], [27].

§ 1. Определение случайного процесса

В главах 3 и 6 мы рассматривали конечные, либо счетные семейства случайных величин, определенных на одном вероятностном пространстве. В этом параграфе будет исследован случай, когда семейство случайных величин может быть континуальным.

Определение 1. *Случайным процессом называется семейство случайных величин* $\xi(t) = \xi(t, \omega)$, *заданных на одном вероятностном пространстве* $(\Omega, \mathcal{A}, \mathrm{P})$ *и зависящих от параметра* t, *принимающего значения из некоторого множества* T.

Пример 1. Последовательности независимых случайных величин $\xi_1, \xi_2, \ldots$, рассматриваемые в предыдущих параграфах, являются случайными процессами, для которых $T = \{1,\ 2,\ 3, \ldots\}$. То же самое справедливо и относительно сумм $S_1, S_2, \ldots$ слагаемых $\xi_1, \xi_2, \ldots$. Случайными процессами являются также последовательности случайных величин, связанных в цепь Маркова. Такие процессы, у которых множество T можно отождествить со всей или с частью последовательности $\{\ldots, -1, 0, 1, \ldots\}$, обычно называются *процессами с дискретным временем* или *случайными последовательностями*.

П р и м е р 2. Если T совпадает с некоторым числовым интервалом $T=[a;b]$, $-\infty \le a < b \le +\infty$, то совокупность $\{\xi(t),\ t\in T\}$ называют *процессом с непрерывным временем*. Интерпретация параметра t как времени, конечно, не обязательна. Она возникла исторически, поскольку в большинстве естественнонаучных задач, которые привели к появлению понятия случайного процесса, параметр t был временем, а значение $\xi(t)$ было тем, что наблюдалось в момент времени t.

Как случайный процесс можно рассматривать, например, движение молекулы газа во времени, уровень воды в водохранилище, колебание крыла самолета и т. п. Процессом с непрерывным временем является случайная функция

$$\xi(t)=\sum_{k=0}^{\infty}\sin kt\cdot\frac{\xi_k}{2^k},\ t\in[0,2\pi],$$

где случайные величины ξ_k независимы и одинаково распределены.

По аналогии со случайными величинами (скалярными и векторными) приведем еще одно определение случайного процесса (как измеримого по ω отображения), эквивалентное определению 1. Для этого зафиксируем $\omega\in\Omega$. Функцию $\xi(t), t\in T$, будем называть *выборочной функцией* или *траекторией процесса*. Через X обозначим пространство функций $x(t),\ t\in T$, в котором лежат траектории $\xi(t)$. Пусть далее $\mathscr{B}_X^T$ – σ- алгебра подмножеств из X, порожденная множествами вида

$$C=\{x\in X:\ x(t_1)\in B_1,\ ...,\ x(t_n)\in B_n\} \tag{1}$$

для любых $n\in N$, любых $t_1,...,t_n$ из T и любых борелевских множеств $B_1,...,B_n$ из R. Множества вида (1) называются *цилиндрическими*.

Случайный процесс $\xi(t)$ будем трактовать как измеримое отображение $(\Omega,\mathscr{A})$ в $(X,\mathscr{B}_X^T)$. Это отображение индуцирует распределение P_ξ на измеримом пространстве $(X,\mathscr{B}_X^T)$, определяемое равенствами

$$P_\xi(B)=P(\xi^{-1}(B)),\ \forall B\in\mathscr{B}_X^T.$$

Тройка $(X,\mathscr{B}_X^T,P_\xi)$ называется *выборочным вероятностным пространством*. В этом пространстве элементарный исход «ω» отождествляется с траекторией процесса, а мера P_ξ называется *распределением процесса* ξ.

П р и м е р 3 . В качестве пространства X в теории случайных процессов с континуальным множеством T значений времени t чаще других рассматриваются следующие пространства функций:

1) Пространство всех функций на T

$$X = R^T = \prod_{t \in T} R_t,$$

где $R_t = (-\infty, +\infty)$. Это пространство обычно рассматривают в паре с σ-алгеброй $\mathscr{B}_X^T$.

2) Пространство всех непрерывных на T функций

$$X = C(T).$$

В этом пространстве σ-алгебра $\mathscr{B}_C^T$ совпадает с σ-алгеброй $\mathscr{B}(C(T))$ – борелевской σ-алгеброй, порожденной множествами, открытыми относительно равномерной метрики

$$\rho(x, y) = \sup_{t \in \mathrm{T}} \left| y(t) - x(t) \right|, x, y \in C(T).$$

(см., например, [9]).

3) Пространство $D(T)$ функций, у которых в каждой точке $t \in T$ существуют пределы $x(t-0)$ и $x(t+0)$ и значение функции $x(t)$ совпадает с одним из них. Если $T = [a, b]$, то предполагается также, что $x(a) = x(a+0)$, $x(b) = x(b-0)$. Через $D_+(T)$ $(D_-(T))$ будем обозначать подпространство функций из $D(T)$ непрерывных справа (слева). В качестве σ-алгебры подмножеств $D(T)$ будем рассматривать $\mathscr{B}_D^T$.

О п р е д е л е н и е 2 . *Пусть* X *– заданное пространство функций и* $\mathscr{G}$ *есть σ-алгебра его подмножеств, содержащая σ-алгебру* $\mathscr{B}_X^T$. *Случайным процессом* $\xi(t) = \xi(t, \omega)$ *называется измеримое по* ω *отображение* $(\Omega, \mathscr{A}, \mathrm{P})$ *на* $\left(X, \mathscr{G}, \mathrm{P}_\xi\right)$ *(каждому* ω *ставится в соответствие* $\xi(t) = \xi(t, \omega)$ *так, что* $\xi^{-1}(C) \in \mathscr{A}$ *для любого* $C \in \mathscr{G}$*)*

З а м е ч а н и е 1 . Условие $\mathscr{B}_X^T \subset \mathscr{G}$ нужно для того, чтобы были определены вероятности цилиндрических множеств и, в частности, $\mathrm{P}(\xi(t) \in B)$, $B \in \mathscr{B}(R)$, означающие, что $\xi(t)$ есть случайные величины, $t \in T$.

З а м е ч а н и е 2 . Из теории для случайных величин известно, что задать случайную величину – значит задать ее распределение вероятностей. Случайные процессы пытаются описать теми или иными сведениями о его распределении. Можно задавать, например, *конечномерные распределения*, т. е. вероятности $\mathrm{P}(C)$ множеств C вида (1).

Из теоремы Колмогорова о согласованных распределениях [5] следует, что задание согласованных конечномерных распределений однозначно определяет распределение процесса на $\left(R^T, \mathcal{B}_{R^T}^T\right)$.

Однако пространство $\left(R^T, \mathcal{B}_{R^T}^T\right)$ не очень удобно для изучения случайных процессов, т. к. далеко не все часто употребляемые в анализе соотношения для функций принадлежат σ-алгебре $\mathcal{B}_{R^T}^T$. Множество $\left\{\sup\limits_{t\in T}\xi(t) < C\right\}$, например, может и не быть событием, так как мы знаем лишь его в виде пересечения несчетного числа измеримых множеств

$$\bigcap_{t\in T}\left(\xi(t) < C\right),$$

если T – интервал числовой оси.

Имеет место и другое неудобство – распределение P_ξ на $\left(R^T, \mathcal{B}_{R^T}^T\right)$ не определяет однозначно свойств траекторий ξ. Этот факт связан с тем, что пространство R^T и принадлежность цилиндрическому множеству вида (1) не несут в себе никакой информации о поведении $x(t)$ в точках t, отличных от $t_1, ..., t_n$, что подтверждает следующий пример.

П р и м е р 4. Случайные процессы $\xi(t)$ и $\eta(t)$ будем называть *стохастически эквивалентными*, если $\mathrm{P}\left(\xi(t) = \eta(t)\right) = 1$ для всех $t \in T$. Процесс η при этом называют *модификацией* ξ. Легко видеть, что если процессы стохастически эквивалентные, то их конечномерные распределения совпадают, но не наоборот. Что касается траекторий, то они у стохастически эквивалентных процессов могут быть различными. Например, пусть $T = [0;1]$, τ – абсолютно непрерывная случайная величина, $0 < \tau < 1$. Положим $\xi(t) \equiv 0$, $\eta(t) = 0$, если $t \neq \tau$, $\eta(t) = 1$, если $t = \tau$. Тогда

$$\mathrm{P}\left(\xi(t) \neq \eta(t)\right) = \mathrm{P}(\tau = t) = 0$$

для любых $t \in T$, т. е. ξ и η – стохастически эквивалентны.

В связи с замечанием 2 отметим также, что

$$\mathrm{P}\left(\sup_{t\in T}\xi(t) \le \frac{1}{2}\right) = 1,$$

$$\mathrm{P}\left(\sup_{t\in T}\eta(t) \le \frac{1}{2}\right) = 0.$$

На приведенном примере 4 легко понять, что множество всех непрерывных функций $C(T)$, множество $\left\{\sup\limits_{t\in T}\xi(t)\le x\right\}$ и многие другие не принадлежат $\mathscr{B}_{R^T}^T$. Простейший способ преодоления этих трудностей состоит в том, чтобы задавать процессы в пространствах $C(T)$ или $D(T)$, если это возможно. Если, например, $\xi(t),\eta(t)\in C(T)$, и они стохастически эквивалентны, то траектории этих процессов будут с вероятностью 1 совпадать, так как

$$\bigcap_{t\in Q}\{\xi(t)=\eta(t)\}=\bigcap_{t\in T}\{\xi(t)=\eta(t)\}=\{\xi(t)=\eta(t),\forall t\in T\},$$

где вероятность события левой части определена и равна 1. Аналогично обстоит дело с событиями

$$\left\{\sup_{t\in T}\xi(t)\le c\right\}=\bigcap_{t\in T}\{\xi(t)\le c\}.$$

Точно такие же рассуждения справедливы для пространства $D(T)$, поскольку каждый элемент $x(\cdot)$ из D однозначно определяется значениями $x(t)$ на счетном всюду плотном множестве значений t.

Для того, чтобы совсем не иметь дела с неудобным пространством $\left(R^T,\mathscr{B}_{R^T}^T\right)$, можно задавать распределение рассматриваемого процесса ξ на суженном пространстве $\left(C(T),\mathscr{B}_C^T\right)$. Если ξ и η эквивалентны на $\left(R^T,\mathscr{B}_{R^T}^T\right)$ и мы рассматриваем η на $\left(C(T),\mathscr{B}_C^T\right)$, то мы тем самым построили процесс η – *непрерывную модификацию* ξ.

Чтобы осуществить эту конструкцию, надо уметь выяснять по распределению процесса ξ, существует для него непрерывная модификация η или нет (все приведенные выше рассуждения аналогично переносятся на $D(T)$).

Очень простой критерий существования непрерывной модификации, принадлежащий Колмогорову, основан на знании лишь двумерных распределений $\xi(t)$.

Теорема 1 [5] (Колмогоров). *Пусть* $\xi(t)$ *– случайный процесс на* $[0;1]$. *Если при всех* $t, t+h$ *из отрезка* $[0;1]$

$$\mathrm{M}\left|\xi(t+h)-\xi(t)\right|^a\le C\left|h\right|^{1+b}\tag{2}$$

при каких-нибудь $a > 0, b > 0, c < \infty$, *то* $\xi(t)$ *имеет непрерывную модификацию.*

Пример 5. Допустим, что случайный процесс $\xi(t)$ имеет вид

$$\xi(t) = \sum_{k=1}^{r} \xi_k \varphi_k(t),$$

где $\varphi_k(t)$ удовлетворяют условию Гёльдера

$$\left|\varphi_k(t+h) - \varphi_k(t)\right| \le C|h|^{\alpha}, \alpha > 0,$$

и при некотором $l > 1/\alpha$

$$\mathrm{M}\left|\xi_k\right|^l < \infty,\ k = \overline{1,r}.$$

Тогда процесс $\xi(t)$ (он является, очевидно, непрерывным) удовлетворяет условию (2).

Критерий существования модификации, принадлежащей пространству $D(T)$, формулируется сложнее и связан с более слабыми условиями на процесс. Ограничимся лишь формулировкой следующего утверждения.

Теорема 2 [22] (Колмогоров – Ченцов). *Если при некоторых* $\alpha \ge 0, \beta \ge 0, b > 0$ *и при всех* $t - h_1 \le t \le t + h_2$ *из отрезка* $[0;1]$

$$\mathrm{M}\left|\xi(t) - \xi(t-h_1)\right|^{\alpha}\left|\xi(t+h_2) - \xi(t)\right|^{\beta} \le Ch^{1+b},\ h = h_1 + h_2, \tag{3}$$

то существует модификация $\xi(t)$ *из* $D(0;1)$.

Пример 6. Пусть γ – равномерно распределенная на $[0;1]$ случайная величина, $\xi(t) = 0$ при $t < \gamma$, $\xi(t) = 1$ при $t \ge \gamma$. Тогда

$$\mathrm{M}\left|\xi(t+h) - \xi(t)\right|^l = \mathrm{P}\big(\gamma \in (t, t+h)\big) = h.$$

Здесь условие (2) не выполнено, хотя

$$\xi(t+h) - \xi(t) \xrightarrow[h \to 0]{\mathrm{P}} 0.$$

Условие (3), очевидно, выполнено, так как

$$\mathrm{M}\left|\xi(t) - \xi(t-h_1)\right|\left|\xi(t+h_2) - \xi(t)\right| = 0.$$

В общем случае, когда не располагают данными для построения процесса из пространств $C(T)$ или $D(T)$, опираются на понятие сепарабельности процесса.

Определение 3. *Случайный процесс* $\xi(t)$ *называется сепарабельным, если существует счетное всюду плотное в* T *множество* S *такое, что*

$$\mathrm{P}\left(\varlimsup_{u \to t, t \in S} \xi(u) \ge \xi(t) \ge \varliminf_{u \to t, t \in S} \xi(u), \forall t \in T\right) = 1. \tag{4}$$

З а м е ч а н и е 3. Условие (4) эквивалентно тому, что для любого интервала $I \subset T$

$$\mathrm{P}\left(\sup_{u\in I\cap S}\xi(u)=\sup_{u\in I}\xi(u);\ \inf_{u\in I\cap S}\xi(u)=\inf_{u\in I}\xi(u)\right)=1.$$

Известно (см., напр., [12]), что *любой случайный процесс имеет сепарабельную модификацию.*

Можно также отметить, что для сепарабельных процессов такие множества, как множество всех неубывающих функций, множества $C(T)$, $D(T)$ являются событиями. Процессы из $C(T)$, $D(T)$ автоматически будут сепарабельными.

Наряду с непрерывностью выборочных траекторий существует еще один способ характеризовать свойство непрерывности случайного процесса.

О п р е д е л е н и е 4. *Случайный процесс $\xi(t)$ называется стохастически непрерывным, если при всех t, $t+h\in T$, $h\to 0$*

$$\xi(t+h)-\xi(t)\xrightarrow{\mathrm{P}}0.$$

Ясно, что все процессы с непрерывными траекториями являются стохастически непрерывными. Но не только. Разрывный процесс из примера 6 также является стохастически непрерывным.

В дальнейшем нам понадобится и такое понятие: *случайный процесс $\xi(t)$ называется непрерывным в среднем порядка r (в среднеквадратичном при $r=2$), если для любых t, $t+h\in T$ при $h\to 0$*

$$\mathrm{M}\left|\xi(t+h)-\xi(t)\right|^r\to 0.$$

Разрывный процесс из примера 6 будет непрерывным в среднем любого порядка.

§ 2. Случайные процессы с независимыми приращениями

Одним из важнейших классов случайных процессов является класс случайных процессов с независимыми приращениями. К этому классу, как будет показано ниже, принадлежат процесс броуновского движения, пуассоновский случайный процесс и др.

О п р е д е л е н и е 1. *Случайный процесс $\xi(t)$, $t\in T$, называется процессом с независимыми приращениями, если любого $n\in N$ и для всех $t_0<t_1<\ldots<t_n$ из T случайные величины $\xi(t_0)$, $\xi(t_1)-\xi(t_0)$, …, $\xi(t_n)-\xi(t_{n-1})$ являются независимыми.*

З а м е ч а н и е 1. Так как случайный вектор $\left(\xi(t_1),...,\xi(t_n)\right)$ получается невырожденным линейным преобразованием из вектора $\left(\xi(t_1)-\xi(t_0),...,\xi(t_n)-\xi(t_{n-1})\right)$, то для того, чтобы задать конечномерные распределения случайного процесса с независимыми приращениями, достаточно задать распределение процесса в одной точке $\xi(t)$ и распределения приращений $\xi(s)-\xi(t)$ при $s>t$.

Пусть $T=[0,\infty]$. Процесс $\xi(t)$, $t\geq 0$, называется *однородным*, если $\xi(0)=0$ (п.н.) и распределение величины $\xi(t+h)-\xi(t)$, $t\geq 0$, $h\geq 0$ не зависит от t.

Введем *характеристическую функцию однородного процесса* $\xi(t)$

$$\varphi(t,z)=\mathrm{M}\exp\left\{iz\cdot\xi(t)\right\}, \tag{1}$$

где $z\in R$.

У т в е р ж д е н и е 1. *Характеристическая функция* $\varphi(t,z)$ *удовлетворяет следующему функциональному уравнению*

$$\varphi(t+s,z)=\varphi(t,z)\cdot\varphi(s,z),\ t,s>0 \tag{2}$$

Д о к а з а т е л ь с т в о. Учитывая то, что $\xi(t)$ – однородный случайный процесс с независимыми приращениями, имеем

$$\mathrm{M}\exp\left\{i\left(z,\xi(t+s)\right)\right\}=\mathrm{M}\exp\left\{i\left(z,\xi(t+s)-\right.\right.$$

$$\left.\left.-\xi(s)\right)\right\}\exp\left\{i\left(z,\xi(s)\right)\right\}=\mathrm{M}\exp\left\{i\left(z,\xi(t+s)-\right.\right.$$

$$\left.\left.-\xi(s)\right)\right\}\mathrm{M}\exp\left\{i\left(z,\xi(s)\right)\right\}=\mathrm{M}\exp\left\{i\left(z,\xi(t)\right)\right\}\mathrm{M}\exp\left\{i\left(z,\xi(s)\right)\right\}.$$

Утверждение 1 доказано.

У т в е р ж д е н и е 2. *Пусть* $\xi(t)$, $t\in T$, *– стохастически непрерывный однородный процесс. Тогда* $\varphi(t,z)$ *непрерывна по переменной* $t\geq 0$.

Д о к а з а т е л ь с т в о вытекает из соотношения (2) и того, что

$$\lim_{s\to 0}\varphi(s,z)=1. \tag{3}$$

Для доказательства (3) заметим, что для любого $\varepsilon>0$

$$\left|\varphi(s,z)-1\right|\leq M\left|\exp iz\xi(s)-1\right|\leq\varepsilon+2\mathrm{P}\left(\left|z\cdot\xi(s)\right|>\varepsilon\right).$$

Последняя вероятность стремиться к нулю при $s\to 0$ в силу того, что ξ – стохастически непрерывный процесс.

Утверждение 2 доказано.

Теорема 1. *Для стохастически непрерывного однородного случайного процесса* $\xi(t)$, $t \in T$, *существуют такие* $a \in R$, $b > 0$, *и неубывающая функция ограниченной вариации* $\bar{G}$ *на* R, *что*

$$\varphi(t,z) = \exp\left\{ t\left(iaz - \frac{1}{2}bz^2 + \int_R \left(e^{izx} - 1 - \frac{izx}{1+x^2} \right) \frac{1+x^2}{x^2} d\bar{G}(x) \right) \right\}. \quad (4)$$

З а м е ч а н и е 2. Пусть

$$G(x) = \bar{G}(x) + b\varepsilon(x),$$

где $\varepsilon(x) = 0$ при $x < 0$, $\varepsilon(x) = 1$ при $x > 0$. Доопределив функцию

$$\left(e^{izx} - 1 - \frac{izx}{1+x^2} \right) \frac{1+x^2}{x^2}$$

в точке $x = 0$ по непрерывности значением $-\frac{z^2}{2}$, соотношение (4) можно записать так

$$\varphi(t,z) = \exp\left(t\left(iaz + \int_R \left(e^{ixz} - 1 - \frac{izx}{1+x^2} \right) \frac{1+x^2}{x^2} dG(x) \right) \right) \quad (5)$$

Д о к а з а т е л ь с т в о теоремы 1. Из соотношения (3) и равенства

$$\varphi(t,z) = \varphi^n\left(\frac{t}{n}, z \right)$$

вытекает, что $\varphi(t,z) \neq 0$ для всех t. Далее

$$\varphi\left(\frac{k}{n}, z \right) = \left[\varphi(1,z) \right]^{k/n}$$

(берется главное значение корня n-ой степени из $\varphi(1,z)$). Переходя к пределу при $\frac{k}{n} \to t$ и обозначая $k(z) = \ln \varphi(1,z)$ главное значение логарифма, имеем

$$\varphi(t,z) = \exp(tk(z)) \quad (6)$$

Функция $k(z)$ называется *кумулянтой* процесса ξ и может быть вычислена по формуле

$$k(z) = \lim_{h \to 0} \frac{1}{h} \left[\varphi(h,z) - 1 \right] \quad (7)$$

Пусть F_h – функция распределения случайной величины $\xi(h)$. Положим

$$k_h(z)=\frac{1}{h}\int\left(e^{izx}-1\right)dF_h(x)=$$

$$=\int\left(e^{izx}-1-\frac{izx}{1+x^2}\right)\frac{1+x^2}{x^2}dG_h(x)+\frac{1}{h}\int\frac{izx}{1+x^2}dF_h(x),$$

где

$$dG_h(x)=\frac{1}{h}\frac{x^2}{1+x^2}dF_h(x).$$

Поскольку для всех $z\in R$ существует предел (7), то существует и предел

$$\lim_{h\to 0}\int(\cos zx-1)\frac{1+x^2}{x^2}dG_h(x)=\operatorname{Re}k(z). \tag{8}$$

Из непрерывности $k(z)$ и формулы (6) вытекает, что предел в (7) существует равномерно по z на любом компакте. Следовательно, и предел в (8) существует равномерно по $|z|\le c$. Тогда

$$\int_R\left(1-\frac{\sin x\delta}{\delta x}\right)\frac{1+x^2}{x^2}dG_h(x)=$$

$$=\frac{1}{2\delta}\int_{-\delta}^{\delta}\int_R(1-\cos\lambda x)\frac{1+x^2}{x^2}dG_h(x)d\lambda\to\frac{-1}{2\delta}\int_{-\delta}^{\delta}\operatorname{Re}k(\lambda)d\lambda.$$

Поэтому

$$\overline{\lim_{h\to 0}}\int_{|x|>\frac{1}{\delta}}\frac{1+x^2}{x^2}dG_h(x)\le-\frac{1}{\inf\limits_{|x|>\frac{1}{\delta}}\left(1-\frac{\sin\delta x}{\delta x}\right)}\frac{1}{2\delta}\int_{-\delta}^{\delta}\operatorname{Re}k(\lambda)d\lambda.$$

Обозначим

$$\inf_{\lambda\ge\alpha}\left(1-\frac{\sin\lambda}{\lambda}\right)=\Psi(\alpha).$$

Тогда

$$\inf_{|x|>\frac{1}{\delta}}\left(1-\frac{\sin\delta x}{\delta x}\right)\ge\Psi(1).$$

Следовательно,

$$\lim_{\delta\to 0}\overline{\lim_{h\to 0}}\int_{|x|>\frac{1}{\delta}} dG_h(x) \le \Psi^{-1}(1)\lim_{\delta\to 0}\frac{1}{2\delta}\int_{-\delta}^{\delta}\operatorname{Re}k(\lambda)\,d\lambda = 0,$$

так как $k(0)=0$. Но тогда множество мер $dG_h(x)$ компактно. Пусть $dG_h(x)$ сходится к $dG(x)$. Имеем

$$\lim_{h\to 0}\int\left(e^{izx}-1-\frac{izx}{1+x^2}+\frac{1}{2}\frac{(zx)^2}{1+x^2}\right)\frac{1+x^2}{x^2}dG_h(x)=$$

$$=\int\left(e^{izx}-1-\frac{izx}{1+x^2}+\frac{1}{2}\frac{(zx)^2}{1+x^2}\right)\frac{1+x^2}{x^2}dG(x).$$

Поскольку подынтегральная функция в выражении справа равна нулю в точке 0, если доопределить ее по непрерывности, то интеграл справа не изменится, когда меру $dG(x)$ заменить на $d\overline{G}(x)$, где $d\overline{G}(A)=dG(A-\{0\})$.Поэтому существует предел

$$\lim_{h\to 0}\left[-\frac{1}{2}\int\frac{(zx)^2}{x^2}dG_h(x)+\frac{1}{h}\int\frac{izx}{1+x^2}dF_h(x)\right]=$$

$$=k(z)-\int\left(e^{izx}-1-\frac{izx}{1+x^2}+\frac{1}{2}\frac{(zx)^2}{1+x^2}\right)\frac{1+x^2}{x^2}d\overline{G}(x).$$

Но тогда существуют пределы вещественной и мнимой частей. Очевидно, что существует $b_1>0$ для которого

$$\lim_{h\to 0}\frac{1}{h}\int\frac{(zx)^2}{1+x^2}dG_h(x)=b_1z^2$$

и $a\in R$, что

$$\lim_{h\to 0}\frac{1}{h}\int\frac{izx}{1+x^2}dF_h(x)=iaz.$$

Поэтому

$$k(z)=iaz-\frac{1}{2}bz^2+\int\left(e^{ixz}-1-\frac{izx}{1+x^2}\right)\frac{1+x^2}{x^2}d\overline{G}(x),$$

где

$$bz^2=b_1z^2-\int\frac{(zx)^2}{x^2}d\overline{G}(x).$$

Теорема 1 доказана.

П р и м е р 1. Однородный процесс Пуассона.

Рассмотрим события, которые могут происходить в каждый момент непрерывно меняющегося времени. К таким событиям можно отнести регистрацию частицы счетчиком, поступление вызова на телефонную станцию, испускание ядром частицы при радиоактивном распаде и т. п. Если принять, что за конечное время происходит лишь конечное число событий, то при весьма общих предположениях число появлений события за любое время t имеет распределение Пуассона.

Пусть $\xi(t)$ – число появлений событий на промежутке $[0,t]$. Будем считать, что $\xi(t)$ определено для всех t и является неотрицательной целочисленной случайной величиной. Тогда для $t_1 < t_2$ число появлений событий на $(t_1,t_2]$ равно $\xi(t_2)-\xi(t_1)$.

Введем некоторые предположения относительно совокупности рассматриваемых событий:

1) числа событий, появившихся на непересекающихся временных промежутках, являются независимыми случайными величинами, т. е. при любых $0 < t_1 < t_2 < ... < t_n$ случайные величины $\xi(t_1)$, $\xi(t_2)-\xi(t_1),\ldots,$ $\xi(t_n)-\xi(t_{n-1})$ независимы между собой;

2) *однородность* процесса появления событий во времени, т. е. в дискретной последовательности события имеют одинаковую вероятность;

3) *ординарность* потока событий – невозможность появления одновременно нескольких событий.

Оформим условия 2) и 3) в математической форме.

Однородность процесса ξ означает, что распределение случайной величины $\xi(t+h)-\xi(t)$ зависит от h, но не зависит от t.

Найдем условие ординарности потока. Пусть B_t – событие, состоящее в том, что на отрезке $[0,t]$ произошло одновременно, по крайней мере, два события. Тогда

$$B_t = \bigcap_{n=1}^{\infty}\bigcup_{k=1}^{n}\left(\xi\left(\frac{kt}{n}\right)-\xi\left(\frac{k-1}{n}t\right)>1\right). \tag{9}$$

Из соотношения (9) вытекает, что

$$\mathrm{P}(B_t) = \lim_{n\to\infty}\mathrm{P}\left(\bigcup_{k=1}^{n}\left(\xi\left(\frac{kt}{n}\right)-\xi\left(\frac{k-1}{n}t\right)>1\right)\right) =$$

$$= \lim_{n\to\infty}\mathrm{P}\left(\bigcup_{k=1}^{n}\bigcap_{i=1}^{k-1}\left(\xi\left(\frac{it}{n}\right)-\xi\left(\frac{i-1}{n}t\right)\le 1\right)\bigcap\left(\xi\left(\frac{k}{n}t\right)-\xi\left(\frac{k-1}{n}t\right)\right)>1\right)$$

$$= \lim_{n\to\infty}\sum_{k=1}^{n}\left(\prod_{i=1}^{k-1}\mathrm{P}\left(\xi\left(\frac{it}{n}\right)-\xi\left(\frac{i-1}{n}t\right)\le 1\right)\cdot\mathrm{P}\left(\xi\left(\frac{i}{n}t\right)-\xi\left(\frac{i-1}{n}t\right)>1\right)\right).$$

Обозначим $\alpha(h)=\mathrm{P}\left(\xi(t+h)-\xi(t)>1\right)$. Тогда

$$\mathrm{P}(B_t)=\lim_{n\to\infty}\sum_{k=1}^{n}\left(1-\alpha\left(\frac{t}{n}\right)\right)^{k-1}\alpha\left(\frac{t}{n}\right)=\lim_{n\to\infty}\left(1-\left(1-\alpha\left(\frac{t}{n}\right)\right)^{n}\right).$$

Так как $\mathrm{P}(B_t)=0$, то

$$\lim_{n\to\infty}\left(1-\alpha\left(\frac{t}{n}\right)\right)^{n}=1,$$

но

$$\lim_{n\to\infty}\left(1-\alpha\left(\frac{t}{n}\right)\right)^{n}=e^{-\lim n\alpha\left(\frac{t}{n}\right)},$$

следовательно,

$$n\alpha\left(\frac{t}{n}\right)\to 0,$$

т. е.

$$\lim_{h\to 0}\frac{1}{h}\mathrm{P}\left(\xi(t+h)-\xi(t)>1\right)=0. \qquad (10)$$

Покажем теперь, что для однородного ординарного потока событий существует такое $\lambda>0$, что

$$\lim_{h\to 0}\frac{1}{h}\mathrm{P}\left(\xi(t+h)-\xi(t)>0\right)=\lambda. \qquad (11)$$

Действительно, пусть $\varphi(t)=\mathrm{P}\left(\xi(t)=0\right)$. Тогда согласно условию 1)

$$\varphi(t+h)=\mathrm{P}\left(\xi(t+h)=0\right)=\mathrm{P}\left(\left(\xi(t)=0\right)\bigcap\left(\xi(t+h)-\xi(t)=0\right)\right)=$$
$$\mathrm{P}\left(\xi(t)=0\right)\mathrm{P}\left(\xi(t+h)-\xi(t)=0\right)=\mathrm{P}\left(\xi(t)=0\right)\mathrm{P}\left(\xi(h)=0\right)=\varphi(t)\,\varphi(h).$$

Так как $\xi(0)=0$ и при $t_n\downarrow 0$

$$1=\varphi(0)=\mathrm{P}\left(\xi(0)=0\right)=\mathrm{P}\left(\bigcap_{n=1}^{\infty}\left(\xi(t_n)=0\right)\right)=\lim_{n\to\infty}\varphi(t_n),$$

то $\varphi(t)$ непрерывна в точке 0 справа. Но так как $\varphi(t+h)=\varphi(t)\cdot\varphi(h)\to\varphi(t)$ при $h\downarrow 0$, то φ непрерывна справа в любой точке t. Поэтому

$$\varphi(1)=\varphi^n\left(\frac{1}{n}\right)=\lim_{n\to\infty}\exp\left\{n\left[\varphi\left(\frac{1}{n}\right)-1\right]\right\}$$

и, следовательно, существует предел

$$\lim_{n\to\infty} n\left(1-\varphi\left(\frac{1}{n}\right)\right)=\lambda .$$

Тогда $\varphi(1)=e^{-\lambda}$ и для всех рациональных t имеем $\varphi(t)=\exp(-\lambda t)$. Из непрерывности справа φ убеждаемся, что эта формула справедлива для всех $t>0$. Из соотношения

$$\mathrm{P}\left(\xi(t+h)-\xi(t)>0\right)=\mathrm{P}\left(\xi(h)>0\right)=1-e^{-\lambda h},\ t>0,\ h>0,$$

вытекает формула (11).

Утверждение 3. *Пусть для семейства целочисленных случайных величин* $\xi(t)$ *выполнено условие* 1), *распределение величины* $\xi(t+h)-\xi(t)$ *не зависит от* t *и выполнены условия* (10) *и* (11). *Тогда*

$$\mathrm{P}\left(\xi(t)=k\right)=\frac{(\lambda t)^k}{k!}e^{-\lambda t},\ k=0,1,2,\ldots$$

Случайный процесс, рассмотренный в примере 1 называется *однородным случайным процессом Пуассона*.

Иллюстрация в Mathematica.

Упражнение 1. Докажите, что если для случайного процесса $\Pi(t)$, $t\ge 0$, выполнены следующие условия:

а) $\Pi(0)=0$ п. н.,

б) для любых $0<t_1<t_2<t_3<\ldots<t_n$ случайные величины $\Pi(t_1),\Pi(t_2)-\Pi(t_1),\ldots,\Pi(t_n)-\Pi(t_{n-1})$ независимы;

в) для любых $t>1$

$$\mathrm{P}\left(\Pi(t)-\Pi(1)=k\right)=\frac{\left[\lambda(t-1)\right]^k}{k!}e^{-\lambda(t-1)},\ \lambda>0,\ k=0,1,2,\ldots,$$

то $\Pi(t)$ – однородный случайный процесс Пуассона.

Пример 2. Обобщенный случайный процесс Пуассона.

Рассмотрим однородный процесс с независимыми приращениями, для которого выполнены следующие условия:

1) $\int\limits_R \frac{1+x^2}{x^2}d\overline{G}(x)<\infty$,

2) $b=0$,

3) $az = \int\limits_R \frac{zx}{|x|^2} d\overline{G}(x)$, $z \in R$,

где $a, b, \overline{G}$ – из формулы (4), в этом случае

$$\varphi(t,z) = \exp(tc\int\limits_R (e^{izx} - 1)dF(x)), \tag{12}$$

причем

$$c = \int\limits_R \frac{1+x^2}{x^2} d\overline{G}(x),$$

$$F(A) = \frac{1}{c}\int\limits_A \frac{1+x^2}{x^2} d\overline{G}(x)$$

есть распределение вероятностей в R. Однородный процесс с независимыми приращениями с характеристической функцией (12) называется *обобщенным случайным процессом Пуассона*. Это название объясняется следующими соображениями. Пусть мера F сосредоточена в точке 1.
Тогда

$$\varphi(t,z) = \exp\left\{tc\left(e^{iz} - 1\right)\right\}. \tag{13}$$

Следовательно, приращение $\xi(t+h) - \xi(h)$ имеет распределение Пуассона с параметром tc, так что процесс $\xi(t)$ есть однородный процесс Пуассона с параметром c.

Покажем теперь, как с помощью процесса Пуассона построить процесс с характеристической функцией (12). Обозначим через $\xi_1, \xi_2, \ldots$ последовательность независимых одинаково распределенных случайных величин с функцией распределения F, которые в совокупности независимы с процессом Пуассона $\upsilon(t)$ с характеристической функцией вида (13). Рассмотрим случайный процесс

$$\eta(t) = \sum_{k=1}^{\upsilon(t)} \xi_k \text{ , } \left(\sum_{k=1}^{0} = 0\right). \tag{14}$$

Этот процесс кусочно-постоянный, его скачки происходят в тех точках, где происходят скачки $\upsilon(t)$, при этом величина k - ого скачка равна ξ_k.

Покажем, что $\eta(t)$ – процесс с независимыми приращениями. Для этого найдем совместную характеристическую функцию величин

$$\eta(t_1), \eta(t_2) - \eta(t_1), \ldots, \eta(t_n) - \eta(t_{n-1}) \tag{15}$$

Пусть $\upsilon(t_1) = k_1, \upsilon(t_2) - \upsilon(t_1) = k_2, \ldots, \upsilon(t_n) - \upsilon(t_{n-1}) = k_n$, тогда

$$\eta(t_1) = \sum_{j=1}^{k_1} \xi_j, \eta(t_2) - \eta(t_1) = \sum_{j=k_1+1}^{k_1+k_2} \xi_j, \ldots,$$

$$\eta(t_n)-\eta(t_{n-1})=\sum_{j=k_1+\ldots+k_{n-1}+1}^{k_1+\ldots+k_n}\xi_j\,.$$

Учитывая независимость ξ_k и то, что эти случайные величины не зависят от $\upsilon(t)$, имеем

$$\mathrm{M}\Big(\exp\big(iz_1\,\eta(t_1)+\ldots+iz_n(\eta(t_n)-\eta(t_1)\big)\Big|\upsilon(t_1)=k_1,\ldots,\upsilon(t_n)-\upsilon(t_{n-1})=k_n\Big)=$$

$$=\big(\mathrm{M}\exp(iz_1,\xi_1)\big)^{k_1}\ldots\big(\mathrm{M}\exp(iz_n\xi_1)\big)^{k_n}.$$

Пусть $\psi(z)=\int\limits_{-\infty}^{\infty}e^{izx}dF(x)$ – характеристическая функция случайной величины ξ_1. Так как

$$\mathrm{P}\big(\upsilon(t_1)=k_1,\ldots,\upsilon(t_n)-\upsilon(t_{n-1})=k_n\big)=\frac{(ct_1)^{k_1}}{k_1!}e^{-ct_1}\ldots\frac{\big(c(t_n-t_{n-1})\big)^{k_n}}{k_n!}e^{-c(t_n-t_{n-1})},$$

то на основании формулы полного математического ожидания

$$\mathrm{M}\exp\big\{iz_1\eta(t_1)+\ldots+iz_n\big(\eta(t_n)-\eta(t_{n-1})\big)\big\}=$$

$$=\sum_{k_1,\ldots,k_n=0}^{\infty}\frac{(ct_1)^{k_1}}{k_1!}\psi^{k_1}(z_1)e^{-ct_1}\frac{\big[c(t_2-t_1)\big]^{k_2}\psi^{k_2}(z_2)e^{-c(t_2-t_1)}}{k_2!}\times\ldots$$

$$\ldots\times\frac{\big[c(t_n-t_{n-1})\big]^{k_n}\psi^{k_n}(z_n)e^{-c(t_n-t_{n-1})}}{k_n!}=\exp\big(ct_1\big(\psi(z_1)-1\big)\big)\times$$

$$\times\exp\big(c(t_2-t_1)\big(\psi(z_2)-1\big)\big)\ldots\exp\big(c(t_n-t_{n-1})\big(\psi(z_n)-1\big)\big).$$

Из этой формулы вытекает, что $\eta(t)$ – процесс с независимыми приращениями и что характеристическая функция величины $\eta(t_2)-\eta(t_1)$ задается выражением

$$\exp\big\{c(t_2-t_1)\big(\psi(z)-1\big)\big\},$$

т. е. распределение этой величины зависит лишь от t_2-t_1. Следовательно, процесс $\eta(t)$ – однородный случайный процесс с независимыми приращениями и характеристической функцией

$$\mathrm{M}\exp\big(iz\eta(t)\big)=\exp\big(ct\big(\psi(z)-1\big)\big)=\exp\left(ct\cdot\int\limits_R\big(e^{izx}-1\big)dF(x)\right).$$

У т в е р ж д е н и е 4. *Для любого обобщенного случайного процесса Пуассона существует стохастически эквивалентный случайный процесс вида* (14).

