Министерство высшего и среднего специального образования РСФСР

ЛЕНИНГРАДСКИЙ ОРДЕНА ЛЕНИНА
ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ
имени М. И. КАЛИНИНА

Б. Е. АКСЕНОВ, И. В. АФОНЬКИН, В. П. ЕВМЕНОВ,
М. И. НЕЧИПОРЕНКО

ОСНОВЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

УЧЕБНОЕ ПОСОБИЕ

Часть 1

Ленинград
1973

Данное пособие является расширенным конспектом курса лекций, читаемого студентам факультета радиоэлектроники ЛПИ им. М. И. Калинина специальностей 0608 и 0646. Пособие состоит из трех частей: «Основы теории вероятностей», «Основы математической статистики» и «Основы теории случайных процессов».

Настоящее издание включает первую часть — «Основы теории вероятностей», в которой излагаются основные понятия и результаты теории: случайные события и вероятность, законы распределения и числовые характеристики случайных величин, предельные теоремы теории вероятностей.

Пособие может быть полезным и для инженеров, желающих познакомиться с теорией вероятностей.

Рецензенты: канд. техн. наук *Н. В. Ращепкин*, канд. техн. наук *И. С. Сребрянский*, ст. преподаватель *Г. К. Дондуа*.

ПРЕДИСЛОВИЕ

Учебные планы подготовки инженеров по специальностям «Электронные вычислительные машины» и «Автоматизированные системы управления» включают предметы, изучение которых требует от студента определенного минимума знаний по теории вероятностей, теории случайных процессов и математической статистике. К настоящему времени в каждой из этих дисциплин накоплен настолько обильный материал, что студентам трудно в нем ориентироваться и выбирать нужные сведения, рассеянные по отдельным учебникам и монографиям. Данное пособие содержит материал в объеме, необходимом для таких дисциплин, как основы теории надежности, основы теории передачи информации, исследование операций и т. д.

Пособие предназначено для первоначального, но достаточно глубокого изучения предмета, требующего от читателя подготовки по математике в пределах программы технического вуза.

Авторы выражают благодарность А. П. Аксенову, Г. К. Дондуа, И. С. Сребрянскому, Г. Н. Черкесову за полезные замечания и советы по рукописи, а также И. Д. Бутомо и В. С. Алмазовой, взявшим на себя труд редакторов.

1*

Введение

В разговорной речи мы часто употребляем слова *возможно, вероятно, сомнительно, безусловно* и т. п., выражая ими степень своей уверенности в том, что явление (событие), о котором идет речь, может произойти или не произойти. Со случайными явлениями человек встречается постоянно и в быту и в сфере производственной или научной деятельности. Окружающий нас мир находится в непрерывном движении, и все происходящие в нем события связаны сложными причинно-следственными отношениями. Если человек выражает неуверенность по поводу возможности появления некоторого события, то это происходит потому, что он не располагает достаточными сведениями о всех причинах, влияющих на появление этого события. Такая ситуация возникает в том случае, если еще не раскрыта совокупность всех необходимых для появления события причин либо причин этих настолько много, что учесть их в полном объеме практически невозможно. Последнее является скорее правилом, чем исключением.

Явления природы, изучаемые наукой и используемые человеком, имеют обычно повторяющийся характер, что является следствием многократного осуществления комплекса причин (условий), определяющих их возникновение. Большое число реализаций некоторого комплекса причин создает массу однотипных явлений. Только относительно явлений, имеющих повторяющийся, массовый характер, у нас могут сложиться объективные представления об их случайности, степени их возможности, и, очевидно, прежде всего такого типа явления представляют интерес для человека.

Практика показывает, что масса случайных явлений очень часто проявляет вполне определенные устойчивые закономерности. Эти закономерности не являются простым повторением индивидуальных особенностей отдельных единичных явлений, составляющих массу. В них отражаются главные, существенные, ведущие причинно-следственные связи, а несущественные, второстепенные сглаживаются и исчезают. Каждое отдельно взятое случайное явление может произойти в той или иной форме или не произойти, но весьма большое скопление случайных явлений приводит к определенности и закономерной необходимости. Многие фундаментальные законы неживой и живой природы являются проявлением массовости случайных событий. Достаточно напомнить о втором законе термодинамики, указывающем на необратимость перехода работы в тепло; механизме естественного отбора, обеспечивающем эволюционное развитие всего живого на основе случайных мутаций молекул дезоксирибонуклеиновых кислот; законе стоимости, согласно которому затраты общественного труда на производство товара в условиях товарного производства принимают форму стоимости. В массовых случайных явлениях воплощается диалектическое единство случайного и необходимого.

Принципиальная невозможность учета полной картины причинно-следственных связей единичного случайного явления во всех подробностях и деталях не является преградой для познания закономерностей массовых явлений. Более того, рассмотрение всех деталей, всех связей, в том числе и несущественных для данного явления, привело бы лишь к запутыванию сути явления и затруднило бы его познание.

Закономерности массовых случайных явлений изучаются теорией вероятностей, теорией случайных процессов и математической статистикой, которые являются разделами высшей математики. Эти дисциплины, как следует из вышесказанного, развивают свой аппарат и методы на основе некоторого упрощения или схематизации природы. Насколько правомерна допущенная схематизация, мы можем судить по согласию теории с опытом, с практикой. История физики свидетельствует, что аппарат теоретико-вероятностных методов оказался хорошо приспособленным к изучению многочисленных явлений природы, а в некоторых случаях он является единственно возможным.

Специалист в области управления и вычислительной техники обязан хорошо владеть основными понятиями и методами теории вероятностей, теории случайных процессов и математической статистики. Во-первых, сама проблема управления любым объектом возникает в связи с тем, что нам необходимо обеспечивать его целенаправленное движение или развитие в условиях случайных воздействий и помех, т. е. задачи управления по своей природе имеют дело со случайными явлениями. Во-вторых, в самой аппаратуре управления и обработки информации возникают случайные явления, а инженер обязан либо бороться с ними, либо использовать их для достижения желаемых характеристик аппаратуры.

И, наконец, очень часто методы управления и обработки информации являются статистическими методами, что отражается на структуре и технических характеристиках аппаратуры. Этими обстоятельствами объясняется, в частности, тот факт, что интенсивное внедрение в практику автоматизированных средств управления и обработки информации сопровождается, с одной стороны, дальнейшим развитием теоретико-вероятностных методов, а с другой — быстрым расширением сферы приложений этих методов для решения научных и практических задач.

ГЛАВА 1

ВЕРОЯТНОСТИ СЛУЧАЙНЫХ СОБЫТИЙ

§ 1.1. Первичные понятия теории вероятностей

Развитие теории вероятностей стимулировалось потребностью решения практических и научных задач в различных областях деятельности человека и происходило на основе экспериментальных наблюдений и теоретических обобщений. В современном состоянии это логически стройная математическая теория, построенная по аксиоматическому принципу.

Аксиоматические теории строятся на основе нескольких не определяемых в самой теории понятий, называемых первичными, и конечного числа высказываний — аксиом о свойствах этих понятий. Построение всех остальных понятий и вывод их свойств осуществляется формально-логическими средствами без обращения к наглядным физическим представлениям и доказательствам на основе «здравого смысла».

Практическая ценность таких математических теорий, в частности теории вероятностей, следует из того, что их основания — первичные понятия и утверждения (аксиомы) являются обобщением длительного человеческого опыта.

Выработка первичных понятий путем обобщения опыта состояла в абстрагировании от частных, специфических деталей и фактов, связанных с физической природой конкретных задач. Применение абстрактной теории требует, во-первых, обратного перехода — наполнения абстрактных понятий конкретным физическим содержанием, связанным с природой задачи, или, как говорят, интерпретации теории в терминах конкретной задачи и, во-вторых, проверки выполнимости, т. е. истинности, аксиом в условиях задачи. Последнее чаще всего осуществляется косвенным путем — сопоставлением выводов, полученных на основе теории, и результатов экспериментальных наблюдений.

Невозможно указать какие-либо правила интерпретации абстрактной теории. В каждом частном случае преодолению этого этапа применения теории вероятностей в немалой степени способствуют интуиция и опыт, которые могут быть развиты только путем упражнений. Поэтому изучение теории должно сопровождаться решением задач.

Первичным понятием теории вероятностей является понятие элементарное событие. Оно не определяется в теории, но смысл и содержание его можно раскрыть на конкретных примерах. С этой целью и для разъяснения общепринятой в теории вероятностей терминологии обратимся вначале к тем интуитивным соображениям, которые оправдывают введение указанного понятия, имея в виду, что никакие заключения из последующих рассуждений ни в коей мере не могут рассматриваться в качестве его определения.

Событием можно назвать любое явление природы. Чаще так называют явления, которые по каким-либо причинам оказываются в сфере нашего внимания, и мы предпринимаем усилия для того, чтобы их исследовать и познать. Изолированных явлений в природе не существует. Для возникновения любого явления необходимо осуществление вполне определенного комплекса условий или причин.

Собственно изучение явлений природы и имеет целью выявить, раскрыть существующие причинно-следственные связи. Всегда, если речь идет о каком-то событии, явно или неявно предполагается наличие вполне определенного, фиксированного комплекса условий, связанного с этим событием. В дальнейшем события будем обозначать заглавными буквами начала латинского алфавита. Совокупность обусловливающих возникновение события A факторов, которые можно реально или по крайней мере принципиально контролировать, назовем основным комплексом условий, связанным с событием A, и обозначим S. Определением «основной» подчеркивается тот факт, что в S входят главные, основные условия, но иногда он может не включать всех условий, необходимо вызывающих появление события A.

Осуществление основного комплекса условий, связанного с какими-либо событиями, называется в теории вероятностей опытом, экспериментом, испытанием или наблюдением. При этом не делается различия между целенаправленными, спланированными экспериментами и пассивными наблюдениями. В целях единообразия терминологии в дальнейшем будем использовать термин *испытание.*

Целенаправленное испытание проводят, например, при определении качества или состояния канала связи, для чего по каналу передают определенное сообщение, т. е. сигнал (тест), и сравнивают его с сообщением, полученным на приемном конце. Аналогичным образом поступают, если желают убедиться в исправности вычислительной машины. С этой целью на машине выполняют программу (тест) с заранее известными результатами. Естественно, что основной комплекс условий в обоих случаях должен включать предусмотренные разработчиками канала связи или вычислительной машины режимы работы аппаратуры. Кроме того, туда же могут входить какие-либо характеристики теста, условия внешней среды, условия обслуживания и т. д.

Опыт эксплуатации каналов связи и вычислительных машин показывает, что соблюдение всех режимов работы аппаратуры, указанных в инструкциях по эксплуатации, не гарантирует прохождение сообщения по каналам связи без искажений или отсутствие ошибок в результатах выполнения программы машиной.

Явления или события, которые фиксирует наблюдатель, проводящий конкретное испытание, могут быть самые разнообразные и определяются только его целями. Так, в случае канала связи наблюдателя могут интересовать следующие события: отсутствие каких-либо искажений в принятом сообщении; возникновение любого искажения; возникновение искажения, препятствующего приему сообщения; возникновение искажения, преобразующего переданное сообщение в другое осмысленное сообщение, и т. д.

В случае испытания вычислительной машины событиями могут быть: выполнение теста машиной без ошибок; возникновение хотя бы одной ошибки в результатах выполнения теста; возникновение ошибки определенного вида, например в результате неправильной работы арифметического устройства машины, и т. д.

Пассивными наблюдениями приходится ограничиваться при изучении, например, закономерностей возникновения и распространения эпидемии гриппа. В данном случае основной комплекс условий может включать место и время года, а событиями, связанными с ним, могут быть: возникновение эпидемии гриппа; отсутствие эпидемии; возникновение эпидемии, вызванной определенным типом возбудителя, и т. д.

Если событие наблюдается при каждом испытании, то его называют достоверным, а невозможным называют событие, которое не может появиться при испытании. Событие называют случайным, если оно при испытании может появиться, а может и не появиться. В таком случае испытание называют испытанием со случайным исходом.

Очень часто встречаются случайные события, обнаруживающие устойчивость частоты появления в сериях испытаний. Серией испытаний называют многократное (одновременное или последовательное во времени) осуществление основного комплекса условий, а частотой $W(A)$ случайного события A — отношение числа испытаний с исходом A к числу испытаний в серии. Если мы установим, что в каждой серии испытаний численное значение $W(A)$ примерно одно и то же, то у нас возникнут основания предполагать существование закономерной связи между комплексом S и событием A. Такую связь при $0 < W(A) < 1$ называют статистической, вероятностной или стохастической, а при $W(A) = 1$ — детерминированной.

Устойчивость частоты появления события в сериях испытаний дает основание ввести для таких событий численную характеристику (меру) возможности появления события при испытании. Эта характеристика и называется вероятностью случайного события. Предметом изучения теории вероятностей являются только такие случайные события, для которых можно говорить о вероятности в указанном смысле.

Анализируя различные конкретные испытания со случайными исходами, мы обнаружим следующее. Во-первых, с каждым испытанием связано вполне определенное множество различающихся между собой возможных исходов, причем каждый исход является случайным событием, а число их в множестве не менее двух. Действительно, если бы испытание имело один единственный возможный исход, то этот исход был бы достоверным событием. Во-вторых, наблюдатель, руководствуясь своею собственной конкретной задачей, иногда отождествляет некоторые различающиеся исходы, рассматривая появление каждого из них как одно и то же случайное событие, хотя в другой задаче те же самые исходы он будет рассматривать как различные случайные события.

Последнее обстоятельство заставляет различать в случае каждого испытания неразложимые исходы и составные исходы и ввести в теорию вероятностей соответствующие им понятия элементарного (неразложимого) события и составного (разложимого) события. А первое обстоятельство заставляет связывать с каждым испытанием пространство (множество) элементарных исходов.

Всякий раз, когда мы хотим применить теорию вероятностей к конкретной задаче, мы должны прежде всего условиться о том, какие исходы испытания в данной задаче будем считать неразложимыми. Каждый неразложимый исход испытания представляется в теории одним и только одним элементарным событием. По определению совокупность всех элементарных событий называется *пространством* элементарных событий, а сами элементарные события — *точками* этого пространства. В теории вероятностей никогда не рассматриваются события и вероятности событий вне связи с каким-нибудь пространством элементарных событий, вне связи с каким-либо реальным или мысленным испытанием.

После того как для некоторого испытания установлены исходы испытания, принимаемые по соглашению в качестве неразложимых, и тем самым определено пространство элементарных событий, последнее становится идеализированной моделью данного испытания в том смысле, что любой исход испытания полностью описывается одним и только одним элементарным событием. О каком-либо произвольном событии (не элементарном) A имеет смысл говорить лишь в том случае, если для любого исхода испытания известно, произошло или не произошло событие A. Подмножество точек пространства элементарных событий, представляющих все те исходы, при появлении которых, по мнению наблюдателя, происходит событие A, полностью описывает это событие. И, наоборот, произвольное подмножество A, содержащее одну или более точек пространства элементарных событий, можно рассматривать как событие. Это событие происходит или не происходит при испытании в зависимости от того, принадлежит или не принадлежит подмножеству A точка, изображающая наблюдаемый исход испытания.

По определению *событием* называется любое подмножество пространства элементарных событий. Если подмножество одноэлементное, то это — элементарное событие. Составными событиями являются многоэлементные подмножества. Очевидно, пустое подмножество является невозможным событием, а все пространство элементарных событий — достоверным событием. Действительно, испытание всякий раз заканчивается каким-либо одним элементарным событием, поэтому событие, описываемое пустым подмножеством, никогда не может наблюдаться, а событие, описываемое всем пространством, наблюдается при каждом испытании. Таким образом, достоверное и невозможное события рассматриваются как частные случаи случайного события.

В зависимости от физической природы объектов исследования пространство элементарных событий может быть конечным, счетным или несчетным.

Пример 1.1. Испытание состоит в трехкратном подбрасывании монеты. Природа испытания подсказывает целесообразность выбора в качестве пространства элементарных событий множества из восьми точек, которые условно можно обозначить: ГГГ, ГГЦ, ГЦГ, ЦГГ, ГЦЦ, ЦГЦ, ЦЦГ, ЦЦЦ. Здесь Ц — означает выпадение цифры, Г — выпадение герба.

Событие E_1 — «число выпадений герба не меньше числа выпадений цифры» есть подмножество $E_1 = \{$ГГГ, ГГЦ, ГЦГ, ЦГГ$\}$. При проведении испытания это событие может наступить, а может и не наступить. Событие E_2 — «число выпадений герба равно числу выпадений цифры» при данном испытании — невозможное событие: $E_2 = V$. Событие E_3 — «хотя бы раз выпадет герб или хотя бы раз выпадет цифра» есть достоверное событие: $E_3=U$.

Пример 1.2. По каналу связи передача ведется двоичными кодами с контролем по четности. Для этого на передающем пункте в одном из разрядов кодового слова, например в последнем, устанавливают 0 или 1 так, чтобы общее число единиц в слове было четным, а на приемном пункте проверяют четность числа единиц. Если в канале связи чаще всего возникают искажения типа одиночных трансформаций 0 в 1 или 1 в 0, то такой способ кодирования, очевидно, повышает достоверность передачи, так как позволяет обнаружить все искажения кодовых слов типа нечетного числа трансформаций.

Пусть при испытании качества связи по каналу передается сообщение, состоящее из трех кодовых слов 0011, 0110, 1001, которые обозначим соответственно буквами a, b, c. На приемный пункт каждое слово может поступить без искажений (состояние 1), с нечетным (состояние 2) либо с четным числом трансформаций (состояние 3). Переданное сообщение превращается в ложное при любом числе трансформаций в любом кодовом слове. Если взять в качестве элементарного события распределение кодовых слов сообщения на приемном пункте по состояниям, то получим пространство элементарных событий, все точки которого перечислены в табл. 1.1. При таком соглашении об элементарном событии исходом любого испытания канала связи является одно из распределений табл. 1.1 и никаких других исходов быть не может. Иные исходы, например такой, как непоступление на приемный пункт целого кодового слова в результате обрыва связи, не рассматриваются либо в связи с тем, что они не могут возникнуть в условиях испытания, либо просто потому, что не интересуют наблюдателя. Отсюда следует, что все пространство элементарных событий, т. е. множество $U=\{1, 2, \ldots, 27\}$, является достоверным событием, а пустое множество — невозможным событием.

Одноэлементное множество $\{1\}$ есть элементарное событие, которое можно назвать «сообщение принято правильно». Множество $\{4\}$ — элементарное событие: «кодовые слова a и b приняты правильно, а в слове c возникло нечетное число трансформаций», т. е. вместо 1001 принято 0001 или 1011 и т. д.

Событие E_1 — «контроль по четности обнаруживает искажение в сообщении» — изображается множеством $E_1=\{2, 4, 5, 6, 10, 11, 12, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27\}$. Событие E_2 — «контроль по четности не обнаруживает искажения в сообщении» — изображается множеством $E_2=\{1, 3, 7, 8, 9, 13, 14, 15\}$. Событие E_3 — «ложное сообщение принима-

Таблица 1.1

№	Состояние слова			№	Состояние слова			№	Состояние слова		
	1	2	3		1	2	3		1	2	3
1	*abc*			10	*a*	*bc*		19		*a*	*bc*
2		*abc*		11	*b*	*ac*		20		*b*	*ac*
3			*abc*	12	*c*	*ab*		21		*c*	*ab*
4	*ab*	*c*		13	*a*		*bc*	22	*a*	*b*	*c*
5	*ac*	*b*		14	*b*		*ac*	23	*a*	*c*	*b*
6	*bc*	*a*		15	*c*		*ab*	24	*b*	*a*	*c*
7	*ab*		*c*	16		*ab*	*c*	25	*b*	*c*	*a*
8	*ac*		*b*	17		*ac*	*b*	26	*c*	*a*	*b*
9	*bc*		*a*	18		*bc*	*a*	27	*c*	*b*	*a*

ется как правильное» имеет вид множества $E_3=\{3, 7, 8, 9, 13, 14, 15\}$. Множество $E_4=\{2, 3, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27\}$ можно назвать событием: «число искаженных слов в сообщении не менее двух». Множество $E_5=\{3, 13, 14, 15\}$ имеет вид: «число искаженных кодовых слов не менее двух и в каждом искаженном слове четное число трансформаций». Множество $E_6=\{7, 8, 9\}$ изображает событие: «искажено одно кодовое слово и число трансформаций четное».

Если наблюдателя интересует, например, событие E_1 (составное), то элементарные события, составляющие множество E_1, он различать не будет, т. е. он зарегистрирует наступление события E_1, как только произойдет любое из них. И, наоборот, наблюдатель отметит непоявление события E_1, если исходом испытания будет распределение, не входящее в множество E_1. То же самое относится и к остальным событиям.

Теперь предположим, что наблюдателя не интересует различие между кодовыми словами сообщения, а интересует только распределение количества кодовых слов по состояниям. В этом случае соглашение об элементарном событии может быть изменено. Например, в качестве пространства элементарных событий могут быть взяты распределения, перечисленные в табл. 1.2.

Здесь событие «контроль по четности обнаруживает искажение в сообщении» изображается подмножеством $\{2, 4, 6, 8, 9, 10\}$.

Пример 1.3. Обозначим через τ время, в течение которого вычислительная машина выполняет программу без сбоя. τ может отсчитываться с любого фиксированного момента

Таблица 1.2

№	Состояние слова			№	Состояние слова			№	Состояние слова		
	1	2	3		1	2	3		1	2	3
1	3	0	0	5	2	0	1	9	0	1	2
2	0	3	0	6	1	2	0	10	1	1	1
3	0	0	3	7	1	0	2				
4	2	1	0	8	0	2	1				

времени, например с момента пуска машины или с момента возникновения очередного сбоя. В качестве пространства элементарных исходов в данном случае можно взять все множество неотрицательных вещественных чисел, т. е. все мыслимые значения величины τ. Таким образом, пространство элементарных событий будет несчетным множеством. Пусть t_1 — некоторое положительное вещественное число. Оно является одноэлементным подмножеством множества положительных вещественных чисел и изображает элементарное событие, которое можно обозначить $\tau = t_1$. Это событие состоит в том, что машина работала до момента времени t_1 без сбоев, а точно в момент времени t_1 возникает сбой. Очевидно, событие $\tau = t_1$ является случайным. Наблюдатель фиксирует появление этого события тогда и только тогда, когда сбой происходит точно в момент времени t_1.

Неравенство $\tau \geqslant t_1$ определяет многоэлементное подмножество пространства элементарных событий, и поэтому его следует рассматривать как составное событие. Смысл этого события: «до момента времени t_1 сбоев в работе машины нет». Наблюдатель зафиксирует появление этого события, если первый сбой наступит в момент времени t_1 или после него, и отсутствие события $\tau \geqslant t_1$, если хотя бы один сбой появится до момента t_1.

Если наблюдателя интересует случайное событие $\tau \geqslant t_1$, а $t_1 < t_2 < t_3$, то события $\tau = t_2$, $\tau = t_3$ или $\tau \geqslant t_3$ он не будет различать в том смысле, что появление любого из них будет отмечать как появление события $\tau \geqslant t_1$.

В реальных условиях точность регистрации времени ограничена. Более того, обычно проверка результатов вычислений не может производиться через интервалы времени, меньшие некоторой величины, например времени выполнения теста. А в некоторых случаях величину τ удобнее измерять количеством выполненных без сбоя команд или программ. Эти ограничения или условия проведения испытания приводят к тому, что пространство элементарных событий превращается в счетное множество.

Точно так же, как в предыдущем примере, в данном испытании все пространство элементарных событий является достоверным событием, а пустое множество — невозможным событием. Это вытекает из предположения, которое оправдывается на практике, что любое испытание закончится сбоем, т. е. τ примет некоторое вещественное значение.

§ 1.2. Поле событий

В каждой вероятностной задаче приходится иметь дело с некоторым классом (множеством) случайных событий, анализируя различные отношения между элементами, т. е. событиями, данного класса. Отношение, в соответствии с которым одному или двум элементам множества сопоставляется один единственный элемент этого же множества, в математике называется алгебраической операцией на данном множестве, соответственно одноместной или двуместной. Такому определению удовлетворяют некоторые отношения между событиями, и поэтому их называют алгебраическими операциями на множестве событий.

Принятая в теории вероятностей точка зрения на события как на подмножества точечного пространства элементарных событий удобна тем, что позволяет отношения между событиями описать в терминах отношений между множествами, т. е. на общепринятом в математике языке.

Вначале напомним основные положения алгебры множеств, а затем интерпретируем их на языке алгебры событий.

Все множества E, E_1, E_2, ..., о которых будет идти речь в последующем, представляют собой события, т. е. подмножества некоторого пространства элементарных событий. Это пространство будем обозначать буквой U, а точки его — строчными буквами латинского алфавита a, b и т. д.

Конечная последовательность n множеств записывается в виде $E_1, E_2, \ldots, E_n$, а счетная — в виде $E_1, E_2, \ldots, E_n, \ldots$ В более общем виде записью $\{E_i : i \in T\}$ обозначают семейство множеств, которым присвоены индексы из множества T. Последнее может быть конечным или счетным.

Если a есть элемент множества E, то это отношение выражают записью $a \in E$ и говорят, что a принадлежит E.

1. Если E_1 и E_2 — такие два множества, что каждая точка из E_1 принадлежит E_2, то E_1 называется *подмножеством* E_2. Это отношение записывается в виде $E_1 \subset E_2$ или $E_2 \supset E_1$.

2. Два множества E_1 и E_2 *равны* в том и только в том случае, если $E_1 \subset E_2$ и $E_2 \subset E_1$. Это отношение записывается в виде $E_1 = E_2$.

3. Пустое множество есть множество, не содержащее никаких точек. Пустое множество будем обозначать буквой V.

Удобно считать, что любое множество $E \subset U$ содержит V в качестве своего подмножества: $V \subset E$.

4. Множество всех тех точек, которые содержатся как в E_1, так и в E_2, называется *пересечением* или *произведением* множеств E_1 и E_2 и обозначается либо $E_1 \cap E_2$, либо E_1E_2.

Очевидно, если E_1 и E_2 не имеют общих точек, то $E_1E_2 = V$. В этом случае говорят, что множества не пересекаются.

Определение операции пересечения множеств можно обобщить на любое семейство множеств. Пересечение совокупности множеств $\{E_i : i \in T\}$ обозначают $\bigcap\limits_{i \in T} E_i$ или просто $\bigcap\limits_i E_i$, если в тексте указано множество значений индекса.

Совокупность множеств $\{E_i : i \in T\}$, такая, что для любых $i, j \in T$ выполняется равенство $E_iE_j = V$, если $i \neq j$, называется совокупностью *попарно непересекающихся* множеств.

5. Множество всех тех точек, которые содержатся по крайней мере в одном из множеств E_1 или E_2, называется объединением E_1 и E_2 и обозначается $E_1 \cup E_2$ или $E_1 + E_2$.

Объединение совокупности множеств $\{E_i : i \in T\}$ обозначается $\bigcup\limits_{i \in T} E_i$ или $\bigcup\limits_i E_i$.

6. Множество, состоящее из всех точек E_1, не содержащихся в E_2, называется *разностью* между E_1 и E_2 и записывается в виде $E_1 - E_2$ или $E_1 \setminus E_2$.

Очевидно, если $E_1 \subset E_2$, то $E_1 - E_2 = V$; если $E_1E_2 = V$, то $E_1 - E_2 = E_1$. Если $E_2 \subset E_1$, то $E_1 - E_2$ называют *собственной разностью* E_1 и E_2 или *дополнением* E_2 относительно E_1. Дополнение E_2 относительно E_1 обозначают $\overline{E_2}(E_1)$. Очень важен случай, когда E_1 совпадает со всем пространством элементарных событий. В таком случае разность $U - E_2$ называют просто дополнением E_2 и обозначают $\overline{E_2}$. Очевидны свойства дополнения:

$$\left.\begin{array}{ll} 1.\ E\overline{E} = V; & 3.\ \overline{U} = V, \\ 2.\ E \cup \overline{E} = U; & 4.\ \overline{\overline{E}} = E \end{array}\right\} \qquad (1.1)$$

и тождество

$$E_1 - E_2 = E_1\overline{E_2}. \qquad (1.2)$$

Для иллюстрации введенных понятий можно воспользоваться геометрическим представлением множеств, состоящих из точек плоскости. Такие представления называют диаграммами Венна. Пусть множество внутренних точек прямоугольника (рис. 1.1) представляет собой пространство U, точки внутри окружности большого радиуса — множество E_1, а внутри окружности малого радиуса — E_2. В таком случае точки внутри контура *1*, *2*, *3*, *4*, *1* образуют $E_1 \cup E_2$; *1*, *6*, *3*, 5, 1 — E_1E_2; 1, 5, 3, 4, 1 — E_1 — E_2; 1, 2, 3, 6, 1 — E_2 — E_1. Точки вне контура 1, 6, 3, 4, 1 образуют $\overline{E}_1$; 1, 2, 3, 5, 1 — $\overline{E}_2$; 1, 6, 3, 5, 1 — $\overline{E_1E_2}$; 1, 2, 3, 4, 1 — $\overline{E_1 \cup E_2}$.

Теперь обратимся к терминологии алгебры событий.

1. Отношение $E_1 \subset E_2$ читается: «событие E_1 влечет E_2» и означает, что наступление события E_1 сопровождается наступлением события E_2, т. е. всякий раз, когда наступает событие E_1, наблюдается и событие E_2. Обратное не всегда верно. В примере 1.2 событие E_3 влечет событие E_2, т. е. $E_3 \subset E_2$, но не наоборот.

Иногда слово «влечет» используют для указания на наличие причинно-следственной связи между явлениями. Следует иметь в виду, что в алгебре событий такое содержание в этот термин не вкладывается.

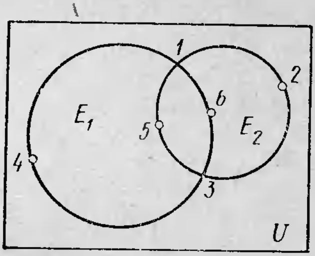

Рис. 1.1

2. События E_1 и E_2 называются *равными* (эквивалентными, равносильными), что записывается как $E_1 = E_2$, в том и только в том случае, если $E_1 \subset E_2$ и $E_2 \subset E_1$. Равенство двух событий означает, что появление любого из них влечет появление другого события. В примере 1.2 событие E_1 — «контроль по четности обнаруживает искажение в сообщении» — равно событию «хотя бы в одном из кодовых слов искаженного сообщения возникло нечетное число трансформаций».

3. $E = V$ означает, что событие E произойти не может, т. е. E — невозможное событие.

4. *Произведением* событий E_1 и E_2 называется событие, которое наблюдается в том и только в том случае, если события E_1 и E_2 наблюдаются совместно. Произведение событий E_1 и E_2 обозначается как $E_1 \cap E_2$ или E_1E_2. События, для которых $E_1E_2 = V$, называются *несовместными*.

В примере 1.2 событие E_5 есть произведение событий E_3 и E_4, т. е. $E_3E_4 = E_5$.

5. *Объединением* или *суммой* событий E_1 и E_2 называется событие, которое наблюдается в том и только в том случае, если наблюдается хотя бы одно из событий E_1 или E_2. Объединение событий E_1 и E_2 обозначается как $E_1 \cup E_2$ или $E_1 + E_2$.

В примере 1.2 событие E_3 есть сумма событий E_5 и E_6, т. е. $E_3 = E_5 \cup E_6$.

6. *Разностью* событий E_1 и E_2 называется событие, которое наступает в том и только в том случае, если наступает

событие E_1 и не наступает событие E_2. Разность событий E_1 и E_2 обозначается как $E_1 - E_2$.

В примере 1.2 событие «искажения в канале при передаче сообщения не возникли» есть $E_2 - E_3$.

7. *Дополнением* события E или событием *противоположным* событию E называется событие, которое наступает всякий раз, когда не наступает событие E. Дополнение события E обозначается как $\overline{E}$.

В примере 1.2 события E_1 и E_2 противоположные: $\overline{E_1} = E_2$ и $\overline{E_2} = E_1$.

8. Очевидно, если $E = V$, то $\overline{E} = U$, т. е. если E — невозможное событие, то $\overline{E}$ — достоверное.

Легко усматриваемая аналогия между двумя группами приведенных определений проявляется в том, что алгебра множеств и алгебра событий имеют одни и те же свойства, поэтому в дальнейшем они не различаются и используется либо та, либо другая терминология в зависимости от удобства.

На основании приведенных определений нетрудно установить справедливость следующих тождеств.

1. $E_1 \cup E_2 = E_2 \cup E_1$; $\qquad E_1E_2 = E_2E_1$.
2. $(E_1 \cup E_2) \cup E_3 = E_1 \cup (E_2 \cup E_3) = E_1 \cup E_2 \cup E_3$;
   $(E_1E_2)E_3 = E_1(E_2E_3) = E_1E_2E_3$.
3. $E_1(E_2 \cup E_3) = E_1E_2 \cup E_1E_3$.
   $E_1 \cup (E_2E_3) = (E_1 \cup E_2)(E_1 \cup E_3)$.

Скобки, как всегда, указывают порядок выполнения операций. В выражениях без скобок с различными операциями по условию сначала выполняется старшая операция, затем — младшая. Принят следующий порядок операций по убыванию старшинства: —, $\cap$, $\cup$, $\supset$. Тождества 1—3 означают, что операции объединения и пересечения коммутативны, ассоциативны и дистрибутивны.

4. $$\left.\begin{aligned} &E_1 \cup E_2 = E_1 \cup (E_2 - E_2E_1); \\ &E_1 \cup E_2 = E_1E_2 \cup (E_1 - E_2E_1) \cup (E_2 - E_2E_1). \end{aligned}\right\} \tag{1.3}$$

Эти тождества означают, что объединение двух любых множеств можно представить в виде объединения непересекающихся множеств.

5. $$E_1 \cup E_2 = \overline{\overline{E}_1\overline{E}_2}; \qquad E_1E_2 = \overline{\overline{E}_1 \cup \overline{E}_2}. \tag{1.4}$$
6. $$\left.\begin{aligned} &E_1E_2 \subset E_1 \quad \text{и} \quad E_1E_2 \subset E_2; \\ &E_1 \cup E_2 \supset E_1 \quad \text{и} \quad E_1 \cup E_2 \supset E_2. \end{aligned}\right\} \tag{1.5}$$

Свойства 4—6 обобщаются на любые совокупности множеств. Для $\{E_i : i \in T\}$:

7. $\bigcap\limits_i E_i = \overline{\bigcup\limits_i \overline{E}_i}$; $\quad \bigcup\limits_i E_i = \overline{\bigcap\limits_i \overline{E}_i}$. (1.6)

8. $\bigcap\limits_i E_i \subset E_j$; $\quad \bigcup\limits_i E_i \supset E_j$, (1.7)

где $j \in T$.

9. Если $E_1, E_2, \ldots, E_n, \ldots$ — конечная или бесконечная последовательность множеств, то множества

E_1;

$E_2 - E_2E_1$;

$E_3 - E_3(E_2 \cup E_1)$;

. . . .

$E_n - E_n(E_{n-1} \cup E_{n-2} \cup \ldots \cup E_1)$

не пересекаются и

$$\bigcup_i E_i = E_1 \cup [E_2 - E_2E_1] \cup [E_3 - E_3(E_2 \cup E_1)] \cup \ldots$$
$$\ldots \cup [E_n - E_n(E_{n-1} \cup E_{n-2} \cup \ldots \cup E_1)] \cup \ldots \quad (1.8)$$

В теории вероятностей в большей мере интересуются не отдельными точками пространства элементарных событий, а определенными классами или множествами подмножеств этого пространства и чаще всего классами, которые называются *полями множеств*.

Непустой класс F подмножеств в U называется *полем* множеств, если он удовлетворяет следующим аксиомам.

П1. Если $E \in F$, то $\overline{E} \in F$;

П2. Если конечная или счетная последовательность $E_1, E_2, \ldots, E_n, \ldots$ принадлежит F, то

$$\bigcup_{i=1}^{\infty} E_i \in F. \quad (1.9)$$

Из этих аксиом вытекают следующие свойства полей.

1. Так как $E \cup \overline{E} = U$, то $U \in F$.

2. Так как $\overline{U} = V$, то $V \in F$.

3. Если $E_1 \in F$ и $E_2 \in F$, то из П1, П2 и (1.4) следует $E_1E_2 \in F$.

Пример 1.4. Возьмем множество всех подмножеств пространства U примера 1.2. Очевидно, это множество удовлетворяет П1 и П2, т. е. является полем.

Пример 1.5. Возьмем в пространстве U примера 1.2 три события E_1, E_2 и E_3. Поле F, содержащее эти события, будет состоять из всех множеств, которые можно получить из E_1,

E_2 и E_3, применяя к ним всеми возможными способами операции из П1 и П2. Нетрудно показать, что таким способом из трех произвольных множеств можно было бы построить 2^8 различных множеств, включая U и V. Но в данном случае выполняются равенства

$$E_1E_2 = E_1E_3 = V, \quad \overline{E}_1 = E_2 \quad \text{и} \quad E_2E_3 = E_3,$$

поэтому получаем поле

$$F = \{U,\ V,\ E_1,\ E_2,\ E_3,\ \overline{E}_3,\ E_2\overline{E}_3,\ \overline{E_2\overline{E}_3}\}.$$

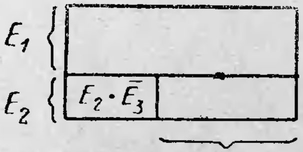

Рис. 1.2

На диаграмме Венна (рис. 1.2) показано взаимное расположение множеств E_1, E_2 и E_3. Все возможные контуры образуют все непустые множества, входящие в поле.

Пример 1.6. Возьмем в пространстве элементарных событий U примера 1.3 все возможные полуоткрытые промежутки $[t_1, t_2)$. Применяя к этому классу операции из П1 и П2, всеми возможными способами (в том числе и счетное число раз) получим поле.

Второй тип класса событий, с которыми часто приходится встречаться в теории вероятностей, носит название *полной группы* событий. Непустой класс $\{E_i : i \in T\}$ событий называется полной группой событий, если $\bigcup_{i \in T} E_i = U$.

Пример 1.7. События E_1, E_2 и E_3 из примера 1.2 составляют полную группу событий. Составляют полную группу и события E_1, E_3 и $E_2\overline{E}_3$, причем эти события попарно несовместны (см. рис. 1.2).

§ 1.3. Классическое и статистическое определения вероятности

Ознакомившись с понятием события и алгебраическими операциями на множестве событий, можно перейти к следующему вопросу — приписыванию событиям вероятностей. Понятие *вероятность события* есть абстракция, которая, как уже говорилось раньше, исходит из идеи об относительной частоте появления события в серии испытаний. Поэтому, прежде чем излагать современную точку зрения, имеет смысл рассмотреть так называемые классическое и статистическое определения вероятности. Хотя для построения логически строгой теории оба этих определения оказались недостаточными, они, с одной стороны, помогают уяснить содержание абстрактного (математического) понятия вероятности, а с другой, необходимы для определения численных значений вероятностей во многих практических ситуациях.

Классическое определение вероятности возникло в связанных с играми задачах и опирается на два условия.

К1. Пространство элементарных событий испытания конечно.

К2. Все элементарные события пространства равновозможны (равновероятны).

Понятие «равновозможность» (равновероятность) при классическом определении вероятности является исходным, неопределяемым. Иногда равновозможность объясняется тем, что события различаются признаками, которые не могут влиять на исход испытания. Например, если при бросании правильной игральной кости, т. е. такой, которая имеет точную форму куба и изготовлена из однородного материала, в качестве пространства элементарных событий взять множества очков $U=\{1, 2, 3, 4, 5, 6\}$, то элементы такого пространства будут равновозможными событиями. Это подтверждается опытом.

Классическое определение вероятности: если пространство элементарных событий U содержит n точек, а событие $E\subset U$ — m точек, то вероятность $P(E)$ события E равна m/n:

$$P(E)=\frac{m}{n}. \tag{1.10}$$

Пример 1.8. Испытание заключается в бросании правильной игральной кости, и в качестве элементарного события взято число выпавших очков. При этих условиях U содержит $n=6$ точек. Событие E_1 — «выпало 2 очка» — имеет вероятность $P(E_1)=1/6$. Событие E_2 — «выпало четное число очков» — представляется множеством $\{2, 4, 6\}$ и поэтому имеет вероятность $P(E_2)=\frac{3}{6}=\frac{1}{2}$.

Пример 1.9. Испытание заключается в том, что три шара a, b, c случайным образом раскладываются по трем ящикам с номерами 1, 2, 3, т. е. трижды случайно выбирается один из трех ящиков, в который помещается очередной шар. Взяв в качестве элементарного события распределение шаров по ящикам, получим пространство элементарных событий, представленное табл. 1.1. Равновозможность элементарных событий, которая является, очевидно, следствием равновозможности выбора ящика для очередного шара, можно обеспечить следующим приемом. Сопоставим каждому ящику двухэлементное множество из числа очков правильной игральной кости так, чтобы множества не имели общих элементов, и будем выбирать ящик в зависимости от числа очков, выпавших при бросании кости.

Поскольку в данном испытании U содержит $n=27$ точек, то вероятность любого элементарного события равна 1/27. Событие A — «второй ящик не пуст» — состоит из $m=19$ точек, так как $A=\{2, 4, 5, 6, 10, 11, 12, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27\}$, поэтому $P(A)=19—27$.

Множество A совпадает с множеством E_1 примера 1.2, однако для утверждения того, что вероятность события E_1 равна 19/27, нет никаких оснований до тех пор, пока мы не убедимся каким-либо образом, что элементарные исходы в испытании примера 1.2 равновозможны.

Из классического определения вероятности следует, что вероятность можно рассматривать как функцию, областью определения которой является множество всех подмножеств пространства элементарных событий, т. е. поле множеств, а областью значений — интервал [0, 1] вещественной оси. Эта функция имеет следующие свойства.

1. Вероятность достоверного события равна единице: $P(U)=1$.

2. Вероятность невозможного события равна нулю: $P(V)=0$.

3. Вероятность события $E\subset U$ находится в интервале [0, 1]: $0\leqslant P(E)\leqslant 1$.

4. Вероятность суммы двух несовместных событий E_1 и E_2 равна сумме вероятностей этих событий: $P(E_1 \cup E_2)=$ $=P(E_1)+P(E_2)$, если $E_1E_2=V$.

Последнее утверждение вытекает из того, что при указанном условии число точек в множестве $E_1 \cup E_2$ равно сумме точек множеств E_1 и E_2. Отсюда

$$P(E_1 \cup E_2)=\frac{m_1+m_2}{n}=\frac{m_1}{n}+\frac{m_2}{n}=P(E_1)+P(E_2),$$

где m_1 — число точек в множестве E_1, а m_2 — в множестве E_2.

Классическое определение вероятности позволяет найти численное значение вероятности события, но оно неудовлетворительно по той причине, что класс испытаний, отвечающих условиям К1 и К2, слишком ограничен. Очень часто встречаются испытания, с которыми невозможно связать какое-либо конечное пространство равновозможных элементарных событий. Типичными испытаниями такого рода являются испытания, упомянутые в примерах 1.1 и 1.2. Однако всякий раз, когда условия К1 и К2 соблюдаются, можно использовать классическое определение вероятности.

Классическое определение вероятности можно распространить на некоторые задачи с бесконечными пространствами элементарных событий, если последние равновозможны. Так возникло понятие геометрической вероятности, применимое к задачам, которые сводятся к следующей формулировке. Пусть имеется область G, а в ней содержится другая область g с квадрируемой границей. В область G бросается наугад точка, и требуется определить вероятность попадания ее в область g. Выражение «бросается наудачу» понимается в том смысле, что брошенная точка может попасть в любую точку области G, а вероятность попадания в каждую часть области G пропорциональна мере (размеру) этой части и не зависит от ее формы и расположения в G.

По определению вероятность попадания в область g при бросании наудачу точки в область G равна

$$P = \frac{\mu(g)}{\mu(G)},$$

где $\mu(S)$ обозначает меру, например длину, площадь или объем области S.

Обратимся теперь к статистическому определению вероятности. Это определение возникло в задачах, связанных с естественно-научными и техническими исследованиями. Именно в задачах такого характера условия К1 и К2, необходимые для классического определения вероятности, очень часто не выполняются. В то же время длительные наблюдения за случайными событиями показывают, что для широкого класса явлений при постоянном комплексе условий S и больших n относительная частота события $W(E) = m/n$, где m — число появлений события E в n независимых испытаниях, почти постоянна, т. е. мало изменяется в различных сериях испытаний. Значительные расхождения между значениями $W(E)$ разных серий испытаний встречаются тем реже, чем больше n. Кроме того, экспериментально было установлено, что для тех случайных событий, к которым применимо классическое определение вероятности, относительная частота события $W(E)$ в сериях испытаний колеблется около вероятности $P(E)$, рассчитанной по классическому определению. В табл. 1.3 приведены результаты экспериментов с бросанием монет [4].

Таблица 1.3

Экспериментатор	Число бросаний	Число выпадений герба	Относительная частота
Бюффон	4 040	2 048	0,5080
К. Пирсон	12 000	6 019	0,5016
К. Пирсон	24 000	12 012	0,5005

Для однородной и правильной формы монеты пространство элементарных событий состоит из двух равновозможных событий — выпадение герба и цифры. Поэтому вероятность выпадения герба в соответствии с классическим определением $P(\Gamma)=1/2$. Как видно из табл. 1.3, относительные частоты в трех сериях испытаний близки к 1/2.

Указанные эмпирические факты послужили основой для выработки статистического определения вероятности. Таким образом, это определение относится только к тем событиям, которые удовлетворяют следующим условиям:

1) можно, по крайней мере принципиально, провести любое число испытаний, в каждом из которых может появиться или не появиться данное событие E;

2) в результате достаточно многочисленных испытаний установлено, что относительная частота события E почти для каждой большой серии испытаний лишь незначительно отклоняется от некоторой (вообще говоря неизвестной) постоянной величины.

По определению эта постоянная величина и является вероятностью случайного события E.

Точное численное значение вероятности чаще всего остается неизвестным, но в качестве ее приближенного значения может быть принята относительная частота события $W(E)$, полученная при большом число испытаний, или какое-либо число, близкое к этой частоте.

Из статистического определения вероятности вытекает следующее.

1. Вероятность достоверного события равна единице, так как $W(U)=1$.

2. Вероятность невозможного события равна нулю, так как $W(V)=0$.

3. Вероятность случайного события E лежит в интервале [0, 1], так как $0 \leqslant W(E) \leqslant 1$.

4. Вероятность суммы двух несовместных событий равна сумме вероятностей этих событий, так как

$$W(E_1 \cup E_2)=\frac{m_1+m_2}{n}=\frac{m_1}{n}+\frac{m_2}{n}=W(E_1)+W(E_2).$$

Здесь m_1 — число появлений события E_1, а m_2 — события E_2 в n испытаниях. Число появлений события $E_1 \cup E_2$ в тех же n испытаниях равно, по определению суммы событий, m_1+m_2, поскольку события несовместимы.

Как видно, статистическое определение вероятности не является формально-логическим (математическим) определением, а по существу представляет собой указание процедуры, которую необходимо выполнить для того, чтобы выяснить, можно ли говорить о вероятности данного события, и оценить ее численное значение. Такого типа определение полезно и необходимо, поскольку оно подчеркивает объективность вероятности случайного события, т. е. независимость этой характеристики явления от наблюдателя и от его эксперимента. Однако попытки построить единую для всех случайных событий теорию на базе этого определения встретились с непреодолимыми трудностями. Отсутствие же такой теории лишает исследователя обоснованных методов расчета неизвестных вероятностей событий, над которыми невозможно провести эксперимент, по известным вероятностям других событий.

§ 1.4. Аксиомы теории вероятностей

Аксиоматическое определение вероятности преодолевает недостатки классического и статистического определений и в то же время учитывает основные свойства вероятности, вытекающие из этих определений и указанные в предыдущем параграфе. Таким образом, аксиоматический подход включает как частные случаи и классический и статистический методы. Подход к построению теории, который впервые был предложен и развит академиком А. Н. Колмогоровым и в настоящее время является общепризнанным, по существу сводится к следующим предположениям:

1) в качестве вероятностей событий в некотором начальном классе относительно простых событий могут быть выбраны числа, принадлежащие интервалу [0, 1];

2) эти начальные вероятности вместе с правилами определения вероятностей более сложных событий должны давать возможность определять вероятность любого события из достаточно широкого класса, чтобы охватить события, представляющие интерес в исходной вероятностной задаче.

Таким образом, в центре внимания теории находятся не численные значения вероятностей, а методы оперирования с вероятностями, которые являются следствием свойств вероятности и, естественно, не должны зависеть ни от численных значений вероятностей, ни от способа их определения (оценки). Математическая теория вероятностей в предложенной А. Н. Колмогоровым формулировке начинается после выбора вероятностей.

Естественно, возникает вопрос, правильно ли выбраны вероятности начальных событий в какой-либо конкретной ситуации. Дать исчерпывающий ответ на этот вопрос — одна из основных задач математической статистики. Как она решается, будет изложено в третьей части настоящего пособия. Здесь же следует отметить, что начальные вероятности во всех случаях выбираются либо на основании классического определения, либо статистического.

Абстрагирование в теории от численных значений каких-то величин не является спецификой теории вероятностей, а, наоборот, обычное явление, обеспечивающее универсальность теории. Так, например, теория электрических цепей на основе уравнений электромагнитного поля дает все соотношения между любыми электрическими параметрами процесса в цепи. Но эта теория не отвечает на вопрос о численных значениях параметров электрического процесса в цепи до тех пор, пока не будут заданы или измерены численные значения некоторых начальных в нашей терминологии параметров процесса. К тому же количество начальных параметров должно быть таким, чтобы обеспечить однозначность решений уравнений электромагнитного поля.

Аналогично механика развивает свой аппарат, не опираясь на численные значения масс и сил и способы их измерения.

При формулировке аксиом теории вероятностей предполагается, что задано некоторое поле событий F и любое событие, упоминаемое в тексте, является элементом поля F, даже если это не оговорено специально. Вероятность определяется следующими аксиомами.

А1. Каждому случайному событию E из поля событий F ставится в соответствие неотрицательное число $P(E)$, называемое его вероятностью.

А2. Если $E_1, E_2, \ldots, E_n, \ldots$ — конечная или счетная последовательность попарно несовместных событий из F, то

$$P(\bigcup_i E_i) = \sum_i P(E_i). \tag{1.11}$$

А3. $P(U) = 1$.

Иногда утверждение, высказанное в А2, формулируют отдельно для конечной последовательности событий и счетной и называют в первом случае *аксиомой сложения*, а во втором — *расширенной аксиомой сложения*.

Следствия из аксиом:

1. Поскольку V и любое событие E несовместны, то, применяя А2 к равенству $E = E \cup V$, получим

$$P(E) = P(E) + P(V).$$

Откуда

$$P(V) = 0, \tag{1.12}$$

т. е. вероятность невозможного события равна нулю.

2. Поскольку $\bar{E}E = V$, то, применяя А2 и А3 к тождеству $E \cup \bar{E} = U$, получим

$$P(E) + P(\bar{E}) = P(U) = 1 \tag{1.13}$$

или

$$P(\bar{E}) = 1 - P(E). \tag{1.14}$$

3. Если $E_1 \subset E_2$, то, применяя А2 к очевидному равенству $E_2 = E_1 \cup \overline{E}_1 E_2$, имеем

$$P(E_2) = P(E_1) + P(\overline{E}_1 E_2).$$

Так как согласно А1 $P(\overline{E}_1 E_2) \geqslant 0$, то

$$P(E_1) \leqslant P(E_2). \tag{1.15}$$

4. Поскольку для любого E $E \subset U$ и $V \subset E$, то, учитывая (1.12), А3 и (1.15), получим

$$0 \leqslant P(E) \leqslant 1. \tag{1.16}$$

5. На основании (1.7) и (1.15) для $\{E_i : i \in T\}$ имеем неравенства:

$$P(\bigcap_i E_i) \leqslant P(E_j); \qquad P(\bigcup_i E_i) \geqslant P(E_j),$$

где $j \in T$.

6. Так как события E_1 и $(E_2 - E_2 E_1)$ несовместны, то, применяя А2 к (1.3), получим

$$P(E_1 \cup E_2) = P(E_1) + P(E_2 - E_2 E_1).$$

Очевидно,

$$E_2 = E_1 E_2 \cup (E_2 - E_2 E_1),$$

поэтому
$$P(E_2) = P(E_1 E_2) + P(E_2 - E_2 E_1)$$

или
$$P(E_2 - E_2 E_1) = P(E_2) - P(E_1 E_2).$$

Отсюда для произвольных E_1 и E_2

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 E_2). \tag{1.17}$$

Применяя А2 к (1.8), получим

$$P(\bigcup_i E_i) = P(E_1) + P(E_2 - E_2 E_1) + P[E_3 - E_3(E_2 \cup E_1)] +$$
$$+ \ldots + P[E_n - E_n(E_{n-1} \cup E_{n-2} \cup \ldots \cup E_1)] + \ldots \tag{1.18}$$

Но

$$P(E_2 - E_2 E_1) = P(E_2) - P(E_1 E_2);$$
$$P[E_3 - E_3(E_2 \cup E_1)] = P(E_3) - P[E_3(E_2 \cup E_1)] =$$
$$= P(E_3) - P(E_3 E_2 \cup E_3 E_1) = P(E_3) - P(E_3 E_2) -$$
$$- P(E_3 E_1) + P(E_1 E_2 E_3) \text{ и т. д.}$$

Используя метод индукции, можно доказать, что

$$P(\bigcup_i E_i) = \sum_{i_1=1}^{\infty} P(E_{i_1}) - \sum_{i_1 i_2} P(E_{i_1} E_{i_2}) + \sum_{i_1 i_2 i_3} P(E_{i_1} E_{i_2} E_{i_3}) +$$
$$+ \ldots + (-1)^{n-1} \sum_{i_1 i_2 \ldots i_n} P(E_{i_1} E_{i_2} \ldots E_{i_n}) + \ldots \tag{1.19}$$

Это равенство называют *теоремой сложения* вероятностей.