П р и м е р 3. С л у ч а й н ы й   п р о ц е с с   б р о у н о в с к о г о   д в и ж е н и я.

Рассмотрим частицу, которая движется в некоторой среде (жидкости или газе) под влиянием взаимодействий с частицами среды. Если среда однородна, то движение частицы из каждой точки среды в какой бы момент мы туда не попали, будет одинаковым в вероятностном смысле, т. е. смещение частицы $\xi(t+h)-\xi(t)$ за время h не должно зависеть ни от поведения до момента t, ни от момента t. Независимость движения после момента t от скорости в момент t противоречит, на первый взгляд, законам механики, однако на самом деле за очень малое время скорость так часто меняет свою величину, а движение – направление, что частица успевает «забыть» значение скорости в исходный момент времени. Очевидно также, что траектория двигающейся частицы $\xi(t)$ – непрерывная функция t. Непрерывный процесс с независимыми приращениями, на основании приведенных соображений, называется *процессом броуновского движения*. Мы рассмотрим однородное во времени броуновское движение, т. е. однородный непрерывный процесс с независимыми приращениями.

Иллюстрация в Mathematica типичной траектории блуждающей частицы на плоскости.

Естественно попытаться построить непрерывный процесс из обобщенного процесса Пуассона с помощью предельного перехода, при котором число скачков процесса будет расти, а сами скачки будут уменьшаться в размере. Пусть $\upsilon_c(t)$ – пуассоновский процесс с параметром c, а $\xi_1^{\,c},\ldots,\xi_n^c,\ldots$ – последовательность независимых одинаково распределенных случайных величин, не зависящих от $\upsilon_c^{(t)}$ и имеющих распределение F^c. Положим

$$\xi_c(t)=\sum_{k=1}^{\upsilon_c(t)}\xi_k^c$$

Исследуем поведение этого процесса при $c\to\infty$ (c – среднее число скачков в единицу времени). Величины ξ_k^c при этом стремятся к 0.

Характеристическая функция процесса $\xi_c(t)$ имеет вид

$$\varphi_c(t,z)=\exp\left(ct\int_R\left(e^{izx}-1\right)dF^c(x)\right)$$

Используя соответствующие свойства характеристических функций, можно убедиться (см., также [10, с. 273]), что при $c\to\infty$ характеристическая функция предельного процесса имеет вид

$$\varphi(t,z)=\exp\left\{t\left(iaz-\frac{1}{2}bz^2\right)\right\},\ a\in R,\ b>0. \qquad (16)$$

Соображения, изложенные в примере 3, справедливы для любого непрерывного процесса, что подтверждает

Теорема 2. *Если* $\xi(t)$ *– непрерывный однородный случайный процесс с непрерывными траекториями и независимыми приращениями, то его характеристическая функция имеет вид (16).*

Доказательство. Покажем сначала, что для любого $\varepsilon > 0$

$$\lim_{h\to 0}\frac{1}{h}\mathrm{P}\left(\left|\xi(h)\right| > \varepsilon\right) = 0 \tag{17}$$

Из непрерывности траекторий $\xi(t)$ вытекает их равномерная непрерывность на любом конечном промежутке. Следовательно,

$$\max_{kh<1}\left|\xi(kh)-\xi\left((k-1)h\right)\right| \to 0$$

с вероятностью 1 при $h\to 0$. Тогда для любого $\varepsilon > 0$

$$1 = \lim_{h\to 0}\mathrm{P}\left(\max_{kh<1}\left|\xi(kh)-\xi\left((k-1)h\right)\right| \le \varepsilon\right) = \lim_{h\to 0}\left(\mathrm{P}\left(\left|\xi(h)\right| \le \varepsilon\right)\right)^{\left[\frac{1}{h}\right]}$$

(здесь $\left[\frac{1}{h}\right]$ – целая часть), но

$$\lim_{h\to 0}\left(\mathrm{P}\left(\left|\xi(h)\right| \le \varepsilon\right)\right)^{\left[\frac{1}{h}\right]} = \lim_{h\to 0}\left(1-\mathrm{P}\left(\left|\xi(h)\right| > \varepsilon\right)\right)^{\left[\frac{1}{h}\right]} =$$

$$= \exp\left\{-\lim_{h\to 0}\left[\frac{1}{h}\right]\mathrm{P}\left(\left|\xi(h)\right| > \varepsilon\right)\right\} = \exp\left\{-\lim_{h\to 0}\frac{1}{h}\mathrm{P}\left(\left|\xi(h)\right| > \varepsilon\right)\right\}.$$

Поскольку это выражение равно 1, то предел под знаком exp равен 0, т. е. справедливо (17).

Как было показано в теореме 1, мера $\overline{G}$ имеет вид

$$\overline{G}(A) = G\left(A-\{0\}\right),$$

где G – слабый предел меры G_h при $h\to 0$. Из (17) вытекает, что для всех $A\subset\left\{x : |x| > \varepsilon\right\}$

$$\lim_{h\to 0}\int_A dG_h(x) = 0.$$

Следовательно, $\overline{G}(A) = 0$ для любого множества A, не содержащего точки 0. Кроме того, по построению $\overline{G}\left(\{0\}\right) = 0$. Таким образом, $\overline{G}(A) = 0$ для всех A и из теоремы 1 вытекает доказательство теоремы 2.

Однородный случайный процесс с независимыми приращениями, имеющий характеристическую функцию вида (16) мы будем называть *процессом броуновского движения*. Если в формуле (16) $a=0$, $b=1$, то случайный процесс с такой характеристической функцией назовем *стандартным процессом броуновского движения*.

Иллюстрация в Mathematica.

У п р а ж н е н и е 2. Докажите, что если случайный процесс $W(t)$, $t \ge 0$ удовлетворяет следующим условиям:

1) $W(0)=0$ п. н.;

2) для любых $0<t_1<\ldots<t_n$ случайные величины $W(t_1),\ W(t_2)-W(t_1),\ldots,W(t_n)-W(t_{n-1})$ независимы;

3) для любых $t>s$

$$W(t)-W(s)\in N(0,t-s),$$

то он является стандартным процессом броуновского движения.

Теорема 3. *Процесс броуновского движения имеет непрерывную модификацию.*

Д о к а з а т е л ь с т в о непосредственно вытекает из теоремы 1 предыдущего параграфа.

§ 3. Корреляционная теория случайных процессов

В настоящем параграфе будут рассмотрены случайные процессы, которые принимают комплексные значения, причем

$$\mathrm{M}\left|\xi(t)\right|^2<\infty,\ \forall t\in[a;b].$$

Значения процесса $\xi(t)$ в момент времени t мы интерпретируем, как элемент пространства $\mathcal{L}_2=\mathcal{L}_2(\Omega,\mathcal{A},\mathrm{P})$. С этой точки зрения процесс $\xi(t),t\in T$, можно рассматривать как некоторую кривую в пространстве $\mathcal{L}_2$.

О п р е д е л е н и е 1. *Функция*

$$B(t,s)=\mathrm{M}\xi(t)\overline{\xi(s)}=\left(\xi(t),\overline{\xi(s)}\right),\ t,s\in T,$$

называется ковариацией случайного процесса $\xi(t),t\in T$.

У т в е р ж д е н и е 1. *Ковариация случайного процесса* $\xi(t)$ *обладает следующими свойствами:*

1) $B(t,t)=\mathrm{M}\left|\xi(t)\right|^2\ge 0,\ t\in T;\ \ B(t,s)=\overline{B(s,t)},\ t,s\in T;$

2) $\left|B(t,s)\right|^2\le B(t,t)B(s,s),\ t,s\in T;$

3) *B – положительно определенная функция, т. е. для любых комплексных чисел* $c_1,\dots,c_n$ *и любых* $t_1,\dots,t_n \in T$

$$\sum_{k,r=1}^{n} c_k \overline{c}_r B(t_k,t_r) \geq 0.$$

Доказательство второго пункта вытекает из неравенства Коши – Буняковского, а для доказательства третьего пункта заметим, что

$$\sum_{k,r=1}^{n} c_k \overline{c}_n B(t_k,t_r) = \mathrm{M}\left(\sum_{k=1}^{n} c_k \xi(t_k) \overline{\sum_{r=1}^{n} c_r \xi(t_r)}\right).$$

Определение 2. *Корреляционной функцией случайного процесса* $\xi(t), t \in T$, *называют функцию*

$$R(t,s) = \mathrm{M}\big(\xi(t) - \mathrm{M}\xi(t)\big)\overline{\big(\xi(t) - \mathrm{M}\xi(t)\big)}.$$

Определение 3. *Случайный процесс* $\xi(t), t \in T$, *называют*

а) непрерывным в среднеквадратическом в точке t, *если*

$$\underset{|h|\to 0}{\mathrm{l.i.m.}}\, \xi(t+h) = \xi(t),$$

т. е. $\lim\limits_{|h|\to 0} \mathrm{M}\left|\xi(t+h) - \xi(t)\right|^2 = 0.$

б) дифференцируемым в среднеквадратическом в точке t, *если существует предел*

$$\underset{|h|\to 0}{\mathrm{l.i.m.}}\, \frac{\xi(t+h) - \xi(t)}{h} = \xi'(t).$$

Если процесс непрерывен (дифференцируем) в среднеквадратичном в каждой точке $t \in T$, то его называют *непрерывным (дифференцируемым) в среднеквадратическом случайным процессом.*

Упражнение 1. Докажите, что

1) для непрерывности в среднеквадратическом процесса $\xi(t)$ в точке t необходимо и достаточно, чтобы

$$\Box B = \Box_{(t',t')} B(t,t) = B(t',t') - B(t',t) - B(t,t') + B(t,t) \to 0$$

при $t' \to 0$;

2) для среднеквадратической дифференцируемости процесса $\xi(t)$ в точке t необходимо и достаточно, чтобы существовал предел

$$\underset{t'\to t,\, t''\to t}{\mathrm{l.i.m.}}\, \frac{\Box_{(t',t'')} B(t,t)}{(t'-t)(t''-t)} = \left.\frac{\partial^2 B(t,s)}{\partial t \partial s}\right|_{s=t},$$

где $\square_{(t',t'')} B(t,t) = B(t',t'') - B(t',t) - B(t,t'') + B(t,t)$.

Определение 4. *Взаимной ковариацией (корреляционной функцией) двух процессов* $\xi(t)$ *и* $\eta(t), t \in T$, *называют функцию*

$$B_{\xi\eta}(t,s) = \mathrm{M}\xi(t)\overline{\eta(s)} \quad \left(R_{\xi\eta}(t,s) = \mathrm{M}\big(\xi(t) - \mathrm{M}\xi(t)\big)\overline{\big(\eta(t) - \mathrm{M}\eta(t)\big)}\right).$$

Определение 5. *Случайный процесс* $\xi(t)$, $t \in T$, *называют процессом с ортогональными приращениями, если для любых* $t_1 < t_2 \le t_3 < t_4$ *из* T

$$\mathrm{M}\big(\xi(t_4) - \xi(t_3)\big)\overline{\big(\xi(t_2) - \xi(t_1)\big)} = 0.$$

Если $\xi_1(t) = \xi(t) - \mathrm{M}\xi(t)$ – процесс с ортогональными приращениями, то случайный процесс $\xi(t)$, $t \in T$, называют *процессом с некоррелированными приращениями.*

Пусть $\xi(t)$, $t \in T$, – процесс с ортогональными приращениями и для некоторого $t_0 \in T$ $\xi(t_0) \equiv 0$. Тогда для $t_0 < s < t$ из T

$$B(t,s) = \mathrm{M}\xi(t)\overline{\xi(s)} = \mathrm{M}\big(\xi(t) - \xi(s)\big)\overline{\big(\xi(s) - \xi(t_0)\big)} +$$

$$+\mathrm{M}\left|\xi(s)\right|^2 = \mathrm{M}\left|\xi(s)\right|^2 = B(s,s).$$

Отсюда, учитывая, что

$$B(t,s) = \overline{B(s,t)},$$

получим следующее выражение для ковариации $B(t,s)$, $t \ge t_0$, $s \ge s_0$

$$B(t,s) = B\big(\min(t,s)\big), \text{ где } B(t) = B(t,t) = \mathrm{M}\left|\xi(t)\right|^2.$$

Заметим, что функция $B(t)$ монотонно не убывает. Действительно, пусть $t > s > t_0$. Тогда, в силу ортогональности процесса $\xi(t)$

$$B(t) = \mathrm{M}\left|\xi(t)\right|^2 = \mathrm{M}\left|\xi(t) - \xi(s) + \xi(s) - \xi(t_0)\right|^2 =$$

$$= \mathrm{M}\left|\xi(t) - \xi(s)\right|^2 + \mathrm{M}\left|\xi(s)\right|^2 = \mathrm{M}\left|\xi(t) - \xi(s)\right|^2 + B(s),$$

т. е. $B(t) \ge B(s)$.

Замечание 1. Несложно видеть, что процесс с ортогональными приращениями $\xi(t), t \in T$, будет непрерывным в среднеквадратическом тогда и только тогда, когда непрерывна функция $B(t)$, $t \in T$. Кроме того, процесс $\xi(t)$ в среднеквадратическом дифференцируем в точке t тогда и только тогда, когда $B'(t) = 0$, так что, вообще говоря, процесс с ортогональными приращениями не дифференцируем в среднеквадратическом.

Упражнение 2. Выясните, являются ли дифференцируемыми в среднеквадратическом процессы Пуассона и броуновского движения.

Интегрирование случайных процессов

Во многих вопросах бывает нужным определить понятие интеграла.

Определение 6. *Интегралом*

$$\int_a^b \xi(t)\,dt \tag{1}$$

от процесса $\xi(t)$, $t \in T = [a;b]$, *называют предел в среднеквадратическом* l.i.m. *при* $\lambda \to 0$ *интегральных сумм*

$$S'_\lambda = \sum_{k=1}^{n} \xi(t_k)\cdot \Delta t_k\,,$$

где $a = t_0 < t_1 < \ldots < t_n = b,\ \Delta t_k = t_k - t_{k-1},\ \lambda = \max\limits_{1\le k\le n} \Delta t_k$.

Замечание 2. Из определения и свойств ковариации случайного процесса непосредственно вытекают следующие утверждения:

1) *для существования интеграла* (1) *достаточно, чтобы функция* $B(t,s)$ *была интегрируема по Риману на квадрате* $a \le s, t \le b$;

2) $\mathrm{M}\left[\int_a^b \xi(t)\,dt\right]^2 = \int_a^b\int_a^b B(t,s)\,dtds$;

3) *если*

$$\eta(t) = \int_a^t \xi(u)\,du, t \in T, \tag{2}$$

то

$$B_\eta(t,s) = \int_a^t\int_a^s B_\xi(u,v)\,dudv.$$

Утверждение 2. *Пусть случайные процессы* $\xi(t)$ *и* $\eta(t)$, $t \in T$, *связаны соотношением* (2). *В каждой точке* t, *в которой функция* $B_\xi(t,t)$ *непрерывна, процесс* $\eta(t)$ *дифференцируем в среднеквадратическом и* $\eta'(t) = \xi(t)$.

Доказательство. Имеем

$$\delta(t,h)=\frac{\eta(t+h)-\eta(t)}{h}-\xi(t)=\frac{1}{h}\int\limits_t^{t+h}\left[\xi(u)-\xi(t)\right]du.$$

Поэтому по неравенству Коши – Буняковского

$$\mathrm{M}\left|\delta(t,h)\right|^2=\frac{1}{h^2}\left|\int\limits_t^{t+h}\int\limits_t^{t+h}\mathrm{M}\Big(\big(\xi(u)-\xi(t)\big)\overline{\big(\xi(v)-\xi(t)\big)}\Big)dudv\right|\le$$

$$\le\frac{1}{h^2}\left(\int\limits_t^{t+h}\left(\mathrm{M}\left|\xi(u)-\xi(t)\right|^2\right)^{1/2}du\right)^2\le\frac{1}{h}\int\limits_t^{t+h}\mathrm{M}\left|\xi(u)-\xi(t)\right|^2du.$$

Из непрерывности функции $B_\xi(t,t)$ в точке t вытекает, что при $u\to t$

$$\mathrm{M}\left|\xi(u)-\xi(t)\right|^2\to 0,$$

следовательно,

$$\lim_{h\to 0}\mathrm{M}\left|\delta(t,h)\right|^2=0.$$

Утверждение 2 доказано.

Далее введем в рассмотрение интеграл

$$\mathcal{I}(f)=\int\limits_a^b f(t)\,d\xi(t),$$

где f – неслучайная функция, $\xi(t)$, $t\in T=[a;b]$, – процесс с ортогональными приращениями. Потребуем, чтобы он обладал основными свойствами интегралов Римана – Стилтьеса или Лебега – Стилтьеса, а именно:

1) если f_1, f_2 – интегрируемые функции, а c_1, c_2 – константы, то $c_1f_1+c_2f_2$ – интегрируемая функция и

$$\mathcal{I}(c_1f_1+c_2f_2)=c_1\mathcal{I}(f_1)+c_2\mathcal{I}(f_2),$$

т. е. $\mathcal{I}$ – линейный функционал;

2) если $f(t)=I_{[c,d]}(t)$ и точки c, d – точки непрерывности $\xi(t)$, то

$$\mathcal{I}\left(I_{[c,d]}\right)=\xi(d)-\xi(c);$$

3) при определенных условиях можно переходить к пределу под знаком интеграла, т. е. если f_n – некоторая сходящаяся последовательность интегрируемых функций, то при некоторых предположениях

$$\lim_{n\to\infty} \mathscr{I}(f_n) = \mathscr{I}\left(\lim_{n\to\infty} f_n\right).$$

Для любых $c,\ d\ (a \le c < d < b)$ положим по определению

$$F(c,d) = \mathrm{M}\left|\xi(d) - \xi(c)\right|^2.$$

Несложно видеть, что эта функция аддитивна, т. е. если $c_1 < c_2 < c_3$, то $F(c_1,c_3) = F(c_1,c_2) + F(c_2,c_3)$.

Поэтому $F(c,d)$ как функция от $d,\ d > c$, монотонно не убывает, а как функция от $c,\ c < d$, монотонно не возрастает. Следовательно, множество точек разрыва функции $F(c,d)$ по каждому аргументу в отдельности не более чем счетно, и односторонние пределы $F(c\pm, d\pm)$ существуют при любых c и d. Отсюда вытекает, что при каждом $t \in (a,b)$ существуют пределы в среднеквадратическом

$$\xi(t-) = \underset{s\uparrow t}{\mathrm{l.i.m.}}\,\xi(s), \quad \xi(t+) = \underset{s\downarrow t}{\mathrm{l.i.m.}}\,\xi(s),$$

причем процессы $\xi(t-)$ и $\xi(t+)$ являются процессами с ортогональными приращениями. При этом процессы $\xi(t-)$ и $\xi(t+)$ отличаются от $\xi(t)$ не более чем на счетном множестве значений и процесс $\xi(t+)$ ($\xi(t-)$) непрерывен в среднеквадратическом справа (слева). Поэтому можно считать, что процесс $\xi(t)$ непрерывен в среднеквадратическом, например, справа. Тогда и функция $F(c,d)$ будет непрерывна по каждому из своих аргументов. В дальнейшем будем считать процесс $\xi(t)$ определенным на R (если процесс $\xi(t)$ определен на (a,b) то его естественно можно продолжить на $(-\infty,+\infty)$, положив $\xi(t) = \xi(a+)$ при $t \le a$ и $\xi(t) = \xi(b-)$ при $t \ge b$ и при этом сформулированные ранее предположения остаются в силе), а функцию $F(c,d)$ – конечной на каждом конечном интервале. Вместо $F(c,d)$ будем писать $F(c,d]$.

Назовем функцию $f(t)$ *простой*, если

$$f(t) = \sum_{k=1}^{n} c_k I_{\Delta_k}(t),$$

где $\Delta_k = (a_k, b_k],\ k = \overline{1,n}$, – не пересекаются, $\coprod_{k=1}^{n} \Delta_k$ – конечный интервал, вне которого $f(t) \equiv 0$.

Положим по определению (в соответствии с общими свойствами 1), 2) интеграла)

$$\mathcal{I}(f)=\int_{-\infty}^{+\infty} f(t)\,d\xi(t)=\sum_{k=1}^{n} c_k\left[\xi(b_k)-\xi(a_k)\right]. \tag{3}$$

Очевидно, что это определение корректно, причем

$$\mathrm{M}\mathcal{I}(f_1)\overline{\mathcal{I}(f_2)}=\sum_{k=1}^{n} c_k^1\overline{c_k^2}F(a_k,b_k], \tag{4}$$

так как для любых простых функций f_1 и f_2 полуинтервалы Δ_k можно сделать одинаковыми.

По непрерывной справа аддитивной функции $F(a,b]$ можно построить на R конечную или σ- конечную меру $F(\cdot)$, определенную на пополнении σ- алгебры борелевских множеств на R и совпадающую с $F(a,b]$ на полуинтервалах $(a,b]$.

Равенство (4) можно записать следующим образом:

$$\mathrm{M}\mathcal{I}(f_1)\overline{\mathcal{I}(f_2)}=\int_{-\infty}^{+\infty} f_1(t)\overline{f_2(t)}dF(t). \tag{5}$$

Обозначим через $\mathcal{L}_0^{\Delta\xi}$ множество всех случайных величин вида

$$\sum_{k=1}^{n} c_k\left[\xi(b_k)-\xi(a_k)\right], -\infty<a_1<b_1\le a_2<b_2\ldots<b_n<\infty,$$

а через $\mathcal{L}_0$ – множество всех простых функций на R. Пространство $\mathcal{L}_0^{\Delta\xi}$ является линейным подпространством $\mathcal{L}_2$. Обозначим его замыкание в $\mathcal{L}_2$ через $\mathcal{L}_2^{\Delta\xi}$; оно является гильбертовым пространством со скалярным произведением $(\xi,\eta)=\mathrm{M}\xi\overline{\eta}$.

Кроме того, рассмотрим замыкание линейного пространства $\mathcal{L}_0$ по норме

$$\|f\|_F=\left(\int_{-\infty}^{+\infty}|f(t)|^2\,dF(t)\right)^{1/2};$$

обозначим его $\mathcal{L}_2^F$. Как известно, $\mathcal{L}_2^F$ состоит из всех $\mathcal{B}_F$- измеримых функций $f(t)$, для которых

$$\int_{-\infty}^{+\infty}|f(t)|^2\,dF(t)<\infty,$$

где $\mathscr{B}_F$ – пополнение относительно меры F борелевской σ- алгебры $\mathscr{B}(R)$, и также является гильбертовым пространством. Формула (3) устанавливает линейное соответствие между $\mathscr{L}_0$ и $\mathscr{L}_0^{\Delta\xi}$. Оно взаимно однозначно. Действительно, из (5) следует равенство

$$\mathrm{M}\left|\mathscr{I}(f_1)-\mathscr{I}(f_2)\right| = \int_{-\infty}^{+\infty}\left|f_1(t)-f_2(t)\right|^2 dF(t). \tag{6}$$

Соответствие $f \to \mathscr{I}(f)$ по формуле (6) изометрично. Следовательно, это соответствие можно по непрерывности продолжить до взаимно однозначного отображения $\mathscr{L}_2^F$ в $\mathscr{L}_2^{\Delta\xi}$. Возьмем произвольную функцию $f \in \mathscr{L}_2^F$ и рассмотрим последовательность $f_n \in \mathscr{L}_0$ такую, что при $n \to \infty$

$$\left\|f - f_n\right\|_F \to 0.$$

Тогда последовательность f_n фундаментальна в $\mathscr{L}_2^F$ и такой же будет последовательность $\eta_n = \mathscr{I}(f_n)$ в $\mathscr{L}_2^{\Delta\xi}$. Положим

$$\mathscr{I}(f) = \underset{n\to\infty}{\mathrm{l.i.m.}}\,\mathscr{I}(f_n).$$

Этот предел не зависти от выбора последовательностей f_n, следовательно, корректно

Определение 7. *Случайную величину $\mathscr{I}(f)\left(f \in \mathscr{L}_2^F\right)$ называют стохастическим интегралом и пишут*

$$\mathscr{I}(f) = \int_{-\infty}^{+\infty} f(t)\,d\xi(t). \tag{7}$$

Утверждение 3. *Стохастический интеграл* (7) *обладает следующими свойствами:*

a) $\int_{-\infty}^{+\infty} I_{(a,b]}(t)\,d\xi(t) = \xi(b)-\xi(a)$;

b) $\int_{-\infty}^{+\infty}\left(\alpha_1 f_1(t)+\alpha_2 f_2(t)\right)d\xi(t) = \alpha_1\int_{-\infty}^{+\infty} f_1(t)\,d\xi(t)+\alpha_2\int_{-\infty}^{+\infty} f_2(t)\,d\xi(t)$;

c) $\mathrm{M}\int_{-\infty}^{+\infty} f_1(t)\,d\xi(t)\overline{\int_{-\infty}^{+\infty} f_2(t)\,d\xi(t)} = \int_{-\infty}^{+\infty} f_1(t)\overline{f_2(t)}\,dF(t)$;

d) *если* $\left\|f - f_n\right\|_F \to 0$ *при* $n \to \infty$, *то*

$$\mathrm{l.i.m.}\,\mathscr{I}(f_n) = \mathscr{I}(f)$$

и устанавливает взаимно однозначное соответствие между пространствами $\mathscr{L}_2^F$ *и* $\mathscr{L}_2^{\Delta\xi}$.

Стационарные случайные процессы

Важный класс случайных образуют *стационарные случайные процессы*. Так называют процессы, теоретико-вероятностные характеристики которых не меняются со временем. Можно еще сказать, что стационарные процессы – это процессы, которые протекают в не изменяющихся со временем условиях. Более точно это означает следующее.

Определение 8. *Случайный процесс* $\xi(t), t \in T$, *называется стационарным, если для любого* $n \in N$ *и любых* $t, t_1, \ldots, t_n$ *таких, что* $t + t_k \in T$, *совместное распределение случайных величин* $\xi(t+t_1), \ldots,$ $\xi(t+t_n)$ *не зависит от* t.

В тех задачах, где имеют значение только моменты первого и второго порядков рассматриваемых процессов, условие стационарности можно ослабить.

Определение 9. *Случайный процесс* $\xi(t), t \in T$, *называют слабо стационарным (стационарным в широком смысле, стационарным второго порядка) процессом, если*

$$\mathrm{M}\xi(t) = m = const,\ \mathrm{M}\big(\xi(t)-m\big)\overline{\big(\xi(s)-m\big)} = R(t-s).$$

Функцию $R(t)$ *называют корреляционной функцией процесса* $\xi(t), t \in T$.

Пример 1. Колебания со случайными параметрами. Во многих вопросах теории колебания рассматривают процессы вида

$$\xi(t) = \sum_{k=1}^{n} \gamma_k e^{iu_k t}. \tag{8}$$

Совокупность частот $u_k, k = \overline{1,n}$, рассматриваемых как множество точек на R, называют *спектром колебания процесса* $\xi(t)$.

Введем некоторые термины, используемые при физической интерпретации случайных процессов. Если $|\xi(t)|$ обозначает силу тока в момент времени t и имеется в виду энергия, рассеиваемая этим током на единичном сопротивлении, то естественными являются следующие определения.

Энергией, переносимой случайным процессом $\xi(t)$ в течение промежутка времени (t_1, t_2) называется величина

$$\int_{t_1}^{t_2} \left|\xi(t)\right|^2 dt,$$

а *средней мощностью* случайного процесса называется предел

$$\lim_{T\to\infty} \int_{-T}^{T} \left|\xi(t)\right|^2 dt,$$

если он существует.

Нетрудно вычислить среднюю мощность, переносимую случайным процессом (8). Имеем

$$\frac{1}{2T}\int_{-T}^{T} \left|\xi(t)\right|^2 dt = \frac{1}{2T}\int_{-T}^{T} \sum_{k,r=1}^{n} \gamma_k \overline{\gamma}_r e^{it(u_k-u_r)} dt =$$

$$= \sum_{k=1}^{n} \left|\gamma_k\right|^2 + \sum_{\substack{k,r=1\\k\neq r}}^{n} \gamma_k \overline{\gamma}_r \frac{\sin T(u_k-u_r)}{T(u_k-u_r)} \xrightarrow[T\to\infty]{} \sum_{k=1}^{n} \left|\gamma_k\right|^2 .$$

В дальнейшем будем считать комплексные амплитуды γ_k случайными величинами. Тогда $\xi(t)$ – случайный комплекснозначный процесс, $t \in R$.

Найти распределение $\xi(t)$ по известным распределениям γ_k сложно, но вычислить моменты первых двух порядков нетрудно. Пусть

$$\mathrm{M}\gamma_k = 0,\ \mathrm{M}\gamma_k \overline{\gamma}_r = 0,\ k \neq r,\ \mathrm{M}\left|\gamma_k\right|^2 = c_k^2, \tag{9}$$

тогда

$$\mathrm{M}\xi(t) = 0,\ R(t_1,t_2) = \sum_{k=1}^{n} c_k^2 e^{iu_k(t_1-t_2)} .$$

Таким образом,

$$R(t_1,t_2) = R(t_1-t_2),$$

где

$$R(t) = \sum_{k=1}^{n} c_k^2 e^{iu_k t}. \tag{10}$$

Следовательно, при выполнении условий (9), $\xi(t)$ – стационарный в широком смысле случайный процесс, и его корреляционная функция задается формулой (10). Эту формулу называют *спектральным представлением корреляционной функции*. Она определяет спектр случайного процесса, т. е. совокупность частот u_k, $k = \overline{1,n}$, гармонических колебаний, составляющих процесс $\xi(t)$ и математические ожидания c_k^2 средних мощностей.

Назовем спектральной функцией $F(u)$ процесса (8) функцию

$$F(u) = \sum_{\substack{k\\u_k\le u}} c_k^2 .$$

С помощью спектральной функции корреляционная функция запишется в виде

$$R(t)=\int_{-\infty}^{+\infty} e^{iut}\,dF(u). \qquad (11)$$

П р и м е р 2. Пусть $F(x), x\in R,$ – произвольная монотонно неубывающая непрерывная справа функция, $F(-\infty)=0, F(+\infty)=\sigma^2>0$. Введем случайную величину ξ с функцией распределения $\frac{1}{\sigma^2}F(x)$ и определим случайный процесс

$$\xi(t)=\sigma\exp\{i(t\xi+\varphi)\},$$

где φ – равномерно распределенная на $[0,2\pi]$ случайная величина, независимая с ξ. Тогда

$$\mathrm{M}\xi(t)=\int_{-\infty}^{\infty}\int_{0}^{2\pi}\sigma e^{i(tu+\tau)}\frac{1}{\sigma^2}dF(u)\frac{d\tau}{2\pi}=0,$$

$$R(t_1,t_2)=\mathrm{M}\xi(t_1)\overline{\xi(t_2)}=\sigma^2\int_{-\infty}^{\infty}e^{i(t_1-t_2)u}\frac{1}{\sigma^2}dF(u)=R(t_1-t_2),$$

где $R(t)$ имеет вид (11).

Теорема 1 (Хинчин). *Для того, чтобы функция $R(t)$ была корреляционной функцией непрерывного в среднеквадратическом слабо стационарного процесса необходимо и достаточно, чтобы она допускала представление* (11)*, где F – ограниченная монотонно неубывающая непрерывная справа функция.*

Д о к а з а т е л ь с т в о. Н е о б х о д и м о с т ь. Из утверждения 1 и упражнения 1 вытекает, что $R(t)$ положительно определенная функция, следовательно она по теореме Бохнера – Хинчина (см. гл. 5 § 1), с точностью до множителя, является характеристической функцией некоторого распределения, т. е. имеет место представление (11).

Д о с т а т о ч н о с т ь. Если функция $R(t)$ имеет вид (11), то она является корреляционной функцией некоторого слабо стационарного процесса, например, процесса из примера 2, который непрерывен в среднеквадратическом в силу упражнения 1.

Теорема 1 доказана.

З а м е ч а н и е 3. Функция $F(u)$, $u\in R$, в формуле (11) определяется через функцию $R(t)$ однозначно (см. теорему единственности из § 2 гл. 5).

О п р е д е л е н и е 10. *Функцию $F(u)$, $u\in R$, из формулы* (11) *называют спектральной функцией слабо стационарного процесса.*

Если $F(u)$, $u\in R$, – абсолютно непрерывная функция, т. е.

$$F(u)=\int\limits_{-\infty}^{u} f(v)\,dv,$$

то функцию $f(u)$, $u\in R$, *называют спектральной плотностью процесса.*

З а м е ч а н и е 4. Если

$$\int\limits_{-\infty}^{\infty}\left|R(t)\right|dt<\infty,$$

то (см., напр., гл.5, § 2),

$$f(u)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{-iut}R(t)\,dt,$$

т. е. спектральная плотность совпадает с обратным преобразованием Фурье корреляционной функции.

З а м е ч а н и е 5. Понятие стационарности и слабой стационарности можно, очевидным образом, применить и к последовательности случайных величин.

Так последовательность $\xi_1,\xi_2,\ldots$ называют *слабо стационарной*, если

$$\mathrm{M}\xi_k=m,\quad \mathrm{M}\left(\xi_{k+n}-m\right)\overline{\left(\xi_k-m\right)}=R_n,\ k=1,2,\ldots;\ n=1,2,\ldots$$

В этом случае последовательность чисел $R_n, n\geq 1,$ называют *корреляционной функцией* исходной последовательности. Несложно показать (см., также, теорему 12, §7, гл. VI из [10]), что

$$R_n=\int\limits_{0}^{2\pi} e^{inu}\,dF(u),$$

где F – непрерывная справа монотонно неубывающая ограниченная функция. Ее называют *спектральной функцией*, а ее производную (если она существует) – *спектральной плотностью* исходной последовательности.

Теорема 3. [10]. *Пусть* $\xi(t), t\in R,$ *– слабо стационарный случайный процесс со спектральной функцией* $F(u)$, $u\in R$, $M\xi(t)\equiv 0$. *Тогда существует процесс с ортогональными приращениями* $\varsigma(u)$, $u\in R$, *такой, что*

$$\mathrm{M}\left|\varsigma(u)\right|^2=F(u),\ u\in R,$$

и

$$\xi(t) = \int_{-\infty}^{\infty} e^{iut} d\varsigma(u). \tag{12}$$

Формулу (12) называют *спектральным разложением* процесса $\xi(t)$, а функцию $\varsigma(u,v) = \varsigma(v) - \varsigma(u)$, $u < v$, – *стохастической спектральной мерой* процесса $\xi(t), t \in R$.

§ 4. Марковские случайные процессы

Марковские процессы с дискретным временем (цепи Маркова) рассматривались в главе 2. Напомним, что главным в определении таких процессов было свойство независимости «будущего» процесса от его «прошлого», если фиксировано его «настоящее». Этот же принцип лежит в основе определения марковского процесса и в общем случае.

Пусть $(\Omega, \mathcal{A}, \mathrm{P})$ – вероятностное пространство, на котором задан случайный процесс $\xi(t), t \geq 0$.

Обозначим

$$\mathcal{F}_t = \sigma\{\xi(u); u \leq t\}, \mathcal{F}_{[t,\infty)} = \sigma\{\xi(u); u \geq t\},$$

так что величина $\xi(u)$ измерима относительно $\mathcal{F}_{\square}$ при $u \leq t$ и относительно $\mathcal{F}_{[t,\infty)}$ при $u \geq t$, σ-алгебра $\sigma\left\{\mathcal{F}_t, \mathcal{F}_{[t,\infty)}\right\}$ может совпадать с $\square$.

Определение 1. *Случайный процесс* $\xi(t), t \geq 0$, *называется марковским, если для любых* $t \geq 0$, $A \in \mathcal{F}_t$, $B \in \mathcal{F}_{[t,\infty)}$

$$\mathrm{P}\left(AB \middle| \xi(t)\right) = \mathrm{P}\left(A \middle| \xi(t)\right) \cdot \mathrm{P}\left(B \middle| \xi(t)\right) \tag{1}$$

Приведем еще одно эквивалентное определение марковского процесса (доказательство эквивалентности можно найти, например, в [2]).

Определение 2. *Случайный процесс* $\xi(t), t \geq 0$, *называется марковским, если для любой ограниченной борелевской функции* f *и любых* $t_1 < t_2 < .. < t_n \leq t$

$$\mathrm{M}\left(f\left(\xi(t)\right) \middle| \xi(t_1), \ldots, \xi(t_n)\right) = \mathrm{M}\left(f\left(\xi(t)\right) \middle| \xi(t_n)\right) \tag{2}$$

Пусть по определению

$$P(s,x,t,B) = \mathrm{P}\left(\xi(t) \in B \mid \xi(s) = x\right)$$

и эта функция является абсолютно непрерывной, т. е. допускает представление

$$P(s,x,t,B)=\int_B p(s,x,t,y)\,dy, \forall B\in\mathscr{B}(R).$$

Утверждение 1. *Если $\xi(t), t\ge 0$, – марковский случайный процесс, то*

$$P(s,x,t,B)=\int_R P(s,x,u,dy)\cdot P(u,y,t,B) \tag{3}$$

и

$$p(s,x,t,z)=\int_R p(s,x,u,y)\cdot p(u,y,t,z)\,dy \tag{4}$$

Уравнения (3), (4) носят название *уравнений Колмогорова–Чепмена*, функция $P(s,x,t,B)$ называется *переходной функцией* (или *вероятностью*) *процесса* $\xi(t)$, а функция $p(s,x,t,y)$ – *переходной плотностью*.

Доказательство. В силу марковости процесса $\xi(t)$ имеем для $u>s$

$$P(s,x,t,B)=\mathrm{P}\big(\xi(t)\in B\,\big|\,\xi(s)=x\big)=$$

$$=\int_R \mathrm{P}\Big(\big(\xi(t)\in B\cap\xi(u)=y\big)\Big|\,\xi(s)=x\Big)dy=$$

$$=\int_R \frac{\mathrm{P}\big(\xi(t)\in B\cap\xi(u)=y\cap\xi(s)=x\big)}{\mathrm{P}\big(\xi(s)=x\big)}dy=$$

$$=\int_R \mathrm{P}\big(\xi(t)\in B\,\big|\,\xi(u)=y\cap\xi(s)=x\big)\cdot\frac{\mathrm{P}\big(\xi(u)=y\cap\xi(s)=x\big)}{\mathrm{P}\big(\xi(s)=x\big)}dy=$$

$$=\int_R \mathrm{P}\big(\xi(t)\in B\,\big|\,\xi(u)=y\big)\cdot\mathrm{P}\big(\xi(u)=y\,\big|\,\xi(s)=x\big)dy=$$

$$=\int_R P(s,x,u,dy)\cdot P(u,y,t,B).$$

Равенство (3) доказано. Аналогично доказывается равенство (4).

Выясним теперь, какими свойствами должна обладать функция $P(s,x,t,B)$, чтобы нашелся марковский процесс $\xi(t)$, для которого

$$P(s,x,t,B)=\mathrm{P}\big(\xi(t)\in B\,\big|\,\xi(s)=x\big).$$

Определение 3. *Функция $P(s,t,x,B)$ называется переходной функцией на измеримом пространстве $(X,\mathscr{B}_X)$, если она удовлетворяет следующим условиям:*

1) $P(s,x,t,B)$ *как функция от* B *есть распределение вероятностей при каждых* $s \le t$, $x \in X$;

2) $P(s,x,t,B)$ *измерима по* x *при каждых* $s \le t$ *и* $B \in \mathscr{B}_X$;

3) *при* $0 \le s < u < t$ *и при всех* x, B *из* $\mathscr{B}_X$ *справедливо равенство* (3);

4) $P(s,x,t,B) = I_X(B)$ *при* $s = t$.

В этом определении свойства 1) и 2) обеспечивают функции $P(s,x,t,B)$ возможность быть условным распределением.

Зададим теперь с помощью $P(s,x,t,B)$ конечномерные распределения некоторого процесса $\xi(t)$ с начальным условием $\xi(0) = a$ по формуле

$$\begin{aligned} &\mathrm{P}\left(\xi(t_1) \in dy_1, \ldots, \xi(t_n) \in dy_n\right) = \\ &= P(0,a,t_1,dy_1) \cdot P(t_1,y_1,t_2,dy_2) \ldots P(t_{n-1},y_{n-1},t_n,dy_n). \end{aligned} \tag{5}$$

В силу свойств 3) и 4) эти распределения согласованы и, значит, определяют $\xi(t)$ на $\left(R^T, \mathscr{B}_{R^T}^T\right)$.

По формуле (5) и правилу (2)

$$\mathrm{P}\left(\xi(t_n) \in B_n \middle| \left(\xi(t_1), \ldots, \xi(t_{n-1})\right) = (y_1, \ldots, y_{n-1})\right) = P(t_{n-1}, y_{n-1}, t_n, B_n)$$

Итак, в силу определения 2 мы построили марковский процесс, для которого

$$P(s,x,t,B) = \mathrm{P}\left(\xi(t) \in B \middle| \xi(s) = x\right).$$

О п р е д е л е н и е 4. Марковский процесс $\xi(t)$ называется *однородным*, если $P(s,x,t,B)$ как функция s и t зависит лишь от разности $t - s$:

$$P(s,x,t,B) = P(t-s,x,B).$$

Это есть вероятность за время $t - s$ попасть из x в B.

Нетрудно видеть, что пуассоновский процесс и процесс броуновского движения являются однородными марковскими процессами.

Рассмотрим один важный класс марковских процессов с непрерывными траекториями.

О п р е д е л е н и е 5. *Однородный марковский процесс* $\xi(t)$ *в* $(R, \mathscr{B}_R)$ *с переходной функцией* $P(t,x,B)$ *называется диффузионным, если*

1) $\lim\limits_{\Delta\to 0}\dfrac{1}{\Delta}\int\limits_R y\cdot P(\Delta,x,dy)=a(x)$;

2) $\lim\limits_{\Delta\to 0}\dfrac{1}{\Delta}\int\limits_R y^2\cdot P(\Delta,x,dy)=b^2(x)$;

3) *при каких-нибудь* $\delta>0$, $c<\infty$

$$\int\limits_R y^{2+\delta}P(\Delta,x,dy)<c\cdot\Delta^{1+\frac{\delta}{2}}.$$

Обозначим $\Delta\xi(t)=\xi(t+\Delta)-\xi(t)$. Тогда приведенные условия можно записать в виде

$$\mathrm{M}\Big(\Delta\xi(t)\Big|\xi(t)=x\Big)\sim\Delta\cdot a(x),$$

$$\mathrm{M}\Big(\big(\Delta\xi(t)\big)^2\Big|\xi(t)=x\Big)\sim\Delta\cdot b^2(x),$$

$$\mathrm{M}\Big(\big(\Delta\xi(t)\big)^{2+\delta}\Big|\xi(t)=x\Big)\sim c\cdot\Delta^{1+\frac{\delta}{2}}.$$

Коэффициенты $a(x)$ и $b(x)$ называют соответственно *коэффициентами сноса* и *диффузии*. Условие 3) есть некий аналог условия Ляпунова. Его можно заменить на условие

3*а*) $\mathrm{M}\Big(\big(\Delta\xi(t)\big)^2;\big|\xi(t)\big|>\varepsilon\Big)=o(\Delta)$ *при любом* $\varepsilon>0$.

Из условия 3) и теоремы Колмогорова непосредственно вытекает, что диффузионный процесс $\xi(t)$ можно рассматривать как процесс с непрерывными траекториями.

Пример 1. Стандартный процесс броуновского движения $W(t)$ является диффузионным процессом, так как для него

$$P(t,x,B)=\frac{1}{\sqrt{2\pi t}}\int\limits_B e^{-\frac{(x-y)^2}{2t}}dy,$$

$\mathrm{M}\Delta W(t)=0$, $\mathrm{M}\big|\Delta W(t)\big|^2=\Delta$, $\mathrm{M}\big|\Delta W(t)\big|^4=3\Delta^2$.

Отметим, что распределение диффузионного процесса однозначно определяется коэффициентами сноса и диффузии. Это доказывает следующая теорема.

Теорема 1 (Обратное уравнение Колмогорова). *Если переходные вероятности* $P(t,x,B)$ *диффузионного процесса дважды непрерывно дифференцируемы по* x, *то они дифференцируемы по* t *и удовлетворяют уравнению*

$$\frac{\partial P}{\partial t} = a \cdot \frac{\partial P}{\partial x} + \frac{b^2}{2} \cdot \frac{\partial^2 P}{\partial x^2} \tag{6}$$

и начальному условию $P(0, x, B) = I_B(x)$.

З а м е ч а н и е 1. Условия теоремы на гладкость переходных функций P на самом деле могут быть доказаны в предположении непрерывности a и b и условий $b > 0$, $|a| \le c \cdot (|x| + 1)$, $b^2 \le c \cdot (|x| + 1)$.

Д о к а з а т е л ь с т в о. Обозначим для краткости через P'_t, P'_x, P''_{x^2} соответствующие частные производные функции P и воспользуемся разложением

$$\begin{gathered} P(t, y, B) - P(t, x, B) = (y - x) \cdot P'_x + \frac{(y-x)^2}{2} \cdot P''_{x^2} + \\ + \frac{(y-x)^2}{2} \left[P''_{x^2}(t, y_0, B) - P''_{x^2}(t, x, B) \right], y_0 \in (x, y). \end{gathered} \tag{7}$$

Тогда в силу уравнения Колмогорова–Чепмена

$$\begin{aligned} & P(t + \Delta, x, B) - P(t, x, B) = \\ & = \int_R P(\Delta, x, dy) \cdot \left(P(t, y, B) - P(t, x, B) \right) = \\ & = a(x) \cdot P'_x \cdot \Delta + \frac{b^2(x)}{2} \cdot P''_{x^2} \cdot \Delta + o(\Delta) + R_\Delta, \end{aligned} \tag{8}$$

где

$$\begin{aligned} R_\Delta &= \int_R \frac{(y-x)^2}{2} \left[P''_{x^2}(t, y_0, B) - P''_{x^2}(t, x, B) \right] P(\Delta, x, dy) = \\ &= \int_{|y-x| \le \varepsilon} \frac{(y-x)^2}{2} \left[P''_{x^2}(t, y_0, B) - P''_{x^2}(t, x, B) \right] P(\Delta, x, dy) + \\ &+ \int_{|y-x| > \varepsilon} \frac{(y-x)^2}{2} \left[P''_{x^2}(t, y_0, B) - P''_{x^2}(t, x, B) \right] P(\Delta, x, dy). \end{aligned}$$

Первый интеграл здесь в силу непрерывности функции P''_{x^2} не превосходит

$$\delta(\varepsilon) \cdot \left[\frac{b^2(x)}{2} + o(\Delta) \right],$$

$\delta(\varepsilon) \to 0$ при $\varepsilon \to 0$; второй интеграл по условию 3а) есть $o(\Delta)$ при $\Delta \to 0$. Таким образом, в силу произвольности ε

$$R_\Delta = o(\Delta), \Delta \to 0,$$

и из вышеприведенного следует, что существует

$$\lim_{\Delta \to 0} \frac{P(t+\Delta, x, B) - P(t, x, B)}{\Delta} = a(x) \cdot P'_x + \frac{b^2(x)}{2} \cdot P''_{x^2}.$$

Теорема доказана.

Замечание 2. Из теории дифференциальных уравнений известно, что при широких предположениях о коэффициентах a и b и при $B = (-\infty, z]$, $z \in R$, задача Коши (6) имеет единственное решение P, бесконечное число раз дифференцируемое по t, x и z. Отсюда в частности следует, что $P(t, x, B)$ имеет плотность $p(t, x, z)$, являющуюся фундаментальным решением (6).

Замечание 3. Из теоремы 1 нетрудно получить, что вместе с $P(t, x, B)$ уравнению (6) будет удовлетворять функция

$$u(t, x) = \int_R \varphi(z) \cdot P(t, x, dz)$$

для любой гладкой функции φ с конечным носителем.

Мы рассматривали в доказательстве теоремы 1 приращение времени, предшествующее основному интервалу. В связи с этим уравнения (6) называются *обратными уравнениями Колмогорова*. Аналогичным образом могут быть получены *прямые уравнения Колмогорова*.

Теорема 2. (Прямое уравнение Колмогорова) *Пусть переходная плотность $p(t, x, y)$ такова, что существуют непрерывные производные*

$$\frac{\partial}{\partial y}\left[a(y) \cdot p(t, x, y)\right], \ \frac{\partial^2}{\partial y^2}\left[b^2(y) \cdot p(t, x, y)\right].$$

Тогда $p(t, x, y)$ удовлетворяет уравнению

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial y}\left[a(y) \cdot p(t, x, y)\right] + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[b^2(y) \cdot p(t, x, y)\right]. \tag{9}$$

Доказательство. Пусть $\varphi(y)$ – гладкая функция с ограниченным носителем,

$$u(t, x) = \int_R \varphi(y) \cdot p(t, x, y)\, dy.$$

Тогда

$$u(t+\Delta,x)-u(t,x)=$$
$$=\int_R p(t,x,z)\cdot\left[\int_R p(\Delta,z,y)\cdot\varphi(y)\,dy-\int_R p(\Delta,z,y)\cdot\varphi(z)\,dy\right]dz \qquad (10)$$

Если воспользоваться разложением в ряд $\varphi(y)-\varphi(z)$, то для выражения в квадратных скобках из (10) , аналогично тому, как это было в доказательстве теоремы 1, получим

$$\left[a(z)\cdot\varphi'(z)+\frac{b^2(z)}{2}\cdot\varphi''(z)\right]\cdot\Delta+o(\Delta),\Delta\to 0.$$

Отсюда следует, что существует частная производная функции u по переменной t и

$$\frac{\partial u}{\partial t}=\int_R p(t,x,z)\cdot\left[a(z)\cdot\varphi'(z)+\frac{1}{2}b^2(z)\cdot\varphi''(z)\right]dz.$$

Используя интегрирование по частям, получим

$$\frac{\partial u}{\partial t}=\int_R\left(-\frac{\partial}{\partial z}\left(a(z)\,p(t,x,z)\right)+\frac{1}{2}\frac{\partial^2}{\partial z^2}\left(b^2(z)\,p(t,x,z)\right)\right)\varphi(z)\,dz.$$

Из последнего соотношения в силу произвольности функции φ приходим к доказательству теоремы 2.

П р и м е р 2 . *Процесс Орнштейна–Уленбека.* Пусть

$$\xi(t)=x\,e^{at}+\sigma\,e^{at}\,W\left(\frac{1-e^{-2at}}{2a}\right),\ t\ge 0,$$

где $W(t)$ – стандартный процесс броуновского движения. Этот процесс является однородным диффузионным процессом с переходной плотностью

$$p(t,x,y)=\frac{1}{\sigma(t)\sqrt{2\pi}}\exp\left\{-\frac{\left(y-x\,e^{at}\right)^2}{2\sigma^2\,t}\right\},$$

где

$$\sigma^2(t)=\frac{\sigma^2}{2a}\left(e^{2at}-1\right).$$

Несложно убедиться в том, что переходная плотность удовлетворяет прямому и обратному уравнениям Колмогорова с коэффициентами $a(x)=ax$, $b(x)=\sigma^2=const$.

ДОПОЛНЕНИЯ К ГЛАВЕ 7

§ 1. Обобщенные случайные процессы

Обобщенный случайный процесс. Одной из целей теории случайных процессов является построение различных моделей случайных процессов, исходя из простейших процессов. В некоторых отношениях простейшим процессом является процесс, называемый «белым шумом». К сожалению, даваемые определения этого процесса часто, особенно в технической литературе, бывают неудовлетворительными. Сам термин «белый шум» должен означать, что этот процесс является стационарным, все гармонические составляющие которого входят с одинаковой интенсивностью, т. е. если $F(u)$ – спектральная функция процесса, то должно быть $F(u) = Jdu$. Однако в этом случае формула, дающая спектральное представление корреляционной функции процесса, лишена смысла. Часто определяют «белый шум» как действительный стационарный процесс с некоррелированными значениями, полагая

$$\mathrm{M}\xi(t)\xi(s) = 0 \text{ при } t \neq s$$

$$\mathrm{M}|\xi(t)|^2 = \infty.$$

Это определение также неудовлетворительно. Иногда говорят, что «белый шум» – это производная от винеровского процесса. Но винеровский процесс не дифференцируем. Все же, несмотря на явную логическую противоречивость приведенных определений, они содержат в себе существенную и важную информацию о «белом шуме», и можно дать разумное определение «белого шума», при котором он обладает тремя перечисленными выше качествами. Разумеется, определенный при этом объект не является случайным процессом в обычном смысле этого слова, а обобщенным случайным процессом, обобщающим понятие случайного процесса в том же смысле, в котором понятие обобщенной функции (или распределения по Л. Шварцу) обобщает понятие обычной функции.

В дальнейшем мы ограничимся лишь самыми необходимыми определениями и свойствами обобщенных случайных процессов.

Пусть $\varphi(t)$ – функция, определенная на вещественной прямой R. *Носителем* функции $\varphi(t)$ называют наименьшее замкнутое множество, вне которого $\varphi(t)$ равно нулю. Введем пространство $\mathscr{D}$ – пространство всех комплекснозначных бесконечно дифференцируемых функций на R с ограниченным носителем. В пространстве $\mathscr{D}$ следующим образом введем *топологию*:

последовательность φ_n, $\varphi_n \in \mathscr{D}$, $n = 1, 2, \ldots$, *сходится к* $\varphi \in \mathscr{D}$, *если:*

а) *все носители функций* φ_n *содержатся в некотором ограниченном множестве;*

б) φ_n *вместе со всеми производными* $\varphi_n^{(k)}$ *равномерно сходятся к* φ *и к соответствующим производным* $\varphi^{(k)}$, $k = 1, 2, \ldots$.

Подпространство $\mathscr{D}$, состоящее из действительных функций φ, обозначим $\mathscr{D}_r$.

Определение 1. *Обобщенным случайным процессом* $\Lambda(\varphi)$ *или* $\langle\Lambda,\varphi\rangle$ *называют случайный линейный непрерывный функционал, определенный на* $\mathscr{D}$. *Это означает, что каждому* $\varphi\in\mathscr{D}$ *поставлена в соответствие некоторая случайная величина* $\Lambda(\varphi)$ *(вообще говоря, комплекснозначная), определенная на фиксированном вероятностном пространстве* $(\Omega,\mathscr{A},\mathrm{P})$, *так что:*

а) $\Lambda(c_1\varphi_1+c_2\varphi_2)=c_1\Lambda(\varphi_1)+c_2\Lambda(\varphi_2)$ *для любых* $\varphi_i\in\mathscr{D}$ *и констант* c_i, $i=1,2$.

б) если φ_n *сходится к* φ *в* $\mathscr{D}$, *то* $\Lambda(\varphi_n)$ *сходится по вероятности к* $\Lambda(\varphi)$.

В настоящей главе рассматриваются только обобщенные случайные процессы $\Lambda(\varphi)$, удовлетворяющие условиям:

а) для всех $\varphi\in\mathscr{D}$ $\mathrm{M}\left|\Lambda(\varphi)\right|^2<\infty$,

б) если $\varphi_n\to\varphi$ в $\mathscr{D}$, то $\mathrm{M}\left|\Lambda(\varphi_n)-\Lambda(\varphi)\right|^2\to 0$, т. е. $\Lambda(\varphi)=\mathrm{l.i.m.}\Lambda(\varphi_n)$.

Поэтому, говоря в дальнейшем об обобщенном случайном процессе, мы будем всегда предполагать, даже если это явно не выражено, что условия а), б) выполнены.

Пусть $\xi(t)$ непрерывный в среднеквадратическом процесс. Положим

$$\langle\xi,\varphi\rangle=\int\limits_{-\infty}^{\infty}\xi(t)\,\varphi(t)\,dt. \tag{1}$$

Этот интеграл имеет смысл, причем

$$\mathrm{M}\left|\langle\xi,\varphi\rangle\right|^2=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\varphi(t)\,B(t,s)\overline{\varphi(s)}\,dt\,ds,$$

где $B(t,s)$ – ковариация процесса $\xi(t)$. Если $\varphi_n\to\varphi$ в $\mathscr{D}$, то

$$\mathrm{M}\left|\langle\Lambda_\xi,\varphi-\varphi_n\rangle\right|^2\le\int\limits_{-a}^{a}B(t,t)\,dt\int\limits_{-a}^{a}\left|\varphi-\varphi_n\right|^2dt\to 0,$$

где $(-a,a)$ – интервал, содержащий все носители функций φ_n. Итак, каждый непрерывный в среднеквадратическом случайный процесс порождает обобщенный процесс $\langle\xi,\varphi\rangle$, удовлетворяющий условиям а) и б). Отметим, что обобщенный процесс $\langle\xi,\varphi\rangle$ однозначно определяет процесс $\xi(t)$. Действительно, если существуют два непрерывных в среднеквадратическом процесса $\xi(t)$ и $\eta(t)$, которым соответствует, согласно формуле (1), один и тот же обобщенный процесс $\langle\xi,\varphi\rangle$, то

$$\int\limits_{-\infty}^{\infty}\zeta(t)\,\varphi(t)\,dt=0$$

для каждой $\varphi\in\mathscr{D}$, где $\zeta(t)=\xi(t)-\eta(t)$. Отсюда

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(t)\,B_\zeta(t,s)\overline{\varphi(s)}\,dt\,ds=0 \tag{2}$$

для любого $\varphi\in\mathcal{D}$.

Если при некотором t_0 выполняется неравенство $B_\zeta(t_0,t_0)>0$, то, на основании непрерывности функции $B_\zeta(t,s)$, можно найти такое h, что $\operatorname{Re}B_\zeta(t,s)>0$ при $|t-t_0|<h,\ |s-t_0|<h.$ Существует функция $\varphi_0\in\mathcal{D}$, такая, что $\varphi(t)\equiv 0$ при $|t-t_0|\geq h$ и $\varphi(t)>0$ при $|t-t_0|<h$. Тогда

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(t)\operatorname{Re}B_\zeta(t,s)\varphi(s)\,dt\,ds>0,$$

что противоречит (2). Итак, $B_\zeta(t,t)=0$ для любых $t>0$ и $\zeta(t)=\xi(t)-\eta(t)=0$ с вероятностью 1 при каждом t. Следовательно, обобщенный процесс Λ, если он имеет вид $\langle\xi,\varphi\rangle$, однозначно определяет процесс $\xi(t)$. В дальнейшем мы часто будем отождествлять непрерывный в среднеквадратическом случайный процесс $\xi(t)$ с порождаемым им обобщенным случайным процессом $\xi(\varphi)=\langle\xi,\varphi\rangle$ и говорить в этом случае, что обобщенный процесс $\xi(\varphi)$ является процессом обычного типа.

Для обобщенного случайного процесса $\Lambda(\varphi)$ введем два функционала на $\mathcal{D}$. Первый из них

$$m(\varphi)=\mathrm{M}\Lambda(\varphi)$$

называют *средним значением процесса* Λ. Очевидно, он линеен и непрерывен на $\mathcal{D}$.

Ковариацией обобщенного случайного процесса называют функционал

$$B(\varphi,\psi)=\mathrm{M}\Lambda(\varphi)\overline{\Lambda(\psi)}.$$

Он билинеен

$$B(a_1\varphi_1+a_2\varphi_2,\psi)=a_1B(\varphi_1,\psi)+a_2B(\varphi_2,\psi),$$

$$B(\varphi,a_1\psi_1+a_2\psi_2)=\overline{a_1}B(\varphi,\psi_1)+\overline{a_2}B(\varphi,\psi_2),$$

симметричен

$$B(\varphi,\psi)=\overline{B(\psi,\varphi)}$$

и непрерывен в $\mathcal{D}$: из $\varphi_n\to\varphi$ и $\psi_n\to\psi$ в $\mathcal{D}$ вытекает

$$B(\varphi_n,\psi_n)\to B(\varphi,\psi).$$

Для доказательства непрерывности ковариации замечаем, что при $\varphi_n\to\varphi$, $\psi_n\to\psi$ (в $\mathcal{D}$)

$$\left|B(\varphi_n,\psi_n)-B(\varphi,\psi)\right|=\left|\mathrm{M}\left(\Lambda(\varphi_n)\overline{\Lambda(\psi_n)}-\Lambda(\varphi)\overline{\Lambda(\psi)}\right)\right|=$$

$$=\Big|\mathrm{M}\Big[\left(\Lambda(\varphi)-\Lambda(\varphi_n)\right)\overline{\left(\Lambda(\psi)-\Lambda(\psi_n)\right)}-\Lambda(\varphi)\overline{\left(\Lambda(\psi)-\Lambda(\psi_n)\right)}-$$

$$-\left(\Lambda(\varphi)-\Lambda(\varphi_n)\right)\overline{\Lambda(\psi)}\Big]\Big|\le$$
$$\le\left(\mathrm{M}\left|\Lambda(\varphi)-\Lambda(\varphi_n)\right|^2\mathrm{M}\left|\Lambda(\psi)-\Lambda(\psi_n)\right|^2\right)^{1/2}+$$
$$+\left(\mathrm{M}\left|\Lambda(\varphi)\right|^2\mathrm{M}\left|\Lambda(\psi)-\Lambda(\psi_n)\right|^2\right)^{1/2}+$$
$$+\left(\mathrm{M}\left|\Lambda(\varphi)-\Lambda(\varphi_n)\right|^2\mathrm{M}\left|\Lambda(\psi)\right|^2\right)^{1/2}\to 0$$

на основании свойства непрерывности обобщенного процесса.

Аналогичными свойствами обладает корреляционный функционал $R(\varphi,\psi)$ обобщенного случайного процесса,

$$R(\varphi,\psi)=\mathrm{M}\left(\Lambda(\varphi)-m(\varphi)\right)\left(\Lambda(\psi)-m(\psi)\right)=$$
$$=B(\varphi,\psi)-m(\varphi)\overline{m(\psi)}.$$

Если обобщенный процесс $\Lambda(\varphi)$ является процессом обычного типа, $\Lambda(\varphi)=\langle\xi,\varphi\rangle$, то для ковариационного функционала имеем следующее выражение:

$$B(\varphi,\psi)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(t)B(t,s)\overline{\psi(s)}\,dt\,ds, \tag{3}$$

где $B(t,s)$–ковариация процесса $\xi(t)$. Заметим еще, что билинейный симметрический функционал $B(\varphi,\psi)$ однозначно восстанавливается по квадратическому $B(\varphi,\varphi)$.

Характеристический функционал обобщенного случайного процесса. *Характеристическим функционалом действительного обобщенного случайного процесса* $\Lambda(\varphi)$ называют обобщенную функцию $\Phi(\varphi)$ в $\mathscr{D}$, определяемую равенством

$$\Phi(\varphi)=\mathrm{M}\exp\left(i\Lambda(\varphi)\right),$$

где $\varphi\in\mathscr{D}$ –действительная функция $(\varphi\in\mathscr{D}_r)$.

Очевидно, характеристический функционал однозначно определяет всевозможные совместные распределения величин

$$\Lambda(\varphi_1),\Lambda(\varphi_2),\ldots,\Lambda(\varphi_n). \tag{4}$$

В самом деле, характеристическая функция $\Phi_n(u_1,u_2,\ldots,u_n)$ совместного распределения последовательности случайных величин (4) равна $\Phi(u_1\varphi_1+\ldots+u_n\varphi_n)$, и она однозначно определяет распределение этих величин.

Действительный обобщенный случайный процесс называют *гауссовским*, если

$$\Phi(\varphi)=\exp\left(im(\varphi)-\frac{1}{2}R(\varphi,\varphi)\right),$$

где $m(\varphi)$ – некоторый линейный непрерывный в $\mathcal{D}$ функционал, а $R(\varphi,\psi)$ – действительный билинейный симметричный непрерывный в $\mathcal{D}$ функционал, такой, что $R(\varphi,\varphi)\geq 0$ для всех $\varphi\in\mathcal{D}$. Если $\xi(t)$ – обычный действительный гауссовский непрерывный в среднеквадратическом процесс с

$$\mathrm{M}\xi(t)=m(t),$$

$$\mathrm{M}\big(\xi(t)-m(t)\big)\big(\xi(s)-m(s)\big)=R(t,s),$$

то функции $m(t)$ и $R(t,s)$ непрерывны (и обратно, если $m(t)$ и $R(t,s)$ непрерывны, то процесс $\xi(t)$ непрерывен в среднеквадратическом), и порождаемый им обобщенный процесс $\langle\xi,\varphi\rangle$ является гауссовским.

Дифференцирование обобщенных процессов. Определим понятие производной $\frac{d\Lambda}{dt}$ обобщенного случайного процесса.

О п р е д е л е н и е 2. *Производной* $\frac{d\Lambda}{dt}$ *называют обобщенный случайный процесс*

$$\left\langle\frac{d\Lambda}{dt},\varphi\right\rangle=-\left\langle\Lambda,\frac{d\varphi}{dt}\right\rangle, \tag{5}$$

и вообще k-й производной $\frac{d^k\Lambda}{dt^k}$ *обобщенного процесса* Λ *называют процесс*

$$\left\langle\frac{d^k\Lambda}{dt^k},\varphi\right\rangle=(-1)^k\left\langle\Lambda,\frac{d^k\varphi}{dt^k}\right\rangle. \tag{6}$$

Очевидно, если Λ – обобщенный процесс, то все его производные также являются обобщенными случайными процессами.

Предположим, что обобщенный случайный процесс $\Lambda(\varphi)=\langle\xi,\varphi\rangle$ является процессом обычного типа, причем, процесс $\xi(t)$ непрерывно дифференцируем. Нетрудно убедиться, что для рассматриваемых интегралов остается справедливой формула интегрирования по частям

$$\int_{-\infty}^{+\infty}\xi(t)\frac{d\varphi}{dt}dt=\int_{-\infty}^{+\infty}\xi'(t)\varphi(t)dt.$$

Следовательно, в рассматриваемом случае

$$\left\langle\frac{d\Lambda_\xi}{dt},\varphi\right\rangle=-\int_{-\infty}^{+\infty}\xi'(t)\varphi(t)\,dt,$$

т. е. обобщенный процесс $\frac{d\Lambda_\xi}{dt}$ порождается процессом $\xi'(t)$. Таким образом, определение производной обобщенного случайного процесса согласовано с определением среднеквадратической производной обычного среднеквадратического дифференцируемого случайного процесса. С другой стороны, произвольный обобщенный случайный процесс всегда обладает производными любого сколь угодно высокого порядка, которые также являются обобщенными случайными процессами.

Из определения гауссовского обобщенного процесса непосредственно вытекает, что его производные любого порядка также являются гауссовскими обобщенными процессами.

Приведенное определение обобщенного случайного процесса может быть применено и к тому случаю, когда случайные величины $\Lambda(\varphi)$ вырождаются в константы: $\Lambda(\varphi) = \mathrm{M}\Lambda(\varphi) = L(\varphi)$. Такой вырожденный обобщенный случайный процесс называют *обобщенной функцией*. Итак, обобщенная функция $L = L(\varphi) = \langle L, u \rangle$ – это линейный непрерывный функционал на $\mathcal{D}$. Если

$$L(\varphi) = \int\limits_{-\infty}^{+\infty} l(t)\varphi(t)\,dt,$$

где $l(t)$ – непрерывная (или даже произвольная измеримая функция, интегрируемая на произвольном конечном отрезке), то $L(\varphi)$ называют *обобщенной функцией обычного типа* и отождествляют с $l(t)$.

Производные от обобщенных функций определяются равенствами (5) и (6) точно так же, как и производные обобщенных случайных процессов.

Простейшим примером обобщенной функции, не являющейся функцией обычного типа, является δ - функция Дирака, она определяется равенством

$$\delta(\varphi) = \varphi(0).$$

δ -функцией в точке a называют обобщенную функцию δ_a, для которой

$$\delta_a(\varphi) = \varphi(a).$$

Часто обобщенную функцию $\delta_a(\varphi)$ записывают в виде

$$\delta_a(\varphi) = \int\limits_{-\infty}^{+\infty} \delta(t-a)\varphi(t)\,dt. \qquad (7)$$

При этом интеграл в правой части равенства имеет чисто символическое значение, так как действительной функции, для которой выполняется равенство (7), не существует. В соответствии с общим определением производных производные от δ -функции определяются равенствами

$$\langle \delta', \varphi \rangle = -\varphi'(0), \ \langle \delta^{(n)}, \varphi \rangle = (-1)^n \varphi^{(n)}(0).$$

Обозначим через $Y(t)$ единичный импульс, $Y(t) = 0$ при $t < 0$, $Y(t) = 1$ при $t \geq 0$ и через $Y(\varphi)$ соответствующую обобщенную функцию

$$Y(\varphi)=\int\limits_0^{+\infty}\varphi(t)\,dt.$$

Тогда $Y'(\varphi)=Y(\varphi')=\varphi(0)$. Таким образом, $Y'=\delta$.

Возвратимся к случайным процессам.

«Белый шум». Назовем процесс с некоррелированными приращениями стандартным, если он удовлетворяет условиям

$$\xi(0)=0,\qquad \mathrm{M}\xi(t)=0,\qquad \mathrm{M}|\xi(t)|^2=t,\ \ t\ge 0, \tag{8}$$

При $t<0$ будем считать $\xi(t)=0$. Если $t>s>0$, то $R(t,s)=\mathrm{M}\xi(t)\overline{\xi(s)}=$
$=\mathrm{M}\left(\xi(t)-\xi(s)\right)\overline{\left(\xi(s)-\xi(0)\right)}+\mathrm{M}|\xi(s)|^2=s.$

Итак, при $t,s>0$ $R(t,s)=\min(t,s)$. Выше уже отмечалось, что этот процесс в L_2 не дифференцируем. Найдем корреляционный функционал производной от процесса $\xi(t)$, рассматриваемой как обобщенный процесс. Сначала преобразуем выражение для корреляционного функционала процесса $\xi(t)$. Имеем

$$R(\varphi,\varphi)=\int\limits_0^\infty\int\limits_0^\infty \min(t,s)\varphi(t)\overline{\varphi(s)}dtds=$$

$$=\int\limits_0^\infty \varphi(t)\left(\int\limits_0^t s\overline{\varphi(s)}ds+t\int\limits_t^\infty \overline{\varphi(s)}ds\right)dt=$$

$$=\int\limits_0^\infty t\varphi(t)\int\limits_t^\infty \overline{\varphi(s)}dsdt+\int\limits_0^\infty s\overline{\varphi(s)}\int\limits_s^\infty \varphi(t)\,dtds.$$

Положим

$$\psi(s)=\int\limits_s^\infty \varphi(t)\,dt.$$

Тогда

$$R(\varphi,\varphi)=-\int\limits_0^\infty t\left(\psi'(t)\overline{\psi(t)}+\overline{\psi'(t)}\psi(t)\right)dt=-\int\limits_0^\infty td\left(\psi(t)\overline{\psi(t)}\right),$$

откуда следует, что

$$R(\varphi,\varphi)=\int\limits_0^\infty |\psi(t)|^2\,dt.$$

Обозначим через $R_0(\varphi,\varphi)$ корреляционный функционал обобщенного процесса $\dfrac{d\xi}{dt}$. Тогда

$$R_0(\varphi,\varphi)=\mathrm{M}\left|\frac{d\xi}{dt}(\varphi)\right|^2=\mathrm{M}|\xi(\varphi')|^2=R(\varphi',\varphi')$$

и, следовательно,

$$R_0(\varphi,\varphi)=\int_0^\infty |\varphi(t)|^2\,dt. \tag{9}$$

Формулу (9) можно записать в виде

$$R_0(\varphi,\varphi)=\int_0^\infty\int_0^\infty \delta(t-s)\varphi(t)\overline{\varphi(s)}dtds.$$

Это равенство можно интерпретировать следующим образом: корреляционная функция слабого «белого шума» $R(t,s)=\delta(t-s)$, т. е. слабый белый шум является стационарным процессом с некоррелированными значениями.

Определение 3. *Обобщенный процесс* $\frac{d\xi}{dt}$, *где* ξ – *стандартный процесс с некоррелированными приращениями, называют слабым белым шумом. Если* $\xi(t)$ – *действительный процесс с независимыми приращениями, удовлетворяющий условиям* (8), *то обобщенный процесс* $\frac{d\xi}{dt}$ *называют «белым шумом».*

В частности, гауссовский белый шум – это производная от винеровского процесса. Она является обобщенным гауссовским процессом с характеристическим функционалом

$$\Phi_0(\varphi)=\exp\left\{-\frac{1}{2}\int_0^\infty \varphi^2(t)\,dt\right\},\quad \varphi\in\mathscr{D}. \tag{10}$$

Укажем еще вид характеристического функционала пуассоновского белого шума. По определению, пуассоновский «белый шум» – это производная от центрированного пуассоновского процесса

$$\Pi_0(t)=\Pi(t)-t,\quad t\ge 0,$$

где $\Pi(t)$ – пуассоновский процесс со средним $\mathrm{M}\Pi(t)=t$. Рассмотрим более общий случай процесса

$$\overline{\Pi(t)}=\Pi(t)-at,$$

где $\Pi(t)$ – пуассоновский процесс со средним $\mathrm{M}\Pi(t)=at$. Заметим, что

$$\int_0^\infty \varphi(t)\overline{\Pi(t)}\,dt=-\psi(t)\overline{\Pi(t)}\Big|_0^\infty+\int_0^\infty \psi(t)\,d\overline{\Pi(t)}=\int_0^\infty \psi(t)\,d\overline{\Pi(t)},$$

где $\psi(t)=\int_t^\infty \varphi(s)\,ds.$ Для характеристического функционала процесса $\Pi(t)$ получаем, учитывая, что процесс $\overline{\Pi(t)}$ является процессом с независимыми приращениями, следующие соотношения:

$$\overline{\Phi}(\varphi)=\mathrm{M}\exp\left\{i\int_0^\infty \varphi(t)\overline{\Pi}(t)\,dt\right\}=\mathrm{M}\exp\left\{i\int_0^\infty \psi(t)\,d\overline{\Pi}(t)\right\}=$$

$$=\lim \mathrm{M}\exp\left\{i\sum_{k=1}^n \psi(t_k)\Delta\overline{\Pi}(t_k)\right\}=\lim\prod_{k=1}^n \mathrm{M}\exp\left\{i\psi(t_k)\Delta\overline{\Pi}(t_k)\right\}.$$

При этом надо иметь в виду, что функция $\psi(t)$, начиная с некоторого $t>0$, обращается в 0, $\psi(t)=0$ при $t \geq b$, а знак $\lim$ означает предел по последовательности разбиений отрезка $[0;b]$ точками $0<t_1<t_2<\ldots<t_n=b$, $\Delta\overline{\Pi}(t_k)=\overline{\Pi}(t_k)-\overline{\Pi}(t_{k-1})$. Далее, для центрированного пуассоновского процесса имеем

$$\mathrm{M}\exp\left\{iu\left[\overline{\Pi}(t+\Delta t)-\overline{\Pi}(t)\right]\right\}=e^{a\Delta t\left(e^{iu}-1-iu\right)}.$$

Таким образом,

$$\overline{\Phi}(\varphi)=\lim\prod_{k=1}^{n}\exp\left\{ia\Delta t_k\left(e^{i\psi(t_k)}-1-i\psi(t_k)\right)\right\},$$

или

$$\overline{\Phi}(\varphi)=\exp\left\{ia\int_0^\infty\left[e^{i\psi(t)}-1-i\psi(t)\right]dt\right\}. \qquad (11)$$

Из последней формулы непосредственно получаем выражение для характеристического функционала $\Phi_\Pi(\varphi)$ пуассоновского «белого шума»

$$\Phi_\Pi(\varphi)=\exp\left\{ia\int_0^\infty\left(e^{i\varphi(t)}-1-i\varphi(t)\right)dt\right\}. \qquad (12)$$

Хотя для обобщенного процесса понятие значения процесса в фиксированный момент времени не определено, все же можно дать естественное определение обобщенного процесса с независимыми значениями и показать, что белый шум является таковым.

Определение 4. *Действительный обобщенный процесс* $\Lambda(\varphi)$ *называют процессом с независимыми значениями, если для любых* $\varphi_i \in \mathscr{D}_r$, $i=1,2$, *для которых* $\varphi_1\varphi_2 \equiv 0$, *случайные величины* $\Lambda(\varphi_1)$ *и* $\Lambda(\varphi_2)$ *независимы.*

Теорема 1. *Белый шум является процессом с независимыми значениями.*

Доказательство. Пусть $\varphi_1(t)\varphi_2(t)\equiv 0$, $\varphi_i(t)\in\mathscr{D}_r$. Множество точек, где $\varphi_i(t)>0$, представляет собой сумму не более чем счетного множества интервалов $\Delta_{ik}=(a_{ik},b_{ik})$, $i=1,2$, $k=1,2,\ldots$ попарно без общих точек. При этом ни один из интервалов Δ_{1k} не имеет общих точек ни с одним из Δ_{2k}. Следовательно, $\varphi_i(t)=\sum_k\varphi_{ik}(t)$, $i=1,2$, где $\varphi_{ik}(t)\in\mathscr{D}_r$ и $\varphi_{ik}(t)=0$ при $t\notin\Delta_{ik}$. Обозначим через γ белый шум, $\gamma=\frac{d\xi}{dt}$, где ξ – процесс с независимыми приращениями. Заметим, что

$$\xi(\varphi'_{ik})=\int_{a_{ik}}^{b_{ik}}\varphi'_{ik}(t)\xi(t)\,dt=\varphi_{ik}(t)\xi(t)\Big|_{a_{ik}}^{b_{ik}}-$$

$$-\int_{a_{ik}}^{b_{ik}} \varphi_{ik}(t)\xi(dt) = \int_{a_{ik}}^{b_{ik}} \varphi_{ik}(t)\xi(dt).$$

Поэтому семейство $\{\xi(\varphi'_{ik}),\ i=1,2,\ k=1,2,...\}$ состоит из независимых случайных величин. Следовательно, величины

$$\xi(\varphi'_1) = \sum_k \xi(\varphi'_{1k}) \text{ и } \xi(\varphi'_2) = \sum_k \xi(\varphi'_{2k})$$

независимы. Учитывая это, имеем

$$\mathrm{M}\exp\{i(u_1\gamma(\varphi_1)+u_2\gamma(\varphi_2))\} = \mathrm{M}\exp\{-i(u_1\xi(\varphi'_1)+u_2\xi(\varphi'_2))\} =$$
$$= \mathrm{M}\exp\{-iu_1\xi(\varphi'_1)\}\cdot\exp\{-iu_2\xi(\varphi'_2)\} =$$
$$= \mathrm{M}\exp\{iu_1\gamma(\varphi_1)\}\cdot\mathrm{M}\exp\{iu_2\gamma(\varphi_2)\}.$$

Полученное равенство доказывает независимость случайных величин $\gamma(\varphi_1)$ и $\gamma(\varphi_2)$. Таким образом, γ является процессом с независимыми значениями.