Заметим, что

$$[E_2 - E_2E_1] \subset E_2;$$
$$[E_3 - E_3(E_2 \cup E_1)] \subset E_3;$$
$$\cdots\cdots\cdots$$
$$[E_n - E_n(E_{n-1} \cup E_{n-2} \cup \ldots \cup E_1)] \subset E_n;$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

поэтому на основании (1.15) и (1.18)

$$P(\bigcup_i E_i) \leqslant \sum_i P(E_i). \qquad (1.20)$$

Теорема непрерывности. *Если последовательность событий $E_1, E_2, \ldots, E_n, \ldots$ такова, что каждое последующее влечет предыдущее и произведение всех событий E_n есть невозможное событие, то*

$$\lim_{n \to \infty} P(E_n) = 0.$$

Доказательство. Пусть последовательность $E_1, E_2, \ldots, E_n, \ldots$ такова, что $E_1 \supset E_2 \supset E_3 \supset \ldots \supset E_n \supset \ldots$ и при любом $n \geqslant 1$ $\bigcap_{k \geqslant n} E_k = V$.

Очевидно, что

$$E_n = \left[\bigcup_{k=n}^{\infty} E_k\bar{E}_{k+1}\right] \cup \left[\bigcap_{k \geqslant n}^{\infty} E_k\right].$$

Поскольку слагаемые в правой части попарно несовместны, то на основании A2

$$P(E_n) = \sum_{k=n}^{\infty} P(E_k\bar{E}_{k+1}) + P\left(\bigcap_{k \geqslant n}^{\infty} E_k\right),$$

но

$$P\left(\bigcap_{k \geqslant n}^{\infty} E_k\right) = P(V) = 0,$$

поэтому

$$P(E_n) = \sum_{k=n}^{\infty} P(E_k\bar{E}_{k+1}),$$

т. е. $P(E_n)$ есть остаток сходящегося ряда

$$\sum_{k=1}^{\infty} P(E_k\bar{E}_{k+1}) = P(E_1).$$

Следовательно,

$$\lim_{n \to \infty} P(E_n) = 0.$$

Можно показать [4], что утверждения теоремы непрерывности и аксиомы сложения (случай конечной последовательности событий) эквивалентны утверждению расширенной аксиомы сложения. По традиции эту теорему чаще называют аксиомой.

Свойства вероятности, формулируемые А1—А3 и следствием (1.16), позволяют рассматривать вероятность как функцию, областью определения которой является поле F (т. е. как функцию множеств), а областью значений — вещественные числа из промежутка [0, 1]. Эту функцию иногда называют вероятностной мерой на поле F, а тройку $< U, F, P >$, где U — пространство элементарных событий, F — поле, образованное некоторой совокупностью событий из U, и P — вероятностная мера на F, — *вероятностным пространством*. Как уже отмечалось, при построении теории вероятностей на основе аксиом не рассматриваются способы определения или построения вероятностной меры на поле F, а предполагается, что она уже существует, т. е. система аксиом теории вероятностей *неполная*. Неполнота системы аксиом обеспечивает универсальность теории в приложениях.

С функциями точечных множеств, свойства которых совпадают со свойствами вероятностной меры, читатель уже хорошо знаком, но они известны ему под другими наименованиями: длина отрезка, площадь поверхности, объем пространства, масса тела.

Действительно, условившись, например, считать массу какого-либо тела равной единице, мы удовлетворим А3, а физические свойства массы удовлетворяют требованиям остальных аксиом: любое множество E точек тела имеет массу $m(E) \geqslant 0$ (нуль относится к пустому множеству) — А1; если $E_1, E_2, \ldots, E_n, \ldots$ — последовательность непересекающихся множеств из точек тела, то $m\left(\bigcup_i E_i\right) = \sum_i m(E_i)$ — А2.

Любое следствие каких-либо свойств справедливо для каждой функции, обладающей этими свойствами. Этим обстоятельством удобно пользоваться для уяснения существа абстрактных рассуждений и понятий, обращаясь к интерпретации в привычных физических представлениях. Так, вероятность события можно интерпретировать как объем некоторого множества точек пространства или как массу некоторого множества точек тела. Кроме того, существование реальных объектов, обладающих свойствами А1—А3, позволяет утверждать, что эта система аксиом *непротиворечива*.

§ 1.5. Основные теоремы теории вероятностей

Условная вероятность

Пусть E и E_1 — события из поля F, причем $P(E)>0$. *Условной вероятностью* события E_1 при гипотезе E (при условии, что событие E произошло) называется отношение

$$P(E_1/E) = \frac{P(E_1E)}{P(E)}. \qquad (1.21)$$

Это отношение имеет смысл, так как $P(E) \neq 0$. Нетрудно проверить, что $P(E_1/E)$ удовлетворяет аксиомам 1—3 и, следовательно, является вероятностью или вероятностной мерой на F. Действительно,

1) так как $P(E_1/E) \geqslant 0$ и $P(E)>0$, то $P(E_1/E) \geqslant 0$;

2) если $E_1, E_2, \ldots, E_n, \ldots$ — попарно несовместные события из F, то и $(E_1E), (E_2E), \ldots$ — попарно несовместные события из F, поэтому

$$P\left(\bigcup_{i=1}^{\infty} E_i \Big/ E\right) = \frac{P\left[\left(\bigcup\limits_{i=1}^{\infty} E_i\right)E\right]}{P(E)} = \frac{P\left[\bigcup\limits_{i=1}^{\infty}(E_iE)\right]}{P(E)} = \frac{\sum\limits_{i=1}^{\infty} P(E_iE)}{P(E)},$$

или

$$P\left(\bigcup_{i=1}^{\infty} E_i \Big| E\right) = \sum_{i=1}^{\infty} P(E_i/E); \qquad (1.22)$$

3) если $E_1=U$, то

$$P(U/E) = \frac{P(UE)}{P(E)} = \frac{P(E)}{P(E)} = 1. \qquad (1.23)$$

Для выяснения содержания понятия условная вероятность обратимся, например, к статистическому определению вероятности. Поле событий всегда предполагает существование фиксированного комплекса условий S, общего для всех событий из F. Если кроме комплекса S никакие другие условия не учитываются, то вероятности случайных событий из F называются *безусловными* вероятностями и обозначаются символом $P(E)$, как в предыдущем параграфе.

Но иногда кроме основного комплекса условий при рассмотрении вероятностей событий из F приходится учитывать дополнительное условие — наступление некоторого события E из того же поля F. Естественно ожидать, что такое дополнение как-то изменит основной комплекс условий, а следовательно, может привести к изменению вероятностей появления событий из F. Обозначим измененный комплекс условий S'.

Пусть в серии n реализаций комплекса S событие E наблюдалось m раз, а событие E_1E — k раз. Очевидно, что число реализаций комплекса S' равно m и при m испытаниях событие E_1 появилось k раз. Отсюда относительная частота события E_1 при комплексе S' равна k/m. Но

$$\frac{k}{m}=\frac{kn}{mn}=\frac{W(E_1E)}{W(E)}, \tag{1.24}$$

где $W(E_1E)=k/n$ — относительная частота события E_1E при комплексе S; $W(E)=m/n$ — частота события E при комплексе S.

Если в качестве вероятностей событий принять их относительные частоты, то (1.24) превращается в (1.21). Таким образом, переход от безусловных вероятностей событий к условным по формуле (1.21) можно рассматривать как косвенное изменение основного комплекса условий.

Аналогичный результат можно получить и для классического определения вероятности.

Пример 1.10. В примере 1.2 найдем условную вероятность $P(E_1/E_7)$ события E_1 («контроль по четности обнаруживает искажение в сообщении») при условии, что наступило событие $E_7=\{4, 5, 6, 7, 8, 9\}$: «искажение произошло только в одном кодовом слове сообщения». Так как элементарные события несовместны, то

$$P(E_7)=P(4)+P(5)+P(6)+P(7)+P(8)+P(9);$$

$$P(E_1E_7)=P(4, 5, 6)=P(4)+P(5)+P(6).$$

Отсюда

$$P(E_1/E_7)=\frac{P(4)+P(5)+P(6)}{P(4)+P(5)+P(6)+P(7)+P(8)+P(9)}.$$

Пусть каким-либо образом удалось установить, что элементарные события в примере 1.2 равновозможны и, следовательно, их вероятности равны 1/27. В таком случае безусловная вероятность события E_1, согласно классическому определению вероятности,

$$P(E_1)=\frac{19}{27},$$

а условная вероятность

$$P(E_1/E_7)=\frac{3/27}{6/27}=\frac{1}{2}.$$

Теорема умножения вероятностей

Пусть E_1 и E_2 — два события с ненулевыми вероятностями. На основании определения (1.21)

$$P(E_1/E_2)=\frac{P(E_1E_2)}{P(E_2)}; \tag{1.25}$$

$$P(E_2/E_1)=\frac{P(E_1E_2)}{P(E_1)}. \tag{1.26}$$

Эти равенства эквивалентны следующим двум:

$$\left.\begin{aligned} P(E_1E_2) &= P(E_1)\,P(E_2/E_1); \\ P(E_1E_2) &= P(E_2)\,P(E_1/E_2), \end{aligned}\right\} \qquad (1.27)$$

которые оказываются справедливыми и для событий с нулевыми вероятностями. Действительно, если, например, $P(E_1)=0$, то $P(E_1E_2)=0$ и $P(E_1/E_2)=0$. Равенства (1.27) называются **теоремой умножения** вероятностей: *вероятность произведения двух событий равна произведению вероятности одного из этих событий на условную вероятность второго при условии, что первое наступило.*

Эту теорему нетрудно обобщить на случай любого конечного числа событий

$$P(E_1E_2\ldots E_n) = P(E_1)\,P(E_2/E_1)\,P(E_3/E_1E_2)\ldots$$
$$\ldots P(E_n/E_1E_2\ldots E_{n-1}). \qquad (1.28)$$

Здесь $P(E_i/E_1E_2\ldots E_{i-1})$, $i=1, 3, \ldots, n$ — условная вероятность события E_i при гипотезе $E_1E_2\ldots E_{i-1}$.

Независимость событий

Говорят, что событие E_1 *независимо* (статистически) от события E_2, если имеет место равенство

$$P(E_1/E_2) = P(E_1). \qquad (1.29)$$

Это равенство означает, что если E_1 независимо от E_2, то вероятность события E_1 не меняется от того, что наступает событие E_2.

Из (1.27) и (1.29) следует, что $P(E_2/E_1)=P(E_2)$, т. е., если событие E_1 независимо от E_2, то и событие E_2 независимо от E_1. Свойство независимости событий взаимно.

Так как очевидно, что

$$P(E_1/E_2) + P(\overline{E}_1/E_2) = 1,$$

то при выполнении условия (1.29)

$$P(E_1) + P(\overline{E}_1/E_2) = 1,$$

$$P(\overline{E}_1/E_2) = 1 - P(E_1) = P(\overline{E}_1).$$

Поэтому если события E_1 и E_2 независимы, то независимы и события $\overline{E}_1$ и E_2, а следовательно, E_1 и $\overline{E}_2$, $\overline{E}_1$ и $\overline{E}_2$ также независимы.

Для двух независимых событий E_1 и E_2 теорема умножения вероятностей принимает простую форму

$$P(E_1E_2) = P(E_1)\,P(E_2). \qquad (1.30)$$

События $E_1, E_2, \ldots, E_n$ называются независимыми в совокупности, если для любого α и произвольных $\beta_1, \beta_2, \ldots, \beta_k$ из множества $\{1, 2, \ldots, n\}$, $\alpha \neq \beta_i$ события E_α и $(E_{\beta_1} E_{\beta_2} \ldots E_{\beta_k})$ взаимно независимы.

Для независимых в совокупности событий

$$P\left(\bigcap_{i=1}^{n} E_i\right) = \prod_{i=1}^{n} P(E_i). \qquad (1.31)$$

Формула полной вероятности

Если $E_1, E_2, \ldots, E_n, \ldots$ — полная группа попарно несовместных событий (конечная или счетная), имеющих ненулевые вероятности, и E — произвольное событие, то

$$P(E) = \sum_i P(E_i) P(E/E_i). \qquad (1.32)$$

Действительно, при указанных условиях

$$\bigcup_i E_i = U \quad \text{и} \quad (EE_i)(EE_j) = V, \quad i \neq j.$$

А так как

$$E = EU = E(\bigcup_i E_i) = \bigcup_i (EE_i),$$

то

$$P(E) = P[\bigcup_i (EE_i)] = \sum_i P(EE_i) = \sum_i P(E_i) P(E/E_i).$$

Равенство (1.32) называется *формулой полной вероятности.* Эта формула играет большую роль при определении неизвестных вероятностей сложных событий.

Формулы для вероятности гипотез (формулы Бейеса)

Пусть событие E может наступить только совместно с одним из случайных событий $E_1, E_2, \ldots, E_n$, составляющих полную группу и попарно несовместных. Другими словами, выполняются условия предыдущего пункта.

Предположим, что до испытания нам по каким-либо причинам известны вероятности $P(E_1), P(E_2), \ldots, P(E_n)$, которые называются иногда априорными, и условные вероятности $P(E/E_1), P(E/E_2), \ldots, P(E/E_n)$, а испытание завершилось событием E. Чему равна вероятность того, что справедлива гипотеза: «событие E произошло совместно с событием E_i, $(i=1, 2, \ldots, n)$»? — или, другими словами, чему равна условная вероятность события E_i, $i=\overline{1,n}$ при условии, что наступило событие E.

В подобных задачах сами события $E_1, E_2, \ldots, E_n$ обычно называют гипотезами, а их условные вероятности $P(E_i/E)$,

$i=\overline{1,n}$ — вероятностями гипотез или апостериорными вероятностями событий E_i, $i=\overline{1,n}$.

По теореме умножения вероятностей

$$P(E_iE)=P(E)\,P(E_i/E)=P(E_i)\,P(E/E_i).$$

Отсюда

$$P(E_i/E)=\frac{P(E_i)\,P(E/E_i)}{P(E)}.$$

Используя формулу полной вероятности (1.32) для $P(E)$, получаем окончательно

$$P(E_i/E)=\frac{P(E_i)\,P(E/E_i)}{\sum\limits_{i=1}^{n}P(E_i)P(E/E_i)}, \qquad (1.33)$$

где $i=1, 2, \ldots, n$.

Выражения (1.33) называются *формулами вероятностей гипотез* или формулами Бейеса.

§ 1.6. Элементы комбинаторного анализа

Два ограничения — конечность пространства элементарных событий и равновозможность их препятствуют построению универсальной теории на основе классического определения вероятностей. Тем не менее практические задачи, в которых эти условия выполняются, встречаются достаточно часто. В некоторых случаях равновозможность элементарных событий приходится принимать в качестве гипотезы и после теоретического рассмотрения задачи проверять эту гипотезу статистическими методами. Конечность пространства элементарных событий — естественное условие для большинства практических задач. Так, например, конечно число элементов в аппаратуре, число разрядов в коде или команде вычислительной машины, конечно число сигналов в передаваемом по каналу связи сообщений и т. д.

Классический подход удобен тем, что позволяет без эксперимента получить численное значение вероятности события. Определением вероятности (1.10) этот расчет сводится к подсчету числа элементов в конечных множествах — в пространстве элементарных событий и в некотором заданном подмножестве этого пространства. Подмножество задается, как правило, словесной формулировкой существа события, т. е. словесная формулировка указывает некоторое свойство точек, которое позволяет выделить их из множества остальных точек. И задача определения численного значения вероятности состоит в том, чтобы по словесной формулировке события найти способ подсчета точек соответствующего подмножества. Эту процедуру иногда называют «определением числа исходов, благоприятствующих заданному событию».

Систематические методы подсчета числа элементов конечных множеств составляют предмет рассмотрения комбинаторного анализа.

В этом параграфе мы рассмотрим наиболее простой и чаще всего используемый аппарат и познакомимся с соответствующей терминологией. Для более обстоятельного изучения комбинаторного анализа можно обратиться к [7].

Все понятия и практические приемы можно определить на языке множеств, однако используемая в комбинаторике терминология имеет традиционную специфику. В частности, испытание трактуется как случайный выбор элемента из некоторого множества, которое называют *генеральной совокупностью*. Определение «случайный» понимается как «равновозможный» выбор любого элемента, т. е. при выборе ни один элемент генеральной совокупности не имеет никаких преимуществ по сравнению с другими элементами. Множество, которое образуют элементы, полученные повторением r раз выбора из генеральной совокупности, называется *выборкой объема* r.

Различают выборки без повторения и с повторением элементов. Выборкой без повторения элементов называют такую выборку, в которую каждый ее элемент входит только один раз, т. е. все элементы выборки различимы между собой. Выборка без повторений извлекается из генеральной совокупности, состоящей из различающихся элементов, если очередной выбранный элемент исключается из нее и не участвует в последующих выборках. Выбор при таком условии называют выбором без возвращения.

Выборкой объема r с повторениями называют такую выборку, в которой каждый ее элемент может входить любое число раз до r включительно, т. е. в которой могут содержаться неразличимые элементы. Такая выборка получается при многократном выборе, если каждый очередной выбор происходит из одной и той же неизменной по составу генеральной совокупности. Это условие выполняется, если генеральную совокупность из элементов трактовать как бесконечное множество, в котором каждый из элементов представлен бесконечное число раз, либо считать, что каждый выбранный элемент перед следующим выбором возвращается в генеральную совокупность или заменяется идентичным. В связи с последней трактовкой выбор при таком условии называют выбором с возвращением. Очевидно, при выборе с возвращением число r может быть больше n.

В некоторых задачах важен только состав выборки, т. е. набор элементов, составляющих выборку, в других требуется учитывать и состав и порядок расположения элементов в выборке.

Определение 1. *Перестановкой r из n называется упорядоченная выборка объема r из генеральной совокупности n элементов.*

Определение 2. *Сочетанием r из n называется выборка объема r из генеральной совокупности n элементов без учета порядка.*

Из этих определений следует, что две перестановки r из n равны (неразличимы), если у них одинаковы состав и расположение элементов. Для равенства (неразличимости) двух r сочетаний достаточно совпадения составов.

Пример 1.11. Пусть генеральная совокупность состоит из цифр 1, 2, 3, 4 ($n=4$).

Перестановки 2 из 4 с повторением образуют множество

$$A=\left\{\begin{matrix} 11, & 12, & 13, & 14, & 21, & 22, & 23, & 24 \\ 31, & 32, & 33, & 34, & 41, & 42, & 43, & 44 \end{matrix}\right\};$$

перестановки 2 из 4 без повторения образуют множество

$$B=\left\{\begin{matrix} 12, & 13, & 14, & 21, & 23, & 24 \\ 21, & 31, & 41, & 12, & 32, & 42 \end{matrix}\right\};$$

сочетания 2 из 4 образуют множество

$$C=\{12,\ 13,\ 14,\ 23,\ 24,\ 34\}.$$

Различные комбинаторные формулы для вычисления числа элементов множеств основываются чаще всего на двух следующих правилах.

Правило суммы. Если элемент a может быть выбран m способами, а элемент b — другими n способами, то выбор «либо a, либо b» может быть осуществлен $n+m$ способами (союз либо, как всегда, предполагает исключение: выбор a исключает выбор b и наоборот).

Правило произведения. Если элемент a может быть выбран m способами и после каждого из таких выборов элемент b в свою очередь может быть выбран n способами, то выбор пары «a и b» в указанном порядке может быть осуществлен mn способами.

Справедливость этих правил становится совершенно очевидной, если перейти на язык множеств. Сформулированные в правилах условия выполняются, когда $a\in A$ и $b\in B$, причем

$$A=\{a_1, a_2\ \ldots, a_m\};$$

$$B=\{b_1, b_2, \ldots, b_n\}.$$

Первое правило есть просто утверждение, что $A \cup B$ содержат $m+n$ элементов, если $AB=V$. Второе правило утверждает, что множество всех возможных пар (a, b), в которых первый член — элемент A, а второй — B, содержит mn элементов (такое множество называют декартовым произведением множеств A и B и обозначают $A \times B$). Действительно, с каждым элементом множества A можно построить n различающихся вторыми членами пар, а так как A содержит m элементов, то общее число пар равно mn.

Оба правила просто распространяются на любое конечное число комбинируемых элементов. Например, если a_1 можно выбрать n_1 способами, a_2 — n_2 способами, ..., a_k — n_k способами, то комбинация элементов $(a_1, a_2, \dots, a_k)$ может быть выбрана в указанном порядке $n_1 n_2 \dots n_k$ способами.

Перестановки

Для получения любой перестановки r из n нужно произвести r выборов. При построении перестановки без повторений первый элемент выбирается из n возможных, второй — из $(n-1)$ и т. д., последний — из $(n-r+1)$. Правило произведения дает число различных перестановок r из n без повторения:

$$(n)_r = n(n-1)(n-2)\ .\ .\ .\ (n-r+1); \qquad r \leqslant n. \tag{1.34}$$

При $r=n$

$$(n)_n = n! \tag{1.35}$$

Число перестановок r из n с повторением определяется, очевидно, формулой

$$u(n, r) = n^r, \tag{1.36}$$

так как каждый элемент выбирается из n возможных.

Множество всех перестановок r из n можно разбить на два непересекающихся подмножества — подмножество перестановок r, не содержащих некоторый фиксированный элемент генеральной совокупности, и содержащих этот элемент. Очевидно, число элементов первого подмножества $(n-1)_r$. Число элементов второго подмножества подсчитывается следующим образом. Если фиксированный элемент входит в перестановку, то он занимает в ней одно из r мест. Число перестановок $(r-1)$ из $(n-1)$ элементов, не содержащих фиксированный элемент, равно $(n-1)_{r-1}$. Применяя правило произведения к двум выборам — выбор места и выбор перестановки $(r-1)$ из $(n-1)$, получим искомое число $r(n-1)_{r-1}$. Таким образом, используя правило суммы, приходим к следующему рекуррентному соотношению:

$$(n)_r = (n-1)_r + r(n-1)_{r-1}. \tag{1.37}$$

Сочетания

Взяв любое сочетание r из n в качестве новой генеральной совокупности, можно получить $r!$ различных перестановок r из r. Другими словами, любое r сочетание из n можно упорядочить $r!$ различными способами. Если число r сочетаний из n обозначим символом C_n^r (используется также обозначение $\binom{n}{r}$ и др.), то

$$(n)_r = r!\ C_n^r, \tag{1.38}$$

поскольку каждую перестановку r из n можно получить двумя выборами: вначале выбрать сочетание r из n, а затем — порядок расположения элементов в этом сочетании r. Отсюда

$$C_n^r = \frac{(n)_r}{r!} = \frac{n(n-1)\ .\ .\ .\ (n-r+1)}{r!} = \frac{n!}{r!\,(n-r)!}. \tag{1.39}$$

Числа C_n^r называются биномиальными коэффициентами, так как они появляются в разложении бинома $(a+b)^n$.

На основании (1.39) получим

$$C_n^r = \frac{n!}{r!\,(n-r)!} = \frac{n(n-1)\ .\ .\ .\ (r+1)}{(n-r)!}$$

или

$$C_n^r = C_n^{n-r}. \tag{1.40}$$

Уравнение (1.40) равносильно утверждению, что выбор r элементов из n однозначно определяет остальные $n-r$. Формула (1.39) имеет смысл для $0<r\leqslant n$, однако ее распространяют и на случай $r=0$, полагая формально, что

$$0! = 1; \quad (n)_0 = 1 \quad \text{и} \quad C_n^0 = 1, \tag{1.41}$$

хотя эти равенства не имеют никакого комбинаторного смысла. Аналогично не имеет комбинаторного смысла число C_{-n}^r, но формально, согласно (1.39),

$$C_{-n}^r = \frac{(-n)(-n-1)\ .\ .\ .\ (-n-r+1)}{r!} =$$

$$= (-1)^r \frac{n(n+1)\ .\ .\ .\ (n+r-1)}{r!} = (-1)^r C_{n+r-1}^r. \tag{1.42}$$

Зафиксируем некоторый элемент генеральной совокупности и разобьем все множество сочетаний r из n на два непересекающихся подмножества таким образом, чтобы одно подмножество составляли сочетания r, содержащие фиксированный элемент, а другое — не содержащие его. Число элементов первого подмножества C_{n-1}^{r-1}, так как все сочетания r из n с одним одинаковым элементом отличаются остальными $r-1$ элементами, выбранными из $n-1$ возможных.

Число элементов во втором подмножестве C_{n-1}^{r}. Отсюда получим важное рекуррентное соотношение

$$C_n^r = C_{n-1}^{r-1} + C_{n-1}^{r}. \qquad (1.43)$$

Пользуясь этим соотношением, легко составить таблицу типа табл. 1.4 для чисел C_n^r. Табл. 1.4 называется треугольником Паскаля.

Таблица 1.4

n	r					
	0	1	2	3	4	5. . . . . .
0	1					
1	1	1				
2	1	2	1			
3	1	3	3	1		
4	1	4	6	4	1	
5	1	5	10	10	5	
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.

Размещения

Очень часто к анализу самых разнообразных явлений может быть применена так называемая урновая модель — случайное размещение r шаров по n урнам. В практических задачах шарами и урнами можно назвать различные объекты, например сбои вычислительной машины (шары) и ее функциональные узлы (урны), сообщения, поступившие в центр сбора информации (шары), и корреспонденты (урны) и т. д.

Каждое размещение r различимых шаров в n различимых урнах характеризуется:

1) набором чисел $\{r_1, r_2, \ldots, r_n\}$, в котором r_i — число шаров в урне с номером i, $i=\overline{1,n}$,

$$r_1 + r_2 + \ldots + r_n = r; \qquad (1.44)$$

2) порядком чисел в наборе $\{r_1, r_2, \ldots, r_n\}$;

3) составом множеств $A_1, A_2, \ldots, A_n$, где A_i — множество шаров в урне с номером i, $i=\overline{1,n}$.

Число размещений, различающихся по всем трем признакам, равно n^r, поскольку для размещения r шаров по урнам для каждого из них нужно выбрать одну из n урн, т. е. образовать перестановку r из n с повторением.

Число размещений с фиксированным набором чисел $\{r_1, r_2, \ldots, r_n\}$ равно числу способов, которыми r элементов могут быть разделены на n подмножеств так, чтобы первое содержало r_1 элементов, второе — r_2 и т. д. Это число определяется формулой

$$\binom{r}{r_1, r_2, \ldots, r_n} = \frac{r!}{r_1!\, r_2! \ldots r_n!}. \quad (1.45)$$

Действительно, способ подобного разделения реализуется последовательностью n выборов. Вначале выбирается перестановка r_1 из r, затем перестановка r_2 из $(r-r_1)$, затем перестановка r_3 из $(r-r_1-r_2)$ и т. д. Последней выбирается перестановка r_n из $r-r_1-r_2-\ldots-r_{n-1}$. Применяя правило умножения, получим искомое число

$$\binom{r}{r_1, r_2, \ldots, r_n} = \frac{r!}{r_1!\,(r-r_1)!} \cdot \frac{(r-r_1)!}{r_2!\,(r-r_1-r_2)!} \cdots$$

$$\cdots \frac{(r-r_1-r_2-\ldots-r_{n-1})!}{r_n! \cdot 0!} = \frac{r!}{r_1!\, r_2! \ldots r_n!}.$$

Числа (1.45) называются полиномиальными коэффициентами, так как они появляются в разложении полинома $(a_1+a_2+\ldots+a_n)^r$. Формула (1.39) является частным случаем (1.45).

Если считать, что шары неразличимы, то размещения будут различаться только признаками 1 и 2. В этом случае число различных размещений r шаров по n урнам, т. е. число различных решений уравнений (1.44), определяется формулой

$$A(r, n) = C^r_{n+r-1} = C^{n-1}_{n+r-1}, \quad (1.46)$$

причем C^{n-1}_{r-1} размещений обладают тем свойством, что ни одна из урн не пуста.

Для доказательства (1.46) будем изображать урны промежутками между вертикальными черточками, а шары — звездочками. Так, символ |***|**||*|**| обозначает размещение $r=8$ шаров в 5 урнах, причем урны содержат 3, 2, 0, 1, 2 шара соответственно. Для n урн число черточек равно $n+1$, причем положение двух черточек в каждом распределении фиксировано, а остальные $n-1$ располагаются между звездочками. Между крайними черточками имеется $r+n-1$ мест, на каждом из которых может находиться звездочка или черточка. Размещение полностью определится, если выбрать r мест для звездочек или $n-1$ мест для черточек из $n+r-1$ возможных. Таким образом, приходим к формуле (1.46).

В размещении не будет пустых урн, если каждая черточка заключена между двумя звездочками. Так как число промежутков между звездочками $r-1$, а свободных черточек $n-1$, то число размещений без пустых урн равно числу $(n-1)$ сочетаний из $(r-1)$, т. е. C^{n-1}_{r-1}.

Пример 1.12. В генеральной совокупности, состоящей из n шаров, имеется n_1 красных шаров и n_2 черных, $n=n_1+n_2$. Найти вероятность P_k события A_k: «выборка объема r содержит ровно k красных шаров, $k \leqslant \min\{n_1, r\}$». Здесь, очевидно, порядок расположения шаров в выборке роли не играет.

Число точек в пространстве элементарных событий равно числу r сочетаний из n, т. е. C_n^r. Выборку объема r с k красными и $r-k$ черными шарами можно получить двумя способами: вначале выбрать k красных шаров из n_1 возможных, затем $(r-k)$ черных из n_2. Поэтому число исходов, благоприятствующих событию A_k, равно

$$C_{n_1}^k C_{n_2}^{r-k},$$

а вероятность этого события

$$p_k = \frac{C_{n_1}^k C_{n_2}^{r-k}}{C_n^r}. \qquad (1.47)$$

Совокупность чисел p_k, $k=\overline{0, n_1}$, называется гипергеометрическим распределением.

Производящие функции числа сочетаний и перестановок

Для выполнения подсчетов комбинаторного характера удобно использовать несложный формальный аппарат, называемый методом производящих функций.

Пусть генеральная совокупность состоит из трех элементов x_1, x_2, x_3. Образуем произведение

$$(1+x_1t)(1+x_2t)(1+x_3t). \qquad (1.48)$$

Раскрыв скобки и расположив слагаемые в порядке возрастания степеней t, получим

$$1 + a_1t + a_2t^2 + a_3t^3 =$$

$$= 1 + (x_1 + x_2 + x_3)t + (x_1x_2 + x_1x_3 + x_2x_3)t^2 + x_1x_2x_3t^3.$$

Нетрудно заметить, что число слагаемых каждого коэффициента a_r, $r=1, 2, 3$ равно числу сочетаний из трех по r. Следовательно, число таких сочетаний будет равно a_r, если все x_k приравнять единице, т. е.

$$(1+t)^3 = \sum_{r=0}^{3} a_r t^r = \sum_{r=0}^{3} C_3^r t^r.$$

Для генеральной совокупности из n различных объектов, обозначенных символами $x_1, x_2, \ldots, x_n$, имеем при $x_k=1, k=\overline{1, n}$

$$(1+t)^n = \sum_{r=0}^{n} a_r t^r = \sum_{r=0}^{n} C_n^r t^r, \qquad (1.49)$$

т. е. коэффициентом при t^r, $r=\overline{0, n}$ в разложении $(1+t)^n$ стоит число C_n^r, о чем уже упоминалось раньше. По этой причине выражение $(1+t)^n$ называют производящей (перечисляющей) функцией числа сочетаний из n различных элементов (без повторения).

Эффективность метода производящих функций продемонстрируем на доказательстве равенства

$$C_n^r = \sum_{k=0}^{r} C_{n-m}^k C_m^{r-k}. \tag{1.50}$$

Поскольку для любого m, $0<m<n$

$$(1+t)^n = (1+t)^{n-m}(1+t)^m,$$

то, приравняв коэффициенты при t^r в левой и правой частях этого равенства, получим из (1.50)

$$C_n^r = C_{n-m}^0 C_m^r + C_{n-m}^1 C_m^{r-1} + \ldots + C_{n-m}^r C_m^0.$$

Обратимся к сумме, которая получается после раскрытия скобок в произведении

$$(1+x_1 t)(1+x_2 t)\ldots(1+x_n t), \tag{1.51}$$

если не группировать слагаемые с равными степенями t. Каждое слагаемое этой суммы образуется (согласно дистрибутивному закону умножения относительно сложения) перемножением элементов, выбираемых по одному из каждого множителя $(1+x_k t)$, $k=\overline{1,n}$. Коэффициент у t^r в каком-либо слагаемом, содержащем t^r, образуется выбором первого члена (единицы) в $(n-r)$ множителях и второго члена в оставшихся r множителях. Поэтому коэффициент у t^r в этом слагаемом соответствует определенному r сочетанию из n. Здесь речь идет именно о сочетаниях, поскольку в связи с коммутативностью умножения произведения $x_i x_j$ и $x_j x_i$ не различаются, т. е. не различается порядок расположения элементов. В коэффициент при t^r каждый из элементов x_k, $k=\overline{1,n}$ либо не входит, либо входит, но не более одного раза. Такая конструкция коэффициента определяется биномиальной формой множителей (1.51) и тем, что первое слагаемое в них равно единице.

Сопоставление коэффициентов при t^r с r сочетаниями из n указывает путь перехода от производящей функции числа сочетаний без повторения к производящей функции числа сочетаний с повторениями, которые образуются при выборке с возвращениями либо если в генеральной совокупности некоторые элементы представлены несколькими экземплярами.

Пусть, например, элемент x_k в r сочетании из n может появляться любое число раз от 0 до s включительно. Чтобы получить производящую функцию для числа сочетаний из n с повторениями элемента x_k, необходимо изменить k-й множитель в (1.51) так, чтобы после раскрытия скобок в сумме появились слагаемые, содержащие t^r с коэффициентами, соответствующими искомым сочетаниям. Коэффициент при t^r будет соответствовать r сочетанию с i повторениями элемента x_k, $i=\overline{0,s}$, если в этот коэффициент входят сомножители x_k^i и $r-i$ других символов x_j, $j \neq k$.

Но такой член в сумме может появиться, если k-й множитель в (1.51) будет содержать слагаемое $x_k^i t^i$. Таким образом, производящая функция числа сочетаний из n с повторением элемента x_k 0, 1, ..., s раз получится из (1.51), если мы множитель $1+x_k t$ заменим множителем

$$1 + x_k t + x_k^2 t^2 + x_k^3 t^3 + \ldots + x_k^s t^s$$

и все x_k, $k=\overline{1,n}$ приравняем единице.

Множители производящей функции можно совершенно независимо друг от друга приспосабливать к требованиям задачи. Например, если x_k должен всегда входить в выборку четное число раз, но не более $2s$, то k-й множитель в производящую функцию должен входить в виде

$$1 + x_k^2 t^2 + x_k^4 t^4 + \ldots + x_k^{2s} t^{2s}.$$

В качестве примера найдем число r сочетаний из n различных элементов, каждый из которых может появиться любое число раз от 0 до r. Производящей функцией для сочетаний с неограниченным повторением n различных элементов без ограничения на число появления любого из них будет

$$(1+t+t^2+\ldots)^n = (1-t)^{-n}.$$

Здесь можно считать, что $|t|<1$.

Применяя к этому выражению формулу разложения бинома и учитывая (1.42), получим:

$$(1-t)^{-n} = \sum_{r=0}^{\infty} C_{-n}^{r}(-t)^r = \sum_{r=0}^{\infty} C_{n+r-1}^{r} t^r.$$

Таким образом, искомое число сочетаний выражается формулой

$$f(n, r) = C_{n+r-1}^{r}. \tag{1.52}$$

Пример 1.13. Сочетания 2 из четырех элементов 1, 2, 3, 4 с повторением образуют множество {11, 12, 13, 14, 22, 23, 24, 33, 34, 44}.

По формуле (1.52) получим

$$f(4, 2) = \frac{5!}{2!\,3!} = 10.$$

Разложение (1.49) можно представить в иной форме, если учесть (1.39):

$$(1+t)^n = \sum_{r=0}^{n} (n)_r \frac{t^r}{r!}. \tag{1.53}$$

При такой записи коэффициентом при $t^r/r!$ в разложении $(1+t)^n$ стоит число перестановок r из n. По этой причине бином $(1+t)^n$ можно назвать также производящей функцией перестановок r из n.

Эта производящая функция дает числа перестановок r без повторения элементов в перестановках. Рассуждения, аналогичные предыдущим, указывают путь обобщения на случай перестановок с повторением элементов.

Пусть генеральная совокупность состоит из n различных элементов $x_1, x_2, \ldots, x_n$, причем x_1 представлен s_1 экземплярами, $x_2 - s_2$ экземплярами и т. д., $x_n - s_n$ экземплярами, $s_1 + s_2 + \ldots + s_n = m$.

Составим произведение

$$\prod_{i=1}^{n}\left(1 + x_i\frac{t}{1!} + x_i^2\frac{t^2}{2!} + \ldots + x_i^{s_i}\frac{t^{s_i}}{s_i!}\right) \tag{1.54}$$

и раскроем скобки. Произведение $x_1^{r_1}x_2^{r_2}\ldots x_n^{r_n}$, $r_1 + r_2 + \ldots + r_n = r$, входящее сомножителем в одно из слагаемых полученной суммы, соответствует перестановке r, в которую входит r_1 экземпляров элемента x_1, r_2 экземпляров элемента x_2 и т. д., r_n экземпляров элемента x_n. Коэффициент при этом произведении вычисляется следующим образом:

$$\frac{t^{r_1}}{r_1!}\frac{t^{r_2}}{r_2!}\ldots\frac{t^{r_n}}{r_n!} = \frac{t^r}{r_1!\,r_2!\ldots r_n!} = \frac{r!}{r_1!\,r_2!\ldots r_n!}\frac{t^r}{r!}. \tag{1.55}$$

Положив в (1.54) $x_i = 1$, $i=\overline{1, n}$, получим производящую функцию чисел перестановок, в которых элемент x_1 может появиться до s_1 раз, x_2 — до s_2 раз и т. д., x_n до s_n раз. Коэффициент при $t^r/r!$, как видно из (1.55), совпадает с (1.45), т. е. число перестановок r из n элементов различных типов с повторениями равно числу размещений r различных элементов по

n различным урнам. Действительно, перестановку, в которую входит r_1 элементов типа x_1, r_2 элементов типа x_2 и т. д., r_n элементов типа x_n, можно трактовать как такое размещение r различных элементов по n урнам, при котором урна x_1 содержит r_1 элементов, урна x_2 — r_2 элементов и т. д., урна x_n — r_n элементов.

Производящие функции с разложениями вида (1.53) называются экспоненциальными производящими функциями, что оправдывается аналогией с рядом

$$\exp at = \sum_{r=0}^{\infty} a^r \frac{t^r}{r!}.$$

Пример 1.14. Для перестановок r из n различных элементов с неограниченными повторениями, т. е. любой элемент может появляться в перестановке произвольное число раз, производящей функцией является

$$\left(1 + \frac{t}{1!} + \frac{t^2}{2!} + \ldots\right)^n,$$

но

$$\left(1 + \frac{t}{1!} + \frac{t^2}{2!} + \ldots\right)^n = \exp nt = \sum_{r=0}^{\infty} n^r \frac{t^r}{r!},$$

т. е. число перестановок r из n различных элементов с неограниченными повторениями равно n^r, что совпадает с (1.36).

ГЛАВА 2

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

§ 2.1. Случайные величины

Явлениям природы свойственна качественная и количественная определенность, а основная задача всякого исследования в области точных наук состоит в установлении количественных соотношений, выражающих объективные законы природы. Это замечание в полной мере относится и к случайным событиям. Исследователя обычно интересует не только факт появления какого-то события, но и численные значения тех физических величин, которые связаны с этим событием. А чаще всего сам факт появления случайного события устанавливается потому, что некоторая физическая величина приняла определенное значение или ряд физических величин приняли ряд определенных значений.

Связанные со случайными событиями величины при испытании принимают конкретные численные значения, но в связи со случайностью событий заранее предсказать эти значения невозможно, поэтому такие величины называются случайными. Случайной величиной является длительность выполнения машиной программы вычислений, если последняя содержит циклы с переменным числом повторений, количество ошибок, возникающих в канале связи при передаче сообщения, отклонение истинного сопротивления резистора от указанного номинала.

Случайные величины очень разнообразны. Множество возможных значений случайной величины может быть конечным, счетным или несчетным. Для характеристики случайной величины недостаточно описать это множество. Необходимо каким-либо способом указывать вероятности появления при испытаниях конкретных численных значений ее. Конечно, желательно иметь единый для всех разнообразных по своей природе случайных величин способ. С этой целью в теории вероятностей введено понятие *функции распределения* случайной величины.

Соотношение между элементарным событием $e \in U$ и численным значением какой-либо характеристики этого события можно рассматривать как однозначную функцию, заданную на множестве U со значениями в множестве вещественных чисел $R: x = \varphi(e)$.

Возьмем в множестве вещественных чисел некоторый промежуток $A \subset R$. Может оказаться, что промежутку A соответствует в пространстве U такое подмножество B, что как только $\varphi(e) \in A$, то $e \in B$. В этом случае множество A можно рассматривать как своего рода определение события B и считать, что $P(B) = P(A)$. Если A есть бесконечный полуоткрытый промежуток $[-\infty, x)$, то он, как и соответствующее ему множество $B \subset U$, однозначно определяется числом x и вероятность события B есть функция x:

$$P(B) = P(A) = P\left(\varphi(e) \in A\right) = F(x).$$

Эта функция независимой переменной x получила название функции распределения случайной величины $\varphi(e)$.

Условимся в дальнейшем случайные величины обозначать прописными буквами конца латинского алфавита, их частные значения — строчными буквами начала алфавита, а «текущие» точки множества вещественных чисел (т. е. аргументы) — строчными буквами конца алфавита. Так, запись $X = a$ означает событие: «случайная величина X принимает фиксированное значение a»; запись $X < x$ означает событие: «численное значение случайной величины X меньше некоторого числа x».

Определение. *Случайной величиной называется однозначная вещественная функция $\varphi(e)$, $e \in U$, для которой существует функция распределения:*

$$F(x) = P(X < x). \tag{2.1}$$

Значение функции распределения в точке x равно вероятности события: «численное значение случайной величины X меньше x». Иногда функцию распределения случайной величины (2.1) или какую-либо иную характеристику поведения случайной величины называют *законом распределения* случайной величины.

Пример 2.1. Обратимся к примеру 1.2. Пусть $\varphi(e)$ — число кодовых слов в сообщении с нечетным числом трансформаций, e — номер элементарного события по табл. 1.1: $e = 1, 2, \ldots, 27$. Очевидно, $\varphi(e)$ — вещественная, однозначная функция, определенная на пространстве U с множеством значений $\{0, 1, 2, 3\}$. Пусть $A_i \subset R$ — множество, состоящее из одного числа i, $i = 0, 1, 2, \ldots$, а B_i — подмножество пространства U, которое определяется по правилу: если $\varphi(e) = i$, то $e \in B_i$, и если $x(e) \neq i$, то $e \overline{\in} B_i$ ($\overline{\in}$ — не принадлежит). Допустим, что элементарные события пространства U равновозможны, поэтому можно воспользоваться классическим определением вероятностей. При этих условиях (см. табл. 1.1)

$$B_0 = \{1, 3, 7, 8, 9, 13, 14, 15\}; \qquad P(B_0) = P(A_0) = 8/27.$$

$$B_1 = \{4, 5, 6, 19, 20, 21, 22, 23, 24, 25, 26, 27\};$$

$$P(B_1) = P(A_1) = 12/27.$$

$$B_2 = \{10, 11, 12, 16, 17, 18\}; \qquad P(B_2) = P(A_2) = 6/27.$$

$$B_3 = \{2\}; \qquad P(B_3) = P(A_3) = 1/27.$$

$$B_4 = B_i = V, \ i = 5, 6, \ldots \qquad P(B_4) = P(A_4) = P(A_i) = 0.$$

Теперь возьмем множество вещественных чисел $A = [-\infty, x)$, x — произвольное число. Это множество в пространстве U определяет подмножество $B = \bigcup_{i<x} B_i$, так как если $e \in B$, то $\varphi(e) < x$ и, следовательно, $\varphi(e) \in A$. Поэтому

$$P(B) = P\left(\varphi(e) < x\right) = P\left(\varphi(e) \in A\right) = P(A),$$

а так как события B_i несовместны, то

$$P(A) = P(B) = P\left(\bigcup_{i<x} B_i\right) = \sum_{i<x} P(B_i). \qquad (2.2)$$

$P(A)$ можно подсчитать для любого x. Согласно определению (2.1) эта вероятность является функцией распределения случайной величины $\varphi(e)$. На рис. 2.1, *а* эта функция показана при условии равновозможности элементарных событий пространства U. Естественно, что это допущение мы приняли только в целях упрощения примера и оно совершенно не влияет на существование вероятности (2.1). График функции $F(x)$ на рис. 2.1 изменится, если это условие не будет выполняться.

Пример 2.2. Случайная величина X подчиняется биномиальному закону распределения (закону Бернулли), если ее функция распределения выражается формулой

$$F(x)=\begin{cases} 0 & \text{для } x\leqslant 0; \\ \sum\limits_{k=0}^{m} C_n^k p^k (1-p)^{n-k} & \text{для } 0<x\leqslant n; \\ 1 & \text{для } x>n, \end{cases} \qquad (2.3)$$

где $0<p<1$, $m<x\leqslant m+1$.

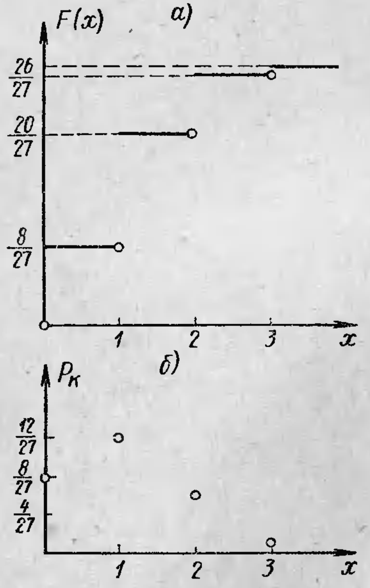

Рис. 2.1

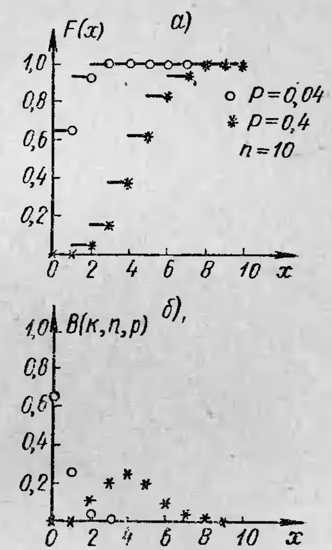

Рис. 2.2

Два графика этой функции для $n=10$, $p=0{,}04$ и 0,4 показаны на рис. 2.2. Графики имеют характер ступенчатых кривых, претерпевающих разрывы в точках 0, 1, 2, ..., 10. В точке $x=k$ скачок равен:

$$\left.\begin{aligned} b(k, n, p) &= C_n^k p^k (1-p)^{n-k} = C_n^k p^k q^{n-k}; \\ q &= 1-p. \end{aligned}\right\} \qquad (2.4)$$

При $p=0{,}04$ функция очень быстро достигает значений, близких к единице. Так, в интервале $3\leqslant x<4$ при $p=0{,}04$ $F(x)=0{,}9996$, а при $p=0{,}4$ $F(x)=0{,}382$.

Пример 2.3. Случайная величина X распределена по закону Пуассона, если ее функция распределения выражается формулой:

$$F(x)=\begin{cases} 0 & \text{для } x<0; \\ \sum\limits_{k=0}^{m} \dfrac{a^k}{k!}\exp(-a) & \text{для } x \geqslant 0, \end{cases} \tag{2.5}$$

где $a>0$, k — целое число, $m<x\leqslant m+1$.

Графики функции распределения Пуассона показаны на рис. 2.3 для $a=1$ и 10. Как следует из 2.5 и видно по графи-

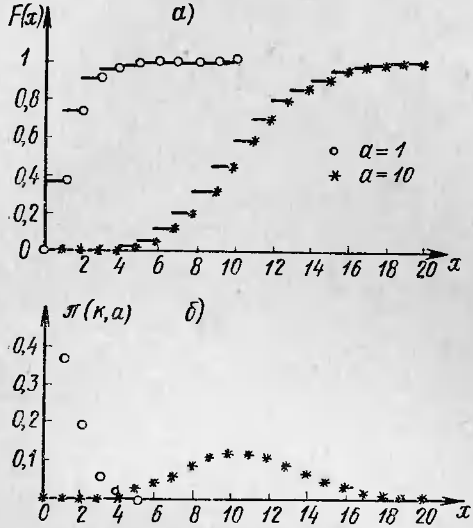

Рис. 2.3

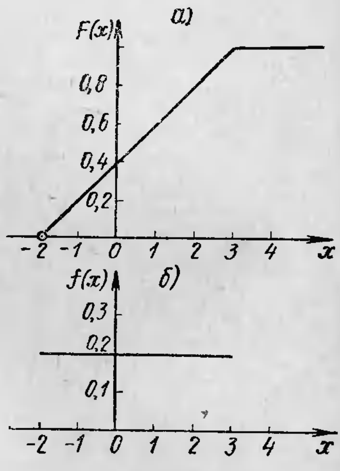

Рис. 2.4

кам, функция разрывна в целочисленных точках оси абсцисс. В точке $x=k$, $k=0, 1, \ldots, n, \ldots$ скачок функции равен:

$$\pi(k, a)=\frac{a^k}{k!}\exp(-a). \tag{2.6}$$

Пример 2.4. Случайная величина X распределена равномерно в промежутке $[a, b)$, если ее функция распределения выражается формулой

$$F(x)=\begin{cases} 0 & \text{для } x<a; \\ \dfrac{x-a}{b-a} & \text{для } a\leqslant x<b; \\ 1 & \text{для } x\geqslant b. \end{cases} \tag{2.7}$$

График этой функции распределения показан на рис. 2.4 для $a=-2$ и $b=3$. Как следует из (2.7) и видно по графику,

$F(x)$ — кусочно-линейная функция и в промежутке $[a, b)$ имеет непрерывную производную

$$f(x) = \frac{dF(x)}{dx} = \frac{1}{b-a}, \quad a \leqslant x < b. \qquad (2.8)$$

Пример 2.5. Случайная величина X распределена по нормальному закону, если ее функция распределения выражается формулой

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} \exp\left[-\frac{(t-m)^2}{2\sigma^2}\right] dt, \qquad (2.9)$$

где m может иметь любое вещественное значение, а $\sigma > 0$.

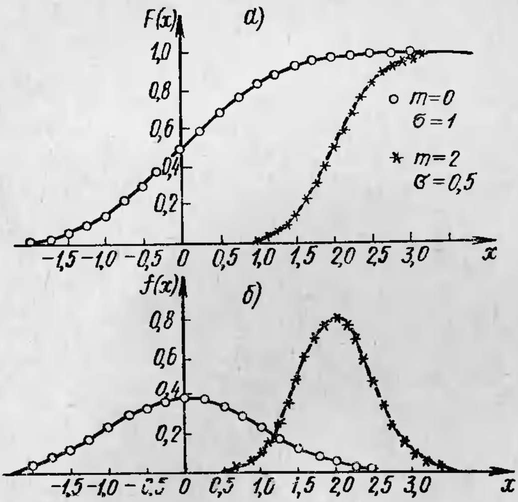

Рис. 2.5

Два графика этой функции распределения для значений параметров $m=0$, $\sigma=1$ и $m=2$, $\sigma=0{,}5$ показаны на рис. 2.5, *а*. При любых x функция $F(x)$ непрерывна и имеет непрерывную производную

$$f(x) = \frac{dF(x)}{dx} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right]. \qquad (2.10)$$

Функция распределения (2.9) имеет две асимптоты — ось абсцисс и прямую $y=1$, а ее производная $f(x)$ — одну асимптоту — ось абсцисс.

§ 2.2. Свойства функции распределения вероятностей

Функция распределения случайной величины позволяет вычислить вероятность того, что случайная величина X принимает значение из полуоткрытого промежутка $[a, b)$ для любых a и b, т. е. вероятность неравенства $a \leqslant X < b$.

Введем обозначения $A = (X < a)$, $B = (a \leqslant X < b)$ и $C = (X < b)$. События A и B несовместны, поэтому, применяя аксиому сложения вероятностей к очевидному равенству $C = A \cup B$, получим: $P(C) = P(A) + P(B)$, но $P(A) = F(a)$ и $P(C) = F(b)$, поэтому

$$P(a \leqslant X < b) = F(b) - F(a). \tag{2.11}$$

Пример 2.6. Пусть X_1 распределена по закону Пуассона (2.5) с параметром $a_1 = 1{,}0$, а X_2 — по этому же закону, но с параметром $a_2 = 10$. По кривым на рис. 2.3 находим:

$$P(6{,}5 \leqslant X_1 < 14{,}5) = F_1(14{,}5) - F_1(6{,}5) \approx 1 - 1 = 0.$$

$$P(6{,}5 \leqslant X_2 < 14{,}5) = F_2(14{,}5) - F_2(6{,}5) = 0{,}9 - 0{,}13 = 0{,}77.$$

Поскольку $P(a \leqslant X < b) \geqslant 0$, то из (2.11) вытекает, что функция распределения любой случайной величины есть неубывающая функция: для любых a и b из $a < b$ следует, что

$$F(a) \leqslant F(b). \tag{2.12}$$

Поскольку $F(x)$ есть вероятность некоторого события, то согласно (1.16) для всех x

$$0 \leqslant F(x) \leqslant 1. \tag{2.13}$$

Для любой случайной величины событие $X < -\infty$ является невозможным, а событие $X < +\infty$ — достоверным, поэтому

$$\lim_{x \to -\infty} F(x) = F(-\infty) = 0; \tag{2.14}$$

$$\lim_{x \to +\infty} F(x) = F(+\infty) = 1. \tag{2.15}$$

Пример 2.7. Пусть X распределена по нормальному закону (2.9). Найдем $F(+\infty)$. Согласно (2.15)

$$F(+\infty) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp\left[-\frac{(t-m)^2}{2\sigma^2}\right] dt =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy. \tag{2.16}$$

Здесь

$$y = \frac{(t-m)}{\sigma}, \quad \sigma dy = dt. \tag{2.17}$$

Известно [2], что

$$\int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy = \sqrt{2\pi}, \tag{2.18}$$

поэтому

$$F(+\infty) = 1. \tag{2.19}$$

Из (2.12) и (2.13) следует, что функция распределения может иметь разрывы только типа скачков. Множество точек разрыва может быть не более чем счетно [4].