Преобразование Фурье обобщенных процессов. Пусть $\xi(t)$ – непрерывный в среднеквадратическом случайный процесс, для которого $\int_{-\infty}^{\infty}\sqrt{B(t,t)}dt < \infty$, где $B(t,s)$ – ковариация процесса $\xi(t)$. Для такого процесса определено преобразование Фурье

$$\tilde{\xi}(t) = \int_{-\infty}^{\infty} e^{itu}\xi(u)\,du, \tag{13}$$

Действительно,

$$\left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{M}e^{itu}\xi(u)e^{-itv}\overline{\xi(v)}dudv\right| \equiv \left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{it(u-v)}B(u,v)dudv\right| \le$$
$$\le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sqrt{B(u,u)B(v,v)}dudv = \left|\int_{-\infty}^{\infty}\sqrt{B(u,u)}du\right|^2 < \infty,$$

так что условие существования интеграла (13) выполнено.

Процесс $\tilde{\xi}(t)$, в свою очередь, непрерывен в среднеквадратическом. Действительно, рассуждения , аналогичные предыдущим, дают

$$\mathrm{M}\left|\tilde{\xi}(t+h)-\tilde{\xi}(t)\right|^2 \le \left(\int_{-\infty}^{\infty}\sqrt{B(u,u)}\left|e^{ihu}-1\right|du\right)^2,$$

причем,

$$\mathrm{M}\left|\tilde{\xi}(t)\right|^2 \le \left(\int_{-\infty}^{\infty}\sqrt{B(u,u)}du\right)^2.$$

Обобщенный случайный процесс $\tilde{\xi}(\varphi) = \langle\tilde{\xi},\varphi\rangle$, порожденный процессом $\tilde{\xi}(t)$, определяется формулой

$$\tilde{\tilde{\xi}}(\varphi)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{itu}\xi(u)\,du\varphi(t)\,dt=\int_{-\infty}^{\infty}\xi(u)\tilde{\varphi}(u)\,du,$$

или

$$\tilde{\tilde{\xi}}(\varphi)=\xi(\tilde{\varphi}), \qquad (14)$$

где $\tilde{\varphi}$ – преобразование Фурье функции $\varphi(t)\in\mathcal{D}$. Полученная формула наталкивает на мысль определить преобразование Фурье произвольного обобщенного процесса с помощью формулы (14). Однако такое определение лишено смысла, так как функция $\tilde{\varphi}(u)$ не принадлежит $\mathcal{D}$ (если $\tilde{\varphi}(u)\neq 0$). Чтобы преодолеть возникшее затруднение, сузим понятие обобщенного случайного процесса.

О п р е д е л е н и е 5. *Пространство S состоит из тех и только тех комплекснозначных функций* $\varphi(t)$, $t\in(-\infty,\infty)$, *которые:*

а) бесконечно дифференцируемы,

б) на бесконечности убывают вместе со своими производными сколь угодно высокого порядка быстрее, чем $|t|^{-n}$, *где n – любое положительное число.*

Последнее требование означает, что для любых целых положительных m и n функция $t^n\varphi^m(t)$ ограничена и интегрируема в любой степени.

Лемма 1. *Если* $\varphi(t)\in S$, *то и* $\tilde{\varphi}(t)\in S$, *где* $\tilde{\varphi}(t)$ *– преобразование Фурье функции* $\varphi(t)$.

Действительно, в формуле

$$\tilde{\varphi}(t)=\int_{-\infty}^{\infty}e^{itu}\varphi(u)\,du$$

можно дифференцировать под знаком интеграла любое число раз,

$$\tilde{\varphi}^{(m)}(t)=\int_{-\infty}^{\infty}(iu)^m e^{itu}\varphi(u)\,du, \qquad (15)$$

так как полученная под знаком интеграла функция непрерывна и интегрируема. Таким образом, функция $\tilde{\varphi}(t)$ бесконечно дифференцируема. С другой стороны, интегрируя n раз по частям, получим равенство

$$\int_{-\infty}^{\infty}e^{itu}\frac{d^n}{du^n}\varphi(u)\,du=(-it)^n\int_{-\infty}^{\infty}e^{itu}\varphi(u)\,du=(-it)^n\tilde{\varphi}(t), \qquad (16)$$

так как все производные функции $\varphi(u)$ стремятся к 0 при $u\to\pm\infty$. Последняя формула показывает, что $\tilde{\varphi}(t)\to 0$ при $t\to\pm\infty$ быстрее любой степени t. Эти же соображения применимы и к любой производной $\tilde{\varphi}^{(m)}(t)$.

Введем теперь в S топологию.

О п р е д е л е н и е 6. *Последовательность* $\varphi_n(t)$, $\varphi_n\in S$, *сходится к* φ *в S, если для любых целых m и p* $\left(1+|t|^p\right)\left(\varphi^{(m)}(t)-\varphi_n^{(m)}(t)\right)\to 0$ *равномерно по t.*

Нетрудно заметить, что из $\varphi_n \to \varphi$ в S вытекает, что преобразования Фурье $\tilde{\varphi}_n$ сходятся к $\tilde{\varphi}$ в S.

Пусть $(\Omega, \mathscr{A}, \mathrm{P})$ – некоторое вероятностное пространство. Предположим, что каждому $\varphi \in S$ поставлена в соответствие случайная величина $\Lambda(\varphi)$, такая, что:

а) $\Lambda(c_1\varphi_1 + c_2\varphi_2) = c_1\Lambda(\varphi_1) + c_2\Lambda(\varphi_2)$, c_i – константы, $\varphi_i \in S$, $i = 1, 2$,

б) из $\varphi_n \to \varphi$ в S следует $\operatorname{l.i.m.} \Lambda(\varphi_n) = \Lambda(\varphi)$. Это соответствие называют *умеренным обобщенным случайным процессом*. Очевидно, что сужение умеренного обобщенного процесса на $\mathscr{D}$ является обобщенным случайным процессом.

Пусть $\Lambda(\varphi) = \int\limits_{-\infty}^{\infty} \lambda(s)\varphi(s)\,ds$, где $\lambda(s)$ – обычный непрерывный в среднеквадратическом процесс, $\varphi \in S$. Этот интеграл существует, если ковариация $B(t,s)$ процесса $\lambda(t)$ не очень быстро растет при $t, s \to \infty$, например, если для некоторых констант c и m $|B(t,s)| \le c\left(1 + |t|^m\right)$. Действительно, в этом случае

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \varphi(t) B(t,s) \overline{\varphi(s)}\, dt\, ds < \infty,$$

а выполнение этого условия достаточно для существования рассматриваемого интеграла. Нетрудно заметить, что условие б) в определении умеренного обобщенного процесса выполнено. Слово «умеренный» как раз и означает, что момент второго порядка $B(t,t) = \mathrm{M}|\xi(t)|^2$ при $|t| \to \infty$ должен возрастать не очень быстро.

Определение 7. *Преобразованием Фурье* $\tilde{\Lambda}(\varphi)$ *умеренного обобщенного процесса называют обобщенный процесс*

$$\tilde{\Lambda}(\varphi) = \Lambda(\tilde{\varphi}).$$

Из сказанного выше следует, что это определение имеет смысл, и $\tilde{\Lambda}(\varphi)$ также является умеренным процессом.

Аналогично определим преобразование $\overset{\infty}{\Lambda}(\varphi)$, положив

$$\overset{\infty}{\Lambda}(\varphi) = \Lambda\left(\overset{\infty}{\varphi}\right),$$

где

$$\overset{\infty}{\varphi}(t) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{-itu}\varphi(u)\,du.$$

Так как $\overset{\infty}{\tilde{\varphi}} = \overset{\tilde{\infty}}{\varphi} = \varphi$, то

$$\overset{\infty}{\tilde{\Lambda}}(\varphi) = \overset{\tilde{\infty}}{\Lambda}(\varphi) = \Lambda(\varphi),$$

т. е. преобразование $\overset{\infty}{\Lambda}$ является обратным к преобразованию $\tilde{\Lambda}$: если $\Lambda_1 = \tilde{\Lambda}$, то $\Lambda = \overset{\infty}{\Lambda}_1$. Эти формулы являются д в о й с т в е н н ы м и ф о р м у л а м и Ф у р ь е для обобщенных случайных процессов.

Обобщенные стационарные процессы. Пусть $\xi(t)$ – среднеквадратический непрерывный слабо стационарный процесс,

$$\xi(t) = \int_{-\infty}^{\infty} e^{itu}\zeta(u)\,du$$

– его спектральное разложение. Заметим, что порождаемый $\xi(t)$ обобщенный процесс $\xi(\varphi)$ можно представить в виде

$$\xi(\varphi) = \int_{-\infty}^{\infty} \tilde{\varphi}(u)\zeta(du). \tag{17}$$

Здесь $\tilde{\varphi}(u)$ – преобразование Фурье функции $\varphi(t)$; $\zeta(u)$, $u \in (-\infty, \infty)$ – процесс с ортогональными приращениями,

$$\mathrm{M}\left|\zeta(du)\right|^2 = F(du),\ F(-\infty) = 0,\ F(+\infty) < \infty,$$

$F(u)$ – спектральная функция процесса $\xi(t)$.

Введем понятие слабой стационарности для произвольных обобщенных процессов. Пусть θ_h – оператор сдвига аргумента функции

$$\theta_h\varphi(t) = \varphi(t+h).$$

Если $\varphi \in S(\mathscr{D})$, то $\theta_h\varphi \in S(\mathscr{D})$. Определение оператора сдвига аргумента расширим на обобщенные процессы, положив

$$\theta_h\Lambda(\varphi) = \Lambda(\theta_{-h}\varphi).$$

Назовем обобщенный процесс *слабо стационарным*, если:

$$\mathrm{M}\theta_h\xi(\varphi) = m(\varphi),\ \mathrm{M}\left|\theta_h\xi(\varphi)\right|^2 = B(\varphi,\varphi),$$

где $m(\varphi)$ и $B(\varphi,\varphi)$ не зависят от h.

Приведем общий прием построения умеренных обобщенных слабо стационарных процессов.

Пусть $\zeta(u)$, $u \in (-\infty, \infty)$, – процесс с ортогональными приращениями, со средним $\mathrm{M}\zeta(u) = 0$ и структурной функцией $F(u)$:

$$\mathrm{M}\left|\zeta(u_2) - \zeta(u_1)\right|^2 = F(u_2) - F(u_1)\ \ (u_1 < u_2).$$

Мы не предполагаем, что функция $F(u)$ ограничена, но пусть при некотором $p > 0$

$$\int_{-\infty}^{\infty} \frac{F(du)}{1+|u|^p} < \infty. \tag{18}$$

Положим

$$\Lambda(\varphi)=\int_{-\infty}^{\infty}\tilde{\varphi}(u)\zeta(du),\ \varphi\in S, \tag{19}$$

где $\tilde{\varphi}(u)$ – преобразование Фурье функции $\varphi(t)$. Интеграл в правой части равенства существует, причем

$$B(\varphi,\varphi)=\mathrm{M}\left|\Lambda(\varphi)\right|^{2}=\int_{-\infty}^{\infty}\left|\tilde{\varphi}(u)\right|^{2}F(du)<\infty. \tag{20}$$

Если $\varphi_n\to\varphi$ в S, то и $\tilde{\varphi}_n\to\tilde{\varphi}$ в S, и поэтому

$$\mathrm{M}\left|\Lambda(\varphi)-\Lambda(\varphi_n)\right|^{2}=\int_{-\infty}^{\infty}\left|\tilde{\varphi}(u)-\tilde{\varphi}_n(u)\right|^{2}F(du)\to 0$$

при $u\to\infty$. Таким образом, $\Lambda(\varphi)$ является умеренным обобщенным процессом. Он слабо стационарен. Действительно, заметим, что

$$\left(\widetilde{\theta_{-h}\varphi}\right)(u)=\int_{-\infty}^{\infty}e^{itu}\varphi(t-h)\,dt=e^{iuh}\tilde{\varphi}(u).$$

Следовательно,

$$\mathrm{M}\left|\theta_{+h}\Lambda(\varphi)\right|^{2}=\mathrm{M}\left|\Lambda(\theta_{-h}\varphi)\right|^{2}=$$

$$=\int_{-\infty}^{\infty}\left|e^{iuh}\tilde{\varphi}(u)\right|^{2}F(du)=\int_{-\infty}^{\infty}\left|\tilde{\varphi}(u)\right|^{2}F(du)$$

не зависит от *h*. Кроме того, $\mathrm{M}\theta_h\Lambda(\varphi)=0$. Таким образом, для произвольного процесса $\zeta(t)$ с ортогональными приращениями, структурная функция которого удовлетворяет условию (16), $\mathrm{M}\zeta(t)=0$, т. е. обобщенный процесс $\Lambda(\varphi)$, определяемый формулой (19), является слабо стационарным умеренным обобщенным процессом с корреляционным функционалом (20).

Можно показать и обратное, а именно: произвольный слабо стационарный умеренный обобщенный процесс имеет вид (19), где $\zeta(u)$ – процесс с ортогональными приращениями, структурная функция которого удовлетворяет при некотором $p>0$ условию (18). Формулу (19) при этом называют *спектральным разложением процесса*, а (20) – *спектральным представлением корреляционного функционала*, процесс $\zeta(u)$ – *спектральной стохастической мерой*, функция $F(u)$ – *спектральной функцией процесса.*

Если $F(-\infty)>-\infty$, $F(+\infty)<+\infty$, то $\Lambda(\varphi)$ является процессом обычного типа. Действительно, в этом случае

$$\Lambda(\varphi)=\int_{-\infty}^{\infty}\tilde{\varphi}(u)\zeta(du)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{itu}\varphi(t)\,dt\,\zeta(du)=\int_{-\infty}^{\infty}\varphi(t)\xi(t)\,dt,$$

где

$$\xi(t)=\int_{-\infty}^{\infty}e^{itu}\zeta(du).$$

Задачи

1. Пусть A, η и φ – случайные величины, причем A, η – неотрицательны и имеют произвольное совместное распределение, а φ не зависит от них и имеет равномерное распределение на отрезке $[0, 2\pi]$. Рассмотрим случайный процесс

$$\xi(t) = A \cdot \cos(\eta \cdot t + \varphi),\ t \in T = R.$$

Докажите, что его конечномерное распределения не меняются при любом сдвиге времени, т. е. он является стационарным в узком смысле случайным процессом.

2. Пусть все траектории пуассоновского процесса непрерывны справа. Докажите, что тогда почти все траектории – неубывающие целочисленные функции, возрастающие только скачками величины 1.

3. Произведем независимое от пуассоновского процесса $\Pi(t)$ бросание монеты и определим случайный процесс $\eta(t)$ следующим образом: если монета выпадает гербом, положим $\eta(t) = (-1)^{\Pi(t)}$, а если решкой, то $\eta(t) = (-1)^{\Pi(t)+1}$. Найдите конечномерное распределение процесса $\eta(t), t \geq 0$.

4. Пусть $B(t)$, $t \in T = [0; a]$, – процесс броуновского движения, $0 = t_0 < t_1 < ... < t_n = a$ – разбиение отрезка T. Докажите, что для любого положительного числа A

$$P\left(\sum_{i=0}^{n-1} \left|B(t_{i+1}) - B(t_i)\right| > A\right) \to 1$$

при $n \to \infty$ и $\max\limits_{0 \leq i \leq n}(t_{i+1} - t_i) \to 0$.

5. Найдите математические ожидания и корреляционные функции случайных процессов Пуассона и броуновского движения.

6. Найдите корреляционную функцию случайного процесса

$$\xi(t) = \alpha_1 \cdot f_1(t) + ... + \alpha_n \cdot f_n(t), t \in T = R,$$

где $f_1, ..., f_n$ – произвольные числовые функции, а $\alpha_1, ..., \alpha_n$ – независимые случайные величины с дисперсиями $d_1, ..., d_n$.

7. Пусть $\xi(t)$ – стационарный в широком смысле случайный процесс с корреляционной функцией

$$K(t) = \left(1 + |t| + \frac{t^2}{3}\right) \cdot e^{-|t|}.$$

Сколько раз он дифференцируем в среднем квадратическом?

8. Найдите спектральное представление случайного процесса $\xi(t)$ из задачи 1.

9. Пусть $\xi(t)$ – стационарный случайный процесс с математическим ожиданием m и спектральной плотностью $f(\lambda)$; $\eta(t) = \xi(t) \cdot \cos(ct + \varphi)$, где c – константа, а φ – независимая от $\xi(t)$ случайная величена, равномерно распределен-

ная на $[0;2\pi]$. Напишите спектральное представление для $\eta(t)$ и найдите соответствующую спектральную функцию.

10. Является ли процесс броуновского движения марковским? Найдите соответствующую переходную функцию.

11. Найдите переходную функцию для процесса $\xi(t)=B(-t)$, $-\infty<t\le 0$.

12. Докажите, что любой случайный процесс с независимыми приращениями является марковским случайным процессом.

13. $\Pi(t), t\ge 0$, – пуассоновский процесс. Покажите, что в смысле сходимости по вероятности

1) Процесс $\Pi(t)$, $t\ge 0$, непрерывен;

2) Существует производная $\Pi'(t)$ и $\Pi'(t)=0$ п. н.

14. Пусть $B(t)$, $t\ge 0$, – стандартный процесс броуновского движения, g – ограниченная измеримая функция на прямой. Покажите, что

$$\mathrm{M}\int_0^c g(B(t))dt=\int_0^c\left(\frac{1}{\sqrt{2\pi t}}\cdot\int_{-\infty}^{\infty} g(y)\cdot e^{\frac{-y^2}{2t}}dy\right)dt=\int_0^c \mathrm{M}g(B(t))dt.$$

15. Пусть $B(t)$, $t\ge 0$, – стандартный процесс броуновского движения и c – некоторое положительное число. Покажите, что

$$\mathrm{M}\left[\sum_{k=0}^{n-1}\left[B(kh+h)-B(kh)\right]^2-c\right]^2\to 0$$

при $n\to\infty$, где $h=\dfrac{c}{n}$.

16. *Процесс Пуассона*. Пусть некоторая физическая система может находиться в одном и счетного множества состояний E_1, E_2, E_3, ..., причем в любой момент времени t она может сменить свое состояние, перейдя в состояние с номером на единицу большим. Если Δt достаточно мало, то вероятность перехода системы за промежуток времени $(t,\ t+\Delta t)$ из состояния E_n в E_{n+1} равна $\lambda\Delta t+o(\Delta t)$, где λ – положительная постоянная. Составьте систему дифференциальных уравнений описывающих этот процесс, и, решив ее, найдите вероятность $p_n(t)$ $(n=1,2,...)$ того, что система в момент времени t окажется в состоянии E_n, если известно, что в начальный момент $t_0=0$ система находиться в состоянии E_0.

17. *Процесс Юла*. В области G имеются частицы, способные размножаться (путем деления или иначе) и остающиеся в этой области в дальнейшем. За малый промежуток времени $(t,t+\Delta t)$ каждая частица с вероятностью $\lambda\Delta t+o(\Delta t)$ производит новую частицу независимо от остальных частиц.

1) Составьте систему дифференциальных уравнений, определяющих этот процесс.

2) Решите эту систему.

3) Найдите математическое ожидание и дисперсию распределения, определяемого этой системой.

18. *Процесс чистой гибели*. В области G в начальный момент времени $t=0$ находилось k частиц. Независимо друг от друга частицы могут исчезать из этой области, причем каждая за малый промежуток времени Δt может исчезнуть с вероятностью $\lambda\Delta t + o(\Delta t)$. Новые частицы в области появиться не могут.

1) Найдите дифференциальные уравнения, описывающие этот процесс.

2) Найдите вероятность $p_n(t)$ – решения этих уравнений.

3) Найдите математическое ожидание и дисперсию числа частиц в области G к моменту t.

19. *Процесс Пойа*. В области G появляются некоторые частицы и в дальнейшем остаются в этой области. Если к моменту t в области имелось n частиц, то вероятность (условная) увеличения их числа на единицу за малый промежуток времени $(t, t+\Delta t)$ равна

$$\frac{1+an}{1+at}\cdot\Delta t + o(\Delta t),$$

где a – некоторая положительная постоянная. Вероятность увеличения числа частиц за то же время на две или более равна $o(\Delta t)$.

1) Составьте дифференциальные уравнения, определяющие этот процесс (т. е. вероятность $p_n(t)$ того, что к моменту t в области G будет n частиц).

2) Найдите явные выражения $p_n(t)$.

3) Найдите математическое ожидание и дисперсию числа частиц в области.

20. *Процесс размножения и гибели*. В области G имеются однородные частицы (например, бактерии), которые могут порождать такие же новые частицы, например, посредством деления на две, а также могут погибать (исчезать). Если Δt мало, то вероятность для каждой частицы (независимо от наличия и поведения остальных частиц) породить одну новую равна $\lambda\Delta t + o(\Delta t)$, а вероятность погибнуть равна $\mu\Delta t + o(\Delta t)$, где λ и μ – некоторые положительные постоянные.

1) Составьте систему обыкновенных дифференциальных уравнений, определяющих этот процесс.

2) При условии, что $p_n(t)$ $(n = 0,1,2,...)$ – решение этой системы, т. е. вероятность того, что к моменту t в области G окажется n частиц, составьте дифференциальное уравнение в частных производных, которому удовлетворяет производящая функция этих вероятностей, т. е. функция

$$F(x,t) = \sum_{n=0}^{\infty} p_n(t)\cdot x^n.$$

3) Считая, что в начальный момент $t=0$ в области G находилась одна частица, решить это дифференциальное уравнение и найти явное выражение производящей функции $F(x,t)$.

Глава 8. ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Математическая статистика – это прикладная математическая дисциплина, родственная теории вероятностей. Она базируется на понятиях и методах последней, но решает свои специфические задачи собственными методами.

Математические модели случайных явлений, изучаемых в теории вероятностей, основываются на понятии вероятностного пространства $(\Omega, \mathscr{A}, \mathrm{P})$. Однако на практике при изучении конкретного эксперимента вероятность P редко бывает известна полностью. Часто можно априори утверждать лишь, что P является элементом некоторого заданного класса (семейства) вероятностей $\mathscr{P}$. Если задан класс $\mathscr{P}$, то говорят, что имеется *статистическая модель*, понимая под этим набор $(\Omega, \mathscr{A}, \mathscr{P})$. Итак, статистическая модель описывает такие ситуации, когда в вероятностной модели изучаемого эксперимента имеется та или иная неопределенность в задании вероятности Р, и задача математической статистики состоит в том, чтобы уменьшить эту неопределенность, используя информацию, содержащуюся в статистических данных. В большинстве случаев исходные статистические данные – результат наблюдения некоторой конечной совокупности случайных величин $X = (X_1, X_2, \ldots, X_n)$, характеризующей исход изучаемого эксперимента. Обычно в этих случаях говорят, что эксперимент состоит в проведении n испытаний, в которых результат i-го испытания описывается случайной величиной X_i, $i = \overline{1,n}$. Совокупность наблюдаемых случайных величин $X = (X_1, X_2, \ldots, X_n)$ называется *выборкой*, сами величины X_i, $i = \overline{1,n}$, – *элементами выборки*, а их число n – *объемом выборки*. *Реализации выборки* X будем обозначать соответствующими строчными буквами $\vec{x} = (x_1, x_2, \ldots, x_n)$. Множество всех возможных значений $\mathscr{X} = \{X\}$ выборки X называется *выборочным пространством*. Далее предполагается, что статистическая модель определяется выборочным пространством $\mathscr{X}$ и семейством функций распределения $\mathscr{F}$, которому принадлежит неизвестная функция распределения $F_X(x_1, x_2, \ldots, x_n)$ выборки X.

В §1 – 4 гл. 8 рассматривается ситуация, когда компоненты выборки X независимы и все распределены так же, как и некоторая случайная величина ξ. Этот случай соответствует эксперименту, в котором проводятся повторные независимые наблюдения над случайной величиной ξ (распределение ξ обозначается символом $\mathscr{L}(\xi)$). Здесь

$$F_{X_i}(x_i) = F_\xi(x_i), \quad i = \overline{1,n}, \quad F_X(x_1, x_2, \ldots, x_n) = F_\xi(x_1) \cdots F_\xi(x_n).$$

В приложениях математической статистики предположения о независимости и одинаковой распределенности компонент X не всегда выполняются. В § 5 этой главы рассмотрен важный случай таких ситуаций, которые описываются линейной регрессионной моделью.

Иногда множество возможных значений ξ с функцией распределения F_ξ называют *генеральной совокупностью*. Статистическая модель для повторных независимых наблюдений обозначается кратко в виде $\mathscr{F} = \{F_\xi\}$, т. е. указывается лишь класс допустимых функций распределения исходной случайной величины ξ. Если функции распределения из класса $\mathscr{F}$ заданы с точностью до значений некоторого параметра θ (скалярного или векторного) с множеством возможных значений Θ, то такая модель обозначается $\mathscr{F} = \{F(x, \theta), \theta \in \Theta\}$ и называется *параметрической*. Модель $\mathscr{F} = \{F_\xi\}$ называется *абсолютно непрерывной* или *дискретной*, если таковыми соответственно являются все составляющие класс $\mathscr{F}$ функции распределения. Мы будем рассматривать только модели этих двух типов, которые наиболее распространены в приложениях.

Материал этой главы условно можно разбить на два направления – теорию оценивания параметров и теорию проверки гипотез. Первое направление представлено в § 2 и 5, второе – в § 3 и 4.

§ 1. Основные понятия и элементы выборочной теории

1. Вариационный ряд, эмпирическая функция распределения, гистограмма, полигон частот

Пусть $X = (X_1, X_2, \ldots, X_n)$ – выборка объема n из распределения $\mathscr{L}(\xi)$ и $\vec{x} = (x_1, x_2, \ldots, x_n)$ – наблюдавшееся значение X. Каждой реализации $\vec{x}$ выборки X можно поставить в соответствие упорядоченную последовательность

$$x_{(1)} \le x_{(2)} \le \ \ldots \ \le x_{(n)}, \qquad (1)$$

где $x_{(1)} = \min(x_1, \ldots, x_n)$, $x_{(2)}$ – второе по величине значение среди $x_1, \ldots, x_n$ и т. д., $x_{(n)} = \max(x_1, \ldots, x_n)$.

Обозначим через $X_{(k)}$ случайную величину, которая для каждой реализации $\vec{x}$ выборки X принимает значение $x_{(k)}$, $k=\overline{1,n}$. Так по выборке X определяют новую последовательность случайных величин $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$, называемых *порядковыми статистиками* выборки; при этом $X_{(k)}$ – k-я порядковая статистика, а $X_{(1)}$ и $X_{(n)}$ – *экстремальные значения выборки*.

Последовательность случайных величин

$$X_{(1)} \le X_{(2)} \le \ \ldots \ \le X_{(n)}$$

называют *вариационным рядом* выборки.

Определим для каждого действительного x случайную величину $\mu_n(x)=\left|\{j : X_j \le x\}\right|$, где через $|A|$ обозначено число элементов конечного множества A. Функция

$$F_n(x)=\frac{\mu_n(x)}{n}$$

называется *эмпирической функцией распределения* (соответствующей выборке X). Функцию распределения $F(x)$ наблюдаемой случайной величины ξ в этом случае называют иногда *теоретической функцией распределения.*

По своему определению эмпирическая функция распределения – случайная функция: для каждого $x \in R$ значение $F_n(x)$ – случайная величина, реализацией которой являются числа

$$0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{k}{n}, \ldots, \frac{n}{n}=1, \tag{2}$$

а так как $\mathscr{L}(\mu_n(x))=B_i(n,p)$, где $B_i(n,p)$ – распределение биномиальной случайной величины и $p=\mathrm{P}(\xi \le x)=F(x)$, то

$$\mathrm{P}\left(F_n(x)=\frac{k}{n}\right)=\mathrm{P}(\mu_n(x)=k)=C_n^k F^k(x)(1-F(x))^{n-k}. \tag{3}$$

Таким образом, для каждой реализации $\vec{x}$ выборки X функция $F_n(x)$ однозначно определена и обладает всеми свойствами функции распределения: изменяется от 0 до 1, не убывает, непрерывна справа. При этом она кусочно-постоянна и возрастает только в точках последовательности $X_{(1)} \le X_{(2)} \le \ \ldots \ \le X_{(n)}$.

Важнейшее свойство эмпирической функции распределения состоит в том, что при увеличении числа испытаний над случайной величиной ξ происходит сближение этой функции с теоретической функцией распределения.

Теорема 1. *Пусть* $F_n(x)$ *– эмпирическая функция распределения, построенная по выборке* $X = (X_1, X_2, \ldots, X_n)$ *из распределения* $\mathscr{L}(\xi)$, *и* $F(x)$ *соответствующая теоретическая функция распределения. Тогда для любого* $x \in R$ *и любого* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}\left(\left|F_n(x) - F(x)\right| \le \varepsilon\right) = 1. \qquad (4)$$

Доказательство. Из равенства (3) следует, что $F_n(x)$– относительная частота события $\{\xi \le x\}$ («успеха») в n испытаниях схемы Бернулли с вероятностью «успеха» $F(x)$. В соответствии с законом больших чисел (теоремой Бернулли)

$$F_n(x) \xrightarrow[n\to\infty]{\mathrm{P}} F(x),$$

т. е. имеет место равенство (4).

Справедлив следующий более сильный результат, полученный В. И. Гливенко (1933 г.).

Теорема 2 [14]. *В условиях теоремы* 1

$$\mathrm{P}\left(\lim_{n\to\infty} \sup_{-\infty<x<\infty} \left|F_n(x) - F(x)\right| = 0\right) = 1. \qquad (5)$$

Соотношение (5) означает, что отклонение

$$\mathscr{D}_n = \mathscr{D}_n(X) = \sup_{-\infty<x<\infty} \left|F_n(x) - F(x)\right|$$

эмпирической функции распределения от теоретической на всей прямой R с вероятностью единица будет сколь угодно мало при достаточном объеме выборки.

Из изложенного следует, что эмпирическая функция распределения позволяет делать выводы о распределении наблюдаемой случайной величины ξ, когда оно неизвестно.

Существуют и другие способы наглядного представления статистических данных. Одним из них является построение гистограммы. В этом случае область значений наблюдаемой случайной величины ξ разбивают на равные интервалы и для заданной реализации $\vec{x} = (x_1, \ldots, x_n)$ выборки X из $\mathscr{L}(\xi)$ подсчитывают число координат x_i, попавших в соответствующие интервалы, и на каждом интервале, как на основании, строят прямоугольник с высотой ν/nh, где h – длина интервала, ν – число выборочных точек в данном интервале. Получаемую при этом фигуру называют *гистограммой*. Таким образом, площадь каждого прямоугольника равна ν/n, т. е. относительной частоте попадания выборочных значений в соответствующий интервал, поэтому по теореме Бернулли она будет сходиться по вероятности при $n \to \infty$ к вероятности попадания значения случайной величины ξ в соответствующий интервал. Если длина интервалов h достаточно мала, а плотность $f(x)$ случайной величины ξ непрерывна, то последняя вероятность приблизительно равна $f(z)h$, где z – середина соответствующего интервала. Таким образом, при большом значении объема выборки n и достаточно малом h верхнюю границу гистограммы можно рассматривать как статистический аналог плотности распределения наблюдаемой случайной величины (применяется только для непрерывных случайных величин!). Этот способ представления имеет очевидные недостатки: неопределенность в способе построения интервалов, потеря информации при группировке данных, поэтому гистограмму рекомендуется применять только на предварительном этапе анализа статистических данных.

В методе гистограмм неизвестная плотность распределения приближается кусочно-постоянным графиком. Если плотность достаточно гладкая, то значительно лучше ее можно приблизить кусочно-линейными графиками. Отсюда вытекает более точная методика, которая основана на построении полигона частот. *Полигон частот* – это ломаная, которую строят так: если построена гистограмма, то ординаты, соответствующие средним точкам интервалов, последовательно соединяют отрезками.

Иллюстрация в Mathematica

2. Выборочные характеристики

Пусть $X = (X_1, X_2, \ldots, X_n)$ – выборка из распределения $\mathscr{L}(\xi)$, а $F(x)$ и $F_n(x)$ – соответственно теоретическая и эмпирическая функции распределения. Точно так же, как функции $F(x)$ ставят в соответствие $F_n(x)$, любой теоретической характеристике

$$g = \int g(x)dF(x)$$

можно поставить в соответствие ее статистический аналог $G_n = G_n(X)$, вычисляемый по формуле

$$G_n = \int_R g(x)\,dF_n(x) = \frac{1}{n}\sum_{i=1}^{n} g(X_i).$$

Если $g(x) = x^k$, то G_n – *выборочный момент k-го порядка*, который будем обозначать A_{nk},

$$A_{nk} = A_{nk}(X) = \frac{1}{n}\sum_{i=1}^{n} X_i^k. \tag{6}$$

При $k = 1$ величину A_{nk} называют *выборочным средним* и обозначают символом $\overline{X}_n$ (или $\overline{X}$):

$$\overline{X} = \overline{X}_n = A_{n1} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Значения случайной величины A_{nk} и $\overline{X}$ для заданной реализации $\vec{x}$ выборки X будем обозначать соответствующими строчными буквами a_{nk} и $\overline{x}$. Аналогично, *выборочным центральным моментом k-го порядка* называют случайную величину

$$M_{nk} = M_{nk}(X) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^k.$$

При $k = 2$ величину M_{nk} называют *выборочной дисперсией* и обозначают символом S_n^2 (или S^2):

$$S^2 = M_{n2} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

Значения случайных величин M_{nk} и S^2 для заданной реализации $\vec{x}$ выборки X будем обозначать соответствующими строчными буквами m_{nk} и s^2.

Аналогично вводят *выборочные абсолютные моменты, выборочные семиинварианты* и т. д.

При исследовании свойств графика плотности распределения непрерывной случайной величины ξ часто рассматривают такие характеристики, как *коэффициенты асимметрии* γ_1 и *эксцесса* γ_2, определяемые по формулам:

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}}, \qquad \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3, \tag{7}$$

где $\mu_k = \mathrm{M}\left(\xi - \mathrm{M}\xi\right)^k$.

Если график плотности распределения симметричен, то $\gamma_1 = 0$ и по значению γ_1 судят о степени отклонения от симметрии. Аналогично, для нормального распределения $\gamma_2 = 0,$ поэтому о кривых плотности распределения с $\gamma_2 = 0$ говорят, что они имеют нормальный эксцесс.

Если задана выборка $X = \left(X_1, X_2, \ldots, X_n\right)$ из абсолютно непрерывного распределения $\mathscr{L}(\xi),$ то *выборочные коэффициенты асимметрии* Γ_{n1} и *эксцесса* Γ_{n2} определяются по формулам

$$\Gamma_{n1} = \frac{M_{n3}}{S^3}, \qquad \Gamma_{n2} = \frac{M_{n4}}{S^4} - 3.$$

Из следующих очевидных соотношений

$$\mathrm{M}\, A_{nk} = \frac{1}{n}\sum_{i=1}^{n} \mathrm{M}\, X_i^k = \mathrm{M}\,\xi^k = \alpha_k,$$

$$\mathrm{D}\, A_{nk} = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{D}\, X_i^k = \frac{1}{n}\mathrm{D}\,\xi^k = \frac{1}{n}\left(\mathrm{M}\,\xi^{2k} - \left(\mathrm{M}\,\xi^k\right)^2\right) \tag{8}$$

на основании неравенства Чебышева следует, что

$$A_{nk} \xrightarrow[n\to\infty]{\mathrm{P}} \alpha_k.$$

Аналогичные утверждения справедливы и для любых выборочных характеристик, которые имеют вид непрерывных функций от конечного числа величин A_{nk}.

Будем говорить, что случайная величина η_n *асимптотически нормальна с параметрами* μ_n *и* σ_n^2 (или *асимптотически нормальна* $N\left(\mu_n, \sigma_n^2\right)$), если

$$\mathscr{L}\left(\frac{\eta_n - \mu_n}{\sigma_n}\right) \xrightarrow[n\to\infty]{} N(0,1).$$

Теорема 3. *Выборочный момент* A_{nk} *асимптотически нормален*

$$N\left(\alpha_k, \frac{\alpha_{2k} - \alpha_k^2}{n}\right).$$

Д о к а з а т е л ь с т в о. Так как согласно (8)

$$\mathrm{M}\, X_i^k = \alpha_k, \qquad \mathrm{D}\, X_i^k = \alpha_{2k} - \alpha_k^2,$$

то по центральной предельной теореме

$$\mathscr{L}(\eta_n) \xrightarrow[n\to\infty]{} N(0, 1),$$

где

$$\eta_n = \frac{1}{\sqrt{n\left(\alpha_{2k} - \alpha_k^2\right)}} \left(\sum_{i=1}^{n} X_i^k - n\alpha_k \right).$$

Следовательно, случайная величина A_{nk} асимптотически нормальна с параметрами

$$\alpha_k \quad \text{и} \quad \frac{\alpha_{2k} - \alpha_k^2}{n}.$$

Теорема доказана.

Данная теорема позволяет оценивать для больших выборок вероятности заданных отклонений значений выборочных моментов от теоретических, так как при любом фиксированном $t > 0$ и $n \to \infty$

$$\mathrm{P}\left(\sqrt{\frac{n}{\alpha_{2k} - \alpha_k^2}} \left| A_{nk} - \alpha_k \right| \le t \right) \to \frac{1}{\sqrt{2\pi}} \int_{-t}^{t} e^{-\frac{x^2}{2}} dx.$$

Для центральных выборочных моментов также можно доказать утверждение об их асимптотической нормальности. Так, например, для выборочной дисперсии S_n^2 при $n \to \infty$ имеет место соотношение

$$\mathscr{L}\left(S_n^2\right) - N\left(\mu_2, \ \frac{\mu_4 - \mu_2^2}{n} \right) \to 0,$$

где $\mu_k = \mathrm{M}\left(\xi - \mathrm{M}\xi\right)^k$ (см., напр., [14]).

Для $p \in (0, 1)$ *p-квантилью* ζ_p распределения $\mathscr{L}(\xi)$ называется корень уравнения

$$F\left(\zeta_p\right) = p. \tag{9}$$

Если функция $F(x)$ строго монотонна, то это уравнение имеет единственный корень; в противном случае при некоторых p уравнению (9) удовлетворяют многие значения ζ (наполняющие целый интервал), тогда в качестве ζ_p берут минимальное из значений ζ, удовлетворяющих (9).

Выборочной p-квантилью $Z_{n,p}$ будем называть порядковую статистику $X_{([np]+1)}$, где $[a]$ – целая часть числа a:

$$Z_{n,p} = \begin{cases} X_{([np]+1)} & \text{при } np \text{ дробном}, \\ X_{np} & \text{при } np \text{ целом}. \end{cases}$$

Ясно, что $Z_{n,p}$ – это элемент выборки, левее которого находится доля

$$\frac{[np]}{n} \le p$$

наблюдений, при этом $Z_{n,p}$ – порядковая статистика с максимальным номером, обладающая этим свойством. Следовательно, $Z_{n,p}$ можно рассматривать как статистический аналог ζ_p.

Величина $\zeta_{\frac{1}{2}}$ называется *медианой* распределения $\mathscr{L}(\xi)$, а $Z_{n,\frac{1}{2}}$ – *выборочной медианой*.