Пусть в точке a функция распределения имеет разрыв и $x_1 < x_2 < \ldots < x_n < \ldots$ — любая возрастающая последовательность, сходящаяся к a. Обозначим через A_n событие $x_n \leqslant X < a$. Совокупность событий A_n, $n = 1, 2, \ldots$ такова, что $A_1 \supset A_2 \supset \ldots$ и не имеет ни одной точки, общей для всех A_n, т. е.

$$\bigcap_{n=1}^{\infty} A_n = V, \tag{2.20}$$

поэтому

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = 0. \tag{2.21}$$

На основании теоремы непрерывности

$$\lim_{n\to\infty} P(A_n) = 0.$$

Поэтому

$$\lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} [F(a) - F(x_n)] = F(a) - \lim_{n\to\infty} F(x_n) = 0,$$

или

$$\lim_{n\to\infty} F(x_n) = F(a). \tag{2.22}$$

Этим доказана непрерывность функции распределения слева: левое предельное значение функции $F(x)$ в точке разрыва a существует и равно $F(a)$.

Если взять последовательность чисел $x_1 > x_2 > \ldots$ $\ldots > x_n > \ldots$, сходящуюся к точке a справа, то последовательность событий $B_n = (a \leqslant X < x_n)$, $n = 1, 2, \ldots$ содержит общую точку, поэтому $\lim\limits_{n\to\infty} P(B_n) \neq 0$ и равенство, аналогичное (2.22), не имеет места. Справа в точке a функция $F(x)$ претерпевает разрыв.

Скачок $F(x)$ в точке a, равный $F(a+0) - F(a)$, есть вероятность события $(a \leqslant X < a+0) = (X = a)$. Если $F(x)$ непрерывна в a, то

$$P(X = a) = F(a+0) - F(a) = 0.$$

В промежутке непрерывности функции распределения отличную от нуля вероятность могут иметь только события типа $x_1 \leqslant X < x_2$, $x_1 \neq x_2$. С первого взгляда представляется парадоксальным утверждение о том, что для случайной величины X и некоторого числа a из множества ее значений вероятность события $X=a$, т. е. вероятность того, что случайная величина принимает одно из возможных значений, равна нулю. Парадоксальность усугубляется и тем обстоятельством, что каждое испытание непременно завершается появлением одного из таких событий с нулевой вероятностью. Такое поведение случайной величины связано с тем, что множество ее значений в промежутке непрерывности $F(x)$ несчетно, и совершенно аналогично тому, как несчетное множество геометрических точек, т. е. объектов, не имеющих размера, образует отрезок линии, т. е. объект, имеющий размер. В нашем случае несчетное множество событий нулевой вероятности образует событие ненулевой вероятности.

§ 2.3. Дискретные и непрерывные случайные величины

Большинство случайных величин, с которыми приходится иметь дело, принадлежат одному из двух типов: дискретному или непрерывному. Для них существуют способы представления законов распределения иногда более удобные, чем функции распределения (2.1).

Определение 1. *Случайная величина называется дискретной, если множество возможных значений ее конечно или счетно.*

Вероятностные свойства дискретной случайной величины X, принимающей n значений, полностью характеризуются табл. 2.1, сопоставляющей каждому возможному значению случайной величины X вероятность его появления $p_i = P(X=a_i)$. Совокупность несовместных событий $X=a_i$, $i=\overline{1,n}$ образует полную группу, поэтому

$$\sum_{i=1}^{n} p_i = 1. \tag{2.23}$$

Таблица 2.1

x_i	a_1	a_2	a_3	. . . .	a_n
p_i	p_1	p_2	p_3		p_n

Если $a_1 < a_2 < \ldots < a_n$, то на основании аксиомы сложения вероятностей

$$F(x) = P(X < x) = \begin{cases} 0, & \text{если } x < a; \\ \sum\limits_{i=1}^{k} p_i, & \text{если } a_k < x \leqslant a_{k+1}. \end{cases} \tag{2.24}$$

Сумма (2.24) представляет собой функцию распределения X, а табл. (2.1) является законом распределения случайной величины. Эта форма закона распределения называется *рядом распределения* дискретной случайной величины X. Выражение (2.24) справедливо и при счетном множестве значений X, а в равенстве (2.23) в этом случае появляется бесконечная сумма

$$\sum_{i=1}^{\infty} p_i = 1. \qquad (2.25)$$

Из (2.22) и (2.24) следует, что функция распределения любой дискретной случайной величины X разрывна, возрастает скачками в тех точках x, которые являются возможными значениями X. Если два возможных значения a_i и a_j разделены промежутком, в котором нет других возможных значений X, то в этом промежутке, как следует из (2.11), $F(x)$ постоянна. Для любых x_1 и x_2, таких, что $a_i < x_1 < x_2 < a_j$,

$$P(x_1 \leqslant X < x_2) = F(x_2) - F(x_1) = 0. \qquad (2.26)$$

При конечном числе возможных значений, например n, функция $F(x)$ является ступенчатой функцией с $n+1$-м промежутком постоянства. Счетное же множество возможных значений X может быть всюду плотным, и в этом случае функция распределения дискретной случайной величины, оставаясь разрывной, не будет иметь промежутков постоянства, так как в любом сколь угодно малом интервале будут находиться ее возможные значения.

Случайные величины, подчиняющиеся биномиальному закону распределения (2.3) и закону Пуассона (2.5), — типичные примеры дискретных случайных величин.

Функция распределения биномиального закона с параметрами n и p (см. рис. 2.2, *а*) изменяется скачками в точках $x=k$, $k=\overline{0,n}$ на величину $b(k, n, p)$ из (2.4).

В промежутках между этими точками $F(x)$ не меняется, поэтому целые числа 0, 1, 2, ..., n — возможные значения X, а совокупность вероятностей $b(k, n, p)$, $k=\overline{0,n}$ есть ряд распределения X. Справедливость (2.23) для этого ряда устанавливается следующим образом:

$$\sum_{k=0}^{n} b(k, n, p) = \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} = (p+1-p)^n = 1.$$

Два ряда биномиального распределения для $n=10$, $p=0{,}4$ и $n=10$, $p=0{,}04$ показаны точками на рис. 2.2, *б*.

Функция распределения Пуассона (2.5) (см. рис. 2.3) претерпевает разрывы во всех целочисленных точках, оставаясь неизменной в промежутках между ними. Поэтому множеством значений случайной величины является счетное множество целых чисел 0, 1, 2, ..., а рядом распределения — счетное множество вероятностей $\pi(k, a)$ (2.6).

Справедливость (2.25) устанавливается просто, если учесть, что

$$\exp a = \sum_{k=0}^{\infty} \frac{a^k}{k!};$$

$$\sum_{k=0}^{\infty} \pi(k, a) = \sum_{k=0}^{\infty} \frac{a^k}{k!} \exp(-a) = \exp a \cdot \exp(-a) = 1.$$

Два ряда распределения Пуассона для параметров $a=1$ и 10 показаны точками на рис. 2.3, *б*.

Определение 2. *Случайная величина X с функцией распределения $F(x)$ называется непрерывной (непрерывно распределенной), если существует такая неотрицательная функция $f(x)$, что для любого x*

$$F(x) = \int_{-\infty}^{x} f(t)\, dt. \tag{2.27}$$

Функция $f(x)$ является законом распределения и называется *плотностью распределения вероятностей* случайной величины (в дальнейшем просто плотностью распределения).

В точках непрерывности плотности распределения

$$f(x) = \frac{dF(x)}{dx}, \tag{2.28}$$

и поэтому, если пренебречь бесконечно малыми высших порядков,

$$P(x \leqslant X < x + dx) = dF(x) = f(x)\, dx. \tag{2.29}$$

Не ограничивая существенно общность изложения, по крайней мере в отношении задач прикладного характера, будем в дальнейшем считать, что непрерывные случайные величины имеют непрерывные (или кусочно-непрерывные) плотности распределения на всем множестве возможных значений случайной величины.

Из (2.27) и свойств функции распределения вытекают следующие свойства плотности распределения:

1. $P(a \leqslant X < b) = \int_{a}^{b} f(x)\, dx.$ (2.30)

2. $\int_{-\infty}^{\infty} f(x)\, dx = 1.$ (2.31)

3. $f(x)$ имеет размерность, обратную размерности случайной величины.

Примерами непрерывно распределенных случайных величин являются случайные величины с равномерным законом распределения (2.7) и нормальным (2.9). Соответствующие плотности распределения выражаются формулами (2.8) и (2.10), а их графики показаны на рис. 2.4, *б* и 2.5, *б*.

Любая неотрицательная функция, удовлетворяющая условию (2.31), может рассматриваться в качестве плотности распределения вероятностей некоторой непрерывной случайной величины.

Иногда функцию $F(x)$ называют интегральной функцией распределения, а $f(x)$ — дифференциальной функцией распределения случайной величины.

Естественно ввести в рассмотрение случайные величины смешанного типа с функциями распределения, подобными функции $F(x)$, показанной на рис. 2.6. Случайная величина с такой функцией распределения принимает численные значения x_1, x_2, x_3 с конечными вероятностями, а в промежутках (x_0, x_1), (x_1, x_2), (x_2, x_3), $(x_3, +\infty)$ распределена непрерывно.

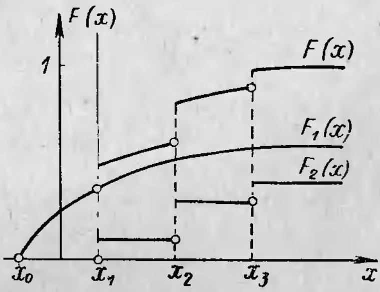

Рис. 2.6

Как видно из рис. 2.6, функцию распределения такой случайной величины можно представить в виде суммы двух функций

$$F(x) = F_1(x) + F_2(x), \qquad (2.32)$$

из которых одна непрерывная, а вторая ступенчатая.

Изложение теории вероятностей удобно вести именно для этого случая, что позволяет избежать некоторых повторений, поскольку дискретные и непрерывные распределения, как это следует из (2.32), являются частными случаями распределений смешанного типа. Но такого характера обобщение требует и обобщенного понятия интеграла. В дальнейшем используется интеграл Стилтьеса. Необходимые для наших целей свойства интеграла этого типа приведены в следующем параграфе. Желающие познакомиться с теорией интеграла Стилтьеса глубже могут обратиться, например, к [10].

§ 2.4. Интеграл Стилтьеса

Положим, что в конечном промежутке $[a, b]$ определены две функции $f(x)$ и $g(x)$, причем $f(x)$ непрерывна в этом промежутке, а $g(x)$ не убывает на $[a, b]$. Разобьем этот промежуток на n частичных промежутков точками

$$a=x_0<x_1<x_2<\ldots<x_{k-1}<x_k<\ldots<x_n=b$$

и составим сумму

$$\sigma_n=\sum_{k=1}^{n} f(\xi_k)\,[g(x_k)-g(x_{k-1})], \tag{2.33}$$

где ξ_k — произвольная точка в частичном промежутке $[x_{k-1}, x_k]$: $x_{k-1}\leqslant\xi_k\leqslant x_k$. Эта сумма называется суммой Римана—Стилтьеса. Обозначим символом Δ наибольшую из разностей x_k-x_{k-1}, $k=\overline{1,n}$:

$$\Delta=\max_{1\leqslant k\leqslant n}(x_k-x_{k-1}).$$

Пусть число n точек x_k неограниченно возрастает, причем так, что $\Delta\to 0$. Если сумма σ_n при этом стремится к конечному пределу, который не зависит от выбора точек ξ_k, то этот предел называют интегралом Стилтьеса на промежутке $[a, b]$ от $f(x)$ с интегрирующей функцией $g(x)$ и пишут

$$\lim_{\substack{n\to\infty\\ \Delta\to 0}}\sum_{k=1}^{n} f(\xi_k)\,[g(x_k)-g(x_{k-1})]=\int_a^b f(x)\,dg(x). \tag{2.34}$$

В интеграле Римана $g(x)=x$, поэтому интеграл (2.34) обладает многими свойствами интеграла Римана, и доказательство этих свойств проводится совершенно так же, как и для интеграла Римана. Приведем эти свойства, считая, что все интегралы, входящие в формулы, существуют.

$$\left.\begin{aligned}
&\int_a^b\left[\sum_{k=1}^{n} a_k f_k(x)\right]dg(x)=\sum_{k=1}^{n} a_k\int_a^b f_k(x)\,dg(x);\\
&\int_a^b f(x)\,d\left[\sum_{k=1}^{n} a_k g(x)\right]=\sum_{k=1}^{n} a_k\int_a^b f(x)\,dg_k(x);\\
&\int_a^b f(x)\,dg(x)=\int_a^c f(x)\,dg(x)+\int_c^b f(x)\,dg(x);\\
&\int_a^b dg(x)=g(b)-g(a).
\end{aligned}\right\} \tag{2.35}$$

Здесь a_k — постоянные.

Если в точке a функция $g(x)$ имеет скачок $g(a+0)-$ $-g(a-0)>0$ и $f(a)\neq 0$, то интеграл Стилтьеса по промежутку, стягивающемуся к точке a, не равен нулю. Действительно,

$$\int_{a-0}^{b} f(x)\,dg(x) = \lim_{\substack{n\to\infty \\ \Delta\to 0}} \sum_{k=1}^{n} f(\xi_k)\,[g(x_k) - g(x_{k-1})] =$$

$$= \lim_{\substack{n\to\infty \\ \Delta\to 0}} \sum_{k=2}^{n} f(\xi_k)\,[g(x_k) - g(x_{k-1})] +$$

$$+ \lim_{x_1\to a} f(\xi_1)\,[g(x_1) - g(a-0)] =$$

$$= \int_{a+0}^{b} f(x)\,dg(x) + f(a)\,[g(a+0) - g(a-0)];$$

$$\int_{a-0}^{a+0} f(x)\,dg(x) = \int_{a-0}^{b} f(x)\,dg(x) - \int_{a+0}^{b} f(x)\,dg(x) =$$

$$= f(a)\,[g(a+0) - g(a-0)] \neq 0. \qquad (2.36)$$

Имеет место следующая т е о р е м а существования интеграла Стилтьеса: *если $f(x)$ непрерывна внутри промежутка интегрирования и ограничена, а $g(x)$ — неубывающая функция и непрерывна на концах промежутка, то интеграл Стилтьеса от $f(x)$ по $g(x)$ в промежутке $[a, b]$ существует.* Теорема справедлива и для бесконечного промежутка $[-\infty, +\infty]$.

Для неограниченной и непрерывной на промежутке $[-\infty, +\infty]$ функции $f(x)$ для любых конечных a и b существует интеграл Стилтьеса $f(x)$ по $g(x)$ на $[a, b]$, если $g(x)$ по-прежнему не убывает и ограничена. Если при стремлении a к $-\infty$ и b к $+\infty$ этот интеграл имеет конечный определенный предел, то этот предел принимается за величину интеграла на бесконечном промежутке и называется несобственным интегралом Стилтьеса:

$$\lim_{\substack{a\to-\infty \\ b\to+\infty}} \int_{a}^{b} f(x)\,dg(x) = \int_{-\infty}^{\infty} f(x)\,dg(x). \qquad (2.37)$$

Интеграл (2.37) существует и называется абсолютно сходящимся, если существует интеграл

$$\int_{-\infty}^{\infty} |f(x)|\,dg(x) = \lim_{\substack{a\to-\infty \\ b\to+\infty}} \int_{a}^{b} |f(x)|\,dg(x). \qquad (2.38)$$

Отметим два частных случая интеграла Стилтьеса — случай ступенчатой функции $g(x)$ и непрерывной $f(x)$.

Пусть в точках $a = c_0 < c_1 < \ldots < c_{n-1} < c_n = b$ промежутка $[a, b]$ функция $g(x)$ претерпевает разрывы типа скачков

$g(c_k+0) - g(c_k-0) = \alpha_k$, а внутри каждого промежутка $[c_{k-1}, c_k]$ сохраняет постоянное значение g_k. α_0 и α_n — скачки в точках a и b соответственно. Разделим $[a, b]$ на части так, чтобы точки c_i не являлись точками деления. В сумме (2.33) все слагаемые, у которых x_{k-1} и x_k лежат внутри одного и того же промежутка $[c_{i-1}, c_i]$, будут равны нулю, так как $g(x_k) = g(x_{k-1})$. Если промежуток $[x_{k-1}, x_k]$ содержит точку разрыва c_m, то $f(\xi_k)$ при беспредельном возрастании n и $\Delta \to 0$ будет стремиться к $f(c_m)$, а $g(x_{k-1}) - g(x_k)$ — к α_m, поэтому в пределе сумма (2.33) обратится в конечную сумму

$$\int_a^b f(x)\, dg(x) = \sum_{m=0}^{n} f(c_m)\, \alpha_m. \tag{2.39}$$

Если точка разрыва c_m является точкой деления $[a, b]$ на части, то надо рассматривать оба интервала, разделенные точкой c_m, и результат будет тем же.

Рассмотрим второй частный случай. Пусть $f(x)$ и $g(x)$ непрерывны на $[a, b]$ и $g(x)$ имеет внутри $[a, b]$ ограниченную производную $g'(x)$. Применяя к разностям $g(x_k) - g(x_{k-1})$ формулу Лагранжа, можно сумму (2.33) записать в виде

$$\sum_{k=1}^{n} f(\xi_k)\, [g(x_k) - g(x_{k-1})] = \sum_{k=1}^{n} f(\xi_k)\, g'(\xi_k')\, [x_k - x_{k-1}],$$

где

$$x_{k-1} \leqslant \xi_k \leqslant x_k \quad \text{и} \quad x_{k-1} \leqslant \xi_k' \leqslant x_k.$$

В силу равномерной непрерывности функции $f(x)$ на $[a, b]$ пределы суммы при $n \to \infty$ и $\Delta \to 0$ существуют и равны пределу сумм Римана

$$\sum_{k=1}^{n} f(\xi_k)\, g'(\xi_k)\, [x_k - x_{k-1}],$$

поэтому

$$\int_a^b f(x)\, dg(x) = \int_a^b f(x)\, g'(x)\, dx. \tag{2.40}$$

В дальнейшем изложении интегрирующими функциями будут функции распределения случайных величин, которые удовлетворяют условиям теорем существования интеграла Стилтьеса. Поскольку результат интегрирования по Стилтьесу может зависеть, как следует из (2.36), от того, входят или не входят границы в интервал интегрирования, условимся, учитывая определение функции распределения (2.1) и ее свойство (2.11), левую границу всегда включать в интервал интегрирования, а правую исключать. При таком соглашении если $F(x)$ есть функция распределения случайной величины X, то

$$\int_a^b dF(x) = F(b) - F(a) = P(a \leqslant X < b). \qquad (2.41)$$

Очевидны также следующие равенства:

$$F(x) = \lim_{a \to -\infty} \int_a^x dF(x) = \int_{-\infty}^x dF(x); \qquad (2.42)$$

$$\int_{-\infty}^{\infty} dF(x) = 1, \qquad (2.43)$$

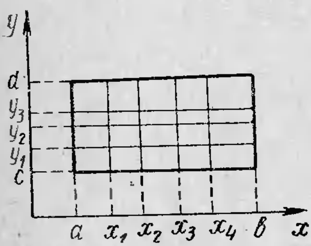

Рис. 2.7

причем интегралы Стилтьеса превращаются в суммы для дискретных случайных величин, а для непрерывных могут быть преобразованы в интегралы Римана.

Понятие интеграла Стилтьеса можно обобщить на случай плоскости, трехмерного пространства или вообще n-мерного пространства. Мы рассмотрим случай плоскости, а для остальных пространств построения будут аналогичны. Прежде всего необходимо обобщить понятия промежутка и разбиения промежутка на частичные промежутки для случая двух измерений. Возьмем на координатных осях xy плоскости промежутки $S_x=[a, b)$ и $S_y=[c, d)$. Множество S всех точек (x, y) плоскости, для которых $a \leqslant x < b$ и $c \leqslant y < d$, называют полуоткрытым промежутком на плоскости. Если промежутки S_x и S_y разбить какими-либо способами на частичные промежутки, то промежуток S также разобьется на частичные промежутки. Каждый частичный промежуток из S определяется одним из частичных промежутков из S_x и одним из частичных промежутков из S_y (рис. 2.7).

Последовательность промежутков $S_1, S_2, \ldots, S_n, \ldots$ называется исчезающей, если $S_1 \supset S_2 \supset \ldots \supset S_n \supset \ldots$ и ни одна точка плоскости не является общей точкой всех промежутков последовательности. Нетрудно установить структуру промежутков исчезающей последовательности. Если S_i определяется интервалами $[a_i, b_i)$ и $[c_i, d_i)$, то $a_i \leqslant a_{i+1}$, $b_i \geqslant b_{i+1}$ для всех i и $(b_i - a_i) \to 0$ при беспредельном возрастании, поэтому монотонные последовательности a_i и b_i имеют общий предел x^*, причем $a_i \leqslant x^* \leqslant b_i$ для любого i. Поскольку ни одна точка, в том числе и x^*, не может быть общей для всех S_i, то, очевидно, для всех достаточно больших i точка x^* должна попасть на правый открытый конец промежутка, т. е. для достаточно больших i все промежутки на оси x имеют вид $[a_i, x^*)$. Аналогичную структуру имеют промежутки и на оси y.

Говорят, что на промежутке S_0 определена функция промежутка $G(S)$, если любому частичному промежутку из S_0 поставлено в соответствие вещественное число. Здесь слово «любому» означает, что частичные промежутки в S_0 могут быть определены одномерными промежутками любого вида: (a, b), $[a, b)$, $(a, b]$, $[a, b]$, в том числе и отдельной точкой, которую следует трактовать как промежуток $[a, a]$.

Будем рассматривать функции промежутков, обладающие следующими тремя свойствами:

1. Функция $G(S)$ неотрицательная.

2. Функция $G(S)$ аддитивная, т. е. если промежуток S из S_0 разбит на конечное число попарно непересекающихся промежутков $S_1, S_2, \ldots, S_n$, т. е. $S = \bigcup_{i=1}^{n} S_i$, то

$$G(S) = \sum_{i=1}^{n} G(S_i).$$

3. Функция $G(S)$ стремится к нулю на исчезающей последовательности промежутков.

Свойство 3 называют нормальностью функции $G(S)$, а функцию, обладающую этим свойством, — нормальной.

Функцию $G(S)$, определенную только для промежутков, принадлежащих S_0, легко распространить вообще на все промежутки плоскости, причем функция останется неотрицательной, аддитивной и нормальной. Действительно, если S — любой промежуток плоскости, то согласно (1.5) $S \cap S_0 \subset S_0$, и мы получим упомянутое расширение функции $G(S)$, если положим, что $G(S) = G(S \cap S_0)$ и $G(V) = 0$.

Аргументом функции $G(S)$ может быть отдельная точка P из S_0 или отрезок линии l, входящей в состав S_0. Здесь термин «линия» обозначает полуоткрытый промежуток плоскости, для которого либо $x = \text{const}$, либо $y = \text{const}$. Если $G(P) = 0$ во всех точках P, принадлежащих S_0, то функция $G(S)$ называется непрерывной в S_0. Если же $G(P) > 0$, то точка P называется точкой разрыва $G(S)$. Можно показать, что если $G(S)$ имеет точки разрыва, то множество их конечно или счетно. Аналогично, если $G(l) > 0$, то l — линия разрыва. Множество линий разрыва функции $G(S)$, если они существуют, конечно или счетно.

Нетрудно разложить функцию $G(S)$ на функцию скачков и непрерывную часть. Пусть, например, P_k, $k=1, 2, \ldots$ — точки разрыва непрерывности $G(S)$. Определим функцию скачков $G_2(S)$ следующим образом: $G_2(S)$ равна сумме значений $G(P_k)$ в тех точках P_k, которые принадлежат S. Если таких точек счетное множество, то получаемый ряд будет сходиться, поскольку

$$\sum_{k=1}^{\infty} G(P_k) \leqslant G(S_0).$$

Разность $G(S) - G_2(S) = G_1(S)$ уже не имеет точек разрыва непрерывности. Таким образом,

$$G(S) = G_1(S) + G_2(S) \tag{2.44}$$

и обе функции $G_1(S)$ и $G_2(S)$ неотрицательны, аддитивны и нормальны. В дальнейшем из чисто практических соображений будем считать, что $G(S)$ имеет только точки разрыва непрерывности.

Приведенные определения становятся совершенно ясными, если толковать $G(S)$ как массу, которую имеет промежуток S при распределении материи на промежутке S_0, или как площадь промежутка S.

Пусть на конечном промежутке S_0 определена непрерывная ограниченная функция $f(x, y)$ и неотрицательная, аддитивная и нормальная функция промежутков $G(S)$. Разобьем S_0 на частичные промежутки $S_1, S_2, \ldots, S_n$, в каждом промежутке возьмем произвольную точку (ξ_k, η_k) и составим сумму произведений

$$\sigma_n = \sum_{k=1}^{n} f(\xi_k, \eta_k)\, G(S_k). \tag{2.45}$$

Напомним, что диаметром $d(A)$ множества A называется точная верхняя грань расстояний $\rho(x, y)$, если точки x и y независимо друг от друга пробегают все множество A. Обозначим символом Δ наибольший из диаметров промежутков $S_1, S_2, \ldots, S_n$:

$$\Delta = \max_{1 \leqslant k \leqslant n} d(S_k).$$

Пусть число n частичных промежутков S_k, $k = \overline{1, n}$ неограниченно возрастает, причем так, что $\Delta \to 0$. Если при этом и любом выборе точек (ξ_k, η_k) сумма (2.45) стремится к определенному и конечному пределу, то этот предел называют интегралом Стилтьеса на промежутке S_0 от $f(x, y)$ по функции $G(S)$ и пишут

$$\lim_{\substack{n\to\infty\\ \Delta\to 0}} \sum_{k=1}^{n} f(\xi_k, \eta_k)\, G(S_k) = \iint_{S_0} f(x, y)\, G(dS). \tag{2.46}$$

От функции промежутков $G(S)$ всегда можно перейти к функции точки. Действительно, пусть $G(S)$ определена на всей плоскости. Будем считать, что $g(x, y)$ равна значению $G(S)$ на промежутке S_{xy}, который определяется следующими одномерными промежутками: $[-\infty, x)$, $[-\infty, y)$, т. е. $g(x, y) = G(S_{xy})$. При таком определении и указанных выше свойствах функции промежутков $G(S)$ функция точки $g(x, y)$ не убывает по каждой переменной при любом фиксированном значении другой переменной. Кроме того, для промежутка S_k (рис. 2.8), определяемого одномерными промежутками $[x_{k-1}, x_k)$ и $[y_{k-1}, y_k)$,

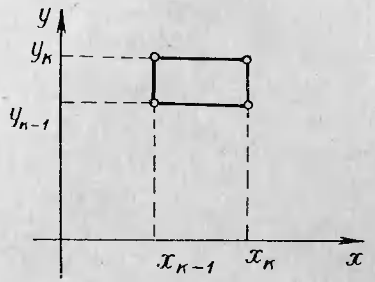

Рис. 2.8

$$G(S_k) = \Delta g(x_k, y_k) = g(x_k, y_k) - g(x_k, y_{k-1}) - g(x_{k-1}, y_k) + \\ + g(x_{k-1}, y_{k-1}) \geqslant 0. \tag{2.47}$$

Если функция $f(x, y)$ непрерывна в промежутке $S_0(a \leqslant x < b;\ c \leqslant y < d)$ и для $g(x, y)$ справедливо (2.47), то интеграл Стилтьеса можно определить как предел, к которому стремится сумма

$$\sigma_{n,m} = \sum_{k=1}^{n} \sum_{l=1}^{m} f(\xi_k, \eta_l)\,[g(x_k, y_l) - g(x_{k-1}, y_l) - \\ - g(x_k, y_{l-1}) + g(x_{k-1}, y_{l-1})], \tag{2.48}$$

когда числа n и m точек деления промежутков $[a, b)$, $[c, d)$ неограниченно возрастают, а наибольшая из разностей

$$(x_k - x_{k-1}),\ (y_l - y_{l-1}),\quad k = \overline{1, n},\quad l = \overline{1, m}$$

стремится к нулю:

$$\lim_{\substack{n,\, m\to\infty\\ \Delta\to 0}} \sigma_{n,m} = \iint_{S_0} f(x, y)\, dg(x, y). \tag{2.49}$$

Здесь

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b;$$

$$c = y_0 < y_1 < \ldots < y_{m-1} < y_m = d;$$

$$x_{k-1} \leqslant \xi_k \leqslant x_k; \quad y_{l-1} \leqslant \eta_l \leqslant y_l.$$

Можно показать, что предел $\sigma_{n,m}$ существует, если $f(x, y)$ непрерывна и ограничена внутри S_0, а функция $G(S)$ кроме условий 1—3 удовлетворяет дополнительно следующему условию: если S_i — промежутки из S_0, которые расширяются и стремятся к S_0 так, что любая внутренняя точка S_0 попадет в S_i при достаточно больших i, то

$$G(S_i) \to G(S_0). \tag{2.50}$$

Из (2.48) следует, что свойства двукратного интеграла Стилтьеса совершенно аналогичны указанным раньше свойствам однократного интеграла. В частности, если промежуток S разбит на частичные непересекающиеся промежутки $S_1, S_2, \ldots, S_n$, т. е.

$$S = \bigcup_{i=1}^{n} S_i,$$

то

$$\iint\limits_{S} f(x, y)\, dg(x, y) = \sum_{k=1}^{n} \iint\limits_{S_k} f(x, y)\, dg(x, y).$$

Областью интегрирования может быть не только промежуток S, но и более сложная область Q плоскости, которая представляет собой конечную сумму промежутков. Мы можем совершать сколь угодно мелкие разбиения на частичные промежутки, составлять суммы (2.48) и переходить к пределу. Интеграл по Q сводится к конечной сумме интегралов по промежуткам, на которые можно разбить Q, и он не зависит от способа разбиения Q на промежутки.

Несобственный интеграл Стилтьеса, когда промежутком интегрирования является вся плоскость, определяют обычным путем. Рассматривают интеграл по произвольному конечному промежутку $[a, b)$, $[c, d)$, затем величины a, b, c, d произвольным образом устремляют к $-\infty$ и $+\infty$; если при этом существует предел

$$\lim_{\substack{a, c \to -\infty \\ b, d \to +\infty}} \int_a^b \int_c^d f(x, y)\, dg(x, y),$$

то этот предел называют несобственным интегралом Стилтьеса от функции $f(x, y)$ по функции $g(x, y)$ и обозначают

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dg(x, y).$$

При вычислениях двукратные интегралы Стилтьеса (как и двукратные интегралы Римана) сводятся к повторным интегралам. Если $f(x, y)=1$, то из (2.47) следует, что

$$\int_a^b \int_c^d dg(x, y) = g(b, d) - g(b, c) - g(a, d) + g(a, c). \quad (2.51)$$

Если $g(x, y)=g_1(x)g_2(y)$, то (2.47) преобразуется к виду

$$\Delta g(x_k, y_k) = [g_1(x_k) - g_1(x_{k-1})]\,[g_2(y_l) - g_2(y_{l-1})] =$$
$$= \Delta g_1(x_k)\,\Delta g_2(y_l),$$

поэтому

$$\int_a^b \int_c^d f(x, y)\, d[g_1(x)\, g_2(y)] = \int_a^b \int_c^d f(x, y)\, dg_1(x)\, dg_2(y). \quad (2.52)$$

Отметим, как и для одной переменной, два частных случая двукратного интеграла Стилтьеса — случай ступенчатой функции $g(x, y)$ и непрерывной.

Ступенчатая функция $g(x, y)$ получается как частный случай из (2.44) при $G_1(S)\equiv 0$. Используя (2.48), нетрудно найти, что при непрерывной в S_0 функции $f(x, y)$

$$\int_a^x \int_c^y f(t_1, t_2)\, dg(t_1, t_2) = \sum_k f(x_k, y_k)\, [g\,(x_k+0, y_k+0) -$$
$$- g\,(x_k+0, y_k-0) - g\,(x_k-0, y_k+0) +$$
$$+ g\,(x_k-0, y_k-0)], \quad (2.53)$$

где (x_k, y_k) — координаты точки P_k разрыва непрерывности функции $G(S)$, и суммирование распространяется на все те точки, для которых $a\leqslant x_k<x$ и $c\leqslant y_k<y$.

Непрерывную функцию $g(x, y)$ мы получим, если в (2.44) $G_2(S)\equiv 0$.

Пусть функция $f(x, y)$ непрерывна в S_0, а функция $g(x, y)$ в S_0 имеет непрерывные частные производные до второго порядка включительно. Возьмем приращения Δx и Δy такими, чтобы они не выводили точку $(x+\Delta x,\ y+\Delta y)$ из области S, и запишем (2.47) в следующем виде:

$$\Delta g(x, y) = g\,(x+\Delta x, y+\Delta y) - g\,(x+\Delta x, y) -$$
$$- g\,(x, y+\Delta y) + g(x, y).$$

При фиксированных y и Δy разность $g(x, y+\Delta y)-g(x, y)=$ $=\varphi(x)$ можно рассматривать как функцию одной переменной x, причем $\Delta g(x, y)=\Delta\varphi(x)$. Применяя к $\Delta\varphi$ формулу Лагранжа, получим

$$\Delta g(x, y)=\Delta\varphi(x)=\varphi_x^{(1)}(x+\theta\Delta x)\,\Delta x=$$

$$=[g_x^{(1)}(x+\theta\Delta x, y+\Delta y)-g_x^{(1)}(x+\theta\Delta x, y)]\,\Delta x.$$

Здесь $0\leqslant\theta\leqslant 1$. Частная производная $g_x^{(1)}(x, y)$ — дифференцируемая в S_0 функция, поэтому

$$\Delta g(x, y)=[g_{xx}^{(2)}(x, y)\,\theta\Delta x+g_{xy}^{(2)}(x, y)\,\Delta y-$$

$$-g_{xx}^{(2)}(x, y)\,\theta\Delta x+\alpha_1\theta\Delta x+\beta_1\Delta y-\alpha_1\theta\Delta x]\,\Delta x=$$

$$=[g_{xy}^{(2)}(x, y)+\beta]\,\Delta x\Delta y, \tag{2.54}$$

где α_1, β_1, β — бесконечно малые более высокого порядка, чем Δx и Δy. Используя (2.54) в сумме (2.48), можно показать, что

$$\int_a^b\int_c^d f(x, y)\,dg(x, y)=\int_a^b\int_c^d f(x, y)\,g_{xy}^{(2)}(x, y)\,dxdy. \tag{2.55}$$

§ 2.5. Многомерные случайные величины

Случайное событие может характеризоваться совокупностью нескольких физических величин, каждая из которых принимает при испытании определенное значение, зависящее от наблюдаемого события. Например, паспортные данные транзистора содержат «типичные» значения девяти электрических параметров для определенного режима питания: h_{11}, h_{12}, h_{22}, h_{21}, $C_к$, f_α, $F_ш$, $I_{к0}$, $I_{э0}$, (см. табл. 3.1). Измерив эти параметры у любого образца, выбор которого есть случайное событие, мы получим девять чисел, несколько отличных от «типичных» значений. Ни одно из этих чисел невозможно заранее предсказать каким-либо способом.

Упорядоченная совокупность случайных величин называется *случайным вектором*, каждая случайная величина — *компонентой вектора*, а число их — *размерностью вектора*.

Для n-мерного вектора X будем пользоваться принятым в математике обозначением

$$X=(X_1, X_2, \ldots, X_n). \tag{2.56}$$

Удобно обратиться к геометрическим представлениям и рассматривать упорядоченную совокупность чисел $a\triangleq(a_1, a_2, \ldots$ $\ldots, a_n)$ как координаты точки a в n-мерном пространстве. В этих обозначениях элементарное событие испытания, связанного с наблюдением n-мерного случайного вектора, записывается в виде

$$(X=a)=[(X_1=a_1)\cap(X_2=a_2)\cap\ldots\cap(X_n=a_n)],$$

или короче

$$(X=a)=(X_1=a_1;\,X_2=a_2;\ldots;X_n=a_n).$$

Определение. *n-мерной функцией распределения (совместной функцией распределения) $F(x_1, x_2, \ldots, x_n)$ случайного n-мерного вектора X называется вероятность события* $\bigcap\limits_{i=1}^{n}(X_i<x_i)$:

$$F(x_1, x_2, \ldots, x_n)=P(X_1<x_1;\,X_2<x_2;\ldots;X_n<x_n)=$$
$$=P\left[\bigcap_{i=1}^{n}(X_i<x_i)\right]. \qquad (2.57)$$

Рассмотрим свойства многомерной функции распределения, ограничившись из соображений простоты случаем $n=2$. Пусть $F(x, y)$ — функция распределения случайного вектора (X, Y). Каждая компонента этого вектора может быть дискретной или непрерывной случайной величиной, и ее можно рассматривать отдельно, независимо от второй компоненты. В таком случае все сказанное в предыдущих параграфах об одномерных случайных величинах справедливо для X и Y в отдельности. Совместная функция распределения $F(x, y)$ помимо характеристики индивидуальных свойств компонент содержит характеристику свойств, специфических для совокупности случайных величин.

На основании теоремы умножения вероятностей (1.27) для фиксированной точки (x_k, y_k) имеем

$$F(x_k, y_k)=P(X<x_k;\,Y<y_k)=P(X<x_k)\,P(Y<y_k/X<x_k)=$$
$$=P(Y<y_k)\,P(X<x_k/Y<y_k). \qquad (2.58)$$

Отсюда вытекает следующее.

1. Значение $F(x, y)$ в точке (x_k, y_k) равно (см. рис. 2.8) вероятности попадания случайной точки с координатами (X, Y) в промежуток $S(-\infty\leqslant x<x_k;\ -\infty\leqslant y<y_k)$.

2. Возрастание (уменьшение) x_k или y_k увеличивает (уменьшает) площадь S, что согласно аксиоме сложения вероятностей не может уменьшить (увеличить) вероятность попадания случайной точки в этот промежуток, т. е. функция распределения $F(x, y)$ есть неубывающая функция по каждому аргументу при любом фиксированном значении другого аргумента.

3. Поскольку события $(X<-\infty)$ и $(Y<-\infty)$ невозможные, то $P(X<-\infty)=P(Y<-\infty)=0$, и поэтому для любых x и y, как следует из (2.58),

$$P(-\infty, y)=P(x, -\infty)=P(-\infty, -\infty)=0. \quad (2.59)$$

4. События $(X<+\infty)$ и $(Y<+\infty)$ достоверные, поэтому согласно (1.23)

$$P(X<+\infty/Y<y_k)=P(Y<+\infty/X<x_k)=1.$$

Отсюда и из (2.58) для любых x и y

$$F(x, +\infty)=P(X<x)=F_1(x); \quad (2.60)$$

$$F(+\infty, y)=P(Y<y)=F_2(y) \quad (2.61)$$

и, учитывая (2.15),

$$F(+\infty, -\infty)=1. \quad (2.62)$$

Здесь $F_1(x)$ и $F_2(y)$ — одномерные функции распределения случайных величин X и Y соответственно.

5. Воспользовавшись аксиомой сложения вероятностей, можно получить формулу для вероятности попадания случайной точки в промежуток $S_k(x_{k-1}\leqslant x<x_k;\ y_{k-1}\leqslant y<y_k)$ (см. рис. 2.8):

$$P(x_{k-1}\leqslant X<x_k;\ y_{k-1}\leqslant Y<y_k)=$$
$$=F(x_k, y_k)-F(x_k, y_{k-1})-F(x_{k-1}, y_k)+F(x_k, y_k)\geqslant 0.$$

Это выражение совпадает с (2.47), поскольку вероятность $P(x_{k-1}\leqslant X<x_k;\ y_{k-1}\leqslant Y<y_k)$ есть неотрицательная, аддитивная и нормальная функция множеств. Нормальность вытекает из аксиом вероятностей [4], а (2.62) означает выполнимость условия (2.50), поэтому $F(x, y)$ можно использовать в качестве интегрирующей функции в интегралах Стилтьеса. Из формулы (2.51) получим

$$P(a\leqslant X<b;\ c\leqslant Y<d)=\int_a^b\int_c^d dF(x, y), \quad (2.63)$$

а для области Q более общего вида, но такой, для которой имеет смысл интегрирование по Стилтьесу,

$$P[(X, Y)\in Q]=\iint_Q dF(x, y). \quad (2.64)$$

В частности, для промежутка, образованного одномерными промежутками $[-\infty, x)$, $[-\infty, y)$,

$$F(x, y)=\int_{-\infty}^{x}\int_{-\infty}^{y} dF(t_1, t_2). \quad (2.65)$$

Для (2.60) — (2.62) получим:

$$F_1(x)=\int_{-\infty}^{x}\int_{-\infty}^{\infty} dF(t_1, t_2)=\int_{-\infty}^{x} dF_1(t_1); \tag{2.66}$$

$$F_2(y)=\int_{-\infty}^{\infty}\int_{-\infty}^{y} dF(t_1, t_2)=\int_{-\infty}^{y} dF_2(t_2); \tag{2.67}$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dF(t_1, t_2)=1. \tag{2.68}$$

Если для случайного вектора (X, Y) существует такая неотрицательная функция $f(x, y)$, что для любых x и y

$$F(x, y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(t_1, t_2)\,dt_1 dt_2, \tag{2.69}$$

то эта функция называется двумерной (совместной) *плотностью распределения вероятности* случайного вектора (X, Y). В дальнейшем из чисто практических соображений будем считать, что $f(x, y)$ всегда непрерывна (или кусочно-непрерывна) на всем множестве возможных значений вектора (X, Y), и называть ее просто плотность распределения вектора (X, Y). Случайный вектор, для которого существует плотность распределения, будем называть непрерывным случайным вектором. С учетом оговорки о непрерывности $f(x, y)$ из (2.69) следует, что

$$f(x, y)=\frac{\partial^2 F(x, y)}{\partial x \partial y}. \tag{2.70}$$

На основании (2.54) и (2.70) с точностью до бесконечно малых величин высшего порядка получим

$$\Delta F(x, y)=P\,(x\leqslant X<x+dx;\ y\leqslant Y<y+dy)= \\ =f(x, y)\,dxdy. \tag{2.71}$$

Для непрерывного случайного вектора интегралы Стилтьеса в формулах (2.64)—(2.68) можно преобразовать в соответствии с (2.55) в интегралы Римана. Плотность распределения имеет следующие свойства:

1) по определению $f(x, y)\geqslant 0$ для любых x и y;

2) дифференцируя (2.66), (2.67) и учитывая (2.28), получим:

$$f_1(x)=\frac{dF_1(x)}{dx}=\int_{-\infty}^{\infty} f(x, t_2)\,dt_2; \tag{2.72}$$

$$f_2(y)=\frac{dF_2(y)}{dy}=\int_{-\infty}^{\infty} f(t_1, y)\,dt_1; \tag{2.73}$$

3) переходя к интегралу Римана в (2.64), имеем

$$P[(X, Y) \in Q] = \iint\limits_{Q} f(x, y)\, dx dy, \tag{2.74}$$

в частности,

$$P(a \leqslant X < b;\ c \leqslant Y < d) = \int\limits_{a}^{b} \int\limits_{c}^{d} f(x, y)\, dx dy.$$

Если компоненты вектора (X, Y) — дискретные случайные величины, то интегралы Стилтьеса в выражениях (2.64)—(2.68) преобразуются в соответствии с (2.53) в конечные или бесконечные суммы. Этот случай рассмотрен в следующем параграфе.

Может оказаться, что в выражении (2.58) для любых x и y условные вероятности равны безусловным вероятностям:

$$P(X < x / Y < y) = P(X < x) \quad \text{и} \quad P(Y < y / X < x) = P(Y < y).$$

Совместная функция распределения вектора (X, Y) при этом равна произведению одномерных функций распределения:

$$F(x, y) = P(X < x)\, P(Y < y) = F_1(x)\, F_2(y). \tag{2.75}$$

Случайные величины с функцией распределения такого вида называются *независимыми* случайными величинами.

Если (X, Y) — непрерывный случайный вектор, то, дифференцируя (2.75) сначала по x, а затем по y, получим в соответствии с (2.70)

$$f(x, y) = \frac{\partial F_1(x)}{\partial x} \frac{\partial F_2(y)}{\partial y} = f_1(x)\, f_2(y). \tag{2.76}$$

Относительно n-мерной функции распределения (2.57) кроме того, что сказано о двумерной функции распределения, необходимо добавить следующее.

1. Из n-мерной функции распределения (2.57) можно получить r-мерную функцию распределения $F(x_{i_1}, x_{i_2}, \ldots, x_{i_r})$ случайного вектора $(X_{i_1}, X_{i_2}, \ldots, X_{i_r})$, $r \leqslant n$, составленного из любой комбинации r компонент вектора (2.56). Для этого необходимо все аргументы функции распределения (2.57), номера которых не представлены среди номеров компонент r-мерного вектора, устремить к $+\infty$. В частности, одномерная функция распределения случайной величины X_i, $i = \overline{1, n}$

$$F_i(x) = F(c_1, c_2, \ldots, c_n),$$

где

$$c_k = \begin{cases} +\infty, & \text{если } k \neq i; \\ x, & \text{если } k = i, \quad i = \overline{1, n}. \end{cases}$$

2. Для непрерывного n-мерного вектора (2.70) превращается в смешанную производную n-го порядка:

$$f(x_1, x_2, \ldots, x_n) = \frac{\partial^n F(x_1, x_2, \ldots, x_n)}{\partial x_1 \partial x_2 \ldots \partial x_n}. \tag{2.77}$$

3. Случайные величины $X_1, X_2, \ldots, X_n$ называются *независимыми в совокупности* случайными величинами, если для любого набора целых чисел $\beta_1 < \beta_2 < \ldots < \beta_k$, $\beta_k \leqslant n$ и произвольных $x_{\beta_1}, x_{\beta_2}, \ldots, x_{\beta_k}$ имеет место равенство

$$P[X_{\beta_1} < x_{\beta_1}; X_{\beta_2} \leqslant x_{\beta_2}; \ldots; X_{\beta_k} < x_{\beta_k}] =$$

$$= P(X_{\beta_1} < x_{\beta_1}) P(X_{\beta_2} < x_{\beta_2}) \ldots P(X_{\beta_k} < x_{\beta_k}). \tag{2.78}$$

Из этого определения следует, что для n независимых в совокупности случайных величин

$$F(x_1, x_2, \ldots, x_n) = F_1(x_1) F_2(x_2) \ldots F_n(x_n); \tag{2.79}$$

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1) f_2(x_2) \ldots f_n(x_n). \tag{2.80}$$

Здесь $F_i(x_i)$ и $f_i(x_i)$ — функция распределения и плотность распределения (если она существует) случайной величины X_i. Можно доказать и обратное: если выполняется (2.79) или (2.80), то выполняется и (2.78).

Пример 2.8. Случайный вектор (X, Y) распределен равномерно в промежутке $a \leqslant x < b$; $c \leqslant y < d$, если его совместная функция распределения имеет вид

$$F(x, y) = \begin{cases} 0 & \text{для } -\infty \leqslant x < a;\ -\infty \leqslant y < c; \\ \dfrac{x-a}{b-a} \dfrac{y-c}{d-c} & \text{для } a \leqslant x < b;\ c \leqslant y < d; \\ \dfrac{x-a}{b-a} & \text{для } a \leqslant x < b;\ d \leqslant y < +\infty; \\ \dfrac{y-c}{d-c} & \text{для } b \leqslant x < +\infty;\ c \leqslant y < d. \end{cases} \tag{2.81}$$

Сравнивая (2.81) с (2.7), видим, что $F(x, y) = F_1(x) F_2(y)$ и в промежутке $a \leqslant x < b$; $c \leqslant y < d$

$$f(x, y) = \frac{1}{b-a} \frac{1}{d-c}. \tag{2.82}$$

Таким образом, (X, Y) есть непрерывный случайный вектор с независимыми компонентами.

Пример 2.9. Случайный вектор (X, Y) распределен нормально, если его функция распределения имеет вид

$$F(x, y) = C\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y}\exp\left\{-\left[\frac{(t_1-m_1)^2}{2A^2} - r\,\frac{(t_1-m_1)(t_2-m_2)}{AB} + \right.\right.$$

$$\left.\left. + \frac{(t_2-m_2)^2}{2B^2}\right]\right\}dt_1 dt_2, \qquad (2.83)$$

где m_1 и m_2 — любые вещественные числа; A и B — неотрицательные вещественные числа; $-1<r<1$. Постоянная C должна быть определена так, чтобы при известных остальных параметрах выполнялось (2.62). Преобразуем показатель экспоненты, выделив в нем квадрат суммы двух слагаемых:

$$\frac{(t_1-m_1)^2}{2A^2} - \frac{2r\,(t_1-m_1)(t_2-m_2)}{\sqrt{2}A\sqrt{2}B} + \frac{(t_2-m_2)^2}{2B^2} =$$

$$= \frac{(t_1-m_1)^2}{2A^2} - \frac{2r\,(t_1-m_1)(t_2-m_2)}{\sqrt{2}A\sqrt{2}B} - \frac{r^2(t_1-m_1)^2}{2A^2} +$$

$$+ \frac{r^2(t_1-m_1)^2}{2A^2} + \frac{(t_2-m_2)^2}{2B^2} = (1-r^2)\frac{(t_1-m_1)^2}{2A^2} +$$

$$+ \left[\frac{r\,(t_1-m_1)}{\sqrt{2}A} - \frac{(t_2-m_2)}{\sqrt{2}B}\right]^2;$$

$$F(x, y) = C\int\limits_{-\infty}^{x}\exp\left\{-(1-r^2)\frac{(t_1-m_1)^2}{2A^2}\right\}\int\limits_{-\infty}^{y}\exp\left\{-\left[\frac{r\,(t_1-m_1)}{\sqrt{2}A} + \right.\right.$$

$$\left.\left. + \frac{t_2-m_2}{\sqrt{2}B}\right]^2\right\}dt_1 dt_2. \qquad (2.84)$$

Условие (2.62) для этого случая принимает вид

$$C\int\limits_{-\infty}^{\infty}\exp\left\{-(1-r^2)\frac{(t_1-m_1)^2}{2A^2}\right\}\int\limits_{-\infty}^{\infty}\exp\left\{-\left[\frac{r\,(t_1-m_1)}{\sqrt{2}A} + \right.\right.$$

$$\left.\left. + \frac{t_2+m_2}{\sqrt{2}B}\right]^2\right\}dt_1 dt_2 = 1.$$

Заменив переменные

$$\frac{\sqrt{1-r^2}\,(t_1-m_1)}{\sqrt{2}A} = z_1; \quad dt_1 = dz_1\frac{\sqrt{2}A}{\sqrt{1-r^2}}; \qquad (2.85)$$

$$\frac{r\,(t_1-m_1)}{\sqrt{2}A} + \frac{t_2-m_2}{\sqrt{2}B} = z_2; \quad dt_2 = dz_2\sqrt{2}B, \qquad (2.86)$$

получим

$$C\int\limits_{-\infty}^{\infty}\exp(-z_1^2)\int\limits_{-\infty}^{\infty}\exp(-z_2^2)\frac{2AB}{\sqrt{1-r^2}}dz_1 dz_2 = 1.$$

Поскольку оба интеграла согласно (2.18) равны $\sqrt{\pi}$, то из этого равенства следует, во-первых, что (2.83) имеет смысл только при $r \neq \pm 1$ и, во-вторых, что $C = \frac{\sqrt{1-r^2}}{2\pi AB}$.

Если $r \neq \pm 1$, то введем обозначения

$$A = \sigma_1 \sqrt{1-r^2} \quad \text{и} \quad B = \sigma_2 \sqrt{1-r^2}. \tag{2.87}$$

Тогда (2.83) принимает следующий вид:

$$F(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \int_{-\infty}^{x} \int_{-\infty}^{y} \exp\left\{ -\frac{1}{2(1-r^2)} \left[\frac{(t_1-m_1)^2}{\sigma_1^2} - \right.\right.$$
$$\left.\left. -2r\frac{(t_1-m_1)(t_2-m_2)}{\sigma_1\sigma_2} + \frac{(t_2-m_2)^2}{\sigma_2^2} \right] \right\} dt_1 dt_2. \tag{2.88}$$

Используя (2.66) и (2.67), можно из (2.88) получить одномерные законы распределения для X и Y:

$$F_1(x) =$$
$$= C \int_{-\infty}^{x} \exp\left[-(1-r^2)\frac{(t_1-m_1)^2}{2A^2} \right] \int_{-\infty}^{\infty} \exp(-z_2^2)\, dz_2 \sqrt{2} B\, dt_1 =$$
$$= \sqrt{2\pi} CB \int_{-\infty}^{x} \exp\left[-(1-r^2)\frac{(t_1-m_1)^2}{2A^2} \right] dt_1,$$

или

$$F_1(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \int_{-\infty}^{x} \exp\left[-\frac{(t_1-m_1)^2}{2\sigma_1^2} \right] dt_1. \tag{2.89}$$

Аналогично

$$F_2(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \int_{-\infty}^{y} \exp\left[-\frac{(t_2-m_2)^2}{2\sigma_2^2} \right] dt_2. \tag{2.90}$$

Сравнивая эти выражения с (2.9), видим, что компоненты двумерного нормально распределенного вектора имеют нормальные законы распределения с параметрами m_1, σ_1 и m_2, σ_2 соответственно.