Теорема 4 [14]. *Пусть функция распределения* $F(x), x \in R$, *непрерывна и уравнение* $F(x) = p,\ p \in (0, 1)$, *имеет единственное решение* $x = \zeta_p$. *Тогда для почти всех* $\omega \in \Omega$

$$Z_{n,p}(\omega) \xrightarrow[n\to\infty]{} \zeta_p.$$

Если, кроме того, функция F *дифференцируема в точке* ζ_p *и* $F'(\zeta_p) > 0$, *то случайная величина* $Z_{n,p}(\omega) - \zeta_p$ *асимптотически нормальна с параметрами*

$$\left(0,\ \frac{p(1-p)}{n\left[F'(\zeta_p)\right]}\right).$$

§ 2. Оценивание неизвестных параметров

1. Статистические оценки и требования к ним

Пусть $X = (X_1,\ \ldots,\ X_n)$ – выборка из распределения $\mathscr{L}(\xi) \in \mathscr{F} = \{F(x,\theta),\ \theta \in \Theta\}$. В общем виде задача формулируется так: используя статистическую информацию, содержащуюся в выборке X, сделать статистические выводы об истинном значении θ^0 неизвестного параметра θ, т. е. оценить точку θ^0 заданного параметрического пространства Θ.

Статистикой называется всякая случайная величина, являющаяся функцией от выборки X, т. е. статистика – это борелевская функция от выборки X .

При *точечном оценивании* ищут статистику $T(X)$, значение которой при заданной реализации $\vec{x} = (x_1,\ \ldots,\ x_n)$ выборки X принимают за приближенное значение параметра θ^0. В этом случае говорят, что *статистика* $T(X)$ *оценивает* θ или, что *статистика* $T(X)$ *есть оценка* θ.

Ясно, что для оценивания θ можно использовать различные оценки и, чтобы выбрать лучшую из них, надо иметь критерий сравнения качества оценок.

Любая оценка $T = T(X)$ является случайной величиной, поэтому общим требованием к построению оценок является требование концентрации (в том или ином смысле) распределения T около истинного значения оцениваемого параметра. Предположим, что θ – скалярный параметр. Статистика $T(X)$ называется *несмещенной оценкой* для параметра θ, если для любого $\theta \in \Theta$

$$\mathrm{M}_\theta T(X) = \theta. \tag{1}$$

Величину

$$\mathrm{M}_\theta \left(T(X) - \theta\right)^2 = \mathrm{D}_\theta T(X) + b^2(\theta), \tag{2}$$

где $b(\theta) = \mathrm{M}_\theta\left(T(X) - \theta\right)$, называют *средним квадратом ошибки* или *среднеквадратической ошибкой* оценки T, а $b(\theta)$ – *смещением* оценки T. Если для любого $\theta \in \Theta$ $b(\theta) \to 0$ при $n \to \infty$, то статистика T называется *асимптотически несмещенной оценкой* для параметра θ.

Аналогично, статистика $T = T(X)$ является *несмещенной оценкой* для $\tau(\theta)$ (τ – некоторая функция), если для любого $\theta \in \Theta$

$$\mathrm{M}_\theta T(x) = \tau(\theta). \tag{3}$$

Требование несмещенности интуитивно привлекательно, однако в некоторых случаях оно может оказаться слишком «жестким» и привести к нежелательным результатам: либо несмещенные оценки вообще не существуют, либо они существуют, но практически бесполезны. Таким образом, требование несмещенности нельзя рассматривать как универсальное, тем не менее во многих встречающихся на практике случаях оно уместно и обоснованно и далее в основном рассматриваются именно несмещенные оценки.

Пусть требуется оценить заданную параметрическую функцию $\tau = \tau(\theta)$ в модели $\mathscr{F} = \{F(x, \theta),\ \theta \in \Theta\}$ по статистической информации, доставляемой соответствующей выборкой $X = (X_1, X_2, \ldots, X_n)$. Обозначим класс всех несмещенных оценок в данной задаче через $\mathscr{T}_\tau$.

Дополнительно предположим, что дисперсии всех оценок из класса $\mathscr{T}_\tau$ конечны, т. е. для любых $T \in \mathscr{T}_\tau$ и $\theta \in \Theta$

$$\mathrm{D}_\theta T = \mathrm{M}_\theta \left(T - \tau(\theta)\right)^2 < \infty.$$

Оценка $\tau^* \in \mathscr{T}_t$, для которой при любых $\theta \in \Theta$ выполняется условие

$$\mathrm{D}_\theta \tau^* = \inf_{T \in \mathscr{T}_\tau} \mathrm{D}_\theta T, \tag{4}$$

называется *оптимальной*.

Пусть $f(x, \theta)$ – плотность распределения наблюдаемой случайной величины ξ (или вероятность в дискретном случае).

Функция

$$L(\vec{x};\theta) = f(x_1, \theta) \cdots f(x_n, \theta) \tag{5}$$

является, очевидно, плотностью распределения случайного вектора X (или вероятностью того, что $X = \vec{x}$ в дискретном случае), здесь $\vec{x} = (x_1, \ldots, x_n)$ – реализация X, т. е. она задает вероятность получения при извлечении выборки объема n именно наблюдений $x_1, \ldots, x_n$ (или величину, пропорциональную вероятности получения выборочных значений в непосредственной близости от точки $\vec{x}$ в непрерывном случае). Поэтому, чем больше значение $L(\vec{x};\theta)$, тем правдоподобнее (или более вероятна) система наблюдений $\vec{x} = (x_1, \ldots, x_n)$ при заданном значении параметра θ. Отсюда и название функции $L(\vec{x};\theta)$, рассматриваемой при фиксированном $\vec{x}$ как функция параметра $\theta \in \Theta$, – *функция правдоподобия*.

Предположим, что $L(\vec{x};\theta) > 0$ для любых $\vec{x} \in \mathscr{X}$ и $\theta \in \Theta$ и дифференцируема по θ. Случайная величина

$$U(X;\theta) = \frac{\partial \ln\left(L(X;\theta)\right)}{\partial\theta} = \sum_{i=1}^{n} \frac{\partial \ln\left(f(X_i, \theta)\right)}{\partial\theta}, \quad \theta \in \Theta, \tag{6}$$

называется *вкладом* (или *функцией вклада*) выборки X. В дальнейшем предполагаем, что для любых $\theta \in \Theta$

$$\mathrm{M}_\theta U(X;\theta) = 0, \quad 0 < \mathrm{M}_\theta U^2(X;\theta) < \infty.$$

Функция

$$i_n(\theta) = \mathrm{D}_\theta U(X;\theta) = \mathrm{M}_\theta U^2(X;\theta) \tag{7}$$

называется *функцией информации Фишера* или просто *информацией Фишера* о параметре θ, содержащейся в выборке X. Величину

$$i(\theta) = i_1(\theta) = \mathrm{M}_\theta \left(\frac{\partial \ln\left(f\left(X_1; \theta \right) \right)}{\partial \theta} \right)^2 \tag{8}$$

называют также *количеством (фишеровской) информации, содержащейся в одном наблюдении*. Из (7) и (8) непосредственно следует, что $i_n(\theta) = ni(\theta)$.

Теорема 1 (**Неравенство Рао – Крамера**). *Пусть функция* $\tau(\theta)$ *дифференцируема, тогда для любой оценки* $T = T(X) \in \mathscr{T}_\tau$ *справедливо неравенство*

$$\mathrm{D}_\theta T \geq \frac{\left[\tau'(\theta) \right]^2}{ni(\theta)}, \quad \theta \in \Theta. \tag{9}$$

Равенство здесь имеет место тогда и только тогда, когда T *– линейная функция вклада выборки, т. е.*

$$T(X) - \tau(\theta) = a(\theta)\, U(X; \theta), \tag{10}$$

где $a(\theta)$ *– некоторая функция от* $\theta \in \Theta$.

Доказательство. По условию для любых $\theta \in \Theta$

$$\mathrm{M}_\theta T(X) = \int\limits_{R^n} T(\vec{x})\, L(\vec{x}; \theta)\, d\vec{x} = \tau(\theta). \tag{11}$$

Так как

$$\int\limits_{R^n} L(\vec{x}; \theta)\, d\vec{x} = 1$$

для любых $\theta \in \Theta$, то

$$0 = \int\limits_{R^n} \frac{\partial L(\vec{x}; \theta)}{\partial \theta} d\vec{x} = \int\limits_{R^n} \frac{\partial \ln\left(L(\vec{x}; \theta) \right)}{\partial \theta} L(\vec{x}; \theta)\, d\vec{x} = \mathrm{M}_\theta U(X; \theta), \ \theta \in \Theta. \tag{12}$$

Дифференцируя равенство (11) по θ и учитывая (12), получаем

$$\tau'(\theta) = \int\limits_{R^n} T(\vec{x}) \frac{\partial \ln\left(L(\vec{x}; \theta) \right)}{\partial \theta} L(\vec{x}; \theta)\, d\vec{x} = \mathrm{M}_\theta \left(T(X) U(X; \theta) \right) =$$

$$= \mathrm{cov}_\theta \left(T(X),\ U(X; \theta) \right). \tag{13}$$

Используя неравенство Коши – Буняковского и определение (7) из (13), получаем (9), причем неравенство обращается в равенство тогда и только тогда, когда $T(X)$ и $U(X; \theta)$ (при каждом θ) линейно связаны.

Таким образом, теорема доказана.

Если существует оценка $T^* \in \mathscr{T}_t$, для которой нижняя граница Рао – Крамера достигается, то ее называют *эффективной*. Очевидно, что эффективная оценка является оптимальной.

Представление (10) является *критерием эффективности оценки*.

Замечание 1. Несложно видеть, что представление (10) эквивалентно следующему равенству

$$L(x;\theta) = \exp\left\{T(\vec{x})\psi_1(\theta) + \psi_2(\theta) + \psi_3(\vec{x})\right\}, \quad \vec{x} \in \mathscr{X}, \quad \theta \in \Theta.$$

Для этого достаточно равенство (10) проинтегрировать по θ.

Модель $\mathscr{F} = \left\{F(x, \theta), \theta \in \Theta\right\}$ – *экспоненциальная*, если соответствующая $f(x, \theta)$ имеет вид

$$f(x, \theta) = \exp\left\{A(\theta)\, B(x) + C(\theta) + D(x)\right\}.$$

Данный класс параметрических моделей важен тем, что эффективные оценки существуют для некоторых функций $\tau(\theta)$ только в случае таких моделей (см. замеч. 1).

Пример 1. Из равенств (8) § 1 видно, что выборочные моменты A_{nk} k-го порядка являются несмещенными оценками для теоретических моментов $\alpha_k = \mathrm{M}\,\xi^k$ k-го порядка, $k = 1, 2, \ldots$.

Оценка $T(X)$ называется *состоятельной* (*сильно состоятельной*) оценкой для параметра $\theta \in \Theta$, если при $n \to \infty$ $T(X) \longrightarrow \theta$ по вероятности (почти наверное).

Из тех же равенств (8) § 1 вытекает, что A_{nk} являются состоятельными оценками для α_k, $k = 1, 2, \ldots$. Используя усиленный закон больших чисел, можно показать, что A_{nk} – сильно состоятельные оценки для α_k. (Покажите это самостоятельно).

Пример 2. Исследуем на несмещенность выборочную дисперсию

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

Напомним, что случайные величины $X_1, \ldots, X_n$ независимы и имеют то же распределение, что и исследуемая случайная величина ξ. Так как $\mathrm{M}\left(X_i - \overline{X}\right) = 0$, то для любых $i = \overline{1,n}$

$$\mathrm{M}\left(X_i - \overline{X}\right)^2 = \mathrm{D}\left(X_i - \overline{X}\right) = \mathrm{D}\left(X_i - \frac{1}{n}\left(X_1 + \ldots + X_n\right)\right) = \frac{n-1}{n}\mathrm{D}\,\xi,$$

а поэтому

$$\mathrm{M}\,S^2 = \frac{1}{n}\sum_{i=1}^{n}\mathrm{M}\left(X_i - \overline{X}\right)^2 = \frac{n-1}{n}\mathrm{D}\,\xi.$$

Полученная формула показывает, что выборочная дисперсия является смещенной оценкой теоретической дисперсии. Однако смещение этой оценки

$$\mathrm{M}\, S^2 - \mathrm{D}\,\xi = -\frac{1}{n}\mathrm{D}\,\xi \to 0, \qquad n \to \infty,$$

так что выборочная дисперсия асимптотически несмещена. Кроме того, легко видоизменить величину S^2 так, чтобы она была несмещенной. Положим

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

Здесь S^2 уже несмещенная оценка теоретической дисперсии. При этом она остается сильно состоятельной (покажите это самостоятельно).

П р и м е р 3. Пусть $X = \left(X_1,\ X_2,\ \ldots,\ X_n\right)$ выборка из распределения Коши с плотностью

$$f\left(x,\ \theta\right) = \frac{1}{\pi}\frac{1}{1+\left(x+\theta\right)^2}, \qquad x \in R.$$

Точка $x = -\theta$ является центром симметрии для плотности. Таким образом, значение $x = -\theta$ играет роль «среднего значения» распределения, математическое ожидание которого не существует. Кажется заманчивым в качестве оценки параметра θ предложить выборочное среднее

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Однако такая оценка неудовлетворительна. Покажем это. Найдем распределение случайной величины $\overline{X}$. Так как характеристическая функция распределения Коши имеет вид

$$f_\xi\left(t\right) = \exp\left\{i\theta t - \left|t\right|\right\},$$

то для характеристической функции $f_{\overline{X}}\left(t\right)$ величины $\overline{X}$ имеем

$$f_{\overline{X}}\left(t\right) = \left[f_\xi\left(\frac{t}{n}\right)\right]^n = \exp\left\{i\theta t - \left|t\right|\right\} = f_\xi\left(t\right).$$

Отсюда можно сделать вывод, что случайная величина $\overline{X} - \theta$ имеет распределение, не зависящее от n, т. е. оценка $\overline{X}$ несостоятельна.

С другой стороны, значение θ является медианой данного распределения,

$$\int_{-\infty}^{\theta} f\left(x,\ \theta\right)dx = \int_{0}^{\infty}\frac{1}{\pi}\frac{1}{1+x^2}dx = \frac{1}{2},$$

и по теореме 4 § 1 оценка $Z_{n,\frac{1}{2}}$ является сильно состоятельной оценкой для θ, причем величина

$$\frac{2}{\pi}\sqrt{n}\left(Z_{n,\frac{1}{2}} - \frac{1}{2}\right)$$

асимптотически (0, 1)-нормальна.

П р и м е р 4. *Параметры нормального распределения.* Рассмотрим задачу оценивания параметров нормального распределения с плотностью

$$f(x,\ a,\ \sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}},\quad a\in R,\quad \sigma^2>0.$$

Плотность распределения выборки X (функция правдоподобия) равна

$$L(\vec{x};a;\sigma)=\left(2\pi\sigma^2\right)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-a)^2\right\}.$$

Предположим, что дисперсия σ^2 известна и оценивается математическое ожидание $a=\theta_1$. Имеем

$$\frac{\partial}{\partial\theta_1}\ln L\left(\vec{x};a;\sigma^2\right)=\frac{1}{\sigma^2}\sum_{k=1}^{n}(x_k-\theta_1).$$

Таким образом, выполнено условие (10) существования несмещенной эффективной оценки, и она имеет вид

$$T_1(X)=\frac{1}{n}\sum_{k=1}^{n}X_k=\bar{X}.$$

Следовательно, выборочное среднее $\bar{X}$ является несмещенной оценкой с минимальной дисперсией математического ожидания нормального распределения. Дисперсия этой оценки равна

$$\mathrm{D}\left(T_1(x)\right)=\frac{\sigma^2}{n}.$$

Предположим теперь, что среднее a известно и оценивается дисперсия $\sigma^2=\theta_2$. Имеем

$$\frac{\partial}{\partial\theta_2}\ln L(\vec{x};a;\theta_2)=\sum_{k=1}^{n}\frac{(x_k-a)^2}{2\theta_2^2}-\frac{n}{2}\frac{1}{\theta_2}.$$

Условие (10) существования эффективной оценки выполнено. Таким образом, оценка

$$T_2(X)=\frac{1}{n}\sum_{k=1}^{n}(X_k-a)^2$$

является эффективной. Ее дисперсия равна

$$\mathrm{D}\left(T_2(X)\right)=\frac{2\theta_2^2}{n}.$$

П р и м е р 5. *Оценка параметра биномиального распределения.* Имеем

$$f(x,\ \theta)=C_N^x\,\theta^x(1-\theta)^{N-x},\quad x=0,\ 1,\ \ldots,\ N,\quad \theta\in(0;1),$$

поэтому

$$L(\vec{x};\theta)=C_N^{x_1}\cdots C_N^{x_n}\,\theta^{x_1+\ldots+x_n}(1-\theta)^{nN-x_1-\ldots-x_n},$$

и

$$\frac{\partial\ln L(\vec{x};\theta)}{\partial\theta}=\frac{x_1+\ldots+x_n}{\theta}-\frac{Nn-x_1-\ldots-x_n}{1-\theta}=\frac{n}{\theta(1-\theta)}(\bar{x}-N\theta).$$

Таким образом, оценка

$$T(X) = \frac{1}{N}\bar{X}$$

является несмещенной эффективной оценкой θ.

П р и м е р 6. *Оценка параметра распределения Пуассона.* Пусть случайная величина ξ имеет распределение Пуассона с неизвестным параметром $\lambda = \theta > 0$. Тогда

$$f(x,\ \theta) = \mathrm{P}(\xi = x) = \frac{\theta^x}{x!}\, e^{-\theta},\ x = 0,\ 1,\ \ldots;\ \ \theta > 0,$$

$$L(\vec{x};\theta) = \prod_{k=1}^{n} f(x_k,\ \theta) = \frac{1}{x_1!\ \ldots\ x_n!}\, e^{-n\theta}\,\theta^{x_1 + \ldots + x_n}$$

и

$$\frac{\partial \ln L(\vec{x};\theta)}{\partial\theta} = \frac{1}{\theta}\sum_{k=1}^{n} x_k - n = \frac{n}{\theta}(\bar{x} - \theta).$$

Следовательно, в силу критерия эффективности (10) выборочное среднее $\bar{X}$ является несмещенной эффективной оценкой параметра θ с дисперсией

$$\mathrm{D}\,\bar{X} = \frac{1}{n}\mathrm{D}\,\xi = \frac{\theta}{n}.$$

П р и м е р 7. *Оценка параметра распределения Коши.* Пусть случайная величина ξ имеет распределение Коши с плотностью

$$f(x,\theta) = \frac{1}{\pi}\,\frac{1}{1+(x-\theta)^2},\quad \theta \in R,$$

тогда

$$\frac{\partial \ln L(\vec{x};\theta)}{\partial\theta} = 2\sum_{k=1}^{n}\frac{(x_k - \theta)}{1+(x_k-\theta)^2}.$$

В этом случае ни для какой функции $\tau(\theta)$ параметра θ не существует эффективной оценки. Заметим, что случайные величины

$$\frac{(X_k - \theta)}{1+(X_k-\theta)^2}$$

имеют средние нуль и дисперсии равные $\frac{\pi}{8}$. Отсюда мы получаем значение для фишеровского количества информации

$$i_n(\theta) = \mathrm{M}_\theta\left(\frac{\partial}{\partial\theta}L(\vec{x};\theta)\right)^2 = \frac{n}{2},$$

и нижняя граница Рао – Крамера дисперсии несмещенных оценок параметра θ в плотности Коши равна $\frac{2}{n}$. В прим. 3 было показано, что оценка $Z_{n,\frac{1}{2}}$ является состоятельной оценкой для параметра θ. Для дисперсии $Z_{n,\frac{1}{2}}$ при больших n имеем

$$D_\theta Z_{n,\frac{1}{2}} \approx \frac{\pi^2}{4n}, \quad \theta \in R.$$

Так как $\frac{\pi^2}{2} > 2,$ то оценка $Z_{n,\frac{1}{2}}$ не является асимптотически эффективной, т. е. в пределе при $n \to \infty$ не является эффективной.

2. Достаточные статистики

Прежде чем перейти к методам нахождения оценок неизвестных параметров, рассмотрим одно понятие, так называемых, достаточных статистик, которое во многих случаях полезно при оценивании параметров.

Основная идея достаточных статистик заключается в следующем. Пусть имеется выборка $X = (X_1, X_2, \ldots, X_n)$ объема n. Величина n может быть очень большой и поэтому будет большим объем экспериментальных данных. Возникает задача приведения этих данных в более компактный вид без потери информации, содержащейся в выборке, т. е., найти такие статистики $T_1 = T_1(X), \ldots, T_k = T_k(X),$ чтобы k было меньше $n,$ и которые давали бы столько же информации о неизвестном исследуемом параметре, сколько несет выборка $X.$ Такие функции $T = (T_1, \ldots, T_k)$ будем называть *достаточными статистиками*. Заметим, что сама выборка $X,$ естественно, является достаточной статистикой.

А теперь перейдем к строгому изложению этого материала.

Статистика $T = T(X)$ называется *достаточной* для модели $\mathscr{F} = \{F(x, \theta), \theta \in \Theta\}$ (или достаточной для параметра $\theta,$ когда ясно, о какой модели идет речь), если условная плотность распределения (или вероятность в дискретном случае) $L(\vec{x} \mid t; \theta)$ случайного вектора $X = (X_1, X_2, \ldots, X_n)$ при условии $T(X) = t$ не зависит от параметра $\theta.$ Это свойство статистики T означает, что она содержит всю информацию о параметре $\theta,$ имеющуюся в выборке, и поэтому все заключения об этом параметре, которые можно сделать при наблюдении $\vec{x},$ зависят только от $t = T(\vec{x}).$ Достаточная статистика, следовательно, дает оптимальный в определенном смысле способ представления статистических данных, что особенно важно при обработке больших массивов статистической информации.

Приведем критерий факторизации, позволяющий в каждом конкретном случае определить, существует ли достаточная статистика, и одновременно установить ее вид.

Теорема 2 (**Критерий факторизации**). *Для того, чтобы статистика* $T(X)$ *была достаточной для* θ, *необходимо и достаточно, чтобы функция правдоподобия* $L(\vec{x};\theta)$ *имела вид*

$$L(\vec{x};\theta)=g\big(T(\vec{x});\theta\big)h(\vec{x}),\tag{14}$$

где множитель g *может зависеть от* θ, *а от* $\vec{x}$ *зависит лишь через* $T(\vec{x})$, *множитель* h *от параметра* θ *не зависит.*

Доказательство (проведем для случая дискретной модели). Необходимость. Так как статистика T является достаточной, то $L(\vec{x}\,|\,t;\theta)=h(\vec{x},t)$. Пусть $T(\vec{x})=t$, тогда $\{X=\vec{x}\}\subseteq\{T(X)=t\}$, поэтому

$$L(\vec{x};\theta)=\mathrm{P}_\theta(X=\vec{x})=\mathrm{P}_\theta\big(X=\vec{x},\ T(X)=t\big)=\mathrm{P}_\theta\big(T(X)=t\big)\times$$
$$\times\mathrm{P}_\theta\big(X=\vec{x}\,|\,T(X)=t\big)=g(t;\theta)L(\vec{x}\,|\,t;\theta)=g\big(T(\vec{x});\theta\big)h(\vec{x},\ t).$$

Достаточность. Пусть имеет место (14), тогда для любых $\vec{x}$, таких, что $T(\vec{x})=t$

$$L(\vec{x}\,|\,t;\theta)=\mathrm{P}_\theta\big(X=\vec{x}\,|\,T(X)=t\big)=\frac{\mathrm{P}_\theta\big(X=\vec{x},\ T(X)=t\big)}{\mathrm{P}_\theta\big(T(X)=t\big)}=\frac{L(\vec{x};\theta)}{\sum\limits_{\vec{x}':T(\vec{x}')=t}L(\vec{x}';\theta)}=$$
$$=\frac{g(t;\theta)\,h(\vec{x})}{\sum\limits_{\vec{x}':T(\vec{x}')=t}g(t;\theta)\,h(\vec{x})}=\frac{h(\vec{x})}{\sum\limits_{\vec{x}':T(\vec{x}')=t}h(\vec{x}')},$$

т. е. $L(\vec{x}\,|\,t;\theta)$ не зависит от θ. Если же для $\vec{x}$ имеет место соотношение $T(\vec{x})\neq t$, то $L(\vec{x}\,|\,t;\theta)=0$. Таким образом, в любом случае условная вероятность $L(\vec{x}\,|\,t;\theta)$ не зависит от θ. Теорема доказана.

Достаточная статистика $T=T(X)$ называется *полной*, если для любой функции $\varphi(t)$ из того, что $M_\theta\varphi(T)=0$ для любых $\theta\in\Theta$ следует $\varphi(t)\equiv 0$ на всем множестве значений статистики T.

Замечание 2. Пусть в рассматриваемой модели $\mathscr{F}$ существует полная достаточная статистика T и требуется оценить заданную параметрическую функцию $\tau(\theta)$, тогда

1) если нет несмещенных оценок вида $H(T)$ (H – некоторая функция), то класс $\mathscr{T}_\tau$ несмещенных оценок пуст;

2) оптимальная оценка (когда она существует) всегда является функцией, зависящей от T;

3) оптимальную оценку τ^* можно искать по формуле $\tau^* = H(T) = \mathrm{M}_\theta(T_1 \mid T)$, исходя из любой несмещенной оценки T_1 функции $\tau(\theta)$ [14].

П р и м е р 8. Пусть исследуемая случайная величина – экспоненциальная случайная величина с параметром $\lambda = \dfrac{1}{\theta} > 0$. Тогда

$$f(x,\ \theta) = \frac{1}{\theta}\exp\left\{-\frac{x}{\theta}\right\}, \quad x \ge 0,$$

$$L(\vec{x};\theta) = \frac{1}{\theta^n}\exp\left\{-\frac{1}{\theta}\sum_{k=1}^{n} x_k\right\}.$$

Итак, функция правдоподобия $L(\vec{x};\theta)$ зависит только от

$$\sum_{k=1}^{n} x_k,$$

следовательно, статистика

$$T(X) = \sum_{k=1}^{n} X_k$$

в силу критерия факторизации (14) является достаточной статистикой.

П р и м е р 9. Пусть ξ – нормальная случайная величина с неизвестными параметрами $\theta_1 = a$, $\theta_2 = \sigma^2$. Тогда

$$L(\vec{x};\theta_1;\theta_2) = \frac{1}{\sqrt{2\pi\theta_2}}\ \exp\left\{-\frac{1}{2\theta_2}\sum_{k=1}^{n} x_k^2 + \frac{\theta_1}{\theta_2}\sum_{k=1}^{n} x_k - \frac{n\theta_1^2}{2\theta_2}\right\}$$

и по критерию факторизации статистики

$$T_1(X) = \sum_{k=1}^{n} X_k \qquad \text{и} \qquad T_2(X) = \sum_{k=1}^{n} X_k^2$$

являются достаточными.

П р и м е р 10. Пусть случайная величина ξ равномерно распределена на отрезке $[\theta_1;\theta_2]$. Тогда

$$L(\vec{x};\theta_1;\theta_2) = \begin{cases} \dfrac{1}{b-a}, & \theta_1 \le x_{(1)} < x_{(n)} \le \theta_2, \\ 0, & \text{в противном случае,} \end{cases}$$

где $x_{(1)} = \min\{x_1,\ \ldots,\ x_n\}$, $x_{(n)} = \max\{x_1,\ \ldots,\ x_n\}$, т. е. порядковые статистики $T_1(X) = X_{(1)}$ и $T_2(X) = X_{(n)}$ являются достаточными в данном случае.

3. Методы оценивания неизвестных параметров

1) М е т о д м а к с и м а л ь н о г о п р а в д о п о д о б и я

Метод был предложен Р. Фишером. Изложим суть этого метода.

Пусть имеется выборка $X = (X_1, X_2, \ldots, X_n)$ из распределения $\mathscr{L}(\xi) \in \mathscr{F} = \{F(x, \theta), \theta \in \Theta\}$ и $L(\vec{x};\theta)$ – функция правдоподобия для реализации $\vec{x} = (x_1, \ldots, x_n)$ выборки X. *Оценкой максимального правдоподобия* $\hat{\theta}$ параметра θ называется такая точка параметрического множества Θ, в которой функция правдоподобия $L(\vec{x};\theta)$ при заданном $\vec{x}$ достигает максимума, т. е.

$$L(\vec{x};\hat{\theta}) = \sup_{\theta\in\Theta} L(\vec{x};\theta).$$

Если для каждого $\vec{x}$ из выборочного пространства $\mathscr{X}$ максимум $L(\vec{x};\theta)$ достигается во внутренней точке Θ и $L(\vec{x};\theta)$ дифференцируема по θ, то оценка максимального правдоподобия $\hat{\theta}$ удовлетворяет уравнению

$$\frac{\partial L(\vec{x};\theta)}{\partial\theta} = 0,$$

или

$$\frac{\partial \ln(L(\vec{x};\theta))}{\partial\theta} = 0. \qquad (15)$$

Если $\vec{\theta}$ – векторный параметр: $\vec{\theta} = (\theta_1, \ldots, \theta_n)$, то это уравнение заменяется системой

$$\frac{\partial \ln L(\vec{x};\vec{\theta})}{\partial\theta_i} = 0, \quad i = \overline{1,n}. \qquad (16)$$

Уравнения (15) и (16) называются *уравнениями правдоподобия*.

З а м е ч а н и е 3 (см. [14]). Отметим свойства оценки максимального правдоподобия $\hat{\theta}$:

1) если существует эффективная оценка $T(X)$ для скалярного параметра θ, то $\hat{\theta} = T(X)$;

2) если имеется достаточная статистика $T = T(X)$, и оценка максимального правдоподобия $\hat{\theta}$ существует и единственна, то она является функцией от T;

3) $\hat{\theta}$ является состоятельной оценкой параметра θ, т. е.

$$\hat{\theta} \xrightarrow[n\to\infty]{P_\theta} \theta;$$

4) $\hat{\theta}$ является асимптотически несмещенной оценкой, т. е. ее смещение (см. (2)) при $n \to \infty$ стремится к 0;

5) $\hat{\theta}$ является асимптотически эффективной оценкой, т. е. в пределе при $n \to \infty$ – эффективная оценка.

2) М е т о д м о м е н т о в

Исторически первым общим методом точечного оценивания неизвестных параметров является метод моментов, предложенный К. Пирсоном. Суть этого метода состоит в следующем.

Пусть $X = \left(X_1, X_2, \ldots, X_n\right)$ – выборка из распределения $\mathscr{L}\left(\xi\right) \in \mathscr{F} = \left\{F(x;\vec{\theta}),\ \vec{\theta} \in \Theta\right\}$, где $\vec{\theta} = \left(\theta_1, \ldots, \theta_r\right)$ и $\Theta \subseteq R^r$. Предположим, что у наблюдаемой случайной величины ξ существуют первые r моментов $\alpha_k = \mathrm{M}\xi^k$, $k = \overline{1,r}$. Они являются функциями от неизвестных параметров $\vec{\theta}: \alpha_k = \alpha_k\left(\vec{\theta}\right)$. Рассмотрим соответствующие выборочные моменты $A_{nk}\left(X\right)$ (см. (6) § 1). Пусть a_{nk} значения этих величин для наблюдавшейся реализации $\vec{x}$ выборки *X*. Тогда метод моментов состоит в приравнивании значений a_{nk} и теоретических моментов:

$$\alpha_k\left(\vec{\theta}\right) = a_{nk}, \quad k = \overline{1,r}. \tag{17}$$

Решая эти уравнения относительно $\theta_1, \ldots, \theta_r$ получаем значения оценок параметров по методу моментов.

Использование начальных моментов не обязательно, можно использовать также центральные, абсолютные моменты и т. п.

З а м е ч а н и е 4. Метод моментов при определенных условиях приводит к состоятельным оценкам, при этом уравнения (17) во многих случаях просты и их решения не связано с большими вычислительными трудностями. Отметим, однако, что метод моментов неприменим, когда теоретические моменты нужного порядка не существуют. Кроме того, оценки метода моментов, вообще говоря, не эффективны, поэтому их часто используют только в качестве первых приближений.

П р и м е р 11. Пусть ξ – случайная величина Пуассона с неизвестным параметром $\lambda = \theta > 0$, т. е.

$$f\left(x,\ \theta\right) = \mathrm{P}\left(\xi = x\right) = \frac{\theta^x}{x!}e^{-\theta}, \quad x = 0, 1, 2, \ldots; \quad \theta > 0,$$

$$L(\vec{x};\theta)=\frac{1}{x_1!\dots x_n!}\theta^{x_1+\dots+x_n}e^{-n\theta}.$$

Оценим неизвестный параметр θ, используя метод максимального правдоподобия. Запишем уравнение правдоподобия

$$\frac{\partial\ln L(\vec{x};\theta)}{\partial\theta}=0$$

или

$$(x_1+\dots+x_n)\frac{1}{\theta}-n=0.$$

Отсюда находим решение

$$\theta=\frac{x_1+\dots+x_n}{n}=\bar{x},$$

т. е. оценка максимального правдоподобия $\hat{\theta}_1$ равна $\hat{\theta}_1=\overline{X}$.

Оценим неизвестный параметр θ, используя метод моментов. Для этого приравняем математическое ожидание $\mathrm{M}_\theta\,\xi$ и реализацию выборочного среднего $\bar{x}$. Так как

$$\mathrm{M}_\theta\xi=\sum_{k=0}^{\infty}k\,\mathrm{P}(\xi=k)=\sum_{k=0}^{\infty}k\,\frac{\theta^k}{k!}e^{-\theta}=\theta,\quad \theta>0,$$

то искомое уравнение

$$\theta=\bar{x}$$

и, следовательно, оценка, полученная методом моментов, совпадает с оценкой максимального правдоподобия:

$$\hat{\theta}_2=\overline{X}.$$

П р и м е р 12. Используя метод максимального правдоподобия и метод моментов, оценим неизвестный параметр θ, если случайная величина ξ равномерно распределена на отрезке $[0;\theta]$, $\theta>0$. Имеем

$$f(x,\,\theta)=\begin{cases}\dfrac{1}{\theta}, & x\in[0;\theta],\\ 0, & x\notin[0;\theta]\end{cases}$$

и

$$L(\vec{x};\theta)=\begin{cases}\dfrac{1}{\theta^n}, & 0\le x_1,\,\dots,\,x_n\le\theta,\\ 0, & \text{в противном случае.}\end{cases}$$

Функция правдоподобия $L(\vec{x};\theta)$ не является дифференцируемой по θ, тем не менее, несложно видеть, что при заданном $\vec{x}$ она достигает максимума в точке $\theta=\max(x_1,\,\dots,\,x_n)$. Таким образом оценка максимального правдоподобия имеет вид

$$\hat{\theta}_1=\max(X_1,\,\dots,\,X_n)=X_{(n)}.$$

Применим метод моментов. Для этого приравняем теоретическое математическое ожидание

$$M_\theta \xi = \int_{-\infty}^{\infty} x f(x, \theta)\, dx = \int_0^\theta \frac{x}{\theta}\, dx = \frac{\theta}{2}$$

и реализацию выборочного среднего $\bar{x}$, тогда

$$\frac{\theta}{2} = \bar{x}$$

или $\theta = 2\bar{x}$. Следовательно, оценка, полученная методом моментов, имеет следующий вид

$$\hat{\theta}_2 = 2\bar{X}.$$

У п р а ж н е н и е 1. Проверьте свойства оценок $\hat{\theta}_1$ и $\hat{\theta}_2$, учитывая замеч. 3 и 4. Какая из них лучше?

4. Интервальное оценивание

Выше были рассмотрены точечные оценки для параметра θ исследуемой модели $\mathscr{L}(\xi) \in \mathscr{F} = \{F(x, \theta),\ \theta \in \Theta\}$ и функций этого параметра. Любая точечная оценка представляет собой функцию $T = T(X)$ выборки $X = (X_1, X_2, \ldots, X_n)$, т. е. является случайной величиной, и при каждой реализации $\vec{x}$ выборки X эта функция определяет единственное значение $t = T(\vec{x})$ оценки, принимаемое за приближенное значение оцениваемой характеристики. При этом в каждом конкретном случае значение оценки может отличаться от значения параметра, следовательно, полезно знать и возможную погрешность, возникающую при использовании предлагаемой оценки. Например, указать такой интервал (или область в случае векторного параметра), внутри которого с высокой вероятностью γ находится точное значение оцениваемого параметра. В этом случае говорят об *интервальном* или *доверительном оценивании*, а соответствующий интервал называется *доверительным*. Величину γ выбирают заранее.

Пусть θ – скалярный параметр. При интервальном оценивании ищут две такие статистики $T_1 = T_1(X)$ и $T_2 = T_2(X)$, что $T_1 < T_2$, для которых при заданном $\gamma \in (0;1)$ выполняется условие

$$P_\theta\left(T_1(X) < \theta < T_2(X)\right) \geq \gamma \tag{18}$$

для любого $\theta \in \Theta$.

В этом случае интервал $(T_1(X); T_2(X))$ называют γ-*доверительным интервалом* (для θ), число γ – *доверительным уровнем,* или *доверительной вероятностью*, $T_1(X)$ и $T_2(X)$ – *нижней* и *верхней* границами соответственно.

На практике обычно используют значения доверительного уровня γ из небольшого набора заранее выбранных, достаточно близких к единице значений, например, $\gamma = 0{,}9;\ 0{,}95;\ 0{,}99$ и т. д. (при этом $\gamma \neq 1$).

Аналогично определяется доверительный интервал для отдельной компоненты (например, θ_1) в случае многомерного параметра $\overline{\theta}$:

$$\mathrm{P}_{\overline{\theta}}\left(T_1(X) < \theta_1 < T_2(X)\right) \geq \gamma$$

для любого $\overline{\theta} \in \Theta$, а также доверительный интервал для скалярной параметрической функции $\tau(\theta)$ (θ может быть как скаляром, так и вектором),

$$\mathrm{P}_{\theta}\left(T_1(X) < \tau(\theta) < T_2(X)\right) \geq \gamma$$

для любого $\theta \in \Theta$.

Общий прием, с помощью которого в ряде случаев можно построить доверительный интервал, состоит в следующем: пусть модель $\mathscr{F}$ абсолютно непрерывна и существует случайная величина $G(X,\ \theta)$ такая, что:

1) распределение $G(X, \theta)$ не зависит от θ;

2) при каждом $\vec{x} \in \mathscr{X}$ функция $G(\vec{x},\ \theta)$ непрерывна и строго монотонна по θ.

Такую случайную величину называют *центральной статистикой* (для θ). Аналогично определяется центральная статистика для отдельной компоненты в случае многомерного параметра $\vec{\theta}$, а также для скалярной параметрической функции. Далее будем рассматривать для краткости лишь случай скалярного параметра θ.

Пусть для модели $\mathscr{F}$ построена центральная статистика $G(X,\ \theta)$ и $f_G(g)$ – ее плотность распределения. Функция $f_G(g)$ от параметра θ не зависит [условие 1)], поэтому для любого $\gamma \in (0;1)$ можно выбрать величины $g_1 < G(\vec{x},\ \theta) < g_2$ (вообще говоря, многими способами; в некоторых моделях единственность выбора обеспечивается требованием минимальности длины отрезка $[g_1; g_2]$) так, чтобы

$$\mathrm{P}_\theta\left(g_1 < G(X,\,\theta) < g_2\right) = \int\limits_{g_1}^{g_2} f_G(g)\,dg = \gamma. \tag{19}$$

Определим теперь при каждом $\vec{x} \in \mathscr{X}$ числа $T_1(\vec{x})$ и $T_2(\vec{x})$, где $T_1(\vec{x}) < T_2(\vec{x})$, как решения относительно θ уравнений $G(\vec{x},\,\theta) = g_1$, $G(\vec{x},\,\theta) = g_2$; однозначность определения этих чисел обеспечивается условием 2). Тогда неравенства $g_1 < G(\vec{x},\,\theta) < g_2$ эквивалентны неравенствам $T_1(\vec{x}) < \theta < T_2(\vec{x})$ и, следовательно, (19) можно переписать в виде

$$\mathrm{P}_\theta\left(T_1(X) < \theta < T_2(X)\right) = \gamma$$

для любого $\theta \in \Theta$.