Из определения (2.83) непосредственно следует существование для вектора (X, Y) плотности распределения

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \exp\left\{ -\frac{1}{2(1-r^2)} \left[\frac{(x-m_1)^2}{\sigma_1^2} - \right.\right.$$
$$\left.\left. -2r\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2} + \frac{(y-m_2)^2}{\sigma_2^2} \right] \right\}. \tag{2.91}$$

При $r=0$

$$f(x, y)=\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(x-m_1)^2}{2\sigma_1^2}\right]\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left[-\frac{(y-m_2)^2}{2\sigma_2^2}\right]=$$
$$=f_1(x)f_2(y), \tag{2.92}$$

что означает согласно (2.80) независимость случайных величин X и Y.

§ 2.6. Условные законы распределения

До сих пор предполагалось, что в определениях функций распределения (2.1) и (2.57) справа от знака равенства стоят безусловные вероятности некоторых случайных событий, в связи с чем эти функции распределения можно назвать безусловными. Однако такое ограничение необязательно. Если вероятность случайного события B не равна нулю, то

$$P(X<x/B)=F(x/B) \tag{2.93}$$

есть условная функция распределения случайной величины X при гипотезе B. Аналогичное определение можно ввести и для n-мерной функции распределения. Все, что до сих пор было сказано о безусловных функциях распределения, справедливо и для условных.

С условными законами распределения приходится иметь дело при изучении случайных векторов, когда событие B состоит в том, что некоторые компоненты вектора имеют заранее заданные значения. Рассмотрим подробнее случаи дискретного и непрерывного двумерного вектора (X, Y).

Пусть X и Y — дискретные случайные величины, причем $A=\{a_1, a_2, \ldots, a_n\}$ — множество возможных значений X, $B=\{b_1, b_2, \ldots, b_m\}$ — множество возможных значений Y и

$$p_{ij}=P(X=a_i;\ Y=b_j),\quad i=\overline{1, n},\ j=\overline{1, m} \tag{2.94}$$

есть вероятность наблюдать вектор (a_i, b_j).

Совокупность nm вероятностей p_{ij}, представленная в табл. 2.2, называется совместным рядом распределения случайного вектора (X, Y). Совместный ряд распределения полностью характеризует случайные свойства вектора.

Представим (2.94), учитывая (1.27), в следующем виде:

$$p_{ij}=P(X=a_i)P(Y=b_j/X=a_i)=$$
$$=P(Y=b_j)P(X=a_i/Y=b_j). \tag{2.95}$$

Событие $X=a_i$ может наблюдаться только совместно с одним из событий $Y=b_j$, $j=\overline{1, m}$, которые попарно несовместны и составляют полную группу событий. Эти условия позволяют использовать формулу полной вероятности (1.32) для вычисления вероятности события $X=a_i$:

$$P(X=a_i)=\sum_{j=1}^{m}P(Y=b_j)\,P(X=a_i/Y=b_j)=\sum_{j=1}^{m}p_{ij}. \qquad (2.96)$$

Повторяя рассуждения, получим

$$P(Y=b_j)=\sum_{i=1}^{n}p_{ij}. \qquad (2.97)$$

Условные вероятности можем получить из (2.95):

$$P(X=a_i/Y=b_j)=\frac{p_{ij}}{P(X=a_i)}=\frac{p_{ij}}{\sum\limits_{j=1}^{m}p_{ij}}; \qquad (2.98)$$

$$P(Y=b_j/X=a_i)=\frac{p_{ij}}{P(Y=b_j)}=\frac{p_{ij}}{\sum\limits_{i=1}^{n}p_{ij}}. \qquad (2.99)$$

Совокупность вероятностей (2.98) для $i=\overline{1,n}$ составляет условный ряд распределения вектора X при гипотезе $Y=b_j$. Аналогично совокупность вероятностей (2.99) для $j=\overline{1,m}$ составляет условный ряд распределения вектора Y при гипотезе $X=a_i$.

Из (2.97) имеем

$$\sum_{j=1}^{m}\sum_{i=1}^{n}p_{ij}=\sum_{j=1}^{m}P(Y=b_j)=1, \qquad (2.100)$$

поскольку $Y=b_j$, $j=\overline{1,m}$ — полная группа несовместных событий. Это равенство есть частный случай свойства (2.62). Совокупность вероятностей (2.96), представленная в последнем столбце табл. 2.2, является одномерным рядом распределения случайной величины X. Функция распределения X выражается формулой

$$F_1(x)=\sum_{a_i<x}P(X=a_i)=\sum_{a_i<x}\sum_{j=1}^{m}p_{ij}, \qquad (2.101)$$

которая есть частный случай формулы (2.66). Аналогичное замечание справедливо и для Y.

Формула (2.63) превращается в данном случае в двойную сумму:

$$P(a\leqslant X<b;\ c\leqslant Y<d)=\sum_{i}\sum_{j}p_{ij}. \qquad (2.102)$$

Таблица 2.2

	X	Y						
		b_1	b_2	b_3	...	b_{m-1}	b_m	$\sum_j p_{ij}$
		1	2	3	...	$m-1$	m	
1	a_1	p_{11}	p_{12}	p_{13}	...	$p_{1\,m-1}$	p_{1m}	$P(X=a_1)$
2	a_2	p_{21}	p_{22}	p_{23}	...	$p_{2\,m-1}$	p_{2m}	$P(X=a_2)$
3	a_3	p_{31}	p_{32}	p_{33}	...	$p_{3\,m-1}$	p_{3m}	$P(X=a_3)$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$n-1$	a_{n-1}	$p_{n-1\,1}$	$p_{n-1\,2}$	$p_{n-1\,3}$	...	$p_{n-1\,m-1}$	$p_{n-1\,m}$	$P(X=a_{n-1})$
n	a_n	p_{n1}	p_{n2}	p_{n3}	...	$p_{n\,m-1}$	p_{nm}	$P(X=a_n)$
	$\sum_i p_{ij}$	$P(Y=b_1)$	$P(Y=b_2)$	$P(Y=b_3)$	...	$P(Y=b_{m-1})$	$P(Y=b_m)$	

Здесь суммирование распространяется на все те p_{ij}, для которых выполняются неравенства $a \leqslant a_i < b$; $c \leqslant b_j < d$.

Теперь обратимся к непрерывному случайному вектору (X, Y) с функцией распределения $F(x, y)$ и плотностью распределения $f(x, y)$.

Применим теорему умножения вероятностей (1.27) к (2.71):

$$f(x, y)\,dxdy =$$
$$= P(x \leqslant X < x + dx)\,P(y \leqslant Y < y + dy/x \leqslant X < x + dx) =$$
$$= P(y \leqslant Y < y + dy)\,P(x \leqslant X < x + dx/y \leqslant Y <$$
$$< y + dy). \qquad (2.103)$$

В пределе событие $(x \leqslant X < x + dx)$ можно считать эквивалентным событию $(X = x)$. Точно так же событие $(y \leqslant Y < y + dy)$ эквивалентно $(Y = y)$. Поэтому в пределе

$$P(x \leqslant X < x + dx/y \leqslant Y < y + dy) =$$
$$= P(x \leqslant X < x + dx/Y = y) = f_1(x/y)\,dx \qquad (2.104)$$

и

$$P(y \leqslant Y < y + dy/X = x) = f_2(y/x)\,dy. \qquad (2.105)$$

Здесь символом $f_1(x/y)$ обозначена условная (одномерная) плотность распределения случайной величины X при гипотезе $Y=y$, т. е. при условии, что случайная величина Y имеет фиксированное значение y. Аналогично $f_2(y/x)$ обозначает условную (одномерную) плотность распределения случайной величины Y при гипотезе $X=x$. Если $f_1(x)$ и $f_2(y)$ — безусловные плотности распределения соответственно X и Y, то (2.103) можно представить в следующем виде:

$$f(x, y) = f_1(x) f_2(y/x) = f_2(y) f_1(x/y). \tag{2.106}$$

Отсюда

$$f_1(x/y) = \frac{f(x, y)}{f_2(y)}; \tag{2.107}$$

$$f_2(y/x) = \frac{f(x, y)}{f_1(x)}. \tag{2.108}$$

А если учесть (2.73) и (2.72), то

$$f_1(x/y) = \frac{f(x, y)}{\int\limits_{-\infty}^{\infty} f(x, y)\,dx}; \tag{2.109}$$

$$f_2(y/x) = \frac{f(x, y)}{\int\limits_{-\infty}^{\infty} f(x, y)\,dy}. \tag{2.110}$$

Нетрудно видеть, что (2.110) и (2.109) подобны (2.98) и (2.99). Полезно эти формулы сравнить и с (1.25), (1.26).

Если случайные величины X и Y независимы, то из формул (2.76) и (2.106) следует

$$f_1(x/y) = f_1(x) \quad \text{и} \quad f_2(y/x) = f_2(y), \tag{2.111}$$

т. е. условные плотности распределения равны безусловным.

Пример 2.10. Случайный вектор (X, Y) распределен по нормальному закону (2.88). Найдем условный закон распределения случайной величины X при гипотезе $Y=y_0$. Дифференцирование (2.90) дает плотность распределения случайной величины Y

$$f_2(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(y-m_2)^2}{2\sigma_2^2}\right].$$

Подставив это выражение и (2.91) при $y=y_0$ в (2.107), получим условную плотность распределения величины X

$$f_1(x/y_0) = \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-r^2}} \exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_1)^2}{\sigma_1^2} - \right.\right.$$

$$\left.\left. - 2r\frac{(x-m_1)\cdot(y_0-m_2)}{\sigma_1\sigma_2} + \frac{(y_0-m_2)^2}{\sigma_2^2}\right] + \frac{(y_0-m_2)^2}{2\sigma_2^2}\right\}.$$

Преобразуем показатель экспоненты

$$-\frac{1}{2(1-r^2)}\left[\frac{(x-m_1)^2}{\sigma_1^2}-2r\frac{(x-m_1)(y_0-m_2)}{\sigma_1\sigma_2}+\frac{(y_0-m_2)^2}{\sigma_2^2}\right]+$$

$$+\frac{(y_0-m_2)^2}{2\sigma_2^2}=-\frac{1}{2(1-r^2)}\left[\frac{(x-m_1)^2}{\sigma_1^2}-2r\frac{(x-m_1)(y_0-m_2)}{\sigma_1\sigma_2}+\right.$$

$$\left.+\frac{r^2(y_0-m_2)^2}{\sigma_2^2}\right]=-\frac{1}{2\sigma_1^2(1-r^2)}\left[x_1-m_1-\frac{\sigma_1 r}{\sigma_2}(y_0-m_2)\right]^2.$$

Введем обозначения

$$\sigma_1\sqrt{1-r^2}=\sigma \quad \text{и} \quad m_1+\frac{\sigma_1 r}{\sigma_2}(y_0-m_2)=m. \qquad (2.112)$$

Окончательно получим

$$f(x/y_0)=\frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-r^2}}\exp\left\{-\frac{1}{2\sigma_1^2(1-r^2)}\left[x_1-m_1-\right.\right.$$

$$\left.\left.-\frac{\sigma_1 r}{\sigma_2}(y_0-m_2)\right]^2\right\}=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-m)^2}{2\sigma^2}\right]. \qquad (2.113)$$

Это выражение показывает, что условный закон распределения X при гипотезе $Y=y_0$ является нормальным законом распределения, причем параметры этого закона зависят от параметров совместного закона распределения и значения y_0.

§ 2.7. Законы распределения функций случайных аргументов

В приложениях теории вероятностей возникают задачи определения по известному закону распределения совокупности случайных величин закона распределения другой совокупности случайных величин, которая получается из первой некоторым заданным функциональным преобразованием. Так, например, может возникнуть вопрос о законе распределения $\sin X$, e^X или $\sum_{k=0}^{n} a_k X^k$, если известен закон распределения случайной величины X. Мы ограничимся рассмотрением нескольких частных, но важных случаев для непрерывных случайных величин.

Пусть G — множество возможных значений величины X с функцией распределения $F_x(x)$ и плотностью распределения $f_x(x)$. На множестве G определена вещественная функция $y=g(x)$, значения которой образуют множество H. Говорят, что $g(x)$ отображает множество G на множество H, множество H является образом множества G, а множество G — прообразом H.

Если $Y=g(X)$ — случайная величина в смысле определения из § 2.1, то обозначим ее функцию распределения символом $F_y(y)$, а плотность распределения (если таковая существует) $f_y(y)$. По определению (2.1)

$$F_y(y)=P(Y<y)=P(g(X)<y)=P(X\in A_y), \qquad (2.114)$$

где $A_y\subset G$ — множество тех значений x, для которых $g(x)<y$ (значение y фиксировано).

Рис. 2.9 иллюстрирует сказанное. Здесь $G=[a, b]$, $H=[c, d]$, $A_y=[a, x_1)\cup(x_2, x_3)\cup(x_4, b]$. Точки x_1, x_2, x_3, x_4 следует исключить из A_y, если $g(x)$ — непрерывная функция, поскольку в этих точках неравенство $g(x_i)<y$ не выполняется. Очевидно, события $(X\in A_y)$ и $(Y<y)$ эквивалентны, поэтому справедливо (2.114).

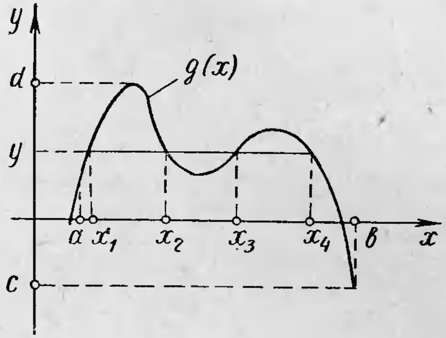

Рис. 2.9

Пусть $g(x)$ строго монотонна и имеет непрерывную не обращающуюся в нуль производную во всех точках G. При этих условиях ее обратная функция, которую обозначим символом $x=g^{-1}(y)$, определена и однозначна во всех точках H. Возьмем в H две точки y и $y+\Delta y$, $\Delta y>0$. Событие $y<Y<y+\Delta y$ эквивалентно событию $g^{-1}(y)<X<g^{-1}(y+\Delta y)$, если $g(x)$ — возрастающая функция, и событию $g^{-1}(y+\Delta y)<X<g^{-1}(y)$, если $g(x)$ — убывающая функция. Поэтому

$$P(y\leqslant Y<y+\Delta y)=P[g^{-1}(y)\leqslant X<g^{-1}(y+\Delta y)],$$

или

$$P(y\leqslant Y<y+\Delta y)=P[g^{-1}(y+\Delta y)<X\leqslant g^{-1}(y)].$$

Объединим эти равенства и воспользуемся формулой (2.11):

$$P(y\leqslant Y<y+\Delta y)=$$
$$=\pm[F_x(g^{-1}(y+\Delta y))-F_x(g^{-1}(y))]. \qquad (2.115)$$

Знаки перед скобкой следует выбирать так, чтобы разность была положительной.

Разделим обе части (2.115) на Δy, умножим и разделим правую часть на $g^{-1}(y+\Delta y)-g^{-1}(y)$ и перейдем к пределу при $\Delta y\to 0$

$$\lim_{\Delta y\to 0}\frac{P(y\leqslant Y<y+\Delta y)}{\Delta y}=$$
$$=\lim_{\Delta y\to 0}\left[\frac{F_x(g^{-1}(y+\Delta y))-F_x(g^{-1}(y))}{g^{-1}(y+\Delta y)-g^{-1}(y)}\right]\left[\frac{g^{-1}(y+\Delta y)-g^{-1}(y)}{\Delta y}\right].$$

При указанных выше свойствах функции $g(x)$ предел правой части равенства существует и является непрерывной функцией y. Предел левой части равен значению плотности распределения случайной величины Y в точке y. Поэтому

$$f_y(y) = \pm f_x(g^{-1}(y)) \frac{dg^{-1}(y)}{dy} = f_x(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| \quad (2.116)$$

и

$$P(a \leqslant Y < b) = \int_a^b f_y(y)\,dy =$$

$$= \int_{g^{-1}(a)}^{g^{-1}(b)} f_x(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| dy. \quad (2.117)$$

Равенство (2.114) можно обобщить на многомерные случайные величины. Действительно, пусть X есть n-мерный случайный вектор, а $g(x) = y$ обозначает упорядоченную совокупность m функций от n аргументов (так называемую вектор-функцию):

$$\left.\begin{array}{l} y_1 = g_1(x_1, x_2, \ldots, x_n); \\ y_2 = g_2(x_1, x_2, \ldots, x_n); \\ \cdots\cdots\cdots\cdots \\ y_m = g_m(x_1, x_2, \ldots, x_n), \end{array}\right\} \quad (2.118)$$

каждая из которых определена на множестве n-мерных точек G. Точка $x \in G$ преобразуется системой (2.118) в точку m-мерного пространства $y = (y_1, y_2, \ldots, y_m)$. Образ множества G в m-мерном пространстве обозначим символом H. Символом Y обозначим совокупность случайных величин $(Y_1, Y_2, \ldots$ $\ldots, Y_m) = Y$, где $Y_i = g_i(X_1, X_2, \ldots, X_n)$, $i = \overline{1, m}$. Неравенство $Y < y$ будем понимать как неравенство

$$(Y < y) = (Y_1 < y_1) \cap (Y_2 < y_2) \cap \ldots \cap (Y_m < y_m).$$

При такой трактовке символов по определению (2.57) выражение (2.114) является m-мерной функцией распределения случайного вектора Y.

Рассмотрим подробнее случай двумерного непрерывного вектора (X_1, X_2) с плотностью распределения $f(x_1, x_2)$. Пусть функции

$$y_1 = g_1(x_1, x_2) \quad \text{и} \quad y_2 = g_2(x_1, x_2) \quad (2.119)$$

отображают G — множество значений вектора (X_1, X_2) на множество H плоскости y_1y_2 (рис. 2.10). Предположим, что это отображение удовлетворяет следующим условиям.

1. Отображение взаимно однозначно, т. е. существуют решения уравнений (2.119) относительно x_1 и x_2:

$$x_1 = g^{-1}(y_1, y_2); \quad x_2 = g_2^{-1}(y, y_2), \tag{2.120}$$

однозначные на всем множестве H.

2. Функции (2.119) и (2.120) непрерывны и имеют непрерывные частные производные первого порядка.

3. Якобиан

$$\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = \begin{vmatrix} \frac{\partial g_1(x_1, x_2)}{\partial x_1} & \frac{\partial g_1(x_1, x_2)}{\partial x_2} \\ \frac{\partial g_2(x_1, x_2)}{\partial x_1} & \frac{\partial g_2(x_1, x_2)}{\partial x_2} \end{vmatrix} \tag{2.121}$$

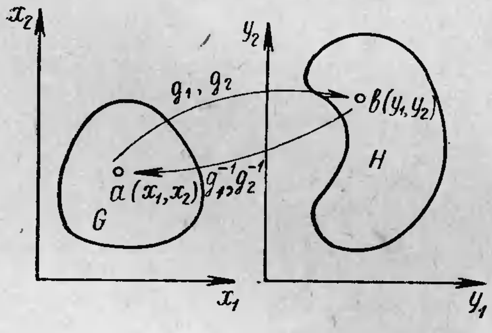

Рис. 2.10

всюду в G отличен от нуля и, следовательно, сохраняет постоянный знак, так как входящие в якобиан производные непрерывны.

Якобиан обратного преобразования (2.120) связан с (2.121) уравнением

$$\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = 1.$$

Как известно из курса математического анализа [2], при выполнении указанных условий для непрерывной в G функции $\psi(x_1, x_2)$ имеет место равенство

$$\iint\limits_G \psi(x_1, x_2)\, dx_1 dx_2 = \iint\limits_H \psi[g_1^{-1}(y_1, y_2), g_2^{-1}(y_1, y_2)] \times$$

$$\times \left| \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} \right| dy_1 dy_2. \tag{2.122}$$

Возьмем в H произвольную точку y и некоторую ее окрестность $\sigma_1 \subset H$. Пусть в G этой точке соответствует точка x, а окрестности σ_1 — окрестность σ_2 (такая окрестность существует, так как преобразование непрерывно по условию). События $Y \in \sigma_1$ и $X \in \sigma_2$ эквивалентны, поэтому $P(Y \in \sigma_1) = P(X \in \sigma_2)$ или

$$\iint\limits_{\sigma_1} f_y(y_1, y_2)\, dy_1 dy_2 = \iint\limits_{\sigma_2} f_x(x_1, x_2)\, dx_1 dx_2,$$

где $f_y(y_1, y_2)$ — плотность распределения Y.

Заменим переменные в правой части этого равенства в соответствии с (2.122)

$$\iint\limits_{\sigma_1} f_y(y_1, y_2)\, dy_1 dy_2 = \iint\limits_{\sigma_1} f_x(g_1^{-1}, g_2^{-1}) \left| \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} \right| dy_1 dy_2$$

и далее, применяя теорему о среднем, получим

$$\sigma_1 f_y(y_1^*, y_2^*) = \sigma_1 \left[f_x(g_1^{-1}, g_2^{-1}) \left| \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} \right| \right]_{(y_1^*, y_2^*)},$$

где (y_1^*, y_2^*) — некоторая точка окрестности σ_1. Если эта окрестность стягивается к точке $y = (y_1, y_2)$, то в силу непрерывности функций получаем значение плотности распределения случайного вектора Y в этой точке

$$f_y(y_1, y_2) = f_x\,[g_1^{-1}(y_1, y_2), g_2^{-1}(y_1, y_2)] \left| \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} \right|_{(y_1, y_2)}. \quad (2.123)$$

Нетрудно видеть, что (2.116) является частным случаем (2.123).

Найдем плотность распределения $f_y(y)$ суммы двух случайных величин X_1 и X_2 с совместной плотностью распределения $f_x(x_1, x_2)$. С этой целью воспользуемся преобразованием

$$y = y_1 = x_1 + x_2; \qquad y_2 = x_2,$$

которое удовлетворяет перечисленным выше условиям. Обратное преобразование

$$x_1 = y_1 - y_2; \qquad x_2 = y_2$$

имеет якобиан

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$$

По (2.123) находим плотность распределения вектора (Y_1, Y_2)

$$f_y(y_1, y_2) = f_x(y_1 - y_2, y_2).$$

На основании (2.72)

$$f_1(y_1) = \int\limits_{-\infty}^{\infty} f_x(y_1 - y_2, y_2)\, dy_2.$$

Заменив $y_1=y$ и $y_2=x_2$, окончательно имеем

$$f(y)=\int\limits_{-\infty}^{\infty} f_x(y-x_2, x_2)\,dx_2. \qquad (2.124)$$

Несколько изменив рассуждение, приходим к выражению

$$f(y)=\int\limits_{-\infty}^{\infty} f_x(x_1, y-x_1)\,dx_1. \qquad (2.125)$$

Для независимых величин (X_1, X_2) с плотностью распределения $f(x_1, x_2)=f_1(x_1)f_2(x_2)$ получим

$$f(y)=\int\limits_{-\infty}^{\infty} f_1(y-x_2)\,f_2(x_2)\,dx_2=\int\limits_{-\infty}^{\infty} f_1(x_1)\,f_2(y-x_1)\,dx_1. \qquad (2.126)$$

Используя (2.27), находим функцию распределения случайной величины $Y=X_1+X_2$

$$F(y)=\int\limits_{-\infty}^{y}\int\limits_{-\infty}^{\infty} f_1(t-x_2)\,f_2(x_2)\,dx_2dt=$$

$$=\int\limits_{-\infty}^{y}\int\limits_{-\infty}^{\infty} f_1(x_1)\,f_2(t-x_1)\,dx_1dt. \qquad (2.127)$$

В этом выражении можно перейти к интегралам Стилтьеса

$$F(y)=\int\limits_{-\infty}^{\infty} F_1(y-x_2)\,dF_2(x_2)=\int\limits_{-\infty}^{\infty} F_2(y-x_1)\,dF_1(x_1). \qquad (2.128)$$

В таком виде эта формула оказывается справедливой и для дискретных независимых случайных величин с функциями распределения $F_1(x)$ и $F_2(x)$.

Функция $\psi(y)$, получаемая в результате интегрального преобразования

$$\psi(y)=\int\limits_{-\infty}^{\infty}\varphi(y-x)\,g(x)\,dx \quad \text{или} \quad \psi(y)=\int\limits_{-\infty}^{\infty}\varphi(y-x)\,dg(x),$$

как и само преобразование, называется *сверткой* (композицией) функции $\varphi(x)$ по $g(x)$. Это преобразование записывается в виде $\psi(y)=\varphi(x) * g(x)$. Таким образом, из (2.126) и (2.128) следует, что функция (плотность) распределения суммы независимых случайных величин есть свертка (композиция) функций (плотностей) распределения слагаемых:

$$F(y)=F_1(x_1) * F_2(x_2). \qquad (2.129)$$

Найдем функцию распределения произведения двух случайных величин, для чего воспользуемся преобразованием

$y_1 = x_1 x_2$; $y_2 = x_2$. Обратное преобразование $x_1 = y_1/y_2$; $x_2 = y_2$ (значение $y_2 = 0$ должно быть исключено) имеет якобиан

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{1}{y_2} & -\frac{1}{y_2^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{y_2},$$

поэтому по (2.123) имеем

$$f_y(y_1, y_2) = f_x\left(\frac{y_1}{y_2}, y_2\right)\left|\frac{1}{y_2}\right|.$$

Отсюда, переходя к обозначениям $y_2 = x_2$ и $y_1 = y$, получим плотность распределения произведения случайных величин

$$f(y) = \int_{-\infty}^{0} f_x\left(\frac{y}{x_2}, x_2\right)\frac{dx_2}{|x_2|} + \int_{0}^{\infty} f_x\left(\frac{y}{x_2}, x_2\right)\frac{dx_2}{|x_2|} =$$

$$= -\int_{-\infty}^{0} f_x\left(\frac{y}{x_2}, x_2\right)\frac{dx_2}{x_2} + \int_{0}^{\infty} f_x\left(\frac{y}{x_2}, x_2\right)\frac{dx_2}{x_2}. \quad (2.130)$$

Найдем плотность распределения $f(y)$ частного $Y = X_1/X_2$ для случайных величин X_1 и X_2 с совместной плотностью распределения $f_x(x_1, x_2)$, изменив несколько ход рассуждения. На рис. 2.11 показаны прямая $x_1/x_2 = y$ для фиксированного y и области, в которых выполняется неравенство $x_1/x_2 < y$. Событие $(Y < y)$ наступит в том и только в том случае, когда случайная точка (X_1, X_2) попадет в заштрихованную область. Интегрируя $f_x(x_1, x_2)$ по этой области, получим функцию распределения Y

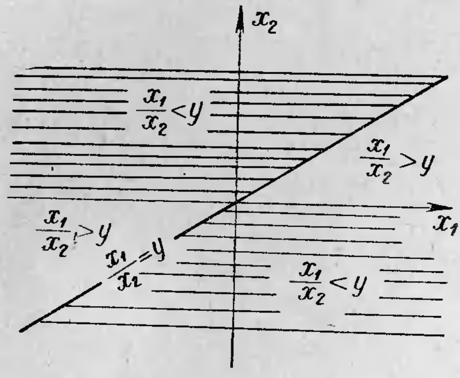

Рис. 2.11

$$F(y) = P(Y < y) = \int_{0}^{\infty}\int_{-\infty}^{x_2 y} f_x(x_1, x_2)\,dx_1 dx_2 +$$

$$+ \int_{-\infty}^{0}\int_{x_2 y}^{\infty} f_x(x_1, x_2)\,dx_1 dx_2.$$

Отсюда плотность распределения частного двух случайных величин

$$f(y)=\frac{dF(y)}{dy}=\int_0^\infty x_2 f_x[(x_2y),x_2]\,dx_2-$$

$$-\int_{-\infty}^0 x_2 f_x[(x_2y),x_2]\,dx_2. \tag{2.131}$$

Для независимых случайных величин

$$f(y)=\int_0^\infty x_2f_1(x_2y)f_2(x_2)\,dx_2-\int_{-\infty}^0 x_2f_1(x_2y)f_2(x_2)\,dx_2. \tag{2.132}$$

Конечно, этот же результат получается и по формуле (2.123).

Пример 2.11. Найдем закон распределения частного двух случайных величин (X_1, X_2), распределенных по нормальному закону с параметрами $m_1=m_2=0$, $r=0$, $\sigma_1=\sigma_2=1$, т. е. по закону

$$f(x_1,x_2)=\frac{1}{2\pi}\exp\left[-\frac{1}{2}(x_1^2+x_2^2)\right].$$

По (2.131) имеем

$$f(y)=\frac{1}{2\pi}\int_0^\infty x_2\exp\left\{-\frac{1}{2}[(x_2y)^2+x_2^2]\right\}dx_2-$$

$$-\frac{1}{2\pi}\int_{-\infty}^0 x_2\exp\left\{-\frac{1}{2}[(x_2y)^2+x_2^2]\right\}dx_2.$$

Заменим переменные

$$z=\frac{1}{2}x^2(1+y^2);\quad dz=x(1+y^2)\,dx$$

$$f(y)=\frac{1}{2\pi(1+y^2)}\left[\int_0^\infty\exp(-z)\,dz-\int_{-\infty}^0\exp(-z)\,dz\right].$$

Окончательно

$$f(y)=\frac{1}{\pi}\cdot\frac{1}{1+y^2}. \tag{2.133}$$

Этот закон распределения называется законом Коши. Функцию распределения его получим по формуле (2.27):

$$F_y(y)=\frac{1}{\pi}\int_{-\infty}^y\frac{1}{1+t^2}\,dt=\frac{1}{\pi}\left[\operatorname{arctg}y+\frac{\pi}{2}\right].$$

Заметим следующее. Если X — угол с равномерным распределением в промежутке $[-\pi/2, \pi/2]$, т. е., согласно (2.8),

$$f_x(x) = \frac{1}{\pi}; \qquad -\frac{\pi}{2} \leqslant x < \frac{\pi}{2},$$

а $Y = \operatorname{tg} X$, то по формуле (2.116)

$$f_y(y) = f_x(\operatorname{arctg} Y) \left| \frac{d}{dy} \operatorname{arctg} Y \right| = \frac{1}{\pi} \left| \frac{1}{1+y^2} \right| = \frac{1}{\pi} \cdot \frac{1}{1+y^2}.$$

Таким образом, по закону Коши распределен й тангенс угла, если последний имеет равномерное распределение в промежутке $[-\pi/2, \pi/2]$.

ГЛАВА 3

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

§ 3.1. Числовые характеристики случайных величин

Законы распределений — функции распределений, плотности распределений, ряды распределений — полные исчерпывающие характеристики случайных величин. Но практическое использование этих характеристик осложняется, во-первых, громоздким математическим аппаратом и, во-вторых, трудностями, связанными с выявлением этих законов. В то же время постановка задачи часто не требует исчерпывающих характеристик случайных величин, а позволяет обойтись менее полными, но более простыми характеристиками, отражающими обобщенно, суммарно и компактно наиболее существенные вероятностные свойства случайных величин, а точнее их законов распределения. Так, например, при расчете транзисторной схемы не пользуются законами распределения электрических параметров компонентов схемы, хотя каждый из параметров любого компонента — случайная величина. В справочниках по транзисторам (так же как и по любым другим радиотехническим изделиям) указываются не законы распределения параметров транзистора, а их «типичные» значения и пределы возможных отклонений действительных значений от «типичных». В табл. 3.1. приведены справочные данные транзистора П13 [6]. В подавляющем большинстве случаев этих данных достаточно для расчета схемы.

Существует большое количество различных обобщенных характеристик случайных величин, имеющих различное назначение и различные области применения. Каждая из них получается по определенным правилам из законов распреде-

Таблица 3.1

Основные параметры	Значения		
	минимальное	типичное	максимальное
h_{11} *ом*	24	28	32
h_{12} *ом*	10^{-4}	10^{-3}	$5\cdot10^{-3}$
h_{21} *ом*	—0,92	—0,94	—0,98
h_{22} *мкмо*	0,3	1,0	3,3
$C_к$ *пф*	25	35	50
f_α *Мгц*	0,5	1,0	2,0
$F_ш$ *дб*	8,0	15	33
$I_{к0}$ *мка*	0,5	3	15
$I_{э0}$ *мка*	0,5	3	15

ления и является функцией параметров, от которых зависит закон распределения случайной величины. При фиксированных значениях этих параметров, выделяющих конкретный закон в классе однотипных законов распределения, каждая обобщенная характеристика принимает постоянное численное значение. По этой причине они называются числовыми характеристиками случайных величин.

В табл. 3.1 три числа дают две характеристики случайной величины — ожидаемое значение, или, как говорят, положение случайной величины на числовой оси, и ее разброс. Числовые характеристики положения и разброса употребляются чаще остальных, поэтому на их рассмотрении мы в основном и остановимся.

§ 3.2. Математическое ожидание случайной величины

Наиболее важной характеристикой положения является *математическое ожидание* случайной величины, для обозначения которого будем использовать символ $\mathbf{M}X$. В литературе встречаются также следующие обозначения: $\mathbf{E}X$, m_x, $\overline{X}$.

Пусть $g(x)$ — некоторая функция вещественного переменного, а X — случайная величина с функцией распределения $F(x)$.

Определение 1. *Если существует интеграл*

$$\int_{-\infty}^{\infty} |g(x)|\,dF(x), \tag{3.1}$$

то интеграл

$$\mathbf{M}g(X) = \int_{-\infty}^{\infty} g(x)\,dF(x) \tag{3.2}$$

называется математическим ожиданием функции $g(X)$.

Из определения интеграла Стилтьеса следует, что (3.2) существует всякий раз, когда существует (3.1), поэтому

$\mathbf{M}g(X)$ существует тогда и только тогда, когда существует $\mathbf{M}|g(X)|$.

Для дискретной случайной величины X, принимающей значения $x_1, x_2, \ldots, x_n, \ldots$ с вероятностями p_i, $i=1, 2, \ldots$ математическое ожидание функций $g(X)$, если оно существует, равно

$$\mathbf{M}g(X)=\sum_{i=1}^{\infty} g(x_i)\,p_i \tag{3.3}$$

и условием существования $\mathbf{M}g(X)$ является абсолютная сходимость ряда (3.3).

Для непрерывной случайной величины X с плотностью распределения $f(x)$

$$\mathbf{M}g(X)=\int_{-\infty}^{\infty} g(x)\,f(x)\,dx \tag{3.4}$$

и условием существования $\mathbf{M}g(X)$ является абсолютная сходимость интеграла (3.4).

При $g(x)=x$ интеграл

$$\mathbf{M}X=\int_{-\infty}^{\infty} x dF(x), \tag{3.5}$$

если он сходится абсолютно, называется математическим ожиданием случайной величины X.

В этом случае (3.3) для $i=\overline{1,n}$ принимает вид

$$\mathbf{M}X=\sum_{i=1}^{n} x_i p_i. \tag{3.6}$$

Этим выражением удобно воспользоваться для того, чтобы выяснить содержание числовой характеристики $\mathbf{M}X$ случайной величины. Поскольку теоретическая вероятность события $X=x_i$ близка к относительной частоте этого события, т. е. $p_i \approx m_i/N$, где m_i — число появления события $X=x_i$ в серии из N испытаний, то

$$x_i p_i \approx \frac{x_i m_i}{N}.$$

Отсюда

$$\mathbf{M}X=\sum_{i=1}^{n} x_i p_i \approx \sum_{i=1}^{n} \frac{x_i m_i}{N}=\frac{1}{N}\sum_{i=1}^{n} x_i m_i.$$

Но $\sum\limits_{i=1}^{n} x_i m_i$ — сумма всех значений случайной величины, полученных в N испытаниях, поэтому

$$\mathbf{M}X \approx \frac{1}{N}\sum_{i=1}^{n} x_i m_i=\frac{1}{N}\sum_{k=1}^{N} x_k. \tag{3.7}$$

Другими словами, если теоретические вероятности близки к относительным частотам событий, то (3.2) дает число, близкое к среднему арифметическому значению случайной величины $g(X)$ в серии испытаний, которым очень широко пользуются на практике в качестве «типичного» значения этой случайной величины. Воспользовавшись выражением (2.29), нетрудно получить тот же результат для непрерывной случайной величины. А если учесть, что относительная частота близка к численному значению вероятности при классическом определении, то ясно, что все рассуждения остаются в силе и для классического определения вероятности.

Пример 3.1. Найдем математическое ожидание случайной величины X, распределенной по биномиальному закону (2.3). Поскольку вероятность появления события $X=k$, $k=\overline{0,n}$ выражается формулой (2.4), то

$$\mathbf{M}X=\sum_{k=0}^{n} k\frac{n!}{k!\,(n-k)!}p^k q^{n-k}=np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)!\,(n-k)!}p^{k-1}q^{n-k}=$$

$$=np\sum_{m=0}^{n-1}\frac{(n-1)!}{m!\,(n-1-m)!}\,p^m q^{n-1-m}=np(p+q)^{n-1}=np.$$

Окончательно $$\mathbf{M}X=np. \tag{3.8}$$

Пример 3.2. Найдем математическое ожидание случайной величины X, распределенной по равномерному закону (2.7):

$$\mathbf{M}X=\int_{-\infty}^{\infty} x\,dF(x)=\int_{-\infty}^{\infty}\frac{x}{b-a}\,dx=\frac{b+a}{2}. \tag{3.9}$$

Пример 3.3. Случайная величина X распределена по нормальному закону (2.10):

$$\mathbf{M}X=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x\exp\left[-\frac{(x-m)^2}{2\sigma^2}\right]dx.$$

Заменим переменную по формуле

$$y=\frac{x-m}{\sigma\sqrt{2}};\quad dx=\sigma\sqrt{2}\,dy \tag{3.10}$$

и, учитывая (2.18), получим

$$\mathbf{M}X=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(y\sigma\sqrt{2}+m)\exp(-y^2)\,\sigma\sqrt{2}\,dy=$$

$$=\frac{\sigma\sqrt{2}}{\sqrt{\pi}}\int_{-\infty}^{\infty} y\exp(-y^2)\,dy+\frac{m}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp(-y^2)\,dy=m.$$

$$\mathbf{M}X=m. \tag{3.11}$$

Пример 3.4. Случайная величина Y распределена по закону Коши (2.133)

$$MY = \int_{-\infty}^{\infty} y f_y(y)\,dy = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{y\,dy}{1+y^2} =$$

$$= \lim_{A\to\infty}\frac{1}{2\pi}\int_{-A}^{A}\frac{2y\,dy}{1+y^2} = \lim_{A\to\infty}\ln|1+y^2|\,\Big|_{-A}^{A}.$$

Этот интеграл расходится, поэтому случайная величина, распределенная по закону Коши, не имеет математического ожидания.

Если $F(x/B)$ — условная функция распределения X при гипотезе B, то

$$M(X/B) = \int_{-\infty}^{\infty} x\,dF(x/B) \tag{3.12}$$

при условии, что интеграл сходится абсолютно, называется условным математическим ожиданием X относительно события B.

Пусть $B_1, B_2, \ldots, B_n$ — полная группа попарно несовместных событий и $F_1(x/B_1), F_2(x/B_2), \ldots, F_n(x/B_n)$ — соответствующие этим событиям условные функции распределения X. По формуле полной вероятности (1.32) для (2.1) получим

$$F(x) = \sum_{i=1}^{n} P(B_i)\,F_i(x/B_i).$$

Отсюда

$$MX = \int_{-\infty}^{\infty} x\,dF(x) = \int_{-\infty}^{\infty} x\,d\left[\sum_{i=1}^{n} P(B_i)\,F(x/B_i)\right] =$$

$$= \sum_{i=1}^{n} P(B_i)\int_{-\infty}^{\infty} x\,dF(x/B_i) = \sum_{i=1}^{n} P(B_i)\,M[X/B_i].$$

Формально это равенство можно представить в виде

$$MX = M[M(X/B_i)]. \tag{3.13}$$

Для n-мерного случайного вектора $(X_1, X_2, \ldots, X_n)$ с функцией распределения $F(x_1, x_2, \ldots, x_n)$ математическое ожидание функции $g(X_1, X_2, \ldots, X_n)$ определяется так же, как и для одномерного.

Определение 2. *Если существует интеграл*

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}|g(x_1, x_2, \ldots, x_n)|\,dF(x_1, x_2, \ldots, x_n),$$

то интеграл

$$\mathrm{M}g(X_1, X_2, \ldots, X_n) =$$

$$= \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} g(x_1, x_2, \ldots, x_n)\, dF(x_1, x_2, \ldots, x_n) \quad (3.14)$$

называется математическим ожиданием функции $g(X_1, X_2, \ldots, X_n)$. В частности, если $g(x_1, x_2, \ldots, x_n) = x_k$, то (3.14) дает математическое ожидание компоненты X_k. Действительно, учитывая (2.66) для случая n-мерного вектора, имеем

$$\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} x_k dF(x_1, x_2, \ldots, x_n) =$$

$$= \int\limits_{-\infty}^{\infty} x_k \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} dF(x_1, x_2, \ldots, x_n) =$$

$$= \int\limits_{-\infty}^{\infty} x_k dF_k(x_k) = \mathrm{M}X_k. \quad (3.15)$$

Математическим ожиданием случайного вектора $(X_1, X_2, \ldots, X_n)$ называется вектор

$$(\mathrm{M}X_1, \mathrm{M}X_2, \ldots, \mathrm{M}X_n). \quad (3.16)$$

Пример 3.5. В примере 2.9 показано, что обе компоненты нормально распределенного вектора (X, Y) имеют нормальные законы распределения (одномерные), а в примере 3.3 показано, что математическим ожиданием случайной величины, распределенной по нормальному закону (2.10), является параметр m. Таким образом, если вектор (X, Y) имеет нормальный закон распределения (2.88), то согласно (2.89), (2.90), (3.15) и (3.16) математическое ожидание его есть вектор (m_1, m_2).

Теоремы о математических ожиданиях случайных величин

Теорема 1. *Математическое ожидание постоянной равно этой постоянной.*

Доказательство. Постоянную a можно рассматривать как случайную величину X с рядом распределения $P(X=a)=1$, поэтому согласно (3.6)

$$\mathrm{M}a = aP(X = a) = a. \quad (3.17)$$

Теорема 2. *Математическое ожидание суммы случайных величин равно сумме математических ожиданий.*

Доказательство. Если $Y=X_1+X_2$ и $F(x_1, x_2)$ — совместная функция распределения вектора (X_1, X_2), то по определению (3.14)

$$MY=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1+x_2)\,dF(x_1, x_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_1dF(x_1\ x_2)+$$

$$+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_2dF(x_1, x_2)=\int_{-\infty}^{\infty}x_1\int_{-\infty}^{\infty}dF(x_1, x_2)+$$

$$+\int_{-\infty}^{\infty}x_2\int_{-\infty}^{\infty}dF(x_1, x_2)=\int_{-\infty}^{\infty}x_1dF_1(x)+\int_{-\infty}^{\infty}x_2dF_2(x_2).$$

Здесь мы воспользовались (2.66). Окончательно

$$MY=M(X_1+X_2)=MX_1+MX_2. \qquad (3.18)$$

Методом индукции (3.18) распространяется на сумму любого конечного числа слагаемых

$$M\left[\sum_{k=1}^{n}X_k\right]=\sum_{k=1}^{n}MX_k. \qquad (3.19)$$

Иногда удобно рассматривать вместо случайной величины ее отклонение от математического ожидания. Это отклонение называют центрированной случайной величиной. Центрированную случайную величину будем обозначать той же буквой, но с ноликом сверху

$$\overset{0}{X}=X-MX. \qquad (3.20)$$

Поскольку MX — постоянное число, то, учитывая (3.17), по формуле (3.18) получим

$$M\overset{0}{X}=M[X-MX]=MX-MX=0. \qquad (3.21)$$

Математическое ожидание центрированной случайной величины равно нулю.

Теорема 3. *Математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий:*

$$M(X_1X_2)=MX_1MX_2. \qquad (3.22)$$

Доказательство. Пусть вектор (X_1, X_2) распределен по закону $F(x_1, x_2)=F_1(x_1)F_2(x_2)$. По определению (3.14)

$$M(X_1X_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_1x_2dF(x_1, x_2).$$

Принимая во внимание (2.52), имеем

$$\mathrm{M}(X_1X_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_1x_2dF_1(x_1)\,dF_2(x_2)=$$

$$=\int_{-\infty}^{\infty}x_1dF_1(x_1)\int_{-\infty}^{\infty}x_2dF_2(x_2)=\mathrm{M}X_1\mathrm{M}X_2.$$

Методом полной математической индукции (3.22) распространяется на любое конечное число независимых сомножителей:

$$\mathrm{M}\left[\prod_{k=1}^{n}X_k\right]=\prod_{k=1}^{n}\mathrm{M}X_k. \tag{3.23}$$

Если a — постоянное число, то на основании (3.17) и (3.22)

$$\mathrm{M}(aX)=a\mathrm{M}X, \tag{3.24}$$

т. е. постоянный множитель можно выносить за знак математического ожидания. В частности, поскольку $i=\sqrt{-1}$ — постоянное число в множестве комплексных чисел, то (3.24) и (3.18) позволяют распространить понятие математического ожидания на комплексные случайные величины. Если (X, Y) — случайный вектор, то сумму $X+iY$ называют комплексной случайной величиной:

$$\mathrm{M}(X+iY)=\mathrm{M}X+i\mathrm{M}Y. \tag{3.25}$$

§ 3.3. Моменты случайной величины

Определение 1. *Моментом k-го порядка случайной величины X относительно a называется математическое ожидание случайной величины $Y=(X-a)^k$:*

$$\mathrm{M}\,(X-a)^k=\int_{-\infty}^{\infty}(x-a)^k dF(x). \tag{3.26}$$

Если $a=0$, то момент k-го порядка называется *начальным* и обозначается в дальнейшем α_k:

$$\alpha_k=\mathrm{M}X^k. \tag{3.27}$$

Если $a=\mathrm{M}X$, то момент k-го порядка называется *центральным* и обозначается в дальнейшем μ_k:

$$\mu_k=\mathrm{M}\,[X-\mathrm{M}X]^k=\mathrm{M}\overset{\circ}{X}{}^k. \tag{3.28}$$

Из определения начального момента k-го порядка следует, что математическое ожидание есть начальный момент первого порядка: $\alpha_1=\mathrm{M}X$.

Для дискретной случайной величины

$$\alpha_k = \sum_{i=1}^{n} x_i^k P(X = x_i); \tag{3.29}$$

$$\mu_k = \sum_{i=1}^{n} (x_i - \alpha_1)^k P(X = x_i). \tag{3.30}$$

Для непрерывной случайной величины

$$\alpha_k = \int_{-\infty}^{\infty} x^k f(x)\, dx; \tag{3.31}$$

$$\mu_k = \int_{-\infty}^{\infty} (x - \alpha_1)^k f(x)\, dx. \tag{3.32}$$

По определению математического ожидания функция случайного аргумента $\mathbf{M}X^k$ существует в том и только в том случае, если существует $\mathbf{M}|X|^k$, т. е. если сходится интеграл $\int_{-\infty}^{\infty} |x|^k dF(x)$. А из этого условия вытекает существование всех моментов порядка $r \leqslant k$, $r > 0$. Действительно, для всех $x > 1$ и $r < k$ справедливо неравенство $|x|^r < |x|^k$. Отсюда

$$\int_{-\infty}^{\infty} |x|^r\, dF(x) = \int_{-\infty}^{-1} |x|^r\, dF(x) + \int_{-1}^{+1} |x|^r\, dF(x) +$$

$$+ \int_{+1}^{\infty} |x|^r\, dF(x) \leqslant \int_{|x|>1} |x|^k dF(x) + \int_{-1}^{+1} |x|^r\, dF(x).$$

Первый интеграл справа от знака неравенства существует по условию, а второй конечен в силу конечности интервала интегрирования. Из ограниченности интеграла $\int_{-\infty}^{\infty} |x|^r dF(x)$ вытекает его существование и существование момента $\mathbf{M}X^r$. Учитывая теоремы о математическом ожидании, можно установить связь между центральными и начальными моментами случайной величины:

$$\mu_k = \mathbf{M}\,(X - \alpha_1)^k = \mathbf{M}\left[\sum_{r=0}^{k} (-1)^r C_k^r X^{k-r} \alpha_1^r\right] =$$

$$= (-1)^k \alpha_1^k + (-1)^{k-1} k \alpha_1^k + (-1)^r \sum_{r=1}^{k-2} C_k^r \alpha_1^r \alpha_{k-r}.$$

Окончательно

$$\mu_k = (-1)^{k-1} (k-1)\, \alpha_1^k + (-1)^r \sum_{r=1}^{k-2} C_k^r \alpha_1^r \alpha_{k-r}. \tag{3.33}$$

Для первых пяти моментов имеем:

$$\left.\begin{aligned} &1.\ \mu_0 = 1; \\ &2.\ \mu_1 = 0; \\ &3.\ \mu_2 = \alpha_2 - \alpha_1^2; \\ &4.\ \mu_3 = \alpha_3 - 3\alpha_2\alpha_1 + 2\alpha_1^3; \\ &5.\ \mu_4 = \alpha_4 - 4\alpha_3\alpha_1 + 6\alpha_2\alpha_1^2 - 3\alpha_1^4. \end{aligned}\right\} \quad (3.34)$$

Обратимся к механической интерпретации вероятностной меры. Будем считать, что $f(x)dx$ есть масса элемента $[x, x+dx]$ стержня, причем ось x направлена вдоль стержня. При такой интерпретации $f(x)$ представляет собой закон изменения плотности (линейной) вдоль оси стержня (рис. 3.1). Произведение $x^k f(x)dx$ есть механический момент k-го порядка элемента $[x, x+dx]$ стержня относительно начала координат, а интеграл (3.31), следовательно,—начальный механический момент k-го порядка всего стержня (при ускорении силы тяжести $g=1$). Примем массу всего стержня за единицу, т. е. $\int\limits_{-\infty}^{\infty} f(x)dx=1$ (вне стержня $f(x)=0$), и пусть x_0 — абсцисса центра тяжести стержня. В таком случае

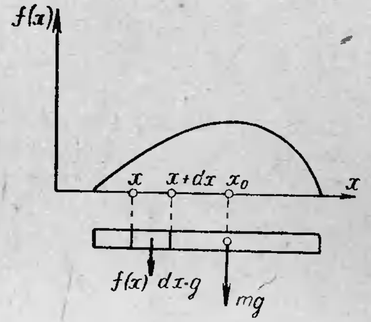

Рис. 3.1

$$\alpha_1 = \int\limits_{-\infty}^{\infty} x f(x)\,dx = x_0 \int\limits_{-\infty}^{\infty} f(x)\,dx = x_0,$$

т. е. начальный момент первого порядка численно равен абсциссе центра тяжести стержня.

Таким образом, центральные моменты согласно (3.28) являются моментами относительно центра тяжести стержня.

Моменты в механике используются в качестве обобщенных характеристик расположения масс относительно начала координат или центра тяжести. Аналогичную роль выполняют моменты случайных величин в теории вероятностей.

Для случайного n-мерного вектора (2.56) выражение (3.26) определяет моменты для отдельных компонент вектора, причем $F(x)$ — одномерная функция распределения соответствующей компоненты, определяемая (2.60). Но кроме этих численных характеристик в многомерном случае можно рассматривать так называемые смешанные центральные и начальные моменты. Ограничимся двумерным вектором.

Определение 2. *Начальным смешанным моментом порядка k, s вектора* (X_1, X_2) *называется математическое ожидание произведения* $X_1^k X_2^s$:

$$\alpha_{ks} = \mathbf{M}\,[X_1^k X_2^s]. \tag{3.35}$$

Определение 3. *Центральным смешанным моментом порядка k, s вектора* (X_1, X_2) *называется математическое ожидание произведения* $[X_1 - \mathbf{M}X_1]^k [X_2 - \mathbf{M}X_2]^s$:

$$\mu_{ks} = \mathbf{M}\,[X_1 - \mathbf{M}X_1]^k\,[X_2 - \mathbf{M}X_2]^s. \tag{3.36}$$

Наиболее часто используется смешанный центральный момент порядка 1,1, который называется корреляционным моментом и обозначается символом K_{12} или $K_{X_1X_2}$:

$$K_{12} = \mu_{11} = \mathbf{M}\,[X_1 - \mathbf{M}X_1]\,[X_2 - \mathbf{M}X_2] = \mathbf{M}[\mathring{X}_1\mathring{X}_2]. \tag{3.37}$$

§ 3.4. Дисперсия случайной величины

Среди числовых характеристик разбросанности случайной величины важнейшей характеристикой является центральный момент второго порядка. В связи с особым значением этой характеристики для нее используется специальное наименование *дисперсия* (что означает рассеивание). В дальнейшем дисперсию будем обозначать символом $\mathbf{D}X$:

$$\mathbf{D}X = \mu_2 = \mathbf{M}\,(X - \alpha_1)^2 = \mathbf{M}\mathring{X}^2 = \int\limits_{-\infty}^{\infty} (x - \alpha_1)^2\,dF(x). \tag{3.38}$$

При $k=2$ (3.30) и (3.32) дают выражения для дисперсии соответственно дискретной и непрерывной случайной величин.