Таким образом, построенный интервал $\left(T_1(\vec{x}); T_2(\vec{x})\right)$ является γ-доверительным интервалом для θ.

З а м е ч а н и е 5 (см. [14]). Пусть по выборке

$$X = (X_1,\, X_2,\, \ldots,\, X_n)$$

требуется построить доверительные интервалы для μ и σ^2 в модели $N(\mu,\,\sigma^2)$. Применяя вышепредложенную методику, получим:

1) если μ известно, а требуется оценить σ^2, то промежуток

$$\left(\zeta^{-1}_{\frac{1+\gamma}{2}} \sum_{i=1}^{n} (X_i - \mu);\; \zeta^{-1}_{\frac{1-\gamma}{2}} \sum_{i=1}^{n} (X_i - \mu)\right) \tag{20}$$

является γ-доверительным интервалом для σ^2; здесь ζ_γ – γ-квантиль распределения χ^2_n (χ^2_n – хи-квадрат распределение с n степенями свободы);

2) если σ^2 известно, то

$$\left(\overline{X} - \sigma\frac{\zeta_{\frac{1+\gamma}{2}}}{\sqrt{n}};\; \overline{X} + \sigma\frac{\zeta_{\frac{1+\gamma}{2}}}{\sqrt{n}}\right) \tag{21}$$

есть γ-доверительный интервал для μ; здесь ζ_γ – γ-квантиль распределения $N(0,\,1)$, где $N(0,\,1)$ – нормальное распределение с параметрами $(0,\,1)$;

3) если оба параметра μ и σ^2 неизвестны, то

$$\left(\bar{X} - \zeta_{\frac{1+\gamma}{2}} \frac{S}{\sqrt{n-1}};\ \bar{X} + \zeta_{\frac{1+\gamma}{2}} \frac{S}{\sqrt{n-1}} \right) \qquad (22)$$

является γ-доверительным интервалом для μ, здесь ζ_γ – γ-квантиль t_{n-1}-распределения (t_{n-1}-распределение Стьюдента с $n-1$ степенью свободы), S^2 – выборочная дисперсия, а

$$\left(\frac{nS^2}{\zeta_{\frac{1+\gamma}{2}}};\ \frac{nS^2}{\zeta_{\frac{1-\gamma}{2}}} \right) \qquad (23)$$

есть γ-доверительный интервал для σ^2, здесь ζ_γ – γ-квантиль распределения χ^2_{n-1}.

З а м е ч а н и е 6. Значения γ-квантилей распределений хи-квадрат, Стьюдента и нормального протабулированы (см. Приложения, табл. 6, 5, 3).

П р и м е р 13. Д о в е р и т е л ь н ы е и н т е р в а л ы д л я в е р о я т н о с т и у с п е х а в с х е м е Б е р н у л л и. Пусть случайная величина μ – число «успехов» при n испытаниях в схеме Бернулли. Функция распределения

$$F(m,\ \theta) = \mathrm{P}(\mu \le m) = \sum_{k=0}^{m} C_n^k \theta^k (1-\theta)^{n-k},$$

рассматриваемая в целых точках $m = 0, 1, 2, \ldots, n,$ убывает с ростом $\theta,$ так как

$$\frac{d}{d\theta} F(m,\ \theta) = -n C_{n-1}^m \theta^m (1-\theta)^{n-m-1} < 0.$$

Обозначим $m_\gamma(\theta)$ такое наименьшее целое число, для которого

$$F\left(m_\gamma(\theta),\ \theta\right) \ge \gamma,$$

тогда $m_\gamma(\theta) - 1$ есть такое наибольшее целое число, что

$$F\left(m_\gamma(\theta) - 1,\ \theta\right) < \gamma.$$

Пусть $\alpha_1 + \alpha_2 = 1 - \gamma$, тогда с вероятностью большей либо равной γ

$$m_{1-\alpha_1}(\theta) \le \mu \le m_{\alpha_2}(\theta).$$

Обозначим решение уравнения $y = m_\gamma(\theta)$ относительно θ через $m_\gamma^{-1}(y)$. Тогда неравенства

$$m_{\alpha_2}^{-1}(\mu) < \theta < m_{1-\alpha_1}^{-1}(\mu)$$

задают доверительный интервал для θ с уровнем доверия γ. Для нахождения границ доверительных интервалов можно пользоваться таблицами биномиального распределения, но такие таблицы громоздки и не очень распространены. Поэтому часто используют предельную теорему Муавра – Лапласа об асимптотической нормальности случайной величины

$$\eta = \frac{\mu - n\theta}{\sqrt{n\theta(1-\theta)}}$$

для построения приближенных доверительных интервалов. Неравенство

$$|\eta| \le \zeta_{\frac{1-\gamma}{2}}$$

при больших n выполняется с вероятностью примерно равной γ, а если учесть, что при $n \to \infty$

$$\frac{\mu}{n} \xrightarrow{п.н.} \theta,$$

т. е.

$$\theta \approx \frac{\mu}{n},$$

то доверительный интервал для θ имеет вид

$$\frac{\mu}{n} - \zeta_{\frac{1-\gamma}{2}} \sqrt{\frac{\mu}{n^2}\left(1 - \frac{\mu}{n}\right)} < \theta < \frac{\mu}{n} + \zeta_{\frac{1-\gamma}{2}} \sqrt{\frac{\mu}{n^2}\left(1 - \frac{\mu}{n}\right)},$$

здесь $\zeta_{\frac{1-\gamma}{2}}$ – $\frac{1-\gamma}{2}$-квантиль $(0, 1)$- нормального распределения.

Существуют и другие подходы к построению приближенного доверительного интервала для θ (см. [34]).

Пример 14. Найдем доверительный интервал с доверительной вероятностью γ для параметра θ случайной величины Пуассона ξ. Имеем

$$\mathrm{P}(\xi = k) = \frac{\theta^k}{k!} e^{-\theta}, \quad \theta > 0.$$

Пусть $X = (X_1, X_2, \ldots, X_n)$ – выборка из распределения Пуассона. Найдем центральную статистику для θ. При этом используем очевидный факт, что кратчайший доверительный интервал обеспечивает эффективная оценка (у нее наименьшая дисперсия; см., также [7]). Поскольку

$$L(\vec{x}, \theta) = e^{-\theta n} \prod_{k=1}^{n} \frac{\theta^{x_k}}{x_k!},$$

то

$$\frac{\partial}{\partial \theta} \ln L(\vec{x}, \theta) = \frac{n}{\theta}(\bar{x} - \theta).$$

В качестве статистики возьмем

$$G(X, \theta) = \sqrt{\frac{n}{\theta}}(\bar{X} - \theta).$$

Так как она асимптотически нормальна с параметрами $(0, 1)$, то пользуясь нормальным распределением, определим приближенные γ-доверительные границы

g_1, g_2 (см. (19)). При этом учитываем, что доверительный интервал $(g_1; g_2)$ будет иметь наименьшую длину, если $g_2 = -g_1 = g$, т. е. имеем

$$\frac{1}{\sqrt{2\pi}}\int_{-g}^{g} e^{-\frac{x^2}{2}}dx = \gamma$$

или

$$\frac{1}{\sqrt{2\pi}}\int_{0}^{g} e^{-\frac{x^2}{2}}dx = \frac{\gamma}{2},$$

таким образом $g = \zeta_{\frac{1+\gamma}{2}}$ – квантиль нормального распределения.

Решая неравенство

$$\sqrt{\frac{n}{\theta}}\left|X-\theta\right| \le \zeta_{\frac{1+\gamma}{2}}$$

относительно θ, получаем

$$X + \frac{\zeta^2_{\frac{1+\gamma}{2}}}{2n} - \zeta_{\frac{1+\gamma}{2}}\sqrt{\frac{X}{n} + \frac{\zeta^2_{\frac{1+\gamma}{2}}}{4n^2}} < \theta < X + \frac{\zeta^2_{\frac{1+\gamma}{2}}}{2n} + \zeta_{\frac{1+\gamma}{2}}\sqrt{\frac{X}{n} + \frac{\zeta^2_{\frac{1+\gamma}{2}}}{4n^2}}$$

есть γ-доверительный интервал для параметра θ.

Пример 15. Задача о двух нормальных выборках. Пусть $X = (X_1, X_2, \ldots, X_n)$ и $Y = (Y_1, Y_2, \ldots, Y_m)$ две независимые нормальные выборки с параметрами (a_1, σ_1^2) и (a_2, σ_2^2) соответственно. На практике часто надо будет знать, насколько значима разность между средними в этих выборках. Иными словами, нас будут интересовать доверительные интервалы для разности $a_1 - a_2$. Эту задачу называют *задачей о сравнении двух средних*. Рассмотрим два варианта этой задачи.

а) Пусть известны дисперсии σ_1^2 и σ_2^2. Так как случайные величины

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \overline{Y} = \frac{1}{m}\sum_{j=1}^{m} Y_j$$

независимы и нормальны с параметрами

$$\left(a_1, \frac{1}{n}\sigma_1^2\right) \text{ и } \left(a_2, \frac{1}{m}\sigma_2^2\right),$$

то разность $\overline{X} - \overline{Y}$ – нормальна с математическим ожиданием $a_1 - a_2$ и дисперсией

$$\left(\frac{1}{n}\sigma_1^2 + \frac{1}{m}\sigma_2^2\right).$$

Поэтому доверительные интервалы для $a_1 - a_2$ имеют вид

$$\overline{X} - \overline{Y} - t\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < a_1 - a_2 < \overline{X} - \overline{Y} + t\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}},$$

где t – γ-доверительная граница для стандартной нормальной случайной величины η, т. е. $\mathrm{P}\left(|\eta|<t\right)=\gamma$ (сравните с п. 2 замеч. 5).

б) Пусть дисперсии выборок неизвестны, но известно, что $\sigma_1^2=\sigma_2^2=\sigma^2$. Положим

$$S_1^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2, \quad S_2^2=\frac{1}{m-1}\sum_{j=1}^{m}\left(Y_j-\bar{Y}\right)^2.$$

Несложно видеть, что случайная величина

$$\frac{1}{\sigma^2}\left[(n-1)S_1^2+(m-1)S_2^2\right]$$

имеет χ^2 распределение с $n+m-2$ степенями свободы и не зависит от случайных величин $\bar{X}$ и $\bar{Y}$ (см. [7]). Учитывая, что разность $\left(\bar{X}-a_1\right)-\left(\bar{Y}-a_2\right)$ является нормальной случайной величиной с параметрами $\left(0,\ \sigma^2\left(\frac{1}{n}+\frac{1}{m}\right)\right)$, видим, что статистика

$$\tau=\frac{\bar{X}-\bar{Y}-a_1+a_2}{\sqrt{(n-1)S_1^2+(m-1)S_2^2}}\sqrt{\frac{nm}{m+n}}$$

имеет распределение Стьюдента с $n+m-2$ степенями свободы. Поэтому γ-доверительный интервал для разности a_1-a_2 имеет вид

$$\bar{X}-\bar{Y}-t\sqrt{(n-1)S_1^2+(m-1)S_2^2}\sqrt{\frac{m+n}{mn}}<a_1-a_2<$$

$$<\bar{X}-\bar{Y}+t\sqrt{(n-1)S_1^2+(m-1)S_2^2}\sqrt{\frac{m+n}{mn}},$$

где t находится из уравнения $\mathrm{P}\left(|\tau|<t\right)=\gamma$.

Другой задачей, связанной с двумя выборками, является *задача о сравнении дисперсий*. В этом случае интересуются доверительными интервалами для отношения дисперсий $\left(\sigma_1^2/\sigma_2^2\right)$. Оценкой такого отношения естественно принять случайную величину S_1^2/S_2^2.

Так как отношение

$$Z=\frac{nS_1^2}{\sigma_1^2}:\frac{mS_2^2}{\sigma_2^2}$$

равно отношению двух независимых случайных величин χ_{n-1}^2 и χ_{m-1}^2, то случайная величина Z имеет распределение Фишера с $(n-1,\ m-1)$ степенями свободы. Найдя для величины Z доверительные границы c_1 и c_2 так, чтобы

$$\mathrm{P}\left(Z\le c_1\right)=\frac{1+\gamma}{2}, \quad \mathrm{P}\left(Z\ge c_2\right)=\frac{1+\gamma}{2},$$

получим для отношения $\left(\sigma_1^2/\sigma_2^2\right)$ следующий интервал

$$\frac{1}{c_2}\frac{nS_1^2}{mS_2^2} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{1}{c_1}\frac{nS_1^2}{mS_2^2}.$$

§ 3. Проверка статистических гипотез

1. Понятие статистической гипотезы и статистического критерия

Статистической гипотезой называют любое утверждение о виде или свойствах распределения наблюдаемых в эксперименте случайных величин. Если для исследуемого явления (процесса, ситуации и т. д.) сформулирована та или иная гипотеза (обычно ее называют *основной* или *нулевой гипотезой* и обозначают символом H_0), то задача состоит в том, чтобы сформулировать такое правило, которое бы позволило по результатам соответствующих наблюдений применять или отклонить эту гипотезу. Правило, согласно которому проверяемая гипотеза H_0 принимается или отвергается, называется *статистическим критерием* проверки гипотезы H_0. Разработка таких правил и их обоснование с точки зрения требований оптимальности и составляет предмет теории проверки статистических гипотез.

В дальнейшем будет формулироваться только одна гипотеза и необходимо будет проверять, согласуются ли имеющиеся статистические данные с этой гипотезой или же они ее опровергают. Соответствующие статистические критерии называют *критериями согласия*. Если гипотеза H_0 однозначно фиксирует распределение наблюдений, то ее называют *простой*, в противном случае – *сложной*.

Рассмотрим общий метод построения критериев согласия. Пусть о распределении случайной величины $X = (X_1, X_2, \ldots, X_n)$, описывающей результат изучаемого эксперимента, сформулирована некоторая гипотеза H_0. Чтобы построить критерий проверки этой гипотезы, пытаются найти такую статистику $T = T(X)$, характеризующую отклонение эмпирических данных от соответствующих (гипотезе H_0) гипотетических значений (обычно такие статистики неотрицательны), распределение которой в случае справедливости H_0 можно было бы определить. Предположим, что такая статистика и ее распределение при гипотезе H_0 найдены.

Пусть $\mathscr{T} = \{t : t = T(\vec{x}), x \in \mathscr{X}\}$, определим для фиксированного заранее достаточно малого числа $\alpha > 0$ подмножество $\mathscr{T}_{1\alpha} \subset \mathscr{T}$ так, чтобы

$$\mathrm{P}\left(T(X) \in \mathscr{T}_{1\alpha} \middle| H_0\right) \le \alpha. \quad (1)$$

Тогда правило гипотезы H_0 можно сформулировать следующим образом. Пусть $\vec{x}$ – наблюдавшаяся реализация случайной величины X и $t = T(\vec{x})$. *Если $t \in \mathscr{T}_{1\alpha}$, гипотеза H_0 должна быть отвергнута как противоречащая статистическим данным, если $t \notin \mathscr{T}_{1\alpha}$, то нет оснований отказываться от принятой гипотезы*.

В описанной методике статистику T называют *статистикой критерия*, $\mathscr{T}_{1\alpha}$ – *критической областью* для гипотезы H_0, число α – *уровнем значимости* критерия (в конкретных случаях величину α выбирают равной 0,1; 0,05; 0,01; и т. д.)

Для проверки одной и той же гипотезы H_0 можно строить различные критерии согласия, основываясь на разных статистиках $T(X)$ и, чтобы выбрать в конкретной ситуации тот или иной критерий, надо иметь принципы сравнения различных критериев.

Любое распределение F наблюдаемой случайной величины X, которое может оказаться истинным, но отличающееся от распределения при гипотезе H_0, называют *альтернативным распределением* или *альтернативой*. Совокупность всех альтернативных распределений называют *альтернативной гипотезой* и обозначают H_1. Величину

$$W(F) = W(\mathscr{T}_{1\alpha}, F) = \mathrm{P}\left(T(X) \in \mathscr{T}_{1\alpha} \middle| F - \text{истинно}\right)$$

называют *функцией мощности критерия*. Для распределения $F \in H_0$ в силу (1) справедливо неравенство $W(F) \le \alpha$ для любого F.

Если $F \in H_1$, то значение $W(F)$ называют *мощностью критерия при альтернативе F*.

В терминах функции $W(F)$ можно сказать, что критерий тем лучше («тем мощнее»), чем больше его мощность при альтернативах.

2. Проверка гипотезы о виде распределения

Пусть $X = (X_1, X_2, \ldots, X_n)$ – выборка из распределения $\mathscr{L}(\xi)$ с неизвестной функцией распределения $F_\xi(x)$, $x \in R$, о которой выдвинута гипотеза

$$H_0 : F_\xi(x) = F(x), x \in R.$$

2.1. Критерий согласия Колмогорова применяют в тех случаях, когда функция распределения $F(x)$ непрерывна. Статистика критерия

$$\mathrm{D}_n = \mathrm{D}_n(X) = \sup_{-\infty < X < +\infty} \left| F_n(X) - F(X) \right| \tag{2}$$

представляет собой максимальное отклонение эмпирической функции распределения $F_n(X)$ от гипотетической функция распределения $F(X)$. Из теоремы 2 § 1 видно, что, по крайней мере при больших n, в тех случаях, когда гипотеза H_0 истинна, значение D_n не должно существенно отклоняться от нуля. Отсюда следует, что критическую область критерия, основанного на статистике D_n, следует задавать в виде $\mathscr{T}_{1\alpha} = \{t \geq t_\alpha\}$. Отметим особенности статистики D_n:

1) ее распределение при гипотезе H_0 не зависит от вида функции $F(X)$. Действительно, полагая в (2) $x = F^{-1}(u)$, $0 \leq u \leq 1$, получаем

$$\mathrm{D}_n = \sup_{0 \leq u \leq 1} \left| F_n\left(F^{-1}(u)\right) - u \right|,$$

но

$$F_n\left(F^{-1}(u)\right) = \frac{1}{n} \sum_{i=1}^{n} I_{\left\{X_i \leq F^{-1}(u)\right\}} = \Phi_n(u),$$

где $\Phi_n(u)$ – эмпирическая функция распределения для выборки из равномерного на $[0;1]$ распределения, поэтому

$$\mathrm{D}_n = \sup_{0 \leq u \leq 1} \left| \Phi_n(u) - u \right|.$$

Тем самым достаточно вычислить и протабулировать распределение D_n только один раз, а именно для выборки из равномерного распределения;

2) из теоремы Колмогорова [14], утверждающей, что для непрерывной $F(x)$ при любом $t > 0$

$$\lim_{n \to \infty} \mathrm{P}\left(\sqrt{n}\, \mathrm{D}_n \leq t\right) = K(t) = \sum_{i=-\infty}^{+\infty} (-1)^j e^{-2j^2t^2}, \tag{3}$$

следует, что распределение статистики D_n при достаточно больших n (уже при $n \geq 20$) практически от n не зависит и имеет вид (3). Отсюда следует, что при $n \geq 20$ критическую границу $t_\alpha = t_\alpha(n)$ можно полагать равной $\frac{\lambda_\alpha}{\sqrt{n}}$, где $K(\lambda_\alpha) = 1 - \alpha$. Действительно, в этом случае

$$\mathrm{P}\left(\mathrm{D}_n \in \mathcal{T}_{1\alpha} \mid H_0\right) = \mathrm{P}\left(\sqrt{n}\;\; \mathrm{D}_n \geq \lambda_\alpha \mid H_0\right) \approx 1 - K\left(\lambda_\alpha\right) = \alpha.$$

Таким образом, при заданном уровне значимости α число λ_α определяют из соотношения $K\left(\lambda_\alpha\right) = 1 - \alpha$ и в этом случае правило проверки гипотезы H_0 формулируется (при $n \geq 20$) следующим образом: *если наблюдавшееся значение* $t = \mathrm{D}_n\left(\vec{x}\right)$ *статистики* (2) *удовлетворяет неравенству* $t\sqrt{n} \geq \lambda_\alpha$, *то гипотезу* H_0 *отвергают; в противном случае считают, что статистические данные не противоречат гипотезе.*

Для небольших значений n точное распределение D_n протабулировано и для расчета можно пользоваться соответствующими таблицами (см. [4]).

На практике вычисление статистики D_n – трудоемкая задача, поэтому часто применяют другой критерий, называемый критерием χ^2.

2.2. К р и т е р и й   с о г л а с и я χ^2 К. П и р с о н а можно использовать для любых распределений, в том числе и многомерных. Чтобы им воспользоваться, выборочные данные предварительно группируют следующим образом. Разобьем множество $\mathcal{E}$ возможных значений наблюдаемой случайной величины ξ на N непересекающихся подмножеств $\mathcal{E}_1, \ldots, \mathcal{E}_N$. Пусть $v = \left(v_1, \ldots, v_N\right)$ – вектор частот попадания выборочных точек в соответствующие интервалы группировки $\mathcal{E}_1, \ldots, \mathcal{E}_N$ $\left(v_1 + \ldots + v_N = n\right)$ и $p^0 = \left(p_1^0, \ldots, p_N^0\right)$, где $p_j^0 = \mathrm{P}\left(\xi \in \mathcal{E}_j \mid H_0\right)$, $j = \overline{1, N}$.

В качестве статистики выбирают величину

$$X_n^2 = X_n^2\left(\nu\right) = \sum_{j=1}^{N} \frac{\left(\nu_j - np_j^0\right)^2}{np_j^0} = \sum_{j=1}^{N} \frac{\nu_j^2}{np_j^0} - n, \tag{4}$$

а критическую область задают в виде $\mathcal{T}_{1\alpha} = \left\{t \geq t_\alpha\right\}$. Точное распределение $\mathcal{L}\left(X_n^2 \mid H_0\right)$ неудобно для вычисления критической границы t_α.

Л е м м а 1 [14]. *Если* $0 < p_j^0 < 1$, $j = \overline{1, N}$, *то при* $n \to \infty$

$$\mathcal{L}\left(X_n^2 \mid H_0\right) \to \chi_{N-1}^2.$$

Учитывая лемму 1, на практике с хорошим приближением уже при $n \geq 50$, $v_j \geq 5$ можно использовать распределение χ_{N-1}^2. Действительно, в этом случае

$$\mathrm{P}\left(X_n^2 \in \mathscr{T}_{1\alpha} \mid H_0\right) = \mathrm{P}\left(X_n^2 \ge \chi^2_{1-\alpha,\,N-1} \mid H_0\right) \approx \int\limits_{\chi^2_{1-\alpha,\,N-1}}^{\infty} k_{N-1}(x)\,dx = \alpha.$$

Здесь $k_{N-1}(x)$ – плотность распределения χ^2_{N-1}.

Таким образом, критерий согласия χ^2 формулируется в виде: *пусть заданы уровень значимости* α *и объем выборки* n, *и наблюдавшиеся значения* $h = (h_1, \ldots, h_N)$ *вектора частот* $\nu = (\nu_1, \ldots, \nu_N)$ *удовлетворяют условиям:* $|h| = h_1 + \ldots + h_N \ge 50$, $h_j \ge 5$, $j = \overline{1, N}$, *тогда, если наблюдавшиеся значения* $t = X_n^2(h)$ *статистики* (4) *удовлетворяют неравенству* $t \ge \chi^2_{1-\alpha, N-1}$, *то гипотезу* H_0 *отвергают; в противном случае* H_0 *не противоречит результатам испытаний.*

2.3. Проверка гипотезы однородности. Одной из важных прикладных задач математической статистики является задача проверки однородности статистического материала. Пусть $X = (X_1, X_2, \ldots, X_n)$ – выборка из распределения $\mathscr{L}(\xi)$ с неизвестной функцией распределения $F_1(x)$, а $Y = (Y_1, Y_2, \ldots, Y_m)$ – выборка из распределения $\mathscr{L}(\eta)$ с неизвестной функцией распределения $F_2(x)$. Требуется проверить гипотезу однородности

$$H_0 : F_1(x) \equiv F_2(x), \quad x \in R.$$

Критерий однородности Смирнова применяют в случае непрерывных распределений. Этот критерий основан на статистике

$$\mathrm{D}_{nm} = \sup_{-\infty < x < +\infty} \left| F_{1n}(x) - F_{2m}(x) \right|,$$

где $F_{1n}(x)$ и $F_{2m}(x)$ – эмпирические функции распределения, построенные по выборкам X и Y соответственно.

Имеет место следующее утверждение.

Теорема 1 [14]. *Если теоретическая функция распределения* $F(x)$ *непрерывна, то*

$$\lim_{n,m\to\infty} \mathrm{P}\left(\sqrt{\frac{nm}{n+m}}\ \mathrm{D}_{nm} \le t\right) = K(t),$$

где $K(t)$ *из* (3).

Отсюда следует, что при больших n и m статистика D_{nm} не должна существенно отклоняться от 0. Таким образом, разумно выбирать критическую область в виде $\mathscr{T}_{1\alpha} = \{t \geq t_\alpha(n, m)\}$, где

$$t_\alpha(n,m) = \sqrt{\frac{nm}{n+m}}\,\lambda_\alpha,\; K(\lambda_\alpha) = 1-\alpha.$$

Действительно, в этом случае

$$\mathrm{P}\left(\mathrm{D}_{nm} \in \mathscr{T}_{1\alpha} \mid H_0\right) = \mathrm{P}\left(\sqrt{\frac{nm}{n+m}}\ \mathrm{D}_{nm} \geq \lambda_\alpha \mid H_0\right) \approx 1 - K\left(\lambda_\alpha\right) = \alpha.$$

Сформулируем критерий однородности Смирнова: *если объем выборок достаточно велик, то, вычислив по выборочным данным значение* t *статистики* D_{nm}, *принимают решение отвергнуть гипотезу* H_0 *в том и только в том случае, когда*

$$t \geq \lambda_\alpha \sqrt{\frac{nm}{n+m}}.$$

2.4. Проверка гипотезы независимости. Пусть $\left((X_1, Y_1), \ldots, (X_n, Y_n)\right)$ – выборка из двумерного распределения $\mathscr{L}\left(\xi = (\xi_1, \xi_2)\right)$ с неизвестной функцией распределения $F_\xi(x, y)$, для которой требуется проверить гипотезу

$$H_0:\ F_\xi(x, y) = F_{\xi_1}(x)\, F_{\xi_2}(y).$$

Критерий независимости χ^2. Эту методику применяют для дискретных моделей с конечным числом исходов, поэтому условимся считать, что случайная величина ξ_1 принимает конечное число s некоторых значений $a_1, \ldots, a_s$, а ξ_2 – k значений $b_1, \ldots, b_k$. Если исходная модель имеет другую структуру, то множество значений ξ_1 разбивается на s интервалов $\mathscr{E}_1^{(1)}, \ldots, \mathscr{E}_s^{(1)}$, множество значений ξ_2 на k интервалов $\mathscr{E}_1^{(2)}, \ldots, \mathscr{E}_k^{(2)}$.

Обозначим через ν_{ij} число наблюдений пары (a_i, b_j) (число элементов выборки, принадлежащих прямоугольнику $\mathscr{E}_i^{(1)} \times \mathscr{E}_j^{(2)}$, если данные группируются) так, что

$$\sum_{i=1}^{s}\sum_{j=1}^{k} \nu_{ij} = n.$$

В дальнейшем будет использована методика χ^2, изложенная в гл. 8, § 3, п. 2. В качестве статистики возьмем

$$\hat{X}_n^2 = n\sum_{i=1}^{s}\sum_{j=1}^{k}\frac{\left(\nu_{ij} - \nu_i\nu_j/n\right)^2}{\nu_i\ \nu_j}, \tag{5}$$

где

$$\nu_i = \sum_{j=1}^{k}\nu_{ij},\ \nu_j = \sum_{i=1}^{s}\nu_{ij}.$$

Оказывается, что при $n \to \infty$ $\mathscr{L}\left(\hat{X}_n^2 \mid H_0\right) \to \chi^2_{(s-1)(k-1)}$ (см. [14]).

Поэтому, при достаточно больших n можно использовать следующее правило проверки гипотезы: *гипотезу H_0 отвергают тогда и только тогда, когда вычисленное по фактическим данным значение t статистики* (5) *удовлетворяет неравенству*

$$t \geq \chi^2_{1-\alpha,\,(s-1)(k-1)}.$$

§ 4. Параметрические гипотезы

1. Понятие параметрической гипотезы. Критерии проверки гипотез

Пусть класс допустимых распределений наблюдаемой случайной величины ξ имеет вид $\mathscr{F} = \left\{F\left(x, \vec{\theta}\right),\ \vec{\theta} \in \Theta,\ \Theta \subseteq R^n\right\}$. *Параметрическая гипотеза* задается указанием некоторого подмножества $\Theta_0 \subseteq \Theta$, элементом которого является, по предположению, неизвестная параметрическая точка $\overline{\theta}$. Записывается это так:

$$H_0 : \overline{\theta} \in \Theta_0.$$

Альтернативная гипотеза имеет вид

$$H_1 : \overline{\theta} \in \Theta_1 = \Theta \setminus \Theta_0;$$

точки $\overline{\theta} \in \Theta_1$ называются *альтернативами*. Если множество $\Theta_0\left(\Theta_1\right)$ состоит из одной точки, то гипотезу H_0 (альтернативу H_1) называют *простой*, в противном случае гипотезу (или альтернативу) называют *сложной*.

П р и м е р 1. Пусть $\vec{\theta} = \left(\theta_1, \theta_2\right)$, $\theta_1, \theta_2 \in R$, $\theta_2 > 0$

$$f\left(x, \vec{\theta}\right) = \frac{1}{\theta_2\sqrt{2\pi}}\exp\left\{-\frac{\left(x-\theta_1\right)^2}{2\theta_2^2}\right\}$$

– плотность (θ_1, θ_2^2)-нормального распределения. Гипотеза $H_0 : (\theta_1, \theta_2) = (0, 1)$ является простой, а гипотеза $H_0 : \theta_1 = a$, где a – фиксированная точка из R, – сложной.

П р и м е р 2. Пусть

$$f(x, \theta) = C_n^x \theta^x (1-\theta)^{n-x}$$

– вероятность x «успехов», $x = 0, 1, \ldots, n$, в n независимых испытаниях схемы Бернулли, $0 < \theta < 1$. Примером простой гипотезы служит гипотеза $H_0 : \theta = \frac{1}{2}$, а примером сложной – $H_0 : \theta > \frac{1}{2}$.

Пусть $X = (X_1, \ldots, X_n)$ – выборка из распределения $\mathscr{L}(\xi) \in \mathscr{F}$, о котором сформулирована некоторая гипотеза $H_0 : \overline{\theta} \in \Theta_0$. Требуется выяснить верна или нет гипотеза H_0, т. е. надо построить критерий, который позволял бы для каждой реализации $\vec{x}$ выборки X принять гипотезу H_0 или отклонить ее. Тем самым каждому критерию соответствует разбиение выборочного пространства $\mathscr{X}$ на 2 подмножества $\mathscr{X}_0$ и $\mathscr{X}_1$ $\left(\mathscr{X}_0 \bigcap \mathscr{X}_1 = \varnothing,\ \mathscr{X}_0 \bigcup \mathscr{X}_1 = \mathscr{X}\right)$, где $\mathscr{X}_0$ состоит из точек x, для которых H_0 принимается, а $\mathscr{X}_1$ – из точек, для которых H_0 отвергается. Множество χ_0 называют *областью принятия гипотезы* H_0, а $\mathscr{X}_1$ – *критической областью*. Критерий, определяемый критической областью $\mathscr{X}_1$ часто для краткости называют *критерием* $\mathscr{X}_1$.

В процессе проверки гипотезы H_0 можно прийти к правильному решению или совершить *ошибку первого рода* – отклонить H_0, когда она верна, или *ошибку второго рода* – принять H_0, когда она ложна. Вероятности этих ошибок можно выразить через функцию мощности $W(\theta)$ критерия $\mathscr{X}_1$

$$W(\theta) = W(\mathscr{X}_1, \theta) = \mathrm{P}_\theta(X \in \mathscr{X}_1), \quad \forall \theta \in \Theta,$$

а именно: вероятности ошибок первого и второго рода равны соответственно $W(\theta)$, $\theta \in \Theta_0$ и $1 - W(\theta), \theta \in \Theta_1$.

Желательно провести проверку гипотезы так, чтобы свести к минимуму вероятности обоих типов ошибок. Однако при данном числе испытаний n в общем случае невозможно ни при каком выборе $\mathscr{X}_1$ одновременно обе эти вероятности сделать как угодно малыми. Рациональный способ выбора критической области можно сформулировать следующим образом: *при заданном числе испытаний n устанавливается граница для вероятности ошибки первого рода и при этом выбирается та критическая область $\mathscr{X}_1$, для которой вероятность ошибки второго рода минимальна,* т. е. выбирается число α между 0 и 1 (обычно для α выбирают одно из следующих стандартных значений: 0,005; 0,01; 0,05 и т. д.) и налагается условие

$$W(\theta) \le \alpha, \quad \forall\theta \in \Theta_0; \tag{1}$$

при этом желательно сделать максимальной (за счет выбора $\mathscr{X}_1$)

$$W(\theta), \ \forall\theta \in \Theta_1. \tag{2}$$

Величину α в (1) называют *уровнем значимости*, а тот факт, что критерий $\mathscr{X}_1$ имеет уровень значимости α, часто подчеркивают обозначением $\mathscr{X}_{1\alpha}$.

Пусть $\mathscr{X}_{1\alpha}$ и $\mathscr{X}_{1\alpha}^*$ – два критерия одного и того же уровня значимости α для гипотезы H_0. Если

$$W\left(\mathscr{X}_{1\alpha}^*, \theta\right) \le W\left(\mathscr{X}_{1\alpha}, \theta\right), \ \forall\theta \in \Theta_0; \tag{3}$$

$$W\left(\mathscr{X}_{1\alpha}^*, \theta\right) \ge W\left(\mathscr{X}_{1\alpha}, \theta\right), \ \forall\theta \in \Theta_1; \tag{4}$$

причем строгое неравенство в (4) имеет место хотя бы при одном значении θ, то говорят, что критерий $\mathscr{X}_{1\alpha}^*$ *равномерно мощнее* критерия $\mathscr{X}_{1\alpha}$. Если соотношения (3) и (4) выполняются для любого критерия $\mathscr{X}_{1\alpha}$, то $\mathscr{X}_{1\alpha}^*$ называют *равномерно наиболее мощным критерием*. В случае, когда гипотеза H_1 простая, вместо термина равномерно наиболее мощный критерий используют термин *наиболее мощный критерий.*

Обычно критическая область задается с помощью некоторой статистики $T(X)$ и, как правило, имеет следующий вид:

$$\mathscr{X}_1 = \left\{\vec{x} : T(\vec{x}) \ge c\right\}, \text{ или } \mathscr{X}_1 = \left\{\vec{x} : T(\vec{x}) \le c\right\}, \text{ или } \mathscr{X}_1 = \left\{\vec{x} : \left|T(\vec{x})\right| \ge c\right\};$$

функцию $T(X)$ называют в этом случае *статистикой критерия*.

2. Выбор из двух простых гипотез. Критерий Неймана – Пирсона

Пусть $\Theta = \left\{\theta_0, \theta_1\right\}$, $H_0 : \theta = \theta_0$, $H_1 : \theta = \theta_1$, т. е. допустимыми распределениями случайной величины ξ являются только два: $F_0(x) = F(x, \theta_0)$ и $F_1(x) = F(x, \theta_1)$. Требуется по выборке

$X=(X_1,\dots,X_n)$ из распределения $\mathscr{L}(\xi)$ определить, какое из этих двух распределений истинно.

Предположим, что уровень значимости α выбран, распределения F_0 и F_1 абсолютно непрерывны и соответствующие плотности $f_0(x)$ и $f_1(x)$ удовлетворяют условию $f_j(x)>0,\ j=0,1.$

Рассмотрим *статистику отношения правдоподобия*

$$l(X)=\frac{L(X;\theta_1)}{L(X;\theta_0)}=\frac{\prod_{i=1}^{n} f_1(X_i)}{\prod_{i=1}^{n} f_0(X_i)}$$

и определим функцию $\psi(c)=\mathrm{P}_{\theta_0}\big(l(X)\ge c\big)$. С ростом c эта функция может только убывать; кроме того, $\psi(0)=1$. Далее, так как

$$\mathrm{P}_{\theta_1}\big(l(X)\ge c\big)=\int\limits_{\vec{x}:l(\vec{x})\ge c} L(\vec{x};\theta_1)\,d\vec{x}\ \ge c\int\limits_{\vec{x}:l(\vec{x})\ge c} L(\vec{x};\theta_0)\,d\vec{x}=$$

$$=c\,\mathrm{P}_{\theta_0}\big(l(X)\ge c\big)=c\,\psi(c),$$

то

$$\psi(c)\le\frac{1}{c}$$

и, следовательно, $\psi(c)\to 0$ при $c\to\infty$. Будем предполагать, что существует такое $c=c(\alpha)$, что $\psi(c)=\alpha$.

Теорема 1 (**Нейман – Пирсон**). *При сделанных предположениях существует наиболее мощный критерий проверки гипотезы* H_0. *Этот критерий задается критической областью*

$$\mathscr{X}_{1\alpha}^{*}=\{\vec{x}:\ l(\vec{x})\ge c\},$$

где критическая граница c *определяется из условия* $\psi(c)=\alpha$.

Доказательство. Рассмотрим любой другой критерий $\mathscr{X}_{1\alpha}$ уровня значимости α. Тогда

$$W(\mathscr{X}_{1\alpha},\theta_1)=\int\limits_{\mathscr{X}_{1\alpha}} L(\vec{x};\theta_1)d\vec{x}=\int\limits_{\mathscr{X}_{1\alpha}\mathscr{X}_{1\alpha}^{*}} L(\vec{x};\theta_1)d\vec{x}+\int\limits_{\mathscr{X}_{1\alpha}\overline{\mathscr{X}}_{1\alpha}^{*}} L(\vec{x};\theta_1)d\vec{x},$$

$$W(\mathscr{X}_{1\alpha}^{*},\theta_1)=\int\limits_{\mathscr{X}_{1\alpha}\mathscr{X}_{1\alpha}^{*}} L(\vec{x};\theta_1)d\vec{x}+\int\limits_{\overline{\mathscr{X}}_{1\alpha}\mathscr{X}_{1\alpha}^{*}} L(\vec{x};\theta_1)d\vec{x}.$$

Отсюда

$$W(\mathscr{X}_{1\alpha},\theta_1)=W(\mathscr{X}_{1\alpha}^{*},\theta_1)+\int\limits_{\mathscr{X}_{1\alpha}\overline{\mathscr{X}}_{1\alpha}^{*}} l(\vec{x})\,L(\vec{x};\theta_0)\,d\vec{x}-\int\limits_{\overline{\mathscr{X}}_{1\alpha}\mathscr{X}_{1\alpha}^{*}} l(\vec{x})\,L(\vec{x};\theta_0)\,d\vec{x}.$$

Согласно определению множества $\mathscr{X}^*_{1\alpha}$, если $\vec{x} \notin \mathscr{X}^*_{1\alpha}$, то $l(\vec{x}) < c$, а если $\vec{x} \in \mathscr{X}^*_{1\alpha}$, то $-l(\vec{x}) \le -c$, поэтому из последнего соотношения вытекает неравенство

$$W(\mathscr{X}_{1\alpha}, \theta_1) < W\left(\mathscr{X}^*_{1\alpha}, \theta_1\right) + c\left(\int\limits_{\mathscr{X}_{1\alpha}\overline{\mathscr{X}}^*_{1\alpha}} L\left(\vec{x}, \theta_0\right) d\vec{x} - \int\limits_{\overline{\mathscr{X}}_{1\alpha}\mathscr{X}^*_{1\alpha}} L\left(\vec{x}, \theta_0\right) d\vec{x}\right) =$$
$$= W\left(\mathscr{X}^*_{1\alpha}, \theta_1\right),$$

так как по условию

$$\int\limits_{\mathscr{X}_{1\alpha}} L\left(\vec{x}, \theta_0\right) d\vec{x} = \int\limits_{\mathscr{X}^*_{1\alpha}} L\left(\vec{x}, \theta_0\right) d\vec{x}.$$

Теорема доказана.

Построенный критерий проверки гипотезы H_0 называют *критерием Неймана – Пирсона*.

Соответствующие рассуждения можно провести и для дискретных распределений.

Пример 3. Наиболее мощные критерии для проверки гипотез о параметрах нормального распределения. Пусть $X = \left(X_1, \ldots, X_n\right)$ – выборка из нормального распределения с параметрами $\left(\theta_1, \theta_2^2\right)$, а $\vec{x} = \left(x_1, \ldots, x_n\right)$ – ее реализация. Предположим, что параметр θ_2 известен, $\theta_2 = \sigma$, а относительно параметра θ_1 есть две простые гипотезы – $H_0:\ \theta_1 = a_0$, $H_1:\ \theta_1 = a_1$. Построим оптимальный критерий Неймана – Пирсона. В этом случае

$$L\left(\vec{x}; a_j\right) = \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i - a_j\right)^2\right\},\ j = 0, 1,$$

и

$$l\left(\vec{x}\right) = \frac{L\left(\vec{x}; a_1\right)}{L\left(\vec{x}; a_0\right)} = \exp\left\{n\bar{x}\left(a_1 - a_0\right) - \frac{n}{2\sigma^2}\left(a_1^2 - a_0^2\right)\right\}.$$

Из последнего соотношения следует, что область значений $\vec{x}$, для которых

$$\frac{L\left(\vec{x}; a_1\right)}{L\left(\vec{x}; a_0\right)} > c,$$

определяется неравенством $\bar{x} > C_1$ при некотором C_1. Как известно, выборочное среднее $\bar{X}$ в нашем случае распределено нормально с параметрами

$$\left(\theta_1, \frac{\sigma}{\sqrt{n}}\right).$$

Определим теперь ошибки первого и второго рода:

$$\alpha_1 = \mathrm{P}\left(\overline{X} > C_1 \mid H_0\right) = \frac{1}{\sqrt{2\pi}} \int\limits_{\frac{C_1 - a_0}{\sigma}\sqrt{n}}^{\infty} e^{-\frac{t^2}{2}} dt = 1 - \Phi\left(\frac{C_1 - a_0}{\sigma}\sqrt{n}\right),$$

$$\alpha_2 = \mathrm{P}\left(\overline{X} \le C_1 \mid H_1\right) = \Phi\left(\frac{C_1 - a_1}{\sigma}\sqrt{n}\right),$$

где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Пусть ζ_α – α-квантиль нормального распределения с параметрами $(0, 1)$, тогда

$$\frac{C_1 - a_0}{\sigma}\sqrt{n} = \zeta_{1-\alpha_1}, \quad \frac{C_1 - a_1}{\sigma}\sqrt{n} = \zeta_{\alpha_2},$$

$$C_1 = a_0 + \zeta_{1-\alpha_1}\frac{\sigma}{\sqrt{n}} = a_1 + \zeta_{\alpha_2}\frac{\sigma}{\sqrt{n}}$$

и

$$n = \frac{\sigma^2\left(\zeta_{1-\alpha_1} + \zeta_{\alpha_2}\right)^2}{\left(a_1 - a_0\right)^2}.$$

Последнее равенство дает тот объем выборки, который при наиболее мощном критерии обеспечивает ошибки первого и второго рода α_1 и α_2 (если правая часть в этом равенстве не целая, то за n надо брать ближайшее большее целое число).

Рассмотрим теперь следующие две гипотезы:

$H_0 : \theta_1 = 0,\ \theta_2 = \sigma_0,\ \ H_1 : \theta_1 = 0,\ \theta_2 = \sigma_1 > \sigma_0.$

В этом случае отношение правдоподобия

$$l(\vec{x}) = \frac{L(\vec{x}; \sigma_1)}{L(\vec{x}; \sigma_0)} = \frac{\sigma_0^n}{\sigma_1^n} \exp\left\{\frac{1}{2}\left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)\sum_{j=1}^{n} x_j^2\right\}$$

приводит к критическому множеству

$$\sum_{j=1}^{n} x_j^2 \ge C_1.$$

Поскольку случайная величина

$$\chi_n^2 = \sum_{j=1}^{n} \frac{X_j^2}{\sigma^2}$$

имеет в этом случае χ^2-распределение с n степенями свободы с функцией распределения

$$K_n(x) = \int\limits_{0}^{x} k_n(t)\, dt$$

и плотностью

$$k_n(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{\frac{n}{2}-1} e^{-\frac{x^2}{2}}, \quad x > 0,$$

то ошибки первого и второго рода определяются из равенств

$$\alpha_1 = \mathrm{P}\left(\frac{1}{\sigma_0^2}\sum_{j=1}^{n} x_j^2 > \frac{C_1}{\sigma_0^2}\right) = 1 - K_n\left(\frac{C_1}{\sigma_0^2}\right),$$

$$\alpha_2 = \mathrm{P}\left(\frac{1}{\sigma_1^2}\sum_{j=1}^{n} x_j^2 \le \frac{C_1}{\sigma_1^2}\right) = K_n\left(\frac{C_1}{\sigma_1^2}\right).$$

Пример 4. Наиболее мощный критерий для проверки гипотез о параметре в схеме Бернулли. Пусть $0 < p_0 < p_1 < 1$. Рассмотрим следующие гипотезы:

$$H_0 : f_0(x) = C_n^x\, p_0^x \left(1 - p_0\right)^{n-x},$$

$$H_1 : f_1(x) = C_n^x\, p_1^x \left(1 - p_1\right)^{n-x}, \; x = 0, 1, ..., n.$$

Наиболее мощный критерий для проверки гипотезы H_0 против альтернативной гипотезы H_1 будем строить исходя из неравенства

$$\frac{f_1(x)}{f_0(x)} = \left[\frac{p_1(1-p_0)}{p_0(1-p_1)}\right]^x \frac{(1-p_1)^n}{(1-p_0)^n} \ge C,$$

которое равносильно неравенству $x \le C_1$ при некотором C_1. Для вычисления ошибок первого и второго рода воспользуемся тем, что число «успехов» x в n испытаниях схемы Бернулли асимптотически нормально с параметрами

$$\left(np, \sqrt{np(1-p)}\right).$$

Поэтому

$$\alpha_1 = \mathrm{P}\left(x \ge C_1 \mid H_0\right) = \mathrm{P}\left(\frac{x - np_0}{\sqrt{np_0(1-p_0)}} \ge \frac{C_1 - np_0}{\sqrt{np_0(1-p_0)}}\right),$$

$$\alpha_2 = \mathrm{P}\left(x < C_1 \mid H_1\right) = \mathrm{P}\left(\frac{x - np_1}{\sqrt{np_1(1-p_1)}} \ge \frac{C_1 - np_1}{\sqrt{np_1(1-p_1)}}\right),$$

и, следовательно, при заданных α_1 и α_2

$$C_1 \approx np_0 + \zeta_{1-\alpha_1}\sqrt{np_0(1-p_0)} \approx np_1 + \zeta_{\alpha_2}\sqrt{np_1(1-p_1)},$$

а необходимый объем выборки

$$n = \left[\frac{\left(\zeta_{1-\alpha_1}\sqrt{np_0(1-p_0)} + \zeta_{\alpha_2}\sqrt{np_1(1-p_1)}\right)^2}{\left(p_1 - p_0\right)^2}\right] + 1,$$

где ζ_α – α-квантиль нормального распределения с параметрами $(0, 1)$, а $[a]$ – целая часть числа a.

3. Метод отношения правдоподобия проверки сложных гипотез

Суть *метода отношения правдоподобия* заключается в следующем. Для проверки гипотезы $H_0 : \theta \in \Theta_0 \subset \Theta$ против альтернативы $H_1 : \theta \in \Theta_1 = \Theta \setminus \Theta_0$ вводится *статистика отношения правдоподобия*

$$\Lambda_n = \Lambda_n(X; \Theta_0) = \sup_{\theta \in \Theta_1} L(X; \theta) / \sup_{\theta \in \Theta_0} L(X; \theta),$$

и критическую область задают в виде

$$\mathscr{X}_1 = \left\{ \vec{x} : \Lambda_n(\vec{x}; \Theta_0) \le c \right\} \qquad (5)$$

Критическую границу c в (5) следует выбрать так, чтобы критерий имел заданный уровень значимости α:

$$\mathrm{P}_\theta \left(X \in \mathscr{X}_1 \right) = \int_{\mathscr{X}_1} L(\vec{x}; \theta) d\vec{x} = \mathrm{P}_\theta \left(\Lambda_n \left(X; \Theta_0 \right) \le c \right) \le \alpha, \ \forall \theta \in \Theta_0.$$

§ 5. Линейная регрессия и метод наименьших квадратов

До сих пор рассматривались статистические выводы для моделей, соответствующих повторным независимым наблюдениям над некоторой случайной величиной ξ. В этих случаях исходные статистические данные представляют собой реализацию случайного вектора $X = \left(X_1, \ \ldots, \ X_n \right)$, компоненты которого независимы и одинаково распределены, а именно: $F_{X_i} = F_\xi, \ \ i = \overline{1, n}$. В приложениях математической статистики предположения о независимости и одинаковой распределенности компонент X_i не всегда выполняются. Здесь рассмотрен важный случай таких ситуаций, которые описываются в терминах *линейной регрессионной модели*.

Пусть

$$X = A\theta + \varepsilon, \quad \mathrm{M}\varepsilon = 0, \qquad (1)$$

$$\mathrm{D}(X) = \mathrm{D}(\varepsilon) = \mathrm{M}(\varepsilon\varepsilon') = \sigma^2 I_n, \qquad (2)$$

где $\theta = \left(\theta_1, \ \ldots, \ \theta_s \right)$ – неизвестный параметр, $\varepsilon = \left(\varepsilon_1, \ \ldots, \ \varepsilon_n \right)$ – случайный вектор ошибок наблюдения, $\mathrm{D} X_i = \mathrm{D}\varepsilon_i = \sigma^2 > 0, \ i = \overline{1, n}, \ I_n$ – единичная матрица порядка $n \times n$,, матрица A порядка $n \times s$ считается известной. Если выполняются условия (1) и (2), то говорят, что имеет место *модель линейной регрессии*.

Метод наименьших квадратов состоит в нахождении $\hat{\theta}$ из условия

$$\min_{\theta}\left(\varepsilon\varepsilon^{T}\right)=\min_{\theta}\left|x-A\theta\right|^{2}=\left|x-A\hat{\theta}\right|^{2}, \tag{3}$$

где $|\cdot|$ – норма в R^{n}.

Такая оценка $\hat{\theta}$ называется *оценкой методом наименьших квадратов*.

Чаще всего задачи регрессионного анализа решают в предположении, что наблюдения подчиняются нормальному закону распределения; в этом случае к условиям (1) и (2) добавляют третье условие:

$$\mathscr{L}(\varepsilon)=N\left(\theta,\sigma^{2}I_{n}\right). \tag{4}$$

Если выполнены условия (1) и (4) (условие (2) вытекает из (4)), то говорят о *нормальной регрессии*.

Для нормальной модели оценки наименьших квадратов коэффициентов регрессии θ_1, …, θ_s совпадают с оценками максимального правдоподобия этих параметров, что несложно видеть, учитывая вид функции правдоподобия. В общем случае метод наименьших квадратов не совпадает с методом максимального правдоподобия.

Исследуем некоторые свойства оценок, полученных методом наименьших квадратов, в общем случае.

Пусть $R=(X-A\theta)^{T}(X-A\theta)$, где T – знак транспонирования. Будем искать оценки методом наименьших квадратов, для этого запишем условие

$$\frac{\partial R}{\partial\theta_i}=0,\ i=\overline{1,s}.$$

В результате чего придем к системе уравнений

$$A^{T}\left(X-A\hat{\theta}\right)=0$$

и, следовательно, оценки $\hat{\theta}$ параметра θ имеют вид

$$\hat{\theta}=\left(A^{T}A\right)^{-1}A^{T}X. \tag{5}$$

Докажем, что $\hat{\theta}$ – несмещенная оценка. Действительно, так как $\mathrm{M}\varepsilon=0$, то из (1) следует, что $\mathrm{M}X=A\theta$ и

$$\mathrm{M}\hat{\theta}=\left(A^{T}A\right)^{-1}A^{T}\,\mathrm{M}X=\left(A^{T}A\right)^{-1}A^{T}A\theta=\theta.$$

Найдем матрицу вариаций оценки $\hat{\theta}$. Подставим в (5) формулу (1), тогда

$$\hat{\theta}=\left(A^{T}A\right)^{-1}A^{T}(A\theta+\varepsilon)=\left(A^{T}A\right)^{-1}A^{T}A\theta+\left(A^{T}A\right)^{-1}A^{T}\varepsilon=\theta+\left(A^{T}A\right)^{-1}A^{T}\varepsilon,$$

т. е. $\hat{\theta}-\theta=\left(A^{T}A\right)^{-1}A^{T}\varepsilon$.

По определению матрица вариаций $\hat{\theta}$ равна

$$V = \mathrm{M}\left(\hat{\theta}-\theta\right)\left(\hat{\theta}-\theta\right)^T,$$

поэтому

$$V = \mathrm{M}\left\{\left(A^T A\right)^{-1} A^T \varepsilon \varepsilon^T A\left(A^T A\right)^{-1}\right\} =$$

$$= \left(A^T A\right)^{-1} A^T \mathrm{M}\left(\varepsilon \varepsilon^T\right) A\left(A^T A\right)^{-1} = \sigma^2 \left(A^T A\right)^{-1}.$$

Итак

$$V = \sigma^2 \left(A^T A\right)^{-1}. \tag{6}$$

На самом деле значение σ^2 изначально неизвестно и его приходится оценивать по результатам измерений. Найдем несмещенную оценку σ^2.

Минимальное значение величины R равно

$$R_{\min} = \left(X - A\hat{\theta}\right)^T \left(X - A\hat{\theta}\right),$$

где $\hat{\theta}$ – оценка параметра θ, полученная методом наименьших квадратов. Учитывая (1) и (5) имеем

$$R_{\min} = \varepsilon^T \left(I_n - A\left(A^T A\right)^{-1} A^T\right)^2 \varepsilon.$$

Несложно проверить, что из последнего равенства следует

$$R_{\min} = \varepsilon^T \left(I_n - A\left(A^T A\right)^{-1} A^T\right)^2 \varepsilon.$$

Найдем математическое ожидание от $R_{\min}$, учитывая, что для любой $n \times n$-матрицы $B = \left[b_{ij}\right]$

$$\mathrm{M}\left(\varepsilon^T B \varepsilon\right) = \sigma^2 \sum_{i=1}^{n} b_{ii} = \sigma^2 Sp\left(B\right).$$

Поэтому

$$\mathrm{M}\left(R_{\min}\right) = \sigma^2 \left(Sp\left(I_n\right) - Sp\left(A\left(A^T A\right)^{-1} A^T\right)\right).$$

Но $Sp\left(I_n\right) = n.$ Далее, так как $Sp\left(AB\right) = Sp\left(BA\right)$, то

$$Sp\left(A\left(A^T A\right)^{-1} A^T\right) = Sp\left(\left(A^T A\right)^{-1}\left(A^T A\right)\right) = Sp\left(I_s\right) = s,$$

следовательно,

$$\mathrm{M}\left(R_{\min}\right) = \sigma^2 \left(n - s\right).$$

Сейчас можем провести несмещенную оценку для σ^2:

$$\hat{\sigma}^2 = \frac{R_{\min}}{n-s}$$

и для матрицы вариаций (см. (6))

$$\hat{V} = \left(A^T A\right)^{-1} \frac{R_{\min}}{n-s}.$$

Оптимальные свойства оценок по методу наименьших квадратов утверждает

Теорема 1 [13]. *Оценки по методу наименьших квадратов имеют наименьшую дисперсию в классе линейных несмещенных оценок.*

П р и м е р 1. Пусть теоретическая зависимость y от x имеет вид $y = ax + b$. Если задать x равным $x_1, \ldots, x_n$ и измерять y равным $y_1, \ldots, y_n$ соответственно, то из-за ошибок измерений на самом деле будем иметь $y_i = a + bx_i + \varepsilon_i$. В данном случае неизвестными параметрами будут a и b, т. е.

$$\theta = (a, b)^T, \quad X = (x_1, \ldots, x_n)^T, \quad Y = (y_1, \ldots, y_n)^T$$

и

$$A = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}.$$

Используя формулу (5), несложно получить, что

$$\hat{\theta} = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \frac{1}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \begin{bmatrix} \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i \\ n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i \end{bmatrix},$$

т. е. оценки параметров $\hat{a}$ и $\hat{b}$, полученные методом наименьших квадратов, имеют вид

$$\hat{a} = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}, \quad \hat{b} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}.$$

У п р а ж н е н и е 1. Найдите оценки для дисперсий оценок $\hat{a}$ и $\hat{b}$.

П р и м е р 2. Пусть (ξ, η) – двумерная случайная величина, например ξ и η – сигналы на входе и выходе некоторого устройства. Во многих случаях бывает важно уметь представить одну из величин, например η, как функцию $g(\xi)$ другой. Точное представление $\eta = g(\xi)$, как правило, невозможно, поэтому возникает задача о возможности приближенного представления: $\eta \approx g(\xi)$. Итак, требуется найти среди всех функций $g(\xi)$ случайной величины ξ такую, которая являлась бы *наилучшим приближением случайной величины* η. Термину «наилучшее приближение» следует придать точный смысл, что можно сделать различными способами. Самым удобным и общепринятым является приближение по методу наименьших квадратов. По определению будем считать, что случайная величина $g(\xi)$ является наилучшим приближением случайной величины η в смысле метода наименьших квадратов, если $\mathrm{M}(\eta - g(\xi))^2$ принимает наименьшее возможное значение, при этом величина $g(\xi)$ называется *средней квадратической регрессией величины* η *на величину* ξ.

Ограничимся нахождением линейной средней квадратической регрессии величины η на величину ξ, т. е. случаем, когда $g(\xi) = a + b\xi$. Пусть $m_1 = \mathrm{M}\xi$, $\sigma_1^2 = \mathrm{D}\xi$, $m_2 = \mathrm{M}\eta$, $\sigma_2^2 = \mathrm{D}\eta$, $\rho(\xi, \eta) = \rho$, $\Phi(a, b) = \mathrm{M}(\eta - a - b\xi)^2$.

Найдем такие a и b при которых функция $\Phi(a,b)$ принимает наименьшее из возможных значений. Имеем

$$\frac{\partial \Phi}{\partial a} = -2(m_2 - a - bm_1) = 0,$$

$$\frac{\partial \Phi}{\partial b} = 2\left[b(\sigma_1^2 + m_1^2) - (\rho\sigma_1\sigma_2 + m_2 m_1) + am_1\right] = 0,$$

откуда $b = \rho(\sigma_2/\sigma_1)$, $a = m_2 - \rho(\sigma_2/\sigma_1)m_1$. В то же время

$$\begin{vmatrix} \dfrac{\partial^2 \Phi}{\partial a^2} & \dfrac{\partial^2 \Phi}{\partial a \partial b} \\ \dfrac{\partial^2 \Phi}{\partial a \partial b} & \dfrac{\partial^2 \Phi}{\partial b^2} \end{vmatrix} = \begin{vmatrix} 2 & 2m_1 \\ 2m_1 & 2(\sigma_1^2 + m_1^2) \end{vmatrix} = 4\sigma_1^2 > 0.$$

Таким образом

$$g(\xi) = m_2 - \rho(\sigma_2/\sigma_1)(\xi - m_1)$$

– *линейная регрессия величины* η *на величину* ξ. Отметим, что прямая, определяемая уравнением

$$(y - m_2)/\sigma_2 = \rho(x - m_1)/\sigma_1$$

называется *прямой линейной регрессии величины* η *на величину* ξ.

Аналогично можно найти *линейную регрессию величины* ξ *на величину* η:

$$h(\eta) = m_1 - \rho(\sigma_1/\sigma_2)(\eta - m_2),$$

и *прямую линейной регрессии величины* ξ *на величину* η:

$$(y - m_2)/\sigma_2 = (1/\rho)(x - m_1)/\sigma_1.$$

Отметим, что оценки коэффициентов уравнений линейной регрессии найдены в прим. 1.

Задачи

1. Из конечной генеральной совокупности $(X_1, \ldots, X_N)$ берутся последовательно две бесповторные выборки $(x_1, \ldots, x_{n_1})$ и $(y_1, \ldots, y_{n_2})$, $n_1 + n_2 \le N$. Найдите ковариацию и коэффициент корреляции между выборочными средними $\bar{x}$ и $\bar{y}$.

2. Найдите математическое ожидание выборочной дисперсии для бесповторной выборки объема n из конечной генеральной совокупности $(X_1, \ldots, X_N)$.

3. Вычислите математическое ожидание и дисперсию k-ой порядковой статистики для выборки объема n из равномерного на $[0; a]$ распределения.

4. Имеется выборка $(X_1, \ldots, X_n)$. По гипотезе H_0 все X_i равномерно распределены на $[0; 2]$, по гипотезе H_1 – на $[1; 3]$. Постройте критерий с наименьшей величиной $\max(\alpha, \beta)$, где α и β – вероятности ошибок 1-го и 2-го рода.

5. Случайная величина ξ подчиняется биномиальному закону распределения. Найдите несмещенные оценки математического ожидания и дисперсии этого распределения, считая параметр p известным.

6. По независимым выборкам $(X_1, \ldots, X_{n_1})$ и $(Y_1, \ldots, Y_{n_2})$ из двух нормальных распределений с параметрами (a_1, σ_1) и (a_2, σ_2) соответственно постройте доверительный интервал с доверительной вероятностью α для разности $a_1 - a_2$, если σ_1 и σ_2 известны.

7. Пусть x есть число «успехов» в n испытаниях схемы Бернулли с вероятностью «успеха» θ в каждом испытании. Покажите, что статистика

$$T(x) = \sum_{k=0}^{n} a_k \frac{x^{(k)}}{n^{(k)}}$$

является эффективной оценкой для параметрической функции $\tau_1(\theta) = a_0 + + a_1\theta + a_2\theta^2 + \ldots + a_n\theta^n$. Здесь $x^{(k)} = x(x-1)\ldots(x-k+1)$. Докажите, что не существует несмещенной оценки для $\tau_2(\theta) = \theta^N$ $(N > n)$.

8. Пусть $X = (X_1, \ldots, X_n)$ – выборка из распределения Пуассона с неизвестным параметром θ. Покажите, что статистика

$$T(X_1, \ldots, X_n) = \sum_{k=0}^{x} a_k \frac{x^{(k)}}{n^k},$$

где $x = X_1 + \ldots + X_n$, $x^{(k)} = x(x-1)\ldots(x-k+1)$ является эффективной оценкой для $\tau_1(\theta) = a_0 + a_1\theta + a_2\theta^2 + \ldots$. Докажите, что для $\tau_2(\theta) = 1/\theta$ не существует несмещенной оценки.

9. Пусть $X = (X_1, \ldots, X_n)$ – выборка из равномерного на $[0; \theta]$ распределения, $\hat{\theta}_n$ – оценка максимального правдоподобия параметра θ. Докажите, что $\hat{\theta}_n$ – состоятельная оценка и

$$M_\theta \hat{\theta}_n = \frac{n}{n+1}\theta, \ D_\theta \hat{\theta}_n = \frac{n}{(n+1)^2(n+2)}\theta^2, \quad \theta > 0.$$

10. Пусть $X = (X_1, \ldots, X_n)$ – выборка из нормального с параметрами $(\theta, 2\theta)$, $\theta > 0.$ Покажите, что статистика

$$T(X) = \sqrt{1 + \frac{1}{n}\sum_{i=1}^{n} X_i^2} - 1$$

является состоятельной оценкой неизвестного параметра θ.

11. Пусть $X = (X_1, \ldots, X_n)$ – выборка из равномерного на $[0;\theta]$ распределения и $T(X) = \max(X_1, \ldots, X_n).$ Докажите, что интервал $(T(X); T(X)\varepsilon^{-1/n})$ есть доверительный интервал параметра θ уровня ε $(0 < \varepsilon < 1)$.

12. Предположим, что наблюдения $x_1, \ldots, x_n$ представляются в виде

$$x_k = \theta_0 + \theta_1 t_k + \varepsilon_k, \ k = \overline{1, n},$$

где $t_k = k,$ $\theta = (\theta_0, \theta_1)$ – неизвестный параметр; случайные ошибки наблюдения $\varepsilon_1, \ldots, \varepsilon_n$ независимы и нормально распределены с параметрами $(0, \sigma^2)$. Пусть $\hat{\theta} = (\hat{\theta}_0, \hat{\theta}_1)$ – оценка параметра θ, полученная методом наименьших квадратов. Докажите, что

$$D\hat{\theta}_0 = \frac{4}{n}\left[1 + \frac{3}{2(n-1)}\right]\sigma^2, \ D\hat{\theta}_1 = \frac{12}{n(n^2-1)}\sigma^2.$$

ПРИЛОЖЕНИЯ

Таблица 1. **Основные вероятностные распределения**

Распределение	Обозначение	Параметры и область их изменения	Плотность распределения (дискрет-ное распределение вероятностей)	Характеристическая функция	Математическое ожидание	Мода	Дисперсия
Бернулли	$Bi(1, p)$	$0 \le p \le 1$	$p^x (1-p)^{1-x},\ x \in \{0, 1\}$	$1 + p\left(e^{it} - 1\right)$	p	$U(p - 1/2)$	$p(1-p)$
Биномиальное	$Bi(N, p)$	N – натуральное число, $0 \le p \le 1$	$C_N^x p^x (1-p)^{N-x},\ x \in \{0, 1, \ldots, N\}$	$\left(1 + p\left(e^{it} - 1\right)\right)^N$	Np	$(N+1)p - 0{,}5 + \varepsilon,\ 0 \le \lvert\varepsilon\rvert \le 0{,}5$	$Np(1-p)$
Отрицательное биномиальное	$\overline{Bi}(r, p)$	r – натуральное число, $0 \le p \le 1$	$C_{x+r-1}^x p^r (1-p)^x,\ x \in \{0, 1, \ldots\}$	$\left(p / \left(1 - (1-p)e^{it}\right)\right)^r$	$r(1-p)/p$	λ и $\lambda - 1$, если λ – целое, иначе $[\lambda]$; $\lambda = (r-1) \times$ $\times (1-p)/p$	$r(1-p)/p^2$
Пуассона	$\Pi(\lambda)$	$\lambda > 0$	$\lambda^x e^{-\lambda}/x!,\ x \in \{0, 1, \ldots\}$	$\exp\left(\lambda\left(e^{it} - 1\right)\right)$	λ	λ и $\lambda - 1$, если λ – целое, иначе $[\lambda]$	λ
Геометрическое	$G(p)$	$0 \le p \le 1$	$p(1-p)^x,\ x \in \{0, 1, \ldots\}$	$p / \left(1 - (1-p)e^{it}\right)$	$(1-p)/p$	0	$(1-p)/p^2$
Гипергеометрическое	$H(N, L, n)$	N, L, n – натуральные числа, $L \le n$	$C_L^x C_{N-L}^{n-x} / C_N^n,\ x \in \{0, 1, \ldots, L\}$	$\sum_{k=0}^{L} e^{itk} C_L^k C_{N-L}^{n-k} / C_N^n$	nL/N	–	$\dfrac{nL(N-L)(N-n)}{N^2(N-1)}$
Одномерное нормальное (гауссовское)	$N(\mu, \sigma^2)$	$\mu \in R,\ \sigma > 0$	$\dfrac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\dfrac{(x-\mu)^2}{2\sigma^2}\right\},\ x \in R$	$\exp\left(it\mu - \dfrac{1}{2}t^2\sigma^2\right)$	μ	μ	σ^2
Логнормальное	$L(\mu, \sigma^2)$	$\mu \in R,\ \sigma > 0$	$\dfrac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\dfrac{\left(\ln(x/\mu)\right)^2}{2\sigma^2}\right)\ x \ge 0$	$\sum_{k=0}^{\infty} \dfrac{(it)^k}{k!} \exp\left(k \ln\mu + \dfrac{k^2\sigma^2}{2}\right)$	$\mu\sqrt{\omega}$, где $\omega = e^{\sigma^2}$	μ/ω	$\mu^2\omega(\omega - 1)$
Многомерное нормальное (гауссовское)	$N_N(\mu, \Sigma)$	$\mu = (\mu_i) \in R^N$, $\Sigma = (\sigma_{ij}) = \Sigma^T$	$(2\pi)^{-\frac{1}{2}N} \lvert\Sigma\rvert^{-\frac{1}{2}} \times$ $\times \exp\left\{\dfrac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right\}$ $x \in R^N$	$\exp\left(it^T\mu - t^T\Sigma t/2\right)$	μ	μ	ковариационная матрица Σ
Равномерное	$R(a, b)$	$a, b \in R, a < b$	$\dfrac{1}{b-a},\ a \le x \le b$	$\dfrac{e^{itb} - e^{ita}}{i(b-a)t}$	$\dfrac{a+b}{2}$	–	$\dfrac{(b-a)^2}{12}$

Продолжение таблицы 1

Распределение	Обозна-чение	Параметры и область их изменения	Плотность распределения (дискрет-ное распределение вероятностей)	Характеристическая функ-ция	Математи-ческое ожидание	Мода	Дисперсия
Гамма-распределение	$\gamma(\alpha,\lambda)$	$\alpha,\ \lambda>0$	$\lambda^{\alpha}x^{\alpha-1}\exp(-\lambda x)/\Gamma(\alpha),\ x\ge 0$	$(1-it/\lambda)^{-\alpha}$	α/λ	$(\alpha-1)/\lambda$, если $\lambda\ge 1$	α/λ^2
Экспоненци-альное	$\mathcal{E}(\lambda)$	$\lambda>0$	$\lambda e^{-\lambda x},\ \ x>0$	$(1-it/\lambda)^{-1}$	λ^{-1}	0	λ^{-2}
Коши	$K(a,b)$	$a\in R,\ \ b>0$	$1/\left(\pi b\left(1+\left((x-a)/b\right)^2\right)\right)\ \ x\in R$	$\exp\left(ita-b\lvert t\rvert\right)$	не сущест-вует	a	не существует
Лапласа	$EE(a,\lambda)$	$a\in R,\ \lambda>0$	$\dfrac{\lambda}{2}\exp\left(-\lambda\lvert x-a\rvert\right),\ \ x\in R$	$\lambda^2\left(t^2+\lambda^2\right)^{-1}e^{ita}$	a	a	$2\lambda^{-2}$
χ^2-рас-пределение	χ^2_{υ}	υ – целое по-ложительное число	$\dfrac{x^{\upsilon/2-1}\exp(-x/2)}{2^{\upsilon/2}\Gamma(\upsilon/2)},\ \ x\ge 0$	$(1-2it)^{-\upsilon/2}$	υ	$\upsilon-2$, если $\upsilon\ge 2$	2υ
t-распределение Стьюдента	t_{υ}	υ – целое по-ложительное число	$\dfrac{\Gamma\left((\upsilon+1)/2\right)}{\sqrt{\pi\upsilon}\,\Gamma(\upsilon/2)}\left(1+\dfrac{x^2}{\upsilon}\right)^{-\frac{\upsilon+1}{2}}\ \ x\in R$	$\dfrac{\sqrt{\pi}\,\Gamma\left((\upsilon+1)/2\right)\exp\left(-\sqrt{\upsilon}\lvert t\rvert\right)}{\Gamma(\upsilon/2)\,2^{2(n-1)}(n-1)!}\times$ $\times\sum_{k=0}^{n-1}(2k)!\,C_{n-1+k}^{2k}\times$ $\times\left(2\sqrt{\upsilon}\lvert t\rvert\right)^{n-1-k}$, если $n=(\upsilon+1)/2$ – целое	0	0	$\upsilon/(\upsilon-1)$, если $\upsilon>2$
F- распределе-ние Фишера	F_{mn}	m,n – целые положитель-ные числа	$m^{\frac{m}{2}}n^{\frac{n}{2}}\dfrac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\dfrac{x^{\frac{m}{2}-1}}{(n+mx)^{\frac{m+n}{2}}}$ $x\ge 0$	$\left(\dfrac{n}{m}\right)^{it/2}\left(\Gamma\left(\dfrac{m}{2}\right)\Gamma\left(\dfrac{n}{2}\right)\right)^{-1}\times$ $\times\Gamma\left(\dfrac{m+it}{2}\right)\Gamma\left(\dfrac{n-it}{2}\right)$	$\dfrac{n}{n-2}$, если $n>2$	$\dfrac{n}{n+2}\left(1-\dfrac{2}{m}\right)$, если $m>2$	$\dfrac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$, если $n>4$
Логистическое	$l(a,\sigma)$	$a\in R,\ \ \sigma>0$, $k=\sqrt{3}\sigma/\pi$	$\dfrac{\exp\left((x-a)/k\right)}{k\left(1+\exp\left((x-a)/k\right)\right)^2},\quad x\in R$	$e^{ita}\Gamma(1-itk)\Gamma(1+itk)$	a	a	σ^2
Вейбулла – Гне-денко	$W(\alpha,\lambda)$	$\alpha,\ \lambda>0$	$\alpha\lambda x^{\alpha-1}\exp\left(-\lambda x^{\alpha}\right),\ \ x\ge 0$	$\sum_{k=0}^{\infty}\dfrac{(it)^k\,\Gamma(1+k/\alpha)}{k!\,\lambda^{k/\alpha}}$	$\lambda^{-1/\alpha}\times$ $\times\Gamma\left(1+\alpha^{-1}\right)$	0, если $\alpha\le 1$; $\left(\dfrac{1-1/\alpha}{\lambda}\right)^{1/\alpha}$	$\lambda^{-2/\lambda}\left(2\alpha^{-1}\Gamma\left(2\alpha^{-1}\right)-\right.$ $\left.-\alpha^{-2}\Gamma^2\left(\alpha^{-1}\right)\right)$
Бета-распределение	$\beta(\upsilon,\omega)$	$\upsilon,\ \ \omega>0$	$\dfrac{\Gamma(\upsilon+\omega)}{\Gamma(\upsilon)\Gamma(\omega)}x^{\upsilon-1}(1-x)^{\omega-1},\ \ 0\le x\le 1$	$\dfrac{\Gamma(\upsilon+\omega)}{\Gamma(\upsilon)}\sum_{k=0}^{\infty}\dfrac{\Gamma(\upsilon+k)}{\Gamma(\upsilon+\omega+k)}\dfrac{it^k}{k!}$	$\dfrac{\upsilon}{\upsilon+\omega}$	$\dfrac{\upsilon-1}{\upsilon+\omega-2}$, если $\upsilon,\ \ \omega>1$	$\dfrac{\upsilon\omega}{(\upsilon+\omega)^2(\upsilon+\omega+1)}$

Таблица 2. **Значения интеграла Лапласа** $\Phi_0(x)=\frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{t^2}{2}}dt$

x	Сотые доли										*x*
	0	1	2	3	4	5	6	7	8	9	
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0200	0,0239	0,0279	0,0319	0,0359	0,0
0,1	398	438	478	517	557	596	636	675	714	753	0,1
0,2	793	832	871	910	948	987	0,1026	0,1064	0,1103	0,1141	0,2
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	406	443	480	517	0,3
0,4	554	591	628	664	700	736	772	808	844	879	0,4
0,5	915	950	985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224	0,5
0,6	0,2257	0,2291	0,2324	357	389	422	454	486	517	549	0,6
0,7	580	611	642	673	703	734	764	794	823	852	0,7
0,8	881	910	939	967	995	0,3023	0,3051	0,3078	0,3106	0,3133	0,8
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	289	315	340	365	389	0,9
1,0	413	437	461	485	508	531	554	577	599	621	1,0
1,1	643	665	686	708	729	749	770	790	810	830	1,1
1,2	849	869	888	907	925	944	962	980	997	0,4015	1,2
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177	1,3
1,4	192	207	222	236	251	265	279	292	306	319	1,4
1,5	332	345	357	370	382	394	406	418	429	441	1,5
1,6	452	463	474	484	495	505	515	525	535	545	1,6
1,7	554	564	573	582	591	599	608	616	625	633	1,7
1,8	641	649	656	664	671	678	686	693	699	706	1,8
1,9	713	719	726	732	738	744	750	756	761	767	1,9
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817	2,0
2,1	821	826	830	834	838	842	846	850	854	857	2,1
2,2	861	864	868	871	875	878	881	884	887	890	2,2
2,3	893	896	898	901	904	906	909	911	913	916	2,3
2,4	918	920	922	925	927	929	931	932	934	936	2,4
2,5	938	940	941	943	945	946	948	949	951	952	2,5
2,6	953	955	956	957	959	960	961	962	963	964	2,6
2,7	965	966	967	968	969	970	971	972	973	974	2,7
2,8	974	975	976	977	977	978	979	979	980	981	2,8
2,9	981	982	982	983	984	984	985	985	985	986	2,9
3,0	987	987	987	988	988	989	989	989	990	990	3,0

Таблица 3. **Значения функции** u_α

Функция u_α определяется равенством $\alpha=\frac{1}{\sqrt{2\pi}}\int\limits_{u_\alpha}^{\infty} e^{-\frac{t^2}{2}}dt$

α	0.001	0.005	0.010	0.015	0.020	0.025	0.030	0.035	0.040	0.045	0.050
u_α	3,0902	2,5758	2,3263	2,1701	2,0537	1,9600	1,8808	1,8119	1,7507	1,6954	1,6449

Таблица 4. **Значения функции** $p_k(\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$

k \ λ	0,1	0,2	0,3	0,4	0,5
0	0.90484	0.81873	0.74082	0.67032	0.60653
1	0,09048	0,16375	0,22225	0,26813	0,30327
2	0,00452	0,01638	0,03334	0,05363	0,07582
3	0,00015	0,00109	0,00333	0,00715	0,01264
4		0,00006	0,00025	0,00072	0,00158
5			0,00002	0,00006	0,00016
6					0,00001

k \ λ	0,6	0,7	0,8	0,9
0	0.54881	0.49659	0.44933	0.40657
1	0,32929	0,34761	0,35946	0,36591
2	0,09879	0,12166	0,14379	0,16466
3	0,01976	0,02839	0,03834	0,04940
4	0,00296	0,00497	0,00767	0,01112
5	0,00036	0,00070	0,00123	0,00200
6	0,00004	0,00008	0,00016	0,00030
7		0,00001	0,00002	0,00004

k \ λ	1,0	2,0	3,0	4,0	5,0
0	0.36788	0.13534	0.04979	0.01832	0.00674
1	0,36788	0,27067	0,14936	0,07326	0,03369
2	0,18394	0,27067	0,22404	0,14653	0,08422
3	0,06131	0,18045	0,22404	0,19537	0,14037
4	0,01533	0,09022	0,16803	0,19537	0,17547
5	0,00307	0,03609	0,10082	0,15629	0,17547
6	0,00051	0,01203	0,05041	0,10419	0,14622
7	0,00007	0,00344	0,02160	0,05954	0,10445
8	0,00001	0,00086	0,00810	0,02977	0,06528
9		0,00019	0,00270	0,01323	0,03627
10		0,00004	0,00081	0,00529	0,01813
11		0,00001	0,00022	0,00193	0,00824
12			0,00006	0,00064	0,00343
13			0,00001	0,00020	0,00132
14				0,00006	0,00047
15				0,00002	0,00016
16					0,00005
17					0,00001

Т а б л и ц а 5. **Значения функции** $t_{\alpha,n}$

Функция $t_{\alpha,n}$ определяется равенством

$$P\left(\tau_n > t_{\alpha,n}\right) = \alpha,$$

где случайная величина τ_n имеет распределение Стьюдента с n степенями свободы. Плотность распределения τ_n равна

$$p_{\tau_n}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}.$$

n \ 2α	0,10	0,05	0,02	0,01
5	2,015	2,571	3,365	4,032
6	1,943	2,447	3,143	3,707
7	1,895	2,365	2,998	3,499
8	1,860	2,306	2,896	3,355
9	1,833	2,262	2,821	3,250
10	1,812	2,228	2,764	3,169
12	1,782	2,179	2,681	3,055
14	1,761	2,145	2,624	2,977
16	1,746	2,120	2,583	2,921
18	1,734	2,101	2,552	2,878
20	1,725	2,086	2,528	2,845
22	1,717	2,074	2,508	2,819
30	1,697	2,042	2,457	2,750
∞	1,645	1,960	2,326	2,576

Т а б л и ц а 6. **Значения функции** $\chi^2_{\alpha,m}$

Функция $\chi^2_{\alpha,m}$ определяется равенством

$$P\left(\chi^2_m > \chi^2_{\alpha,m}\right) = \alpha,$$

где случайная величина χ^2_m имеет χ^2- распределение с m степенями свободы.