Из (3.38) видно, что дисперсия есть математическое ожидание квадрата отклонений X от $\mathbf{M}X$, поэтому дисперсия есть величина неотрицательная и имеет размерность квадрата размерности случайной величины. Квадратный корень из дисперсии

$$\sigma(x) = \sqrt{\mathbf{D}X} \tag{3.39}$$

имеет размерность случайной величины и также зависит от ее разброса. Эта числовая характеристика разброса называется *средним квадратическим отклонением* или *стандартом* случайной величины X.

Для вычисления дисперсии удобно пользоваться формулой 3 из (3.34), которая в других обозначениях имеет вид

$$\mathbf{D}X = \mathbf{M}X^2 - [\mathbf{M}X]^2. \tag{3.40}$$

Пример 3.6. Найдем дисперсию случайной величины X, распределенной по биномиальному закону (2.3). Учитывая (2.4), получим согласно (3.40) и (3.29)

$$\mathbf{D}X+[\mathbf{M}X]^2=\sum_{k=0}^{n}k^2\frac{n!}{k!\,(n-k)!}p^kq^{n-k}=$$

$$=\sum_{k=1}^{n}k\frac{n!}{(k-1)!\,(n-k)!}p^kq^{n-k}=np\sum_{i=0}^{n-1}(i+1)\frac{(n-1)!}{i!\,(n-i-1)!}p^iq^{n-1-i}=$$

$$=np\sum_{i=0}^{n-1}i\frac{(n-1)!}{i!\,(n-i-1)!}p^iq^{n-i-1}+np\sum_{i=0}^{n-1}\frac{(n-1)!}{i!\,(n-i-1)!}p^iq^{n-i-1}=$$

$$=n^2p^2+np(1-p). \tag{3.41}$$

Если учесть (3.8), то

$$\mathbf{D}X=np(1-p)=npq,\qquad q=1-p. \tag{3.42}$$

Пример 3.7. Случайная величина X распределена по равномерному закону (2.7):

$$\mathbf{D}X+[\mathbf{M}X]^2=\int_{-\infty}^{\infty}x^2f(x)\,dx=\frac{1}{b-a}\int_a^b x^2dx=\frac{b^2+ab+a^2}{3}.$$

Учитывая (3.9), получим

$$\mathbf{D}X=\frac{b^2+ab+a^2}{3}-\frac{(b+a)^2}{4}=\frac{(b-a)^2}{12}. \tag{3.43}$$

Из этого выражения видно, что $\mathbf{D}X$ тем больше, чем больше интервал возможных значений X, т. е. чем больше разброс случайной величины k.

Пример 3.8. Случайная величина X распределена по нормальному закону (2.9):

$$\mathbf{D}X+[\mathbf{M}X]^2=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}t^2\exp\left[-\frac{(t-m)^2}{2\sigma^2}\right]dt.$$

Замена переменной по формуле (2.17) дает

$$\mathbf{D}X+[\mathbf{M}X]^2=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}(\sigma\sqrt{2}y+m)^2\exp(-y^2)\,dy=$$

$$=\frac{2\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{\infty}y^2\exp(-y^2)\,dy+\frac{2\sigma m\sqrt{2}}{\sqrt{\pi}}\int_{-\infty}^{\infty}y\exp(-y^2)\,dy+\frac{m^2\sqrt{\pi}}{\sqrt{\pi}}.$$

Так как

$$\int_{-\infty}^{\infty}y^2\exp(-y^2)\,dy=\frac{1}{2}\int_{-\infty}^{\infty}\exp(-y^2)\,dy=\frac{\sqrt{\pi}}{2},$$

поэтому

$$\mathbf{D}X+[\mathbf{M}X]^2=\sigma^2+m^2.$$

Учитывая (3.11), окончательно получим

$$\mathrm{D}X=\sigma^2. \tag{3.44}$$

Таким образом, параметрами нормального закона распределения (2.9) являются математическое ожидание случайной величины и ее дисперсия:

$$m=\mathrm{M}X; \qquad \sigma^2=\mathrm{D}X.$$

Графики на рис. 2.5 иллюстрируют связь дисперсии с разбросом случайной величины. Найдем вероятность наблюдать отклонения X от $\mathrm{M}X$ на величину, большую $\beta>0$:

$$P(|\overset{\circ}{X}|>\beta)=P[(\overset{\circ}{X}<-\beta)\cup(\overset{\circ}{X}>\beta)]=$$
$$=P(\overset{\circ}{X}<-\beta)+P(\overset{\circ}{X}>\beta)=P[(X-\mathrm{M}X)<-\beta]+$$
$$+P[(X-\mathrm{M}X)>\beta]=P(X<-\beta+\mathrm{M}X)+$$
$$+P[X>\beta+\mathrm{M}X]=P(X<-\beta+\mathrm{M}X)+$$
$$+1-P(X<\beta+\mathrm{M}X)=1+F(-\beta+\mathrm{M}X)-F(\beta+\mathrm{M}X).$$

При $\beta=1{,}0$ по кривым рис. 2.5 находим для случая $m=0$ и $\sigma=1$

$$P(|\overset{\circ}{X}|>1{,}0)=F(-1{,}0)+1-F(1{,}0)=$$
$$=1+0{,}16-0{,}83=0{,}33$$

и для случая $m=2$ и $\sigma=0{,}5$

$$P(|\overset{\circ}{X}|>1{,}0)=1+F(1{,}0)-F(3{,}0)=$$
$$=1+0{,}02-0{,}98=0{,}04.$$

Таким образом, уменьшение среднего квадратического отклонения вдвое уменьшило вероятность отклонения случайной величины на $\pm 1{,}0$ от $\mathrm{M}X$ почти в 10 раз.

§ 3.5. Дисперсия случайного вектора

Определение. *Дисперсией n-мерного случайного вектора (или дисперсионной матрицей) называется матрица*

$$\|K_{ij}\|=\begin{Vmatrix} K_{11} & K_{12} & \dots & K_{1n} \\ K_{21} & K_{22} & \dots & K_{2n} \\ \dots & \dots & \dots & \dots \\ K_{n1} & K_{n2} & \dots & K_{nn} \end{Vmatrix}, \tag{3.45}$$

элементы которой определяются формулой

$$K_{ij}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty}(x_i-\mathrm{M}X_i)(x_j-\mathrm{M}X_j)\,dF(x_1,x_2,\dots,x_n). \tag{3.46}$$

Если учесть (2.66) для случая n-мерного вектора, то (3.46) можно представить в виде

$$K_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - \mathbf{M}X_i)(x_j - \mathbf{M}X_j)\, dF(x_i, x_j). \qquad (3.47)$$

Из (3.47) следует:

1. $K_{ij} = \mathbf{M}[X_i - \mathbf{M}\,X_i]\,[X_j - \mathbf{M}X_j] = \mathbf{M}\,[\overset{\circ}{X}_i \overset{\circ}{X}_j] = \mu_{1,1}$, (3.48)

т. е. K_{ij} согласно (3.37) есть корреляционный момент случайных величин X_i и X_j;

2. $K_{ii} = \mathbf{D}X_i$, $i = \overline{1, n}$, т. е. диагональными элементами дисперсионной матрицы являются дисперсии соответствующих компонент n-мерного вектора;

3. $K_{ij} = K_{ji}$, т. е. дисперсионная матрица симметрична.

В связи с тем что недиагональные элементы матрицы (3.45) представляют собой моменты корреляции компонент вектора, эту матрицу называют также *корреляционной* матрицей. Поскольку эта матрица симметрична, то ее изображают иногда наполовину заполненной таблицей

$$\|K_{ij}\| = \begin{Vmatrix} K_{11} & K_{12} & K_{13} & \ldots & K_{1n} \\ & K_{22} & K_{23} & \ldots & K_{2n} \\ & & \ldots & \ldots & \ldots \\ & & & \ldots & \\ & & & & K_{nn} \end{Vmatrix}. \qquad (3.49)$$

Численные значения элементов дисперсионной матрицы определяются, как следует из (3.47), разбросом компонент вектора X, но, кроме того, на величину недиагональных элементов влияет и зависимость между соответствующими парами компонент. Покажем, что для независимых случайных величин момент корреляции равен нулю. Пусть X_1 и X_2 — случайные величины, совместная функция распределения которых

$$F(x_1, x_2) = F_1(x_1)\, F_2(x_2).$$

По формуле (3.47), учитывая (2.52), имеем

$$K_{12} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - \mathbf{M}X_1)(x_2 - \mathbf{M}X_2)\, dF_1(x_1)\, dF_2(x_2) =$$

$$= \int_{-\infty}^{\infty} (x_1 - \mathbf{M}X_1)\, dF_1(x_1) \int_{-\infty}^{\infty} (x_2 - \mathbf{M}X_2)\, dF_2(x_2) = 0. \qquad (3.50)$$

Случайные величины, момент корреляции которых равен нулю, называются *некоррелированными*. Из (3.50) следует, что независимость случайных величин влечет их некоррелированность. Но обратное утверждение неверно. Иногда зависимые величины могут быть некоррелированными, т. е. из некоррелированности случайных величин не следует, вообще говоря, их независимость.

Пример 3.9. Пусть случайная величина X_1 имеет симметричное распределение относительно начала координат, т. е. плотность распределения ее характеризуется тем, что $f(x) = f(-x)$. Случайная величина $X_2 = |X_1|$ зависит от случайной величины X_1, поскольку численное значение X_1 однозначно определяет численное значение X_2. Найдем момент корреляции этих зависимых случайных величин.

$$\mathbf{M} X_1 = \int_{-\infty}^{\infty} xf(x)\,dx = \int_{-\infty}^{0} xf(x)\,dx + \int_{0}^{\infty} xf(x)\,dx =$$

$$= \int_{0}^{\infty} xf(x)\,dx - \int_{0}^{\infty} xf(x)\,dx = 0.$$

$$\mathbf{M}[X_1 X_2] = \mathbf{M}[X_1 |X_1|] = \int_{-\infty}^{\infty} x|x|f(x)\,dx =$$

$$= \int_{-\infty}^{0} x|x|f(x)\,dx + \int_{0}^{\infty} x|x|f(x)\,dx =$$

$$= \int_{0}^{\infty} x|x|f(x)\,dx - \int_{0}^{\infty} x|x|f(x)\,dx = 0.$$

Поэтому

$$K_{12} = \mathbf{M}[X_1 - \mathbf{M}X_1][X_2 - \mathbf{M}X_2] = \mathbf{M}[X_1 X_2] - \mathbf{M}X_1 \mathbf{M}X_2 = 0,$$

т. е. при симметричном законе распределения случайной величины X_1 зависимые случайные величины X_1 и X_2 некоррелированы.

Для характеристики только коррелированности компонент случайного n-мерного вектора безотносительно к их рассеиванию пользуются вместо корреляционной матрицы $\|K_{ij}\|$ *нормированной* корреляционной матрицей $\|r_{ij}\|$, составленной из безразмерных коэффициентов:

$$r_{ij} = \frac{K_{ij}}{\sqrt{K_{ii}}\sqrt{K_{jj}}} = \frac{K_{ij}}{\sigma_i \sigma_j}, \tag{3.51}$$

где

$$\sigma_i = \sqrt{\mathbf{D}X_i} \text{ и } \sigma_j = \sqrt{\mathbf{D}X_j}.$$

Коэффициент r_{ij} называется *коэффициентом корреляции* случайных величин X_i и X_j. Покажем, что для любых X_i и X_j коэффициент корреляции лежит в промежутке [1, —1] и принимает значения ± 1 в том и только в том случае, если между X_i и X_j существует линейная функциональная зависимость.

Найдем дисперсию случайной величины $Y=\frac{1}{\sigma_i}X_i \mp \frac{1}{\sigma_j}X_j$:

$$\mathbf{D}Y=\mathbf{M}\left[\frac{1}{\sigma_i}\overset{\circ}{X}_i \mp \frac{1}{\sigma_j}\overset{\circ}{X}_j\right]^2=$$

$$=\mathbf{M}\left[\frac{1}{\sigma_i^2}\overset{\circ}{X}{}_i^2+\frac{1}{\sigma_j^2}\overset{\circ}{X}{}_j^2 \pm 2\frac{1}{\sigma_i\sigma_j}\overset{\circ}{X}_i\overset{\circ}{X}_j\right]=$$

$$=\frac{1}{\sigma_i^2}\mathbf{M}\overset{\circ}{X}_i+\frac{1}{\sigma_j^2}\mathbf{M}\overset{\circ}{X}{}_j^2 \pm 2\frac{1}{\sigma_i\sigma_j}\mathbf{M}\overset{\circ}{X}_i\overset{\circ}{X}_j=2\pm 2r_{ij}.$$

Поскольку для любой случайной величины $\mathbf{D}Y \geqslant 0$, то

$$-1 \leqslant r_{ij} \leqslant 1. \qquad (3.52)$$

Если $r_{ij}=\pm 1$, то $\mathbf{D}Y=0$, что может быть только в том случае, если Y не является случайной величиной, поскольку только при этом условии отклонения Y от постоянной $\mathbf{M}Y$ равны нулю. Отсюда

$$\frac{1}{\sigma_i}X_i \mp \frac{1}{\sigma_j}X_j=a=\text{const},$$

т. е. X_i и X_j связаны линейной функциональной зависимостью.

Теперь пусть $X_i=aX_j+b$, a и b — постоянные:

$$\overset{\circ}{X}_i=aX_j+b-\mathbf{M}[aX_j+b]=a\overset{\circ}{X}_j,$$

$$K_{ij}=\mathbf{M}\overset{\circ}{X}_i\overset{\circ}{X}_j=a\mathbf{M}\overset{\circ}{X}{}_i^2=a\sigma_i^2,$$

$$K_{ij}=\mathbf{M}\overset{\circ}{X}_i\overset{\circ}{X}_j=\frac{1}{a}\mathbf{M}\overset{\circ}{X}{}_j^2=\frac{1}{a}\sigma_j^2.$$

Отсюда $\sigma_j^2=a^2\sigma_i^2$. Поскольку $\sigma_j^2 \geqslant 0$, то $\sigma_j=|a|\sigma_i$. Следовательно,

$$r_{ij}=\frac{K_{ij}}{\sigma_i\sigma_j}=\frac{a\sigma_i^2}{\sigma_i|a|\sigma_i}=\operatorname{sign} a\cdot 1. \qquad (3.53)$$

Можно сказать, что величина коэффициента корреляции характеризует близость к линейной зависимости случайных величин. При независимости X_i и X_j коэффициент корреляции $r_{ij}=0$, а чем сильнее проявляется связь между случайными величинами, тем ближе r_{ij} к $+1$ или -1. Но при нелинейной связи между X_i и X_j коэффициент корреляции может быть равным нулю.

Пример 3.10. Найдем корреляционную матрицу вектора (X, Y), распределенного по нормальному закону (2.91).

Для K_{11} по формуле (3.47), учитывая (2.66) и (3.44), имеем:

$$K_{11}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mathbf{M}X)^2 f(x, y)\,dxdy=$$

$$=\int_{-\infty}^{\infty}(x-\mathbf{M}X)^2\int_{-\infty}^{\infty}f(x, y)\,dxdy=\int_{-\infty}^{\infty}(x-\mathbf{M}X)^2 f(x)\,dx=$$

$$=\mathbf{D}X=\sigma_1^2.$$

Аналогично $K_{22}=\sigma_2^2$.

Для вычисления момента $K_{12}=K_{xy}$ воспользуемся тем, что $K_{xy}=\mathbf{M}\mathring{X}\mathring{Y}$. Плотность распределения вектора $(\mathring{X}, \mathring{Y})$, которую обозначим $f_0(x, y)$, получим из (2.91) по формуле (2.123), учитывая, что $\mathring{X}=X-m_1$ и $\mathring{Y}=Y-m_2$:

$$f_0(x, y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\times$$

$$\times\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{x^2}{\sigma_1^2}-2r\frac{xy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_2^2}\right]\right\}. \qquad (3.54)$$

Таким образом,

$$K_{xy}=\mathbf{M}\mathring{X}\mathring{Y}=$$

$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xy\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{x^2}{\sigma_1^2}-2r\frac{xy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_2^2}\right]\right\}dxdy.$$

Квадратичную форму в показателе экспоненты приведем к каноническому виду преобразованием

$$\begin{aligned} x&=\sigma_1\left[u\sqrt{1-r^2}-vr\right];\\ y&=-v\sigma_2 \end{aligned} \qquad (3.55)$$

с якобианом

$$\left|\frac{\partial(x, y)}{\partial(u, v)}\right|=\left\|\begin{matrix}\sigma_1\sqrt{1-r^2} & -r\sigma_1\\ 0 & -\sigma_2\end{matrix}\right\|=\sigma_1\sigma_2\sqrt{1-r^2}.$$

Получим

$$\frac{x^2}{\sigma_1^2}-2r\frac{xy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_2^2}=(1-r^2)(u^2+v^2);$$

$$\mathbf{M}\mathring{X}\mathring{Y}=\frac{\sigma_1^2\sigma_2^2\sqrt{1-r^2}}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}-(\sqrt{1-r^2}\,uv-$$

$$-rv^2)\exp\left(-\frac{u^2+v^2}{2}\right)dudv.$$

Отсюда

$$K_{xy} = \sigma_1 \sigma_2 r. \tag{3.56}$$

Теперь можно составить дисперсионную матрицу нормально распределенного вектора (X, Y) полностью:

$$\| K_{ij} \| = \begin{Vmatrix} \sigma_1^2 & \sigma_1 \sigma_2 r \\ \sigma_1 \sigma_2 r & \sigma_2^2 \end{Vmatrix}. \tag{3.57}$$

Нормированная корреляционная матрица имеет следующий вид:

$$\| r_{ij} \| = \begin{Vmatrix} 1 & r \\ r & 1 \end{Vmatrix}. \tag{3.58}$$

Таким образом, параметр r, входящий в формулу двумерного нормального закона распределения (2.91) вектора (X, Y), есть коэффициент корреляции случайных величин X и Y. В примере 2.9 было показано, что при $r=0$ компоненты вектора (X, Y) оказываются (2.92) независимыми случайными величинами. Но, с другой стороны, $r=0$ означает, что случайные величины X и Y некоррелированы. Таким образом, для нормально распределенных случайных величин некоррелированность влечет независимость, а независимость при любом законе распределения, в том числе и при нормальном, влечет некоррелированность, т. е. при нормальном законе распределения свойство независимости равносильно свойству некоррелированности.

§ 3.6. Теоремы о дисперсии

Основные свойства дисперсии случайной величины сформулируем в виде нескольких теорем.

Теорема 1. *Дисперсия постоянной равна нулю*:

$$\mathbf{D}a = 0. \tag{3.59}$$

Доказательство. По определению (3.38), учитывая, что $\mathbf{M}a = a$, получим

$$\mathbf{D}a = \mathbf{M}(a - a)^2 = 0.$$

Теорема 2. *Постоянный множитель выносится за знак дисперсии в квадрате:*

$$\mathbf{D}aX = a^2 \mathbf{D} X. \tag{3.60}$$

Доказательство. Принимая во внимание (3.24), имеем

$$\mathbf{D}aX = \mathbf{M}[aX - \mathbf{M}aX]^2 = \mathbf{M}a^2 [X - \mathbf{M}X]^2 = a^2 \mathbf{D}X.$$

Теорема 3. *Дисперсия суммы (разности) двух случайных величин равна сумме их дисперсий плюс (минус) удвоенный корреляционный момент:*

$$\mathbf{D}(X_1+X_2)=\mathbf{D}X_1+\mathbf{D}X_2+2K_{12}; \tag{3.61}$$

$$\mathbf{D}(X_1-X_2)=\mathbf{D}X_1+\mathbf{D}X_2-2K_{12}. \tag{3.62}$$

Доказательство.

$$\mathbf{D}(X_1+X_2)=\mathbf{M}[(X_1+X_2)-\mathbf{M}(X_1+X_2)]^2=$$

$$=\mathbf{M}[X_1+X_2-\mathbf{M}X_1-\mathbf{M}X_2]^2=\mathbf{M}\overset{\circ}{X}{}_1^2+\mathbf{M}\overset{\circ}{X}{}_2^2+2\mathbf{M}\overset{\circ}{X}_1\overset{\circ}{X}_2=$$

$$=\mathbf{D}X_1+\mathbf{D}X_2+2K_{12}.$$

Аналогично доказывается (3.62). Теорему 3 можно распространить на любое конечное число слагаемых:

$$\mathbf{D}\left[\sum_{i=1}^{n}X_i\right]=\sum_{i=1}^{n}\mathbf{D}X\pm 2\sum_{i<j}K_{ij}. \tag{3.63}$$

Поскольку для независимых случайных величин коэффициент корреляции равен нулю, то дисперсия суммы независимых случайных величин равна сумме дисперсий слагаемых, т. е. если

$$F(x_1, x_2, \ldots, x_n)=\prod_{i=1}^{n}F_i(x_i),$$

то

$$\mathbf{D}\left[\sum_{i=1}^{n}X_i\right]=\sum_{i=1}^{n}\mathbf{D}X_i. \tag{3.64}$$

Теоремы 1—3 используются для вычисления дисперсий случайных величин.

Линейное преобразование

$$Y=\frac{X-\mathbf{M}X}{\sqrt{\mathbf{D}X}}=\frac{X-\mathbf{M}X}{\sigma(X)} \tag{3.65}$$

называется *нормированием* случайной величины X, а случайная величина Y — *нормированной* случайной величиной. Из теорем о математическом ожидании и дисперсии следует, что для нормированной случайной величины $\mathbf{M}Y=0$ и $\mathbf{D}Y=1$.

§ 3.7. Квантили, медиана, мода, коэффициент асимметрии и эксцесс

Квантилью порядка p называется корень уравнения $F(x)=p$:

$$x_p=F^{-1}(p). \tag{3.66}$$

Здесь $F^{-1}(p)$ есть функция, обратная функции $F(x)$.

Некоторые квантили получили специальные наименования. Так, квантиль порядка $p=0{,}5$ называется *медианой* распределения:

$$x_{0,5} = \mathrm{Me}\,X = F^{-1}(0{,}5). \tag{3.67}$$

Квантили порядка $p=0{,}25$ и $0{,}75$ называются *нижней* и *верхней квартилями* соответственно.

Медиана делит область значений случайной величины на две части так, что вероятности попадания случайной величины в каждую из них равны 0,5. Квартили делят область значений случайной величины на 4 интервала таким образом, что вероятность попадания случайной величины в каждый из них равна 0,25.

Модой дискретной случайной величины называется наиболее вероятное значение этой случайной величины. Мода биномиально распределенной случайной величины с параметрами $n=10$ и $p=0{,}04$ равна нулю (см. рис. 2.2, *б*), а с параметрами $n=10$ и $p=0{,}4$ равна 4.

Модой непрерывной случайной величины называется значение случайной величины, при котором плотность распределения максимальная. Мода случайной величины, распределенной по нормальному закону с параметрами $m=2$ и $\sigma=0{,}5$ (см. рис. 2.5, *б*), равна 2.

Распределения могут иметь одну моду (унимодальные) или несколько мод (полимодальные) или вообще не иметь моды. Равномерно распределенная случайная величина моды не имеет.

Для характеристики формы закона распределения можно использовать центральные моменты высших порядков. При симметричном относительно математического ожидания распределении все моменты нечетных порядков случайной величины равны нулю. Действительно, при этом условии в сумму

$$\mu_{2k+1} = \sum_{i=1}^{n} (x_i - \mathrm{M}X)^{2k+1} p_i$$

с каждым слагаемым входит равное ему по абсолютной величине и противоположное по знаку другое слагаемое, поэтому вся сумма равна нулю. Аналогично интеграл

$$\mu_{2k+1} = \int_{-\infty}^{\infty} (x - \mathrm{M}X)^{2k+1} f(x)\, dx$$

равен нулю.

Оба выражения для μ_{2k+1} отличны от нуля при асимметричном распределении. Численные значения и знаки этих центральных моментов зависят от степени и направления «скошенности» распределения, и для характеристики асимметрии можно взять любой момент, но удобнее простейший из них — третий. Его размерность равна кубу размерности случайной величины, поэтому, чтобы получить безразмерный показатель асимметрии, μ_3 делят на куб среднего квадратического отклонения:

$$S_k = \frac{\mu_3}{\sigma^3}. \tag{3.68}$$

Эта численная характеристика называется *коэффициентом асимметрии* распределения. Плотности распределения с

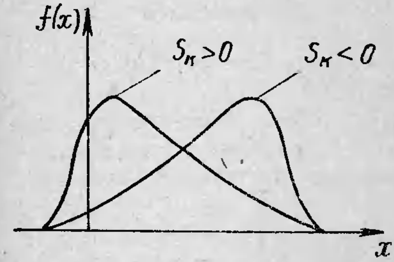

Рис. 3.2

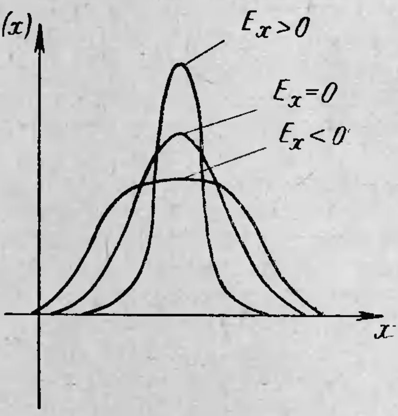

Рис. 3.3

положительной асимметрией $S_k > 0$ и отрицательной асимметрией $S_k < 0$ показаны на рис. 3.2.

Степень заостренности кривой плотности распределения или огибающей ряда распределения характеризуют безразмерным коэффициентом

$$E_x = \frac{\mu_4}{\sigma^4} - 3, \tag{3.69}$$

который называют *эксцессом* распределения. За «эталон» заостренности принята форма кривой плотности распределения нормального закона (2.10). Для этого закона $\mu_4/\sigma_4 = 3$ и, следовательно, $E_x = 0$. Кривые более островершинные, чем кривая нормальной плотности распределения, обладают положительным эксцессом, а кривые более плосковершинные — отрицательным эксцессом (рис. 3.3).

ГЛАВА 4

МЕТОДЫ ОПЕРАЦИОННОГО ИСЧИСЛЕНИЯ В ТЕОРИИ ВЕРОЯТНОСТЕЙ

§ 4.1. Методы операционного исчисления

Операционное исчисление в теории вероятностей, так же как и в других приложениях, служит важным инструментом решения многих практических и теоретических задач. Под операционным исчислением понимают методы, включающие следующий прием.

1. Функции, с которыми необходимо выполнить какое-либо действие и которые называются *оригиналами*, преобразуют по определенным правилам в другие функции — *изображения*.

2. Над изображениями производят операции, соответствующие заданным действиям над оригиналами.

3. Результат операций над изображениями обратным преобразованием переводят в оригинал.

В качестве преобразования используются преобразования Фурье, Карсона или Лапласа. При этих преобразованиях между оригиналами и их изображениями существует взаимно однозначное соответствие.

Эффективность методов операционного исчисления состоит в том, что операции над изображениями во многих случаях значительно проще, чем соответствующие операции над оригиналами.

В теории вероятностей методы операционного исчисления упрощают процедуру нахождения законов распределения сумм случайных величин. Затруднения могут возникнуть на этапах преобразований оригиналов и изображения и обратно. Однако в настоящее время кроме систематических методов прямых и обратных преобразований известны изображения Фурье и Лапласа почти для всех практически более или менее важных функций. Этими данными пользуются так же, как таблицами логарифмов и антилогарифмов при вычислениях.

Преобразования Фурье и Лапласа изучаются в курсах математического анализа, теории функций комплексного переменного и математической физики, поэтому здесь рассматриваются только те свойства этих преобразований и соответствующая терминология, которые специфичны и важны для теории вероятностей. Желающим познакомиться глубже с этим вопросом можно рекомендовать, например [8].

В других приложениях в интегральных преобразованиях Фурье и Лапласа обычно производится интегрирование по Риману. В теории вероятностей удобнее пользоваться интегралом Стилтьеса, применимым как к дискретным, так и к непрерывным распределениям случайных величин. В отличие от обычных преобразований Фурье и Лапласа преобразования с интегрированием по Стилтьесу называются соответственно преобразованиями Фурье—Стилтьеса или Лапласа—Стилтьеса.

§ 4.2. Характеристическая функция случайной величины

Пусть X — случайная величина с функцией распределения $F(x)$.

О п р е д е л е н и е. *Характеристической функцией распределения $F(x)$ (или случайной величины X) называется функция*

$$\varphi(t)=\int_{-\infty}^{\infty}\exp itx\,dF(x)=u(t)+iv(t). \tag{4.1}$$

Здесь $i=\sqrt{-1}$; t — вспомогательная вещественная переменная;

$$u(t)=\int_{-\infty}^{\infty}\cos tx\,dF(x);\quad v(t)=\int_{-\infty}^{\infty}\sin tx\,dF(x). \tag{4.2}$$

Термин «характеристическая функция», в отличие от более широкого термина «преобразование Фурье функции», подчеркивает тот факт, что оригиналом преобразования является функция распределения случайной величины.

Поскольку $|\exp itx|=1$ для всех t, то интеграл (4.1) сходится для любой функции распределения, т. е. для любой случайной величины может быть определена характеристическая функция.

Сравнивая (3.2) и (4.1), отметим, что характеристическую функцию случайной величины X можно определить как математическое ожидание функции $\exp itX$:

$$\varphi(t)=\mathbf{M}\exp itX. \tag{4.3}$$

Можно доказать [4], что характеристическая функция равномерно непрерывна во всем интервале $-\infty<t<+\infty$. Из определения (4.1) непосредственно вытекает: 1. $\varphi(0)=1$; 2. $|\varphi(t)|\leqslant 1$ для всех t; 3. $\varphi(-t)=\overline{\varphi(t)}$. Остальные свойства характеристической функции сформулируем в виде теорем.

Т е о р е м а 1. *Если $Y=aX+b$, где a и b — постоянные, то*

$$\varphi_y(t)=\varphi_x(at)\exp itb, \tag{4.4}$$

где $\varphi_x(t)$ и $\varphi_y(t)$ — характеристические функции величин X и Y соответственно.

Доказательство. Воспользуемся (4.3) и (3.24):

$$\varphi_y(t) = \mathbf{M} \exp itY = \mathbf{M} \exp [it(aX + b)] =$$
$$= \mathbf{M} \exp [i(at) X] \exp itb = \exp itb\, \varphi_x(at).$$

Теорема 2. *Характеристическая функция суммы двух независимых случайных величин равна произведению их характеристических функций.*

Доказательство. Убедимся вначале, что при независимости X_1 и X_2 случайные величины $Y_1 = \exp itX_1$ и $Y_2 = \exp itX_2$ также независимы. Пусть $F_x(x_1, x_2) = F_{1x}(x_1) F_{2x}(x_2)$ — функция распределения вектора (X_1, X_2) и $F_y(y_1, y_2)$ — функция распределения вектора (Y_1, Y_2). Воспользовавшись (2.114), (2.64) и (2.52), получим

$$F_y(y_1, y_2) = P[(X_1, X_2) \in A_y] = \iint\limits_{A_y} dF_x(x_1, x_2) =$$
$$= \iint\limits_{A_y} dF_{1x}(x_1)\, dF_{2x}(x_2) = \int\limits_{A_y} dF_{1x}(x_1) \int\limits_{A_y} dF_{2x}(x_2),$$

откуда $F_y(y_1, y_2)$ есть произведение одномерных функций распределения Y_1 и Y_2.

Теперь можно воспользоваться (3.22) для определения характеристической функции $\varphi_z(t)$ случайной величины $Z = X_1 + X_2$:

$$\varphi_z(t) = \mathbf{M} \exp itZ = \mathbf{M} \exp [it(X_1 + X_2)] =$$
$$= \mathbf{M} [\exp itX_1 \cdot \exp itX_2] = \mathbf{M} \exp itX_1 \mathbf{M} \exp itX_2 = \varphi_1(t) \varphi_2(t).$$

Здесь $\varphi_1(t)$ и $\varphi_2(t)$ — характеристические функции X_1 и X_2 соответственно.

Таким образом,

$$\varphi_z(t) = \varphi_1(t) \varphi_2(t). \tag{4.5}$$

Учитывая (2.129), теорему 2 можно сформулировать иначе: характеристическая функция свертки законов распределения $F_1(x)$ и $F_2(x)$ равна произведению характеристических функций этих законов.

Если $X = \sum\limits_{i=1}^{n} X_i$ и каждое слагаемое в этой сумме независимо от суммы предыдущих слагаемых, то из теоремы 2 вытекает, что характеристическая функция X равна произведению характеристических функций слагаемых:

$$\varphi(t) = \prod_{i=1}^{n} \varphi_i(t). \tag{4.6}$$

Теорема 3. *Если случайная величина X имеет начальный момент n-го порядка, то ее характеристическая функция n раз дифференцируема и при $r \leqslant n$*

$$\mathbf{M} X^r = -i^r \varphi^{(r)}(0) = -i^r \left. \frac{d^r \varphi(t)}{dt^r} \right|_{t=0}. \qquad (4.7)$$

Доказательство. Если существует начальный момент $\mathbf{M}X^n$, то это означает, что существует $\mathbf{M}|X|^n$ и все начальные моменты порядка $r \leqslant n$, $r > 0$. Дифференцируя формально (4.1) r раз по t, имеем

$$\varphi^{(r)}(t) = i^r \int_{-\infty}^{\infty} x^r \exp itx \, dF(x). \qquad (4.8)$$

Этот интеграл существует, поскольку он ограничен. Действительно,

$$\left| \int_{-\infty}^{\infty} x^r \exp itx \, dF(x) \right| \leqslant \int_{-\infty}^{\infty} |x^r \exp itx| \, dF(x) = \int_{-\infty}^{\infty} |x^r| \, dF(x)$$

и последний интеграл существует, так как $r \leqslant n$.

Полагая в (4.8) $t = 0$, получим (4.7).

Эта теорема указывает способ вычисления начальных моментов случайной величины по ее характеристической функции.

Существование начального момента n-го порядка случайной величины X означает, что ее характеристическая функция может быть разложена в ряд Маклорена по крайней мере до члена, содержащего $(it)^n/n!$, причем коэффициентами у этих членов являются начальные моменты X:

$$\varphi(t) = \sum_{r=0}^{n} \frac{\varphi^{(r)}(0)}{r!} t^r + 0(t^n) = 1 + \sum_{r=1}^{n} \mathbf{M} X^r \frac{(it)^r}{r!} + 0(t^n). \qquad (4.9)$$

Формула (4.4) при $a = 1$ и $b = -\alpha_1$ дает характеристическую функцию центрированной случайной величины

$$\overset{\circ}{\varphi}(t) = \varphi(t) \exp(-i\alpha_1 t). \qquad (4.10)$$

Поскольку центральные моменты случайной величины X есть согласно (3.28) начальные моменты $\overset{\circ}{X}$, то применение формулы (4.7) к (4.10) дает центральный момент порядка r случайной величины X, а ряд Маклорена функции (4.10) имеет коэффициентами центральные моменты величины X:

$$\overset{\circ}{\varphi}(t) = 1 + \sum_{r=1}^{n} \mathbf{M} \overset{\circ}{X}{}^r \frac{(it)^r}{r!} + 0(t^n). \qquad (4.11)$$

Наиболее важные моменты случайной величины — математическое ожидание и дисперсию удобно вычислять по логарифму характеристической функции. Действительно,

$$\frac{d\ln\varphi(t)}{dt}=\frac{\varphi'(t)}{\varphi(t)};$$

$$\frac{d^2\ln\varphi(t)}{dt^2}=\frac{\varphi''(t)\varphi(t)-[\varphi'(t)]^2}{\varphi(t)^2}.$$

Принимая во внимание (3.40), (4.7) и $\varphi(0)=1$, получим:

$$\frac{d\ln\varphi(0)}{dt}=\frac{\varphi'(0)}{\varphi(0)}=i\mathbf{M}X; \qquad (4.12)$$

$$\frac{d^2\ln\varphi(0)}{dt^2}=\frac{\varphi''(0)-[\varphi'(0)]^2}{\varphi(0)^2}=i^2\mathbf{M}X^2-[i\mathbf{M}X]^2=-\mathbf{D}X. \qquad (4.13)$$

Производные различного порядка от логарифма характеристической функции в точке $t=0$, умноженные на соответствующие степени i, используются в качестве числовых характеристик случайной величины. Эти числовые характеристики называются *семиинвариантами* (или полуинвариантами) случайной величины. Введем обозначение λ_r для семиинварианта порядка r и $\psi(t)$ для логарифма характеристической функции: $\psi(t)=\ln\varphi(t)$. По определению

$$\lambda_r=-i^r\psi^{(r)}(0). \qquad (4.14)$$

Ряд Маклорена для $\psi(t)$, если семиинвариант порядка n существует, имеет вид

$$\psi(t)=\ln\varphi(t)=\sum_{r=0}^{n}\lambda_r\frac{(it)^r}{r!}+0(t^n). \qquad (4.15)$$

Можно показать, что семиинвариант любого порядка n есть (целая) рациональная функция первых n моментов. В частности,

$$\left.\begin{array}{ll} 1.\ \lambda_1=\alpha_1; & 3.\ \lambda_3=\mu_3; \\ 2.\ \lambda_2=\mu_2; & 4.\ \lambda_4=\mu_4-3\mu_2^2. \end{array}\right\} \qquad (4.16)$$

Семиинвариант суммы независимых случайных величин равен сумме семиинвариантов слагаемых.

Пример 4.1. Найдем характеристическую функцию случайной величины X, распределенной по биномиальному закону (2.3). X принимает значение k, $k=\overline{0,n}$ с вероятностью (2.4), поэтому

$$\varphi(t)=\mathbf{M}\exp iXt=\sum_{k=0}^{n}\exp iktC_n^kp^kq^{n-k}=$$

$$=\sum_{k=0}^{n}C_n^k(p\exp it)^kq^{n-k}=(q+p\exp it)^n.$$

Окончательно

$$\varphi(t) = (q + p \exp it)^n. \tag{4.17}$$

Воспользуемся (4.17), (4.12) и (4.13) для нахождения **M**X и **D**X:

$$\psi(t) = \ln \varphi(t) = n \ln (q + p \exp it);$$

$$\frac{d\psi(t)}{dt} = \frac{npi \exp it}{(q + p \exp it)}.$$

Отсюда **M**$X = np$.

$$\frac{d^2\psi(t)}{dt^2} = \frac{npi^2 \exp it\,(q + p \exp it) - np^2 i^2 \exp 2it}{(q + p \exp it)^2}.$$

$$\mathbf{D}\,X = -\,i^2\,(np - np^2) = npq.$$

Пусть компоненты (X_1, X_2) независимы и распределены по биномиальным законам с параметрами n_1, p_1 и n_2, p_2 соответственно. Характеристическая функция $Y = X_1 + X_2$ согласно (4.5)

$$\varphi_y(t) = \varphi_1(t)\,\varphi_2(t) = (q_1 + p_1 \exp it)^{n_1}(q_2 + p_2 \exp it)^{n_2}.$$

Если $p_1 = p_2 = p$, то

$$\varphi_y(t) = (q + p \exp it)^{n_1 + n_2}. \tag{4.18}$$

Учитывая взаимно однозначное соответствие между характеристическими функциями и функциями распределения, на основании (4.18) заключаем, что Y имеет биномиальное распределение с параметрами $(n_1 + n_2)$, p.

Пример 4.2. Случайная величина X распределена по закону Пуассона (2.5). X принимает значение k, $k = 0, 1, 2, \ldots$ $\ldots, n, \ldots$ с вероятностью (2.6), поэтому

$$\varphi(t) = \mathbf{M} \exp iXt = \sum_{k=0}^{\infty} \exp ikt \frac{a^k}{k!} \exp(-a) =$$

$$= \exp(-a) \sum_{k=0}^{\infty} \frac{1}{k!} (a \exp it)^k = \exp a\,(\exp it - 1).$$

Окончательно

$$\varphi(t) = \exp a\,(\exp it - 1). \tag{4.19}$$

Найдем **M**X и **D**X по (4.19):

$$\psi(t) = a\,(\exp it - 1);$$

$$\psi'(t) = ai \exp it.$$

Отсюда

$$\mathbf{M}X = a. \tag{4.20}$$

$$\psi''(t) = ai^2 \exp it.$$

Отсюда

$$\mathbf{D}X = a. \tag{4.21}$$

Из (4.19) следует, что сумма двух независимых случайных величин X_1 и X_2, каждая из которых распределена по закону Пуассона с параметрами a_1 и a_2 соответственно, также распределена по закону Пуассона, но с параметром $a = a_1 + a_2$. Действительно, характеристическая функция суммы $Y = X_1 + X_2$

$$\varphi_y(t) = \varphi_1(t)\,\varphi_2(t) = \exp\left[(a_1 + a_2)(\exp it - 1)\right].$$

Отсюда

$$P(Y = k) = \frac{(a_1 + a_2)^k}{k!} \exp\left[-(a_1 + a_2)\right] = \frac{a^k}{k!} \exp(-a). \tag{4.22}$$

Пример 4.3. Найдем характеристическую функцию случайной величины Y, распределенной по нормальному закону:

$$f_y(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right). \tag{4.23}$$

$$\varphi_y(t) = \mathbf{M} \exp iYt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(ity - \frac{y^2}{2}\right) dy.$$

Преобразуем показатель экспоненты, выделив в нем квадрат суммы:

$$\frac{y^2}{2} - ity = \frac{y^2}{2} - ity + \frac{(it)^2}{2} - \frac{(it)^2}{2} = \frac{1}{2}(y - it)^2 + \frac{t^2}{2}.$$

$$\varphi_y(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}(y - it)\right] dy =$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \lim_{a \to \infty} \int_{-a}^{a} \exp\left[-\frac{1}{2}(y - it)\right] dy.$$

Переменную интегрирования заменим по формуле

$$z = (y - it); \qquad dz = dy:$$

$$\varphi_y(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \lim_{a \to \infty} \int_{-a-it}^{a-it} \exp\left(-\frac{z^2}{2}\right) dz.$$

Рассмотрим интеграл от комплексной функции $\exp\left(-\frac{z^2}{2}\right)$ по прямоугольному контуру C (рис. 4.1). В силу аналитичности функции $\exp\left(-\frac{z^2}{2}\right)$

$$\int_C \exp\left(-\frac{z^2}{2}\right) dz = \int_{-a}^{a} \exp\left(-\frac{z^2}{2}\right) dz + \int_{a-it}^{-a-it} \exp\left(-\frac{z^2}{2}\right) dz +$$

$$+ \int_{a}^{a-it} \exp\left(-\frac{z^2}{2}\right) dz + \int_{-a-it}^{-a} \exp\left(-\frac{z^2}{2}\right) dz = 0.$$

Очевидно, что сумма двух последних интегралов при $a \to \infty$ сходится к нулю. Следовательно, с учетом (2.18)

$$\lim_{a\to\infty} \int_{-a-it}^{a-it} \exp\left(-\frac{z^2}{2}\right) dz =$$

$$= \lim_{a\to\infty} \int_{-a}^{a} \exp\left(-\frac{z^2}{2}\right) dz =$$

$$= \sqrt{2\pi},$$

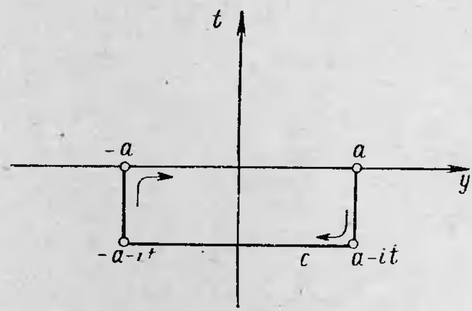

Рис. 4.1

поэтому

$$\varphi_y(t) = \exp\left(-\frac{t^2}{2}\right). \tag{4.24}$$

Теперь положим, что $Y = \frac{X-m}{\sigma}$, т. е. $X = \sigma Y + m$. Согласно (2.116)

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right].$$

Характеристическую функцию X находим по формуле (4.4):

$$\varphi_x(t) = \exp itm\, \varphi_y(t\sigma) = \exp itm \exp\left(-\frac{1}{2}\sigma^2 t^2\right).$$

Таким образом, характеристическая функция нормального закона (2.9)

$$\varphi(t) = \exp itm - \frac{1}{2}\sigma^2 t^2. \tag{4.25}$$

Из (4.25) прежним способом получим, что $\mathbf{M}X = m$ и $\mathbf{D}X = \sigma^2$.

Пример 4.4. Пусть X — число появлений событий A в серии из n испытаний, в каждом из которых вероятность появления A равна p и не зависит от исходов предыдущих испытаний. Исход серии из n испытаний есть n-мерный вектор $(X_1, X_2, \ldots, X_n)$, в котором компонента X_k — число появлений события A в испытании номер k. X_k, $k=\overline{0, n}$ принимает значение нуля с вероятностью $q=1-p$ и единицы с вероятностью p (пара p, q есть ряд распределения X_k), поэтому характеристическая функция X_k, согласно (4.3),

$$\varphi_k(t) = q \exp it0 + p \exp it1 = q + p \exp it.$$

Очевидно, что $X = \sum_{k=1}^{n} X_k$, поэтому на основании (4.6) характеристическая функция X равна:

$$\varphi(t) = \prod_{k=1}^{n} \varphi_k(t) = (q + p \exp it)^n.$$

Сравнение этой характеристической функции с (4.17) показывает, что случайная величина X распределена по биномиальному закону с параметрами n и p.

Для каждой функции распределения можно найти ее характеристическую функцию (4.1). И, наоборот, зная характеристическую функцию какого-либо распределения, можно определить само распределение. Для этого служит формула обращения Фурье. В случае непрерывного распределения формула обращения имеет вид

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-itx)\, \varphi(t)\, dt. \qquad (4.26)$$

Для дискретного распределения, когда случайная величина X принимает значение x_k, $k=1, 2, \ldots$ с вероятностью p_k, формула обращения Фурье имеет вид

$$p_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-itx_k)\, \varphi(t)\, dt, \qquad (4.27)$$

где

$$\varphi(t) = \sum_k p_k \exp itx_k.$$

Действительно,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-itx_k)\, \varphi(t)\, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-itx_k) \left[\sum_n p_n \exp itx_n\right] dt =$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\sum_n p_n \exp\left[it(x_n - x_k)\right]\right] dt.$$

Ряд сходится равномерно, поэтому можно переставить знаки суммирования и интегрирования:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}[\sum_n \exp it(x_n - x_k)p_n]\,dt = \frac{1}{2\pi}\sum_n p_n\int_{-\pi}^{\pi}\exp it(x_n - x_k)\,dt =$$

$$= \frac{1}{2\pi}\left[\sum_{n=1}^{k-1} p_n\int_{-\pi}^{\pi}\exp it(x_n - x_k)\,dt + p_k\int_{-\pi}^{\pi}dt + \right.$$

$$\left. + \sum_{n\geqslant k+1} p_n\int_{-\pi}^{\pi}\exp it(x_n - x_k)\,dt\right].$$

Первый и третий интегралы равны нулю, поэтому

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\exp(-itx_k)\,\varphi(t)\,dt = p_k.$$

Из формул (4.26) и (4.27) вытекает единственность закона распределения для заданной характеристической функции. Говорят, что последовательность неубывающих функций $F_1(x), F_2(x), \ldots, F_n(x), \ldots$ *сходится в основном* к неубывающей функции $F(x)$, если при $n \to \infty$ она сходится к $F(x)$ в каждой ее точке непрерывности. Следующие две теоремы, важные в теоретическом отношении, приведем без доказательств, с которыми можно познакомиться в [4].

Прямая предельная теорема. *Если последовательность функций распределения $F_1(x), F_2(x), \ldots, F_n(x), \ldots$ сходится в основном к функции распределения $F(x)$, то последовательность характеристических функций $\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t), \ldots$ сходится к характеристической функции*

$$\varphi(t) = \int_{-\infty}^{\infty}\exp itx\,dF(x),$$

причем эта сходимость равномерна в каждом конечном интервале t.

Обратная предельная теорема. *Если последовательность характеристических функций $\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t), \ldots$ сходится к непрерывной функции $\varphi(t)$, то последовательность функций распределения $F_1(t), F_2(t), \ldots, F_n(t), \ldots$ сходится в основном к некоторой функции распределения $F(x)$ и в силу прямой предельной теоремы*

$$\varphi(t) = \int_{-\infty}^{\infty}\exp itx\,dF(x).$$

Пример 4.5. Пусть X — случайная величина из примера 4.4, т. е. подчинена биномиальному закону с параметрами

n и p. В примере 4.1 мы нашли, что $\mathbf{M}X=np$ и $\sigma=\sqrt{\mathbf{D}X}=\sqrt{npq}$.

Пусть $Y=\frac{X-np}{\sqrt{npq}}$. Согласно (4.4) характеристическая функция Y равна:

$$\varphi_n(t)=\exp\left(-\frac{np}{\sqrt{npq}}\,it\right)\left[q+p\exp\left(\frac{it}{\sqrt{npq}}\right)\right]^n=$$
$$=\left[q\exp\left(-it\sqrt{\frac{p}{nq}}\right)+p\exp\left(it\sqrt{\frac{q}{np}}\right)\right]^n.$$

Воспользуемся этим выражением для того, чтобы выяснить, как при возрастании n изменяется закон распределения Y. Рассмотрим последовательность $\varphi_n(t)$ при $n\to\infty$. Разложим в ряд Маклорена выражение в скобках:

$$q\exp\left(-it\sqrt{\frac{p}{nq}}\right)+p\exp\left(it\sqrt{\frac{q}{np}}\right)=$$
$$=\sum_{k=0}^{\infty}\frac{(it)^k}{k!}\left[(-1)^k\sqrt{\frac{p^kq^2}{n^kq^k}}+\sqrt{\frac{q^kp^2}{n^kp^k}}\right]=$$
$$=1+\frac{(it)^2}{2n}(1+R_n), \tag{4.28}$$

где

$$R_n=2\sum_{k=3}^{\infty}\frac{(it)^{k-2}}{\sqrt{n^{k-2}}}\,\frac{(-1)^k p^kq+q^kp}{k!\sqrt{(pq)^k}}. \tag{4.29}$$

При $n\to\infty$ $R_n\to 0$, поэтому

$$\varphi_n(t)=\left[1-\frac{t^2}{2n}(1+R_n)\right]^n\to\exp\left(-\frac{t^2}{2}\right). \tag{4.30}$$

Таким образом, последовательность характеристических функций $\varphi_n(t)$ при $n\to\infty$ сходится к функции $\exp(-t^2/2)$, но последняя является характеристической функцией нормального закона (4.23). На основании обратной предельной теоремы заключаем, что

$$F_n(y)=P\left(\frac{X-np}{\sqrt{npq}}<y\right)\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}\exp\left(-\frac{t^2}{2}\right)dt \tag{4.31}$$

при $n\to\infty$.

Это соотношение известно в теории вероятностей как интегральная предельная теорема Муавра—Лапласа.

§ 4.3. Характеристическая функция многомерного случайного вектора

Пусть $F(x_1, x_2, \ldots, x_n)$ есть n-мерная функция распределения вектора $X=(X_1, X_2, \ldots, X_n)$.

Определение. Характеристической функцией распределения $F(x_1, x_2, \ldots, x_n)$ (или случайного вектора X) называется функция

$$\varphi(t_1, t_2, \ldots, t_n)=$$
$$=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\ldots\int\limits_{-\infty}^{\infty}\exp\left[i\sum_{k=1}^{n}t_kx_k\right]dF(x_1, x_2, \ldots, x_n), \qquad (4.32)$$

где $t_1, t_2, \ldots, t_n$ — вспомогательные вещественные переменные.

Интеграл в (4.32) сходится абсолютно, поскольку $|\exp[i\sum\limits_{k=1}^{n}t_kx_k]|<1$ при любых t_k, $k=\overline{1,n}$, поэтому, сопоставив (4.32) с (3.14), можем определить характеристическую функцию n-мерного вектора X как математическое ожидание функции $\exp[i\sum\limits_{k=1}^{n}t_kx_k]$:

$$\varphi(t_1, t_2, \ldots, t_n)=\mathbf{M}\exp\left[i\sum_{k=1}^{n}t_kX_k\right]. \qquad (4.33)$$

Аналогично одномерному случаю n-мерная характеристическая функция может быть найдена для любого распределения, она равномерно непрерывна во всем пространстве $-\infty<t_k<+\infty$, $k=\overline{1,n}$ и

1. $\varphi(0, 0, \ldots, 0)=1$;

2. $|\varphi(t_1, t_2, \ldots, t_n)|\leqslant 1$ для $-\infty<t_k<+\infty$, $k=\overline{1,n}$;

3. $\varphi(-t_1, -t_2, \ldots, -t_n)=\overline{\varphi(t_1, t_2, \ldots, t_n)}$. $\qquad (4.34)$

По характеристической функции $\varphi_1(t_1, t_2, \ldots, t_n)$ n-мерного вектора X можно найти характеристическую функцию r-мерного вектора $(X_{k1}, X_{k2}, \ldots, X_{kr})$, $r\leqslant n$, составленного из любой комбинации r компонент вектора X. Для этого необходимо в функции $\varphi(t_1, t_2, \ldots, t_n)$ приравнять нулю те аргументы, номера которых не представлены среди номеров компонент r-мерного вектора.

Пусть, например, $\varphi(t_1, t_2)$ — характеристическая функция случайного вектора (X, Y). Характеристические функции компонент получим следующим образом:

$$\varphi(t, 0) = \mathbf{M} \exp i (tX + 0Y) \equiv \mathbf{M} \exp itX = \varphi_x(t);$$

$$\varphi(0, t) = \mathbf{M} \exp i (0X + tY) = \mathbf{M} \exp itY = \varphi_y(t).$$

На примере функции $\varphi(t_1, t_2)$ познакомимся еще с некоторыми свойствами многомерных характеристических функций. Их справедливость для произвольного n будет очевидной.