Плотность распределения χ^2_m равна

$$p_{\chi^2_m}(x) = \frac{1}{\Gamma\left(\frac{m}{2}\right)2^{\frac{m}{2}}} x^{\frac{m}{2}-1} e^{-\frac{x}{2}},\ x > 0$$

m \ α	0,10	0,05	0.02	0,01	0,005
1	2,7	3,8	5,4	6,6	7,9
2	4,6	6,0	7,8	9,2	11,6
3	6,3	7,8	9,8	11,3	12,8
4	7,8	9,5	11,7	13,3	14,9
5	9,2	11,1	13,4	15,1	16,3
6	10,6	12,6	15,0	16,8	18,6
7	12,0	14,1	16,6	18,5	20,3
8	13,4	15,5	18,2	20,1	21,9
9	14,7	16,9	19,7	21,7	23,6
10	16,0	18,3	21,2	23,2	25,2
11	17,3	19,7	22,6	24,7	26,8
12	18,5	21,0	24,1	26,2	28,3
13	19.8	22,4	25,5	27,7	29,8
14	21,1	23,7	26,9	29,1	31
15	22,3	25,0	28,3	30,6	32,5
16	23,5	26,3	29,6	32,0	34
17	24,8	27,6	31,0	33,4	35,5
18	26,0	28,9	32,3	34,8	37
19	27,2	30,1	33,7	36,2	38,5
20	28,4	31,4	35,0	37,6	40
21	29,6	32,7	36,3	38,9	41,5
22	30,8	33,9	37,7	40,3	42,5
23	32,0	35,2	39,0	41,6	44,0
24	33,2	36,4	40,3	43,0	45,5
25	34,4	37,7	41,6	44,3	47

ЛИТЕРАТУРА

1. Антоневич А.Б., Радыно Я.В. Функциональный анализ и интегральные уравнения. – Мн.: БГУ, 2002.
2. Бартлет М.С. Введение в теорию случайных процессов. – М.: ИЛ, 1968.
3. Бернштейн С.Н. Теория вероятностей. – М.: Гостехиздат, 1948.
4. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1965.
5. Боровков А.А. Теория вероятностей. – М.: Наука, 1986.
6. Бочаров П.П., Печинкин А.В. Теория вероятностей. Математическая статистика. – М.: Гардарика, 1998.
7. Вентцель А.Д. Курс теории случайных процессов. – М.: Наука, 1978.
8. Гельфанд И.М., Виленкин Н.Я. Некоторые применения гармонического анализа. Оснащенные гильбертовы пространства //Обобщенные функции, вып.IV – М.: Физматгиз, 1961.
9. Гихман И.И., Скороход А.В. Введение в теорию случайных процессов. – М.: Наука, 1977.
10. Гихман И.И., Скороход А.В., Ядренко М.И. Теория вероятностей и математическая статистика. – Киев: Вища школа, 1979.
11. Гнеденко Б.В. Курс теории вероятностей. – М.: Наука, 1988.
12. Дуб Дж.Л. Вероятностные процессы. – М.: ИЛ, 1956.
13. Зуеў М.М., Сячко Ул.Ул. Тэорыя імавернасцей і матэматычная статыстыка. – Мазыр: Белы вецер, 2000.
14. Ивченко Г.И., Медведев Ю.И. Математическая статистика. – М.: Высшая школа, 1984.
15. Ито К., Маккин Г. Диффузионные процессы и их траектории. – М.: Мир, 1968.
16. Климов Г.П. Теория вероятностей и математическая статистика. – М.: МГУ, 1983.
17. Коваленко И.Н., Филиппова А.А. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1982.
18. Козлов М.В., Прохоров А.В. Введение в математическую статистику. – М.: МГУ, 1987.
19. Колмогоров А.Н. Основные понятия теории вероятностей. – М.: Наука, 1974.
20. Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. – М.: Наука, 1972.
21. Крамер Г. Математические методы статистики. – М.: Мир, 1976.
22. Крамер Г., Лидбеттер М. Стационарные случайные процессы, свойства выборочных функций и их приложения. – М.: Мир, 1969.
23. Круглов В.М. Дополнительные главы теории вероятностей. – М.: Высшая школа, 1984.
24. Ламперти Дж. Вероятность. – М.: Наука, 1973.
25. Леви П. Стохастические процессы и броуновское движение. – М.: Мир, 1972.

26. Леман Э. Проверка статистических гипотез. – М.: Наука, 1964.
27. Липцер Р.Ш., Ширяев А.Н. Статистика случайных процессов (нелинейная фильтрация и смежные вопросы). – М.: Наука, 1974.
28. Лоэв М. Теория вероятностей. – М.: ИЛ, 1962.
29. Невё Ж. Математические основы теории вероятностей: Пер. с фр. – М.: Мир, 1969.
30. Партасарати К. Введение в теорию вероятностей и теорию меры. – М.: Мир, 1983.
31. Прохоров Ю.В., Розанов Ю.А. Теория вероятностей. – М.: Наука, 1973.
32. Розанов Ю.А. Случайные процессы. Краткий курс. – М.: Наука, 1979.
33. Розанов Ю.А. Теория вероятностей, случайные процессы, математическая статистика. – М.: Наука, 1985.
34. Севастьянов Б.А. Курс теории вероятностей и математической статистики. – М.: Наука, 1982.
35. Скороход А.В. Случайные процессы с независимыми приращениями. – М.: Наука, 1986.
36. Уиттл П. Вероятность: Пер. с англ. – М.: Наука, 1982.
37. Феллер В. Введение в теорию вероятностей и ее приложения. Т.1,2. – М.: Мир, 1984.
38. Хеннекен П.А., Тортра А. Теория вероятностей и некоторые её приложения. – М.: Наука, 1974.
39. Хида Т. Броуновское движение. – М.: Наука, 1987.
40. Чистяков В.П. Курс теории вероятностей. – М.: Наука, 1987.
41. Ширяев А.Н. Вероятность. – М.: Наука, 1980.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Аксиоматика Колмогорова 15
Алгебра событий 10
Альтернатива 282, 287

Белый шум 241
Блуждание с отражением 56, 58
– – поглощением 56
Борелевская сигма-алгебра 18, 23, 34
– функция 78
Борелевское множество 34
Будущее 55

Вариационный ряд 254
Вектор случайный 80
Вероятностное пространство 15
– – геометрическое 23
– – дискретное 22
– – классическое 20
– – конечное 21
Вероятность 6, 15
– апостериорная 43
– априорная 42
– биномиальная 49
– геометрическая 23
– доверительная 275
– классическая 21
– полиномиальная 50
– условная 40
– частотная (статистическая) 33
Взаимная ковариация 218
Выборка 252
– без возвращения 31
– с возвращением 30
Выборочная функция 199
Выборочное вероятностное пространство 199
Выборочное пространство 252
– среднее 257
Выборочный коэффициент асимметрии 258
– – эксцесса 258
– момент *k*-го порядка 257
– центральный момент *k*-го порядка 257

Генеральная совокупность 253
Гипотеза альтернативная 282
– нулевая 281
– основная 281
– параметрическая 287
– простая 281
– сложная 281
– статистическая 281
Гистограмма 256
Границы доверительные 275

Дисперсия 104
– выборочная 257
Доверительная вероятность 275
Доверительный интервал 274
Достаточная статистика 268

Задача Банаха 37
– Бюффона 38
– о баллотировке 37
– о двух нормальных выборках 279
– о пьянице 38
– о сравнении двух дисперсий 280
– семейная 38
Закон больших чисел 173
– распределения 68
– – устойчивый 160
– редких событий 70
– «0 или 1» Колмогорова 191

Индикатор 63, 93
Интеграл Лебега 116
–Лебега – Стилтьеса 118
– Римана – Стилтьеса 118
Интегральная теорема Муавра – Лапласа 54
– – Пуассона 52
Интервальное оценивание 274
Испытание 46

Квантиль 259
– выборочная 259
Ковариационная матрица 155

Ковариация обобщенного случайного процесса 237
– случайного процесса 216
– случайных величин 105
Количество информации 263
Комбинаторика 26
Конечномерные распределения случайного процесса 200
Корреляционная функция случайного процесса 217
Коэффициент асимметрии 257
– диффузии 231
– корреляции 112
– сноса 231
– эксцесса 257
Критерий независимости
хи - квадрат 286
– Неймана – Пирсона 291
– однородности Смирнова 285
– согласия 283
– – Колмогорова 283
– – хи - квадрат Пирсона 284
– факторизации 269
– эффективности 264
Критическая область 282, 288
Кумулянта 206

Лемма Бореля – Кантелли 189
Линейная регрессионная модель 294
Локальная теорема Муавра – Лапласа 52
– – Пуассона 51

Математическое ожидание 95
Матрица переходных вероятностей 55
Медиана 260
– выборочная 260
Метод максимального правдоподобия 271
– моментов 272
– Монте-Карло188
– наименьших квадратов 294
– отношения правдоподобия 294
Множество упорядоченное 26
– цилидрическое 199
Модель Бозе – Эйнштейна 32
– линейной регрессии 294
– Максвелла – Больцмана 32
– Ферми – Дирака 33
Моменты 104
Моменты абсолютные 104
– – центральные 104
– смешанные154
– центральные 104
Мощность критерия 282

Наиболее мощный критерий 286
Настоящее 55
Независимость классов событий 46
– сигма-алгебр 46
– случайных величин 87
– случайных событий 44
– – – в совокупности 45
– – – попарная 45
некоррелированные случайные величины 113
Неравенство Гаека – Реньи 180
– Гёльдера 111
– Иенсена 110
– Колмогорова 181, 191
– Коши – Буняковского 112
– Ляпунова 111
– Минковского 112
– Рао – Крамера 263
– Чебышева 109
Нормальные числа 188

Обобщенная функция обычного типа 240
Обобщенный случайный процесс 236
– – – Пуассона 211
– – – слабо стационарный 247
– – – умеренный 246
Объединение событий 10
Объем выборки 252
– n-мерный 23
Отклонение среднеквадратическое 104
Отсутствие последействия 73
Оценка 260, 261
– асимптотически несмещенная 261, 272
– асимптотически эффективная 272
– максимального правдоподобия 271
– метода моментов 272
– – наименьших квадратов 295
– несмещённая 261
– оптимальная 262
– сильно состоятельная 264
– состоятельная 264, 271
– эффективная 264

Ошибка второго рода 288
– первого рода 288
– систематическая 163
– случайная 163
– среднеквадратическая 261

Парадокс Бертрана 24
Пересечение событий 10
Перестановки 26
– с повторениями 28
Переходная плотность 229
– функция 229, 230
Переходные вероятности 55
Плотность 71
Полигон частот 256
Полиномы Бернштейна 186
Полная группа событий 41
Порядковые статистики 254
Правило трех сигм 114
– умножения 26
Предел последовательности событий 13
– – – верхний 12
– – – нижний 13
Произведение событий 10
Производная обобщенного случайного процесса 239
Производящая функция 149
Пространство элементарных событий 9
Процесс с дискретным временем 198
– – непрерывным временем 199
Прошлое 55
Прямая линейной регрессии 298
Прямое произведение вероятностей 47
– – вероятностных пространств 47
– – сигма - алгебр 47

Равномерно наиболее мощный критерий 289
Разбиение 41
Размещение 26
Разность событий 11
Распределение абсолютно непрерывное 71
– Бернулли 69, 301
– бета 74, 302
– биномиальное 70, 301
– Вейбулла – Гнеденко 76, 302
– вероятностей случайной величины 67
– гамма 74, 302
Распределение гауссовское 72, 301
– геометрическое 70, 301
– гипергеометрическое 70, 301
– дискретное 68
– Коши 74, 302
– Лапласа 76, 302
– логарифмически - нормальное 75, 301
– логистическое 76, 302
– многомерное 82
– – нормальное 85, 301
– нормальное 72, 301
– отрицательное биномиальное 71, 301
– полиномиальное 82
– процесса 199
– пуассоновское 70, 301
– равномерное 73, 301
– решётчатое 151
– сингулярное 77
– Стьюдента 156, 302
– сферическое 155
– Фишера 157, 302
– хи - квадрат 155, 302
– экспоненциальное 73, 302
Реализации выборки 252
Регрессия нормальная 295
– среднеквадратическая 297

Свёртка 93
Свойства вероятностей 18
– плотности 72
– функции распределения 65
Свойство согласованности 80
Сигма - алгебра 10
– остаточная 190
– порождённая случайной величиной 87
Слабый белый шум 242
Случайная величина 61
– – абсолютно непрерывная 71
– – асимптотически нормальная 258
– – вырожденная 90
– – дискретная 68
– – комплексная 127
– – многомерная 80
– – простая 116
– – сингулярная 77
– – целочисленная 149
Случайная последовательность 198
Случайное событие 10, 16

Случайные числа 31
Случайный вектор 80
– выбор без возвращения 31
– – с возвращением 30
Случайный процесс 198
– – броуновского движения 214
– – диффузионный 230
– – марковский 228
– – непрерывный в среднем 204, 217
– – однородный 205
– – Орнштейна – Уленбека 234
– – Пуассона 208
– – с независимыми значениями 243
– – с независимыми приращениями 204
– – с некоррелированными приращениями 218
– – с ортогональными приращениями 218
– – сепарабельный 203
– – слабо стационарный 224
– – стационарный 224
– – стохастически непрерывный 204
Смещение 261
Событие достоверное 10
– невозможное 10
– противоположное 11
События несовместные 10
Сочетания 27
– с повторениями 29
Спектр колебания процесса 224
Спектральная плотность 227
– стохастическая мера 228
– функция 226
Спектральное представление корреляционного функционала 248
– – корреляционной функции 225
– разложение процесса 228
Среднее значение случайной величины 95
– – процесса 237
Стандартный процесс броуновского движения 216
Статистика 260
– достаточная 268
– критерия 289
– отношения правдоподобия 290
– полная 269
– центральная 275
Статистическая модель 252
– – абсолютно непрерывная 253
– – дискретная 253
Статистическая модель параметрическая 253
– – экспоненциальная 264
Статистический критерий 281
Стохастическая матрица 56
Стохастически эквивалентные случайные процессы 201
Стохастический интеграл 223
Среднеквадратическое отклонение 104
Сумма событий 10
Схема Бернулли 48
– полиномиальная 49
Сходимость в среднем порядка *r* 175
– по вероятности 173
– поточечная 141
– почти наверное 172
– с вероятностью единица 172
– слабая 141, 173

Теорема Бернулли 179
– Бернштейна 186
– Бореля 186
– Бохнера 133
– Гливенко 254
– единственности 140
– Колмогорова 183, 192, 202
– – о «трёх рядах» 193
– Колмогорова – Ченцова 203
– Лебега 77
– Линдеберга 164
– Ляпунова 168
– Маркова 177
– обратная предельная 145
– о «двух рядах» 193
– – мажорируемой сходимости 96
– – монотонной сходимости 96
– – предельных вероятностях 57
– прямая предельная 145
– умножения 40
– Хелли 142, 143
– Хинчина 178, 192
– центральная предельная 162
– Чебышева 178
– Штольца 178
Тождество Вальда 121
Траектория процесса 199
Уравнение Колмогорова прямое 233
– – обратное 231

Уравнения Колмогорова – Чепмена 229
– правдоподобия 271
Уровень значимости 289
Усиленный закон больших чисел 179
Условие Линдеберга 164
– Ляпунова 168
– пренебрежимой малости 169
– равномерной малости 171
Условное математическое ожидание 123

Формула композиции 92
– обращения 139
– полной вероятности 41
– Стирлинга 52
Формулы Байеса 43
Функция вклада 262
– информации Фишера 262
– мощности критерия 282
– правдоподобия 262
– распределения 63
– – многомерная 80
Функция распределения теоретическая 254
– – эмпирическая 254

Характеристическая функция 127
– – многомерная 152
– – однородного процесса 205
Характеристический функционал обобщенного случайного процесса 238

Цепь Маркова 54
– – однородная 55

Частота 33
– относительная 6

Шаг распределения 151
– – максимальный 152

Экстремальные значения выборки 254
Элемент выборки 252
Элементарное событие 9

УКАЗАТЕЛЬ ОБОЗНАЧЕНИЙ

$\equiv$ – тождественно равно

$\bigcap_n A_n$ – произведение случайных событий A_n

$\bigcup_n A_n$ – сумма событий A_n

$\sum_n A_n$ – сумма попарно несовместных событий A_n

$\prod_{k=1}^{n} a_k$ – произведение $a_1 a_2 \ldots a_n$

$A \Rightarrow B$ – импликация, т.е. высказывание "если *A*, то *B*", или "*A* есть достаточное условие для *B*", или "*B* есть необходимое условие для *A*"

$A \Leftrightarrow B$ – эквиваленция, т.е. единая запись двух импликаций $A \Rightarrow B$ и $B \Rightarrow A$

$a^+ = \max(a;0)$

$a^- = -\min(a;0)$

$a \wedge b = \min(a;b)$

$a \vee b = \max(a;b)$

$\operatorname{Im} f$ – мнимая часть функции f

N – множество натуральных чисел

R – множество действительных чисел

$\operatorname{Re} f$ – действительная часть функции f

Z – множество целых чисел

$\int_{-\infty}^{+\infty} f(t)\,d\xi(t)$ 223

$\int_R g(x)\,dF_\xi(x)$ 117, 118

$\int_R g(x)\,dx$ 118

$\int_a^b g(x)\,dF(x)$ 118

$\int_{-\infty}^{+\infty} g(x)\,dF(x)$ 118

$\int_\Omega \xi(\omega)\,d\mathrm{P}(\omega)$ 95, 117

$\int_a^b \xi(t)\,dt$ 219

$(\Omega, \mathscr{A}, \mathrm{P})$ 15

$(\Omega, \mathscr{A}, \mathscr{P})$ 252

$|A|$ 254

$\varnothing$ 10

A^* 12

A_* 12

$\overline{A}$ 11

$A \bigcup B$ 10

$A + B$ 10

$A \bigcap B$ 10

$A \setminus B$ 11

$A \Delta B$ 11

$\mathscr{A}_1 \times \mathscr{A}_2 \times \ldots \times \mathscr{A}_n$ .47

$\mathscr{A}$ 10

$\mathscr{A}_\xi$ 87

A_{nk} 257

a_{nk} 257

$Bi(N,p)$ 301

$\overline{Bi}(r,p)$ 301

$B(t,s)$ 216

$\mathscr{B}(C(T))$200

$\mathscr{B}(R)$ 34

$\mathscr{B}(R^n)$ 35

$\mathscr{B}_X^T$ 199

$C(T)$ 200

$\operatorname{cov}(\xi,\eta)$ 105

$D(T)$ 200

$D_+(T)$ 200

$D_-(T)$ 200
$\mathrm{D}\xi$ 101
$\mathscr{D}$ 235
$F_{m,n}$ 158, 302
$\mathscr{F}_t$ 228
$\mathscr{F}_{[t,\infty)}$ 228
F_ξ 63
$F_{\bar{\xi}}$ 80
$F_\xi(-\infty)$ 65
$F_\xi(+\infty)$ 65
$F_n \underset{n\to\infty}{\Rightarrow} F$ 141
$G(p)$ 301
$G_n(X)$ 256
$H(N,L,n)$ 301
I_A 63
$i(\theta)$ 263
$i_n(\theta)$ 262
$L(\vec{x};\theta)$ 262
$L(\mu,\sigma^2)$ 301
$\mathscr{L}(\xi)$ 253
$\lim$ 13
$\overline{\lim}$ 13
$\underline{\lim}$ 13
$\mathrm{M}\xi$ 95
$\mathrm{M}(\xi\,|\,B)$ 120
$\mathrm{M}(\xi\,|\,\eta)$ 124
$\mathrm{M}(\xi\,|\,\mathscr{F})$ 123
M_{nk} 25
m_{nk} 257
$N(A)$ 6, 33
$N(\mu,\sigma^2)$ 301
$\mathscr{P}(\Omega)=2^\Omega$ 20
$\mathrm{P}(A)$ 6, 15
$\mathrm{P}(A|B)$ 40
$\mathrm{P}_1\times\mathrm{P}_2\times\ldots\times\mathrm{P}_n$ 47
$P(s,x,t,B)$ 229
$p(s,x,t,z)$ 229
p_ξ 71
$p_{\bar{\xi}}$ 84
$\left[p_{ij}\right]$ 55
R^T 200
$R(a,b)$ 301
$R(t,s)$ 217
S^2 257
s^2 257
t_υ 302
$U(X;\theta)$ 262
$\bar{X}$ 257
$X_{(k)}$ 254
$\bar{X}_n$ 257
$Z_{n,p}$ 259
$\xi_n \xrightarrow[n\to\infty]{\text{п.н.}} \xi$ 172
$\xi_n \xrightarrow[n\to\infty]{r} \xi$ 175
$\xi_n \xrightarrow[n\to\infty]{\mathrm{P}} \xi$ 173
$\xi_n \Rightarrow \xi$ 173
$\Pi(\lambda)$ 301
$\rho(\xi,\eta)$ 113
$\sigma(\mathscr{A})$ 34
$\Phi(x)$ 162
χ^2 155, 302
ω 9

Обратная предельная теорема

$\{f_n\}, n \geq 1$ - хар. ф-ция, f непр. в т. 0, $f_n(t) \xrightarrow[n\to\infty]{} f(t)\ \forall t \in \mathbb{R} \Rightarrow$
$\Rightarrow F_n \underset{n\to\infty}{\Rightarrow} F$, $F, F_n, n \geq 1$ - ф-ции распределения.

Доказательство.

1

1 теорема Хелли $\Rightarrow \exists \left\{F_{n_k}\right\} \subset \{F_n\}$, $F_{n_k} \underset{n_k\to\infty}{\Rightarrow} F$

$$P\left(\left|\xi_{n_k}\right| \leq x\right) \geq \frac{\left|\frac{1}{2\tau}\int_{-\tau}^{\tau} f_{n_k}(t)\,dt\right| - \frac{1}{\tau x}}{1 - \frac{1}{\tau x}},\ \tau, x > 0 \Rightarrow$$

$$\Rightarrow \forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}:\ \forall n_k > n_0 \qquad F_{n_k}\left(\frac{2}{\tau}\right) - F_{n_k}\left(\frac{2}{\tau} - 0\right) \geq 1 - \varepsilon$$

$F(+\infty) - F(-\infty) = 1$, F - ф-ция. распр. $\Rightarrow$ (прям. пред. теор.)

$$f_{n_k}(t) \xrightarrow[n_k\to\infty]{} \int_{-\infty}^{+\infty} e^{itx}\,dF(x) \Rightarrow f(t) = \int_{-\infty}^{+\infty} e^{itx}\,dF(x) \text{ - пр. ф-ция.}$$

2

$F_n \underset{n\to\infty}{\Rightarrow} F$ **?** От противного $F_n \underset{n\to\infty}{\not\Rightarrow} F$

$\exists \left\{F_{m_k}\right\} \subset \{F_n\}$, $F_{m_k} \underset{m_k\to\infty}{\Rightarrow} F^*$, $F^* \neq F$, F^*, F - ф-ции распр.

$f_{m_k}(t) \to f^*(t)$
$f_{n_k}(t) \to f(t)$
$f(t) \neq f^*(t)$
противоречие

Прямая предельная теорема

Теорема единственности

Обозначения.

$$a_k = M\xi_k,\ B_n^{\,2} = \sum_{k=1}^{n} D\xi_k,\ F_k(x) = P(\xi_k \le x),\ \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt,\ \xi_{kn} = \frac{\xi_k - a_k}{B_n},$$

$$S_n = \sum_{k=1}^{n}\xi_{kn}, \varphi_{kn}(t) = Me^{it\xi_{kn}}, \varphi_{S_n}(t) = Me^{itS_n}, \gamma_n(\tau) = \frac{1}{B_n^{\,2}}\sum_{k=1}^{n}\int_{|x-a_k|>\tau B_n}(x-a_k)^2\,dF_k(x).$$

Теорема Линдеберга.

$\xi_n, n \in \mathbb{N}$ - независимые случайные величины; $\forall \tau > 0\ \ \gamma_n(\tau) \underset{n\to\infty}{\to} 0$

⇓

$$\forall x \in \mathbb{R} \qquad P\left(\frac{\xi_1 + \xi_2 + \ldots + \xi_n - (a_1 + a_2 + \ldots + a_n)}{B_n} \le x\right) \underset{n\to\infty}{\to} \Phi(x) \qquad \text{В}$$

Доказательство.

$$\ln(1+x) = \sum_{s=1}^{\infty}\frac{(-1)^{s-1}}{s}x^s$$

$$\ln\varphi_{S_n}(t) = \sum_{k=1}^{n}\ln\varphi_{kn}(t)$$

$\ln\varphi_{S_n}(t) = \sum_{k=1}^{n}(\varphi_{kn}(t)-1) + R_n$, где $R_n = \sum_{k=1}^{n}\sum_{s=2}^{\infty}\frac{(-1)^{s-1}}{s}(\varphi_{kn}(t)-1)^s$,

$|R_n| \le \frac{t^2}{2}\max_{1\le k\le n}|\varphi_{kn}(t)-1|$, т.е. $R_n \underset{n\to\infty}{\to} 0$ равномерно в каждом конечном интервале

⇓

$\sum_{k=1}^{n}(\varphi_{kn}(t)-1) = -\frac{t^2}{2} + \rho_n$, где $\rho_n = \frac{t^2}{2} + \sum_{k=1}^{n}\int_{-\infty}^{+\infty}\left(e^{itx} - 1 - itx\right)dF_{kn}(x)$,

$\rho_n \underset{n\to\infty}{\to} 0$ равномерно в каждом конечном интервале

⇓

$\ln\varphi_{S_n}(t) \underset{n\to\infty}{\to} -\frac{t^2}{2}$ равномерно в каждом конечном интервале

⇓

В

Теорема 8 (Колмогоров)

ξ_n - независимы, $n \geq 1$,
$F_{\xi_n}(x) = F_{\xi_1}(x)$, $M\xi_1 = a < +\infty$

$\Downarrow K_1$

$$\frac{1}{n}\sum_{k=1}^{n} \xi_k \xrightarrow[n\to\infty]{п.н.} a$$ B_1

ξ_n - независимы, $n \geq 1$, $F_{\xi_n}(x) = F_{\xi_1}(x)$,
$\exists a \in \mathbb{R}: \frac{1}{n}\sum_{k=1}^{n} \xi_k \xrightarrow[n\to\infty]{п.н.} a$

$\Downarrow K_2$

$M|\xi_1| < +\infty$ B_2

Доказательство K_1.

Обозначение: $\xi_n \in A_1$, $\overline{\xi_n} = \xi_n \cdot I_{\{|\xi_n| \leq n\}} = \begin{cases} \xi_n, |\xi_n| \leq n \\ 0, |\xi_n| > n \end{cases}$

$$\frac{1}{n}\sum_{k=1}^{n} \xi_k - na = \frac{\sum_{k=1}^{n} \xi_k - \sum_{k=1}^{n} \overline{\xi_k}}{n} + \frac{\sum_{k=1}^{n} \overline{\xi_k} - \sum_{k=1}^{n} M\overline{\xi_k}}{n} + \frac{\sum_{k=1}^{n} M\overline{\xi_k} - na}{n} = I_1 + I_2 + I_3$$

Лемма Бореля-Кантелли
Лемма 1 (§1 доп. к гл.6)

$\Rightarrow$ $I_1 \xrightarrow[n\to\infty]{п.н.} 0$ C_1

$\sum_{k=1}^{n} \frac{D\overline{\xi_k}}{k^2} < +\infty$ $\xrightarrow{\text{Теорема 7}}$ $I_2 \xrightarrow[n\to\infty]{п.н.} 0$ C_2

$I_3 = -\frac{1}{n}\sum_{k=1}^{n} M\xi_k \cdot I_{\{|\xi_k| > k\}} \xrightarrow[n\to\infty]{п.н.} 0$ C_3

$A_1 \Rightarrow (C_2 \ C_1 \ C_3) \Rightarrow B_1$

Доказательство K_2.

$$A_2 \Rightarrow \frac{\xi_n}{n} = \frac{\sum_{k=1}^{n} \xi_k}{n} - \frac{n-1}{n} \cdot \frac{\sum_{k=1}^{n-1} \xi_k}{n-1} \xrightarrow[n\to\infty]{п.н.} 0 \Rightarrow B_2$$

Вторая теорема Хелли

$$F_n,\ F \text{ — ф-р},\ n \geq 1;\ F_n \underset{n\to\infty}{\Rightarrow} F \Rightarrow \int_{-\infty}^{+\infty} g(x)\,dF_n(x) \xrightarrow[n\to\infty]{} \int_{-\infty}^{+\infty} g(x)\,dF(x)\ \forall g \in C_B(R)$$

Доказательство.

1

$$a = x_0 < x_2 < \ldots < x_L = b \in C(F),\ g_\varepsilon(x) \overset{df}{=} g(x_k),\ x \in (x_{k-1}, x_k] \Rightarrow$$
$$\forall \varepsilon > 0\ \left|g(x) - g_\varepsilon(x)\right| < \varepsilon$$

$$\left|\int_{-\infty}^{+\infty} g(x)\,dF_n(x) - \int_{-\infty}^{+\infty} g(x)\,dF(x)\right| \leq \int_{|x|>X} \left|g(x)\right| dF_n(x) +$$
$$+\left|\int_{|x|\leq X} g(x)\,dF_n(x) - \int_{|x|\leq X} g(x)\,dF(x)\right| + \int_{|x|>X} \left|g(x)\right| dF(x) \xrightarrow[n\to\infty]{} 0$$

2

$$\exists X, -X \in C(F)\ F(-X) < \frac{\varepsilon}{2}, \qquad 1 - F(X) < \frac{\varepsilon}{2};$$
$$F_n \underset{n\to\infty}{\Rightarrow} F \Rightarrow \exists n_0\ \forall n > n_0\ F(-X) < \varepsilon, 1 - F(X) < \varepsilon.$$

1

$$\left|\int_a^b g(x)\,dF_n(x) - \int_a^b g(x)\,dF(x)\right| \leq \left|\int_a^b \left(g(x) - g_\varepsilon(x)\right) dF_n(x)\right| +$$
$$+\left|\int_a^b g_\varepsilon(x)\,dF_n(x) - \int_a^b g_\varepsilon(x)\,dF(x)\right| + \left|\int_a^b \left(g_\varepsilon(x) - g_\varepsilon(x)\right) dF(x)\right| \xrightarrow[n\to\infty]{} 0$$

Свойства мультипликативности математического ожидания

$$\xi_1,\dots,\xi_n \text{ — независ.; } M\xi_i<\infty, i=\overline{1,n} \Rightarrow \exists M\left(\xi_1,\xi_2,\dots,\xi_n\right)$$
$$M\left(\xi_1,\xi_2,\dots,\xi_n\right)=M\xi_1,M\xi_2,\dots,M\xi_n$$

Доказательство (n=2).

1

$$\xi_1(\omega)=\sum_{k=1}^{m}c_k I_{A_k}(\omega),\quad \xi_2(\omega)=\sum_{i=1}^{l}b_j I_{B_j}(\omega) \text{ — нез.}$$

Критерий независимости дискретных случайных величин

$$P\left(A_k\cap B_j\right)=P\left(A_k\right)P\left(B_j\right),\ k=\overline{1,m},\ j=\overline{1,l}$$

$$M\xi_1\xi_2=\sum_{k=1}^{m}\sum_{j=1}^{l}c_k b_j P\left(A_k\cap B_j\right)=M\xi_1\cdot M\xi_2$$

2

$$\xi_1\ge 0,\xi_2\ge 0 \text{ — нез, } M\xi_i<\infty, i=\overline{1,2}$$

ξ_1,ξ_2 — *нез.,* g_1,g_2 — *борел. ф-ции* $\Rightarrow g_1(\xi_1),g_2(\xi_2)$ — *нез*

A

$$\xi_{in}^{(\omega)}\overset{df}{=}\sum_{k=1}^{n2^2}\frac{k-1}{2^n}I_{\left(\frac{k-1}{2^n}<\xi\le\frac{k}{2^n}\right)}(\omega),\ i=1,2;\ \xi_{1n},\xi_{2n}-\text{нез.},\ \xi_{in}\uparrow\xi_i, n\to\infty,\ i=1,2$$

Теор. о монотон. сходимости

1

$$M\xi_1\xi_2=\lim_{n\to\infty}M\xi_{1\eta}\xi_{2\eta}=M\xi_1\cdot M\xi_2$$

3

$$\xi_1,\xi_2 \text{ нез., } M\xi_i<\infty$$

A

$$\xi_i^+(\omega)\overset{df}{=}\max\left(0,\xi_i(\omega)\right),\xi_i^-(\omega)\overset{df}{=}-\min\left(0,\xi_i(\omega)\right),\xi_1^\pm\ge 0,\xi_2^\pm\ge 0-\text{нез.}$$

2 A

$$M\xi_1\xi_2=M\left(\xi_1^+-\xi_1^-\right)\left(\xi_2^+-\xi_2^-\right)=M\xi_1\cdot M\xi_2$$

Локальная предельная теорема Муавра-Лапласа

$$P_n(m) = P\left(\frac{\mu_n - np}{\sigma} = x\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2}}\left(1+o(1)\right),\ \sigma = \sqrt{npq} \to \infty$$

$$x \in \left\{x \in R : |x| \le C, x = \frac{m-np}{\sigma}, m = 0,1,2,\dots\right\}$$

Доказательство.

$$\ln P_n(m) = \ln n! - \ln m! - \ln(n-m)! + m\ln p + (n-m)\ln q$$

Формула Стирлинга

$$\ln n! = n\ln n + \ln\sqrt{2\pi n} - n + \theta_n$$

$$|\theta_n| < \frac{1}{12n},\ n \to \infty$$

$$\ln P_n(m) = \overset{m\ln n + k\ln n}{\cancel{n\ln n}} + \ln\sqrt{2\pi n} - n +$$
$$+\theta_n - m\ln m - \ln\sqrt{2\pi m} + m + \theta_m -$$
$$-k\ln k - \ln\sqrt{2\pi k} + k + \theta_k + m\ln p + k\ln q$$

$$m = np + x\sigma = np\left(1 + \frac{xq}{\sigma}\right) \xrightarrow[n\to\infty]{} \infty$$

$$k = n - m = nq - x\sigma = nq\left(1 + \frac{xp}{\sigma}\right) \xrightarrow[n\to\infty]{} \infty$$

$$\ln(1+\varepsilon) = o(\varepsilon)$$

$$\ln(1+\varepsilon) = \varepsilon - \frac{\varepsilon^2}{2} + o\left(\varepsilon^3\right), \varepsilon \to 0$$

$$\ln P_n(m) = \ln\frac{1}{\sigma\sqrt{2\pi}} - \frac{x^2}{2} + o\left(\frac{1}{\sigma}\right)$$

Формула замены переменных в интеграле Лебега

$$g \in C(R), Mg(\xi) < \infty \Rightarrow Mg(\xi) = \int\limits_{\Omega} g(\xi(\omega))dP(\omega) = \int\limits_{-\infty}^{\infty} g(x)dF_{\xi}(x)$$

Доказательство.

1

$$g \in C[a,b]; g(x) = 0, x \notin [a,b]; g_n(x) \overset{df}{=} g(x_{nk}), x \in \left(x_{n\,k-1}, x_{nk}\right],$$

$$x_{nk} = a + \frac{b-a}{n}k, k = \overline{0,n} \Rightarrow \left|g(x) - g_n(x)\right| < \varepsilon \ \forall \varepsilon > 0, \exists n_0, \forall n > n_0 \forall x \in [a,b]$$

$\Downarrow$ Теорему Лебега о предельном переходе

$$Mg_n(\xi) = \int\limits_a^b g_n(x)dF_{\xi}(x), Mg_n(\xi) \xrightarrow[n\to\infty]{} Mg(\xi),$$

$$\left|\int\limits_a^b g(x)dF_{\xi}(x) - \int\limits_a^b g_n(x)dF_{\xi}(x)\right| < \varepsilon, \forall n > n_0 \ \forall \varepsilon > 0$$

2

$$g \in C(R), g(x) \ge 0 \ \forall x \in R, g_N(x) \overset{df}{=} \begin{cases} 0, |x| > N \\ g(x), |x| \le N \end{cases} \Rightarrow g_N \underset{N\to\infty}{\uparrow} g$$

$\Downarrow$ 1, Теорема о монотонной сходимости

$$Mg_N(\xi) \xrightarrow[N\to\infty]{} Mg(\xi), Mg_N(\xi) = \int\limits_{-N}^{N} g(x)dF_{\xi}(x)$$

$$\int\limits_{-N}^{N} g(x)dF_{\xi}(x) \xrightarrow[N\to\infty]{} \int\limits_{-\infty}^{\infty} g(x)dF_{\xi}(x)$$

3

$$g \in C(R), g^+(x) \overset{df}{=} \max(0, g(x)), g^-(x) \overset{df}{=} -\min(0, g(x)), \ g(x) = g^+(x) - g^-(x)$$

$\Downarrow$ 2

$$Mg^{\pm}(\xi) = \int\limits_{-\infty}^{\infty} g^{\pm}(x)dF_{\xi}(x)$$

$\Downarrow$ Линейность интеграла Римана-Стилтьеса

$$Mg(\xi) = \int\limits_{-\infty}^{\infty} g(x)dF_{\xi}(x)$$