Если компоненты вектора (X, Y) независимы, т. е.

$$F(x, y) = F_1(x) F_2(y),$$

то

$$\varphi(t_1, t_2) = \mathbf{M} \exp i(t_1 X + t_2 Y) = \mathbf{M} \exp it_1 X \, \mathbf{M} \exp it_2 Y = \\ = \varphi_x(t_1) \varphi_y(t_2). \qquad (4.35)$$

По характеристической функции вектора можно вычислить различные его моменты. Например, если существует $\mathbf{M}[X^k Y^s]$, то

$$\mathbf{M}[X^k Y^s] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k y^s \, dF(x, y) = \\ = \left| \frac{\partial^{(k+s)} \varphi(t_1, t_2)}{\partial^k t_1 \partial^s t_2} \right|_{t_1 = t_2 = 0} (i)^{-(k+s)}. \qquad (4.36)$$

Найдем характеристическую функцию вектора $(Y_1, Y_2) = (a_1 X_1 + b_1;\ a_2 X_2 + b_2)$ по характеристической функции вектора (X_1, X_2). На основании (4.33) и теорем о математическом ожидании имеем:

$$g_y(t_1, t_2) = \mathbf{M} \exp [i (Y_1 t_1 + Y_2 t_2)] = \\ = \mathbf{M} \exp [it_1 (a_1 X_1 + b_1) + it_2 (a_2 X_2 + b_2)] = \\ = \exp [it_1 b_1 + it_2 b_2] \, \mathbf{M} \exp i [a_1 t_1 X_1 + a_2 t_2 X_2].$$

Окончательно

$$g_y(t_1, t_2) = \exp i (t_1 b_1 + t_2 b_2) \varphi_x [(a_1 t_1), (a_2 t_2)]. \qquad (4.37)$$

Если $\varphi(t_1, t_2)$ — характеристическая функция вектора (X_1, X_2), то $\varphi_Y(t) = \varphi(t, t)$ — характеристическая функция случайной величины $Y = X_1 + X_2$. Действительно,

$$g(t, t) = \mathbf{M} \exp (itX_1 + itX_2) = \mathbf{M} \exp it (X_1 + X_2) = \\ = \mathbf{M} \exp itY = \varphi_y(t). \qquad (4.38)$$

Пример 4.6. Найдем характеристическую функцию вектора (X, Y), распределенного по нормальному закону,

$$f(x_1, x_2) = \\ = \frac{1}{2\pi \sqrt{1 - r^2}} \exp \left[-\frac{1}{2(1 - r^2)} (x_1^2 - 2r x_1 x_2 + x_2^2) \right]. \qquad (4.39)$$

По определению

$$g(t_1, t_2)=\frac{1}{2\pi\sqrt{1-r^2}}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\exp\left[-\frac{1}{2(1-r^2)}\Big(x_1^2-2rx_1x_2+\right.$$

$$\left.+x_2^2-2(1-r^2)it_1x_1-2(1-r^2)it_2x_2\Big)\right]dx_1dx_2.$$

Напомним [5], что уравнение линии второго порядка

$$a_{11}x^2+2a_{12}xy+a_{22}y^2+2a_{13}x+2a_{23}y+a_{33}=0$$

переносом начала координат в центр линии преобразуется к виду:

$$a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2+a'_{33}=0,$$

где $a_3'=I_3I_2^{-1}$;

$$I_2=\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix};\quad I_3=\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}.$$

Применяя это преобразование к выражению в квадратных скобках в показателе экспоненты, получим

$$x_1^2-2rx_1x_2+x_2^2-2(1-r^2)it_1x_1-2(1-r^2)it_2x_2=$$
$$=x_1^2-2rx_1x_2+x_2^2-rt_1t_2(1-r^2)+t_1^2(1-r^2)+t_2^2(1-r^2).$$

Отсюда

$$g(t_1, t_2)=\frac{1}{2\pi\sqrt{1-r^2}}\exp\left[-\frac{1}{2}(t_1^2+rt_1t_2+t_2^2)\right]\times$$

$$\times\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\exp\left[-\frac{1}{2(1-r^2)}(x_1^2-2rx_1x_2+x_2^2)\right]dx_1dx_2. \quad (4.40)$$

Подстановкой

$$x_1=u\sqrt{1-r^2}-rv,\quad x_2=-v$$

с якобианом

$$\left|\frac{\partial(x_1, x_2)}{\partial(u, v)}\right|=\left\|\begin{matrix}\sqrt{1-r^2} & -r\\ 0 & -1\end{matrix}\right\|=\sqrt{1-r^2}$$

(4.40) преобразуется к виду

$$g(t_1, t_2)=\frac{1}{2\pi}\exp\left[-\frac{1}{2}(t_1^2+rt_1t_2+t_2^2)\right]\times$$

$$\times\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\exp\left(-\frac{u^2+v^2}{2}\right)du\,dv.$$

Окончательно

$$g(t_1, t_2) = \exp\left[-\frac{1}{2}(t_1^2 + rt_1t_2 + t_2^2)\right]. \qquad (4.41)$$

Теперь воспользуемся (4.41) и (4.37) для того, чтобы найти характеристическую функцию нормального закона (2.91). Для этого положим, что

$$Y_1 = \sigma_1 X_1 + m_1 \quad \text{и} \quad Y_2 = \sigma_2 X_2 + m_2. \qquad (4.42)$$

Обратное преобразование

$$X_1 = \frac{Y_1 - m_1}{\sigma_1} \quad \text{и} \quad X_2 = \frac{Y_2 - m_2}{\sigma_2}$$

имеет якобиан

$$\left|\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)}\right| = \left\|\begin{matrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2} \end{matrix}\right\| = \frac{1}{\sigma_1\sigma_2}.$$

Воспользовавшись (2.116), получим из (4.39) плотность распределения вектора (Y_1, Y_2)

$$f(y_1, y_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(y_1-m_1)^2}{\sigma_1^2} - \right.\right.$$
$$\left.\left. - 2r\frac{(y_1-m_1)(y_2-m_2)}{\sigma_1\sigma_2} + \frac{(y_2-m_2)^2}{\sigma_2^2}\right]\right\}, \qquad (4.43)$$

совпадающую с (2.91). Применяя (4.37) для преобразования (4.42), найдем характеристическую функцию нормального закона распределения (4.43)

$$\varphi(t_1, t_2) = \exp\left[i(m_1t_1 + m_2t_2) - \right.$$
$$\left. - \frac{1}{2}(\sigma_1^2t_1^2 + 2\sigma_1\sigma_2rt_1t_2 + \sigma_2^2t_2^2)\right]. \qquad (4.44)$$

Положив в (4.44) $t_2 = 0$, получим характеристическую функцию случайной величины Y_1

$$\varphi(t_1, 0) = \exp\left(im_1t_1 - \frac{1}{2}\sigma_1^2t_1^2\right),$$

которая совпадает с (4.25).

Положив в (4.44) $t_1 = t_2 = t$, найдем согласно (4.38) характеристическую функцию суммы случайных величин $Y_1 + Y_2$

$$\varphi(t, t) = \exp\left[i(m_1 + m_2)t - \frac{1}{2}(\sigma_1^2 + 2\sigma_1\sigma_2r + \sigma_2^2)t^2\right]. \qquad (4.45)$$

Нетрудно видеть, что (4.45) есть характеристическая функция случайной величины, подчиняющейся нормальному закону распределения с параметрами

$$m = m_1 + m_2. \tag{4.46}$$

$$\sigma = \sqrt{\sigma_1^2 + 2\sigma_1\sigma_2 r + \sigma_2^2}. \tag{4.47}$$

Найдем момент корреляции случайных величин Y_1 и Y_2. Согласно (3.37) $K_{y_1y_2} = \mathbf{M}[\mathring{Y}_1\mathring{Y}_2]$. Характеристическую функцию вектора $(\mathring{Y}_1, \mathring{Y}_2)$ получим из (4.44), приравняв нулю m_1 и m_2, поскольку математические ожидания $\mathring{Y}_1$ и $\mathring{Y}_2$ равны нулю:

$$\mathring{\varphi}(t_1, t_2) = \exp\left[-\frac{i}{2}(\sigma_1^2 t_1^2 + 2\sigma_1\sigma_2 r t_1 t_2 + \sigma_2^2 t_2^2)\right]. \tag{4.48}$$

Формула (4.36) для $k = s = 1$ дает

$$\mathbf{M}[\mathring{Y}_1\mathring{Y}_2] = (i)^{-2}\left[\frac{\partial^2\mathring{\varphi}(t_1, t_2)}{\partial t_1 \partial t_2}\right]_{t_1 = t_2 = 0}.$$

$$\frac{\partial^2\mathring{\varphi}(t_1, t_2)}{\partial t_1 \partial t_2} = (i)^{-2}\left[(\sigma_1^2 t_1 + \sigma_1\sigma_2 r t_2)(\sigma_2^2 t_2 + \sigma_1\sigma_2 r t_1) - \sigma_1\sigma_2 r\right] \times$$
$$\times \exp\left[-\frac{1}{2}(\sigma_1^2 t_1^2 + 2\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2)\right].$$

Отсюда

$$K_{y_1y_2} = \sigma_1\sigma_2 r, \tag{4.49}$$

что совпадает с выражением (3.56).

§ 4.4. Преобразование Лапласа — Стилтьеса

В некоторых приложениях теория вероятностей имеет дело исключительно с положительными случайными величинами, дискретными или непрерывными. Так, например, в теории надежности рассматриваются такие случайные величины, как время безотказной работы изделия или число отказавших за некоторый интервал времени изделий. Теория массового обслуживания имеет дело с такими случайными величинами, как число заявок на обслуживание, поступивших за некоторый интервал времени, или длина очереди. Для положительных случайных величин вместо преобразования Фурье—Стилтьеса, рассматриваемого обычно в курсах теории вероятностей, удобнее пользоваться преобразованием Лапласа — Стилтьеса.

Преобразованием Лапласа — Стилтьеса функции $g(x)$ называется интеграл

$$\varphi(z) = \int_0^\infty \exp(-zx)\, dg(x), \tag{4.50}$$

где $z=\alpha+i\beta$ — вспомогательный комплексный параметр.

Если $g(x)$ — неубывающая функция и растет не быстрее, чем $C \exp ax$, где C и a — некоторые постоянные, т. е. для любого x выполняется условие

$$|g(x)| < C \exp ax, \tag{4.51}$$

то интеграл (4.50) сходится в комплексной полуплоскости $\mathrm{Re}\, z > a$ ($\mathrm{Re}\, z$ — вещественная часть z), причем в этой полуплоскости $\varphi(z)$ является аналитической функцией. Функции распределения неотрицательных случайных величин удовлетворяют условию (5.51).

Изображением по Лапласу—Стилтьесу функции распределения $F(x)$ случайной величины X (или просто случайной величины X) называется интеграл

$$\varphi(z) = \int_0^\infty \exp(-zx)\, dF(x). \tag{4.52}$$

В дальнейшем преобразование Лапласа—Стилтьеса ради краткости будем называть просто преобразованием Лапласа, а в отличие от него преобразование Лапласа—Римана — обычным преобразованием Лапласа.

Сравнение (4.52) и (3.2) с учетом того, что $F(x)=0$ на отрицательной полуоси, показывает, что изображение по Лапласу случайной величины можно определить как математическое ожидание комплексной функции $\exp(-zX)$:

$$\varphi(z) = \mathbf{M} \exp(-zX). \tag{4.53}$$

Если случайная величина X непрерывна с плотностью распределения $f(x)$, то (4.52) превращается в обычное преобразование Лапласа

$$\varphi(z) = \int_0^\infty \exp(-zx) f(x)\, dx. \tag{4.54}$$

Для дискретной случайной величины X, принимающей значение x_n с вероятностью p_n, $n=0, 1, 2, \ldots$, (4.52) превращается в дискретное преобразование Лапласа

$$\varphi(z) = \sum_n \exp(-zx_n)\, p_n. \tag{4.55}$$

В данном случае можно считать, что x_n, $n=0, 1, 2, \ldots$ есть функция $\psi_1(n)=x_n$, т. е. функция, определенная на множестве $n=0, 1, 2, \ldots$ со значениями x_n, $n=0, 1, 2, \ldots$

Если ряд распределения функции $\psi^{-1}(X)$, принимающей значения $n = 0, 1, 2, \ldots$, совпадает с рядом распределения X, то преобразование ряда распределения случайной величины X можно представить в виде

$$\varphi(z) = \sum_n \exp(-zn)\, p_n = \sum_n t^n p_n, \qquad (4.56)$$

где $\exp(-z) = t$.

Преобразование Фурье—Стилтьеса (4.1) функции распределения неотрицательной случайной величины X можно записать в виде

$$\varphi(t) = \int_{-\infty}^{\infty} \exp itx\, dF(x) = \int_0^{\infty} \exp itx\, dF(x). \qquad (4.57)$$

Это выражение совпадает с (4.52), если в последнем считать, что $z = -it$, т. е. изображения по Фурье и Лапласу неотрицательной случайной величины различаются только обозначением вспомогательной переменной. Для произвольных функций преобразование Фурье отличается от преобразования Лапласа тем, что предъявляет к функции более жесткие условия, необходимые для сходимости интеграла. Для функций распределений случайных величин это различие не имеет значения. Отсюда свойства характеристических функций, о которых шла речь в предыдущем параграфе, справедливы (с учетом обозначений вспомогательной переменной) и для преобразований Лапласа.

В частности, для изображений по Лапласу справедливы прямая и обратная предельные теоремы, остается в силе связь между моментами случайной величины и производными изображений:

$$\mathbf{M} X^n = (-1)^n \left[\frac{\partial^n \varphi(z)}{\partial z^n}\right]_{z=0}. \qquad (4.58)$$

Изображение по Лапласу свертки функций распределения $F_1(x)$ и $F_2(x)$ равно произведению изображений этих функций, т. е. если

$$\varphi_1(z) \risingdotseq F_1(x) \quad \text{и} \quad \varphi_2(z) \risingdotseq F_2(x),$$

то

$$F_1(x) * F_2(x) = \int_0^{\infty} F_1(y - x)\, dF_2(x) =$$

$$= \int_0^{\infty} F_2(y - x)\, dF_1(x) \risingdotseq \varphi_1(z)\, \varphi_2(z). \qquad (4.59)$$

Здесь запись $\varphi(z) \risingdotseq F(x)$, используемая в операционном исчислении, означает, что $\varphi(z)$ есть преобразование Лапласа (изображение) функции (оригинала) $F(x)$.

Перечислим без доказательств, которые можно найти в любом руководстве по операционному исчислеиию, остальные свойства (теоремы) преобразования Лапласа, предполагая, что оригиналы — непрерывные функции. Случай дискретных распределений рассмотрен в следующем параграфе.

1. Если a и b — постоянные, то

$$af_1(x) + bf_2(x) \risingdotseq a\varphi_1(z) + b\varphi_2(z) \text{ (теорема линейности);}$$

2. $f(tx) \risingdotseq \frac{1}{t}\varphi\left(\frac{z}{t}\right)$, $t > 0$ (теорема подобия);

3. $f(x)\exp ax \risingdotseq \varphi(z-a)$ (теорема затухания);

4. $f(x-y) \risingdotseq \varphi(z)\exp(-zy)$ (теорема запаздывания);

5. Если $f(x, \alpha) \risingdotseq \varphi(z, \alpha)$, то

$\frac{\partial f(x, \alpha)}{\partial\alpha} \risingdotseq \frac{\partial\varphi(z, \alpha)}{\partial\alpha}$ (теорема о дифференцировании по параметру);

6. $\frac{df(x)}{dx} \risingdotseq z\varphi(z) - f(0)$ (теорема дифференцирования оригинала);

6′. $\frac{d^nf(x)}{dx^n} \risingdotseq z^n\varphi(z) - z^{n-1}f(0) - \ldots - f^{n-1}(0)$;

7. $\int\limits_0^x f(t)\,dt \risingdotseq \frac{1}{z}\varphi(z)$ (теорема интегрирования оригинала);

8. $-xf(x) \risingdotseq d\varphi(z)/dz$ (теорема дифференцирования изображения);

9. $\frac{\varphi(x)}{x} \risingdotseq \int\limits_z^\infty \varphi(y)\,dy$ (теорема интегрирования изображения);

10. Поведение функции $f(x)$ в окрестности нуля определяется поведением преобразования Лапласа в бесконечности, и, наоборот, поведение функции $f(x)$ в бесконечности определяется поведением функции $\varphi(z)$ в окрестности нуля. В частности,

а) $\lim\limits_{z\to\infty} z\varphi(z) = f(0)$; (4.60)

б) если существует $\lim\limits_{x\to\infty} f(x) = f(\infty)$, то

$$\lim_{z\to 0} z\varphi(z) = f(\infty). \tag{4.61}$$

Функция восстанавливается по своему изображению с помощью формулы обращения

$$f(x) = \frac{1}{2\pi i}\int\limits_{d-i\infty}^{d+i\infty} \exp zx\,\varphi(z)\,dz, \tag{4.62}$$

где интеграл берется по вертикальной прямой $\operatorname{Re} z = d$, $d > a$ и a определяется неравенством $|f(x)| < C\exp ax$.

Пример 4.7. Случайная величина X распределена по экспоненциальному (показательному) закону, если ее плотность распределения выражается формулой

$$f(x)=\begin{cases} 0 & \text{при } x<0; \\ \lambda\exp(-\lambda x) & \text{при } x\geqslant 0, \end{cases} \tag{4.63}$$

где $\lambda>0$.

Найдем изображение по Лапласу экспоненциального закона распределения

$$\varphi_1(z)=\int_0^\infty \exp(-zx)\,\lambda\exp(-\lambda x)\,dx=\frac{\lambda}{z+\lambda}. \tag{4.64}$$

По формуле (4.58) имеем

$$\mathbf{M}X=(-1)\left[\frac{d\varphi_1(z)}{dz}\right]_{z=0}=\left[\frac{\lambda}{(z+\lambda)^2}\right]_{z=0}.$$

$$\mathbf{M}X=\frac{1}{\lambda}. \tag{4.65}$$

$$\mathbf{M}X^2=(-1)^2\left[\frac{2\lambda}{(z+\lambda)^3}\right]_{z=0}=\frac{2}{\lambda^2}.$$

Отсюда, используя (3.40), получим

$$\mathbf{D}X=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}. \tag{4.66}$$

Пример 4.8. Случайная величина X подчиняется гамма-распределению, если ее плотность распределения выражается формулой

$$f(x,\nu,\lambda)=\begin{cases} 0 & \text{при } x<0; \\ \dfrac{\lambda^\nu}{\Gamma(\nu)}x^{\nu-1}\exp(-\lambda x) & \text{при } x\geqslant 0. \end{cases} \tag{4.67}$$

Здесь

$$\nu>0,\ \lambda>0 \text{ и } \Gamma(\nu)=\int_0^\infty x^{\nu-1}\exp(-x)\,dx \text{ — гамма-функция.}$$

Преобразование Лапласа функции гамма-распределения дает

$$\varphi_2(z)=\frac{\lambda^\nu}{\Gamma(\nu)}\int_0^\infty x^{\nu-1}\exp[-(z+\lambda)x]\,dx. \tag{4.68}$$

Если ν — натуральное число, то $\Gamma(\nu)=(\nu-1)!$ и последовательным интегрированием по частям из (4.68) получим:

$$\varphi_2(z)=\frac{\lambda^\nu}{(z+\lambda)^\nu}. \tag{4.69}$$

Сопоставив (4.69) с (4.64), видим, что $\varphi_2(z)=[\varphi_1(z)]^{\nu}$. Это равенство означает, что если величина Y распределена по закону (4.67), то $Y=\sum_{k=1}^{\nu} X_k$, причем случайные величины X_k, $k=\overline{1,\nu}$ независимы и каждая из них распределена по экспоненциальному закону с параметром λ.

Используя (4.58), получим

$$MY=(-1)\left[\frac{d\varphi_2(z)}{dz}\right]_{z=0}=(-1)\left[-\frac{\lambda^{\nu}\nu}{(z+\lambda)^{\nu+1}}\right]_{z=0}.$$

$$MY=\frac{\nu}{\lambda}. \tag{4.70}$$

$$MY^2=(-1)^2\left[\frac{d^2\varphi_2(z)}{dz^2}\right]_{z=0}=\left[\frac{\nu\lambda^{\nu}(\nu+1)}{(z+\lambda)^{\nu+2}}\right]_{z=0}=\frac{\nu^2+\nu}{\lambda^2}.$$

Отсюда

$$DY=\frac{\nu^2+\nu}{\lambda^2}-\left[\frac{\nu}{\lambda}\right]^2=\frac{\nu}{\lambda^2}. \tag{4.71}$$

Если учесть (4.65) и (4.66), то выражения (4.70) и (4.71) можно было бы получить на основании теорем о математическом ожидании и дисперсии суммы случайных величин.

§ 4.5. Производящие функции

Важный класс дискретных случайных величин составляют случайные величины, принимающие только целые значения $k=0, 1, 2, \ldots, n, \ldots$ Такие случайные величины называются целочисленными. Операционный метод для целочисленных случайных величин основывается на аппарате производящих функций.

Определение 1. *Обыкновенной производящей функцией для последовательности чисел (конечной или счетной)* $\{a_n\}=\{a_0, a_1, \ldots, a_n, \ldots\}$ *называется сумма.*

$$A(t)=a_0+a_1t+a_2t^2+\ldots+a_nt^n+\ldots \tag{4.72}$$

Определение 2. *Экспоненциальной производящей функцией для той же последовательности* $\{a_n\}$ *называется сумма*

$$E(t)=a_0+a_1t+a_2\frac{t^2}{2!}+\ldots+a_n\frac{t^n}{n!}+\ldots \tag{4.73}$$

Последовательностью чисел $\{a_n\}$, связанной с целочисленной случайной величиной X, может быть ряд распределения X либо последовательность моментов X возрастающих порядков. В первом случае производящая функция называется производящей функцией вероятностей (ПФВ), поскольку

$a_n = P(X=n)$, $n=0, 1, 2, \ldots$, а во втором — производящей функцией моментов (ПФМ), поскольку $a_n = \mathbf{M}X^n$ или $a_n = \mathbf{M}[X - a_1]^n$.

Переменная t в (4.72) и (4.73) играет вспомогательную роль. Если последовательность $\{a_n\}$ ограничена, т. е. существует такое конечное число $K > 0$, что $|a_n| \leqslant K$, $n=0, 1, 2, \ldots$, то ряд (4.72) сходится при $|t| < 1$, а ряд (4.73) сходится для всех значений t. Однако, вообще говоря, нет необходимости считать t действительной или комплексной переменной величиной. Выражения (4.72) и (4.73), если речь идет только о численных значениях коэффициентов при степени t, можно рассматривать как формальные конструкции, а t — как абстрактный или неопределенный символ, роль которого, как и его степеней, сводится к тому, чтобы различать элементы последовательности, объединенные в сумму. При такой трактовке символ t оказывается вспомогательным средством для алгебраических преобразований с последовательностями элементов. Формальные операции над производящими функциями, такие, как сложение, умножение, дифференцирование и интегрирование по t, помогают сформулировать определения и выявить зависимости в алгебре рядов $A(t)$ или $E(t)$, в частности путем приравнивания коэффициентов при одинаковых степенях t после выполнения этих операций.

Обозначим через $A(t)$, $B(t)$ и $C(t)$ обыкновенные производящие функции, соответствующие последовательностям $\{a_n\}$, $\{b_n\}$ и $\{c_n\}$.

Суммой обыкновенных производящих функций $A(t)$ и $B(t)$

$$C(t) = A(t) + B(t) \tag{4.74}$$

называется обыкновенная производящая функция последовательности $c_n = a_n + b_n$, $n=0, 1, 2, \ldots$

Произведением обыкновенных производящих функций $A(t)$ и $B(t)$

$$C(t) = A(t)B(t) \tag{4.75}$$

называется обыкновенная производящая функция последовательности

$$c_n = \sum_{k=0}^{n} a_{n-k} b_k, \quad n = 0, 1, 2 \ldots \tag{4.76}$$

Это выражение получается путем приравнивания коэффициентов при одинаковых степенях t в обеих частях равенства (4.75). Если X и Y — независимые целочисленные случайные величины, причем $\{a_n\}$ — ряд распределения X, а $\{b_n\}$ — ряд распределения Y, то

$$a_{n-k} b_k = P[X = (n-k); \quad Y = k],$$

поэтому

$$\sum_{k=0}^{n} a_{n-k} b_k = \sum_{k=0}^{n} P[X=(n-k);\ Y=k] = P(X+Y=n),$$

т. е. последовательность $\{c_n\}$ есть ряд распределения случайной величины $X+Y$.

Функцию распределения $F(y)=\sum_{n=0}^{s} c_n$, где $s<y\leqslant s+1$, случайной величины $X+Y$ можно получить сверткой (2.129) функций распределения случайных величин X и Y. Пусть

$$F_1(x)=P(X<x)=\sum_{i=0}^{m} a_i;$$

$$F_2(x)=P(Y<x)=\sum_{i=0}^{m} b_i,$$

где $m<x\leqslant m+1$.

По формуле (2.128) получим

$$F(y)=\int_{-\infty}^{\infty} F_1(y-x)\,dF_2(x)=b_0\sum_{i=0}^{s} a_i+b_1\sum_{i=0}^{s-1} a_i+$$

$$+b_2\sum_{i=0}^{s-2} a_i+\ldots+b_s a_0=a_0b_0+(a_1b_0+a_0b_1)+$$

$$+(a_2b_0+a_1b_1+a_0b_2)+\ldots+b_s a_0=\sum_{n=0}^{s} c_n.$$

Если последовательность $\{c_n\}$ есть свертка последовательностей $\{a_n\}$ и $\{b_n\}$, то это записывают $\{c_n\}=\{a_n\} * \{b_n\}$.

Часто оказываются полезными следующие соотношения, полученные формальным дифференцированием функции $A(t)$ по t:

$$\frac{dA(t)}{dt}=\sum_{k=1}^{\infty} k a_k t^{k-1};$$

$$\frac{d^2A(t)}{dt^2}=\sum_{k=2}^{\infty} k(k-1)a_k t^{k-2};$$

. . . .

$$\frac{d^nA(t)}{dt^n}=\sum_{k=n}^{\infty} k(k-1)\ldots(k-n+1)a_k t^{k-n}=\sum_{k=n}^{\infty}(k)_n t^{k-n}.$$

Для сохранения показателей степени у t удобно пользоваться оператором $\theta = t\frac{d}{dt}$:

$$\theta A(t) = \sum_{k=0}^{\infty} k a_k t^k;$$

$$\theta^2 A(t) = \sum_{k=0}^{\infty} k^2 a_k t^k;$$

$$\cdots\cdots\cdots$$

$$\theta^n A(t) = \sum_{k=0}^{\infty} k^n a_k t^k.$$

В частности, если $a_k = P(X=k)$, то

$$\theta^n A(1) = \sum_{k=0}^{\infty} k^n a_k = \mathbf{M} X^n. \tag{4.77}$$

Обозначим через $E(t)$, $F(t)$, $G(t)$ экспоненциальные производящие функции последовательностей $\{a_n\}$, $\{b_n\}$ и $\{c_n\}$.

Суммой экспоненциальных производящих функций $E(t)$ и $F(t)$

$$G(t) = E(t) + F(t) \tag{4.78}$$

называется экспоненциальная производящая функция последовательности $c_n = a_n + b_n$, $n = 0, 1, 2, \ldots$

Произведением экспоненциальных производящих функций $E(t)$ и $F(t)$

$$G(t) = E(t)\,F(t) \tag{4.79}$$

называется экспоненциальная производящая функция последовательности

$$c_n = b_n a_0 + n b_{n-1} a_1 + \ldots + C_n^r b_{k-r} a_r + \ldots + b_0 a_n, \quad n = 0, 1, 2, \ldots \tag{4.80}$$

Это выражение можно получить, приравняв коэффициенты при t^n слева и справа от знака равенства после перемножения $E(t)$ и $F(t)$ по правилам умножения рядов. Но удобнее воспользоваться следующим формальным приемом. Запишем последовательность $\{a_n\}$ в виде $\{a^n\}$ и будем оперировать с верхними индексами как с обычными показателями степени, считая при этом, что $a^0 = 1$ в том и только в том случае, если $a_0 = 1$. После выполнения преобразований вновь вернемся к нижним индексам. При таком условии (4.80) принимает вид

$$c^n = b^n a^0 + n b^{n-1} a^1 + \ldots + C_n^r b^{n-r} a^r + \ldots + b^0 a^n = (b+a)^n, \tag{4.81}$$

где

$$a_n \equiv a^n \text{ и } b_n \equiv b^n,$$

а ряд (4.73)

$$E(t) = a^0 + a^1 t + \frac{a^2t^2}{2!} + \ldots + \frac{a^n t^n}{n!} + \ldots = \exp at.$$

Аналогично $F(t) = \exp bt$ и $G(t) = \exp ct$, поэтому (4.79) можно представить в виде

$$\exp ct = \exp at \exp bt = \exp (a + b) t.$$

Приравняв в разложениях экспонент коэффициенты при t^n, получим (4.81) и, следовательно (4.80). Этот прием удобно использовать и в других преобразованиях производящих функций.

Обыкновенная и экспоненциальная производящие функции связаны преобразованием

$$A(t) = \int_0^\infty \exp(-x) E(tx)\, dx. \tag{4.82}$$

Действительно, учитывая, что для целочисленного n

$$\int_0^\infty x^n \exp(-x)\, dx = \Gamma(n+1) = n!,$$

имеем

$$\int_0^\infty \exp(-x) E(tx)\, dx = \sum_{n=0}^\infty a_n \frac{t^n}{n!} \int_0^\infty \exp(-x) x^n\, dx =$$

$$= \sum_{n=0}^\infty a_n t^n = A(t).$$

Пусть целочисленная случайная величина X имеет ряд распределения $p_0, p_1, \ldots, p_n, \ldots$ Обозначим ПФВ символом $P(t)$:

$$P(t) = p_0 + p_1 t + p_2 t^2 + \ldots + p_n t^n + \ldots \tag{4.83}$$

Производящая функция начальных моментов, т. е. последовательности

$$\alpha_k = \mathbf{M} X^k = \sum_{n=0}^\infty n^k p_n, \quad k = 0, 1, 2, \ldots, n, \ldots,$$

которую обозначим символом $M(t)$, связана с ПФВ формулой

$$M(t) = P(\exp t) = \sum_{k=0}^\infty \alpha_k \frac{t^k}{k!}. \tag{4.84}$$

Действительно, заменив в (4.83) t на $\exp t$, получим

$$P(\exp t)=\sum_{n=0}^{\infty} p_n \exp nt=\sum_{n=0}^{\infty} p_n \sum_{k=0}^{\infty} \frac{(nt)^k}{k!}=$$

$$=\sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{n=0}^{\infty} p_n n^k=\sum_{k=0}^{\infty} \alpha_k \frac{t^k}{k!}.$$

Выражение для $M(t)$ можно представить в виде

$$M(t)=\exp \alpha t, \text{ где } \alpha_k \equiv \alpha^k.$$

Формулу (3.33), связывающую центральные и начальные моменты, представим в следующем виде, переходя к обозначениям $\alpha_k \equiv \alpha^k$ и $\mu_k=\mu^k$:

$$\mu_k=\sum_{r=0}^{k} C_k^r \alpha_{k-r}(-\alpha_1)^r=\sum_{r=0}^{k} C_k^r \alpha^{k-r}(-\alpha_1)^r=(\alpha-\alpha_1)^k,$$

или

$$\mu^k=(\alpha-\alpha_1)^k.$$

Используя это выражение, получим производящую функцию центральных моментов случайной величины X, которую обозначим символом $\mu(t)$:

$$\mu(t)=\exp \mu t=\sum_{k=0}^{\infty} \mu^k \frac{t^k}{k!}=\sum_{k=0}^{\infty}(\alpha-\alpha_1)^k \frac{t^k}{k!}=$$

$$=\exp(\alpha-\alpha_1) t=\exp(-\alpha_1 t) \exp \alpha t=\exp(-\alpha_1 t) \mathbf{M}(t),$$

или

$$\mu(t)=\exp(-\alpha_1 t) \mathbf{M}(t)=\exp(-\alpha_1 t) P(\exp t). \qquad (4.85)$$

Пример 4.9. Найдем ПФВ и ПФМ случайной величины X, распределенной по биномиальному закону (2.3):

$$P(t)=\sum_{k=0}^{n} P(X=k) t^k=\sum_{k=0}^{n} C_n^k p^k q^{n-k} t^k=(q+pt)^n \qquad (4.86)$$

Это выражение можно было бы получить из (4.17), заменив $\exp it$ на t.

Для вычисления моментов X воспользуемся формулой (4.77):

$$\theta P(t)=np(q+pt)^{n-1} t;$$

$$\theta^2 P(t)=npt(q+pt)^{n-1}[1-p(n-1)].$$

Отсюда

$$\mathbf{M} X=\theta P(1)=np;$$

$$\mathbf{M} X^2=\theta^2 P(1)=npq+n^2 p^2,$$

что совпадает с выражениями (3.8) и (3.42), полученными непосредственными вычислениями.

Производящие функции моментов получим по формулам (4.84) и (4.85):

$$M(t)=(q+p\exp t)^n; \tag{4.87}$$

$$\mu(t)=\exp(-npt)(q+p\exp t)^n= \\ =(q\exp(-pt)+p\exp qt)^n. \tag{4.88}$$

Разложив (4.88) в ряд Маклорена, найдем, что коэффициент при $t^2/2$ $\mu_2=npq$, т. е. совпадает с выражением для дисперсии X, полученным в примере 3.6.

Пример 4.10. Для случайной величины X, распределенной по закону Пуассона (2.5), ПФВ $P(t)$ получим заменой в (4.19) $\exp it$ на t:

$$P(t)=\exp a(t-1). \tag{4.89}$$

Отсюда по формуле (4.84)

$$M(t)=\exp[a(\exp t-1)]. \tag{4.90}$$

Коэффициент при t в разложении по степеням t этого выражения

$$\alpha_1=\frac{d}{dt}M(t)\Big|_{t=0}=a\exp t\exp[a(\exp t-1)]\Big|_{t=0}=a,$$

что совпадает с результатом, полученным в примере 4.2.

Производящая функция центральных моментов согласно (4.85)

$$\mu(t)=\exp(-at)\exp[a(\exp t-1)]. \tag{4.91}$$

Производящие функции можно ввести и для последовательности чисел с двойной индексацией. Пусть, например, пара целочисленных случайных величин (X, Y) имеет распределение

$$p_{jk}=P(X=j,\ Y=k),\quad j,k=0,1,2,\ \ldots$$

Производящая функция вероятностей этого вектора есть сумма

$$P(t_1,t_2)=\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}p_{jk}t_1^jt_2^k. \tag{4.92}$$

Отметим свойства этой производящей функции, не останавливаясь на их доказательствах.

1. Производящие функции вероятностей случайных величин X и Y определяются формулами $P_x(t)=P(t,1)$ и $P_y(t)=$ $=P(1,t)$.

2. ПФВ суммы $X+Y$ равна $P(t,t)$.

3. Случайные величины X и Y независимы тогда и только тогда, когда

$$P(t_1, t_2) = P_x(t_1)\, P_y(t_2).$$

В заключение отметим, что в литературе по теории линейных импульсных систем для изучения числовых последовательностей или функций целочисленного аргумента (решетчатых функций) используются операционные исчисления, аналогичные исчислению производящих функций, а потому пригодные и для задач теории вероятностей. Преобразования этих исчислений являются частными случаями преобразования Лапласа—Стилтьеса и различаются только формой записи вспомогательной переменной. Одно из них называется D-преобразованием или просто дискретным преобразованием Лапласа, а второе — z-преобразованием. Естественно, что свойства этих преобразований аналогичны свойствам характеристических и производящих функций.

Дискретное преобразование Лапласа (D-преобразование) последовательности $\{p_n\}$, $n=0, 1, 2, \ldots$, т. е. ряда распределения целочисленной случайной величины X, определяется формулой (4.55), в которой $x_n = n$. В литературе [12] эта формула представляется в следующем виде:

$$F^*(q) = \sum_{n=0}^{\infty} p_n \exp(-qn), \tag{4.93}$$

где q — вспомогательная комплексная переменная.

Выражение для z-преобразования можно получить из (4.93), заменив вспомогательную переменную по формуле $z = \exp q$:

$$F(z) = \sum_{n=0}^{\infty} p_n \frac{1}{z^n}. \tag{4.94}$$

Выражения $F^*(q)$ и $F(z)$, как следует из определения преобразования Лапласа—Стилтьеса, рассматриваются уже не как формальные суммы, а как функции комплексного переменного, поэтому необходима сходимость рядов (4.93) и (4.94). Условие сходимости обоих рядов для произвольной последовательности $\{a_n\}$, $n=0, 1, 2, \ldots$ является частным случаем (4.51): если существуют постоянные $C>0$ и a, такие, что

$$|a_n| < C \exp an, \quad n = 0, 1, 2, \ldots, \tag{4.95}$$

то ряд (4.94) для последовательности $\{a_n\}$ сходится в области $|z| > \exp a$, причем $F(z)$ является аналитической функцией, а ряд (4.93) сходится в полуплоскости $\operatorname{Re} q > a$ и $F^*(q)$ также является аналитической функцией.

Различие в обозначении вспомогательной переменной для

D- и z-преобразований приводит к различным формулам обратных преобразований. Восстановление последовательности $\{a_n\}$ по известному z-преобразованию $F(z)$ сводится к разложению $F(z)$ в ряд Лорана в окрестности бесконечно удаленной точки. Коэффициенты ряда вычисляются по формуле

$$a_n = \frac{1}{2\pi i} \int\limits_C F(z)\, z^{n-1}\, dz, \tag{4.96}$$

где C — любая замкнутая кривая, окружающая все особые точки $F(z)$, например окружность с центром в начале координат и достаточно большого радиуса.

Формула для обратного D-преобразования получается из (4.96), если учесть, что $F^*(q)$ — периодическая вдоль мнимой оси функция:

$$F^*(q + 2\pi i k) = F^*(q),$$

а преобразование $z = \exp q$ отображает окружность C, $|z| > \exp a$, в прямолинейный отрезок

$$[\gamma - i\pi,\ \gamma + i\pi],\ \gamma > a \quad \text{и} \quad dz = \exp q\, dq,$$

поэтому

$$a_n = \frac{1}{2\pi i} \int\limits_{\gamma - i\pi}^{\gamma + i\pi} F^*(q) \exp nq\, dq. \tag{4.97}$$

Отмеченная связь между характеристическими функциями, производящими функциями и изображениями при D- и z-преобразованиях указывает на возможность перехода от одного преобразования к другому, что может упростить решение задачи, особенно на этапах обратного преобразования.

ГЛАВА 5
ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

§ 5.1. Биномиальный закон распределения

Определение. *Повторные независимые испытания, в каждом из которых появление некоторого события A имеет постоянную вероятность $P(A) = p$ и рассматриваются только два несовместных исхода A и $\bar{A}$, называются испытаниями Бернулли или испытаниями по схеме Бернулли.*

События A принято называть «успехом», а дополнительное событие $\bar{A}$ — «неудачей» и обозначать вероятность «неудачи» буквой q: $P(\bar{A}) = 1 - p = q$. Независимость, о которой идет речь в определении, означает независимость вероятности $P(A)$ в каждом отдельном испытании от исходов других испытаний.

Последовательность нескольких испытаний Бернулли, полученная упорядочиванием их каким-либо способом, порождает последовательность успехов и неудач, которую можно представить последовательностью символов A и $\bar{A}$. Например, при трех испытаниях Бернулли можно наблюдать последовательность $A\bar{A}A$. Каждая такая последовательность есть произведение независимых случайных событий, и ее вероятность, следовательно, равна произведению вероятностей составляющих событий, в частности $P(A\bar{A}A)=p^2q$.

Классическим примером испытаний по схеме Бернулли является бросание монеты. Если монета симметричная и однородная, то можно принять $p=q=\frac{1}{2}$, при этом $P(A\bar{A}A)=$ $=\frac{1}{2^3}$. При бросании игральной кости, считая успехом выпадение шести очков, а неудачей — выпадение любого другого числа очков, можно принять $p=\frac{1}{6}$ и $q=\frac{5}{6}$, если кость симметричная и однородная.

Последовательность нескольких испытаний можно рассматривать как одно испытание. Такие испытания называются композициями испытаний. Пространством элементарных исходов композиции n испытаний Бернулли является множество всех последовательностей из символов A и $\bar{A}$ длины n, число которых согласно (1.36) равно 2^n, так как каждая последовательность есть перестановка n из 2 с повторением.

Найдем вероятность события, которое состоит в том, что при n испытаниях Бернулли k испытаний привели к успеху. При таком определении события безразличен, очевидно, порядок расположения символов A и $\bar{A}$ в последовательности, поэтому число элементарных событий, благоприятствующих нашему событию, равно C_n^k. Поскольку вероятность появления каждой последовательности с k успехами и $n—k$ неудачами равна p^kq^{n-k}, то искомая вероятность

$$b(k, n, p)=C_n^kp^kq^{n-k}.$$

Результат этих рассуждений можно сформулировать в виде теоремы.

Пусть $b(k, n, p)$ есть вероятность того, что n испытаний Бернулли с вероятностью успеха p и вероятностью неудачи q привели k раз к успеху и $n—k$ раз к неудаче $0\leqslant k\leqslant n$, тогда

$$b(k, n, p)=C_n^kp^kq^{n-k}. \tag{5.1}$$

Число успехов, а следовательно, и число неудач в последовательности n испытаний Бернулли, согласно определению из § 2.1, есть случайная величина. Формула (5.1) означает, что эта случайная величина распределена по биномиальному закону с параметрами n и p. По этой причине биномиальный закон распределения иногда называют законом Бернулли.

Некоторые свойства биномиального распределения рассмотрены в примерах предыдущих глав. Перечислим полученные там результаты.

1. Множеством значений X является множество целых чисел $k=0, 1, 2, \ldots, n$.

2. Ряд распределения составляют вероятности (5.1) для $k=\overline{0, n}$.

3. Функция распределения определяется формулой

$$F(x)=\begin{cases} 0 & \text{для } x<0; \\ \sum\limits_{k=0}^{s} C_n^k p^k q^{n-k} & \text{для } 0\leqslant x\leqslant n \text{ и } s<x\leqslant s+1; \\ 1 & \text{для } x>n. \end{cases} \tag{5.2}$$

4. Математическое ожидание

$$\mathbf{M}\,X=np. \tag{5.3}$$

5. Дисперсия

$$\mathbf{D}\,X=npq. \tag{5.4}$$

6. Характеристическая функция

$$\varphi(t)=(q+p^{it})^n. \tag{5.5}$$

7. Производящая функция вероятностей

$$P(t)=(q+pt)^n. \tag{5.6}$$

8. Производящая функция начальных моментов

$$M(t)=(q+p\exp t)^n. \tag{5.7}$$

9. Производящая функция центральных моментов

$$\mu(t)=[q\exp(-pt)+p\exp qt]^n. \tag{5.8}$$

10. Интегральная предельная теорема Муавра — Лапласа. *Если X — число наступлений события в n независимых испытаниях, в каждом из которых вероятность этого события p, причем $0<p<1$, то при $n\to\infty$ имеет место соотношение*

$$F(y)=P\left\{\frac{X-np}{\sqrt{npq}}<y\right\}\to\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{y}\exp\left(-\frac{x^2}{2}\right)dx. \tag{5.9}$$

11. Сумма m независимых случайных величин, распределенных по биномиальным законам с параметрами p и n_i, $i=\overline{1, m}$, распределена по биномиальному закону с параметрами p и $n=\sum_{i=1}^{m} n_i$.

Исследуем поведение вероятности $b(k, n, p)$ при постоянных n, p и переменном k, для чего найдем отношение вероятностей при $k+1$ и k

$$\frac{b(k+1, n, p)}{b(k, n, p)} = \frac{n!k!(n-k)!\,p^{k+1}q^{n-k-1}}{(k+1)!(n-k-1)!\,n!\,p^k q^{n-k}} = \frac{p}{q}\cdot\frac{n-k}{k+1}. \qquad (5.10)$$

Из этого соотношения следует:

1) $b(k+1, n, p)>b(k, n, p)$, если $p(n-k)>q(k+1)$, т. е. если $k<np-q$;

2) $b(k+1, n, p)=b(k, n, p)$, если $k=np-q$;

3) $b(k+1, n, p)<b(k, n, p)$, если $k>np-q$;

4) $b(k+1, n, p)<b(k, n, p)$ для всех k, если $np-q<0$.

Мы видим, что вероятность $b(k, n, p)$ при $np-q>0$ с увеличением k сначала возрастает, затем достигает максимума (при значении k близком к математическому ожиданию np) и при дальнейшем росте k убывает (см. рис. 2.2, *б*). Если $np-q$ является целым числом, то максимальное значение вероятность $b(k, n, p)$ принимает для двух значений k: $k_1=np-q$ и $k_2=np-q+1=np+p$. Если же $np-q$ не является целым числом, то максимальное значение вероятность $b(k, n, p)$ достигает при k, равном наименьшему целому числу, превосходящему k_1. Значение k, при котором $b(k, n, p)$ достигает максимального значения, является модой распределения.

Пример 5.1. При считывании двоичных кодов из памяти машины вероятность искажения любого символа кода $p=0{,}005$. Необходимо, чтобы не менее 0,99 всех кодов, поступающих из памяти, были верными. Достаточно ли применить корректирующий код, обнаруживающий и исправляющий одну ошибку, если число разрядов в коде 40?

Вероятность отсутствия искажения в коде согласно (5.1)

$$b(0; 40; 0{,}005) = C_{40}^0 (0{,}005)^0 (0{,}995)^{40} = 0{,}8166.$$

Эта вероятность меньше 0,99, поэтому без специальных мер защиты от ошибок указанное в условии требование не выполняется.

Вероятность одной ошибки

$$b(1; 40; 0{,}005) = C_{40}^1 (0{,}005)^1 (0{,}995)^{39} = 0{,}1641,$$

поэтому

$$P(X \leqslant 1) = 0{,}8166 + 0{,}1641 = 0{,}9807.$$

Эта сумма меньше 0,99, следовательно, применение кода, исправляющего только одну ошибку, недостаточно.

Вероятность двух ошибок

$$b(2; 40; 0{,}005) = C_{40}^{2}(0{,}005)^2 (0{,}995)^{38} = 0{,}0161,$$

поэтому

$$P(X \leqslant 2) = 0{,}9807 + 0{,}0161 = 0{,}9968.$$

Для обеспечения поставленного требования необходимо либо повысить устойчивость работы аппаратуры (т. е. уменьшить p), либо использовать корректирующий код, способный обнаруживать и исправлять две ошибки (при том же числе разрядов).

Схема Бернулли является частным случаем более общей схемы независимых испытаний, в каждом из которых рассматривается некоторая полная группа несовместных событий A_i, $i=\overline{1,m}$ с постоянными вероятностями $p_i = P(A_i)$, $i=\overline{1,m}$. Последовательность n таких испытаний порождает последовательность длиной n, состоящую из событий A_i, $i=\overline{1,m}$, или, другими словами, n-перестановку из генеральной совокупности m различных элементов с повторениями каждого элемента от 0 до n раз включительно.

Вероятность появления любой последовательности из n событий, в которую событие A_1 входит r_1 раз, событие A_2 — r_2 раз и т. д., событие A_m — r_m раз, равна $p_1^{r_1} p_2^{r_2} \ldots p_m^{r_m}$, $r_1+r_2+\ldots+r_m=n$, поскольку события A_i, $i=\overline{1,m}$ независимые. Если не интересоваться порядком появления событий A_i, то вероятность наблюдать при n независимых испытаниях (при композиции n независимых испытаний) r_1 раз событие A_1, r_2 раз событие A_2 и т. д., r_m раз событие A_m равна

$$P_n(r_1, r_2, \ldots, r_m) = \frac{n!}{r_1! r_2! \ldots r_m!} p_1^{r_1} p_2^{r_2} \ldots p_m^{r_m}, \quad (5.11)$$

$$r_1 + r_2 + \ldots + r_m = n.$$

Совокупность вероятностей (5.11) для всех возможных наборов чисел $r_1, r_2, \ldots, r_m$, $r_1+r_2+\ldots+r_m=n$ является законом распределения m-мерного целочисленного случайного вектора $(X_1, X_2, \ldots, X_m)$. Этот закон распределения называется m-мерным полиномиальным законом распределения с параметрами $p_1, p_2, \ldots, p_m$.

§ 5.2. Распределение Пуассона

Представим себе некоторую физическую систему, подверженную мгновенным изменениям или воздействиям, которые обусловлены случайными событиями. Приведем несколько примеров. Совокупность радиоактивных атомов образует систему, претерпевающую мгновенное изменение в результате случайного распада какого-то атома в некоторый момент времени. В системе городского транспорта возникают мгновенные изменения типа автомобильных катастроф. В вычислительной машине возникают изменения типа сбоев вычислений. Запросы на телефонную связь, поступающие на автоматическую телефонную станцию, являются мгновенными воздействиями на эту систему.

Все изменения системы или воздействия на нее считаются подобными друг другу, и мы интересуемся только их общим числом. Каждое изменение можно отмечать точкой на оси времени, в результате чего образуется некоторое случайное распределение точек на оси вещественных чисел. Распределение точек такого вида характеризуется вероятностями $p_k(t)$ того, что в течение интервала времени от 0 до t осуществятся ровно k изменений системы. В частности, $p_0(t)$ есть вероятность того, что не происходит ни одного изменения, а $1-p_0(t)$ — вероятность одного или большего числа изменений.

Во многих случаях физические процессы таковы, что удовлетворяются следующие условия, называемые постулатами процесса Пуассона: каково бы ни было число изменений в период времени от 0 до t, (условная) вероятность того, что в течение промежутка времени $[t, t+\Delta t]$ произойдет изменение, равна $\lambda\Delta t+O(\Delta t)$, а вероятность того, что произойдет более чем одно изменение, есть $O(\Delta t)$. Здесь λ — постоянная, а $O(\Delta t)$ — бесконечно малая величина более высокого порядка, чем Δt.

Из этих условий легко выводятся уравнения для вероятностей $p_k(t)$. Рассмотрим два смежных промежутка времени $[0, t]$ и $[t, t+\Delta t]$, где Δt мало. Если $k\geqslant 1$, то в промежутке $[0, t+\Delta t]$ может произойти ровно k изменений тремя взаимно исключающими друг друга способами:

1) в промежутке $[t, t+\Delta t]$ нет ни одного изменения и k изменений в промежутке $[0, t]$;

2) одно изменение в промежутке $[t, t+\Delta t]$ и $k-1$ изменение за время $[0, t]$;

3) $m\geqslant 2$ изменений в промежутке $[t, t+\Delta t]$ и $n-m$ изменений в промежутке $[0, t]$.

Каждый из этих трех способов — случайное событие, которое в свою очередь есть произведение двух независимых событий. В соответствии с постулатами Пуассона вероятность первого события равна произведению $p_k(t)$ на вероятность того, что в промежутке $[t, t+\Delta t]$ не произойдет ни одного изменения. Эта последняя вероятность равна $1-\lambda\Delta t-O(\Delta t)$. Аналогично вероятность второго события равна $p_{k-1}(t)[\lambda\Delta t+$

$+O(\Delta t)]$. Вероятность третьего события имеет порядок $O(\Delta t)$. Таким образом,

$$p_k(t+\Delta t) = p_k(t)(1-\lambda\Delta t) + p_{k-1}(t)\lambda\Delta t + O_1(\Delta t)$$

или

$$\frac{p_k(t+\Delta t)-p_k(t)}{\Delta t} = -\lambda p_k(t) + \lambda p_{k-1}(t) + \frac{O_1(\Delta t)}{\Delta t}.$$

Переходя к пределу при $\Delta t \to 0$, получим

$$\frac{dp_k(t)}{dt} = -\lambda\left(p_k(t) - p_{k-1}(t)\right), \quad k \geqslant 1. \tag{5.12}$$

Если $k=0$, то вероятности второго и третьего событий равны нулю, поэтому

$$p_0(t+\Delta t) = p_0(t)(1-\lambda\Delta t) + O_2(\Delta t),$$

откуда

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t). \tag{5.13}$$

Поскольку вероятность того, что в промежутке $[0, \Delta t]$ произойдет хотя бы одно изменение, по условию равна $\lambda\Delta t + O(\Delta t)$ и, следовательно, вероятность того, что не произойдет ни одного события, равна $1-\lambda\Delta t - O(\Delta t)$, получим при $\Delta t \to 0$ $p_0(0)=1$ и $p_k(0)=0$ при $k \geqslant 1$.

Учитывая эти условия, интегрированием (5.13), а затем последовательным интегрированием (5.12) находим

$$p_0(t) = \exp(-\lambda t),\ p_1(t) = \lambda t \exp(-\lambda t),\ p_2(t) = \frac{(\lambda t)^2}{2!}, \ldots$$

По индукции

$$p_k(t) = \frac{(\lambda t)^k}{k!}\exp(-\lambda t). \tag{5.14}$$

Если в этой формуле обозначим

$$p_k(t) = P(X=k) = \pi(k, a), \quad a = \lambda t,$$

то она примет вид формулы

$$\pi(k, a) = \frac{a^k}{k!}\exp(-a), \tag{5.15}$$

совпадающей с (2.6).

Следовательно, если некоторая система подвержена мгновенным случайным изменениям таким образом, что выполняются постулаты Пуассона, то число изменений в промежутке времени от 0 до t есть случайная величина X, распределенная по закону Пуассона с параметром $a=\lambda t$. Напомним, что $a=\mathbf{M}X$, поэтому $\lambda = a/t = \mathbf{M}\left[\frac{1}{t}X\right]$ есть математическое ожидание (среднее число) изменений системы в единицу времени. Выражение «мгновенные случайные изменения» следует понимать как изменения, происходящие в моменты времени, предсказать которые заранее невозможно, а длительности их пренебрежимо малы по сравнению с интервалами между соседними изменениями.

Вывод формулы (5.14) из постулатов Пуассона не использует специфических свойств времени. В предыдущих рассуждениях важно не то, что параметр t — время, а то, что объектом рассмотрения является случайное распределение точек на оси вещественных чисел, поэтому t может быть параметром любой физической природы. Упоминание о системах, подверженных мгновенным изменениям, потребовалось, во-первых, для того чтобы отметить физическую реальность условий, сформулированных в постулатах Пуассона, и, во-вторых, уточнить предмет рассмотрения ссылкой на интуитивно понятную картину.

Те же самые рассуждения о распределении случайных событий или точек на оси t применимы и к распределению точек на плоскости или в пространстве. Вместо промежутков длиной t мы можем рассматривать области площадью или объемом t, если примем дополнительное допущение о том, что вероятность иметь k точек в какой-либо определенной области зависит только от площади или объема области, но не от ее формы. Нетрудно для этого случая перефразировать постулаты Пуассона: количества точек в неперекрывающихся областях взаимно независимы и при малом t вероятность иметь более одной точки в области площадью или с объемом t мала по сравнению с t. При этих условиях количество точек в некоторой области есть случайная величина X, распределенная по закону Пуассона с параметром $a=\lambda t$, где λ — средняя плотность точек, а t — площадь или объем области.

Пример 5.2. Табл. 5.1 иллюстрирует случайное распределение точек на плоскости — падений самолетов-снарядов в южной части Лондона во время второй мировой войны [11].

Таблица 5.1

k	0	1	2	3	4	5 и больше
N_k	229	211	93	35	7	1
MX	226,74	211,39	98,54	30,62	7,14	1,57

Вся территория разделена на 576 малых участков площадью $t=1/4$ км^2 каждый. В табл. 5.1 приведены полученные наблюдением числа N_k участков, на которые приходилось по k падений снарядов, и математические ожидания этих чисел

$\mathbf{M}X=N\pi\,(k;\ 0{,}9323)$, рассчитанные по (5.3). Общее число снарядов $T=\Sigma kN_k=537$. Среднее число падений на один участок $a=\lambda t=T/N=0{,}9323$.

Табл. 5.1 показывает, что точки падения снарядов были совершенно случайными. Число попаданий в каждый участок — случайная величина, закон распределения которой хорошо согласуется с законом Пуассона с параметром $a=0{,}9323$.

Перечислим результаты, полученные в примерах предыдущих глав при рассмотрении закона распределения Пуассона.

1. Случайная величина X принимает целочисленные значения из бесконечного множества: $k=0,\ 1,\ 2,\ \dots,\ n,\ \dots$

2. Ряд распределения X составляют вероятности (5.15) при $k=0,\ 1,\ 2,\ \dots,\ n,\ \dots$

3. Функция распределения определяется формулой

$$F(x)=\begin{cases} 0 & \text{для } x<0; \\ \sum\limits_{k=0}^{s}\dfrac{a^k}{k!}\exp(-a) & \text{для } x\geqslant 0 \text{ и } s<x\leqslant s+1. \end{cases} \tag{5.16}$$

4. Математическое ожидание

$$\mathbf{M}X=a. \tag{5.17}$$

5. Дисперсия

$$\mathbf{D}X=a. \tag{5.18}$$

6. Характеристическая функция

$$\varphi(t)=\exp\left[a\,(\exp it-1)\right]. \tag{5.19}$$

7. Производящая функция вероятностей

$$P(t)=\exp a\,(t-1). \tag{5.20}$$

8. Производящая функция начальных моментов

$$M(t)=\exp\left[a\,(\exp t-1)\right]. \tag{5.21}$$

9. Производящая функция центральных моментов

$$\mu(t)=\exp(-at)\exp\left[a\,(\exp t-1)\right]. \tag{5.22}$$

Применяя любой способ вычисления моментов, можно найти:

$$\begin{aligned} &\alpha_0=1; && \mu_0=1; \\ &\alpha_1=a; && \mu_1=0; \\ &\alpha_2=a+a^2; && \mu_2=a; \\ &\alpha_3=a+3a^2+a^3; && \mu_3=a; \\ &\alpha_4=a+7a^2+6a^3+a^4; && \mu_4=a(1+3a). \end{aligned} \tag{5.23}$$

Распределение Пуассона обладает тем свойством, что математическое ожидание, дисперсия и третий центральный момент численно равны между собой: они выражаются одним и тем же числом a, которое является единственным параметром распределения Пуассона, полностью его определяющим.

Показатель асимметрии найдем по формуле (3.68), а эксцесс — по формуле (3.69):

$$S_k = \frac{a}{\sqrt{a^3}} = \frac{1}{\sqrt{a}};$$

$$E_x = \frac{a(1+3a)}{a^2} - 3 = \frac{1}{a} > 0. \qquad (5.24)$$

Найдем производную порядка r от логарифма характеристической функции в точке 0, которая согласно (4.14) определяет семиинвариант λ_r:

$$\frac{d^r}{dt^r}\left[a(\exp it - 1)\right]_{t=0} = i^r a.$$

Отсюда видно, что все семиинварианты распределения Пуассона равны a.

Вернемся к биномиальному распределению с параметрами n и p. Предположим, что параметр n беспредельно увеличивается, принимая значения $n_1 < n_2 < \ldots < n_i < \ldots$, а параметр p уменьшается, принимая значения $p_1 > p_2 > \ldots$ $\ldots > p_i > \ldots$ но так, что произведение $n_i p_i$ имеет конечный предел $a = \lim\limits_{i \to \infty} n_i p_i$. Последовательности пар n_1, p_1; n_2, p_2; ... соответствует последовательность биномиальных законов распределения. Оказывается, что эта последовательность биномиальных законов имеет предельный закон распределения, которым является закон распределения Пуассона. Для доказательства этого проще всего воспользоваться взаимно однозначным соответствием между ПФВ и рядами распределения случайных величин.

ПФВ биномиального закона распределения с параметрами n_i, p_i, $n_i p_i = a_i$ согласно (5.6)

$$P_i(t) = (q_i + p_i t)^{n_i} = (1 - p_i + p_i t)^{n_i} = \left[1 - \frac{a_i(t-1)}{n_i}\right]^{n_i}.$$

Отсюда

$$\lim_{i \to \infty} P_i(t) = \lim_{i \to \infty}\left[1 - \frac{a_i(t-1)}{n_i}\right]^{n_i} = \exp a(t-1), \qquad (5.25)$$

Но $P(t) = \exp a(t-1)$ — производящая функция закона Пуассона с параметром a.

Для уяснения физической стороны этого факта рассмотрим его с другой точки зрения. Обратимся вновь к случайному распределению точек на оси t. Представим себе промежуток $[0, t]$, разделенный на большое число nt частичных промежутков, каждый длиной $1/n$. Произвольный фиксированный частичный промежуток либо пуст, либо содержит по крайней мере одну из случайных точек. Эти две несовместные возможности можно назвать соответственно неудачей и успехом.

Пусть вероятность успеха p_n определяется только длиной промежутка и не зависит от его положения на оси t, а попадания точек в различные промежутки — события независимые. Выполнение этих условий означает, что мы имеем дело со схемой Бернулли, поэтому вероятность k успехов на промежутке $[0, t]$ равна $b(k;\ nt;\ p_n)$. Перейдем к новому разбиению, разделив каждый частичный промежуток пополам. Успех в каждом из частичных промежутков длиной $1/n$ равен сумме успехов (в смысле суммы событий) в его левой и правой половинах, поэтому по формуле (1.17) находим, что $p_n=2p_{2n}-p_{2n}^2$, где p_{2n} — вероятность успеха при новом разбиении. Из этого выражения видно, что $p_n<2p_{2n}$ и, следовательно, $np_n<2np_{2n}$, т. е. произведение np_n монотонно растет при возрастании n, а поэтому необходимо иметь конечный или бесконечный предел. Допущение $np_n\to\infty$ означало бы допущение возможности появления бесконечного числа случайных точек в сколь угодно малом промежутке, так как np_n при законе Бернулли есть математическое ожидание числа точек в промежутке длиной $1/n$. Более естественно предположить, что этот предел конечный, т. е. $\lim\limits_{n\to\infty} np_n=\lambda$.

Введем обозначение

$$\pi(k, \lambda t)=\lim_{n\to\infty} b(k;\ nt;\ p_n)=\lim_{n\to\infty} b(k;\ nt;\ \lambda_n/n)$$

и найдем этот предел. Из формулы (5.10) получим

$$\lim_{n\to\infty}\frac{b(k;\ nt;\ \lambda_n/n)}{b(k-1;\ nt;\ \lambda_n/n)}=\lim_{n\to\infty}\frac{\lambda_n\left(t-\dfrac{k-1}{n}\right)}{k\,(1-\lambda_n/n)}=\frac{\lambda t}{k}.$$

По формуле (5.1) находим:

$$b(0;\ nt;\ \lambda_n/n)=\left(\frac{\lambda_n}{n}\right)^0\left(1-\frac{\lambda_n}{n}\right)^{nt}=\left(1-\frac{\lambda_n}{n}\right)^{nt}. \qquad (5.26)$$

Следовательно,

$$\pi(0;\ \lambda t)=\lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^{nt}=\exp(-\lambda t).$$

Отсюда и из (5.26) находим

$$\lim_{n\to\infty} b(k;\ nt;\ p_n)=\pi(k, \lambda t)=\frac{(\lambda t)^k}{k!}\exp(-\lambda t)=$$

$$=\frac{a^k}{k!}\exp(-a). \qquad (5.27)$$

Это предельное равенство интересно не только в теоретическом отношении, но и в практическом, поскольку указывает, что закон распределения Пуассона является удобным приближением для биномиального распределения с большим n и малым p.

Пример 5.3. Изучение качества продукции, выпускаемой заводом, показало, что доля брака равна 2% всей продукции. Это означает, что вероятность для данного взятого наудачу изделия оказаться бракованным равна $p=0{,}02$. Найдем вероятность того, что в партии, содержащей $n=1000$ изделий, количество бракованных равно 25. По формуле (5.1) имеем

$$b(25;\ 1000;\ 0{,}02)=\frac{1000!}{25!\,975!}\cdot 0{,}02^{25}\cdot 0{,}98^{975}=0{,}045.$$

Используя в качестве приближения формулу (5.27), в которой $a=np=1000\cdot 0{,}02=20$, по таблице [1] для закона распределения Пуассона для $k=25$ и $a=20$ сразу же находим, что $\pi(25;\ 20)=0{,}045$. Но если не пользоваться таблицей, то и в этом случае вычисления по формуле (5.27) значительно проще, чем по формуле (5.1).

§ 5.3. Одномерное нормальное распределение

Нормальный закон распределения занимает среди других законов особое положение в связи с его фундаментальным значением и в теории и в практике. Впервые о нормальном законе распределения упоминается в работах Муавра (1730 г.) в связи с изучением предельного распределения для биномиального закона при $p=0{,}5$ и $n\to\infty$. Впоследствии этот закон распределения исследовался в работах Гаусса (1809 г.) в связи с разработкой основ теории ошибок и Лапласа (1812 г.).

Особое положение нормального закона распределения объясняется двумя обстоятельствами (вообще говоря, первое из них является следствием второго). Во-первых, в природе очень часто встречаются случайные величины, подчиненные нормальному закону распределения. Так, например, по нормальному закону распределены ошибки всех видов измерения, если методика измерения исключает грубые ошибки типа просчетов, отклонения снарядов от цели при артиллерийской стрельбе, значения любых параметров изделий при массовом производстве и стабильном и хорошо контролируемом технологическом процессе, многочисленные характеристики жизненных процессов, протекающих в стабильных условиях, и т. д.

Во-вторых, нормальный закон распределения является предельным законом распределения для многих других зако-

10

нов распределения. Частный случай этого факта выражается предельной теоремой Муавра—Лапласа (4.31). Более подробно этот вопрос рассмотрен в следующей главе.

Перечислим результаты, полученные в примерах предыдущих глав при рассмотрении нормального закона распределения случайной величины X.

1. Функция распределения

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}\exp\left[-\frac{(t-m)^2}{2\sigma^2}\right]dt. \quad (5.28)$$

2. Плотность распределения

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-m)^2}{2\sigma^2}\right]. \quad (5.29)$$

3. Математическое ожидание

$$\mathbf{M}X=m. \quad (5.30)$$

4. Дисперсия

$$\mathbf{D}X=\sigma^2. \quad (5.31)$$

Математическое ожидание и дисперсия нормально распределенной случайной величины однозначно определяют форму кривой функции распределения и плотности распределения.

5. Характеристическая функция

$$\varphi(t)=\exp\left(imt-\frac{1}{2}\sigma^2t^2\right). \quad (5.32)$$

6. Сумма двух нормально распределенных случайных величин X_1 и X_2 с параметрами m_1, σ_1 и m_2, σ_2 и коэффициентом корреляции r распределена нормально с параметрами

$$m=m_1+m_2,$$

$$\sigma=\sqrt{\sigma_1{}^2+2\sigma_1\sigma_2r+\sigma_2{}^2}. \quad (5.33)$$

Заметим, что кривая $y=C\exp(-\beta x^2)$ называется гауссовой кривой, в связи с чем и нормальный закон распределения также иногда называют гауссовым.

Используя (2.116) или (4.4), нетрудно показать, что нормирование (3.65) преобразует случайную величину X, распределенную по нормальному закону (5.28), в случайную величину Y, также распределенную по нормальному закону, но с параметрами $m=0$ и $\sigma=1$, т. е. для нормированной нормально распределенной случайной величины

$$F(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{t^2}{2}\right)dt; \quad (5.34)$$

$$f(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right); \qquad (5.35)$$

$$\varphi(t)=\exp\left(-\frac{t^2}{2}\right). \qquad (5.36)$$

На рис. 2.5, *а* и *б* изображены графики функции (5.34) и (5.35).

Численные расчеты, связанные с нормально распределенными случайными величинами, требуют умения вычислять значения интеграла, входящего в формулу (5.28), при любых значениях x и параметрах m и σ, а также решать обратную задачу — по известному значению $F(x)$ находить значение аргумента x, т. е. квантиль заданного уровня. Но неопределенный интеграл $\int \exp(-t^2/2)\,dt$ не выражается в конечном виде через элементарные функции. При машинных расчетах вероятностей можно использовать различные разложения функции $F(x)$. Вблизи начала координат $F(x)$ имеет участок, близкий к линейному, который хорошо описывается несколькими первыми членами степенного ряда

$$F(x)=\frac{1}{2}+\frac{1}{\sqrt{2\pi}}\left(x-\frac{x^3}{6}+\frac{x^5}{40}-\frac{x^7}{336}+\ldots\right)=$$

$$=\frac{1}{2}+\frac{1}{\sqrt{2\pi}}\sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^{2k-1}}{2^{k-1}(2k-1)(k-1)!}. \qquad (5.37)$$

При достаточно большом аргументе имеет место асимптотическое разложение

$$F(x)\backsim 1-\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\left(\frac{1}{x}-\frac{1}{x^3}+\frac{3}{x^5}-\frac{15}{x^7}+\ldots\right)=$$

$$=1-\frac{1}{x\sqrt{2\pi}}\left[1+\sum_{k=1}^{\infty}(-1)^k\frac{(2k)!}{2^k k!}\cdot\frac{1}{x^{2k+1}}\right]\exp\left(-\frac{x^2}{2}\right). \qquad (5.38)$$

Символ $\backsim$ (асимптотическое равенство) означает, что отношение двух выражений, соединенных этим символом, стремится к единице при неограниченном возрастании x. Ряд (5.38) знакопеременный, поэтому ошибка любой конечной суммы его не превосходит величины первого отброшенного члена. Отсюда

$$1-\frac{1}{x\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)<F(x)<$$

$$<1-\frac{1}{x\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\left(1-\frac{1}{x^2}\right). \qquad (5.39)$$

В справочниках с математико-статистическими таблицами, например [1], приводятся таблицы либо функции распределения нормированной случайной величины

$$F_1(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x}\exp\left(-\frac{t^2}{2}\right)dt, \tag{5.40}$$

либо функции

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{x}\exp\left(-\frac{t^2}{2}\right)dt, \tag{5.41}$$

называемой интегралом вероятностей, либо функции

$$F_2(x)=\frac{2}{\sqrt{\pi}}\int\limits_{0}^{x}\exp(-t^2)\,dt. \tag{5.42}$$

Эти функции обладают следующими свойствами.

1. $$1=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}\exp\left(-\frac{t^2}{2}\right)dt=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{-x}\exp\left(-\frac{t^2}{2}\right)dt+$$
$$+\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x}\exp\left(-\frac{t^2}{2}\right)dt=F_1(-x)+F_1(x),$$

или

$$F_1(-x)=1-F_1(x), \tag{5.43}$$

поэтому функция $F_1(x)$ табулируется только для положительных аргументов.

2. Из (2.18) следует, что

$$\Phi(\infty)=\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{\infty}\exp\left(-\frac{t^2}{2}\right)dt=\frac{1}{2}. \tag{5.44}$$

3. $$\Phi(-x)=\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{-x}\exp\left(-\frac{t^2}{2}\right)dt=$$
$$=-\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{x}\exp\left(-\frac{t^2}{2}\right)dt=-\Phi(x). \tag{5.45}$$

4. $$F_2(\infty)=\frac{2}{\sqrt{\pi}}\int\limits_{0}^{\infty}\exp(-t^2)\,dt=1. \tag{5.46}$$

5. $$F_2(-x)=-F_2(x). \tag{5.47}$$

Для случайной величины X, распределенной по закону (5.28):

1. $F(x)=F_1\left(\frac{x-m}{\sigma}\right)$;

$$P(a\leqslant X<b)=F_1\left(\frac{b-m}{\sigma}\right)-F_1\left(\frac{a-m}{\sigma}\right). \tag{5.48}$$

2. $F(x)=\frac{1}{2}+\Phi\left(\frac{x-m}{\sigma}\right)$;

$$P(a\leqslant X<b)=\Phi\left(\frac{b-m}{\sigma}\right)-\Phi\left(\frac{a-m}{\sigma}\right). \qquad (5.49)$$

3. $F(x)=\frac{1}{2}+F_2\left(\frac{x-m}{\sigma\sqrt{2}}\right)$;

$$P(a\leqslant X<b=\frac{1}{2}\left[F_2\left(\frac{b-m}{\sigma\sqrt{2}}\right)-F_2\left(\frac{a-m}{\sigma\sqrt{2}}\right)\right]. \qquad (5.50)$$

Найдем центральные моменты нормально распределенной случайной величины. С этой целью разложим в степенной ряд характеристическую функцию центрированной случайной величины

$$\varphi(t)=\exp\left(-\frac{1}{2}\sigma^2t^2\right),$$

поскольку этот ряд является производящей функцией центральных моментов X. Функция $\varphi(t)$ — четная по переменной t, поэтому ее степенной ряд содержит только четные степени t. Следовательно, все центральные моменты нечетных порядков нормального распределения равны нулю:

$$\mu_{2k+1}=0, \quad k=0, 1, 2, \ldots \qquad (5.51)$$

Для вычисления коэффициента при t^{2k} воспользуемся формулой Лейбница для производных высшего порядка произведения двух функций:

$$(uv)^{(n)}=\sum_{i=0}^{n}C_n^iu^{(i)}v^{(n-i)}.$$

Получим:

$$\varphi^{(2k)}(t)=\left[\exp\left(-\frac{1}{2}\sigma^2t^2\right)\right]^{(2k)}=\left[-\sigma^2t\exp\left(-\frac{1}{2}\sigma^2t^2\right)\right]^{(2k-1)}=$$

$$=-\sigma^2t\left[\exp\left(-\frac{1}{2}\sigma^2t^2\right)\right]^{(2k-1)}-$$

$$-(2k-1)\sigma^2\left[\exp\left(-\frac{1}{2}\sigma^2t^2\right)\right]^{(2k-2)};$$

$$\varphi^{(2k)}(0)=-(2k-1)\sigma^2\left[\exp\left(-\frac{1}{2}\sigma^2t^2\right)\right]=$$

$$=i^2(2k-1)\sigma^2\varphi^{(2k-2)}(0).$$

Отсюда на основании (4.7) имеем

$$\mu_{2k}=\mathbf{M}\mathring{X}^{2k}=i^{-2k}\varphi^{(2k)}(0)=(2k-1)\sigma^2\mu_{(2k-2)}=$$

$$=(2k-1)(2k-3)\ldots 3\cdot 1\sigma^{2k}\mu_0.$$

Используя очевидное равенство

$$(2k-1)(2k-3)\ .\ .\ .\ 3\cdot 1=\frac{(2k)!}{2^k k!},$$

получим

$$\mu_{2k}=\frac{(2k)!}{2^k k!}\sigma^{2k}. \tag{5.52}$$

Первые четыре центральных момента равны: $\mu_0=1$, $\mu_1=0$, $\mu_2=\sigma^2$, $\mu_3=0$, $\mu_4=3\sigma^4$.

Следовательно, коэффициент асимметрии нормального распределения согласно (3.68) равен нулю, т. е. кривая плотности распределения симметрична относительно математического ожидания. Это видно и непосредственно по формуле (5.29). Симметрия кривой относительно математического ожидания определяется тем, что показателем экспоненты является четная функция отклонения от математического ожидания.

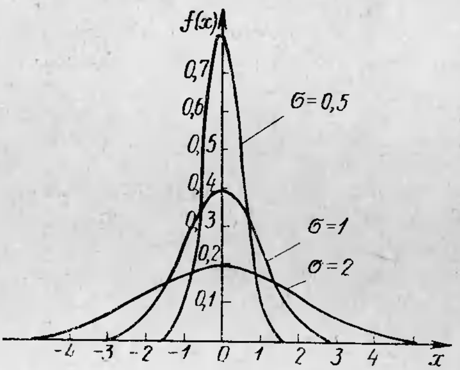

Рис. 5.1

Эксцесс нормального распределения

$$E_x=\frac{\mu_4}{\sigma^4}-3=\frac{3\sigma^4}{\sigma^4}-3=0. \tag{5.53}$$

Нормальное распределение имеет одну моду в точке $x=m$. В этой точке плотность распределения

$$f(m)=f_{\max}=\frac{1}{\sigma\sqrt{2\pi}}. \tag{5.54}$$

Влияние параметра σ на форму кривой плотности распределения показано на рис. 5.1. Нетрудно найти, что абсциссы точек перегиба, в которых кривая плотности распределения имеет максимальную крутизну, равны

$$x=m\pm\sigma. \tag{5.55}$$

Из формул (5.54) и (5.55) видно, что чем больше параметр σ (чем больше дисперсия), тем дальше точка перегиба от моды и тем меньше максимальная плотность вероятности, т. е. тем более разнесены весомые значения $f(x)$. Наоборот, чем меньше параметр σ, тем больше максимальная плотность вероятности и тем ближе точки перегиба к моде, т. е. тем более компактны весомые значения $f(x)$, тем больше вероятность для случайной величины принять значение из интервала, включающего математическое ожидание.

Используя таблицы, можно найти, что в случае нормального распределения случайной величины X

$$\left.\begin{aligned} &P(\mathbf{M}X-\sigma\leqslant X<\mathbf{M}X+\sigma)=0{,}683;\\ &P(\mathbf{M}X-2\sigma\leqslant X<\mathbf{M}X+2\sigma)=0{,}954;\\ &P(\mathbf{M}X-3\sigma\leqslant X<\mathbf{M}X+3\sigma)=0{,}997.\end{aligned}\right\} \qquad (5.56)$$

Отсюда видно, что хотя случайная величина X может отклоняться от математического ожидания на величину сколь угодно большую, однако в действительности почти все наблюдаемые отклонения не будут превосходить трех стандартов.

§ 5.4. Многомерное нормальное распределение

Случайный вектор $X=(X_1, X_2, \ldots, X_n)$ распределен по нормальному закону, если n-мерная плотность распределения его выражается формулой

$$f(x_1, x_2, \ldots, x_n)=\frac{1}{\sigma_1\sigma_2\ldots\sigma_n\sqrt{(2\pi)^n D}}\times$$
$$\times\exp\left[-\frac{1}{2D}\sum_{i=1}^{n}\sum_{k=1}^{n}D_{ik}\frac{(x_i-m_i)(x_k-m_k)}{\sigma_i\sigma_k}\right]. \qquad (5.57)$$

Здесь D — определитель n-го порядка:

$$D=\begin{vmatrix} 1 & r_{12} & \ldots & r_{1n}\\ r_{21} & 1 & \ldots & r_{2n}\\ \cdot & \cdot & \cdot\cdot\cdot & \cdot\\ r_{n1} & r_{n2} & \ldots & r_{nn}\end{vmatrix}; \qquad (5.58)$$

D_{ik} — алгебраическое дополнение элемента r_{ik} определителя D.

Из (5.57) видно, что n-мерная плотность распределения определяется $2n$ параметрами m_i, σ_i и $\frac{1}{2}n(n-1)$ параметрами r_{ik}. Можно показать, что

$$m_i=\mathbf{M}X_i;\quad \sigma_i^2=\mathbf{D}X_i;\quad r_{ik}=\frac{K_{ik}}{\sigma_i\sigma_k},\quad i, k=\overline{1, n},$$

т. е. D есть определитель нормированной корреляционной матрицы вектора X.

Выражение (5.57) можно существенно упростить, а следовательно, упростить и его применение, если воспользоваться матричной записью. Показателем экспоненты в (5.57) является положительно определенная квадратичная форма разностей $(x_i - m_i)$, $i = \overline{1, n}$

$$Q = \sum_{i=1}^{n} \sum_{k=1}^{n} \frac{D_{ik}}{\sigma_i \sigma_k D} (x_i - m_i)(x_k - m_k), \qquad (5.59)$$

которую можно представить в виде

$$Q = (x - m)^T M^{-1} (x - m), \qquad (5.60)$$

где $(x - m)$ — вектор-столбец:

$$(x - m) = \begin{Vmatrix} x_1 - m_1 \\ x_2 - m_2 \\ \cdot\ \cdot\ \cdot\ \cdot \\ x_n - m_n \end{Vmatrix}, \qquad (5.61)$$

$(x - m)^T$ — транспонированный вектор $(x - m)$, т. е. вектор-строка:

$$(x - m)^T = (x_1 - m_1;\ x_2 - m_2;\ \ldots;\ x_n - m_n); \qquad (5.62)$$

M — матрица $n \times n$:

$$M = \| \sigma_i \sigma_k r_{ik} \|, \ i, k = \overline{1, n}; \qquad (5.63)$$

$$\det M = D \prod_{k=1}^{n} \sigma_k^2;$$

M^{-1} — матрица, обратная для матрицы M:

$$M^{-1} = \left\| \frac{D_{ik}}{\sigma_i \sigma_k D} \right\|. \qquad (5.64)$$

Поскольку $r_{ik} = r_{ki}$, $i, k = \overline{1, n}$, то матрица M^{-1} является симметричной. Для всякой симметричной матрицы M^{-1} можно найти такую ортогональную матрицу R, которая приводит матрицу M^{-1} к диагональному виду, т. е. матрица $A = R^{-1} M^{-1} R$, полученная трансформированием матрицы M^{-1} матрицей R, будет диагональной. Поэтому линейное преобразование

$$(x - m) = Ry \qquad (5.65)$$

приводит квадратичную форму (5.59) к каноническому виду

$$Q = \sum_{i=1}^{n} y_i^2,$$

а случайный вектор $(X_1, X_2, \ldots, X_n)$ — к нормированному случайному вектору $(Y_1, Y_2, \ldots, Y_n)$ с плотностью распределения

$$f(y_1, y_2, \ldots, y_n) = \frac{1}{\sqrt{(2\pi)^n}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} y_i^2\right). \qquad (5.66)$$

При $n=2$, $X_1=X$ и $X_2=Y$ (5.58) превращается в (3.58), а (5.57) — в (2.91). Линейным преобразованием (5.65) мы пользовались в примерах предыдущих глав, например преобразование (3.55).

§ 5.5 χ^2-распределение

Пусть компоненты вектора $X=(X_1, X_2, \ldots, X_k)$ — независимые случайные величины, распределенные по одному и тому же нормальному закону с параметрами m и σ, т. е. плотность распределения вектора X имеет вид

$$f(x_1, x_2, \ldots, x_k) = \prod_{i=1}^{k} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i-m)^2}{2\sigma^2}\right] =$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^k \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{k}(x_i-m)^2\right]. \qquad (5.67)$$

Закон распределения случайной величины

$$\chi^2 = \frac{1}{\sigma^2}\sum_{i=1}^{k}(X_i-m)^2 = \sum_{i=1}^{k} Y_i^2 \qquad (5.68)$$

носит название χ^2-распределение.

Здесь $Y_i = \frac{1}{\sigma}(X_i - m)$ — нормированная нормально распределенная случайная величина.

Формулу χ^2-распределения найдем из закона распределения случайной величины

$$Z = \frac{\chi}{\sqrt{k}} = \frac{1}{\sqrt{k}}\sqrt{\sum_{i=1}^{k} Y_i^2}. \qquad (5.69)$$

Случайная величина χ неотрицательная, так как она равна абсолютному значению корня квадратного из суммы (5.68), поэтому функция распределения $F_z(z)$ случайной величины Z равна нулю при $z<0$. Для $z \geqslant 0$ имеем

$$F_z(z) = P\left(\frac{1}{\sqrt{k}}\sqrt{\sum_{i=1}^{k} Y_i^2} < z\right) = P\left(\sum_{i=1}^{k} Y_i^2 < z^2 k\right).$$

Случайный вектор $(Y_1, Y_2, \ldots, Y_k)$ распределен по нормированному нормальному закону

$$f_y(y_1, y_2, \ldots, y_k) = \left(\frac{1}{\sqrt{2\pi}}\right)^k \exp\left(-\frac{1}{2}\sum_{i=1}^{k} y_i^2\right),$$

поэтому

$$F_z(z) = \left(\frac{1}{\sqrt{2\pi}}\right)^k \overbrace{\int\int \ldots \int}^{k}_{y_1^2+y_2^2+\ldots+y_k^2<z^2k} \exp\left(-\frac{1}{2}\sum_{i=1}^{k} y_i^2\right) \times$$

$$\times dy_1 dy_2 \ldots dy_k. \qquad (5.70)$$

Таким образом, $F_z(z)$ есть вероятность попадания случайной точки $(Y_1, Y_2, \ldots, Y_k)$ внутрь шара радиусом $z\sqrt{k}$:

$$\sum_{i=1}^{k} y_i^2 < z^2 k.$$

Для вычисления этого интеграла перейдем к сферическим координатам $\rho, \theta_1, \theta_2, \ldots, \theta_{k-1}$, где $\rho = \sqrt{y_1^2 + y_2^2 + \ldots + y_k^2}$ — длина радиус-вектора, а θ_i, $i = \overline{1, k-1}$ — углы, определяющие его направление относительно осей координат. В общем виде связь между старыми и новыми координатами можно представить формулой

$$y_i = \rho\varphi_i(\theta_1, \theta_2, \ldots \theta_{k-1}), \qquad (5.71)$$

где $\varphi_i(\theta_1, \theta_2, \ldots, \theta_{k-1})$, $i = \overline{1, k}$ — функция углов. Якобиан преобразования (5.71) также представим в общем виде

$$\left|\frac{\partial(y_1, y_2, \ldots, y_k)}{\partial(\rho, \theta_1, \theta_2, \ldots, \theta_{k-1})}\right| = \rho^{k-1} D(\theta_1, \theta_2, \ldots, \theta_{k-1}).$$

Функция $D(\theta_1, \theta_2, \ldots, \theta_{k-1})$ зависит только от углов. При этих обозначениях (5.70) принимает вид

$$F_z(z) = \left(\frac{1}{\sqrt{2\pi}}\right)^k \int_{-\pi/2}^{\pi/2}\int_{-\pi/2}^{\pi/2} \ldots \int_{-\pi/2}^{\pi/2}\int_{0}^{z\sqrt{k}} \exp\left(-\frac{\rho^2}{2}\right)\rho^{k-1} \times$$

$$\times D(\theta_1, \theta_2, \ldots, \theta_{k-1})\, d\theta_1 d\theta_2 \ldots d\theta_{k-1} d\rho =$$

$$= C_k \int_{0}^{z\sqrt{k}} \rho^{k-1} \exp\left(-\frac{\rho^2}{2}\right) d\rho,$$

где постоянная

$$C_k = \int_{-\pi/2}^{\pi/2}\int_{-\pi/2}^{\pi/2} \ldots \int_{-\pi/2}^{\pi/2} D(\theta_1, \theta_2, \ldots, \theta_{k-1})\, d\theta_1 d\theta_2 \ldots d\theta_{k-1}$$

зависит только от k. Для вычисления этой постоянной воспользуемся тем, что $F_z(\infty)=1$:

$$C_k \int_0^\infty \rho^{k-1} \exp\left(-\frac{\rho^2}{2}\right) d\rho = 1.$$

Заменим переменную интегрирования $\rho^2=2x$; $\rho=\sqrt{2x}$; $d\rho = dx(2x)^{-1/2}$:

$$1 = C_k \int_0^\infty (2x)^{\frac{k-1}{2}} \exp(-x)(2x)^{-1/2} dx = C_k 2^{\frac{k}{2}-1} \int_0^\infty x^{\frac{k}{2}-1} \exp(-x) dx.$$

Учитывая, что интеграл в этом выражении есть гамма-функция от $k/2$, получим

$$C_k = \frac{1}{2^{\frac{k}{2}-1}\Gamma\left(\frac{k}{2}\right)}.$$

Таким образом, функция распределения случайной величины (5.69)

$$F_z(z) = \frac{1}{2^{\frac{k}{2}-1}\Gamma\left(\frac{k}{2}\right)} \int_0^{z\sqrt{k}} \rho^{k-1} \exp\left(-\frac{\rho^2}{2}\right) d\rho. \qquad (5.72)$$

Дифференцируя это выражение по z, получим плотность распределения Z

$$f_z(z) = \frac{\sqrt{k}}{2^{\frac{k}{2}-1}\Gamma\left(\frac{k}{2}\right)} (z\sqrt{k})^{k-1} \exp\left[-\frac{(z\sqrt{k})^2}{2}\right]. \qquad (5.73)$$

Теперь нетрудно найти плотность вероятности χ^2-распределения, которую обозначим $f(x, k)$.

Для того чтобы получить случайную величину $X=\chi^2$, необходимо случайную величину Z преобразовать по формуле $x=(z\sqrt{k})^2$, причем $\frac{dz}{dx}=\frac{1}{2\sqrt{kx}}$. Воспользовавшись формулой (2.116), найдем из (5.73)

$$f(x, k) = \frac{\sqrt{k}}{2^{\frac{k}{2}-1}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k-1}{2}} \exp\left(-\frac{x}{2}\right) \frac{1}{2\sqrt{kx}}.$$

Таким образом, плотность вероятности χ^2-распределения равна нулю при $x<0$ и

$$f(x, k) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} \exp\left(-\frac{x}{2}\right) \quad \text{при } x \geqslant 0. \qquad (5.74)$$

Единственный параметр χ^2-распределения равен числу слагаемых в (5.68) и называется степенью свободы χ^2-распределения. Кривые χ^2-распределения для нескольких значений k показаны на рис. 5.2. В справочниках с математико-статистическими таблицами приводятся квантили χ^2-распределения.

Функция χ^2-распределения имеет вид

$$F(x, k) = \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} \int_0^x t^{\frac{k}{2}-1} \exp\left(-\frac{t}{2}\right) dt. \qquad (5.75)$$

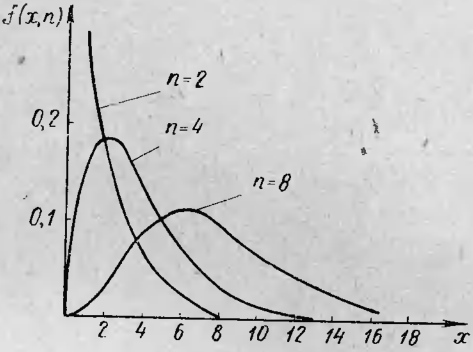

Рис. 5.2

Найдем ее изображение по Лапласу

$$\varphi(z) = \int_0^\infty \exp(-zx)\, dF(x, k) =$$

$$= \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} \int_0^\infty x^{\frac{k}{2}-1} \exp(-zx) \exp\left(-\frac{x}{2}\right) dx =$$

$$= \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} \int_0^\infty x^{\frac{k}{2}-1} \exp\left[-x\left(z + \frac{1}{2}\right)\right] dx.$$

Заменим переменную интегрирования $x\left(z + \frac{1}{2}\right) = y$; $dx \times \left(z + \frac{1}{2}\right) = dy$; $x = \frac{y}{(z + {}^1\!/_2)}$:

$$\varphi(z) = \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right) \left(z + \frac{1}{2}\right)^{\frac{k}{2}}} \int_0^\infty y^{\frac{k}{2}-1} \exp(-y)\, dy.$$

Окончательно

$$\varphi(z) = \frac{1}{2^{\frac{k}{2}}\left(z+\frac{1}{2}\right)^{\frac{k}{2}}} = (2z+1)^{-\frac{k}{2}}. \qquad (5.76)$$

Из формулы (5.76) вытекает, что сумма двух независимых случайных величин, распределенных по закону χ^2 с параметрами k и l, также распределена по закону χ^2 с параметром $k+l$. Действительно, преобразование Лапласа свертки законов распределения этих случайных величин

$$\varphi(z) = \varphi_1(z)\varphi_2(z) =$$

$$= (2z+1)^{-\frac{k}{2}}(2z+1)^{-\frac{l}{2}} = (2z+1)^{-\frac{k+l}{2}}. \qquad (5.77)$$

Впрочем, этот факт и непосредственно следует из формулы (5.68).

Найдем математическое ожидание и дисперсию случайной величины X, подчиненной χ^2-распределению с k-степенями свободы. Для этого воспользуемся формулой (4.58)

$$\frac{d\varphi(z)}{dz} = -k(2z+1)^{-\frac{k+2}{2}};$$

$$\frac{d^2\varphi(z)}{dz^2} = k(k+2)(2z+1)^{-\frac{k+4}{2}}.$$

Отсюда

$$\mathbf{M}X = k; \qquad (5.78)$$

$$\mathbf{M}X^2 = k^2 + 2k;$$

$$\mathbf{D}X = 2k. \qquad (5.79)$$

§ 5.6. Гамма-распределение

Закон распределения χ^2 является частным случаем более общего закона, который называется гамма-распределением. Этот закон зависит от двух неотрицательных параметров ν и λ. Плотность вероятности гамма-распределения определяется формулой

$$f(x, \nu, \lambda) = \begin{cases} 0 & \text{при } x < 0; \\ \dfrac{\lambda^\nu x^{\nu-1}}{\Gamma(\nu)} \exp(-\lambda x) & \text{при } x \geqslant 0, \end{cases} \qquad (5.80)$$

а функция распределения

$$F(x) = \frac{\lambda^\nu}{\Gamma(\nu)} \int_0^x t^{\nu-1} \exp(-\lambda t)\, dt. \qquad (5.81)$$

При $\lambda=1/2$ и $\nu=k/2$ (5.80) превращается в (5.74). Следовательно, воспользовавшись (5.76), сразу же напишем изображение по Лапласу функции (5.80):

$$\varphi(z)=\frac{\lambda^{\nu}}{(z+\lambda)^{\nu}}. \tag{5.82}$$

Отсюда получим для случайной величины X, подчиненной гамма-распределению:

$$\mathbf{M}X=\frac{\nu}{\lambda};\quad \mathbf{D}X=\frac{\nu}{\lambda^2}. \tag{5.83}$$

В качестве частных случаев гамма-распределение содержит также распределение Эрланга и экспоненциальное (или показательное) распределение. Распределением Эрланга называется гамма-распределение, если параметр ν равен натуральному числу. Формула плотности распределения Эрланга для $x>0$

$$f(x,\nu,\lambda)=\frac{\lambda^{\nu}}{(\nu-1)!}x^{\nu-1}\exp(-\lambda x). \tag{5.84}$$

Плотность распределения экспоненциального закона получим из (5.80) при $\nu=1$

$$f(x,\lambda)=\begin{cases}0 & \text{при } x<0,\\ \lambda\exp(-\lambda x) & \text{при } x\geqslant 0.\end{cases} \tag{5.85}$$

Функция распределения

$$F(x)=1-\exp(-\lambda x). \tag{5.86}$$

Из (5.82) и (5.83) для случайной величины, распределенной по экспоненциальному закону,

$$\varphi(z)=\frac{\lambda}{z+\lambda}, \tag{5.87}$$

$$\mathbf{M}X=\frac{1}{\lambda};\quad \mathbf{D}X=\frac{1}{\lambda^2}. \tag{5.88}$$

Можно показать, что экспоненциальному распределению подчиняется интервал между соседними точками, случайно распределенными на вещественной оси t, если выполняются постулаты Пуассона. Действительно, вероятность (5.14) зависит только от длины интервала $[0,\ t]$ и не зависит от выбора точки отсчета. Возьмем в качестве таковой одну из случайных точек. Обозначим символом X интервал между соседними точками. Событие, которое состоит в том, что на интервале $[0,\ t]$ нет ни одной случайной точки, эквивалентно событию, которое состоит в том, что X не меньше t, т. е. $(k=0)=(X\geqslant t)$, поэтому согласно (5.15) $\pi(0;\ \lambda t)=P(X\geqslant t)$. Отсюда найдем по определению функцию распределения случайной величины X:

$$F(t)=P(X<t)=1-P(X\geqslant t)=1-\pi(0;\lambda t),$$

или

$$F(t)=1-\exp(-\lambda t),$$

что совпадает при $x=t$ с (5.86).

Сопоставив (5.86) и (5.82), видим, что сумма ν независимых случайных величин, каждая из которых распределена по экспоненциальному закону, подчиняется распределению Эрланга. Таким образом, сумма нескольких смежных интервалов между точками, случайным образом распределенными на вещественной оси, при выполнении постулатов Пуассона распределена по закону Эрланга, причем ν — число слагаемых, а λ — плотность точек.

Случайные величины, подчиненные гамма-распределению или его частным случаям — распределению Эрланга или экспоненциальному распределению, встречаются в процессах, которые являются предметом изучения теории надежности и теории массового обслуживания.

§ 5.7. Другие часто встречающиеся распределения

Число всевозможных распределений случайных величин очень велико, так как любую неотрицательную функцию $f(x)$ или последовательность неотрицательных чисел $\{p_k\}$, $k=0, 1, 2, \ldots$, удовлетворяющих условиям нормировки (2.31) и (2.23), можно рассматривать как закон распределения некоторой случайной величины. Из всех законов распределения наиболее важными в теоретическом и прикладном отношении являются биномиальный закон, закон Пуассона и нормальный закон. Но при изучении случайных величин встречаются и другие виды распределений.

Логарифмически нормальное распределение

Положительная случайная величина Y имеет логарифмически нормальное распределение, если ее логарифм распределен нормально. На практике встречаются два варианта логарифмического преобразования

$$X_1=\ln Y \quad \text{и} \quad X_2=\lg Y. \tag{5.89}$$

При этом имеет место равенство $X_2=0{,}4343\ X_1$, где 0,4343 — коэффициент перехода от натуральных логарифмов к десятичным.

Плотности распределений X_1 и X_2 имеют вид:

$$f_1(x) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left[-\frac{(x-m_1)^2}{2\sigma_1^2}\right]; \tag{5.90}$$

$$f_2(x) = \frac{1}{\sigma_2 \sqrt{2\pi}} \exp\left[-\frac{(x-m_2)^2}{2\sigma_2^2}\right], \tag{5.91}$$

причем

$$m_2 = 0{,}4343 m_1; \tag{5.92}$$

$$\sigma_2 = 0{,}4343 \sigma_1. \tag{5.93}$$

Из (5.89) и (5.92) следует, что существует такое y_0, что

$$\ln y_0 = m_1 \quad \text{и} \quad \lg y_0 = m_2. \tag{5.94}$$

Используя формулу (2.116), получим из (5.90) и (5.91) формулу плотности распределения случайной величины Y

$$f(y) = \frac{1}{\sigma_1 y \sqrt{2\pi}} \exp\left[-\frac{(\ln y - \ln y_0)^2}{2\sigma_1^2}\right] =$$

$$= \frac{0{,}4343}{\sigma_2 y \sqrt{2\pi}} \exp\left[-\frac{(\lg y - \lg y_0)^2}{2\sigma_2^2}\right]. \tag{5.95}$$

Распределение Вейбулла

Случайная положительная величина распределена по закону Вейбулла, если ее функция распределения выражается формулой

$$F(x) = \begin{cases} 0 & \text{при } x < 0; \\ 1 - \exp\left[-\left(\frac{x}{\sigma}\right)^{\eta}\right] & \text{при } x \geqslant 0, \end{cases} \tag{5.96}$$

где σ и η — положительные постоянные.

Плотность распределения Вейбулла

$$f(x, \eta, \sigma) = \frac{\eta}{\sigma}\left(\frac{x}{\sigma}\right)^{\eta-1} \exp\left[-\left(\frac{x}{\sigma}\right)^{\eta}\right]. \tag{5.97}$$

Нетрудно найти, что

$$\mathbf{M}\, X = \sigma \Gamma\left(\frac{1}{\eta} + 1\right); \tag{5.98}$$

$$\mathbf{D}\, X = \sigma^2 \left\{\Gamma\left(\frac{2}{\eta} + 1\right) - \left[\Gamma\left(\frac{1}{\eta} + 1\right)\right]^2\right\}. \tag{5.99}$$

Распределение Вейбулла используется при изучении разброса срока службы радиоэлектронной аппаратуры.

Распределение Стьюдента

Пусть компоненты вектора (X, Y) — независимые случайные величины, причем X распределена по нормированному нормальному закону, а Y — по закону χ^2 с k степенями свободы.

Закон распределения случайной величины

$$T=\frac{X\sqrt{k}}{\sqrt{Y}} \tag{5.100}$$

называется t-распределением Стьюдента. Необходимость рассмотрения случайной величины T возникает во многих задачах статистики.

Совместная плотность распределения вектора (X, Y) равна произведению плотностей распределения компонент

$$f(x, y)=\frac{1}{2^{\frac{k}{2}}\sqrt{2\pi}\,\Gamma\left(\frac{k}{2}\right)}\,y^{\frac{k}{2}-1}\exp\left[-\frac{1}{2}(x^2+y)\right]. \tag{5.101}$$

Преобразуем (X, Y) в вектор (S, T):

$$S=Y;\quad T=\frac{X\sqrt{k}}{\sqrt{Y}}.$$

Якобиан обратного преобразования равен $\sqrt{\frac{1}{k}S}$, поэтому, применяя к (5.101) формулу (2.123), получим плотность распределения вектора (S, T)

$$f(s, t)=\frac{1}{2\sqrt{k\pi}\,\Gamma\left(\frac{k}{2}\right)}\left(\frac{s}{2}\right)^{\frac{k-1}{2}}\exp\left[-\frac{1}{2}(1+t^2)\,s\right]. \tag{5.102}$$

Интегрируя это выражение по s от $-\infty$ до $+\infty$, получим согласно (2.72) плотность распределения случайной величины T, т. е. плотность распределения закона Стьюдента,

$$f_t(t)=\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}}(1+t^2)^{-\frac{k+1}{2}}. \tag{5.103}$$

Отсюда можно найти, что

$$\mathbf{M}T=0 \quad \text{и} \quad \mathbf{D}T=\frac{k}{k-2}. \tag{5.104}$$

Распределение Релея

Пусть случайный вектор (X, Y) распределен по нормальному закону (2.91). В трехмерном пространстве с координатными осями x, y, z функция (2.91) определяет поверхность колоколообразной формы. Сечение этой поверхности плоскостью $z=\mathrm{const}$ образует линию, вдоль которой плотность распределения сохраняет постоянное значение. Уравнение этой линии получим, приравняв показатель экспоненты некоторой постоянной λ^2:

$$\frac{(x-m_1)^2}{\sigma_1^2}-2r\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2}+\frac{(y-m_2)^2}{\sigma_2^2}=\lambda^2. \quad (5.105)$$

Если учесть, что $\sigma_1>0$, $\sigma_2>0$ и $|r|<1$, то нетрудно убедиться в том, что (5.105) есть уравнение эллипса. Это соответствует высказанному в § 5.4 утверждению о том, что показатель экспоненты многомерного нормального закона — положительно определенная квадратичная форма. Эллипсы (5.105) называются эллипсами равной вероятности. При $\lambda^2=0$ эллипс вырождается в точку с координатами $x=m_1$; $y=m_2$. В этой точке плотность вероятностей максимальна:

$$f(m_1, m_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}. \quad (5.106)$$

По мере увеличения λ^2 размеры эллипса увеличиваются, а плотность вероятности в точках эллипса уменьшается.

Вероятность того, что случайная точка (X, Y), координаты которой распределены по закону (2.91), окажется внутри эллипса равной вероятности при фиксированном λ^2, равна

$$P(\lambda)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\iint\limits_{G(\lambda)}\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_1)^2}{\sigma_1^2}-\right.\right.$$

$$\left.\left.-2r\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2}+\frac{(y-m_2)^2}{\sigma_2^2}\right]\right\}dxdy. \quad (5.107)$$

Здесь $G(\lambda)$ — часть плоскости xy, ограниченная эллипсом (5.105). Подстановкой

$$x=\sigma_1\left[u\sqrt{1-r^2}-vr\right]+m_1;$$

$$y=-\sigma_2 v+m_2$$

с якобианом $\left|\frac{d(x, y)}{d(u, v)}\right|=\sigma_1\sigma_2\sqrt{1-r^2}$ (5.107) приводится к виду:

$$P(\lambda)=\frac{1}{2\pi}\iint\limits_{K(\lambda)}\exp\left[-\frac{1}{2}(u^2+v^2)\right]dudv, \quad (5.108)$$

где $K(\lambda)$ — часть плоскости uv, ограниченная окружностью

$$u^2+v^2=\frac{\lambda^2}{(1-r^2)}=R^2.$$

Перейдем в (5.108) к полярным координатам $\rho=\sqrt{u^2+v^2}$, $\varphi=\mathrm{arctg}\ v/u$ преобразованием $u=\rho\cos\varphi$; $v=$

$=\rho\sin\varphi$ с якобианом $\left|\frac{d(u,v)}{d(\rho,\varphi)}\right|=\begin{Vmatrix}\cos\varphi & -\rho\sin\varphi\\ \sin\varphi & \rho\cos\varphi\end{Vmatrix}=\rho$:

$$P(\lambda)=\frac{1}{2\pi}\int_0^{2\pi}\int_0^R\exp\left(-\frac{1}{2}\rho^2\right)\rho d\rho\, d\varphi=$$

$$=\int_0^R\exp\left(-\frac{1}{2}\rho^2\right)\rho d\rho=1-\exp\left(-\frac{1}{2}R^2\right).$$

Окончательно

$$P(\lambda)=1-\exp\left[-\frac{\lambda^2}{2(1-r^2)}\right]. \qquad (5.109)$$

Как видно из (5.105), при $\sigma_1=\sigma_2=\sigma$ эллипсы равных вероятностей переходят в окружности радиусом $R=\lambda\sigma$. Вероятность того, что точка плоскости, координаты которой являются независимыми нормально распределенными случайными величинами с параметрами m_1, σ_1 и m_2, σ_2 соответственно, находится в круге, ограниченном окружностью радиусом x с центром в точке (m_1, m_2), получим из (5.109) при $r=0$ и $\lambda=\sigma/x$:

$$F(x)=1-\exp\left(-\frac{x^2}{2\sigma^2}\right),\quad x>0. \qquad (5.110)$$

Закон распределения с функцией распределения (5.110) носит название закона распределения Релея. Плотность вероятности закона Релея выражается формулой

$$f(x)=\begin{cases}0 & \text{при } x<0;\\ \frac{x}{\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right) & \text{при } x\geqslant 0.\end{cases} \qquad (5.111)$$

Случайные величины, подчиненные закону Релея, встречаются в задачах статистической радиотехники.

ГЛАВА 6
ЗАКОНЫ БОЛЬШИХ ЧИСЕЛ И ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

§ 6.1. Закон больших чисел

Огромный опыт, накопленный человечеством, учит, что наблюдения за большими массами явлений позволяют установить общие для однотипных явлений особенности, которые в единичных явлениях оказываются замаскированными, скрытыми индивидуальными, частными особенностями. Уже давно было замечено, что факторы, не связанные с существом процесса, а проявляющиеся только в единичных его реализациях, в среднем из большого числа наблюдений взаимно погашаются, взаимно компенсируются. Это эмпирическое наблюдение получило название статистической устойчивости средних или закона больших чисел. Упомянутый в первой главе факт незначительных колебаний относительной частоты события в длинных сериях испытаний, который послужил эмпирическим основанием для введения теоретического понятия вероятности, является частным проявлением этой закономерности. В качестве второго примера проявления той же закономерности можно привести повышение точности и надежности измерений путем хорошо известного приема осреднения результатов многих измерений.

Закон больших чисел имеет фундаментальное и принципиальное значение для теории вероятностей и, в особенности, для ее приложений. Его математическая формулировка требует рассмотрения неограниченной последовательности случайных величин $X_1, X_2, \ldots, X_n, \ldots$ и понятия сходимости по вероятности. Последовательность случайных величин можно связывать либо с последовательностью испытаний таким образом, что исходом испытания с номером k является событие $X_k=x_k$, либо с одним сложным испытанием, исходом которого является неограниченная последовательность чисел $x_1, x_2, \ldots, x_n, \ldots$, такая, что $X_k=x_k$, $k=1, 2, \ldots, n, \ldots$

Определение 1. *Последовательность случайных величин $X_1, X_2, \ldots, X_n, \ldots$ сходится к случайной величине X по вероятности, если для любого $\varepsilon>0$ вероятность неравенства*

$$|X_n - X| < \varepsilon \tag{6.1}$$

при $n \to \infty$ стремится к единице:

$$\lim_{n \to \infty} P(|X_n - X| < \varepsilon) = 1, \tag{6.2}$$

или, другими словами, для любых неотрицательных ε и η при достаточно большом n

$$1 - \eta < P(|X_n - X| < \varepsilon). \tag{6.3}$$

Если сопоставить понятие сходимости по вероятности с понятием обычной сходимости упорядоченной переменной, которое рассматривается в математическом анализе [9], то нетрудно заметить, что при сходимости по вероятности упорядоченной переменной, имеющей предел, является вероятность. Сходимость последовательности случайных величин по вероятности не дает оснований для утверждения о существовании у нее предела. Действительно, если при испытании для некоторого k наступило событие $|x_k - x| < \varepsilon$, то это не означает, что для всех $n > k$ также наступили события $|x_n - x| < \varepsilon$. Вполне возможно, что найдется такое $n > k$, что в данном сложном испытании $|x_k - x| < \varepsilon$, но $|x_n - x| \geqslant \varepsilon$, несмотря на то что неравенство (6.3) выполняется для всех $n > k$, если оно выполняется для k.

Определение 2. *Последовательность случайных величин $X_1, X_2, X_3, \ldots, X_n, \ldots$ подчиняется закону больших чисел, если последовательность случайных величин*

$$Y_n = \frac{1}{n} \sum_{k=1}^{n} (X_k - \mathbf{M} X_k)$$

сходится по вероятности к нулю:

$$\lim_{n \to \infty} P\left(\left| \frac{1}{n} \sum_{k=1}^{n} (X_k - \mathbf{M} X_k) \right| < \varepsilon \right) = 1 \tag{6.4}$$

для любого $\varepsilon > 0$.

Поскольку событие

$$\left| \frac{1}{n} \sum_{k=1}^{n} (X_k - \mathbf{M} X_k) \right| \geqslant \varepsilon$$

противоположно событию

$$\left| \frac{1}{n} \sum_{k=1}^{n} (X_k - \mathbf{M} X_k) \right| < \varepsilon,$$

то условие (6.4) можно заменить равносильным условием

$$\lim_{n \to \infty} P\left(\left| \frac{1}{n} \sum_{k=1}^{n} (X_k - \mathbf{M} X_k) \right| \geqslant \varepsilon \right) = 0. \tag{6.5}$$

Неравенство Чебышева. *Для любой случайной величины X, имеющей конечную дисперсию, при каждом $\varepsilon > 0$ имеет место неравенство*

$$P(|X - \mathbf{M} X| \geqslant \varepsilon) \leqslant \frac{\mathbf{D} X}{\varepsilon^2}. \tag{6.6}$$

Доказательство. Пусть $F(x)$ — функция распределения случайной величины X. Согласно (2.41)

$$P(|X - \mathbf{M} X| \geqslant \varepsilon) = P\left[(X < \mathbf{M} X - \varepsilon) \cup (X \geqslant \mathbf{M} X + \varepsilon)\right] =$$

$$= \int_{-\infty}^{\mathbf{M} X - \varepsilon} dF(x) + \int_{\mathbf{M} X + \varepsilon}^{\infty} dF(x) = \int_{|x - \mathbf{M} X| \geqslant \varepsilon} dF(x).$$

В области интегрирования выполняется неравенство

$$\frac{|x - \mathbf{M} X|}{\varepsilon} \geqslant 1,$$

равносильное неравенству

$$\frac{(x - MX)^2}{\varepsilon^2} \geqslant 1,$$

поэтому

$$\int\limits_{|x - MX| \geqslant \varepsilon} dF(x) \leqslant \frac{1}{\varepsilon^2} \int\limits_{|x - MX| \geqslant \varepsilon} (x - MX)^2 \, dF(x).$$

Это неравенство усиливается, если интегрирование распространяется на все значения x:

$$P(|X - MX| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon^2} \int\limits_{|x - MX| \geqslant \varepsilon} (x - MX)^2 \, dF(x) \leqslant$$

$$\leqslant \frac{1}{\varepsilon^2} \int\limits_{-\infty}^{\infty} (x - MX)^2 \, dF(x) = \frac{1}{\varepsilon^2} DX.$$

Этим завершается доказательство неравенства (6.6).

Если случайная величина X имеет центральный момент порядка $2k$, $k = 1, 2, \ldots$, то, повторяя рассуждения, можно получить неравенство более общего вида

$$P(|X - MX| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon^2} M(X - MX)^{2k}. \qquad (6.7)$$

Неравенство Чебышева хорошо иллюстрирует роль дисперсии как меры рассеяния случайной величины. Представим (6.6) несколько в ином виде, положив $\varepsilon = t\sigma$, где $\sigma = \sqrt{DX}$:

$$P(|X - MX| \geqslant t\sigma) = Q(t) \leqslant \frac{1}{t^2}. \qquad (6.8)$$

Отсюда видно, что убывание вероятности $Q(t)$ при возрастании t при любом законе распределения случайной величины X происходит не медленнее, чем $1/t^2$. Для практических расчетов эта оценка оказывается слишком грубой, но она позволяет получить ряд важных теоретических заключений.

Теорема 1. *Для того чтобы для последовательности $X_1, X_2, X_3, \ldots, X_n, \ldots$ (как угодно зависимых) случайных величин при любом положительном ε выполнялось соотношение*

$$\lim_{n \to \infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} MX_k\right| < \varepsilon\right) = 1, \qquad (6.9)$$

необходимо и достаточно при $n \to \infty$

$$M\frac{\left[\sum\limits_{k=1}^{n}(X_k - MX_k)\right]^2}{n^2 + \left[\sum\limits_{k=1}^{n}(X_k - MX_k)\right]^2} \to 0. \qquad (6.10)$$

Доказательство. Пусть выполняется (6.10). Обозначим символом $\Phi_n(x)$ функцию распределения случайной величины

$$Y_n = \frac{1}{n}\sum_{k=1}^{n}(X_k - \mathrm{M}X_k). \qquad (6.11)$$

При таком обозначении

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k - \frac{1}{n}\sum_{k=1}^{n}\mathrm{M}X_k\right| \geqslant \varepsilon\right) = P(|Y_n| \geqslant \varepsilon) =$$

$$= \int_{|y|\geqslant\varepsilon} d\Phi_n(y); \qquad (6.12)$$

$$\mathrm{M}\,\frac{Y_n^2}{1+Y_n^2} = \mathrm{M}\,\frac{\left[\sum_{k=1}^{n}(X_k - \mathrm{M}X_k)\right]^2}{n^2 + \left[\sum_{k=1}^{n}(X_k - \mathrm{M}X_k)\right]^2}.$$

В области интегрирования в (6.12) для любого y выполняется неравенство $\varepsilon^2 \leqslant y^2$, равносильное неравенству $|y| \geqslant \varepsilon$, поскольку $\varepsilon > 0$. Отсюда

$$\frac{1}{\varepsilon^2} \geqslant \frac{1}{y^2}; \quad 1 + \frac{1}{\varepsilon^2} \geqslant 1 + \frac{1}{y^2};$$

$$\frac{1+\dfrac{1}{\varepsilon^2}}{1+\dfrac{1}{y^2}} = \frac{1+\varepsilon^2}{1+y^2}\,\frac{y^2}{\varepsilon^2} \geqslant 1,$$

поэтому

$$\int_{|y|\geqslant\varepsilon} d\Phi_n(y) \leqslant \frac{1+\varepsilon^2}{\varepsilon^2}\int_{|y|\geqslant\varepsilon}\frac{y^2}{1+y^2}\,d\Phi_n(y).$$

Неравенство усилится, если интегрирование распространить на все значения y. Согласно (3.2)

$$\int_{-\infty}^{\infty}\frac{y^2}{1+y^2}\,d\Phi_n(y) = \mathrm{M}\,\frac{Y_n^2}{1+Y_n^2},$$

поэтому

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k - \frac{1}{n}\sum_{k=1}^{n}\mathrm{M}X_k\right| \geqslant \varepsilon\right) \leqslant \frac{1+\varepsilon^2}{\varepsilon^2}\int_{-\infty}^{\infty}\frac{y^2}{1+y^2}\,d\Phi_n(y) =$$

$$= \frac{1+\varepsilon^2}{\varepsilon^2}\,\mathrm{M}\,\frac{Y_n^2}{1+Y_n^2}. \qquad (6.13)$$

Поскольку по предположению выполняется (6.10), то при достаточно больших n левая часть неравенства (6.13) будет сколь угодно малой при любом $\varepsilon > 0$, что доказывает достаточность условия (6.10) для выполнения (6.9).

Теперь покажем, что условие (6.10) необходимо для (6.9). Нетрудно видеть, что

$$P(|Y_n| \geqslant \varepsilon) = \int\limits_{|y| \geqslant \varepsilon} d\Phi_n(y) \geqslant \int\limits_{|y| \geqslant \varepsilon} \frac{y^2}{1+y^2} d\Phi_n(y) =$$

$$= \int\limits_{-\infty}^{\infty} \frac{y^2}{1+y^2} d\Phi_n(y) - \int\limits_{|y| < \varepsilon} \frac{y^2}{1+y^2} d\Phi_n(y).$$

Так как

$$\int\limits_{|y| < \varepsilon} \frac{y^2}{1+y^2} d\Phi_n(y) \leqslant \frac{\varepsilon^2}{1+\varepsilon^2} \int\limits_{-\varepsilon}^{\varepsilon} d\Phi_n(y) < \varepsilon^2,$$

то

$$P(|Y_n| \geqslant \varepsilon) \leqslant \mathbf{M} \frac{Y_n^2}{1+Y_n^2} - \varepsilon^2$$

или

$$0 \leqslant \mathbf{M} \frac{Y_n^2}{1+Y_n^2} \leqslant P(|Y_n| \geqslant \varepsilon) + \varepsilon^2. \tag{6.14}$$

Поскольку выполняется (6.9), то, выбирая сначала ε сколь угодно малым, а затем n достаточно большим, можно правую часть (6.14) сделать сколь угодно малой, что означает выполнение (6.10).

Приведем несколько теорем, которые для фактической проверки применимости закона больших чисел могут оказаться более удобными, чем предыдущая теорема. Исторически доказательства их были найдены раньше доказательства этой теоремы, но все они могут быть получены как ее следствия.

Теорема Маркова. *Если последовательность случайных величин* $X_1, X_2, \ldots, X_n, \ldots$ такова, что при $n \to \infty$

$$\frac{1}{n^2} \mathbf{D} \left[\sum_{k=1}^{n} X_k \right] \to 0, \tag{6.15}$$

то, каково бы ни было положительное постоянное ε,

$$\lim_{n \to \infty} P\left(\left| \frac{1}{n} \sum_{k=1}^{n} X_k - \frac{1}{n} \sum_{k=1}^{n} \mathbf{M} X_k \right| < \varepsilon \right) = 1. \tag{6.16}$$

Доказательство. Используем обозначение (6.11). При любом n и любых X выполняется неравенство

$$\frac{Y_n^2}{1+Y_n^2} \leqslant Y_n^2 = \frac{1}{n^2} \left[\sum_{k=1}^{n} (X_k - \mathbf{M} X_k) \right]^2$$

или

$$\mathbf{M}\frac{Y_n^2}{1+Y_n^2} \leqslant \frac{1}{n^2}\mathbf{M}\left[\sum_{k=1}^{n} X_k - \sum_{k=1}^{n}\mathbf{M}X_k\right]^2 =$$

$$= \frac{1}{n^2}\mathbf{D}\left[\sum_{k=1}^{n} X_k\right].$$

Отсюда видно, что при выполнении условия теоремы (6.15) выполняется (6.10), а следовательно, и (6.9), т. е. (6.16).

Теорема Чебышева. *Если $X_1, X_2, \ldots, X_n, \ldots$ — последовательность попарно независимых случайных величин, имеющих конечные дисперсии, ограниченные одной и той же постоянной C: $\mathbf{D}X_1 \leqslant C$, $\mathbf{D}X_2 \leqslant C, \ldots, \mathbf{D}X_n \leqslant C, \ldots$, то, каково бы ни было постоянное $\varepsilon > 0$,*

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n}\mathbf{M}X_k\right| < \varepsilon\right) = 1.$$

Доказательство. Если условие теоремы выполняется, то

$$\frac{1}{n^2}\mathbf{D}\left[\sum_{k=1}^{n} X_k\right] = \frac{1}{n^2}\sum_{k=1}^{n}\mathbf{D}X_k$$

и, следовательно, выполняется условие Маркова (6.15).

Теорема Бернулли. *Пусть Y — число наступлений события A в n независимых испытаниях и p есть вероятность наступления события A в каждом из испытаний. Тогда, каково бы ни было $\varepsilon > 0$,*

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}Y - p\right| < \varepsilon\right) = 1. \qquad (6.17)$$

Доказательство. Обозначим через X_k, $k=\overline{1,n}$ число наступлений события A в испытании с номером k. X_k принимает значение единицы с вероятностью p и нуля с вероятностью $q=1-p$. Очевидно,

$$\mathbf{M}X_k = 1p + 0q = p; \quad Y = \sum_{k=1}^{n} X_k;$$

$$\mathbf{M}Y = \mathbf{M}\sum_{k=1}^{n} X_k = \sum_{k=1}^{n}\mathbf{M}X_k = np;$$

$$\mathbf{D}X_k = \mathbf{M}\left[X_k - \mathbf{M}X_k\right]^2 = (1-p)^2 p + (0-p)^2 q = pq,$$

поэтому (6.17) можно представить в виде

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} \mathbf{M}X_k\right| < \varepsilon\right) = 1. \qquad (6.18)$$

Поскольку $\mathbf{D}X_k = pq \leqslant \frac{1}{4}$ для всех k, то выполняется условие теоремы Чебышева, что влечет (6.18), а следовательно, и (6.17).

Теорема Пуассона. *Если в последовательности независимых испытаний вероятность появления события A в k-м испытании равна p_k, то*

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n} Y - \frac{p_1 + p_2 + \ldots + p_n}{n}\right| < \varepsilon\right) = 1,$$

где Y — число появлений события A в первых n испытаниях.

Доказательство. Введем, как и в предыдущей теореме, случайные величины X_k, равные числу появлений события A в k-м испытании, и, заметив, что

$$\mathbf{M}X_k = p_k; \quad \mathbf{D}X_k = p_k q_k \leqslant \frac{1}{4},$$

мы вновь приходим к теореме Чебышева.

Теорема 2. *Если последовательность попарно независимых случайных величин $X_1, X_2, \ldots, X_n, \ldots$ такова, что $\mathbf{M}X_1 = \mathbf{M}X_2 = \ldots = \mathbf{M}X_n = \ldots = a$ и $\mathbf{D}X_1 \leqslant C$; $\mathbf{D}X_2 \leqslant C, \ldots, \mathbf{D}X_n \leqslant C, \ldots$, то при любом $\varepsilon > 0$*

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - a\right| < \varepsilon\right) = 1. \qquad (6.19)$$

Этот частный случай теоремы Чебышева дает основание правилу среднего арифметического. Предположим, что измеряют некоторую физическую величину a. Повторив измерения n раз в одних и тех же условиях, наблюдатель получит не вполне совпадающие результаты $x_1, x_2, \ldots, x_n$. В качестве приближенного значения a принято брать среднее арифметическое

$$\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k \approx a.$$

Если измерения не содержат систематической ошибки, то $\mathbf{M}X_1 = \mathbf{M}X_2 = \ldots = \mathbf{M}X_n = a$ и согласно закону больших чисел при достаточно больших n с вероятностью, сколь угодно близкой к единице, мы можем получить значение $\bar{x}$, сколь угодно близкое к искомой величине a.

§ 6.2. Усиленный закон больших чисел

Как указывалось в предыдущем параграфе, сходимость по вероятности последовательности случайных величин X_1, X_2, ..., X_n, ... к случайной величине X еще не означает, что X есть предел этой последовательности. Поясним это на примере испытаний по схеме Бернулли. Пусть подбрасывается правильная монета, а Y_n — число появлений герба при n бросаниях или буквы Г в последовательности из n букв Г (герб) и Ц (цифра), которая является элементарным событием композиции n испытаний. Если n достаточно велико, то согласно теореме Бернулли с вероятностью, близкой к единице, мы получим последовательность, в которой Y_n близко к $n/2$, например $Y_n=0{,}45\,n$, т. е. частота появления Г равна 0,45. Является ли такой исход n испытаний гарантией того, что при последующих n бросаниях монеты в последовательности из $2n$ букв частота появления Г приблизится к 1/2 или хотя бы сохранится равной 0,45? Конечно, нет. Вполне возможно, что при последующих n бросаниях каждый раз будет выпадать герб, т. е. окажется, что

$$\frac{1}{2n}Y_n=\frac{1}{2n}(0{,}45n+n)=0{,}725.$$

Вероятность такого события при большом n мала, но не равна нулю. Однако в 1909 г. Э. Борель доказал, что вероятность исходов с подобным изменением частоты с ростом n стремится к нулю, т. е., увеличивая число испытаний, мы с вероятностью, близкой к единице, будем наблюдать сходимость в обычном смысле частоты $\frac{1}{n}Y_n$ к 1/2.

Этот факт, более существенный, чем сходимость частоты по вероятности, получил название усиленного закона больших чисел. Точная формулировка усиленного закона больших чисел и доказательство теорем об условиях его выполнимости требует введения еще одного понятия сходимости для последовательности случайных величин — сходимости с вероятностью единица.

Напомним (§ 2.1), что случайная величина есть функция $\varphi(e)$, заданная на пространстве элементарных событий U. Пусть на одном и том же пространстве U определена последовательность случайных величин

$$X_1=\varphi_1(e),\ X_2=\varphi_2(e),\ .\ .\ .\ ,\ X_n=\varphi_n(e),\ .\ .\ . \qquad (6.20)$$

На каждом элементе $e\in U$ последовательность (6.20) превращается в упорядоченную числовую последовательность. Бесконечная упорядоченная последовательность чисел либо сходится к некоторому пределу, либо не имеет предела. Таким образом, U заданием последовательности (6.20) разбивается на два непересекающихся множества (события). На элементах одного множества (6.20) превращается в сходящиеся числовые последовательности, а на элементах второго — в расходящиеся.

Для пояснения сказанного представим себе бесконечно длинную последовательность бросаний монеты. Пространством элементарных событий такого испытания является бесконечное множество всех бесконечно длинных последовательностей из букв Г и Ц. Определим на этом пространстве последовательность случайных величин следующим образом:

$$\varphi_n(e) = \frac{1}{n} Y_n, \quad n = 1, 2, \quad . \ . \ , \tag{6.21}$$

где n — длина начального участка последовательности e, а Y_n — число букв Г на этом начальном участке. На элементе e, начальный участок которого ГЦЦГГЦГГГЦ..., последовательность случайных величин (6.21) превращается в числовую последовательность: 1,00; 0,500; 0,333; 0,500; 0,600; 0,500; 0,574; 0,624; 0,667; 0,600; ...

Пусть A — множество всех элементарных событий $e \in U$, для которых последовательность (6.20) сходится, а $\varphi(e)$ — предел этой последовательности в точке e. Обозначим через A_{nk}^r множество всех тех e, для которых выполняется неравенство

$$|\varphi_{n+k}(e) - \varphi(e)| < \frac{1}{r}. \tag{6.22}$$

Если в точке e последовательность функций $\varphi_n(e)$ сходится, то для любого r существует такое n, начиная с которого для всех k выполняются неравенства (6.22). Поэтому множество A представляется в виде

$$A = \bigcap_{r=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=1}^{\infty} A_{nk}^r. \tag{6.23}$$

Определим случайную величину X следующим образом:

$$X = \begin{cases} \varphi(e), & \text{если } e \in A, \\ 0, & \text{если } e \in \overline{A}. \end{cases}$$

Определение 1. *Если вероятность случайного события A равна единице, то последовательность случайных величин $X_n = \varphi_n(e)$, $n = 1, 2, \ldots$ сходится к случайной величине X с вероятностью единица (или почти наверное).*

Эта сходимость записывается в виде

$$P(X_n \longrightarrow X) = 1 \qquad \text{при } n \to \infty \tag{6.24}$$

или

$$P(X_n \not\longrightarrow X) = 0 \qquad \text{при } n \to \infty. \tag{6.25}$$

Определение 2. *Последовательность случайных величин $X_1, X_2, \ldots, X_n, \ldots$ подчиняется усиленному закону больших чисел, если последовательность случайных величин*

$$Y_n = \frac{1}{n} \sum_{k=1}^{n} (X_k - \mathbf{M} X_k), \quad n = 1, 2, \ldots \tag{6.26}$$

сходится к нулю с вероятностью единица:

$$P\left[\left(\frac{1}{n} \sum_{k=1}^{n} X_k - \frac{1}{n} \sum_{k=1}^{n} \mathbf{M} X_k\right) \to 0\right] = 1. \tag{6.27}$$

Теорема. *Если при любом целом положительном r*

$$\sum_{n=1}^{\infty} P\left[|X_n - X| \geqslant \frac{1}{r}\right] < +\infty, \tag{6.28}$$

то последовательность случайных величин $X_1, X_2, \ldots, X_n, \ldots$ сходится к X с вероятностью единица.

Доказательство. Докажем, что (6.28) влечет (6.25), для чего рассмотрим событие, противоположное событию A (6.23):

$$S = \overline{A} = \overline{\bigcap_{r=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=1}^{\infty}} A_{nk}^r. \tag{6.29}$$

Используя (1.6), преобразуем (6.29) к следующему виду:

$$S = \bigcup_{r=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} \overline{A_{nk}^r}. \tag{6.30}$$

Здесь $\overline{A_{nk}^r}$ — событие, противоположное событию (6.22), т. е. $\overline{A_{nk}^r}$ — событие, состоящее в том, что выполняется неравенство

$$|X_{n+k} - X| \geqslant \frac{1}{r}. \tag{6.31}$$

Пусть $S_n^r = \bigcup_{k=1}^{\infty} \overline{A_{nk}^r}$. На основании (1.20)

$$P(S_n^r) \leqslant \sum_{k=1}^{\infty} P(\overline{A_{nk}^r}) = \sum_{k=1}^{\infty} P\left[|X_{n+k} - X| \geqslant \frac{1}{r}\right] =$$

$$= \sum_{l=n+1}^{\infty} P\left[|X_l - X| \geqslant \frac{1}{r}\right]. \tag{6.32}$$

Отсюда и из условия (6.28) следует, что $P(S_n^r)$ меньше любого сколь угодно малого числа $\varepsilon > 0$, если n достаточно большое число, т. е.

$$\lim_{n \to \infty} P(S_n^r) = 0. \tag{6.33}$$

Пусть $S^r = \bigcap\limits_{n=1}^{\infty} S_n^r$. Из (1.7), (1.15) и (6.33) следует, что $P(S^r)=0$. Но

$$S = \bigcup_{r=1}^{\infty} S^r \quad \text{и} \quad P(S) \leqslant \sum_{r=1}^{\infty} P(S^r),$$

поэтому $P(S)=0$, что и требовалось доказать.

Теорема Бореля. *Пусть Y — число наступлений события A при n независимых испытаниях, в каждом из которых событие A может появиться с вероятностью p. Тогда при $n \to \infty$*

$$P\left(\frac{1}{n} Y \to p\right) = 1. \tag{6.34}$$

Доказательство сводится предыдущей теоремой к доказательству сходимости для любого натурального r ряда

$$\sum_{n=1}^{\infty} P\left(\left|\frac{1}{n} Y - p\right| \geqslant \frac{1}{r}\right). \tag{6.35}$$

Для $k=2$ из (6.7) получим

$$P\left(\left|\frac{1}{n} Y - p\right| \geqslant \frac{1}{r}\right) \leqslant r^4 \mathbf{M}\left(\frac{1}{n} Y - p\right)^4. \tag{6.36}$$

Пусть X_k — число появлений события A в k-м испытании, т. е. $Y = \sum\limits_{k=1}^{n} X_k$. Тогда

$$\frac{1}{n} Y - p = \left[\frac{1}{n} \sum_{k=1}^{n} X_k\right] - p = \frac{1}{n} \sum_{k=1}^{n} (X_k - p);$$

$$\mathbf{M}\left(\frac{1}{n} Y - p\right)^4 = \frac{1}{n^4} \mathbf{M}\left[\sum_{k=1}^{n} (X_k - p)\right]^4 =$$

$$= \frac{1}{n^4} \mathbf{M}\left[\sum_{i=1}^{n} (X_i - p) \sum_{j=1}^{n} (X_j - p) \sum_{k=1}^{n} (X_k - p) \sum_{l=1}^{n} (X_l - p)\right] =$$

$$= \frac{1}{n^4} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{l=1}^{n} \mathbf{M}(X_i - p)(X_j - p)(X_k - p)(X_l - p). \tag{6.37}$$

Заметим, что случайные величины X_i и X_j при $i \neq j$ независимы, поэтому для них имеет место равенство (3.22), и далее, поскольку $\mathbf{M}(X_i - p) = 0$, $i = \overline{1, n}$, то в сумме (6.37) все слагаемые, содержащие хотя бы один множитель $(X_i - p)$ в первой степени, равны нулю. Таким образом, в этой сумме отличны от нуля только слагаемые вида

$$\mathbf{M}(X_i - p)^4 \quad \text{и} \quad \mathbf{M}(X_i - p)^2 (X_j - p)^2.$$

Количество слагаемых первого вида равно n, поскольку i пробегает значения от единицы до n. Для построения слагаемого второго вида нужно произвести три выбора: выбрать значение для i — всего n возможностей, выбрать второй индекс из трех возможных — j, k, l, выбрать значение для второго индекса, отличное от значения первого индекса, — $(n-1)$ возможностей. Применяя правило произведения, получим число слагаемых второго вида, равное $3n(n-1)$. Найдем значения этих слагаемых, учитывая, что $P(X_i=1)=p$, $P(X_i=0)=q$, а ряд распределения вектора (X_i, X_j)

x_j	x_i	
	0	1
0	q^2	qp
1	qp	p^2

По формуле (3.30) получим

$$\mathrm{M}(X_i-p)^4=(1-p)^4p+(0-p)^4q=pq(q^3+p^3).$$

Применение формулы (3.36) для слагаемого второго вида дает

$$\mathrm{M}(X_i-p)^2(X_j-p)^2=\sum_{i=1}^{2}\sum_{j=1}^{2}(x_i-p)^2(x_j-p)^2\times$$

$$\times P(X_i=x_i;\ X_j=x_j)=(1-p)^2(1-p)^2p^2+(1-p)^2p^2pq+$$

$$+p^2(1-p)^2pq+p^2p^2q^2=p^2q^2.$$

Таким образом,

$$\mathrm{M}\left(\frac{1}{n}Y-p\right)^4=\frac{1}{n^4}[npq(q^3+p^3)+3n(n-1)p^2q^2]=$$

$$=\frac{pq}{n^2}\left[\frac{q^3+p^3}{n}+3pq-\frac{3pq}{n}\right]<\frac{1}{4n^2},$$

т. е. при любом r члены ряда (6.35) убывают с возрастанием n не медленнее, чем $1/n^2$, и, следовательно, ряд сходится. Теорема доказана.

Доказательство теоремы Бореля вызвало ряд исследований, направленных на разыскание более широких условий, при выполнении которых имеет место усиленный закон больших чисел. Приведем без доказательства результаты, полученные Колмогоровым [4].

Т е о р е м а К о л м о г о р о в а 1. *Если последовательность взаимно независимых случайных величин* $X_1, X_2, \ldots, X_n, \ldots$ *удовлетворяет условию*

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \mathbf{D} X_n < +\infty, \tag{6.38}$$

то она подчиняется усиленному закону больших чисел.

Из этой теоремы вытекает следствие: если $X_1, X_2, \ldots$ $\ldots, X_n, \ldots$ — последовательность попарно независимых случайных величин, имеющих конечные дисперсии, ограниченные одной и той же постоянной C, то эта последовательность подчиняется усиленному закону больших чисел.

Т е о р е м а К о л м о г о р о в а 2. *Существование математического ожидания является необходимым и достаточным условием для применимости усиленного закона больших чисел к последовательности одинаково распределенных и взаимно независимых случайных величин.*

§ 6.3. Предельные теоремы о функциях распределения

Теоремы о законе больших чисел формулируют условия, при которых последовательность определенного вида случайных величин, а именно средних арифметических многих случайных величин сходится по вероятности или с вероятностью единица к некоторой постоянной величине. Вторая группа очень важных для теории и приложений теорем формулирует условия, при которых последовательность функций распределения последовательности случайных величин сходится к некоторой предельной функции распределения. Случайная величина, которая рассматривается этой группой теорем, есть сумма множества независимых случайных величин, а последовательность образуется упорядочиванием сумм по числу слагаемых.

Необходимость изучать законы распределения сумм большого числа случайных величин вытекает из общего свойства всех явлений природы — массовости причин, обусловливающих каждое явление. Численное значение случайной величины очень часто формируется под действием многочисленных независимых или слабо зависимых случайных факторов. Так, например, при измерениях любого вида на результат неизбежно действует множество факторов, порождающих ошибки. Это ошибки, вызванные состоянием измерительного прибора, которое может незаметно для наблюдателя меняться под влиянием атмосферных, механических или электромагнитных причин. Сюда же относятся и субъективные ошибки наблюдателя, вызванные особенностями его органов чувств и способные также бесконтрольно изменяться в зависимости от психического или физиологического состояния наблюдателя или состояния среды. На измерении сказываются все эти факторы. Наблюдаемая в конечном счете ошибка будет случайной величиной, являющейся суммой большого числа случайных величин. Причем последние при правильно проводимых измерениях малы и приблизительно одного и того же порядка, поскольку всегда и конструктор измерительного прибора и наблюдатель стремятся предусмотреть специальные меры для выявления и компенсации сильно действующих факторов. Величина каждой частной ошибки остается неизвестной, неизвестны и законы распределения их, но суммарное влияние на результат измерений, несмотря на малость каждого слагаемого, может быть заметно и поэтому требует изучения.

Подобные примеры можно привести из самых разнообразных областей естествознания, техники и экономики. Таким образом, возникает задача изучения закономерностей, свойственных суммам большого числа независимых случайных величин. Но в аналогичных случаях в математике применяется оправдавший себя во многих приложениях прием перехода от конечной постановки задачи к предельной. Вместо того чтобы изучать сумму очень большого, но конечного числа слагаемых, рассматривают последовательности сумм с возрастающим числом слагаемых, считая, что предельная функция распределения последовательности функций распределения сумм дает решение задачи. При таком подходе прежде всего возникает вопрос об условиях, при которых последовательность функций распределения имеет предел, и вопрос о виде предельной функции распределения. Условия, при которых предельным оказывается нормальный закон распределения, осуществляются настолько часто, что долгое время этот закон рассматривался как единственный предельный закон распределения, а классической проблемой теории вероятностей являлся поиск наиболее общих условий так называемой центральной предельной теоремы. Эта теорема утверждает, что при определенных условиях

$$\lim_{n\to\infty} P\left[\frac{\sum_{k=1}^{n}(X_k-a_k)}{\sqrt{\sum_{k=1}^{n}\sigma_k^2}}<y\right]=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}\exp\left(-\frac{z^2}{2}\right)dz, \qquad (6.39)$$

где $X_1, X_2, \ldots, X_n, \ldots$ — последовательность независимых случайных величин с математическими ожиданиями $a_1, a_2, \ldots, a_n, \ldots$ и дисперсиями $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2, \ldots$

12

Сопоставив (6.39) и (4.31), нетрудно заметить, что теорема Муавра—Лапласа является частным случаем центральной предельной теоремы.

Достаточные условия для выполнения равенства (6.39), как показали П. А. Чебышев, А. А. Марков и А. И. Ляпунов, заключаются в том, чтобы слагаемые

$$\frac{X_k - \alpha_k}{\sqrt{\sum_{k=1}^{n} \sigma_k^2}} \tag{6.40}$$

были «равномерно малыми», т. е. каждое отдельное слагаемое должно оказывать незначительное влияние на сумму. А. Н. Колмогоров и Б. В. Гнеденко установили, что класс предельных законов не исчерпывается нормальным законом, и нашли условия, которым должны удовлетворять слагаемые, чтобы функции распределения их сумм сходились к тому или иному предельному закону. Таким образом, центральная предельная теорема сейчас является следствием теорем более общего характера.

Ограничимся доказательством достаточности для выполнения (6.39) требований, получивших наименование условий Линдеберга [4]. Для более обстоятельного изучения предельных теорем теории вероятностей следует обратиться к монографии [3].

Пусть случайные величины последовательности X_1, X_2, ..., X_n, ... взаимно независимы, распределены по законам

$$F_1(x),\ F_2(x),\ \ldots,\ F_n(x),\ \ldots,$$

имеют конечные математические ожидания,

$$\alpha_1 = \mathbf{M}X_1;\ \alpha_2 = \mathbf{M}X_2;\ \ldots;\ \alpha_n = \mathbf{M}X_n;\ \ldots$$

и конечные дисперсии

$$\sigma_1^2 = \mathbf{D}X_1;\ \sigma_2^2 = \mathbf{D}X_2;\ \ldots;\ \sigma_n^2 = \mathbf{D}X_n;\ \ldots$$

Введем следующие обозначения:

$$B_n^2 = \mathbf{D}\left[\sum_{k=1}^{n} X_n\right] = \sum_{k=1}^{n} \sigma_k^2; \tag{6.41}$$

$$X_{nk} = \frac{X_k - \alpha_k}{B_n};$$

$$F_{nk}(x) = P\,(X_{nk} < x). \tag{6.42}$$

Нетрудно найти следующее:

1. $F_{nk}(x) = F_k\,(B_n x + \alpha_k)$;

2. $\mathbf{M}X_{nk} = 0$; $\qquad$ (6.43)

3. $\mathrm{D}X_{nk} = \frac{1}{B_n^2}\sigma_k^2$;

4. $\sum\limits_{k=1}^{n} \mathrm{D}X_{nk} = 1.$ (6.44)

Теорема. *Если последовательность взаимно независимых случайных величин $X_1, X_2, \ldots, X_n, \ldots$ при любом постоянном $\tau > 0$ удовлетворяет условию Линдеберга*

$$\lim_{n\to\infty} \sum_{k=1}^{n} \int\limits_{|x-a_k|>\tau B_n} \frac{(x-a_k)^2}{B_n^2}\,dF_k(x) = 0, \tag{6.45}$$

то при $n\to\infty$ равномерно относительно x

$$P\left(\sum_{k=1}^{n} \frac{X_k - a_k}{B_n} < x\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{z^2}{2}\right) dz. \tag{6.46}$$

Прежде чем перейти к доказательству теоремы, заметим следующее. Слева от стрелки в выражении (6.46) стоит функция распределения нормированной случайной величины

$$Y_n = \frac{\sum\limits_{k=1}^{n} (X_k - a_k)}{B_n} = \sum_{k=1}^{n} X_{nk}, \tag{6.47}$$

а справа — функция распределения нормированного нормального закона. Поскольку характеристическая функция нормированного нормального закона распределения имеет вид (5.36), то для доказательства настоящей теоремы достаточно в силу обратной предельной теоремы (см. § 4.2) доказать, что последовательность характеристических функций последовательности случайных величин Y_n сходится при $n\to\infty$ к (5.36), т. е.

$$\lim_{n\to\infty} \varphi_n(t) = \exp\left(-\frac{t^2}{2}\right), \tag{6.48}$$

где $\varphi_n(t)$ — характеристическая функция Y_n. Но (6.48) равносильно равенству

$$\lim_{n\to\infty} \ln \varphi_n(t) = -\frac{t^2}{2}, \tag{6.49}$$

где ln обозначает главное значение логарифма комплексного числа. Далее, обозначив характеристическую функцию случайной величины X_{nk} (6.42) через $\varphi_{nk}(t)$, $n=1, 2, \ldots, k=\overline{1, n}$

и учитывая независимость этих случайных величин, получим в соответствии с (4.6) и (6.47)

$$\varphi_n(t) = \prod_{k=1}^{n} \varphi_{nk}(t),$$

или

$$\ln \varphi_n(t) = \sum_{k=1}^{n} \ln \varphi_{nk}(t). \qquad (6.50)$$

Таким образом, доказательство теоремы сводится к доказательству предельного равенства

$$\lim_{n\to\infty} \sum_{k=1}^{n} \ln \varphi_{nk}(t) = -\frac{t^2}{2}. \qquad (6.51)$$

Заметим также, что при обозначениях (6.41) условие Линдеберга принимает следующую простую форму:

$$\lim_{n\to\infty} \sum_{k=1}^{n} \int\limits_{|x|>\tau} x^2\, dF_{nk}(x) = 0. \qquad (6.52)$$

Доказательство. Прежде всего докажем, что равномерно относительно k, $k=\overline{1,n}$ в каждом конечном интервале $|t| \leqslant T$

$$\lim_{n\to\infty} \varphi_{nk}(t) = 1. \qquad (6.53)$$

Так как

1. $\varphi_{nk}(t) = \int\limits_{-\infty}^{\infty} \exp itx\, dF_{nk}(x);$

2. $\int\limits_{-\infty}^{\infty} dF_{nk}(x) = 1;$

3. $it\mathrm{M}X_{nk} = \int\limits_{-\infty}^{\infty} ixt\, dF_{nk}(x) = 0,$

то

$$\varphi_{nk}(t) - 1 = \int\limits_{-\infty}^{\infty} (\exp itx - 1 - itx)\, dF_{nk}(x). \qquad (6.54)$$

При любом вещественном y

$$|\exp iy - 1| = \left|\int\limits_0^y \exp iy\, dy\right| \leqslant \int\limits_0^y dy,$$

т. е.

$$|\exp iy - 1| \leqslant |y|. \qquad (6.55)$$

На основании этого неравенства получим следующие два неравенства:

$$|\exp iy - 1 - iy| = \left|\int_0^y (\exp iy - 1)\,dy\right| \leqslant \int_0^y |y|\,dy,$$

т. е.

$$|\exp iy - 1 - iy| \leqslant \frac{y^2}{2}, \tag{6.56}$$

и аналогично

$$\left|\exp iy - 1 - iy + \frac{y^2}{2}\right| \leqslant \frac{|y|^3}{6}. \tag{6.57}$$

Используя неравенство (6.56), из выражения (6.54) находим

$$|\varphi_{nk}(t) - 1| \leqslant \int_{-\infty}^{\infty} |\exp ixt - 1 - itx|\,dF_{nk}(x) \leqslant$$

$$\leqslant \frac{t^2}{2}\int_{-\infty}^{\infty} x^2\,dF_{nk}(x). \tag{6.58}$$

Пусть ε — произвольное положительное число, тогда

$$\int_{-\infty}^{\infty} x^2\,dF_{nk}(x) = \int_{|x| \leqslant \varepsilon} x^2\,dF_{nk}(x) + \int_{|x| > \varepsilon} x^2\,dF_{nk}(x) \leqslant$$

$$\leqslant \varepsilon^2 \int_{-\infty}^{\infty} dF_{nk}(x) + \int_{|x| > \varepsilon} x^2\,dF_{nk}(x) = \varepsilon^2 + \int_{|x| > \varepsilon} x^2\,dF_{nk}(x). \tag{6.59}$$

В силу условия Линдеберга второе слагаемое в (6.59) при достаточно больших n будет меньше ε^2.

Таким образом,

$$|\varphi_{nk}(t) - 1| \leqslant t^2\varepsilon^2,$$

т. е. для всех достаточно больших n равномерно относительно k, $k=\overline{1,n}$ и t в любом конечном интервале $|t| \leqslant T$

$$|\varphi_{nk}(t) - 1| \leqslant \varepsilon^2 T^2.$$

В силу произвольности $\varepsilon > 0$ равномерно относительно k, $k=\overline{1,n}$

$$\lim_{n\to\infty} \varphi_{nk}(t) = 1. \tag{6.60}$$

В частности, для всех достаточно больших n при t, лежащих в произвольном конечном интервале $|t| \leqslant T$, выполняется неравенство

$$|\varphi_{nk}(t) - 1| < \frac{1}{2}. \tag{6.61}$$

Теперь перейдем к доказательству (6.51). Из (6.60) следует, что в любом конечном интервале $|t|<T$ для любого k, $k=\overline{1,n}$ справедливо разложение

$$\ln \varphi_{nk}(t) = \ln\left[1 + (\varphi_{nk}(t) - 1)\right] = [\varphi_{nk}(t) - 1] - \frac{1}{2}[\varphi_{nk}(t) - 1]^2 + \frac{1}{3}[\varphi_{nk}(t) - 1]^3 - \ldots = [\varphi_{nk}(t) - 1] + \sum_{s=2}^{\infty} \frac{(-1)^{s-1}}{s}[\varphi_{nk}(t) - 1]^s,$$

поэтому (6.50) можно представить в виде

$$\ln \varphi_n(t) = \sum_{k=1}^{n} [\varphi_{nk}(t) - 1] + R_n, \tag{6.62}$$

где

$$R_n = \sum_{k=1}^{n} \sum_{s=2}^{\infty} \frac{(-1)^{s-1}}{s}[\varphi_{nk}(t) - 1]^s. \tag{6.63}$$

Далее, используя (6.54), представим (6.62) в виде

$$\ln \varphi_n(t) = -\frac{t^2}{2} + \frac{t^2}{2} + \sum_{k=1}^{n} \int_{-\infty}^{\infty} (\exp itx - 1 - itx)\, dF_{nk}(x) + R_n = -\frac{t^2}{2} + \rho_n + R_n, \tag{6.64}$$

где

$$\rho_n = \frac{t^2}{2} + \sum_{k=1}^{n} \int_{-\infty}^{\infty} (\exp ixt - 1 - itx)\, dF_{nk}(x). \tag{6.65}$$

Оценим величину R_n. Учитывая (6.61), получим

$$|R_n| \leqslant \sum_{k=1}^{n} \sum_{s=2}^{\infty} \frac{1}{2} |\varphi_{nk}(t) - 1|^s = \frac{1}{2} \sum_{k=1}^{n} \frac{|\varphi_{nk}(t) - 1|^2}{1 - |\varphi_{nk}(t) - 1|} \leqslant \sum_{k=1}^{n} |\varphi_{nk}(t) - 1|^2.$$

На основании (6.58) и (6.44) имеем

$$\sum_{k=1}^{n} |\varphi_{nk}(t) - 1| \leqslant \sum_{k=1}^{n} \frac{t^2}{2} \int_{-\infty}^{\infty} x^2\, dF_{nk}(x) = \frac{t^2}{2},$$

поэтому

$$|R_n| \leqslant \sum_{k=1}^{n} |\varphi_{nk}(t) - 1|^2 \leqslant \max_{1\leqslant k\leqslant n} |\varphi_{nk}(t) - 1| \sum_{k=1}^{n} |\varphi_{nk}(t) - 1| \leqslant \frac{t^2}{2} \max_{1\leqslant k\leqslant n} |\varphi_{nk}(t) - 1|. \tag{6.66}$$

Следовательно, равномерно относительно t в произвольном конечном интервале $|t| \leqslant T$ при $n \to \infty$

$$R_n \to 0. \tag{6.67}$$

Оценим величину ρ_n (6.65). Пусть задано $\varepsilon > 0$. Преобразуем выражение для ρ_n следующим образом, используя (6.43) и (6.44):

$$\rho_n = \frac{t^2}{2}\sum_{k=1}^{n} \mathrm{D}\,X_{nk} + \sum_{k=1}^{n}\int_{-\infty}^{\infty}(\exp itx - 1 - itx)\,dF_{nk}(x) =$$

$$= \sum_{k=1}^{n}\int_{-\infty}^{\infty}\left(\exp itx - 1 - itx - \frac{(itx)^2}{2}\right)dF_{nk}(x) =$$

$$= \sum_{k=1}^{n}\int_{|x|\leqslant\varepsilon}\left(\exp itx - 1 - itx - \frac{(itx)^2}{2}\right)dF_{nk}(x) +$$

$$+ \sum_{k=1}^{n}\left[\int_{|x|>\varepsilon}(\exp itx - 1 - itx)\,dF_{nk}(x) + \frac{t^2}{2}\int_{|x|>\varepsilon}x^2\,dF_{nk}(x)\right]. \tag{6.68}$$

На основании (6.57)

$$\left|\sum_{k=1}^{n}\int_{|x|\leqslant\varepsilon}\left(\exp itx - 1 - itx - \frac{(itx)^2}{2}\right)dF_{nk}(x)\right| \leqslant$$

$$\leqslant \frac{|t|^3}{6}\sum_{k=1}^{n}\int_{|x|\leqslant\varepsilon}|x|^3\,dF_{nk}(x) \leqslant \frac{|t|^3}{6}\,\varepsilon\sum_{k=1}^{n}\int_{|x|\leqslant\varepsilon}x^2\,dF_{nk}(x). \tag{6.69}$$

На основании (6.56)

$$\left|\sum_{k=1}^{n}\left[\int_{|x|>\varepsilon}(\exp itx - 1 - itx)\,dF_{nk}(x) + \frac{t^2}{2}\int_{|x|>\varepsilon}x^2\,dF_{nk}(x)\right]\right| \leqslant$$

$$\leqslant t^2\sum_{k=1}^{n}\int_{|x|>\varepsilon}x^2\,dF_{nk}(x). \tag{6.70}$$

Таким образом,

$$|\rho_n| \leqslant \frac{|t|^3}{6}\,\varepsilon\sum_{k=1}^{n}\int_{|x|\leqslant\varepsilon}x^2\,dF_{nk}(x) + t^2\sum_{k=1}^{n}\int_{|x|>\varepsilon}x^2\,dF_{nk}(x). \tag{6.71}$$

Используя (6.44), получим

$$\sum_{k=1}^{n}\int_{|x|\leqslant\varepsilon}x^2\,dF_{nk}(x) = \sum_{k=1}^{n}\int_{-\infty}^{\infty}x^2\,dF_{nk}(x) -$$

$$- \sum_{k=1}^{n}\int_{|x|>\varepsilon}x^2\,dF_{nk}(x) = 1 - \sum_{k=1}^{n}\int_{|x|>\varepsilon}x^2\,dF_{nk}(x),$$

поэтому неравенство (6.71) принимает вид

$$|\rho_n| \leqslant \frac{|t|^3}{2}\varepsilon + t^2\left[1 - \frac{|t|}{6}\varepsilon\right] \sum_{k=1}^{n} \int\limits_{|x|>\varepsilon} x^2\,dF_{nk}(x). \quad (6.72)$$

Из условия Линдеберга (6.52) следует, что, выбирая n достаточно большим, можно сделать второе слагаемое в (6.72) при любом $\varepsilon>0$ сколь угодно малым. А поскольку ε произвольно, то его можно выбрать так, чтобы при любом $\eta>0$ и T для всех t из интервала $|t|<T$ выполнялось неравенство $|\rho_n|\leqslant\eta$ при всех $n>n_0$ (ε, η, T), т. е. равномерно в каждом конечном интервале значений t

$$\lim_{n\to\infty} \rho_n = 0. \quad (6.73)$$

Принимая во внимание (6.64), (6.67) и (6.73), приходим к выводу, что равномерно в каждом конечном интервале t

$$\lim_{n\to\infty} \ln \varphi_n(t) = -\frac{t^2}{2},$$

что и требовалось доказать.

Следствие. *Если независимые случайные величины $X_1, X_2, \ldots$ одинаково распределены и имеют конечную отличную от нуля дисперсию, то при $n\to\infty$ равномерно по x*

$$P\left[\frac{1}{B_n}\sum_{k=1}^{n}(X_k - \mathbf{M}X_k) < x\right] \to \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} \exp\left(-\frac{z^2}{2}\right) dz. \quad (6.74)$$

Действительно, пусть случайные величины X_k, $k=1, 2, \ldots$ распределены по закону $F(x)$, $\mathbf{M}X_k=\alpha$, $\mathbf{D}X_k=\sigma^2$, тогда $B_n=\sigma\sqrt{n}$ и сумма в условии Линдеберга принимает следующий вид:

$$\sum_{k=1}^{n} \int\limits_{|x-\alpha|>\tau B_n} \frac{1}{B_n^2}(x-\alpha)^2\,dF(x) =$$

$$= \frac{1}{\sigma^2} \int\limits_{|x-\alpha|>\tau\sigma\sqrt{n}} (x-\alpha)^2\,dF(x). \quad (6.75)$$

Поскольку по условию дисперсия случайных величин X_k конечна, то интеграл в правой части равенства (6.75) при $n\to\infty$ стремится к нулю, и, следовательно, условие Линдеберга выполняется, а значит имеет место (6.74).

Выясним теперь смысл условия Линдеберга.

Пусть A_k, $k=\overline{1, n}$ обозначает событие, состоящее в том, что

$$|X_k - \alpha_k| > \tau B_n, \quad \tau > 0.$$

Оценим вероятность того, что

$$\max_{1\leqslant k\leqslant n} |X_k - \alpha_k| > \tau B_n. \tag{6.76}$$

События

$$\max_{1\leqslant k\leqslant n} |X_k - \alpha_k| > \tau B_n \quad \text{и} \quad \bigcup_{k=1}^{n} A_k$$

равносильны, так как наступление первого из них влечет второе и обратно, поэтому

$$P\left(\max_{1\leqslant k\leqslant n} |X_k - \alpha_k| > \tau B_n\right) = P\left(\bigcup_{k=1}^{n} A_k\right).$$

Воспользовавшись неравенством (1.20), получим

$$P\left(\max_{1\leqslant k\leqslant n} |X_k - \alpha_k| > \tau B_n\right) \leqslant \sum_{k=1}^{n} P(A_k), \tag{6.77}$$

но

$$P(A_k) = \int\limits_{|x-\alpha_k| > \tau B_n} dF_{nk}(x) \leqslant$$

$$\leqslant \frac{1}{\tau^2 B_n^2} \int\limits_{|x-\alpha_k| > \tau B_n} (x-\alpha_k)^2\, dF_k(x). \tag{6.78}$$

Здесь $F_k(x)$ — функция распределения случайной величины X_k, $k=1, 2, \ldots, n$.

Неравенства (6.77) и (6.78) дают оценку вероятности события (6.76)

$$P\left(\max_{1\leqslant k\leqslant n} |X_k - \alpha_k| > \tau B_n\right) \leqslant$$

$$\leqslant \frac{1}{\tau^2 B_n^2} \sum_{k=1}^{n} \int\limits_{|x-\alpha_k| > \tau B_n} (x-\alpha_k)^2\, dF_k(x).$$

Если условие Линдеберга выполняется, то, как следует из этой оценки, каково бы ни было $\tau > 0$, при $n \to \infty$ вероятность события $\max\limits_{1\leqslant k\leqslant n} |X_k - \alpha_k| > \tau B_n$ стремится к нулю. Отсюда видно, что условие Линдеберга выражает в математической форме требование равномерной малости всех слагаемых суммы

$$\sum_{k=1}^{n} \frac{(X_k - \alpha_k)}{B_n}.$$

ЛИТЕРАТУРА

1. Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. М., «Наука», 1965.

2. Будак Б. М., Фомин С. В. Кратные интегралы и ряды. М., «Наука», 1967.

3. Гнеденко Б. В., Колмогоров А. Н. Предельные распределения для сумм независимых случайных величин. Гостехиздат, 1949.

4. Гнеденко Б. В. Курс теории вероятностей. М., «Наука», 1969.

5. Ильин В. А., Позняк Э. Г. Аналитическая геометрия. М., «Наука», 1971.

6. Лабутин В. К. Транзисторы общего назначения. «Энергия», 1964.

7. Риордан Дж. Введение в комбинаторный анализ. Изд. иностр. лит., 1963.

8. Свешников А. Г., Тихонов А. Н. Теория функций комплексного переменного. М., «Наука», 1970.

9. Смирнов В. И. Курс высшей математики, т. 1. Физматгиз, 1967.

10. Смирнов В. И. Курс высшей математики, т. 5. Физматгиз, 1959.

11. Феллер В. Введение в теорию вероятностей и ее приложения, т. 1. М., «Мир», 1967.

12. Ципкин Я. З. Теория импульсных систем. Физматгиз, 1958.

ОГЛАВЛЕНИЕ

Борис Евгеньевич Аксенов, Иван Васильевич Афонькин, Владимир Петрович Евменов, Михаил Иванович Нечипоренко

Основы теории вероятностей

Учебное пособие

Часть 1

Научные редакторы *И. Д. Бутомо, В. С. Алмазова*

Редактор *И. А. Лебедева*

Корректоры *С. Д. Рутковская* и *Н. Н. Тарасова*

М-09594. Подписано к печати 11/XII 1973 г. Формат бумаги $60 \times 90^1/_{16}$.
Объем 11,75 печ. л. Заказ 555. Тираж 2000. Цена 60 коп.

Лаборатория полиграфических машин Ленинградского ордена Ленина политехнического института им. М. И. Калинина
194251, Ленинград, Политехническая ул., 29.

Министерство высшего и среднего специального образования РСФСР

ЛЕНИНГРАДСКИЙ ОРДЕНА ЛЕНИНА
ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ
имени М. И. КАЛИНИНА

Б. Е. АКСЕНОВ, И. В. АФОНЬКИН, В. П. ЕВМЕНОВ,
М. И. НЕЧИПОРЕНКО

ОСНОВЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

УЧЕБНОЕ ПОСОБИЕ

Часть 1

Ленинград
1973